



**HAL**  
open science

# Schémas numérique d'ordre élevé en temps et en espace pour l'équation des ondes du premier ordre. Application à la Reverse Time Migration.

Florent Ventimiglia

► **To cite this version:**

Florent Ventimiglia. Schémas numérique d'ordre élevé en temps et en espace pour l'équation des ondes du premier ordre. Application à la Reverse Time Migration.. Equations aux dérivées partielles [math.AP]. Université de Pau et des Pays de l'Adour, 2014. Français. NNT: . tel-01111039

**HAL Id: tel-01111039**

**<https://theses.hal.science/tel-01111039v1>**

Submitted on 29 Jan 2015

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

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

## THÈSE DE DOCTORAT

présentée à

L'Université de Pau et des Pays de l'Adour  
École doctorale des sciences et leurs applications - ED 211  
par

**Florent VENTIMIGLIA**

pour obtenir le grade de

**DOCTEUR de l'Université de Pau et des Pays de l'Adour**  
**Spécialité Mathématiques Appliquées**

---

# SCHÉMAS NUMÉRIQUES D'ORDRE ÉLEVÉ EN TEMPS ET EN ESPACE POUR L'ÉQUATION DES ONDES DU PREMIER ORDRE.

-

## APPLICATION À LA REVERSE TIME MIGRATION.

---

soutenue le 05 Juin 2014

**Après avis de :**

M. Bruno LOMBARD	Chargé de recherches CNRS - Université de Marseille	Rapporteur
M. Jean VIRIEUX	Professeur - Université de Grenoble	Rapporteur

**et devant la commission d'examen formée de :**

Mme. Hélène BARUCQ	Directrice de Recherche INRIA - INRIA Bordeaux Sud-Ouest	Directrice de thèse
M. Gilles CARBOU	Professeur - LMAP, Université de Pau et des Pays de l'Adour	Examineur
M. Bertrand DENEL	Ingénieur R&D - CSTJF, TOTAL Pau	Examineur
M. Julien DIAZ	Chargé de Recherche INRIA - INRIA Bordeaux Sud-Ouest	Directeur de thèse
M. Luc GIRAUD	Directeur de Recherche INRIA, INRIA Bordeaux Sud-Ouest	Examineur
M. Bruno LOMBARD	Chargé de recherches CNRS - Université de Marseille	Rapporteur
M. Jean VIRIEUX	Professeur - Université de Grenoble	Rapporteur
M. Paul WILLIAMSON	Ingénieur R&D - CSTJF, TOTAL Pau	Examineur

**Équipe Projet INRIA MAGIQUE-3D**, Institut National de Recherche en Informatique et en Automatique (INRIA)

-

**TOTAL SA**, Centre Scientifique et Technique Jean Feger (CSTJF), Pau

-

**Laboratoire de Mathématiques et de leurs Applications de Pau**, Unité Mixte de Recherche CNRS 5142, Université de Pau et des Pays de l'Adour (UPPA)



# Remerciements

Un grand merci à mon jury de thèse pour avoir pris le temps d' assister à ma soutenance. Je remercie chaleureusement Messieurs J. VIRIEUX et B. LOMBARD pour avoir été mes rapporteurs. Cela à été pour moi un honneur. Merci également à mes directeurs de thèse H. Barucq et J. Diaz.

Néanmoins je tenais également à remercier bon nombre de personnes qui ont su faire preuve de patience et de soutiens durant mes jours de doutes et d'incertitudes ... Je pense plus particulièrement à mes parents, ma soeur, ma cousine, ma grand-mère, mes amis : KK, Gaga, Veyo, Nunie, Cycy, Jéjé, François, ROM, Angelito, Juanito, Elies, JJ, mon petit Willy, Mrs Théophile Gautier :p. Grâce à certains d'entre eux j'ai aussi pu enrichir mes "tags" ainsi que mon espagnol loool : "¿ Conoces a Maria ? :p". Et également mes collègues et amis : Vanessa, Marie, Aralar, Simon, Juan, Victor (j'en oublie sûrement ...). Ces trois années d'aventures me feront garder également en mémoire de très belles rencontres et échanges lors de déplacements et conférences.

Une page vient de se tourner, une autre doit s'écrire ...

Je pars vers un avenir qui je l'espère sera prometteur et plein de richesses tant intellectuel, que professionnel et humaine ...

Bon courage aux thésards qui sont actuellement en pleine rédaction (Jéjé, Vanessa, Lionel), ça va le faire " no stress, take it easy :) " (oui il fallait bien que je parle Anglais à un moment donné "Adams-Bashforth ;) ?")

Pour tous ceux que j'aurais oublié, un très grand **MERCI** également.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Équations d'ondes et discrétisation en espace</b>	<b>7</b>
2.1	L'équation des ondes acoustiques . . . . .	8
2.1.1	Formulation standard de l'équation des ondes acoustiques . . . . .	8
2.1.2	Formulation du premier ordre de l'équation des ondes acoustiques . . . . .	8
2.2	Discrétisation en espace . . . . .	9
2.2.1	Formulation DG avec flux centré pour l'équation des ondes acoustiques . . . . .	9
2.2.1.1	Formulation variationnelle continue . . . . .	9
2.2.1.2	Formulation variationnelle discrète . . . . .	14
2.2.2	Formulation IPDG . . . . .	16
<b>3</b>	<b>Schémas d'ordre élevé en temps</b>	<b>21</b>
3.1	Schémas en temps d'ordre élevé pour la formulation du premier ordre de l'équation des ondes acoustiques . . . . .	22
3.1.1	Schéma Leap-Frog . . . . .	22
3.1.2	Schéma ADER . . . . .	23
3.1.3	Schéma $\nabla^m$ . . . . .	27
3.1.3.1	Discrétisation en temps . . . . .	27
3.1.3.2	Discrétisation en espace . . . . .	29
3.1.3.3	Le schéma totalement discrétisé . . . . .	31
3.2	Schémas en temps d'ordre élevé pour la formulation du second ordre de l'équation des ondes acoustiques . . . . .	33
3.2.1	Le schéma LF . . . . .	33
3.2.2	Le schéma ADER . . . . .	34
3.2.3	Le schéma biharmonique . . . . .	35
3.3	Résultats Numériques . . . . .	38
3.3.1	Cas test académique . . . . .	38
3.3.2	Validation sur benchmark : le modèle Marmousi . . . . .	44
3.3.3	Analyse du cas avec une source : le cas de maillages réguliers . . . . .	47
3.4	Conclusion . . . . .	50
3.A	Calcul pratique du terme source . . . . .	50
<b>4</b>	<b>Analyse de dispersion des schémas temporels</b>	<b>53</b>
4.1	Méthodologie générale . . . . .	53
4.2	Condition nécessaire de stabilité des schémas . . . . .	57

4.3	Analyse de dispersion des schémas . . . . .	59
4.4	Schéma avec pénalisation . . . . .	67
4.4.1	Condition nécessaire de stabilité des schémas avec pénalisation . . . . .	70
4.4.2	Analyse de dispersion des schémas avec pénalisation . . . . .	70
4.5	Conclusion . . . . .	81
<b>5</b>	<b>Étude de l'effet du terme de pénalisation en dimension deux</b>	<b>85</b>
5.1	Cas test académique avec condition initiale . . . . .	85
5.1.1	Cas test académique avec condition initiale . . . . .	86
5.1.1.1	Le schéma DG-Flux avec pénalisation . . . . .	86
5.2	Cas de la source ponctuelle . . . . .	89
5.2.1	Le schéma DG-Flux avec pénalisation . . . . .	89
5.2.1.1	Le schéma IPDG . . . . .	90
5.3	Validation sur benchmark : le modèle Marmousi . . . . .	93
5.4	Application à l'imagerie sismique . . . . .	93
5.4.1	Présentation de la Reverse Time Migration . . . . .	96
5.4.2	Condition d'imagerie et stratégie de stockage . . . . .	96
5.4.3	Caractéristiques du modèle . . . . .	97
5.5	Étude de l'effet d'autres facteurs sur la précision du code DIVA . . . . .	101
5.5.1	Effet de la simple précision . . . . .	101
5.5.2	Effet du choix du vecteur de pénalisation . . . . .	102
5.6	Conclusion . . . . .	106
<b>6</b>	<b>Discrétisation en temps avec des pas de temps locaux</b>	<b>107</b>
6.1	Pas de temps local pour une équation hyperbolique d'ordre deux en temps . . . . .	108
6.2	Utilisation de la méthode d'Adams-Bashforth . . . . .	112
6.2.1	Présentation de la méthode . . . . .	112
6.2.2	Méthode d'Adams-Bashforth avec pas de temps local . . . . .	114
6.3	Utilisation du schéma de Störmer-Verlet . . . . .	116
6.3.1	Présentation de la méthode . . . . .	117
6.4	Résultats numériques . . . . .	123
6.4.1	Le schéma SV . . . . .	123
6.4.2	Le schéma AB . . . . .	125
6.5	Conclusion . . . . .	127
	<b>Bibliographie</b>	<b>127</b>

# Chapitre 1

## Introduction

La propagation des ondes intervient dans de nombreux domaines applicatifs et des communautés très différentes, comme par exemple les médecins et les géophysiciens, l'utilisent pour produire des images synthétiques et éviter ainsi des techniques invasives dangereuses ou/et coûteuses. L'objet de cette thèse est de proposer des schémas en temps d'ordre élevé pour améliorer la qualité des champs d'ondes calculés avec une charge de calcul raisonnable. Dans le cadre d'une collaboration avec la compagnie pétrolière Total, nous nous sommes aussi intéressés à l'impact de ces schémas sur une technique d'imagerie du sous-sol qui est appelée Reverse Time Migration ou RTM.

La RTM nécessite de résoudre un grand nombre d'équations d'ondes complètes et s'appuie sur la théorie élaborée par Claerbout dans les années 70 [17]. C'est une technique dont la mise en œuvre se fait en deux étapes. La première, appelée campagne sismique, est expérimentale et fournit un ensemble de mesures des ondes réfléchies qui sont enregistrés par des capteurs positionnés en surface ou en profondeur, selon les zones à explorer. Elle peut être vue comme la phase d'initialisation de la RTM tandis que la deuxième étape va fournir une image numérique du sous-sol après un certain nombre d'itérations. Chacune des itérations a pour objectif de projeter les mesures temporelles (temps d'arrivée des ondes réfléchies) dans le domaine spatial afin de retrouver la position des réflecteurs enfouis dans le domaine exploré. Cette projection est réalisée en résolvant à chaque itération, et pour une source donnée, deux équations d'onde dans le domaine temporel, ce qui génère une charge énorme de calculs. Par exemple, en adoptant une démarche très directe pour imager un bloc de côté 5 km par 5 km, on va positionner un réseau de sources tous les 25 m, ce qui signifie qu'on va résoudre 1000 équations d'ondes uniquement pour simuler la propagation directe des sources. A cela, s'ajoute la rétro-propagation des ondes réfléchies captées par le réseau de capteurs qui est simulée en résolvant 1000 équations d'onde de plus. Pour former l'image du sous-sol, on corrèle champs propagés et rétro-propagés en chaque point de la grille de calcul, ce qui, a priori, nécessite de stocker un nombre très important d'instantanés. A la difficulté de résoudre des équations d'ondes dans des milieux hétérogènes, s'ajoute donc l'encombrement de la mémoire qui peut être très facilement un facteur bloquant dans des cas réalistes. La RTM, comme procédé robuste d'imagerie sismique, a été très longtemps considérée dans une version simplifiée qui s'appuyait sur des équations d'ondes approchées (Kirchhoff, Beam . . .) car les moyens de calcul disponibles n'étaient pas suffisants pour résoudre des équations d'ondes complètes à 3D. Les progrès constants du calcul scientifique ont remis à l'ordre du jour l'utilisation de la RTM mais ils ne suffisent pas toujours pour imager des zones dont la tectonique est de plus en plus complexe. On pourrait bien évidemment attendre que les capacités de calcul augmentent encore mais on peut aussi prendre le parti de développer des méthodes numériques avancées qui, de plus, pourront toujours s'appliquer à des modèles plus complexes que les équations d'ondes élastiques standards.



Dans ce cas, il est préférable de développer des méthodes de calcul d'ordre élevé qui vont fournir des champs d'ondes précis. Mais augmenter l'ordre de la méthode est un objectif qui doit être atteint en pensant à la parallélisation qui va réduire les temps de calcul. C'est un facteur de réussite, notamment dans un contexte RTM où chaque itération requiert la résolution d'un grand nombre d'équations d'ondes en domaine temporel. Le schéma doit donc se prêter à une mise en oeuvre en parallèle efficace et pour cela, il doit être conçu en veillant à ce que le nombre de communications entre les noeuds de calcul soit minimisé. Toujours dans une perspective de RTM, on doit utiliser l'espace mémoire avec soin car on peut vite en atteindre les limites. Pour s'en convaincre, il suffit de penser aux 1000 équations d'ondes à résoudre et à la nécessité de stocker à chaque pas de temps 1000 instantanés en chaque noeud du maillage. En pratique, on adopte une démarche plus élégante qui réduit le nombre d'historiques à stocker mais il est encore facile d'atteindre les limites de la mémoire. Dans l'idée de limiter l'occupation de la mémoire, on choisit donc souvent de privilégier une représentation explicite des champs d'onde. Le prix à payer est de satisfaire une condition de Courant-Friedrichs-Lewy (CFL) qui assure la stabilité du schéma à condition de prendre un pas de temps assez petit. L'espace mémoire est donc privilégié à la durée des calculs, ce qui n'est pas forcément très gênant si on utilise le calcul parallèle. C'est d'ailleurs le choix qui est fait dans cette thèse. Mais avant de parler plus en détail de la discrétisation en temps, expliquons ce qui a guidé notre choix pour la discrétisation en espace.

La méthode des différences finies (FD) est la technique d'approximation en espace la plus répandue dans la communauté géophysique, sa grande popularité étant certainement due à sa facilité de mise en oeuvre dans des codes de simulation. Toutefois, la méthode FD n'est pas vraiment adaptée ni efficace pour des milieux où la topographie varie beaucoup. En effet, pour reproduire correctement des effets de topographie, on doit utiliser des grilles très fines, à moins d'appliquer des FD sur des grilles non structurées. Dans ce dernier cas, on perd toute la simplicité de mise en oeuvre des FD sans profiter de la flexibilité d'autres méthodes comme les méthodes d'éléments finis. La méthode FD peut être confrontée à des problèmes de dispersion numérique, phénomène non physique généré par la discrétisation des équations et qui se traduit par une dépendance de la vitesse par rapport au pas de discrétisation (à la fréquence), (Sei & Symes [58] ; Yang *et al.* [63]). La dispersion numérique survient principalement quand on ne prend pas assez de points par longueur d'onde ou quand les contrastes de vitesse sont très importants. On peut aussi citer quelques travaux (Fei & Lerner [33], Yang *et al.* [63]) qui mentionnent des problèmes de dispersion liés à l'injection de la source. La présence de ce type d'artéfacts numériques a motivé l'élaboration de différentes méthodes de différences finies. Parmi elles, citons le schéma de Virieux, connu aussi sous le nom de méthode Staggered-Grid ou SG (Virieux, [62] ; voir aussi Fornberg, [35] ; Igel *et al.*, [46]) qui, comparé au schéma FD standard d'ordre deux, améliore la précision de la solution tout en préservant les propriétés de stabilité sans augmenter les coûts de calcul et l'occupation de la mémoire. Mais la méthode SG souffre aussi de dispersion numérique et une façon de s'en débarrasser, ou du moins de la limiter, est d'utiliser des schémas d'ordre élevé. Les schémas d'ordre élevé permettent d'améliorer la qualité des résultats mais des problèmes de dispersion numérique subsistent quand on utilise des grilles assez grossières. De plus, les schémas d'ordre élevé ont un support ("stencil") plus large [65], ce qui peut être un obstacle à une mise en oeuvre parallèle réellement efficace des algorithmes. On observe aussi que ces schémas ne permettent pas toujours de prendre en compte certaines conditions aux limites comme les conditions aux limites absorbantes. En vue d'éliminer la dispersion numérique, Fei et Lerner [33] et Yang *et al.* [63] ont proposé la méthode FCT (Flux-Corrected-Transport) mais quand la grille est trop grossière, la méthode FCT peut difficilement récupérer la précision initialement perdue dans la dispersion numérique. On peut alors penser aux méthodes pseudo-spectrales (PS) qui sont exactes dans une bande de fréquence allant jusqu'à la fréquence de Nyquist mais elles utilisent des transformées de Fourier (FFT). Les méthodes PS sont donc excessivement coûteuses en temps de calcul, difficiles à mettre en oeuvre dans le

cas de conditions de bord non périodiques et pas du tout adaptées au calcul parallèle car les accès à la mémoire des FFT ne sont pas locaux (Mizutani et al., [56]). De plus, elles souffrent aussi de dispersion numérique en temps dont les effets sont dramatiques quand le pas de temps est assez grand (Dablain [23]; Sei & Symes [58]; Yang *et al.*, 2006). Il y a une décennie, Yang et ses collaborateurs [64] ont proposé une nouvelle approche, la méthode NAD (Nearly Analytic Discrete) qui avait été initialement développée par Konddoh *et al.* [53] pour résoudre des problèmes paraboliques et hyperboliques. La méthode NAD est un autre moyen efficace pour limiter la dispersion numérique. On observe que cette méthode est beaucoup plus stable que la méthode FD si bien qu'elle permet d'utiliser une grille plus grossière pour obtenir le même niveau de précision. Les temps CPU (Central Processing Unit) des deux méthodes sont donc comparables. Par contre, la méthode NAD n'est que d'ordre deux et ne semble pas adaptée pour simuler des ondes dans des milieux poreux complexes. Une autre façon de résoudre le problème de dispersion numérique est d'utiliser des grilles fines pour augmenter le nombre de points par longueur d'onde. Par exemple, Holberg [45] a montré que pour un schéma d'ordre deux, il faut prendre 20 points sur la longueur d'onde minimale pour obtenir des résultats corrects tandis que 10 points suffisent pour un schéma d'ordre quatre. Dablain [22] a montré qu'à la fréquence de Nyquist, il faut 8 points pour éliminer la dispersion numérique d'une méthode FD du second ordre et 4 points pour une méthode FD d'ordre 4. Mais plus il y a de points par longueur d'onde, plus les coûts de calcul et de stockage sont élevés. On peut donc comprendre que ces méthodes soient difficilement applicables pour réaliser des simulations 3D d'ondes sismiques car elles utilisent trop de CPU et nécessitent trop d'accès directs à la mémoire. Toutefois, les progrès récents et spectaculaires du calcul scientifique, aidés par l'émergence de nouvelles architectures de calcul, permettent aujourd'hui de croire aux méthodes numériques avancées pour résoudre des problèmes d'ondes en milieux réalistes.

Notre objectif étant de développer un code de simulation pour l'imagerie sismique capable de tenir compte des effets de topographie, il nous a semblé plus opportun de privilégier des méthodes d'éléments finis d'ordre élevé. Pendant longtemps, celles-ci n'ont pas été très plébiscitées pour les applications à la géophysique industrielle car leur mise en œuvre n'est pas toujours facile. Il existe bien des méthodes d'éléments finis dont l'implémentation est à peine plus compliquée que celle des FD mais ces méthodes génèrent en général des charges de calcul rédhibitoires car leur matrice de masse n'est pas facile à inverser. Il existe tout de même une méthode d'éléments finis très reconnue et très utilisée par la communauté géophysique. Il s'agit de la méthode des éléments spectraux SEM [13, 52, 51] qui a la propriété remarquable de posséder une matrice de masse diagonale du même ordre (i.e. calculée par une formule de Gauss-Lobatto d'ordre optimal) que la méthode d'éléments finis. Elle est au cœur du code SPECFEM (pour plus d'information sur le code : <http://www.geodynamics.org/>). Le code SPECFEM est un code qui est très utilisé dans le monde entier pour produire des calculs d'aléas sismiques, très souvent dans les minutes qui suivent un tremblement de terre. Toutefois, les éléments spectraux ne semblent pas vraiment adaptés à des calculs dans des milieux dont les caractéristiques oscillent fortement et leur usage justifie le développement de techniques d'homogénéisation, comme celles proposées par exemple par Capdeville [43]. De plus, à 3D, il n'est pas possible d'utiliser des maillages tétraédriques avec des éléments spectraux d'ordre élevé. Cet aspect est à prendre en considération dans le cadre de cette thèse car notre objectif est de développer un solveur utilisable dans l'industrie. Il doit donc dépendre d'outils de maillage facilement accessibles et à notre connaissance, il n'existe pas dans le commerce d'outils permettant de construire facilement des maillages hexaédriques. C'est pourquoi notre attention va se porter plus particulièrement sur les méthodes de Galerkin discontinues (DGM) car elles ont l'avantage de conduire à des formulations quasi-explicites, la matrice de masse étant diagonale par blocs. De plus, les DGM se prêtent parfaitement à une mise en œuvre des calculs en parallèle car elles reposent sur des calculs locaux à l'échelle de l'élément. Enfin, ces méthodes sont bien adaptées à la prise en compte des hétérogénéités du milieu de propagation

et contribuent à limiter les effets de dispersion numérique, offrant la possibilité d'adapter le pas de discrétisation  $h$  et le degré d'approximation polynomiale  $m$ . La thèse de Caroline Baldassari [8] a bien illustré l'intérêt d'utiliser une méthode d'éléments finis discontinus pour résoudre l'équation des ondes acoustiques pour imager des régions hétérogènes. La méthode de discrétisation reposait sur une formulation du second ordre de l'équation des ondes qui est stabilisée via l'introduction d'un paramètre de pénalisation. Cette approche est connue dans la littérature [5, 18] sous l'acronyme IPDGM (Interior Penalty Discontinuous Galerkin Method). Elle s'avère très robuste et a démontré son intérêt dans un contexte RTM [8, 2]. Dans cette thèse, nous avons fait le choix d'utiliser une méthode d'éléments finis discontinus appliquée à l'équation des ondes acoustiques formulée comme un système du premier ordre. Nous aurions pu poursuivre en améliorant le solveur IPDGM mais notre objectif était de développer un code facilement adaptable pour les ondes élastiques et adopter une formulation vectorielle pour les deux types d'ondes nous a semblé pertinent, d'autant plus que les simulations que nous devons réaliser nécessitent l'emploi de conditions de bord absorbantes qui sont plus facilement intégrables dans une formulation du premier ordre que dans une formulation du second ordre. Dans cette thèse, nous avons choisi de mettre en oeuvre une méthode d'ordre élevé en espace dans le but d'utiliser un nombre minimum d'éléments par longueur d'onde. De plus, en appliquant une méthode d'ordre élevé, on espère améliorer les performances de la méthode dans le cas où le milieu est fortement hétérogène. Nous avons été confortés dans cette idée en lisant les travaux de Jameson [47, 48]. Dans [47], on montre qu'il y a un net avantage à utiliser des schémas en espace d'ordre élevé pour résoudre des problèmes comportant des échelles très différentes. Les méthodes d'ordre élevé en espace s'avèrent aussi très efficaces quand les coefficients représentant le milieu de propagation sont discontinus. Dans [48], on trouve une comparaison entre les méthodes adaptatives et les méthodes d'ordre élevé et la conclusion de cette étude est très largement en faveur des méthodes d'ordre élevé. Venons-en à la discrétisation en temps des équations à laquelle nous nous proposons de contribuer. Le schéma le plus utilisé est le schéma Leap-Frog (LF) qui est d'ordre deux. Il ne permet donc pas de profiter de l'ordre élevé de l'approximation en espace puisque l'ordre de convergence de l'approximation sera au mieux deux. Il est donc intéressant de développer des schémas d'ordre élevé en temps avec lesquels on peut aussi espérer relâcher la contrainte CFL et ainsi utiliser un pas de temps plus grand. On peut penser d'abord aux schémas de type Runge-Kutta. Ces schémas nécessitent des calculs auxiliaires qui doivent être stockés à chaque pas de temps et qui sont d'autant plus nombreux que le schéma est d'ordre élevé. Dans un souci de limiter l'occupation de la mémoire, notre choix s'est donc plutôt porté sur les schémas de type ADER (Arbitrary DERivative). Ces schémas ont été proposés par Dumbser et ses coauteurs, d'abord pour l'aéroacoustique [29, 31]. Ils peuvent se formuler comme des schémas de Lax-Wendroff [54], ce qui signifie que l'intégration en temps ne requiert qu'un pas de discrétisation et assure donc que la mémoire sera utilisée avec parcimonie. Pour intégrer en temps, il est nécessaire d'utiliser les degrés de liberté d'un élément ainsi que ceux des voisins directs. Cette propriété est très intéressante pour effectuer des calculs en parallèle. Il s'avère que les méthodes ADER-DG utilisent moins de mémoire que les méthodes RK-DG et on montre aussi qu'elles sont plus rapides. Enfin, les méthodes ADER démontrent les mêmes performances sur des maillages structurés et non structurés. On peut aussi citer les travaux de Bruno Lombard [55] qui propose, pour des équations d'ondes, une analyse numérique complète des schémas ADER d'ordre pair (4 à 10), en notant qu'à l'ordre 2, les schémas ADER et LF coïncident. Nous avons mis en oeuvre le schéma ADER avec deux objectifs. Le premier était de tester son efficacité après intégration dans la plate-forme industrielle DIVA développée à Total. Le second était de comparer ce schéma avec un nouveau schéma d'ordre élevé que nous avons développé pour l'équation des ondes. La construction de ce schéma ne suit pas les règles usuelles car sa mise en oeuvre suppose que la discrétisation en temps est faite avant la discrétisation en espace. Ce nouveau schéma est appelé schéma  $\nabla^m$  où  $m$

désigne l'ordre du schéma et  $\nabla$  indique que le schéma fait intervenir des opérateurs d'ordre plus ou moins élevé définis à partir de l'opérateur gradient. Cette idée a été exploitée dans un premier temps dans [1, 3] en association avec la méthode IPDG et le travail réalisé dans cette thèse est une adaptation, non triviale, au système du premier ordre.

Si on revient à l'application visée par cette thèse, on s'intéresse à la propagation d'ondes dans des milieux hétérogènes pour lesquels il est optimal d'utiliser des maillages non structurés composés de zones dans lesquelles la grille doit être particulièrement fine pour, par exemple, reproduire des détails géométriques petits, ou des variations importantes de topographie. Un schéma explicite ne peut fonctionner qu'avec un pas de temps calé sur les plus petits éléments du maillage. On va donc réaliser un nombre de calculs qui pourrait être minimisé en adaptant le pas de temps à la taille des mailles et on prend le risque d'avoir de la dispersion numérique en autorisant des régions où la condition CFL est "sur-satisfaite". Une façon de remédier à cela est d'utiliser des pas de temps locaux, comme suggéré par [40, 37] ou [25]. Cette idée avait été exploitée avec succès dans la thèse de Caroline Baldassari [8] en considérant l'équation des ondes acoustiques dans sa formulation scalaire. Précisons cependant que la technique des pas de temps locaux n'avait pas été abordée dans la perspective d'une implémentation parallèle. Dans les codes industriels, le partitionnement du maillage est très efficace lorsqu'on utilise des schémas à pas de temps global. Dès qu'on utilise des schémas à pas de temps local, l'équilibrage de charge est beaucoup plus difficile à réaliser. Il est en effet très difficile de gérer des domaines avec des petits pas de temps où on doit faire beaucoup d'itérations et des domaines avec des grands pas de temps et peu d'itérations. Ainsi des algorithmes très efficaces dans un cadre séquentiel peuvent avoir des performances très limitées dans un cadre HPC, à cause de l'équilibrage des charges et de l'augmentation des communications. C'est en particulier le cas des schémas à pas de temps local.

Ce manuscrit est organisé de la façon suivante. Faisant suite à cette introduction, le chapitre 2 est dédié à la présentation des équations et à la méthode de discrétisation en espace que nous avons choisie. Le chapitre 3 est consacré à des schémas en temps explicites et d'ordre élevé. Nous rappelons d'abord les schémas LF et ADER puis nous proposons un nouveau schéma dont la construction repose sur l'approximation d'opérateurs différentiels en espace d'ordre élevé. Notre approche est motivée par le fait que les méthodes DG sont parfaitement adaptées à la discrétisation d'opérateurs d'ordre élevé. Nous menons ensuite une analyse complète d'un nouveau schéma que nous comparons aux schémas LF et ADER. Nous utilisons pour cela des cas tests académiques et notre conclusion est que le nouveau schéma permet d'obtenir des résultats aussi précis qu'avec les schémas ADER pour un coût de calcul plus faible. Nous testons ensuite à la fin de ce chapitre notre schéma espace-temps sur des cas test réalistes fournis par Total. Nous observons l'apparition d'ondes parasites polluant fortement la solution. Comme ces ondes parasites sont également présentes lorsque nous utilisons les schémas ADER ou LF, il semble raisonnable de penser que le problème vient de la méthode de discrétisation en espace utilisée. Pour étayer cette conjecture, nous avons alors défini une expérience académique qui nous a permis de mettre en évidence des problèmes liés à l'injection d'une source ponctuelle.

Au chapitre 4, nous menons une analyse de dispersion des trois schémas pour mettre en évidence les ondes parasites observées. Nous proposons ensuite de supprimer ces ondes parasites en ajoutant un terme de pénalisation, comme cela avait été réalisé par Ainsworth, Monk et Muniz dans [4] sur des maillages cartésiens. L'analyse de dispersion du schéma modifié montre clairement l'influence positive de la pénalisation sur la qualité des solutions. Toutefois, c'est au détriment du pas de temps qui doit être diminué pour garantir la stabilité.

Au chapitre 5, nous reprenons les expériences menées au chapitre 3.3 en ajoutant le terme de pénalisation. Les résultats obtenus montrent à la fois l'intérêt des nouveaux schémas  $\nabla^m$  en terme de coûts de calculs, et l'amélioration apportée par le terme de pénalisation. Nous remarquons ensuite que l'utilisation de calculs en simple précision peut rendre complètement inefficace l'utilisation de

schémas d'ordre élevé. Nous mettons également en évidence la difficulté d'implémentation du nouveau terme de pénalisation sur des maillages non structurés. Nous terminons ce chapitre en présentant des résultats d'imagerie sismique obtenus par un algorithme de Reverse Time Migration couplé aux schémas d'ordre élevé en temps que nous avons étudiés. Pour finir cette thèse, nous nous sommes intéressés à la technique de pas de temps local appliquée au système d'ordre un, même si nous savions qu'elle serait difficilement applicable dans un code parallélisé. Nous avons développé deux schémas à pas de temps locaux. Le premier est une adaptation d'une version à pas de temps local du schéma d'Adams-Bashforth, proposée dans [37] pour l'équation des ondes scalaire. Ce schéma donne de bons résultats, mais il est inapplicable dans un algorithme de Reverse Time Migration pour deux raisons. Tout d'abord sa condition CFL est beaucoup plus restrictive que celle du schéma Leap-Frog classique, ce qui pénalise fortement les temps de calcul. Ensuite, c'est un schéma multi-pas, c'est-à-dire que pour calculer la solution à l'instant  $n + 1$ , il faut connaître ses valeurs non seulement à l'instant  $n$  mais également aux instants  $n - 1$ ,  $n - 2$  et  $n - 3$ . Lorsqu'on utilise un algorithme de RTM optimisé basé sur l'algorithme de Griewank [36], on utilise des *checkpoints*, c'est-à-dire des sauvegardes de la solution à plusieurs instants, afin de pouvoir relancer la simulation en ces différents instants. Avec des méthodes multi-pas, le coût de stockage des *checkpoints* est multiplié par quatre, ce qui rend la méthode inutilisable pour des calculs tridimensionnels. Nous nous sommes donc concentrés sur des méthodes à un pas. Parmi elles, le schéma LF est le plus utilisé mais il a pour inconvénient d'introduire une dissymétrie entre l'inconnue en pression, qui est calculée aux instants  $n$  et l'inconnue en vitesse, qui est calculée aux instants  $n + \frac{1}{2}$ . Ainsi, on ne dispose jamais de la solution complète à un instant donné. Or, les schémas à pas de temps locaux sont plus faciles à mettre en oeuvre si on connaît toute la solution au début de chaque pas de temps global. Nous avons donc pensé au schéma de Stormer-Verlet [61], qui est en fait une extension du schéma LF, et qui permet de calculer la pression et la vitesse à l'instant  $n$ . Nous avons proposé un nouveau schéma à pas de temps local, efficace en temps court, mais qui s'est malheureusement révélé instable en temps long.

## Chapitre 2

# Équations d'ondes et discrétisation en espace

L'imagerie géophysique est utilisée depuis longtemps par les compagnies pétrolières. Elle commence par une campagne d'acquisition qui restitue un volume très important d'enregistrements d'ondes sismiques réfléchies qui ont été générées par des ondes artificielles. Mathématiquement, les ondes réfléchies sont représentées comme les solutions d'équations d'ondes à coefficients variables représentant les caractéristiques physiques du milieu de propagation. Les images du sous-sol peuvent être fabriquées grâce à la résolution de ces équations. Le procédé d'imagerie est itératif car l'image numérique ne correspond pas forcément aux enregistrements obtenus lors de la campagne d'acquisition. Il peut donc être nécessaire de recommencer les simulations numériques plusieurs fois avant d'obtenir une image probante du sous-sol. Ces dix dernières années, l'imagerie sismique a gagné considérablement en précision car elle peut désormais s'appuyer sur les progrès constants du calcul scientifique. Pendant longtemps, les géophysiciens ont utilisé des équations d'ondes approchées comme les modèles one-way car les coûts de calcul engendrés par les équations d'ondes complètes étaient trop élevés pour obtenir des images 3D de milieux fortement hétérogènes. Il est aujourd'hui tout à fait imaginable de résoudre ces problèmes car les compagnies pétrolières se sont dotées de moyens de calcul puissants et en même temps, les méthodes numériques se sont améliorées.

Ce chapitre est dédié à la modélisation et à la discrétisation en espace qui a été choisie pour cette thèse. L'équipe Magique-3D a développé par le passé une méthode numérique avancée pour résoudre l'équation scalaire des ondes acoustiques. Il s'agit de la méthode IPDG (Interior Penalty Discontinuous Galerkin) qui a été couplée avec une méthode de pas de temps locaux [8, 9]. Cette méthode a démontré de très bonnes performances. Toutefois, dans la perspective de résoudre les équations de l'élastodynamique, il n'est pas évident que privilégier une formulation d'ordre deux soit le plus pertinent. En effet, pour effectuer les simulations, il est nécessaire de borner le domaine de calcul et la littérature sur le sujet va plus en faveur de la formulation du premier ordre pour laquelle on dispose de conditions PML (Perfectly Matched Layer) très efficaces. De plus, il s'avère que la formulation du premier ordre donne accès directement aux valeurs discrètes de  $\nabla p$  et de  $\text{div} \mathbf{v}$ ,  $p$  désignant la pression acoustique et  $\mathbf{v}$  le vecteur vitesse. Ces deux quantités sont particulièrement intéressantes pour appliquer la Reverse Time Migration (RTM) qui est la technique d'imagerie que nous proposons pour valider nos schémas numériques. Toutefois, comme nous utiliserons la méthode IPDG pour certains tests de validation, nous avons choisi de la présenter en même temps que la méthode DG (Discontinuous Galerkin) que nous utiliserons.

Ce chapitre est organisé de la façon suivante. Dans une première partie, nous rappelons les équations

des ondes acoustiques en formulation du premier et du second ordre. Puis dans une deuxième partie, nous présentons tout d'abord la DGM qui sera appliquée à une formulation du premier ordre de l'équation des ondes acoustiques. Ensuite, nous décrivons la méthode IPDG appliquée à l'équation des ondes acoustiques.

## 2.1 L'équation des ondes acoustiques

Nous rappelons ici deux formulations de l'équation des ondes acoustiques. La première est la formulation du second ordre. C'est probablement celle qui est la plus utilisée en pratique, car elle ne nécessite que le calcul d'une inconnue scalaire. La seconde est la formulation du premier ordre. Elle est plus coûteuse à mettre en oeuvre car elle nécessite, en plus du calcul de l'inconnue scalaire, le calcul d'une inconnue vectorielle.

### 2.1.1 Formulation standard de l'équation des ondes acoustiques

L'équation des ondes acoustiques décrit le comportement du champ de pression que nous noterons  $p$ . C'est un scalaire qui dépend du temps  $t$  et de la variable d'espace  $\mathbf{x}$ . Le milieu de propagation est caractérisé par une vitesse noté  $c$  qui dépend de la variable d'espace  $\mathbf{x}$  et s'exprime en fonction de la densité  $\rho$  et du module de compressibilité  $\mu$  :  $c = \sqrt{\frac{\mu}{\rho}}$ . Lorsque les coefficients sont variables, l'équation des ondes acoustiques s'écrit sous la forme suivante :

$$(2.1.1) \quad \frac{1}{\mu(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right) = f(\mathbf{x}, t)$$

Le terme  $f(\mathbf{x}, t)$  est un terme source. Nous supposerons lors de la présentation des différents schémas en espace et en temps que ce terme est très régulier. Typiquement, pour les schémas d'ordre élevé, nous devons supposer que  $f$  appartient à l'espace

$$C^2((0, T), L^2(\Omega)) \cap C^1((0, T), H^1(\Omega)) \cap C^0((0, T), H^2(\Omega)).$$

Cependant, dans la plupart des applications géophysiques, la composante en espace est une source ponctuelle telle que  $f(\mathbf{x}, t) = \delta(\mathbf{x}_0)g(t)$ , où  $\mathbf{x}_0$  est le point où la source est appliquée et  $g$  une fonction de  $C^\infty(0, T)$ . Cette fonction ne vérifie donc pas les conditions de régularité demandées. Nous discuterons de ce point dans l'annexe du chapitre 3, et nous expliquerons en détails comment nous prenons en compte ce type de source.

### 2.1.2 Formulation du premier ordre de l'équation des ondes acoustiques

L'équation des ondes acoustiques est généralement formulée comme une équation scalaire mais on peut aussi la formuler comme un système du premier ordre. Il faut pour cela introduire une autre inconnue notée  $\mathbf{v}$  qui représente le champ de vitesse. L'équation des ondes acoustiques du premier ordre s'écrit alors :

$$(2.1.2) \quad \begin{cases} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} + \nabla p(\mathbf{x}, t) = 0 & \text{dans } \Omega \times [0, T] \\ \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} + \nabla \cdot (\mathbf{v}(\mathbf{x}, t)) = f(\mathbf{x}, t) & \text{dans } \Omega \times [0, T] \end{cases}$$

Bien évidemment, les équations (2.1.1) ou le problème (2.1.2) ne suffisent pas pour décrire un phénomène de propagation d'ondes. D'une part, nous avons négligé les conditions initiales pour simplifier la présentation. D'autre part, il suffit généralement de considérer le phénomène dans une région limitée de l'espace, ce qui signifie que les équations peuvent être posées dans un milieu borné. Le modèle mathématique correspondant n'a alors des chances d'être bien posé que si on introduit des conditions aux limites adéquates. Nous parlerons précisément des conditions de bord à la section suivante.

## 2.2 Discrétisation en espace

Nous rappelons maintenant les schémas de type Galerkin Discontinus les plus couramment employés pour la discrétisation en espace des équations précédentes. Nous commençons par décrire la méthode DG avec flux centrés pour la formulation de l'équation des ondes du premier ordre. C'est la formulation qui est actuellement implémentée dans la plateforme DIVA de Total. Nous rappelons ensuite la formulation IPDG (Interior Penalty Discontinuous Galerkin Method) qui est adaptée à la formulation de l'équation des ondes du second ordre (sous-section 2.2.2).

### 2.2.1 Formulation DG avec flux centré pour l'équation des ondes acoustiques

La méthode DG que nous avons choisie a été proposée pour l'équation des ondes élastiques par S. Delcourte et al. [24]. Nous la présentons ici pour le système (2.1.2). Pour être dans le cadre des expériences numériques, on suppose que les calculs vont être réalisés dans un domaine borné que nous notons  $\Omega$  et que nous recouvrons par une collection  $\mathcal{T}_h$  d'éléments  $K$  qui sont des triangles à 2D, des tétraèdres à 3D. On note  $h_K$  le diamètre de l'élément  $K$  et  $h$  le plus petit de ces diamètres. On écrit d'abord que le système (2.1.2) est vérifié dans un élément  $K$  et on fait une formulation variationnelle du problème dans  $K$ . Cette formulation dépend du choix des fonctions test que nous allons faire.

#### 2.2.1.1 Formulation variationnelle continue

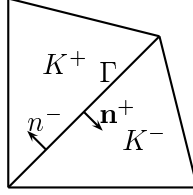
Soient  $q$  et  $\mathbf{w}$  les fonctions test qui sont définies dans tout le domaine de calcul  $\Omega$ . Pour l'instant, nous ne précisons pas dans quel espace elles sont, nous imposons seulement que ces fonctions sont localement continues dans  $\Omega$  au sens suivant. Si  $K^+$  et  $K^-$  désignent deux éléments contigus ayant pour côté (face) commun (commune)  $\Gamma$ , on note  $q^+$  et  $\mathbf{w}^+$  les fonctions  $q$  et  $\mathbf{w}$  restreintes à  $K^+$  et  $q^-$  et  $\mathbf{w}^-$  les fonctions  $q$  et  $\mathbf{w}$  restreintes à  $K^-$ . On introduit la quantité  $[[\cdot]]$ , définie sur  $\Gamma$ , désignant le saut de  $q$  ou  $\mathbf{w}$  à la traversée de  $\Gamma$ . En disant que  $q$  et  $\mathbf{w}$  sont localement continues dans  $\Omega$ , on veut dire que  $q^\pm$  et  $\mathbf{w}^\pm$  sont continues et que :

$$(2.2.1) \quad [[q]] = q^+ - q^- \neq 0 \text{ et } [[\mathbf{w}]] = \mathbf{w}^+ - \mathbf{w}^- \neq 0$$

La définition des sauts est fixée par l'orientation de la normale. Ainsi définis, cela sous-entend que la normale  $\mathbf{n}$  à la surface  $\Gamma$  est orientée de  $K^+$  vers  $K^-$  (voir figure 2.1). Considérons donc le système (2.1.2) que nous testons sur un élément  $K$  par les fonctions test  $q$  et  $\mathbf{w}$ . Il vient :

$$\begin{cases} \int_K \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \int_K \nabla p(\mathbf{x}, t) \cdot \mathbf{w}(\mathbf{x}) = 0 \\ \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \int_K q(\mathbf{x}) \nabla \cdot \mathbf{v}(\mathbf{x}, t) = \int_K f(\mathbf{x}, t) q(\mathbf{x}) \end{cases}$$



FIGURE 2.1 – Les éléments  $K^+$  et  $K^-$ .

On peut alors appliquer la formule de Green-Ostrogradski qui nous donne :

$$\left\{ \begin{array}{l} \int_K \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \int_{\partial K} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}_K - \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}, t) = 0 \\ \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \int_{\partial K} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_K q(\mathbf{x}) - \int_K \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) = \int_K f(\mathbf{x}, t) q(\mathbf{x}) \end{array} \right.$$

où  $\mathbf{n}_K$  représente le vecteur normal unitaire orienté extérieurement à  $K$ .

On peut alors obtenir une formulation variationnelle du problème en sommant sur l'ensemble des éléments. Par exemple, en sommant les contributions sur  $K^+$  et  $K^-$ , on décrit comment les éléments interagissent :

$$\left\{ \begin{array}{l} \int_{K^+} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \int_{\Gamma} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}_{K^+} - \int_{K^+} p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}, t) \\ + \int_{K^-} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \int_{\Gamma} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}_{K^-} - \int_{K^-} p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}, t) \\ + \int_{\partial(K^+ \cup K^-)} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}_{K^+} = 0 \\ \int_{K^+} \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_{K^+} q(\mathbf{x}) - \int_{K^+} \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ + \int_{K^-} \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_{K^-} q(\mathbf{x}) - \int_{K^-} \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ + \int_{\partial(K^+ \cup K^-)} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_{K^+} q(\mathbf{x}) = \int_{K^+ \cup K^-} f(\mathbf{x}, t) q \end{array} \right.$$

Sachant que sur  $\Gamma$  on a  $n_{K^-} = -n_{K^+}$ , on obtient :

$$\left\{ \begin{array}{l} \int_{K^+} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) - \int_{K^+} p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) \\ + \int_{K^-} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) - \int_{K^-} p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) \\ + \int_{\partial(K^+ \cup K^-)} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}_{K^+} + \int_{\Gamma} [[p(\mathbf{x}, t) \mathbf{w}(\mathbf{x})]] \cdot \mathbf{n}_{K^+} = 0 \\ \int_{K^+} \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) - \int_{K^+} \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ + \int_{K^-} \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) - \int_{K^-} \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ + \int_{\partial(K^+ \cup K^-)} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_{K^+} q(\mathbf{x}) + \int_{\Gamma} [[\mathbf{v}(\mathbf{x}, t) q]] \cdot \mathbf{n}_{K^+} = \int_{K^+ \cup K^-} f(\mathbf{x}, t) q. \end{array} \right.$$

En sommant sur l'ensemble des triangles, on peut ainsi reformuler les intégrales sur les surfaces internes à partir des termes de saut. Dans le cas où  $\Gamma$  est une surface située sur la frontière  $\partial\Omega$  de  $\Omega$ , il ne vient bien évidemment aucune contribution de l'extérieur de  $\Omega$ . Avant d'écrire la formulation variationnelle globale il nous faut donc distinguer ces deux types de surface. Nous noterons  $\Gamma_{int}$  l'ensemble des surfaces internes, c'est-à-dire celles qui sont communes à deux éléments du maillages, et  $\Gamma_{ext}$  l'ensemble des surfaces externes, c'est-à-dire celles qui n'appartiennent qu'à un seul élément du maillage situé au bord du domaine. La formulation variationnelle s'écrit alors

$$\left\{ \begin{array}{l} \sum_K \int_K \rho(\mathbf{x}) \frac{\partial v(\mathbf{x}, t)}{\partial t} \mathbf{w}(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[p(\mathbf{x}, t) \mathbf{w}(\mathbf{x})]] \cdot \mathbf{n} \\ + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) = 0 \\ \sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[\mathbf{v}(\mathbf{x}, t) q(\mathbf{x})]] \cdot \mathbf{n} \\ + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}, t) q. \end{array} \right.$$

Dans le cas où  $\Gamma$  est une surface externe,  $\mathbf{n}$  désigne le vecteur normal à  $\Gamma$  dirigé vers l'extérieur de  $\Omega$ .

Nous allons maintenant intégrer les conditions limites dans la formulation variationnelle. Typiquement, on impose soit une condition de Dirichlet, soit une condition de Neumann, soit une condition absorbante.

- Pour une condition aux limites de type Dirichlet, c'est-à-dire  $p = 0$  sur  $\partial\Omega$  : on utilise cette condition pour annuler le terme  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}$ .
- Pour une condition aux limites de type Neumann, c'est-à-dire  $\mathbf{v} \cdot \mathbf{n} = 0$  sur  $\partial\Omega$  : on utilise cette condition pour annuler le terme  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n}$ .
- Pour une condition aux limites absorbante d'ordre 1, de la forme

$$\frac{\partial p}{\partial t} + \sqrt{\frac{\mu}{\rho}} \frac{\partial p}{\partial \mathbf{n}} = 0,$$

on utilise l'équation des ondes (2.1.2) pour reformuler cette condition :

$$\frac{\partial p}{\partial t} - \sqrt{\mu\rho} \frac{\partial \mathbf{v}}{\partial t} \cdot \mathbf{n} = 0.$$

Cette condition devient après intégration en temps, et en supposant que les données initiales sont nulles sur le bord du domaine :

$$p - \sqrt{\mu\rho} \mathbf{v} \cdot \mathbf{n} = 0.$$

Nous avons alors plusieurs possibilités pour intégrer cette condition dans la formulation variationnelle. On peut

1. Remplacer le terme  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} p(\mathbf{x}, t) \mathbf{w}(\mathbf{x}) \cdot \mathbf{n}$  par  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \sqrt{\mu(\mathbf{x})\rho(\mathbf{x})} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}$
2. Remplacer le terme  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n}$  par  $\sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \frac{1}{\mu(\mathbf{x})\rho(\mathbf{x})} p(\mathbf{x}) q(\mathbf{x})$ ;

3. Effectuer à la fois 1. et 2.

Nous avons testé les trois possibilités et nous n'avons pas observé de différences sur les résultats numériques.

Il est aussi pratique d'appliquer une condition aux limites périodiques. Ce dernier cas revient à considérer que les arêtes externes sont en fait des arêtes internes assurant la communication entre les éléments sur le bord gauche (resp. bas) du domaine et les éléments sur le bord droit du domaine (resp. haut). C'est au moment des tests numériques que nous précisons quelle condition est utilisée.

Dans la suite de ce chapitre, nous considérons des conditions de Dirichlet et la formulation variationnelle s'écrit alors

$$\left\{ \begin{array}{l} \sum_K \int_K \rho(\mathbf{x}) \frac{\partial v(\mathbf{x}, t)}{\partial t} \mathbf{w}(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[p(\mathbf{x}, t) \mathbf{w}(\mathbf{x})]] \cdot \mathbf{n} - \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) = 0 \\ \sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[\mathbf{v}(\mathbf{x}, t) q(\mathbf{x})]] \cdot \mathbf{n} \\ + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}, t) q. \end{array} \right.$$

Cette formulation n'est pas complètement satisfaisante car elle n'est pas symétrique. C'est-à-dire que

$$\begin{aligned} & \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[\mathbf{v}(\mathbf{x}, t) p(\mathbf{x})]] \cdot \mathbf{n} + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) p(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(\mathbf{x}, t) \cdot \nabla p(\mathbf{x}) \neq \\ & - \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[p(\mathbf{x}, t) \mathbf{v}(\mathbf{x})]] \cdot \mathbf{n} + \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{v}(\mathbf{x}). \end{aligned}$$

Cette propriété est essentielle pour démontrer la conservation d'une énergie, comme nous le verrons par la suite. Toutefois, en faisant des hypothèses sur la régularité du champ d'ondes solution du problème continu, on peut obtenir une formulation symétrique. Nous supposons par la suite que les données initiales et la source sont suffisamment régulières pour que la solution du problème continu soit globalement continue dans  $\Omega$ . Dans ce cas, on a :

$$(2.2.2) \quad [[p]] = p^+ - p^- = 0 \text{ et } [[\mathbf{v}]] \cdot \mathbf{n} = (\mathbf{v}^+ - \mathbf{v}^-) \cdot \mathbf{n} = 0$$

De plus, on a la relation suivante :

$$(2.2.3) \quad ab - cd = (b - d) \left( \frac{a + c}{2} \right) + (a - c) \left( \frac{b + d}{2} \right)$$

Si on note  $\{\{q\}\}$  et  $\{\{\mathbf{w}\}\}$  les moyennes de  $q$  et  $\mathbf{w}$  qui sont définies par :

$$\{\{q\}\} = \frac{q^+ + q^-}{2} \text{ et } \{\{\mathbf{w}\}\} = \frac{\mathbf{w}^+ + \mathbf{w}^-}{2}$$

l'équation (2.2.3) implique la relation :

$$(2.2.4) \quad [[\mathbf{v}(\mathbf{x}, t) q(\mathbf{x})]] \cdot \mathbf{n} = [[q(\mathbf{x})]] \{\{\mathbf{v}(\mathbf{x}, t)\}\} \cdot \mathbf{n} + [[\mathbf{v}(\mathbf{x}, t)]] \{\{q(\mathbf{x})\}\} \cdot \mathbf{n}$$

qui devient en utilisant (2.2.2) :

$$[[\mathbf{v}(\mathbf{x}, t) q(\mathbf{x})]] = \{\{\mathbf{v}(\mathbf{x}, t)\}\} [[q(\mathbf{x})]].$$

De la même façon, on obtient aussi que

$$[[p(\mathbf{x}, t)\mathbf{w}(\mathbf{x})]] = \{\{p(\mathbf{x}, t)\}\} [[\mathbf{w}(\mathbf{x})]]$$

et la formulation devient alors :

$$(2.2.5) \quad \left\{ \begin{array}{l} \sum_K \int_K \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{\{p(\mathbf{x}, t)\}\} [[\mathbf{w}(\mathbf{x})]] \cdot \mathbf{n} - \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) = 0 \\ \sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{\{\mathbf{v}(\mathbf{x}, t)\}\} [[q(\mathbf{x})]] \cdot \mathbf{n} \\ + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}, t) q. \end{array} \right.$$

Écrite sous cette forme, la symétrie de la formulation n'est pas évidente. Pour la démontrer on vérifie, avec des calculs similaires aux calculs précédents, que

$$\begin{aligned} \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) &= \\ \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} [[\mathbf{w}(\mathbf{x}, t)p(\mathbf{x})]] \cdot \mathbf{n} + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{w}(\mathbf{x}, t)p(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{w}(\mathbf{x}, t) \cdot \nabla p(\mathbf{x}). \end{aligned}$$

En utilisant 2.2.4, on alors

$$(2.2.6) \quad \begin{aligned} \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{\{p(\mathbf{x}, t)\}\} [[\mathbf{w}(\mathbf{x})]] \cdot \mathbf{n} - \sum_K \int_K p(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) &= \\ - \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{\{\mathbf{w}(\mathbf{x})\}\} [[p(\mathbf{x})]] \cdot \mathbf{n} - \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{w}(\mathbf{x})p(\mathbf{x}) \cdot \mathbf{n} + \sum_K \int_K \mathbf{w}(\mathbf{x}) \cdot \nabla p(\mathbf{x}). \end{aligned}$$

Nous pouvons maintenant démontrer la conservation d'une énergie.

**Propriété 2.2.1.** *En l'absence de terme source, la solution  $(p, \mathbf{v})$  de (2.2.5) vérifie la conservation de l'énergie*

$$\sum_K \int_K \frac{1}{\mu(\mathbf{x})} |p(\mathbf{x}, t)|^2 + \sum_K \int_K \rho(\mathbf{x}) \|\mathbf{v}(\mathbf{x}, t)\|^2$$

**Démonstration :** En choisissant comme fonctions test  $q(\mathbf{x}) = p(\mathbf{x}, t)$  et  $\mathbf{w}(\mathbf{x}) = \mathbf{v}(\mathbf{x}, t)$  dans (2.2.5) à chaque instant  $t$ , en additionnant les deux équations et en utilisant (2.2.6), on obtient

$$\sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} p(\mathbf{x}, t) + \sum_K \int_K \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} \cdot \mathbf{v}(\mathbf{x}, t) = 0,$$

ce qui peut se réécrire

$$\frac{1}{2} \frac{d}{dt} \left( \sum_K \int_K \frac{1}{\mu(\mathbf{x})} |p(\mathbf{x}, t)|^2 + \sum_K \int_K \rho(\mathbf{x}) \|\mathbf{v}(\mathbf{x}, t)\|^2 \right) = 0.$$

■

### 2.2.1.2 Formulation variationnelle discrète

Il nous reste maintenant à définir l'espace d'approximation retenu pour discrétiser la formulation variationnelle. Il s'écrit sous la forme

$$V_h^m = \{ \phi \in L(\Omega) \mid \phi|_K \in \mathcal{P}_m(K) \},$$

où  $\mathcal{P}_m(K)$  désigne l'espace des polynômes de degré inférieur ou égal à  $m$  sur  $K$ . La formulation variationnelle discrète s'écrit alors :

Trouver  $(p_h, \mathbf{v}_h)$  dans  $C^1((0, T), V_h^m \times (V_h^m)^2)$  tel que, pour tout  $(q, \mathbf{w})$  dans  $V_h^m \times (V_h^m)^2$ , on ait

$$(2.2.7) \quad \left\{ \begin{array}{l} \sum_K \int_K \rho(\mathbf{x}) \frac{\partial \mathbf{v}_h(\mathbf{x}, t)}{\partial t} \cdot \mathbf{w}(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{ \{ p_h(\mathbf{x}, t) \} \} [[\mathbf{w}(\mathbf{x})]] \cdot \mathbf{n} - \sum_K \int_K p_h(\mathbf{x}, t) \nabla \cdot \mathbf{w}(\mathbf{x}) = 0 \\ \sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial p_h(\mathbf{x}, t)}{\partial t} q(\mathbf{x}) + \sum_{\Gamma \in \Gamma_{int}} \int_{\Gamma} \{ \{ \mathbf{v}_h(\mathbf{x}, t) \} \} [[q(\mathbf{x})]] \cdot \mathbf{n} \\ + \sum_{\Gamma \in \Gamma_{ext}} \int_{\Gamma} \mathbf{v}_h(\mathbf{x}, t) q(\mathbf{x}) \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}_h(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}, t) q. \end{array} \right.$$

On montre, de la même façon que pour la formulation variationnelle continue, que l'énergie

$$\sum_K \int_K \frac{1}{\mu(\mathbf{x})} |p_h(\mathbf{x}, t)|^2 + \sum_K \int_K \rho(\mathbf{x}) \|\mathbf{v}_h(\mathbf{x}, t)\|^2$$

est conservée. Cette propriété est essentielle pour démontrer la convergence du schéma. Nous renvoyons aux travaux de Fezoui *et al* [34] pour l'étude détaillée de cette convergence. Les auteurs montrent en particulier que

$$\|p_h - p\|_{C^0((0, T), L^2(\Omega))} + \|\mathbf{v}_h - \mathbf{v}\|_{C^0((0, T), L^2(\Omega))} \leq CTh^{\min(s, m)} \left( \|p\|_{C^0((0, T), \mathcal{V}_h^s)} + \|\mathbf{v}\|_{C^0((0, T), V_h)} \right),$$

où  $\mathcal{V}_h^s$  est défini par

$$\mathcal{V}_h^s = \{ \phi \in L^2(\Omega) \mid \phi|_K \in H^s(K) \}.$$

Ainsi, si la solution du problème est suffisamment régulière, c'est-à-dire si  $s \geq m$ , la solution discrète converge à l'ordre  $m$ . Cette convergence est sous-optimale car elle est d'ordre inférieur à l'ordre de convergence de l'erreur d'approximation  $\|\Pi_h \phi - \phi\|_{L^2(\Omega)}$ , où  $\Pi_h \phi$  représente la projection de  $p$  sur  $V_h^m$ . L'erreur d'approximation converge en effet à l'ordre  $m + 1$  [16]. Nous perdons un ordre de précision par rapport aux méthodes d'éléments finis classiques ou à d'autres méthodes DG comme la méthode IPDG ou la méthode DG avec flux upwind. Cependant, les méthodes d'éléments finis classiques ne permettent pas d'obtenir des schémas explicites en temps, la méthode IPDG est spécifique aux équations des ondes en formulation d'ordre deux et la méthode DG upwind est dissipative. Dans [14, 15], Engquist et Chung ont proposé une technique permettant d'obtenir l'ordre de convergence optimal et la conservation de l'énergie. Cette méthode se base également sur la formulation variationnelle (2.2.7), mais seul  $p_h$  est cherché dans l'espace  $V_h^m$ . L'inconnue  $\mathbf{v}_h$  est cherchée dans un autre espace construit à partir d'un maillage dual de  $\mathcal{T}_h$ . Les fonctions  $p_h$  et  $\mathbf{v}_h$  ne sont alors plus discontinues sur les mêmes arêtes. L'inconvénient de cette formulation est la complexité de sa mise en œuvre, notamment dans un code tridimensionnel.

Nous discuterons plus en détails de l'ordre de convergence de la méthode lors des chapitres 3 et 4 consacrés respectivement aux schémas d'ordre élevé et à l'analyse de dispersion.

Pour définir une base de l'espace  $V_h^m$ , nous procédons élément par élément, à partir des fonctions de base de Lagrange définies sur un élément de référence  $\hat{K}$ . Par exemple, en 2D, cet élément est le triangle  $(0,0)$   $(0,1)$   $(1,0)$ . Sur ce triangle, nous définissons les  $(m+1)(m+2)/2$  points  $\mathbf{x}_i = (x_j, x_k)_{j+k=1..m+1}$ , avec  $x_j = (j-1)/m$  (voir Fig. 2.2). La numérotation des degrés de liberté, c'est-à-dire le choix de la bijection  $(j, k) \mapsto i$  est arbitraire. Nous choisissons de numérotter d'abord les trois sommets, puis les degrés de liberté sur les arêtes et enfin les degrés de liberté internes. Un exemple de numérotation en 2D dans le cas  $m = 3$  est donné sur la figure 2.2. Les fonctions de base de Lagrange de  $\mathcal{P}^m$  sont les polynômes de degré  $m$   $(\hat{\phi}_i)_{i=1..(m+1)(m+2)/2}$ , tels que  $\hat{\phi}_i(x_j) = \delta_{ij}$ , où  $\delta_{ij}$  représente le symbole de Kronecker. Pour construire les fonctions de base sur un élément  $K$ , nous utilisons une fonction linéaire  $F_K$  qui transforme l'élément  $\hat{K}$  en l'élément  $K$  et nous définissons  $\phi_i^K(\mathbf{x}) = \hat{\phi}_i \circ F_K^{-1}$  pour  $\mathbf{x} \in K$  et  $\phi_i^K(\mathbf{x}) = 0$  sinon. Les images des degrés de liberté par les fonctions  $F_{K^+}$  et  $F_{K^-}$  sont représentées sur la figure 2.2 pour  $m = 3$ . En dimension 2 par exemple, nous

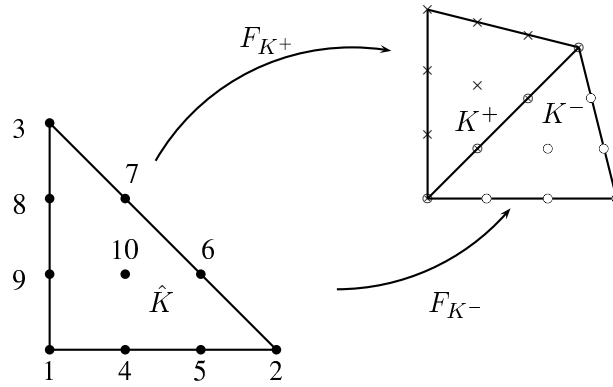


FIGURE 2.2 – Les degrés de liberté pour des EF  $\mathcal{P}^3$  sur les éléments  $K^+$  et  $K^-$ .

pouvons représenter  $p_h$  et  $\mathbf{v}_h$  sous la forme

$$p_h = \sum_K \sum_i P_i^K(t) \phi_i^K \text{ et } \mathbf{v}_h = \sum_K \sum_i V_{x_i}^K(t) \phi_i^K e_x + \sum_K \sum_i V_{y_i}^K(t) \phi_i^K e_y,$$

où  $(e_x, e_y)$  est la base de canonique de  $\mathbb{R}^2$ . En choisissant les fonctions de base comme fonctions test, on obtient le système d'équations différentielles ordinaires suivant :

$$(2.2.8) \quad \begin{cases} \mathcal{M}_v \frac{d\mathbf{V}}{dt} + \mathcal{K} \mathbf{P} = 0 \\ \mathcal{M}_p \frac{d\mathbf{P}}{dt} - \mathcal{K}^T \mathbf{V} = \mathbf{F}(t) \end{cases}$$

où les vecteurs  $\mathbf{P}$  et  $\mathbf{V}$  contiennent les valeurs de  $p_h$  et  $\mathbf{v}_h$  aux degrés de liberté :  $\mathbf{P} = (P_i)_{i=1..N}$  et  $\mathbf{V} = (\mathbf{V}_i)_{i=1..N}$ , où  $N$  désigne le nombre total de degrés de liberté. Les matrices de masses  $\mathcal{M}_p$  et  $\mathcal{M}_v$  sont facilement inversibles car elles sont diagonales par bloc par construction et la matrice de rigidité  $\mathcal{K}$  est creuses. Le terme  $\mathbf{F}(t)$  est le terme source. Plus précisément, en dimension deux, ce

système s'écrit

$$(2.2.9) \quad \begin{cases} \mathcal{M}_{v_x} \frac{d\mathbf{V}_x}{dt} + \mathcal{K}_x \mathbf{P} = 0 \\ \mathcal{M}_{v_y} \frac{d\mathbf{V}_y}{dt} + \mathcal{K}_y \mathbf{P} = 0 \\ \mathcal{M}_p \frac{d\mathbf{P}}{dt} - \mathcal{K}_x^T \mathbf{V}_x - \mathcal{K}_y^T \mathbf{V}_y = \mathbf{F}(t). \end{cases}$$

Les méthodes DG étant naturellement définies élément par élément, il est commode d'utiliser une structure par blocs pour définir les vecteurs et les matrices. Ainsi le vecteur  $\mathbf{P}$  peut-être représenté par l'ensemble des blocs  $\mathbf{P}^K$ ,  $K \in \mathcal{T}_h$ , et chaque bloc  $\mathbf{P}^K$  a pour composante  $(P_i^K)_{i=1..(m+1)(m+2)/2}$ , où  $i$  est un des degrés de liberté de l'élément  $K$ . De la même façon, les vecteurs  $\mathbf{V}_x$  et  $\mathbf{V}_y$  peuvent être décrit par des blocs  $\mathbf{V}_x^K$  et  $\mathbf{V}_y^K$ . Une matrice  $\mathcal{A}$  est décrite par des blocs  $\mathcal{A}^{K,J}$ .

Les matrices de masse  $\mathcal{M}_p$ ,  $\mathcal{M}_{v_x}$  et  $\mathcal{M}_{v_y}$  sont telles que  $\mathcal{M}_p^{K,J} = \mathcal{M}_{v_x}^{K,J} = \mathcal{M}_{v_y}^{K,J} = 0$  si  $K \neq J$  et

$$\mathcal{M}_{p_i,j}^{K,K} = \int_K \frac{1}{\mu(\mathbf{x})} \phi_i^K \phi_j^K \quad \text{et} \quad \mathcal{M}_{v_x i,j}^{K,K} = \mathcal{M}_{v_y i,j}^{K,K} = \int_K \rho(\mathbf{x}) \phi_i^K \phi_j^K.$$

Les matrices de raideurs  $\mathcal{K}_x$  et  $\mathcal{K}_y$  sont telles que  $\mathcal{K}_x^{K,J} = \mathcal{K}_y^{K,J} = 0$  si  $K \cap J = \emptyset$ ,

$$\mathcal{K}_{x i,j}^{K,J} = -\frac{1}{2} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_x^K \quad \text{et} \quad \mathcal{K}_{y i,j}^{K,J} = -\frac{1}{2} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_y^K,$$

si  $K \cap J = \Gamma$ , autrement dit, si  $K$  et  $J$  sont voisins et

$$\begin{aligned} \mathcal{K}_{x i,j}^{K,K} &= \frac{1}{2} \sum_{\Gamma \in \Gamma_{\text{int}} \cap \partial K} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_x^K + \sum_{\Gamma \in \Gamma_{\text{ext}} \cap \partial K} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_x^K - \int_K \phi_j^K \cdot \nabla \phi_i^K, \\ \mathcal{K}_{y i,j}^{K,K} &= \frac{1}{2} \sum_{\Gamma \in \Gamma_{\text{int}} \cap \partial K} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_y^K + \sum_{\Gamma \in \Gamma_{\text{ext}} \cap \partial K} \int_{\Gamma} \phi_i^K \phi_j^K \cdot \mathbf{n}_y^K - \int_K \phi_j^K \cdot \nabla \phi_i^K. \end{aligned}$$

Le vecteur  $\mathbf{F}$  est tel que  $\mathbf{F}_i^K = \int_{\Omega} f(\mathbf{x}, t) \phi_i^K$ .

Cette formulation étant équivalente à la formulation (2.2.7), on vérifie facilement qu'on a la conservation de l'énergie définie par la norme  $L^2$  du graphe, c'est-à-dire

$$\|\mathbf{V}\|_{M_v}^2 + \|\mathbf{P}\|_{M_p}^2.$$

## 2.2.2 Formulation IPDG

La formulation IPDG (Interior Penalty Discontinuous Galerkin) est obtenue à partir de la formulation du second ordre des équations de l'acoustique. Pour plus de détails sur cette méthode, nous renvoyons à [28, 6]. On trouvera dans [4] une analyse de dispersion et dans [41, 38, 39] une étude des propriétés de convergence. On pourra également se référer à [7] pour une comparaison numérique des performances de IPDG par rapport aux méthodes d'éléments spectraux.

Pour obtenir la formulation variationnelle associée à IPDG, on multiplie (2.1.1) par une fonction test  $q$ , localement continue, on intègre sur  $K$  et on somme sur tous les éléments  $K$ . On obtient alors

$$(2.2.10) \quad \begin{aligned} & \sum_K \int_K \frac{q(\mathbf{x})}{\mu(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} + \sum_K \int_K \frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ & - \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left[ \frac{q(\mathbf{x})}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right] \right] \cdot \mathbf{n} - \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \frac{q(\mathbf{x})}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \cdot \mathbf{n} = \int_{\Omega} f(\mathbf{x}, t) q. \end{aligned}$$

Comme pour la formulation du premier ordre, on n'obtient pas immédiatement une formulation symétrique. Pour la symétriser, on va utiliser la relation

$$(2.2.11) \quad \left[ \left[ \frac{q}{\rho} \nabla p \right] \right] = \llbracket q \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla p \right\} \right\} + \{\{q\}\} \left[ \left[ \frac{1}{\rho} \nabla p \right] \right]$$

On suppose que les données initiales et la source sont suffisamment régulières pour que les conditions de transmission classiques, à savoir,

$$(2.2.12) \quad \llbracket p \rrbracket = 0 \quad \text{et} \quad \left[ \left[ \frac{1}{\rho} \nabla p \cdot \mathbf{n} \right] \right] = 0,$$

soient vérifiées. La seconde condition nous permet de supprimer le terme  $\{\{q\}\} \left[ \left[ \frac{1}{\rho} \nabla p \right] \right]$  et la première nous permet d'ajouter le terme  $\llbracket p \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla q \right\} \right\}$  de telle sorte que :

$$(2.2.13) \quad \left[ \left[ q \frac{1}{\rho} \nabla p \right] \right] = \llbracket q \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla p \right\} \right\} + \{\{p\}\} \left[ \left[ \frac{1}{\rho} \nabla q \right] \right].$$

Il reste à considérer les termes de bord posés sur  $\Gamma_{\text{ext}}$  :

$$\sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \frac{q(\mathbf{x})}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \cdot \mathbf{n}.$$

Si on impose des conditions de Neumann, il n'y a pas de difficulté car  $\nabla p(\mathbf{x}, t) \cdot \mathbf{n} = 0$  et ce terme s'annule. La condition absorbante d'ordre un ne pose pas plus de difficulté, on utilise simplement son expression et on a

$$\int_{\Gamma} \frac{q(\mathbf{x})}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \cdot \mathbf{n} = \int_{\Gamma} \mu(\mathbf{x}) q(\mathbf{x}) \frac{\partial p(\mathbf{x}, t)}{\partial t}.$$

Dans le cas d'une condition de Dirichlet, qui est la condition que nous considérons dans ce chapitre, nous ajoutons le terme

$$\sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \frac{p(\mathbf{x})}{\rho(\mathbf{x})} \nabla q(\mathbf{x}, t) \cdot \mathbf{n}$$

pour symétriser la formulation. Ce terme ne modifie pas la formulation de départ car la solution du problème continu vérifie  $p = 0$  sur le bord  $\Gamma_{\text{ext}}$ .

Nous considérerons par la suite des conditions de Dirichlet. La formulation variationnelle s'écrit alors

$$\begin{aligned} & \sum_K \int_K \frac{1}{\mu(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} q(\mathbf{x}) + \sum_K \int_K \frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \cdot \nabla q(\mathbf{x}) \\ & - \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left( \llbracket q(\mathbf{x}) \rrbracket \left\{ \left\{ \frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) \right\} \right\} + \llbracket p(\mathbf{x}, t) \rrbracket \left\{ \left\{ \frac{1}{\rho(\mathbf{x})} \nabla q(\mathbf{x}) \right\} \right\} \right) \cdot \mathbf{n} \\ & - \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \left( q(\mathbf{x}) \frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}, t) + p(\mathbf{x}, t) \frac{1}{\rho(\mathbf{x})} \nabla q(\mathbf{x}) \right) \cdot \mathbf{n} = \int_{\Omega} f(\mathbf{x}, t) q. \end{aligned}$$

Cette nouvelle formulation est bien symétrique, mais la forme bilinéaire associée n'est pas coercive [28, 6]. Pour assurer la coercivité, nous proposons d'ajouter un terme de pénalisation  $\sum_{\Gamma \in \Gamma_{\text{int}}} \int \gamma \llbracket p \rrbracket \llbracket q \rrbracket$ ,



en utilisant à nouveau l'hypothèse  $[[p]] = 0$ . La fonction de pénalisation  $\gamma$  est définie sur chaque face interne par

$$\gamma = \frac{\alpha}{\xi_F}$$

où  $\alpha$  est un paramètre positif. On peut trouver plusieurs définitions de la fonction  $\xi_F$  dans la littérature. Les plus couramment utilisées sont :

- $\xi_F = h(F)$  où  $h(F)$  est le diamètre de  $F$ . C'est la définition utilisée par exemple dans [6, 4, 42]. Remarquons que cette définition n'a pas de sens en dimension 1.
- $\xi_F = \min(h(K^+), h(K^-))$  où  $h(K^\pm)$  est le diamètre de  $K^\pm$ . C'est la définition utilisée par exemple dans [41, 38, 39].
- $\xi_F = \min(\rho(K^+), \rho(K^-))$  où  $\rho(K^\pm)$  est le diamètre du cercle (ou de la sphère) inscrit(e) de  $K^\pm$ . C'est la définition utilisée par exemple dans [59].

L'influence de  $\xi_F$  sur la stabilité de la méthode est discutée dans [1, 3]. Il existe une valeur  $\alpha_0$ , dépendant de la définition de  $\xi_F$ , telle que la stabilité du schéma soit garantie pour tout  $\alpha > \alpha_0$ . Quand nous introduirons la discrétisation en temps de cette formulation, nous verrons notamment que la CFL du schéma numérique décroît quand  $\alpha$  augmente.

Toujours dans le cas d'une condition de Dirichlet, nous ajoutons la pénalisation  $2 \sum_{\Gamma \in \Gamma_{\text{ext}}} \int \gamma p q$  qui ne perturbe pas la condition initiale car la solution du problème continu vérifie  $p = 0$  sur le bord.

Finalement, la formulation variationnelle IPDG continue s'écrit

$$(2.2.14) \quad \sum_K \int_K \frac{1}{\mu} \frac{\partial^2 p}{\partial t^2} q + a(p, q) = \sum_K \int_K f(\mathbf{x}, t) q,$$

où la forme bilinéaire  $a$  définie par

$$\begin{aligned} a(p, q) = & \sum_K \int_K \frac{1}{\rho} \nabla p \cdot \nabla q - \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left( [[q]] \left\{ \left\{ \frac{1}{\rho} \nabla p \right\} \right\} + [[p]] \left\{ \left\{ \frac{1}{\rho} \nabla q \right\} \right\} \right) \cdot \mathbf{n} - \gamma [[p]] [[q]] \right] \\ & - \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \left[ \left( q \frac{1}{\rho} \nabla p + p \frac{1}{\rho} \nabla q \right) \cdot \mathbf{n} - \gamma p q \right] \end{aligned}$$

est symétrique et coercive pour  $\alpha$  suffisamment grand.

Cette formulation préserve également une énergie. Pour le vérifier au moins formellement (on suppose que la solution est assez régulière pour pouvoir effectuer tous les calculs), on pose  $q = \frac{\partial p}{\partial t}$  dans (2.2.14), pour obtenir

$$\sum_K \int_K \frac{1}{\mu} \frac{\partial^2 p}{\partial t^2} \frac{\partial p}{\partial t} + a\left(p, \frac{\partial p}{\partial t}\right) = \sum_K \int_K f(\mathbf{x}, t) \frac{\partial p}{\partial t}.$$

Comme  $a$  est bilinéaire et symétrique, on a

$$a\left(p, \frac{\partial p}{\partial t}\right) = \frac{1}{2} \frac{d}{dt} a(p, p).$$

En l'absence de terme source, on obtient donc la relation

$$\frac{d}{dt} \left( \sum_K \int_K \left| \frac{\partial p}{\partial t} \right|^2 + a(p, p) \right) = 0,$$

et la coercivité de  $a$  nous garantit la positivité de l'énergie

$$E(t) = \sum_K \int_K \left| \frac{\partial p}{\partial t} \right|^2 + a(p, p).$$

Pour la discrétisation de la formulation (2.2.14) on utilisera l'espace d'approximation  $V_h^m$  défini à la section précédente. La formulation IPDG discrète s'écrit donc :

Trouver  $p_h$  dans  $C^1((0, T), V_h^m)$  tel que, pour tout  $q$  dans  $V_h^m$ , on ait

$$(2.2.15) \quad \sum_K \int_K \frac{1}{\mu} \frac{\partial^2 p_h}{\partial t^2} q + a(p_h, q) = \sum_K \int_K f(\mathbf{x}, t) q.$$

On vérifie qu'en l'absence de terme source on a bien la conservation de l'énergie discrète

$$E_h(t) = \sum_K \int_K \left| \frac{\partial p_h}{\partial t} \right|^2 + a(p_h, p_h).$$

Dans [41], Grote, Schneebeli et Schötzau montrent le théorème suivant.

**Théorème 2.2.2.** *On suppose que le domaine  $\Omega$  et les fonctions  $\rho$  et  $\mu$  sont suffisamment régulières pour assurer l'hypothèse de régularité elliptique. C'est-à-dire que pour tout  $\lambda \in L^2(\Omega)$ , la solution  $u$  du problème*

$$-\nabla \cdot (\sqrt{\mu\rho} \nabla z) = \lambda \text{ in } \Omega, \quad z = 0 \text{ on } \Gamma,$$

*appartient à  $H^2(\Omega)$  et il existe une constante de stabilité  $C_S$  telle que*

$$\|z\|_2 \leq C_S \|\lambda\|_0;$$

*Si la solution  $p$  de (2.2.14) est telle que*

$$p \in L^\infty((0, T), H^{1+s}(\Omega)), \quad \frac{\partial u}{\partial t} \in L^\infty((0, T), H^{1+s}(\Omega)), \quad \frac{\partial^2 u}{\partial t^2} \in L^\infty((0, T), H^s(\Omega)),$$

*pour  $s > \frac{1}{2}$  et si  $\alpha$  est suffisamment grand ;*

*alors la solution  $p_h$  de (2.2.15) vérifie*

$$\|p - p_h\|_{L^\infty((0, T), L^2(\Omega))} \leq Ch^{\min(s, m)+1} \left[ \|u\|_{L^\infty((0, T), H^{1+s}(\Omega))} + T \left\| \frac{\partial u}{\partial t} \right\|_{L^\infty((0, T), H^{1+s}(\Omega))} \right].$$

Contrairement au schéma DG avec flux centré appliqué à la formulation du premier ordre, on obtient ici la convergence optimale d'ordre  $m + 1$  pour une solution suffisamment régulière.

En utilisant les fonctions de base de Lagrange classiques, comme à la section précédente, on obtient le schéma semi-discrétisé

$$(2.2.16) \quad \mathcal{M}_p \frac{d^2 \mathbf{P}}{dt^2} + \mathcal{K}_{IP} \mathbf{P} = \mathbf{F}(t)$$

où la matrice de masse  $\mathcal{M}_p$  est identique à celle définie pour le schéma DG d'ordre un. La matrice de rigidité  $\mathcal{K}_{IP}$  est creuse et a un stencil similaire à celui du schéma DG d'ordre un, c'est-à-dire qu'un élément communique avec lui-même et avec ses voisins. Elle est telle que  $\mathcal{K}_{IP}^{K, J} = 0$  si  $K \cap J = \emptyset$ ,

$$\mathcal{K}_{IP}^{K, J} = \frac{1}{2} \int_\Gamma \left( \frac{1}{\rho} \nabla \phi_j^J \phi_i^K + \frac{1}{\rho} \phi_j^J \nabla \phi_i^K \right) \cdot \mathbf{n} - \int_\Gamma \gamma \phi_j^J \phi_i^K,$$

si  $K \cap J = \Gamma$ , et

$$\begin{aligned} \mathcal{K}_{IP_{i,j}^{K,K}} = & \int_K \frac{1}{\rho} \nabla \phi_j^J \cdot \nabla \phi_i^K - \sum_{\Gamma \in \Gamma_{int} \cap \partial K} \left[ \frac{1}{2} \left( \frac{1}{\rho} \nabla \phi_j^J \phi_i^K + \frac{1}{\rho} \phi_j^J \nabla \phi_i^K \right) \cdot \mathbf{n} - \int_{\Gamma} \gamma \phi_j^J \phi_i^K \right] \\ & - \sum_{\Gamma \in \Gamma_{ext} \cap \partial K} \left[ \left( \frac{1}{\rho} \nabla \phi_j^J \phi_i^K + \frac{1}{\rho} \phi_j^J \nabla \phi_i^K \right) \cdot \mathbf{n} - 2 \int_{\Gamma} \gamma \phi_j^J \phi_i^K \right]. \end{aligned}$$

Le vecteur  $\mathbf{F}(t)$  est tel que  $\mathbf{F}_i^K(t) = \int_{\Omega} f(\mathbf{x}, t) \phi_i^K$ .

Cette formulation étant équivalente à la formulation (2.2.15), on vérifie facilement qu'on a la conservation de l'énergie

$$\left\| \frac{d\mathbf{P}}{dt} \right\|_{M_p}^2 + \|\mathbf{P}\|_{\mathcal{K}_{IP}}^2.$$

## Chapitre 3

# Schémas d'ordre élevé en temps

Dans le chapitre 2, nous avons présenté deux méthodes de Galerkin Discontinue qui sont particulièrement intéressantes pour résoudre des équations d'ondes posées dans des milieux hétérogènes. En effet, les DGM sont parfaitement adaptées au calcul parallèle et elles disposent d'une assez grande flexibilité qui autorise notamment la combinaison de différents ordres d'approximation pour résoudre le problème. Cependant, ces méthodes ne permettent de capturer que les variations en espace des équations d'onde et elles doivent être associées à des schémas en temps. La méthode de discrétisation en temps la plus largement utilisée repose sur l'utilisation du schéma Leap-Frog qui est explicite et d'ordre 2 seulement. L'erreur globale est donc pénalisée si on ne combine pas les DGM avec des schémas en temps d'ordre élevé. C'est ainsi que les schémas DG-ADER [29] ont été développés comme une extension de la technique de l'équation modifiée (MET) [22, 60]. L'intégration en temps des schémas DG-ADER n'utilise qu'un seul pas de temps, contrairement par exemple aux schémas d'Adams Bashforth que nous étudions au chapitre 6. Les schémas ADER ne nécessitent donc que le stockage de la solution au pas de temps précédent. Cependant, la mise en oeuvre optimisée des schémas DG-ADER dans un environnement HPC (High Performance Computing) nécessite de définir des variables auxiliaires. On peut alors dépasser les capacités de la mémoire, en particulier pour la résolution des équations d'ondes 3D. Il nous semble donc intéressant de proposer un nouveau schéma de discrétisation en temps qui utilise moins de mémoire que les méthodes DG-ADER pour un même niveau fixé de précision. Pour construire un schéma en temps d'ordre élevé, on peut utiliser la dualité entre dérivée en temps d'ordre élevé et opérateurs en espace d'ordre élevé qui est un héritage de l'hyperbolicité de l'équation des ondes. La MET [22, 60] exploite cette propriété mais la substitution se fait au niveau de l'équation semi-discrétisée en espace. On exprime ainsi la dérivée en temps d'ordre élevé par des produits de matrices, ce qui génère des coûts additionnels. La MET a été élaborée à un moment où on utilisait essentiellement des éléments finis classiques. Or, si on applique une DGM, on peut facilement approcher des opérateurs en espace d'ordre élevé. On se dit alors qu'au lieu de considérer l'équation semi-discrétisée en espace, on pourrait travailler directement sur le problème continu. Dans ce cas, on écrit le schéma en temps avant de discrétiser les opérateurs en espace. Dans ce chapitre, nous allons montrer qu'on peut construire un schéma à un seul pas d'intégration, d'ordre arbitrairement élevé pour des coûts de calcul acceptables. À la section 3.1, nous rappelons la formulation des schémas Leap-Frog (sous-section 3.1.1) et ADER (sous-section 3.1.2), puis nous présentons le nouveau schéma, appelé schéma  $\nabla^m$  3.1.3. Dans la deuxième section, nous rappelons les formulations des trois schémas pour la formulation de l'équation des ondes du second ordre car nous utiliserons cette formulation pour valider certains de nos résultats.

### 3.1 Schémas en temps d'ordre élevé pour la formulation du premier ordre de l'équation des ondes acoustiques

Nous considérons ici l'équation semi-discrétisée en espace obtenue par une méthode DG avec flux centrés comme celle qui a été exposée au chapitre 2 :

$$(3.1.1) \quad \begin{cases} \mathcal{M}_v \frac{d\mathbf{V}}{dt} + \mathcal{K} \mathbf{P} = 0 \\ \mathcal{M}_p \frac{d\mathbf{P}}{dt} - \mathcal{K}^T \mathbf{V} = \mathbf{F}(t) \end{cases}$$

Comme les matrices  $\mathcal{M}_v^{-1}$  et  $\mathcal{M}_p^{-1}$  sont diagonales par bloc, on peut facilement les inverser pour obtenir une formulation simplifiée :

$$(3.1.2) \quad \begin{cases} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P} \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V} + \mathcal{M}_p^{-1} \mathbf{F}(t) \end{cases}$$

où,  $\mathcal{A}_p = \mathcal{M}_v^{-1} \mathcal{K}$  et  $\mathcal{A}_v = -\mathcal{M}_p^{-1} \mathcal{K}^T$ .

#### 3.1.1 Schéma Leap-Frog

Le schéma le plus classique pour discrétiser (3.1.2) en temps est le schéma Leap-Frog (LF). Ce schéma est explicite, d'ordre deux et à un pas d'intégration, ce qui en fait un schéma économe, notamment d'un point de vue stockage puisque nous n'avons pas besoin de stocker de solutions intermédiaires. Il s'écrit

$$(3.1.3) \quad \begin{cases} \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t} = \mathcal{A}_p \mathbf{P}^{n+\frac{1}{2}} \\ \frac{\mathbf{P}^{n+\frac{3}{2}} - \mathbf{P}^{n+\frac{1}{2}}}{\Delta t} = \mathcal{A}_v \mathbf{V}^{n+1} + \mathcal{M}_p^{-1} \mathbf{F}^{n+1} \end{cases}$$

Ce schéma est conditionnellement stable, c'est-à-dire que le pas de temps doit respecter une condition de type CFL (Courant Friedrichs Lewy) pour garantir la stabilité de la solution :  $\Delta t < \beta_{LF} h$  où  $\beta_{LF}$  est un scalaire ne dépendant que de l'ordre de discrétisation et des paramètres physiques des milieux acoustiques. Pour démontrer la stabilité, on vérifie tout d'abord que

$$(3.1.4) \quad \frac{\mathbf{V}^{n+2} - \mathbf{V}^n}{2\Delta t} = \mathcal{A}_p \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2}.$$

On multiplie ensuite la seconde équation de (3.1.3) par  $\mathcal{M}_p \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2}$  et (3.1.4) par  $\mathcal{M}_v \mathbf{V}^{n+1}$  :

$$(3.1.5) \quad \begin{cases} \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+2} - \mathbf{V}^n}{2\Delta t}, \mathbf{V}^{n+1} \right\rangle = \left\langle \mathcal{K} \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2}, \mathbf{V}^{n+1} \right\rangle \\ \left\langle \mathcal{M}_p \frac{\mathbf{P}^{n+\frac{3}{2}} - \mathbf{P}^{n+\frac{1}{2}}}{\Delta t}, \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2} \right\rangle = \left\langle -\mathcal{K}^T \mathbf{V}^{n+1}, \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2} \right\rangle + \left\langle \mathbf{F}^{n+1}, \frac{\mathbf{P}^{n+\frac{3}{2}} + \mathbf{P}^{n+\frac{1}{2}}}{2} \right\rangle \end{cases}$$

En additionnant ces deux équations, et en supposant que  $\mathbf{F}^{n+1} = 0$ , nous obtenons

$$\frac{1}{2\Delta t} \left( \langle \mathcal{M}_v \mathbf{V}^{n+2}, \mathbf{V}^{n+1} \rangle + \langle \mathcal{M}_p \mathbf{P}^{n+\frac{3}{2}}, \mathbf{P}^{n+\frac{3}{2}} \rangle - \langle \mathcal{M}_v \mathbf{V}^n, \mathbf{V}^{n+1} \rangle - \langle \mathcal{M}_p \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \rangle \right) x = 0,$$

ce qui montre que  $E^{n+\frac{3}{2}} = E^{n+\frac{1}{2}}$ , avec

$$(3.1.6) \quad E^{n+\frac{1}{2}} = \langle \mathcal{M}_v \mathbf{V}^n, \mathbf{V}^{n+1} \rangle + \langle \mathcal{M}_p \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \rangle.$$

Il nous reste maintenant à vérifier que  $E^{n+\frac{1}{2}} > 0$ . Nous utilisons d'abord la relation

$$\langle \mathcal{M}_v \mathbf{V}^n, \mathbf{V}^{n+1} \rangle = \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2}, \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2} \right\rangle - \frac{\Delta t^2}{4} \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t}, \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t} \right\rangle.$$

De plus, en utilisant (3.1.3), on a

$$\left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t}, \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t} \right\rangle = \left\langle \mathcal{K} \mathbf{P}^{n+\frac{1}{2}}, \mathcal{M}_v^{-1} \mathcal{K} \mathbf{P}^{n+\frac{1}{2}} \right\rangle = \left\langle \mathcal{K}^T \mathcal{M}_v^{-1} \mathcal{K} \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \right\rangle,$$

de telle sorte que (3.1.6) devient

$$(3.1.7) \quad E^{n+\frac{1}{2}} = \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2}, \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2} \right\rangle + \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \mathcal{K}^T \mathcal{M}_v^{-1} \mathcal{K} \right) \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \right\rangle.$$

La positivité de  $E^{n+\frac{1}{2}}$  est alors assurée si la matrice  $\mathcal{M}_p - \frac{\Delta t^2}{4} \mathcal{K}^T \mathcal{M}_v^{-1} \mathcal{K}$  est positive. Comme  $\mathcal{M}_p$  est une matrice définie positive, il est équivalent d'étudier la positivité de

$$I - \frac{\Delta t^2}{4} \mathcal{M}_p^{-1} \mathcal{K}^T \mathcal{M}_v^{-1} \mathcal{K}.$$

En notant  $\lambda_{\max}$  la plus grande valeur propre de  $\mathcal{M}_p^{-1} \mathcal{K}^T \mathcal{M}_v^{-1} \mathcal{K}$ , une condition suffisante pour que  $E$  soit une énergie est donc

$$1 - \frac{\Delta t^2 \lambda_{\max}}{4} > 0.$$

On obtient alors la condition CFL,  $\Delta t < \beta_{LF} h$ , avec  $\lambda_{\max} \approx \beta_{LF}^2 h^2$  d'après [32]. Nous reviendrons sur cette condition lors de l'étude de dispersion du schéma.

Nous allons maintenant nous intéresser à des méthodes permettant d'obtenir des schémas similaires au schéma Leap-Frog mais d'ordre élevé en temps.

### 3.1.2 Schéma ADER

La méthode ADER, de son nom Arbitrary DERivative polynomials, développée par M. Käser et M. Dumbser [50, 29], est une méthode d'ordre élevé en temps et en espace à un pas d'intégration. Elle est équivalente à la Technique de l'Équation Modifiée (MET) [22] lorsque l'on considère le même pas de temps sur tout le domaine. Le principe de la méthode est basé sur l'utilisation de développements de Taylor dans lesquels les dérivées en temps sont remplacées par 3.1.2.

Par exemple, pour construire un schéma d'ordre quatre en temps, nous commençons par utiliser un développement de Taylor d'ordre 4 sur la pression et la vitesse :

$$(3.1.8) \quad \begin{cases} \frac{\mathbf{V}(t^n + \Delta t) - \mathbf{V}^n(t^n)}{\Delta t} &= \partial_t \mathbf{V}(t^n + \frac{\Delta t}{2}) + \frac{\Delta t^2}{24} \partial_t^3 \mathbf{V}(t^n + \frac{\Delta t}{2}) + O(\Delta t^4) \\ \frac{\mathbf{P}(t^n + \frac{3\Delta t}{2}) - \mathbf{P}^n(t^n + \frac{\Delta t}{2})}{\Delta t} &= \partial_t \mathbf{P}(t^n + \Delta t) + \frac{\Delta t^2}{24} \partial_t^3 \mathbf{P}(t^n + \Delta t) + O(\Delta t^4) \end{cases}$$

On utilise ensuite 3.1.2 pour calculer les dérivées première et troisième en temps. Par exemple pour la pression, on obtient successivement

$$\partial_t^3 P = \partial_t^2 (\mathcal{A}_v \mathbf{V} + \mathcal{M}_p^{-1} \mathbf{F}(t)) = \mathcal{A}_v \mathcal{A}_p \partial_t P + \mathcal{M}_p^{-1} \mathbf{F}''(t) = \mathcal{A}_v \mathcal{A}_p \mathcal{A}_v \mathbf{V} + \mathcal{M}_p^{-1} (\mathcal{A}_v \mathcal{A}_p \mathbf{F}(t) + \mathbf{F}''(t)).$$

Le schéma ADER d'ordre 4 s'écrit alors sous la forme suivante :

$$(3.1.9) \quad \begin{cases} \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t} = \mathcal{A}_p \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} (\mathcal{A}_p \mathcal{A}_v \mathcal{A}_p \mathbf{P}^{n+\frac{1}{2}} + \mathcal{A}_p \mathcal{M}_p^{-1} \mathbf{F}^{n+\frac{1}{2}}) \\ \frac{\mathbf{P}^{n+\frac{3}{2}} - \mathbf{P}^{n+\frac{1}{2}}}{\Delta t} = \mathcal{A}_v \mathbf{V}^n + \mathcal{M}_p^{-1} \mathbf{F}^{n+1} + \frac{\Delta t^2}{24} (\mathcal{A}_v \mathcal{A}_p \mathcal{A}_v \mathbf{V}^{n+1} + \mathcal{A}_v \mathcal{A}_p \mathcal{M}_p^{-1} \mathbf{F}^{n+1} + \mathcal{M}_p^{-1} \mathbf{F}^{n+1}) \end{cases}$$

On introduit alors :

$$(3.1.10) \quad E^{n+\frac{1}{2}} = \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2}, \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2} \right\rangle + \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \mathcal{K}_{ADER}^T \mathcal{M}_v^{-1} \mathcal{K}_{ADER} \right) \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \right\rangle,$$

avec

$$(3.1.11) \quad \mathcal{K}_{ADER} = \mathcal{K} - \frac{\Delta t^2}{24} (\mathcal{K} \mathcal{M}_v \mathcal{K}^T \mathcal{M}_p \mathcal{K}).$$

En effectuant les mêmes calculs que pour le schéma Leap-Frog, on montre que  $E$  est une quantité conservée à chaque pas de temps. De plus, si on note  $\lambda_{ADER, \max}$  la plus grande valeur propre de  $\mathcal{M}_p^{-1} \mathcal{K}_{ADER}^T \mathcal{M}_v^{-1} \mathcal{K}_{ADER}$ , on obtient que  $E$  est une quantité positive si

$$(3.1.12) \quad 1 - \frac{\Delta t^2}{4} \lambda_{ADER, \max} > 0.$$

Dans [12, 32], on montre que cette condition peut s'exprimer en fonction de la valeur propre  $\lambda_{\max}$  définie précédemment pour le schéma Leap-Frog. Cette relation s'obtient en partant de (3.1.11) qui donne :

$$\mathcal{M}_p^{-1} \mathcal{K}_{ADER}^T \mathcal{M}_v^{-1} \mathcal{K}_{ADER} = \mathcal{M}_p^{-1} \mathcal{K} \mathcal{M}_v^{-1} \mathcal{K} - \frac{\Delta t^2}{12} (\mathcal{M}_p^{-1} \mathcal{K} \mathcal{M}_v^{-1} \mathcal{K})^2 + \frac{\Delta t^4}{576} (\mathcal{M}_p^{-1} \mathcal{K} \mathcal{M}_v^{-1} \mathcal{K})^3,$$

et donc

$$\lambda_{ADER, \max} = \lambda_{\max} - \frac{\Delta t^2 \lambda_{\max}^2}{12} + \frac{\Delta t^4 \lambda_{\max}^3}{576}.$$

La condition (3.1.12) peut alors se réécrire

$$(3.1.13) \quad 1 - \frac{\Delta t^2}{4} \left( \lambda_{\max} - \frac{\Delta t^2 \lambda_{\max}^2}{12} + \frac{\Delta t^4 \lambda_{\max}^3}{576} \right) > 0.$$

Ce polynôme admet deux racines complexes et une racine réelle de valeur approximative 8.1072. Pour que le schéma ADER soit stable, il suffit donc que  $\Delta t^2 \lambda_{\max} < 8.1072$ , soit, en utilisant l'expression de  $\lambda_{\max}$  en fonction de  $h$ ,

$$\Delta t < 2.8 \beta_{LF} h.$$

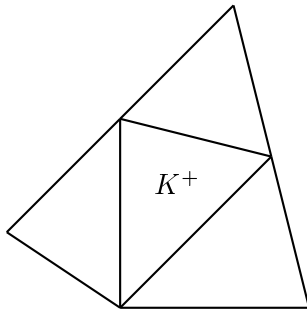
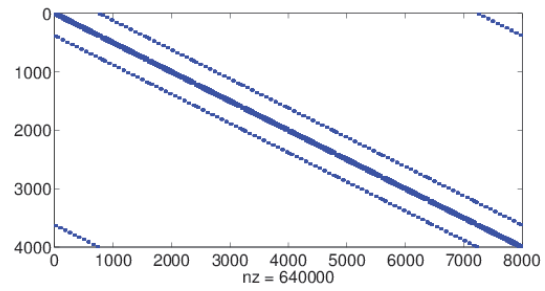
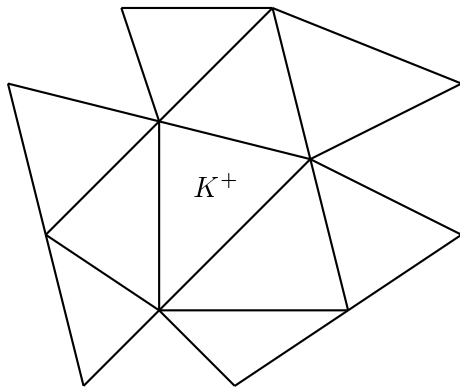
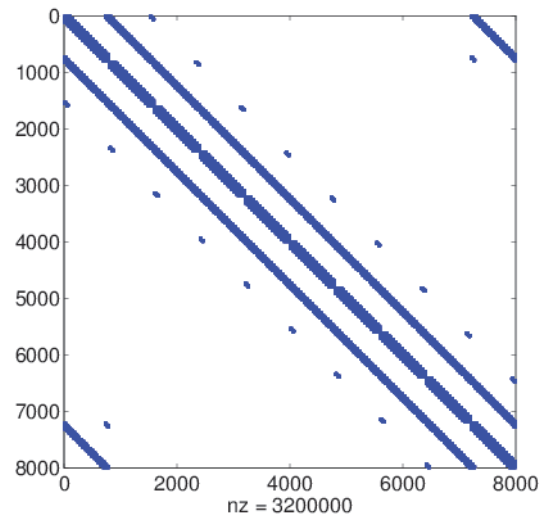
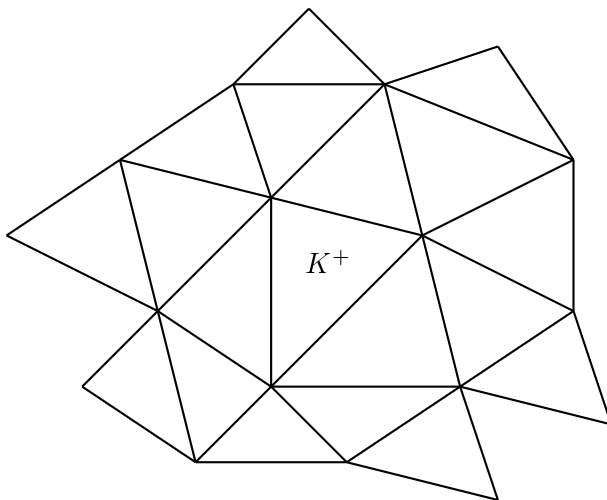
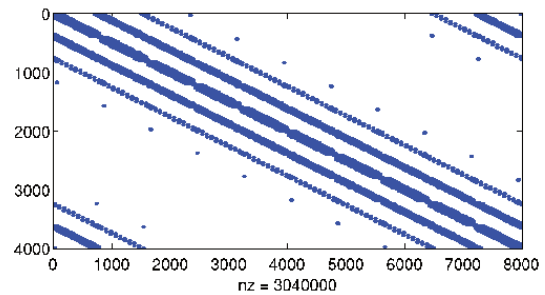
Le nombre de pas de temps requis par le schéma ADER est donc 2.8 fois moins grand que le nombre de temps requis par le schéma Leap-frog. Mais cela ne suffit pas pour comparer les deux schémas. Il faut en effet évaluer le coût de calcul d'une itération en temps.

Au premier abord, on se dit qu'il vaut mieux calculer et stocker  $\mathcal{A}_v + \frac{\Delta t^2}{24} \mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$  et  $\mathcal{A}_p + \frac{\Delta t^2}{24} \mathcal{A}_v \mathcal{A}_p \mathcal{A}_v$  dès le début, pour limiter le nombre de calcul. En fait, le coût de trois multiplications par les matrices  $\mathcal{A}_p$  et  $\mathcal{A}_v$  est sensiblement le même que celui d'une multiplication par la matrice  $\mathcal{A}_v + \frac{\Delta t^2}{24} \mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$  et ceci est dû à la structure de ces matrices. Pour expliquer cette propriété, prenons le cas de la dimension deux. La matrice  $\mathcal{A}_p$  permet de calculer la valeur de  $P$  sur un élément à partir des valeurs de  $V$  sur cet élément et ses trois voisins. Elle est donc constituée de blocs de taille  $N_\Phi \times 2N_\Phi$ ,  $N_\Phi$  étant le nombre de degrés de liberté d'un élément. Pour chaque élément, elle contient un bloc sur sa diagonale correspondant aux interactions de l'élément avec lui-même et trois blocs extra diagonaux correspondant aux interactions de l'élément avec ses trois voisins. La matrice  $\mathcal{A}_v \mathcal{A}_p$  contient beaucoup plus d'éléments non nuls car en plus des interactions avec les voisins, il faut prendre en compte les interactions avec les voisins de second niveau (i.e. les voisins des voisins). Pour la matrice  $\mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$  il faut rajouter les interactions avec les voisins de troisième niveau. Au final, le coût de stockage est également considérablement augmenté, car les matrices globales sont bien moins creuses que les matrices  $\mathcal{A}_v$  et  $\mathcal{A}_p$ . Nous représentons les voisins de premier, deuxième et troisième niveau sur les figures 3.1.a), c) et e). Sur les figures 3.1.b), d) et f), nous représentons les profils respectifs des matrices  $\mathcal{A}_p$ ,  $\mathcal{A}_v \mathcal{A}_p$  et  $\mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$ . Remarquons que, dans un algorithme parallèle, il faut optimiser les communications entre les processeurs, et il est évidemment bien plus simple de communiquer à un processeur les valeurs d'un vecteur sur les voisins du domaine que de communiquer les valeurs de ce vecteur sur les voisins des trois niveaux.

En pratique, ce schéma est résolu en utilisant des variables auxiliaires et en appliquant l'algorithme suivant,

- Calcul de  $\mathbf{P}^{n+\frac{1}{2}}$  :
  - $W = \mathcal{A}_v \mathbf{V}^n$
  - $Q = \mathcal{A}_p W + \mathcal{M}_p^{-1} \mathbf{F}'^n$
  - $\mathbf{P}^{n+\frac{1}{2}} = \mathbf{P}^{n-\frac{1}{2}} + \Delta t \left( W + \frac{\Delta t^2}{24} \mathcal{A}_v Q \right)$
- Calcul de  $\mathbf{V}^{n+1}$  :
  - $Q = \mathcal{A}_p \mathbf{P}^{n+\frac{1}{2}} + \mathcal{M}_p^{-1} \mathbf{F}'^{n+\frac{1}{2}}$
  - $W = \mathcal{A}_v Q$
  - $\mathbf{V}^{n+1} = \mathbf{V}^n + \Delta t \left( Q + \frac{\Delta t^2}{24} \left( \mathcal{A}_p W + \mathcal{M}_p^{-1} \mathbf{F}'^{n+\frac{1}{2}} \Delta t^2 \right) \right)$



(a) L'élément  $K$  et ses voisins de premier niveau(b) Profil de la matrice  $\mathcal{A}_p$ (c) L'élément  $K$  et ses voisins de premier et second niveaux(d) Profil de la matrice  $\mathcal{A}_v \mathcal{A}_p$ (e) L'élément  $K$  et ses voisins de premier, second et troisième niveaux(f) Profil de la matrice  $\mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$ FIGURE 3.1 – Voisins de  $K$  et profils des matrices  $\mathcal{A}_p$ ,  $\mathcal{A}_v \mathcal{A}_p$  et  $\mathcal{A}_p \mathcal{A}_v \mathcal{A}_p$

Même si le schéma ADER est un schéma à un pas, sa mise en œuvre demande donc le stockage de deux variables auxiliaires supplémentaires. En ce qui concerne le nombre d'opérations, il faut réaliser six multiplications par des matrices de rigidité contre deux pour le schéma Leap-Frog classique mais cette augmentation du nombre d'opérations est compensée par l'augmentation de la condition CFL qui est multipliée par 2.8 par rapport à celle du Leap-Frog. Cependant, cette compensation n'est effective que pour les schémas d'ordre quatre et pour les schémas d'ordre supérieur, l'augmentation de la CFL ne suffit pas à compenser les coûts additionnels engendrés par les produits de matrices.

Afin de limiter les coûts de stockage et de calcul, nous allons donc proposer un nouveau schéma qui ne nécessite pas de stocker de variables auxiliaires et dont les coûts de calcul n'augmentent pas avec l'ordre de discrétisation. Ce schéma est présenté à la section suivante.

### 3.1.3 Schéma $\nabla^m$

On propose ici une alternative au schéma DG-ADER, le schéma  $\nabla^m$ . Cette méthode est une adaptation aux systèmes du premier ordre de la méthode développée par Agut *et al* [1, 3] pour l'équation des ondes en formulation du second ordre que nous rappelons à la section suivante. L'idée consiste à inverser l'ordre des discrétisations, c'est-à-dire à discrétiser d'abord l'équation continue en temps, puis à discrétiser le système en espace. La discrétisation en temps est faite en suivant la technique de l'équation modifiée, ce qui fait apparaître des opérateurs d'ordre élevé qui sont les équivalents de la triple multiplication par des matrices de raideur de la méthode ADER. Au premier abord, l'introduction d'opérateurs d'ordre élevé peut constituer une difficulté mais comme nous le verrons par la suite, les techniques de Galerkin Discontinues sont particulièrement bien adaptées pour la discrétisation en espace de ces opérateurs.

#### 3.1.3.1 Discrétisation en temps

Pour la discrétisation en temps, nous effectuons un développement de Taylor à l'ordre quatre des inconnues  $\mathbf{v}$  et  $p$ ,

$$(3.1.14) \quad \begin{cases} \frac{\mathbf{v}(\mathbf{x}, t^n + \Delta t) - \mathbf{v}(\mathbf{x}, t^n)}{\Delta t} = \partial_t \mathbf{v} \left( \mathbf{x}, t^n + \frac{\Delta t}{2} \right) + \frac{\Delta t^2}{24} \partial_t^3 \mathbf{v} \left( \mathbf{x}, t^n + \frac{\Delta t}{2} \right) + O(\Delta t^4) \\ \frac{p \left( \mathbf{x}, t^n + \frac{3\Delta t}{2} \right) - p \left( \mathbf{x}, t^n + \frac{\Delta t}{2} \right)}{\Delta t} = \partial_t p(\mathbf{x}, t^n + \Delta t) + \frac{\Delta t^2}{24} \partial_t^3 p(\mathbf{x}, t^n + \Delta t) + O(\Delta t^4). \end{cases}$$

Nous utilisons alors l'équation des ondes continue (2.1.2) pour calculer les dérivées première et troisième en temps de  $p$  et de  $\mathbf{v}$ . Nous obtenons ainsi le schéma semi-discrétisé en temps dans lequel nous omettons la dépendance en espace des inconnues et des paramètres pour alléger la présentation :

$$(3.1.15) \quad \begin{cases} \rho \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \nabla p^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \left[ \nabla \left( \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right) \right) + \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \right] \\ \frac{1}{\mu} \frac{p^{n+\frac{3}{2}} - p^{n+\frac{1}{2}}}{\Delta t} = \nabla \cdot \mathbf{v}^{n+1} + \frac{\Delta t^2}{24} \left[ \nabla \cdot \left( \frac{1}{\rho} \nabla (\mu \nabla \cdot \mathbf{v}^{n+1}) \right) + \mu \nabla \cdot \frac{1}{\rho} \nabla f^{n+1} + f_{t^2}^{n+1} \right] + f^{n+1}. \end{cases}$$

Ici,  $f^n$ ,  $f_t^n$  et  $f_{t^2}^n$  désignent respectivement les valeur de  $f$ , de sa dérivée en temps et de sa dérivée seconde en temps à l'instant  $t^n = n\Delta t$ .

Nous avons choisi d'appeler ce schéma  $\nabla^3$  car il fait intervenir des opérateurs du troisième ordre qui s'expriment à partir du gradient. Nous ne décrivons pas la méthode pour construire des schémas

d'ordre plus élevé car elle ne présente aucune difficulté supplémentaire. Il suffit de suivre la même démarche en partant d'une formule de Taylor d'ordre plus élevé. On peut donc ainsi construire des schémas d'ordre arbitrairement élevé en temps. Dans la suite de ce travail, nous nous concentrerons uniquement sur les schémas  $\nabla^3$  et nous les appellerons simplement schémas  $\nabla$  pour simplifier les notations.

Si les champs  $p$  et  $\mathbf{v}$  sont suffisamment réguliers, et en l'absence de terme source, on peut montrer que ce schéma conserve la quantité

$$(3.1.16) \quad \mathcal{E}(t) = \int_{\Omega} \rho \mathbf{v}^{n+1} \cdot \mathbf{v}^n + \int_{\Omega} \frac{1}{\mu} \left| p^{n+\frac{1}{2}} \right|^2.$$

On écrit tout d'abord la première équation de (3.1.15) sur deux pas de temps :

$$(3.1.17) \quad \begin{cases} \rho \frac{\mathbf{v}^{n+2} - \mathbf{v}^n}{2\Delta t} = \nabla \frac{p^{n+\frac{3}{2}} + p^{n+\frac{1}{2}}}{2} + \frac{\Delta t^2}{24} \nabla \left( \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^{n+\frac{3}{2}} + p^{n+\frac{1}{2}} \right) \right) \\ \frac{1}{\mu} \frac{p^{n+\frac{3}{2}} - p^{n+\frac{1}{2}}}{\Delta t} = \nabla \cdot \mathbf{v}^{n+1} + \frac{\Delta t^2}{24} \nabla \cdot \left( \frac{1}{\rho} \nabla (\mu \nabla \cdot \mathbf{v}^{n+1}) \right). \end{cases}$$

On multiplie alors la première équation par  $\mathbf{v}^{n+1}$  et la seconde par  $\frac{p^{n+\frac{3}{2}} + p^{n+\frac{1}{2}}}{2}$ , on intègre sur  $\Omega$  et on additionne les deux équations pour obtenir (3.1.16).

Pour étudier la positivité de  $\int_{\Omega} \rho \mathbf{v}^{n+1} \cdot \mathbf{v}^n$ , on utilise l'identité

$$\int_{\Omega} \rho \mathbf{v}^{n+1} \cdot \mathbf{v}^n = \int_{\Omega} \rho \left\| \frac{\mathbf{v}^{n+1} + \mathbf{v}^n}{2} \right\|^2 - \frac{\Delta t^2}{4} \int_{\Omega} \rho \left\| \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} \right\|^2,$$

puis on remplace  $\frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t}$  en utilisant la première équation de (3.1.15). On a alors

$$\mathcal{E}(t) = \int_{\Omega} \rho \left\| \frac{\mathbf{v}^{n+1} + \mathbf{v}^n}{2} \right\|^2 + \int_{\Omega} \frac{1}{\mu} p^{n+\frac{1}{2}^2} - \frac{\Delta t^2}{4} \int_{\Omega} \left[ \nabla p^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \nabla \left( \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right) \right) \right]^2.$$

Contrairement aux schémas que nous avons étudiés précédemment, on peut remarquer qu'il est possible que  $\mathcal{E}(t)$  soit négatif. Pour illustrer ce phénomène, considérons le cas particulier 1D où  $\Omega = [0, 1]$  et  $\rho = \mu = 1$ .  $\mathcal{E}(t)$  s'écrit alors

$$\mathcal{E}(t) = \int_0^1 \rho \left\| \frac{\mathbf{v}^{n+1} + \mathbf{v}^n}{2} \right\|^2 + \int_0^1 p^{n+\frac{1}{2}^2} - \frac{\Delta t^2}{4} \int_0^1 \left[ \nabla p^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \nabla \left( \Delta p^{n+\frac{1}{2}} \right) \right]^2.$$

Si  $\mathbf{v}^{n+1} = \mathbf{v}^n = 0$  et  $p = \sin(kx)$ , alors

$$\mathcal{E}(t) = \int_0^1 \sin^2(kx) dx - \frac{\Delta t^2}{4} \int_0^1 \left[ k \cos(kx) - \frac{\Delta t^2}{24} k^3 \cos^3(kx) \right]^2,$$

soit

$$\mathcal{E}(t) = \int_0^1 \sin^2(kx) dx - k^2 \frac{\Delta t^2}{4} \int_0^1 \left[ \cos(kx) - \frac{\Delta t^2}{24} k^2 \cos^3(kx) \right]^2.$$

Pour  $\Delta t$  fixé, on voit que  $\mathcal{E}(t)$  est positif pour  $k$  suffisamment petit et négatif pour  $k$  suffisamment grand. Le schéma semi-discrétisé est donc instable pour les hautes fréquences. Comme nous le verrons par la suite, le schéma totalement discrétisé ne présente pas ces phénomènes d'instabilité car l'espace d'approximation discret filtre les hautes fréquences. En effet, cet espace ne permet pas de représenter les solutions de fréquence supérieure à  $\beta h$ , où  $\beta$  est un paramètre dépendant de la méthode de discrétisation en espace et de son ordre. On obtiendra donc à nouveau une condition CFL du type  $\Delta t < \beta h$ .

### 3.1.3.2 Discrétisation en espace

Il nous faut maintenant discrétiser les opérateurs d'ordre élevé en espace. Pour cela, nous allons adapter aux opérateurs d'ordre élevé la méthode DG avec flux centrés que nous avons présentée à la section précédente. Comme pour les schémas DG-Flux classiques, nous multiplions chaque équation de (3.1.15) par une fonction test localement continue et nous intégrons élément par élément. Pour les opérateurs d'ordre élevé, nous appliquons trois fois la formule de Green. Nous nous concentrons uniquement sur l'opérateur  $\nabla \mu \nabla \cdot \frac{1}{\rho} \nabla$ , l'autre opérateur étant traité de manière similaire. Pour simplifier la présentation, nous supposons de plus que  $\rho$  et  $\mu$  sont constants par élément et nous noterons  $\rho_K$  et  $\mu_K$  la valeur de  $\rho$  et  $\mu$  sur l'élément  $K$ .

En sommant sur tous les éléments, nous obtenons la forme bilinéaire

$$\begin{aligned} & \sum_K \frac{\mu_K}{\rho_K} \int_K \left( p^{n+\frac{1}{2}} \nabla \cdot \nabla \nabla \cdot \mathbf{w} \right) \\ & - \sum_{\Gamma_{int}} \int_{\Gamma} \left( \left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \cdot \mathbf{w} \right] \right] - \left[ \left[ \frac{\mu}{\rho} \nabla p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} \right] \right] + \left[ \left[ \frac{\mu}{\rho} p^{n+\frac{1}{2}} \nabla \nabla \cdot \mathbf{w} \right] \right] \right) \cdot \mathbf{n} \\ & - \sum_{\Gamma_{ext}} \int_{\Gamma} \left( \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \cdot \mathbf{w} - \frac{\mu}{\rho} \nabla p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} + \frac{\mu}{\rho} p^{n+\frac{1}{2}} \nabla \nabla \cdot \mathbf{w} \right) \cdot \mathbf{n} \end{aligned}$$

Nous nous intéressons d'abord au premier terme de bord intérieur. Nous avons :

$$\left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \mathbf{w} \right] \right] \cdot \mathbf{n} = \left\{ \left\{ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \right\} \right\} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n} + \left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \right] \right] \{ \{ \mathbf{w} \} \} \cdot \mathbf{n}$$

En dérivant deux fois par rapport au temps la condition de transmission  $\llbracket p \rrbracket = 0$  pour obtenir  $\llbracket \partial_t^2 p \rrbracket = 0$  et en utilisant l'équation des ondes, on obtient

$$\left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \right] \right] = 0,$$

pour  $p$  suffisamment régulière, La discrétisation en espace reste donc consistante si on utilise

$$\left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \mathbf{w} \right] \right] \cdot \mathbf{n} = \left\{ \left\{ \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \right\} \right\} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n}.$$

De la même manière, nous avons

$$\left[ \left[ \frac{\mu}{\rho} \nabla p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n} = \left[ \left[ \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right] \right] \{ \{ \mu \nabla \cdot \mathbf{w} \} \} \cdot \mathbf{n} + \left\{ \left\{ \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right\} \right\} \llbracket \mu \nabla \cdot \mathbf{w} \rrbracket \cdot \mathbf{n}$$

En dérivant par rapport au temps la condition de transmission  $\llbracket \mathbf{v} \rrbracket = 0$  pour obtenir  $\llbracket \partial_t \mathbf{v} \rrbracket = 0$  et en utilisant l'équation des ondes, on obtient

$$\left[ \left[ \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right] \right] \cdot \mathbf{n} = 0,$$

pour  $\mathbf{v}$  suffisamment régulière. Nous pouvons alors poser

$$\left[ \left[ \frac{\mu}{\rho} \nabla p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n} = \left\{ \left\{ \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right\} \right\} \llbracket \mu \nabla \cdot \mathbf{w} \rrbracket \cdot \mathbf{n}$$

Il ne nous reste plus qu'à étudier le dernier terme :

$$\left[ \left[ \frac{\mu}{\rho} p^{n+\frac{1}{2}} \nabla \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n} = \left[ \left[ p^{n+\frac{1}{2}} \right] \right] \left\{ \left\{ \frac{\mu}{\rho} \nabla \nabla \cdot \mathbf{w} \right\} \right\} \cdot \mathbf{n} + \left\{ \left\{ p^{n+\frac{1}{2}} \right\} \right\} \left[ \left[ \frac{\mu}{\rho} \nabla \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n},$$

qui devient, en utilisant la condition de transmission  $\llbracket p \rrbracket = 0$ ,

$$\left[ \left[ \frac{\mu}{\rho} p^{n+\frac{1}{2}} \nabla \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n} = \left\{ \left\{ p^{n+\frac{1}{2}} \right\} \right\} \left[ \left[ \frac{\mu}{\rho} \nabla \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n}.$$

Nous choisissons maintenant d'imposer une condition de Dirichlet  $p = 0$  sur le bord extérieur. En dérivant cette condition deux fois par rapport au temps et en utilisant l'équation des ondes, on obtient la relation  $\nabla \cdot \nabla p = 0$  sur le bord extérieur. Nous posons donc

$$\sum_{\Gamma_{ext}} \int_{\Gamma} \left( \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \cdot \mathbf{w} + \frac{\mu}{\rho} p^{n+\frac{1}{2}} \nabla \nabla \cdot \mathbf{w} \right) \cdot \mathbf{n} = 0.$$

Finalement l'équation régissant la vitesse s'écrit

$$(3.1.18) \quad \sum_K \rho_K \int_K \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} \cdot \mathbf{w} = \sum_K \int_K p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ p^{n+\frac{1}{2}} \right\} \right\} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n} \\ + \frac{\Delta t^2}{24} \left( b_{p\nabla}(p, \mathbf{w}) + \int_{\Omega} \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \cdot \mathbf{w} \right)$$

avec

$$b_{p\nabla}(p, \mathbf{w}) = \sum_K \frac{\mu_K}{\rho_K} \int_K \left( p^{n+\frac{1}{2}} \nabla \cdot \nabla \nabla \cdot \mathbf{w} \right) - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \left( \frac{\mu}{\rho} \nabla \cdot \nabla p^{n+\frac{1}{2}} \right) \right\} \right\} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n} \\ + \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \left( \frac{1}{\rho} \nabla p^{n+\frac{1}{2}} \right) \right\} \right\} \llbracket \mu \nabla \cdot \mathbf{w} \rrbracket \cdot \mathbf{n} - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ p^{n+\frac{1}{2}} \right\} \right\} \left[ \left[ \frac{\mu}{\rho} \nabla \nabla \cdot \mathbf{w} \right] \right] \cdot \mathbf{n} \\ + \sum_{\Gamma_{ext}} \int_{\Gamma} \frac{\mu}{\rho} \nabla p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} \cdot \mathbf{n}$$

De la même manière, on obtient l'équation régissant la pression

$$(3.1.19) \quad \sum_K \frac{1}{\mu_K} \int_K \frac{p^{n+\frac{3}{2}} - p^{n+\frac{1}{2}}}{\Delta t} \cdot q = \sum_K \int_K \mathbf{v}^{n+1} \nabla q - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \mathbf{v}^{n+1} \right\} \right\} \llbracket q \rrbracket \cdot \mathbf{n} \\ + \int_{\Omega} f^{n+\frac{1}{2}} q + \frac{\Delta t^2}{24} b_{\mathbf{v}\nabla}(\mathbf{v}, q) \\ + \frac{\Delta t^2}{24} \left( \int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla f^{n+1} q + \int_{\Omega} f_{t^2}^{n+1} q \right)$$

avec

$$b_{\mathbf{v}\nabla}(\mathbf{v}, q) = \sum_K \frac{\mu_K}{\rho_K} \int_K \left( \mathbf{v}^{n+1} \nabla \nabla \cdot \nabla \cdot q \right) - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \frac{\mu}{\rho} (\nabla \nabla \cdot \mathbf{v}^{n+1}) \right\} \right\} \llbracket q \rrbracket \cdot \mathbf{n} \\ + \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \mu (\nabla \cdot \mathbf{v}^{n+\frac{1}{2}}) \right\} \right\} \left[ \left[ \frac{1}{\rho} \nabla q \right] \right] \cdot \mathbf{n} - \sum_{\Gamma_{int}} \int_{\Gamma} \left\{ \left\{ \mathbf{v}^{n+\frac{1}{2}} \right\} \right\} \left[ \left[ \frac{\mu}{\rho} \nabla \cdot \nabla q \right] \right] \cdot \mathbf{n} \\ - \sum_{\Gamma_{ext}} \int_{\Gamma} \frac{\mu}{\rho} (\nabla \nabla \cdot \mathbf{v}^{n+1}) q \cdot \mathbf{n} - \sum_{\Gamma_{ext}} \int_{\Gamma} \mathbf{v}^{n+\frac{1}{2}} \frac{\mu}{\rho} \nabla \cdot \nabla q \cdot \mathbf{n}.$$

On vérifie, en utilisant des formules de Green comme pour le cas du schéma DG flux centrés, que  $b_{p\nabla}(q, \mathbf{w}) = -b_{\mathbf{v}\nabla}(\mathbf{w}, q)$ . Nous pouvons donc simplifier les notations en posant  $b_{\nabla}(q, \mathbf{w}) = b_{p\nabla}(q, \mathbf{w})$  et la formulation variationnelle continue s'écrit

$$(3.1.20) \quad \left\{ \begin{array}{l} \sum_K \int_K \rho_K \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} \cdot \mathbf{w} = \sum_K \int_K p^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} - \sum_{\Gamma_{int}} \int_{\Gamma} \{ \{ p^{n+\frac{1}{2}} \} \} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n} \\ \quad + \frac{\Delta t^2}{24} \left( b_{\nabla}(p, \mathbf{w}) + \int_{\Omega} \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \cdot \mathbf{w} \right) \\ \sum_K \frac{1}{\mu_K} \int_K \frac{p^{n+\frac{3}{2}} - p^{n+\frac{1}{2}}}{\Delta t} \cdot q = \sum_K \int_K \mathbf{v}^{n+1} \nabla q - \sum_{\Gamma_{int}} \int_{\Gamma} \{ \{ \mathbf{v}^{n+1} \} \} \llbracket q \rrbracket \cdot \mathbf{n} \\ \quad + \int_{\Omega} f^{n+1} q - \frac{\Delta t^2}{24} b_{\nabla}(q, \mathbf{v}) \\ \quad + \frac{\Delta t^2}{24} \left( \int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla f^{n+1} q + \int_{\Omega} f_{t^2}^{n+1} q \right). \end{array} \right.$$

La formulation variationnelle discrète s'écrit alors :

Trouver  $(p_h, \mathbf{v}_h)$  dans  $C^1((0, T), V_h^m \times (V_h^m)^2)$  tel que, pour tout  $(q, \mathbf{w})$  dans  $V_h^m \times (V_h^m)^2$ , on ait

$$(3.1.21) \quad \left\{ \begin{array}{l} \sum_K \int_K \rho_K \frac{\mathbf{v}_h^{n+1} - \mathbf{v}_h^n}{\Delta t} \cdot \mathbf{w} = \sum_K \int_K p_h^{n+\frac{1}{2}} \nabla \cdot \mathbf{w} - \sum_{\Gamma_{int}} \int_{\Gamma} \{ \{ p_h^{n+\frac{1}{2}} \} \} \llbracket \mathbf{w} \rrbracket \cdot \mathbf{n} \\ \quad + \frac{\Delta t^2}{24} \left( b_{\nabla}(p_h, \mathbf{w}) + \int_{\Omega} \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \cdot \mathbf{w} \right) \\ \sum_K \frac{1}{\mu_K} \int_K \frac{p_h^{n+\frac{3}{2}} - p_h^{n+\frac{1}{2}}}{\Delta t} \cdot q = \sum_K \int_K \mathbf{v}_h^{n+1} \nabla q - \sum_{\Gamma_{int}} \int_{\Gamma} \{ \{ \mathbf{v}_h^{n+1} \} \} \llbracket q \rrbracket \cdot \mathbf{n} \\ \quad + \int_{\Omega} f^{n+1} q - \frac{\Delta t^2}{24} b_{\nabla}(q, \mathbf{v}_h) \\ \quad + \frac{\Delta t^2}{24} \left( \int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla f^{n+1} q + \int_{\Omega} f_{t^2}^{n+1} q \right). \end{array} \right.$$

### 3.1.3.3 Le schéma totalement discrétisé

En utilisant les fonctions de base de Lagrange classiques, on obtient finalement le schéma totalement discrétisé

$$(3.1.22) \quad \left\{ \begin{array}{l} \mathcal{M}_v \frac{\mathbf{V}^{n+1} - \mathbf{V}^n}{\Delta t} = \mathcal{K} \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3} \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \mathbf{F}_v^{n+\frac{1}{2}} \\ \mathcal{M}_p \frac{\mathbf{P}^{n+\frac{3}{2}} - \mathbf{P}^{n+\frac{1}{2}}}{\Delta t} = -\mathcal{K}^T \mathbf{V}^n - \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3}^T \mathbf{V}^n + \mathbf{F}^{n+1} + \frac{\Delta t^2}{24} \mathbf{F}_p^{n+1} \end{array} \right.$$

où  $\mathcal{K}$  est la matrice associée à la discrétisation de l'opérateur d'ordre un (la même que pour DG-Flux),  $\mathcal{K}_{\nabla^3}$  est la matrice associée à l'opérateur d'ordre trois. Les vecteurs  $\mathbf{F}_v$  et  $\mathbf{F}_p$  sont les vecteurs

obtenus après discrétisation des termes sources d'ordre élevé. Par exemple, en dimension deux, ce système s'écrit

(3.1.23)

$$\begin{cases} \mathcal{M}_{vx} \frac{\mathbf{V}_x^{n+1} - \mathbf{V}_x^n}{\Delta t} = \mathcal{K}_x \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3 x} \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \mathbf{F}_{\mathbf{v}_x}^{n+\frac{1}{2}} \\ \mathcal{M}_{vy} \frac{\mathbf{V}_y^{n+1} - \mathbf{V}_y^n}{\Delta t} = \mathcal{K}_y \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3 y} \mathbf{P}^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \mathbf{F}_{\mathbf{v}_y}^{n+\frac{1}{2}} \\ \mathcal{M}_p \frac{\mathbf{P}^{n+\frac{3}{2}} - \mathbf{P}^{n+\frac{1}{2}}}{\Delta t} = -\mathcal{K}_x^T \mathbf{V}_x^n - \mathcal{K}_y^T \mathbf{V}_y^n - \frac{\Delta t^3}{24} \left( \mathcal{K}_{\nabla^3 x} \mathbf{V}_x^n + \mathcal{K}_{\nabla^3 y} \mathbf{V}_y^n \right) + \mathbf{F}^{n+1} + \frac{\Delta t^2}{24} \mathbf{F}_p^{n+1} \end{cases}$$

Les matrices de raideurs  $\mathcal{K}_{\nabla^3 x}$  et  $\mathcal{K}_{\nabla^3 y}$  sont telles que  $\mathcal{K}_{\nabla^3 x}^{K,J} = \mathcal{K}_{\nabla^3 y}^{K,J} = 0$  si  $K \cap J = \emptyset$ ,

$$\mathcal{K}_{\nabla^3 x}^{K,J} = b_{\nabla}(\phi_j^J, \phi_i^K \mathbf{e}_x), \quad \text{et} \quad \mathcal{K}_{\nabla^3 y}^{K,J} = b_{\nabla}(\phi_j^J, \phi_i^K \mathbf{e}_y)$$

si  $K \cap J = \Gamma$ , autrement dit, si  $K$  et  $J$  sont voisins et

$$\mathcal{K}_{\nabla^3 x}^{K,K} = b_{\nabla}(\phi_j^K, \phi_i^K \mathbf{e}_x), \quad \text{et} \quad \mathcal{K}_{\nabla^3 y}^{K,K} = b_{\nabla}(\phi_j^K, \phi_i^K \mathbf{e}_y).$$

Les vecteurs  $F_{\mathbf{v}_x}^{n+\frac{1}{2}}$ ,  $F_{\mathbf{v}_y}^{n+\frac{1}{2}}$  et  $\mathbf{F}_p^{n+1}$  sont tels que

$$\mathbf{F}_{\mathbf{v}_x}^{n+\frac{1}{2}} = \int_{\Omega} \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \cdot \phi_i^K \mathbf{e}_x, \quad \mathbf{F}_{\mathbf{v}_y}^{n+\frac{1}{2}} = \int_{\Omega} \frac{1}{\rho} \nabla f_t^{n+\frac{1}{2}} \cdot \phi_i^K \mathbf{e}_y$$

et

$$\mathbf{F}_p^{n+1} = \int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla f^{n+1} \phi_i^K + \int_{\Omega} f_t^{n+1} \phi_i^K.$$

À première vue, ce schéma est plus coûteux que le schéma LF, car il nécessite une multiplication supplémentaire par la matrice de raideur  $\mathcal{K}_{\nabla^3}$ . Cependant, cette nouvelle matrice a quasiment le même nombre d'éléments non nuls que la matrice  $\mathcal{K}$ , car un élément ne communique qu'avec ses trois voisins, même dans le cas d'opérateurs d'ordre élevé. En pratique, et contrairement au schéma ADER, on peut donc calculer et stocker directement la matrice  $\mathcal{K} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3}$ , sans augmentation de coût de stockage ou de calcul.

Il nous reste encore à étudier la stabilité du schéma. Des calculs similaires à ceux effectués pour le schéma Leap-Frog permettent de montrer la conservation de

(3.1.24)

$$\begin{aligned} E^{n+\frac{1}{2}} &= \left\langle \mathcal{M}_v \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2}, \frac{\mathbf{V}^{n+1} + \mathbf{V}^n}{2} \right\rangle \\ &+ \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \left( \mathcal{K}^T + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3}^T \right) \right) \mathcal{M}_v^{-1} \left( \mathcal{K} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3} \right) \right\rangle \mathbf{P}^{n+\frac{1}{2}}, \mathbf{P}^{n+\frac{1}{2}} \right\rangle. \end{aligned}$$

L'étude des valeurs propres de

$$\mathcal{M}_p - \frac{\Delta t^2}{4} \left( \mathcal{K}^T + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3}^T \right) \mathcal{M}_v^{-1} \left( \mathcal{K} + \frac{\Delta t^3}{24} \mathcal{K}_{\nabla^3} \right)$$

est beaucoup plus compliquée que pour les études précédentes car ces valeurs propres dépendent de  $\Delta t$ . De plus  $\mathcal{K}^T \mathcal{K}$  et  $\mathcal{K}_{\nabla^3}^T \mathcal{K}_{\nabla^3}$  ne sont pas diagonalisables simultanément (i.e. elles ne sont pas

diagonalisables dans la même base). Nous n'avons donc pas pu calculer une expression analytique de la condition CFL. Dans la pratique, nous avons observé que cette condition était légèrement inférieure, de l'ordre de 10%, à celle du schéma Leap-Frog. Les coûts de calcul d'une itération du schéma  $\nabla$  étant les mêmes que les coûts de calcul d'une itération du Leap-Frog, nous en concluons que le schéma  $\nabla$  est légèrement plus coûteux que le schéma Leap-Frog, mais avec un ordre de précision supérieur. Nous étudierons ces coûts de calcul plus précisément à la section 3.3 et nous donnerons plus de détails sur la condition CFL au chapitre 4.

## 3.2 Schémas en temps d'ordre élevé pour la formulation du second ordre de l'équation des ondes acoustiques

Dans cette section, nous rappelons brièvement les équivalents des schémas LF, ADER et  $\nabla$  pour la formulation du second ordre de l'équation des ondes.

### 3.2.1 Le schéma LF

On s'intéresse à la discrétisation en temps du système.

$$(3.2.1) \quad \mathcal{M}_p \frac{d^2 \mathbf{P}}{dt^2} + \mathcal{K}_{IP} \mathbf{P} = \mathbf{F}(t)$$

Le schéma Leap-Frog associé à ce schéma est obtenu grâce à un développement de Taylor à l'ordre deux. Il s'écrit

$$(3.2.2) \quad \mathcal{M}_p \frac{\mathbf{P}^{n+1} - 2\mathbf{P}^n + \mathbf{P}^{n-1}}{\Delta t^2} = -\mathcal{K}_{IP} \mathbf{P}^n + \mathbf{F}(t)$$

C'est un schéma à deux pas (il faut stocker  $\mathbf{P}^n$  et  $\mathbf{P}^{n-1}$ ), mais le sur-coût est compensé par le fait qu'on ne calcule qu'une inconnue scalaire au lieu d'une inconnue scalaire et une inconnue vectorielle.

Comme pour la formulation d'ordre un, ce schéma est explicite et stable sous une condition CFL. Pour le démontrer, on multiplie (3.2.2) par  $\frac{\mathbf{P}^{n+1} - \mathbf{P}^{n-1}}{\Delta t}$  en supposant l'absence de terme source. On a alors

$$\left\langle \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle_{\mathcal{M}_p}^2 - \left\langle \frac{\mathbf{P}^n - \mathbf{P}^{n-1}}{\Delta t} \right\rangle_{\mathcal{M}_p}^2 + \langle \mathcal{K}_{IP} \mathbf{P}^n, \mathbf{P}^{n+1} \rangle - \langle \mathcal{K}_{IP} \mathbf{P}^n, \mathbf{P}^{n-1} \rangle = 0.$$

La symétrie de  $\mathcal{K}_{IP}$  nous permet d'écrire

$$\langle \mathcal{K}_{IP} \mathbf{P}^n, \mathbf{P}^{n-1} \rangle = \langle \mathcal{K}_{IP} \mathbf{P}^{n-1}, \mathbf{P}^n \rangle$$

et de montrer la conservation de

$$E^{n+\frac{1}{2}} = \left\langle \mathcal{M}_p \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle^2 + \langle \mathcal{K}_{IP} \mathbf{P}^n, \mathbf{P}^{n+1} \rangle.$$

On utilise alors l'identité

$$\langle \mathcal{K}_{IP} \mathbf{P}^n, \mathbf{P}^{n+1} \rangle = \left\langle \mathcal{K}_{IP} \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2}, \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2} \right\rangle - \frac{\Delta t^2}{4} \left\langle \mathcal{K}_{IP} \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t}, \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle,$$



pour obtenir

$$E^{n+\frac{1}{2}} = \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \mathcal{K}_{IP} \right) \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t}, \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle + \left\langle \mathcal{K}_{IP} \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2}, \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2} \right\rangle.$$

La positivité de  $\mathcal{K}_{IP}$  garantit la positivité du second terme de cet expression. Pour assurer la positivité du premier terme, il faut que  $\mathcal{M}_p - \frac{\Delta t^2}{4} \mathcal{K}_{IP}$  soit positive, c'est-à-dire que

$$1 - \frac{\Delta t^2}{4} \lambda_{IP, \max} \geq 0,$$

où  $\lambda_{IP, \max}$  désigne la plus grande valeur propre de  $\mathcal{M}_p^{-1} \mathcal{K}_{IP}$ . On en déduit alors la condition CFL  $\Delta t \leq \beta_{IP} h$ . Dans [3, 1], les auteurs montrent que le coefficient  $\beta_{IP}$  décroît quand le paramètre de pénalisation augmente et fournissent une expression analytique de ce coefficient en fonction de l'ordre des éléments finis et du paramètre de pénalisation. Par exemple, pour un maillage cartésien en dimension  $d$  et avec des éléments de degré 3, une condition nécessaire de stabilité s'écrit :

$$(3.2.3) \quad \sqrt{d} \frac{c \Delta t}{h} \leq \begin{cases} 0.153 & \text{si } 6 \leq \alpha \leq 9.65 \\ \sqrt{2-45 + 6\alpha + \sqrt{405 - 240\alpha + 36\alpha^2}} & \text{si } \alpha > 9.65 \end{cases}$$

où  $\alpha$  est le paramètre de stabilisation décrit au chapitre 2. De plus, ce schéma est inconditionnellement instable pour  $\alpha < \alpha_0 = 6$ . La condition (3.2.3) est uniquement une condition nécessaire de stabilité. Il existe une condition plus précise, également donnée dans [3, 1], mais son calcul est beaucoup plus compliqué et, dans la pratique, la condition (3.2.3) s'avère très proche de la condition nécessaire et suffisante. Pour un paramètre de pénalisation proche du paramètre minimal  $\alpha_0$ , la condition CFL est environ 1.5 fois plus grande que la condition CFL du schéma DG avec flux centré.

### 3.2.2 Le schéma ADER

Les schémas ADER, ou Équation Modifiée, s'écrivent en utilisant des développements de Taylor d'ordre élevé. Nous présentons ici à titre d'exemple la construction du schéma ADER d'ordre quatre. Le développement de Taylor à l'ordre quatre de  $\mathbf{P}$  nous donne

$$(3.2.4) \quad \mathcal{M}_p \frac{\mathbf{P}(t^n + \Delta t) - 2\mathbf{P}(t^n) + \mathbf{P}(t^n - \Delta t)}{\Delta t^2} = \frac{d^2 \mathbf{P}}{dt^2}(t^n + \Delta t) + \frac{\Delta t^2}{12} \frac{d^4 \mathbf{P}}{dt^4}(t^n + \Delta t) + O(\Delta t^4)$$

En utilisant (3.2.1) pour calculer les dérivées d'ordre deux et quatre, on obtient

$$\mathcal{M}_p \frac{d^2 \mathbf{P}}{dt^2} = -\mathcal{K}_{IP} \mathbf{P} \quad \text{et} \quad \mathcal{M}_p \frac{d^4 \mathbf{P}}{dt^4} = -\mathcal{K}_{IP} \frac{d^2 \mathbf{P}}{dt^2} = \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \mathbf{P}$$

et le schéma ADER d'ordre quatre s'écrit

$$(3.2.5) \quad \mathcal{M}_p \frac{\mathbf{P}^{n+1} - 2\mathbf{P}^n + \mathbf{P}^{n-1}}{\Delta t^2} = -\mathcal{K}_{IP} \mathbf{P}^n + \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \mathbf{P}^n + \mathbf{F}^n + \frac{\Delta t^2}{12} (\mathbf{F}''^n + \mathcal{M}_v^{-1} \mathcal{K}_{IP} \mathbf{F}^n)$$

Ce schéma est de la même forme que le schéma Leap-Frog, en remplaçant la matrice  $\mathcal{K}_{IP}$  par  $\mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP}$ . On peut donc montrer comme précédemment que la quantité

$$E^{n+\frac{1}{2}} = \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \left( \mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \right) \right) \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t}, \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle + \left\langle \left( \mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \right) \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2}, \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2} \right\rangle.$$

est conservée. Pour vérifier que c'est une énergie, nous pouvons étudier son signe en considérant chacune des deux quantités qui la définissent. Nous allons voir que la première quantité est toujours positive et la seconde peut l'être sous une condition qui va fixer la CFL.

- Étude de la matrice  $\mathcal{M}_p - \frac{\Delta t^2}{4} \left( \mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \right)$ . La positivité de cette matrice est équivalente à la positivité de

$$\mathcal{I} - \frac{\Delta t^2}{4} \left( \mathcal{M}_p^{-1} \mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP} \right).$$

Elle est donc assurée si et seulement

$$1 - \frac{\Delta t^2}{4} \left( \lambda - \frac{\Delta t^2}{12} \lambda^2 \right) > 0$$

pour toute valeur propre  $\lambda$  de  $\mathcal{M}_p^{-1} \mathcal{K}_{IP}$ . Cette condition peut se réécrire sous la forme

$$(\Delta t^2 \lambda - 6)^2 + 12 > 0,$$

ce qui est bien évidemment toujours vérifié.

- Étude de la matrice  $\mathcal{K}_{IP} - \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{M}_p^{-1} \mathcal{K}_{IP}$ . En multipliant par  $\mathcal{M}_p^{-1}$  qui est symétrique définie positive, on montre que cette matrice est positive si et seulement si

$$\lambda - \frac{\Delta t^2}{12} \lambda^2 > 0,$$

pour toute valeur propre  $\lambda$  de  $\mathcal{M}_p^{-1} \mathcal{K}_{IP}$ . Comme  $\mathcal{M}_p^{-1} \mathcal{K}_{IP}$  est définie positive, cette condition devient  $\Delta t^2 < \frac{12}{\lambda}$ .

Finalement, nous obtenons la condition CFL  $\Delta t^2 < \frac{12}{\lambda}$ , qui s'écrit également

$$\Delta t \leq \sqrt{3} \alpha_{IP} h.$$

La condition CFL de ce schéma est donc  $\sqrt{3}$  fois plus élevée que celle du schéma Leap-Frog classique. L'augmentation des coûts de calcul, générée par la nécessité de multiplier deux fois par la matrice de raideur, est donc compensée.

Ici encore, on pourrait penser qu'il suffit de calculer et de stocker la matrice  $\mathcal{K}_{IP} + \frac{\Delta t^2}{12} \mathcal{K}_{IP} \mathcal{K}_{IP}$ . Pour les mêmes raisons que pour la formulation d'ordre un, on préfère faire le calcul en deux temps en utilisant une variable auxiliaire.

### 3.2.3 Le schéma biharmonique

Nous présentons dans cette section le schéma associé à l'équation des ondes scalaire qui correspond au schéma  $\nabla$  pour l'équation des ondes vectorielles. Ce schéma, appelé schéma biharmonique (ou  $\Delta^2$ ), s'obtient en effectuant la discrétisation en temps avant de considérer la discrétisation en espace. En utilisant un développement de Taylor sur la solution du problème continu (2.1.1) et en remplaçant les dérivées en temps par des opérateurs en espace à l'aide de (2.1.1), on obtient

$$(3.2.6) \quad \begin{aligned} \frac{1}{\mu} \frac{p^{n+1} - 2p^n + p^{n-1}}{\Delta t^2} &= \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) + \frac{\Delta t^2}{12} \nabla \cdot \left( \frac{1}{\rho} \nabla \left[ \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) \right] \right) \\ &+ f^n + \frac{\Delta t^2}{12} \left( f_t^n \nabla \cdot \left( \frac{1}{\rho} \nabla f^n \right) + f_t^n \right). \end{aligned}$$

Pour obtenir la formulation variationnelle associée à ce problème, on multiplie (3.2.6) par une fonction test  $q$ , localement continue, on intègre sur  $K$  et on somme sur tous les éléments  $K$ . On obtient alors, après symétrisation et pénalisation des termes d'ordre deux

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \int_K \frac{1}{\mu} \frac{p^{n+1} - 2p^n + p^{n-1}}{\Delta t^2} q + a(p^n, q) - \frac{\Delta t^2}{12} \sum_{K \in \mathcal{T}_h} \int_K \nabla \cdot \left( \frac{1}{\rho} \nabla \left[ \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) \right] \right) q \\ &= \sum_{K \in \mathcal{T}_h} \int_K \left( f_t^n + \frac{\Delta t^2}{12} \frac{\mu}{\rho} \Delta f_t^n + f_t^{n2} \right) q, \end{aligned}$$

où  $a(p, q)$  est la forme bilinéaire définie précédemment pour discrétiser l'opérateur Laplacien. Pour l'opérateur d'ordre élevé, on utilise deux fois la formule de Green pour obtenir

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \int_K \nabla \cdot \left( \frac{1}{\rho} \nabla \left[ \mu \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) \right] \right) q = \sum_{K \in \mathcal{T}_h} \int_K \mu \left[ \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) \right] \left[ \nabla \cdot \left( \frac{1}{\rho} \nabla q \right) \right] \\ & - \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left[ q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] + \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla q \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] \\ & - \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p + \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} \frac{1}{\rho} \nabla q \mu \nabla \cdot \frac{1}{\rho} \nabla p \end{aligned}$$

On considère ensuite successivement chaque terme de bord.

1. Pour une face interne :

$$\int_{\Gamma} \left[ \left[ q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] = \int_{\Gamma} \llbracket q \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right\} \right\} + \int_{\Gamma} \{\{q\}\} \left[ \left[ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right].$$

En dérivant deux fois par rapport au temps la condition de transmission  $\left[ \left[ \frac{1}{\rho} \nabla p \cdot \mathbf{n} = 0 \right] \right]$  et en utilisant le problème continu (2.1.1) on vérifie que, pour  $p$  suffisamment régulière,

$$\int_{\Gamma} \{\{q\}\} \left[ \left[ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] = 0.$$

La condition de transmission  $\llbracket p \rrbracket = 0$  nous permet de symétriser la formulation en ajoutant le terme

$$\int_{\Gamma} \llbracket p \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q \right\} \right\}.$$

Finalement

$$\int_{\Gamma} \left[ \left[ q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] = \int_{\Gamma} \llbracket q \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right\} \right\} + \int_{\Gamma} \llbracket p \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q \right\} \right\}.$$

2. On montre de même que, pour une face interne :

$$\sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla q \mu \nabla \cdot \frac{1}{\rho} \nabla p \right] \right] = \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla q \right] \right] \left\{ \left\{ \mu \nabla \cdot \frac{1}{\rho} \nabla p \right\} \right\} + \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla p \right] \right] \left\{ \left\{ \mu \nabla \cdot \frac{1}{\rho} \nabla q \right\} \right\}.$$

3. Pour une face externe, si on utilise une condition de Dirichlet, on a

$$\int_{\Gamma} p \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q = 0,$$

ce qui nous permet d'écrire

$$\int_{\Gamma} q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p = \int_{\Gamma} q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p + \int_{\Gamma} p \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q$$

4. En dérivant deux fois la condition de Dirichlet par rapport au temps et en utilisant le problème continu (2.1.1), on vérifie que pour  $p$  suffisamment régulière

$$\int_{\Gamma} \frac{1}{\rho} \nabla q \mu \nabla \cdot \frac{1}{\rho} \nabla p = 0$$

sur une face externe.

Finalement la formulation variationnelle continue s'écrit

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \int_K \frac{1}{\mu} \frac{p^{n+1} - 2p^n + p^{n-1}}{\Delta t^2} q + a(p^n, q) + \frac{\Delta t^2}{12} a_{\Delta^2}(p^n, q) \\ &= \sum_{K \in \mathcal{T}_h} \int_K \left( f_t^n + \frac{\Delta t^2}{12} \frac{\mu}{\rho} \Delta f_t^n + f_{t^2}^n \right) q, \end{aligned}$$

où  $a_{\Delta^2}(p^n, q)$  est la forme bilinéaire symétrique définie par

$$\begin{aligned} a_{\Delta^2}(p^n, q) &= - \sum_{K \in \mathcal{T}_h} \int_K \mu \left[ \nabla \cdot \left( \frac{1}{\rho} \nabla p^n \right) \right] \left[ \nabla \cdot \left( \frac{1}{\rho} \nabla q \right) \right] \\ &+ \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \llbracket q \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p \right\} \right\} + \int_{\Gamma} \llbracket p \rrbracket \left\{ \left\{ \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q \right\} \right\} \\ &- \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{int}}} \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla q \right] \right] \left\{ \left\{ \mu \nabla \cdot \frac{1}{\rho} \nabla p \right\} \right\} + \int_{\Gamma} \left[ \left[ \frac{1}{\rho} \nabla p \right] \right] \left\{ \left\{ \mu \nabla \cdot \frac{1}{\rho} \nabla q \right\} \right\} \\ &+ \frac{\Delta t^2}{12} \sum_{\Gamma \in \Gamma_{\text{ext}}} \int_{\Gamma} q \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla p + \int_{\Gamma} p \frac{1}{\rho} \nabla \mu \nabla \cdot \frac{1}{\rho} \nabla q. \end{aligned}$$

et la formulation variationnelle s'écrit

Trouver  $p_h^{n+1}$  dans  $V_h^m$  tel que, pour tout  $q$  dans  $V_h^m$ , on ait

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \int_K \frac{1}{\mu} \frac{p_h^{n+1} - 2p_h^n + p_h^{n-1}}{\Delta t^2} q + a(p_h^n, q) + \frac{\Delta t^2}{12} a_{\Delta^2}(p_h^n, q) \\ &= \sum_{K \in \mathcal{T}_h} \int_K \left( f_t^n + \frac{\Delta t^2}{12} \frac{\mu}{\rho} \Delta f_t^n + f_{t^2}^n \right) q, \end{aligned}$$

En utilisant les fonctions de base de Lagrange classiques, on obtient le système totalement discrétisé

$$(3.2.7) \quad \mathcal{M}_p \frac{\mathbf{P}^{n+1} - 2\mathbf{P}^n + \mathbf{P}^{n-1}}{\Delta t^2} = \mathcal{K}_{IP} \mathbf{P}^n + \frac{\Delta t^2}{12} \mathcal{K}_{\Delta^2} \mathbf{P}^n + \mathbf{F}^n + \frac{\Delta t^2}{12} \mathbf{F}_{\Delta^2}^n$$

où  $\mathcal{K}_{\Delta^2}$  désigne la matrice obtenue après discrétisation de l'opérateur d'ordre élevé. Elle est telle que  $\mathcal{K}_{IP}^{K,J} = 0$  si  $K \cap J = \emptyset$ ,

$$\mathcal{K}_{IP}^{K,J} = a_{\Delta^2}(\phi_j^J, \phi_i^K),$$

si  $K \cap J = \Gamma$ , et

$$\mathcal{K}_{IP}^{K,K} = a_{\Delta^2}(\phi_j^K, \phi_i^K).$$

Le vecteur  $\mathbf{F}(t)$  est tel que

$$\mathbf{F}_i^K(t) = \int_{\Omega} f_{t^2}^n + \frac{\mu}{\rho} \Delta f^n \phi_i^K.$$

Les matrices  $\mathcal{K}_{IP}$  et  $\mathcal{K}_{\Delta^2}$  ayant des stencils similaires (à nouveau, chaque élément communique uniquement avec lui-même et ses voisins), on peut stocker sans coûts supplémentaires la matrice  $\mathcal{K}_{IP} + \frac{\Delta^2}{12} \mathcal{K}_{\Delta^2}$ . Les résultats présentés dans [1, 3] indiquent que la condition CFL de ce schéma est sensiblement la même que celle du Leap-Frog. Les auteurs obtiennent également ce résultat pour l'ordre six, de sorte que les schémas bi et tri-harmoniques ont le même coût global que le schéma LF.

Dans ce cas, la condition CFL est déterminée numériquement. En effet, on peut montrer la conservation de la quantité

$$\begin{aligned} E^{n+\frac{1}{2}} &= \left\langle \left( \mathcal{M}_p - \frac{\Delta t^2}{4} \left( \mathcal{K}_{IP} + \frac{\Delta^2}{12} \mathcal{K}_{\Delta^2} \right) \right) \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t}, \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} \right\rangle \\ &+ \left\langle \left( \mathcal{K}_{IP} + \frac{\Delta^2}{12} \mathcal{K}_{\Delta^2} \right) \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2}, \frac{\mathbf{P}^{n+1} + \mathbf{P}^n}{2} \right\rangle \end{aligned}$$

mais, comme pour le schéma  $\nabla$ , les valeurs propres dépendent de  $\Delta t$  et les deux matrices ne sont pas diagonalisables dans une même base. Il ne semble donc pas possible de déterminer une expression analytique de la CFL.

### 3.3 Résultats Numériques

Nous allons maintenant analyser les performances des schémas  $\nabla$  et ADER pour différentes configurations. Nous commençons par un cas test académique avec des données initiales régulières (section 3.3.1) puis nous considérons un modèle géophysique de référence, le modèle Marmousi (section 3.3.2) dans lequel on pose les mêmes équations avec cette fois une source ponctuelle et des données initiales identiques à zéro. Ce cas test va nous permettre de mettre en évidence l'apparition d'ondes parasites que nous analysons à la section 3.3.3.

#### 3.3.1 Cas test académique

On résout le système :

$$(3.3.1) \quad \begin{cases} \rho(\mathbf{x}) \frac{\partial \mathbf{v}(\mathbf{x}, t)}{\partial t} + \nabla p(\mathbf{x}, t) = 0 & \text{dans } \Omega \times [0, T] \\ \frac{1}{\mu(\mathbf{x})} \frac{\partial p(\mathbf{x}, t)}{\partial t} + \nabla \cdot (\mathbf{v}(\mathbf{x}, t)) = 0 & \text{dans } \Omega \times [0, T] \end{cases}$$

avec des conditions de bords périodiques :

$$(3.3.2) \quad \begin{cases} p(x, 0, t) = p(x, L, t) & \text{pour } x \in [0, L] \text{ et } t \in [0, T], \\ p(0, y, t) = p(L, y, t) & \text{pour } y \in [0, L] \text{ et } t \in [0, T], \\ v_x(0, y, t) = v_x(L, y, t) & \text{pour } y \in [0, L] \text{ et } t \in [0, T], \\ v_y(y, 0, t) = v_y(y, L, t) & \text{pour } y \in [0, L] \text{ et } t \in [0, T]. \end{cases}$$

Nous avons réalisé des comparaisons 2D entre le schéma LF classique et les schémas DG-ADER et  $\nabla$  d'ordre 4, en utilisant des éléments  $\mathcal{P}^3$ ,  $\mathcal{P}^4$  et  $\mathcal{P}^5$ . Le domaine d'étude est un carré de côté  $L=1000\text{m}$ . On suppose que la vitesse de propagation est constante et on choisit pour module de compressibilité  $\mu = 6.25 \text{ GPa}$  et pour masse volumique  $\rho = 1\text{kg}\cdot\text{m}^{-3}$ . La vitesse de propagation des ondes est ainsi  $c = \sqrt{\mu/\rho} = 2500\text{m}\cdot\text{s}^{-1}$ . Le temps de simulation est de 20 secondes. Le domaine est recouvert par un maillage régulier composé de triangles rectangles isocèles. Le pas de discrétisation est noté  $h$ . Il est d'abord fixé à 50m puis le maillage est raffiné en divisant  $h$  par un facteur noté  $\delta$  qui varie de 1 à 16. Le plus petit pas d'espace est donc égal à 3.125m. Les maillages structurés utilisés pour  $\delta = 1, 2, 4$  et 8 sont représentés sur les figures 3.2.(a) à 3.2.(d). Les données initiales sont définies par

$$p(x, t) = \sin\left(\frac{2\pi}{\lambda_0}(k_x x + k_y y - ct)\right),$$

où  $k_x$  et  $k_y$  sont égaux à  $\frac{\sqrt{2}}{2}$  et  $\lambda_0 = 250\frac{\sqrt{2}}{2} \approx 180\text{m}$  est la longueur d'onde caractéristique du problème. On choisit donc

$$p(x, 0) = \sin\left(\frac{2\pi}{\lambda_0}(k_x x + k_y y)\right),$$

$$\mathbf{v}_x(x, 0) = -\mu\frac{k_x}{c} \sin\left(\frac{2\pi}{\lambda_0}(k_x x + k_y y)\right) \text{ et } \mathbf{v}_y(x, 0) = -\mu\frac{k_y}{c} \sin\left(\frac{2\pi}{\lambda_0}(k_x x + k_y y)\right).$$

Le nombre d'éléments de chaque maillage ainsi que le nombre d'éléments par longueur d'onde sont indiqués dans le tableau 3.1.

$\delta$	Nb. Élément	Nb. Élément / $\lambda$
1	800	3.5
2	3200	7
4	12800	14
8	51200	28
16	204800	56

TABLE 3.1 – Nombre d'éléments constituant les maillages 3.2 ainsi que le nombre d'éléments par longueur d'onde ( $\lambda_0 = 180\text{m}$ ).

Les tableaux 3.2 (resp. 3.3 et 3.4) indiquent les erreurs relatives  $L^2$  et le pas de temps pour un facteur de raffinement  $\delta$  donné ainsi que l'ordre de convergence (CV) entre deux raffinements, pour les schémas LF (a),  $\nabla$  (b) et ADER (c) et des éléments  $\mathcal{P}^3$  (resp.  $\mathcal{P}^4$  et  $\mathcal{P}^5$ ). Remarquons que pour des éléments  $\mathcal{P}^3$ , nous n'obtenons pas la convergence d'ordre quatre attendue. Cela est dû au fait que

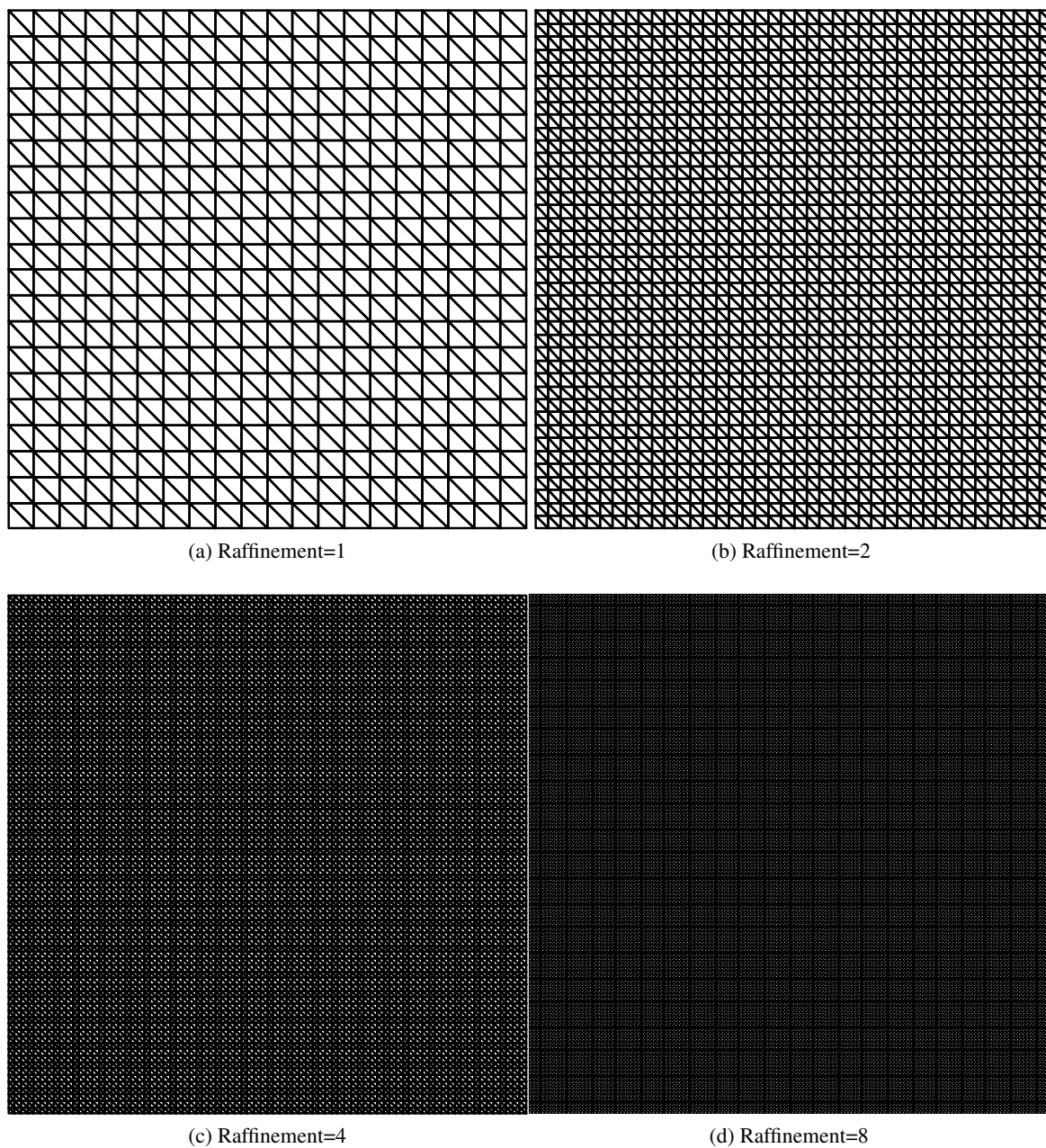


FIGURE 3.2 – Illustration des maillages structurés utilisés.

l'ordre de convergence de la méthode DG avec flux centré n'est pas optimal, comme cela a été illustré et analysé dans [24, 32, 12]. En effet, pour une approximation de degré  $m$ , on s'attend à un ordre de convergence de l'ordre de  $h^{m+1}$  et on observe que l'ordre de convergence est entre 3 et 4. Il suit donc au mieux la courbe  $h^m$ . L'ordre 4 ne peut être donc atteint que si on utilise des polynômes de degré au moins égal à 4. On peut remarquer que les tableaux 3.3 et 3.4 indiquent même un ordre de convergence plus élevé pour le schéma  $\nabla$  sur des maillages grossiers. Pour mieux visualiser la convergence des schémas, on représente également l'erreur en fonction du pas d'espace sur les figures 3.3.(b), 3.4.(b) et 3.5.(b).

Comme nous pouvions nous y attendre, pour un pas d'espace donné, les schémas  $\nabla$  et ADER sont clairement meilleurs que le schéma LF, quel que soit l'ordre des éléments finis. Pour des éléments  $\mathcal{P}^3$ , les schémas ADER et  $\nabla$  démontrent des performances comparables. Le schéma  $\nabla$  donne de meilleurs résultats que le schéma ADER quand l'ordre des éléments augmente. Notamment, pour des éléments  $\mathcal{P}^5$ , le schéma  $\nabla$  est dix fois plus précis que le schéma ADER.

$\delta$	dt	Erreur	CV
1	1.9e-03	1.5e-01	
2	9.3e-04	3.7e-02	2.0
4	4.7e-04	9.1e-03	2.0
8	2.3e-04	2.3e-03	2.0
16	1.2e-04	5.7e-04	2.0

(a) LF

dt	Erreur	CV
1.7e-03	1.6e-04	
8.7e-04	1.3e-05	3.7
4.4e-04	1.2e-06	3.3
2.2e-04	1.3e-07	3.3
1.1e-04	1.1e-08	3.6

(b)  $\nabla$ 

dt	Erreur	CV
5.2e-03	1.4e-04	
2.6e-03	1.3e-05	3.5
1.3e-03	1.3e-06	3.3
6.5e-04	1.5e-07	3.1
3.3e-04	1.8e-08	3.0

(c) ADER

TABLE 3.2 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^3$ . On remarque à nouveau que la convergence n'est que d'ordre trois.

$\delta$	dt	Erreur	CV
1	1.3e-03	6.7e-02	
2	6.3e-04	1.7e-02	2.0
4	3.2e-04	4.2e-03	2.0
8	1.6e-04	1.0e-03	2.0
16	7.9e-05	2.6e-04	2.0

(a) LF

dt	Erreur	CV
1.2e-03	3.5e-06	
6.0e-04	1.6e-07	4.5
3.0e-04	8.8e-09	4.2
1.5e-04	5.3e-10	4.0
7.5e-05	3.3e-11	4.0

(b)  $\nabla$ 

dt	Erreur	CV
3.5e-03	2.4e-05	
1.8e-03	1.5e-06	4.0
8.8e-04	9.2e-08	4.0
4.4e-04	5.8e-09	4.0
2.2e-04	3.2e-10	4.2

(c) ADER

TABLE 3.3 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^4$ . On obtient bien la convergence d'ordre 4.

Cependant, comparer les erreurs sur un maillage donné n'est pas suffisant pour conclure qu'un schéma est meilleur qu'un autre. Il faut aussi tenir compte des coûts de calcul générés par le schéma.



$\delta$	dt	Erreur	CV
1	9.0e-04	3.4e-02	
2	4.5e-04	8.5e-03	2.0
4	2.2e-04	2.1e-03	2.0
8	1.1e-04	5.3e-04	2.0
16	5.6e-05	7.6e-05	2.8

(a) LF

dt	Erreur	CV
8.4e-04	5.3e-07	
4.2e-04	3.3e-08	4.0
2.1e-04	2.0e-09	4.0
1.0e-04	1.3e-10	4.0
5.2e-05	3.3e-12	5.2

(b)  $\nabla$

dt	Erreur	CV
2.5e-03	6.2e-06	
1.3e-03	3.8e-07	4.0
6.3e-04	2.4e-08	4.0
3.1e-04	1.5e-09	4.0
1.6e-04	3.4e-11	5.5

(c) ADER

TABLE 3.4 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^5$ . On obtient bien la convergence d'ordre 4.

Ces coûts dépendent du nombre d'éléments non nuls au sein de la matrice de rigidité mais aussi du nombre de pas de temps à effectuer pour réaliser la simulation. Pour poursuivre l'analyse dans cette direction, on introduit  $N$  le nombre de triangles du maillage et  $N_\phi$  le nombre de degrés de liberté d'un élément. On a  $N = 2 \left( \frac{1000}{h} \right)^2$  et  $N_\phi = \frac{(l+1)(l+2)}{2}$  où  $l$  désigne le degré des polynômes d'approximation. La taille des matrices de rigidité est  $2N_\phi^2 N \times N_\phi^2 N$  pour  $\mathcal{A}_p$  et  $\mathcal{B}_p$ , et  $N_\phi^2 N \times 2N_\phi^2 N$  pour  $\mathcal{A}_v$  et  $\mathcal{B}_v$ . Le nombre d'éléments non nuls est  $4N \times 2N_\phi^2$ . Le facteur 4 vient du fait qu'un élément communique avec lui-même mais aussi avec ses trois voisins. Le facteur 2 correspond aux deux composantes de la vitesse.

On peut estimer à  $2 \times 8NN_\phi^2$  le coût d'une itération du schéma LF et du schéma  $\nabla$  car on doit faire deux multiplications par les matrices de rigidité à chaque itération. Le coût d'une itération pour le schéma ADER est égal à  $6 \times 8NN_\phi^2$  car on doit réaliser six multiplications par les matrices de rigidité à chaque itération. Le coût total de calcul est obtenu en multipliant les coûts élémentaires précédents par le nombre de pas de temps qui vaut  $60/\Delta t$ . Par exemple, le coût de calcul pour atteindre une précision de  $10^{-3}$  (0.1%) est  $1.2 \cdot 10^{11}$  pour le schéma LF- $\mathcal{P}_3$ ,  $2.5 \cdot 10^8$  pour le schéma  $\nabla$ - $\mathcal{P}_3$  et  $4 \cdot 10^8$  pour le schéma ADER- $\mathcal{P}_3$ . Pour atteindre cette fois une précision de l'ordre de  $10^{-4}$ , les coûts numériques sont  $3.2 \cdot 10^{11}$  pour le schéma LF- $\mathcal{P}_4$ ,  $8.1 \cdot 10^8$  pour le schéma  $\nabla$ - $\mathcal{P}_4$  et  $1.3 \cdot 10^9$  pour ADER- $\mathcal{P}_4$ .

Ces résultats nous montrent que les schémas  $\nabla$  génèrent moins de calculs que le schéma LF (jusqu'à 100 fois moins) et que le schéma ADER (jusqu'à deux fois moins). Cet avantage est d'autant plus net qu'on augmente le niveau de précision. Le schéma  $\nabla$ - $\mathcal{P}_5$  est capable d'atteindre un niveau de précision de l'ordre de  $10^{-10}$  pour un coût de  $1.2 \cdot 10^{12}$  alors que le schéma ADER- $\mathcal{P}_5$  a un coût de  $1.5 \cdot 10^{13}$  avec des éléments  $\mathcal{P}_5$ ). Pour mieux mettre en évidence ces résultats, nous représentons l'erreur en fonction des coûts de calcul pour des éléments  $\mathcal{P}_3$  (resp.  $\mathcal{P}_4$  et  $\mathcal{P}_5$ ) sur la figure 3.3.(a) (resp. 3.4.(a) et 3.5.(a)). On remarque que les courbes du schéma  $\nabla$  sont bien en dessous de celles du schéma ADER, ce qui confirme que pour une précision donnée le schéma  $\nabla$  génère moins de calcul.

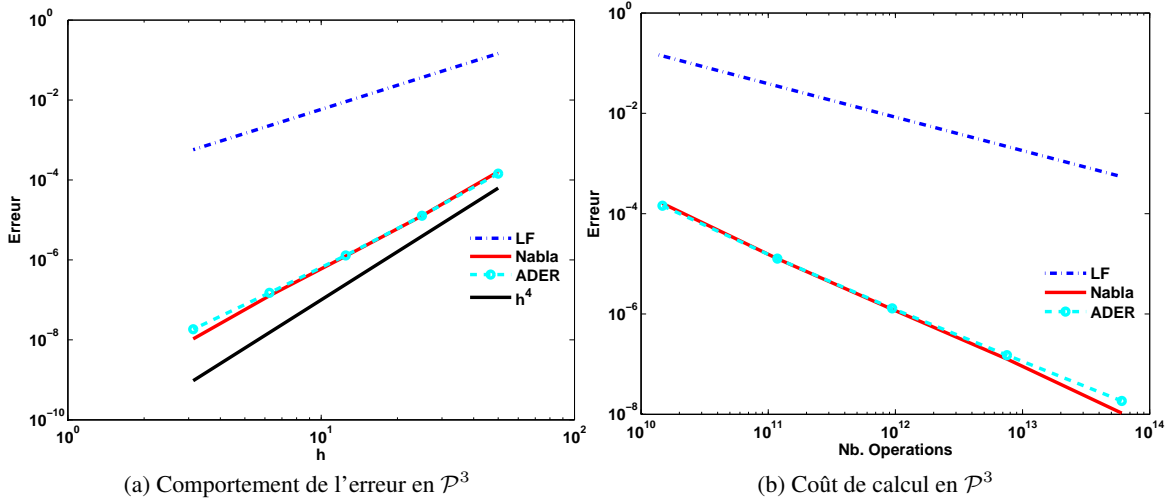


FIGURE 3.3 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 3. On remarque que les schémas d'ordre élevé ne convergent qu'à l'ordre 3 avec des éléments finis de degré 3.

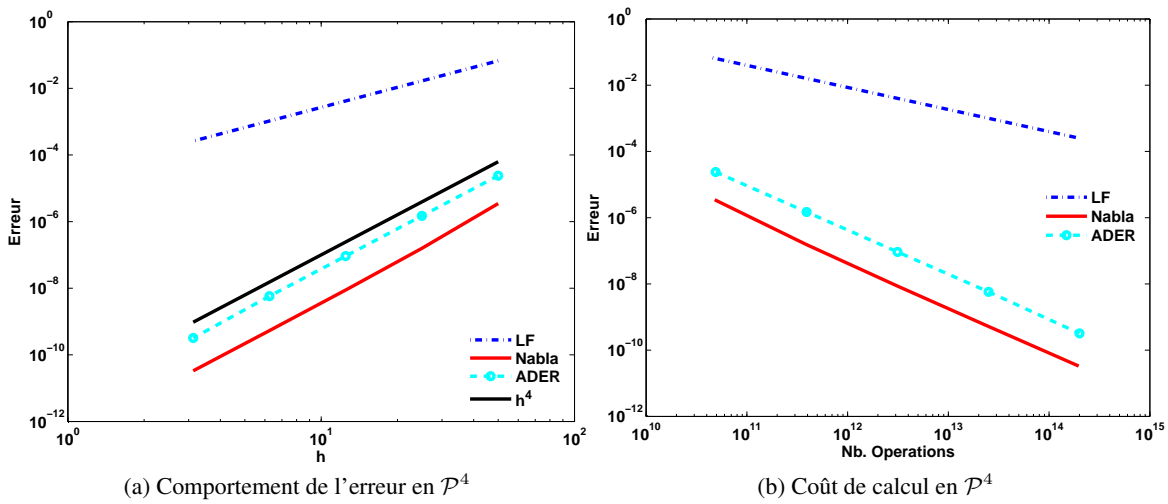


FIGURE 3.4 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 4.

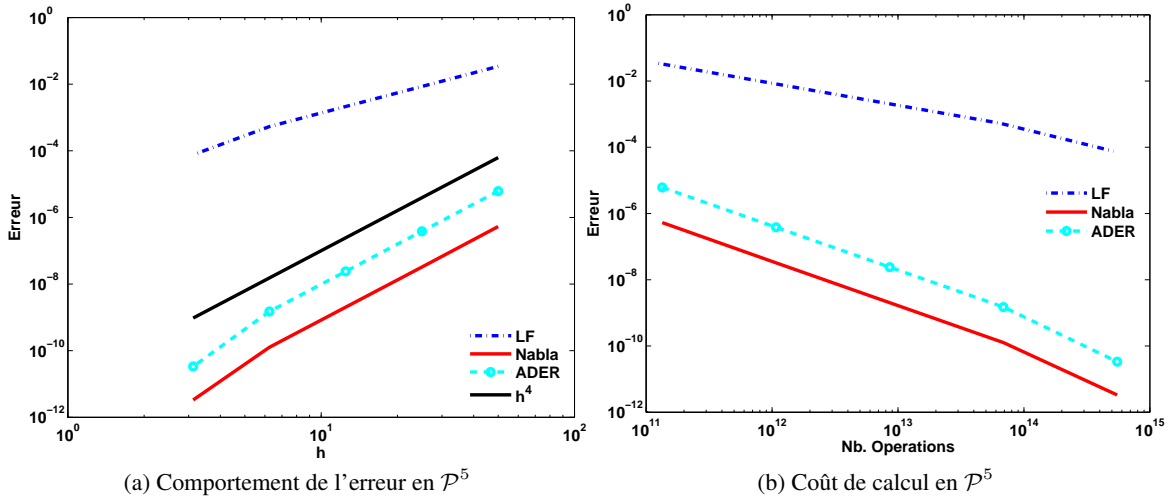


FIGURE 3.5 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 5.

Nous avons obtenu des résultats très prometteurs sur un cas test académique, à l'avantage du schéma  $\nabla$ . Toutefois, cela ne suffit pas pour conclure sur l'intérêt d'appliquer ce schéma pour des applications géophysiques. Il est maintenant nécessaire de considérer un cas plus réaliste et nous avons alors choisi de tester le schéma  $\nabla$  en prenant comme modèle de vitesse le modèle Marmousi. Cependant, les calculs n'ont pas été réalisés immédiatement car il a été nécessaire de porter nos algorithmes dans la plate-forme de calcul DIVA, ce qui a pris un certain temps car DIVA est complètement optimisée. Les algorithmes ont donc été réécrits dans un environnement HPC non trivial.

### 3.3.2 Validation sur benchmark : le modèle Marmousi

Le cas test Marmousi a été créé en 1988 par l'Institut Français du Pétrole (IFP). La géométrie de ce modèle 2D est basée sur un site localisé en Angola dans le bassin de Kwenza. La géométrie ainsi que le modèle de vitesse ont été créés afin de produire des données sismiques complexes dont l'utilisation requiert des méthodes de calcul avancées afin d'obtenir une image correcte du sous-sol. Ce modèle est de nos jours un modèle industriel de validation incontournable.

Le modèle Marmousi contient 158 couches. Le domaine est large de 9.5 km et profond de 3 km. Le modèle de vitesse contient 751 points de données dans la direction  $z$  et 2301 dans la direction  $x$ . Les vitesses varient de  $1500 \text{ ms}^{-1}$  à  $5500 \text{ ms}^{-1}$  et la masse volumique est supposée constante  $\rho = 1 \text{ kgm}^{-3}$ . La source est une fonction de Ricker de la forme  $f(t) = (1 + 2\pi^2 \times f_0^2 t^2) \exp(-\pi^2 \times f_0^2 t^2)$  de fréquence centrale  $f_0=6\text{Hz}$  située au point de coordonnées (4875m,4m). Nous avons utilisé 96 récepteurs positionnés à 4m de profondeur et dont les abscisses varient de 2300m à 4675m. Le temps de simulation est de 6.2s, nous utilisons des conditions de bord de Dirichlet sur le haut du domaine et des PML pour les trois autres bords. Le domaine est recouvert par 21469 triangles.

Si nous effectuons une simple propagation en utilisant le schéma LF classique, les sismogrammes sont très bruités quelle que soit la discrétisation en espace (Figs. 3.7 (a), (b) et (c) pour des éléments  $\mathcal{P}^1$ ,  $\mathcal{P}^2$  et  $\mathcal{P}^3$ ). Remarquons que ce phénomène se manifeste de manière différente suivant l'ordre des éléments finis : le bruit est très fort pour les éléments  $\mathcal{P}^1$  et relativement faible pour des éléments  $\mathcal{P}^2$ . Pour des éléments  $\mathcal{P}^3$ , le bruit se propage plus vite que l'onde physique alors qu'il se propage

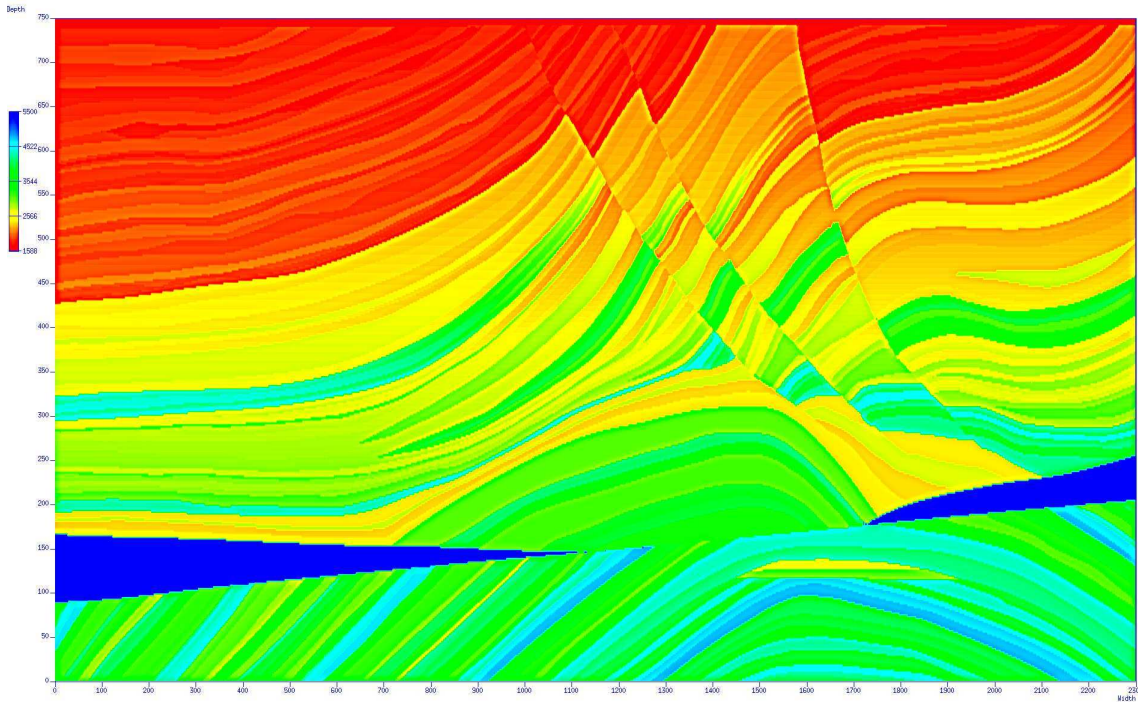


FIGURE 3.6 – Domaine d'étude Marmousi.

plus lentement pour des éléments  $\mathcal{P}^1$  et  $\mathcal{P}^2$ . Ces différences montrent que ce phénomène n'est pas physique mais uniquement dû au schéma numérique. Nous pensions pouvoir supprimer ce bruit en utilisant des schémas d'ordre élevé en temps, mais nous avons obtenu des résultats similaires avec les schémas ADER et  $\nabla$ .

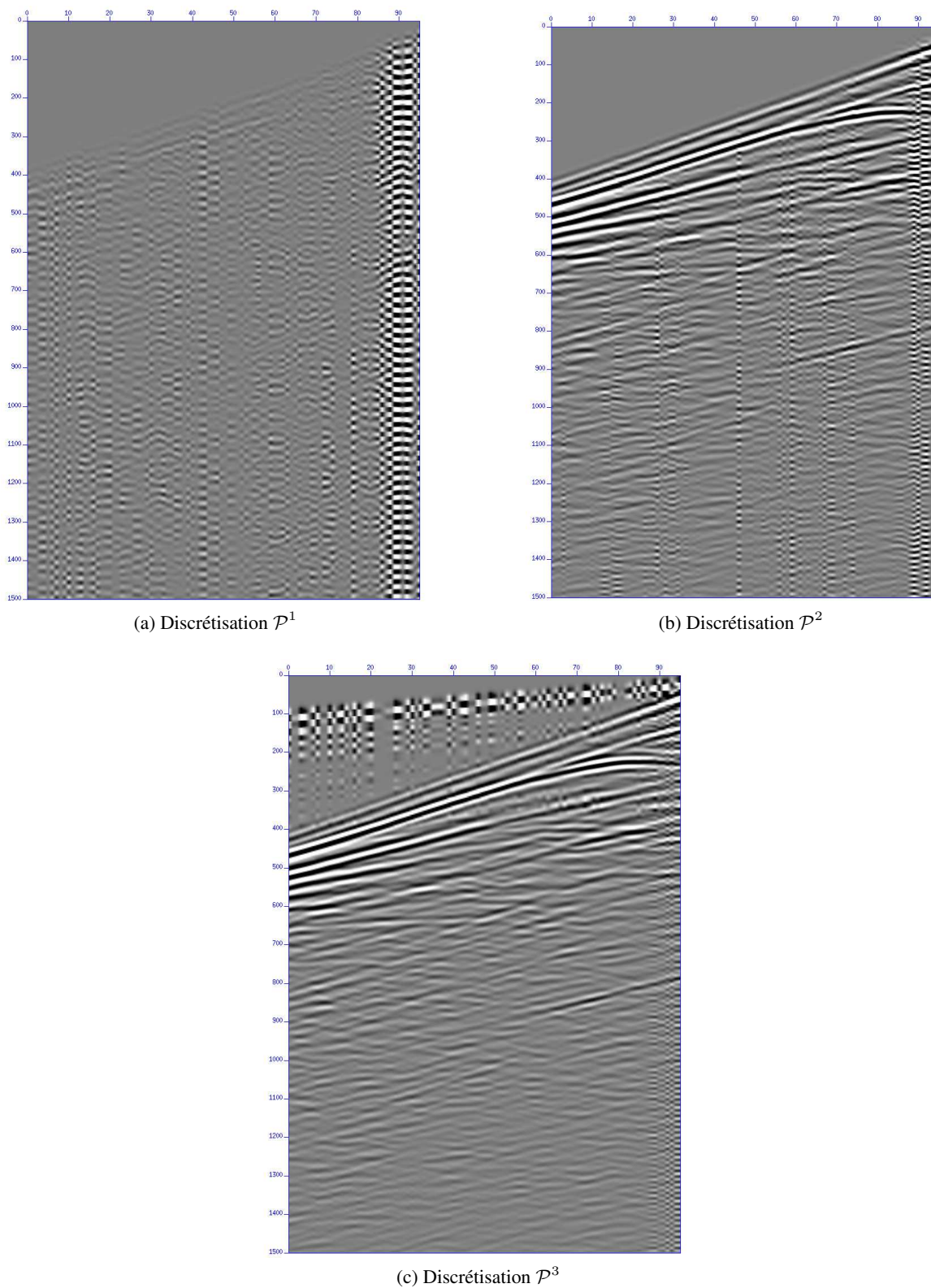


FIGURE 3.7 – Sismogrammes Marmousi avec bruit numérique pour des éléments  $\mathcal{P}^1$  à  $\mathcal{P}^3$  utilisant le schéma LF classique

Comme l'utilisation des deux schémas en temps d'ordre élevé n'a pas contribué à éliminer le bruit numérique, nous avons cherché quelles pouvaient être les causes de ce phénomène de pollution numérique. Nous en avons trouvé quatre possibles : le modèle est fortement hétérogène ; le maillage utilisé est non structuré, l'onde est générée par une source ponctuelle et les calculs sont effectués en simple précision. Nous avons donc choisi de développer un code prototype en milieu homogène sur une grille structurée et en double précision. C'est ce code que nous avons utilisé pour calculer les solutions du problème avec des conditions initiales. Nous l'avons adapté au cas de la source ponctuelle et nous avons mené une analyse de convergence que nous présentons à la section suivante.

### 3.3.3 Analyse du cas avec une source : le cas de maillages réguliers

Nous considérons à nouveau le système (3.3.1) et nous imposons cette fois une source ponctuelle au lieu des conditions initiales. La source est une fonction de Ricker de la forme  $f(t) = (1 + 2\pi^2 f_0^2 t^2) \exp(-\pi^2 f_0^2 t^2)$  ayant pour fréquence centrale  $f_0=20\text{Hz}$  et située au point de coordonnées (500m,500m). Nous choisissons à nouveau  $\rho = 1\text{kg.m}^{-3}$  et  $\mu = 6.25\text{ GPa}$ . Nous étudions plus précisément le signal enregistré par le récepteur placé au point (650m,650m). La longueur d'onde vaut  $\lambda = 2500/20 = 125\text{m}$ . La distance séparant le récepteur de la source étudiée est de 210 m. Nous représentons le sismogramme enregistré sur la figure 3.8.(a) avec des éléments  $\mathcal{P}^3$  et un pas d'espace  $h = 12.5\text{m}$  ( $\delta = 4$ ) et nous le comparons avec la solution analytique calculée à l'aide du logiciel Gar6more-2D (<https://team.inria.fr/magique3d/software/>). On observe du bruit numérique juste après l'onde incidente. Le récepteur a été placé suffisamment loin du bord du domaine pour que ce bruit ne corresponde pas aux réflexions multiples causées par le bord du domaine de calcul. Sur la figure 3.8.(b), on représente la solution obtenue avec un maillage quatre fois plus fin ( $\delta = 16$ ). On voit que le fait de raffiner le maillage ne diminue pas l'amplitude du bruit.

Nous avons représenté les résultats obtenus avec les schémas  $\nabla$  et ADER sur les figures 3.9.(a) et 3.9.(b). Nous remarquons que l'augmentation de l'ordre du schéma en temps ne permet pas non plus de réduire le bruit numérique. Pour mieux mettre cela en évidence, nous représentons sur les figures 3.10.(a), (b) et (c), les courbes de convergences des trois schémas, obtenus respectivement pour des éléments finis  $\mathcal{P}^3$ ,  $\mathcal{P}^4$  et  $\mathcal{P}^5$ . Nous observons que l'erreur est inchangée, quel que soit le pas de maillage et l'ordre du schéma en temps.

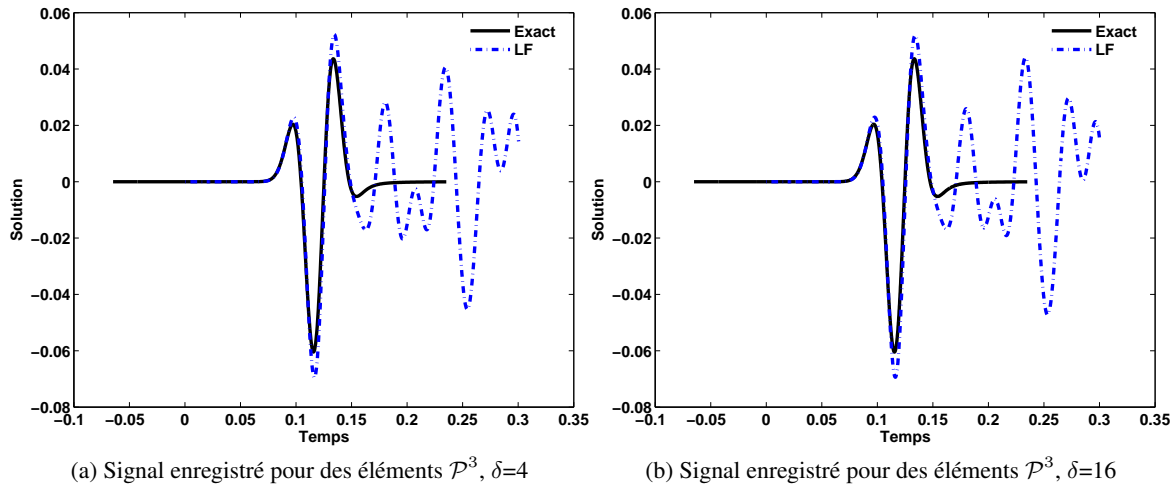
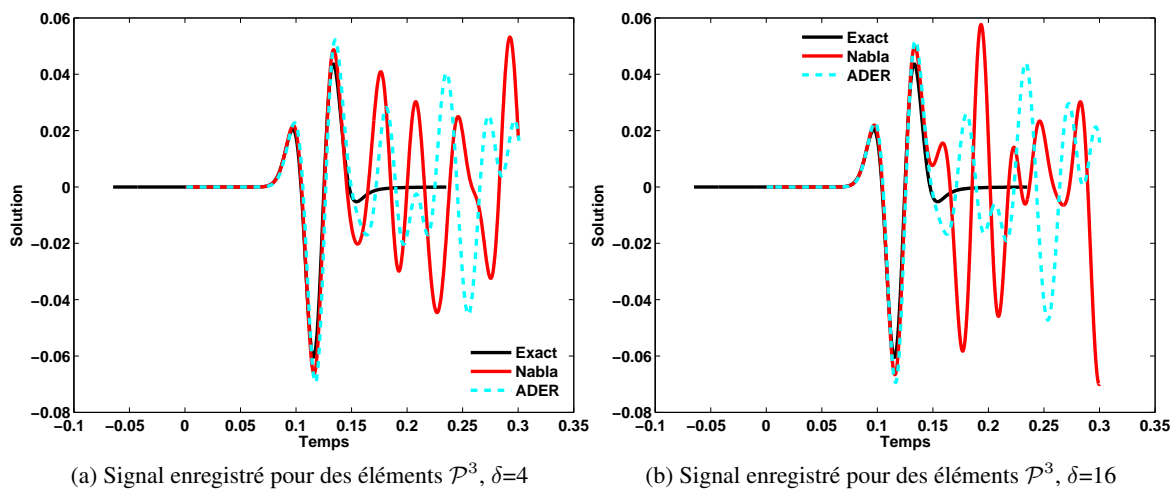
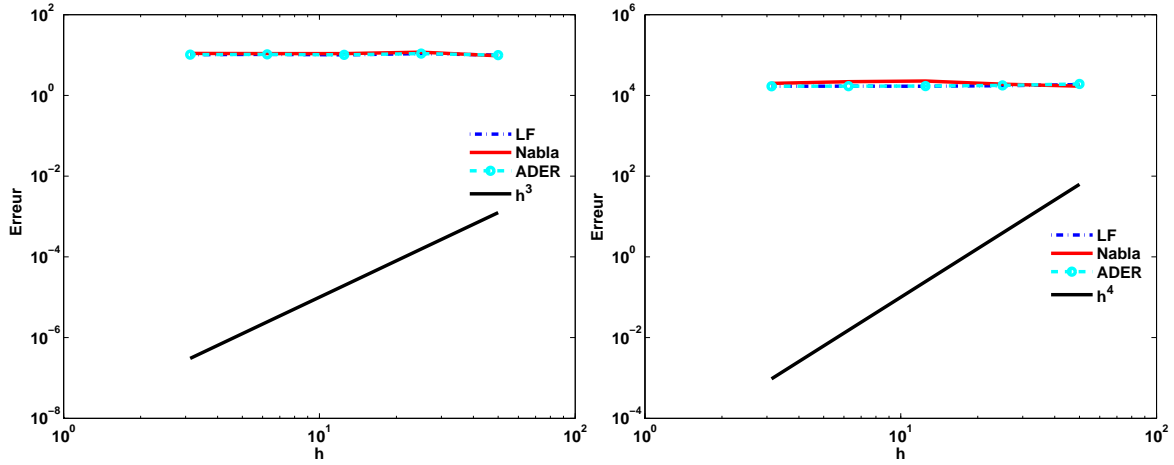


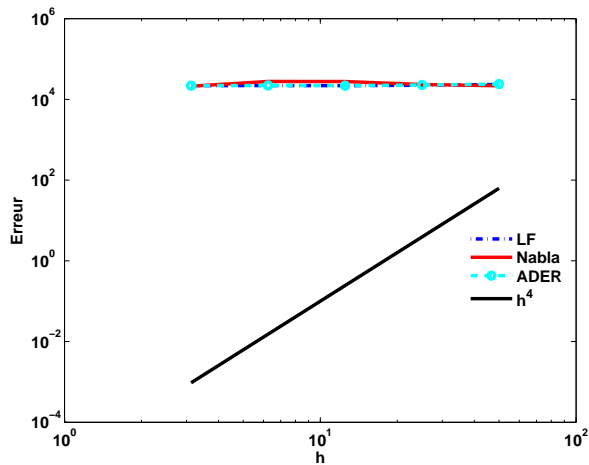
FIGURE 3.8 – Sismogrammes pour le schéma LF et pour des éléments finis de degré 3.

FIGURE 3.9 – Sismogrammes pour le schéma ADER et  $\nabla$  et pour des éléments finis de degré 3.



(a) Comportement de l'erreur pour des éléments  $\mathcal{P}^3$

(b) Comportement de l'erreur pour des éléments  $\mathcal{P}^4$



(c) Comportement de l'erreur en  $\mathcal{P}^5$

FIGURE 3.10 – Courbes d'erreur, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 3, 4 et 5.



### 3.4 Conclusion

Lorsqu'on considère des modèles académiques avec des données initiales régulières, le nouveau schéma d'ordre élevé que nous avons proposé permet d'obtenir des résultats plus précis avec des coûts de calcul plus faibles que le schéma ADER. Nous avons mis en œuvre ce schéma dans la plateforme DIVA dans le but de faire des tests sur des cas plus réalistes. En considérant la propagation d'une source ponctuelle dans le modèle Marmousi, nous observons l'apparition de modes parasites et en appliquant les schémas LF et ADER, on voit aussi des ondes parasites. On en a donc déduit que ces artéfacts viennent certainement de la discrétisation en espace. Pour en être plus sûrs, nous avons refait des calculs dans un modèle académique, mais en utilisant une source ponctuelle au lieu de données initiales régulières. On voit alors que même dans un cas test simple, l'injection d'une source génère des modes parasites. Ceci motive le chapitre suivant dans lequel nous présentons une analyse de dispersion numérique en dimension un pour étudier plus précisément les modes parasites.

### 3.A Calcul pratique du terme source

Comme nous l'avons vu, la plupart des applications géophysiques nécessitent la prise en compte de sources ponctuelles en espace, c'est-à-dire que  $f(\mathbf{x}, t) = \delta_{\mathbf{x}}g(t)$  où  $\delta$  représente la distribution de Dirac et  $g(t)$  est une fonction très régulière (i.e. le plus souvent  $C^\infty$ ). Dans ce cas, il n'est pas évident de prendre en compte les termes source dans la formulation variationnelle. Par exemple, il faut calculer  $\int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla f^{n+1} q$  qui fait intervenir le Laplacien de la mesure de Dirac. Ce calcul peut toutefois se faire au sens des distributions. En effet, on a :

**Propriété 3.A.1.** *Pour toute fonction test  $\phi$  de  $\mathcal{D}(\Omega)$ , on a*

$$\langle \delta_{\mathbf{x}_0}, \phi \rangle = \phi(\mathbf{x}_0), \quad \langle \nabla \delta_{\mathbf{x}_0}, \phi \rangle = \nabla \phi(\mathbf{x}_0), \quad \text{et} \quad \langle \Delta \delta_{\mathbf{x}_0}, \phi \rangle = \Delta \phi(\mathbf{x}_0).$$

où  $\langle \cdot, \cdot \rangle$  désigne le produit de dualité entre  $\mathcal{D}'(\Omega)$  et  $\mathcal{D}(\Omega)$ .

Supposons que  $\mathbf{x}_0$  soit interne à un seul élément  $K_{src}$  de la triangulation  $\mathcal{T}_h$ . Si cela n'est pas le cas,  $\mathbf{x}_0$  est sur une face, une arête ou un sommet et dans ce cas, on propose de déplacer  $\mathbf{x}_0$  très légèrement. Dans la pratique, si le point source appartient à plusieurs éléments, nous choisissons arbitrairement un de ces éléments, ce qui revient à déplacer  $\mathbf{x}_0$  d'une distance inférieure à la précision machine. Ceci nous permet de simplifier la mise en œuvre numérique, notamment lorsque nous devons traiter des problèmes contenant un grand nombre de sources ponctuelles.

Nous notons  $\delta_{h,\mathbf{x}_0}$ ,  $\delta_{h,\mathbf{x}_0}^\nabla$  et  $\delta_{h,\mathbf{x}_0}^\Delta$  les fonctions de  $V_h^m$  vérifiant respectivement

$$\int_{\Omega} \delta_{h,\mathbf{x}_0} \phi_i^K = \phi_i^K(\mathbf{x}_0), \quad \int_{\Omega} \frac{1}{\rho} \delta_{h,\mathbf{x}_0}^\nabla \phi_i^K = \frac{1}{\rho(\mathbf{x}_0)} \nabla \phi_i^K(\mathbf{x}_0)$$

et

$$\int_{\Omega} \frac{\mu}{\rho} \nabla \cdot \nabla \delta_{h,\mathbf{x}_0}^\Delta \phi_i^K = \frac{\mu(\mathbf{x}_0)}{\rho(\mathbf{x}_0)} \nabla \cdot \nabla \phi_i^K(\mathbf{x}_0),$$

pour tout  $K$  de  $\mathcal{T}_h$  et tout  $i = 1..(m+1)(m+2)/2$ . Nous pouvons alors remplacer respectivement

$$\nabla f_t^{n+\frac{1}{2}}, \quad \mu \nabla \cdot \frac{1}{\rho} \nabla f^{n+1}, \quad f_t^{n+1} \quad \text{et} \quad f^{n+1}$$

dans (3.1.15) par

$$\frac{1}{\rho} \nabla \delta_{h, \mathbf{x}_0}^{\nabla} g^{n+\frac{1}{2}}, \quad \mu \nabla \cdot \frac{1}{\rho} \nabla \delta_{h, \mathbf{x}_0}^{\Delta} g^{n+1}, \quad \delta_{h, \mathbf{x}_0} g^{n+1} \quad \text{et} \quad \delta_{h, \mathbf{x}_0} g^{n+1}.$$

Les vecteurs source du schéma  $\nabla$  deviennent donc

$$\mathbf{F}_i^{n+1K} = \phi_i^K(\mathbf{x}_0) g^{n+1},$$

$$\mathbf{F}_{\mathbf{v}_x}^{n+\frac{1}{2}K} = \frac{1}{\rho(\mathbf{x}_0)} \nabla \phi_i^K(\mathbf{x}_0) g^{n+\frac{1}{2}} \cdot \mathbf{e}_x, \quad \mathbf{F}_{\mathbf{v}_y}^{n+\frac{1}{2}K} = \frac{1}{\rho(\mathbf{x}_0)} \nabla \phi_i^K(\mathbf{x}_0) g^{n+\frac{1}{2}} \cdot \mathbf{e}_y$$

et

$$\mathbf{F}_p^{n+1K} = \frac{\mu(\mathbf{x}_0)}{\rho(\mathbf{x}_0)} \nabla \cdot \nabla \phi_i^K(\mathbf{x}_0) g^{n+1} + \phi_i^K(\mathbf{x}_0) g^{n+1}.$$

Remarquons que comme nous supposons que le point source appartient uniquement à  $K_{src}$ , on a :

$$\mathbf{F}_i^{n+1K} = \mathbf{F}_{\mathbf{v}_x}^{n+\frac{1}{2}K} = \mathbf{F}_{\mathbf{v}_y}^{n+\frac{1}{2}K} = \mathbf{F}_p^{n+1K} = 0,$$

si  $K \neq K_{src}$ .

Nous avons ici uniquement présenté la construction des vecteurs source pour le schéma  $\nabla$ , mais nous utilisons la même technique pour les autres schémas présentés dans ce chapitre.



## Chapitre 4

# Analyse de dispersion des schémas temporels

Dans ce chapitre, nous cherchons à expliquer les modes parasites observés au chapitre précédent. Pour cela, nous menons une analyse de dispersion en dimension 1. Le cadre général de l'analyse est donné à la section 4.1. L'analyse que nous menons nous permet d'établir une condition nécessaire de stabilité pour chacun des schémas (section 4.2) et étudier les courbes de dispersion (section 4.3). Les résultats obtenus mettent en évidence l'apparition d'ondes parasites pour certaines fréquences. Ils montrent également une détérioration de l'ordre de convergence des schémas en temps. Pour supprimer ces ondes parasites et restaurer l'ordre de convergence, nous proposons de pénaliser la formulation en espace (section 4.4), ce qui revient à introduire un terme de pénalisation dans les équations variationnelles. Afin de vérifier que le terme de pénalisation rétablit les propriétés de convergence des schémas, nous réalisons une analyse de dispersion des schémas pénalisés qui montrent que la précision des schémas est améliorée.

### 4.1 Méthodologie générale

Pour étudier les modes parasites observés au chapitre précédent, nous allons effectuer une analyse de dispersion en dimension un. Le principe de cette analyse consiste à chercher des solutions particulières du schéma numérique qui sont des ondes planes discrètes et à comparer leurs propriétés (en particulier leur vitesse de propagation) à celles des ondes planes continues solutions du problème et qui sont de la forme  $p(x, t) = e^{i(kx - \omega t)}$  avec  $(k, \omega) \in \mathbb{R}^2$  et  $\frac{\omega}{k} = \pm c$ .

Nous considérons un maillage régulier de  $\mathbb{R}$  de pas  $h$  représenté par des segments  $S_I = [x_I, x_{I+1}]$ ,  $I \in \mathbb{Z}$  avec  $x_I = (I - 1)h$  (voir Fig. 4.1).

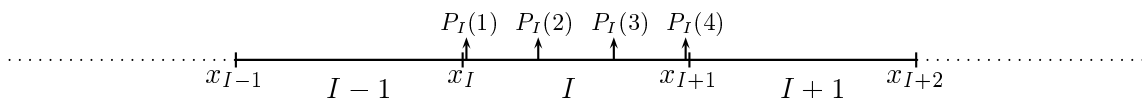


FIGURE 4.1 – Le maillage régulier.

Pour les schémas Leap-Frog et  $\nabla$ , le schéma totalement discrétisé peut se réécrire élément par

élément sous la forme

$$(4.1.1) \quad \begin{cases} hM \frac{V_I^{n+\frac{1}{2}} - V_I^{n-\frac{1}{2}}}{\Delta t} = KP_I^n + K^W P_{I-1}^n + K^E P_{I+1}^n, \\ hM \frac{P_I^{n+1} - P_I^n}{\Delta t} = KV_I^{n+\frac{1}{2}} + K^W V_{I-1}^{n+\frac{1}{2}} + K^E V_{I+1}^{n+\frac{1}{2}}, \end{cases}$$

où  $P_I$  (resp.  $V_I$ ) est un vecteur de taille  $N_\Phi = m + 1$  contenant les valeurs de  $p_h$  (resp.  $v_h$ ) aux degrés de liberté du segment  $S_I$ , la matrice  $M$  est une matrice symétrique de taille  $N_\Phi \times N_\Phi$  définie par

$$M_{ij} = \int_0^1 \hat{\phi}_j(x) \hat{\phi}_i(x) dx,$$

et  $K$ ,  $K^E$  et  $K^W$  sont des matrices de taille  $N_\Phi \times N_\Phi$  définies par

$$K_{ij} = - \int_0^1 \hat{\phi}_j(x) \frac{d\hat{\phi}_i(x)}{dx} dx + \frac{1}{2} \left( \hat{\phi}_j(1) \hat{\phi}_i(1) - \hat{\phi}_j(0) \hat{\phi}_i(0) \right),$$

$$K_{ij}^W = \frac{1}{2} \hat{\phi}_j(0) \hat{\phi}_i(1) \text{ et } K_{ij}^E = -\frac{1}{2} \hat{\phi}_j(1) \hat{\phi}_i(0),$$

dans le cas du schéma Leap-Frog et

$$K_{ij} = - \int_0^1 \hat{\phi}_j(x) \frac{d\hat{\phi}_i(x)}{dx} dx + \frac{1}{2} \left( \hat{\phi}_j(1) \hat{\phi}_i(1) - \hat{\phi}_j(0) \hat{\phi}_i(0) \right),$$

$$- \frac{\Delta t^2}{24h^2} \int_0^1 \hat{\phi}_j(x) \frac{d^3 \hat{\phi}_i(x)}{dx^3} dx + \frac{\Delta t^2}{48h^2} \left( \frac{d^2 \hat{\phi}_j}{dx^2}(1) \hat{\phi}_i(1) - \frac{d^2 \hat{\phi}_j}{dx^2}(0) \hat{\phi}_i(0) \right)$$

$$+ \frac{\Delta t^2}{48h^2} \left( \hat{\phi}_j(1) \frac{d^2 \hat{\phi}_i}{dx^2}(1) - \hat{\phi}_j(0) \frac{d^2 \hat{\phi}_i}{dx^2}(0) - \frac{d\hat{\phi}_j}{dx}(1) \frac{d\hat{\phi}_i}{dx}(1) + \frac{d\hat{\phi}_j}{dx}(0) \frac{d\hat{\phi}_i}{dx}(0) \right)$$

$$K_{ij}^W = \frac{1}{2} \hat{\phi}_j(0) \hat{\phi}_i(1) + \frac{\Delta t^2}{48h^2} \left( \frac{d^2 \hat{\phi}_j}{dx^2}(0) \hat{\phi}_i(1) + \hat{\phi}_j(0) \frac{d^2 \hat{\phi}_i}{dx^2}(1) - \frac{d\hat{\phi}_j}{dx}(0) \frac{d\hat{\phi}_i}{dx}(1) \right).$$

et

$$K_{ij}^E = -\frac{1}{2} \hat{\phi}_j(1) \hat{\phi}_i(0) - \frac{\Delta t^2}{48h^2} \left( \frac{d^2 \hat{\phi}_j}{dx^2}(1) \hat{\phi}_i(0) + \hat{\phi}_j(1) \frac{d^2 \hat{\phi}_i}{dx^2}(0) - \frac{d\hat{\phi}_j}{dx}(1) \frac{d\hat{\phi}_i}{dx}(0) \right)$$

pour le schéma  $\nabla$ . En toute rigueur, nous devrions noter différemment les matrices  $K$  en fonction des schémas utilisés. Nous avons choisi de ne pas le faire afin de limiter les notations qui nous semblent déjà très nombreuses.

Le schéma ADER d'ordre 4 s'écrit sous la forme

$$(4.1.2) \quad \begin{cases} hM \frac{V_I^{n+\frac{1}{2}} - V_I^{n-\frac{1}{2}}}{\Delta t} = KP_I^n + K^W P_{I-1}^n + K^E P_{I+1}^n + \frac{\Delta t^2}{24} \left( K \tilde{P}_I^n + K^W \tilde{P}_{I-1}^n + K^E \tilde{P}_{I+1}^n \right), \\ hM \frac{P_I^{n+1} - P_I^n}{\Delta t} = KV_I^{n+\frac{1}{2}} + K^W V_{I-1}^{n+\frac{1}{2}} + K^E V_{I+1}^{n+\frac{1}{2}} + \frac{\Delta t^2}{24} \left( K \tilde{V}_I^{n+\frac{1}{2}} + K^W \tilde{V}_{I-1}^{n+\frac{1}{2}} + K^E \tilde{V}_{I+1}^{n+\frac{1}{2}} \right), \end{cases}$$

où  $M$  et  $K$  sont les matrices définies pour le schéma Leap-Frog et le vecteur  $\tilde{P}_I^n$  est défini pour tout  $I$  de  $\mathbb{Z}$  par

$$(4.1.3) \quad \begin{cases} hM\tilde{V}_I^{n+\frac{1}{2}} = KP_I^n + K^W P_{I-1}^n + K^E P_{I+1}^n \\ hM\tilde{P}_I^n = K\tilde{V}_I^{n+\frac{1}{2}} + K^W \tilde{V}_{I-1}^{n+\frac{1}{2}} + K^E \tilde{V}_{I+1}^{n+\frac{1}{2}}. \end{cases}$$

Le vecteur  $\tilde{V}_I^{n+\frac{1}{2}}$  est défini pour tout  $I \in \mathbb{Z}$  par

$$(4.1.4) \quad \begin{cases} hM\tilde{P}_I^{n+1} = K\tilde{V}_I^{n+\frac{1}{2}} + K^W \tilde{V}_{I-1}^{n+\frac{1}{2}} + K^E \tilde{V}_{I+1}^{n+\frac{1}{2}}, \\ hM\tilde{V}_I^{n+\frac{1}{2}} = K\tilde{P}_I^n + K^W \tilde{P}_{I-1}^n + K^E \tilde{P}_{I+1}^n. \end{cases}$$

Comme nous l'avons dit en introduction, nous cherchons des ondes planes numériques solutions des différents schémas. Cela se traduit par le fait que, pour tout  $I$  de  $\mathbb{Z}$  et tout  $n$  de  $\mathbb{N}$ , les vecteurs  $P_I$  et  $V_I$  vérifient

$$P_{I+1}^{n+1} = e^{i(kh+\omega_h\Delta t)} P_I^n \text{ et } V_{I+1}^{n+\frac{1}{2}} = e^{i(kh+\omega_h\Delta t)} V_I^{n-\frac{1}{2}},$$

où  $k \in \mathbb{R}$  représente le nombre d'onde de la solution et  $\omega_h$  sa pulsation. Dans le cas des schémas Leap-Frog et  $\nabla$ , l'équation (4.1.1) devient

$$(4.1.5) \quad \begin{cases} \frac{2i}{h\Delta t} \sin\left(\frac{\omega_h\Delta t}{2}\right) e^{i\frac{\omega_h\Delta t}{2}} M V_I^{n+\frac{1}{2}} = \left(K + K^W e^{-ikh} + K^E e^{ikh}\right) P_I^n, \\ \frac{2i}{h\Delta t} \sin\left(\frac{\omega_h\Delta t}{2}\right) e^{-i\frac{\omega_h\Delta t}{2}} M P_I^n = \left(K + K^W e^{-ikh} + K^E e^{ikh}\right) V_I^{n+\frac{1}{2}}. \end{cases}$$

Ce système peut se réécrire uniquement en fonction de  $P_I^n$  sous la forme

$$(4.1.6) \quad \sin^2\left(\frac{\omega_h\Delta t}{2}\right) P_I^n = \frac{\Delta t^2}{4h^2} K^*(kh) P_I^n,$$

avec

$$K^*(kh) = -M^{-1} \left(K + K^W e^{-ikh} + K^E e^{ikh}\right) M^{-1} \left(K + K^W e^{-ikh} + K^E e^{ikh}\right).$$

Par la suite, nous noterons  $K_{LF}^*(kh)$  et  $K_{\nabla}^*(kh)$  les matrices associées respectivement au schéma Leap-Frog et au schéma  $\nabla$ . En procédant de manière similaire dans le cas du schémas ADER d'ordre quatre, on obtient le système

$$(4.1.7) \quad \sin^2\left(\frac{\omega_h\Delta t}{2}\right) P_I^n = \frac{\Delta t^2}{4h^2} K_{ADER}^* P_I^n,$$

avec

$$(4.1.8) \quad K_{ADER}^* = - \left( M^{-1} \left( K + K^W e^{-ikh} + K^E e^{ikh} \right) + \frac{\Delta t^2}{24h^2} \left[ M^{-1} \left( K + K^W e^{-ikh} + K^E e^{ikh} \right) \right]^3 \right)^2.$$

La suite de l'analyse de dispersion dépend du calcul des valeurs propres des matrices  $K^*$ , sous réserve de leur existence, et on va voir que la relation de dispersion met en jeu la racine carrée des valeurs propres des matrices  $K^*$ . Pour que cette relation de dispersion ait un sens, il nous faut donc tout d'abord vérifier que ces valeurs propres existent et l'équation de dispersion n'aura alors un sens que si elles sont réelles et positives. On établit alors le théorème suivant :

**Théorème 4.1.1.** *Les matrices  $M^{-1} (K + K^W e^{-ikh} + K^E e^{ikh})$  sont diagonalisables et leurs valeurs propres sont imaginaires pures.*

**Démonstration :** Pour démontrer ce théorème, nous allons tout d'abord vérifier que la matrice  $K + K^W e^{-ikh} + K^E e^{ikh}$  est antihermitienne, c'est-à-dire que

$$\left( K + K^W e^{-ikh} + K^E e^{ikh} \right) = -\overline{\left( K + K^W e^{-ikh} + K^E e^{ikh} \right)^T}.$$

On vérifie facilement que  $K_{ij}^E = -K_{ji}^W$  et donc que

$$K_{ij}^W e^{-ikh} + K_{ij}^E e^{ikh} = -K_{ji}^E e^{-ikh} - K_{ji}^W e^{ikh} = -\overline{K_{ji}^W e^{-ikh} + K_{ji}^E e^{ikh}}.$$

Il nous reste à montrer que  $K_{ij} = -\overline{K_{ji}}$ . Nous ne le faisons ici que pour le schéma Leap-Frog, la démonstration étant similaire pour le schéma  $\nabla$ . En effectuant une intégration par parties, on remarque que

$$K_{ij} = \int_0^1 \frac{d\widehat{\phi}_j(x)}{dx} \widehat{\phi}_i(x) dx - \left( \widehat{\phi}_j(1)\widehat{\phi}_i(1) - \widehat{\phi}_j(0)\widehat{\phi}_i(0) \right) + \frac{1}{2} \left( \widehat{\phi}_j(1)\widehat{\phi}_i(1) - \widehat{\phi}_j(0)\widehat{\phi}_i(0) \right),$$

soit

$$K_{ij} = \int_0^1 \frac{d\widehat{\phi}_j(x)}{dx} \widehat{\phi}_i(x) dx - \frac{1}{2} \left( \widehat{\phi}_j(1)\widehat{\phi}_i(1) - \widehat{\phi}_j(0)\widehat{\phi}_i(0) \right) = -K_{ji} = -\overline{K_{ij}}.$$

On a donc bien

$$\left( K + K^W e^{-ikh} + K^E e^{ikh} \right)_{ij} = -\overline{\left( K + K^W e^{-ikh} + K^E e^{ikh} \right)_{ji}}.$$

Pour achever la preuve du Théorème 4.1.1, nous utilisons le lemme suivant.

**Lemme 4.1.1.** *Si  $M$  est une matrice symétrique définie positive et  $N$  une matrice antihermitienne,  $M^{-1}N$  est diagonalisable et toutes ses valeurs propres sont imaginaires pures.*

**Démonstration :** Comme  $M$  est symétrique définie positive,  $M^{-\frac{1}{2}}$  existe et  $M^{-\frac{1}{2}}NM^{-\frac{1}{2}}$  est antihermitienne. Soit  $(X_i)_{i=1..N_\Phi}$ ,  $N_\Phi$  vecteurs propre de  $M^{-\frac{1}{2}}NM^{-\frac{1}{2}}$  formant une base de  $\mathbb{C}^{N_\Phi}$ . L'existence de cette base est assurée par le caractère antihermitien de  $M^{-\frac{1}{2}}NM^{-\frac{1}{2}}$ . Soit  $\lambda_i$  la valeur propre (imaginaire pure) associée au vecteur propre  $X_i$ . On a alors

$$M^{-\frac{1}{2}}NM^{-\frac{1}{2}}X_i = \lambda_i X_i.$$

En multipliant cette équation par  $M^{-\frac{1}{2}}$ , on obtient

$$M^{-1}NM^{-\frac{1}{2}}X_i = \lambda_i M^{-\frac{1}{2}}X_i.$$

En posant  $Y_i = M^{-\frac{1}{2}}X_i$ , on vérifie que  $Y_i$  est un vecteur propre de  $M^{-1}N$  associé à  $\lambda_i$ . De plus,  $(Y_i)_{i=1..N_\Phi}$  forme une base de  $\mathbb{C}^{N_\Phi}$  car  $M^{-\frac{1}{2}}$  est symétrique définie positive.

■

■

**Corollaire 4.1.1.** *Les matrices  $K_{LF}^*$ ,  $K_{\nabla}^*$  et  $K_{ADER}^*$  sont diagonalisables et leurs valeurs propres sont réelles et positives.*

**Démonstration :** On pose  $\mathcal{B} = M^{-1} (K + K^W e^{-ikh} + K^E e^{ikh})$ . D'après le théorème précédent,  $-\mathcal{B}$  est diagonalisable et ses valeurs propres sont imaginaires pures. On en déduit que  $-\mathcal{B}^2$  est diagonalisable et ses valeurs propres sont réelles et positives. Donc  $K_{LF}^*$  et  $K_{\nabla}^*$  sont diagonalisables et leurs valeurs propres sont réelles et positives.

Comme

$$K_{ADER}^* = - \left( \mathcal{B} + \frac{\Delta t^2}{24h^2} \mathcal{B}^3 \right)^2,$$

on vérifie que les valeurs propres  $\mathcal{B} + \frac{\Delta t^2}{24h^2} \mathcal{B}^3$  sont imaginaires pures ce qui entraîne que  $K_{ADER}^*$  est diagonalisable et ses valeurs propres sont réelles et positives. Ceci achève la preuve du corollaire 4.1.1

■

Nous pouvons maintenant exprimer la relation (4.1.6) en fonction des valeurs propres  $(\lambda^*(kh))_{i=1..N \leq N_{\Phi}}$  des matrices  $K^*(kh)$  :

$$(4.1.9) \quad \sin\left(\frac{\omega_{hi}(kh)\Delta t}{2}\right) = \frac{\Delta t}{2h} \sqrt{\lambda_i^*(kh)}, \quad i = 1..N_{\Phi}.$$

Ainsi, pour chaque valeur propre  $\lambda_i^*(kh)$  le schéma admet une onde plane numérique de pulsation  $\omega_{hi}(kh)$ . Avant d'effectuer l'analyse de dispersion proprement dite, nous pouvons remarquer que cette relation fait apparaître une condition nécessaire de stabilité des schémas.

## 4.2 Condition nécessaire de stabilité des schémas

Si l'équation (4.1.9) admet des solutions  $\omega_{hi}(kh)$  complexes, ces solutions sont nécessairement complexes conjuguées et l'onde plane associée à la pulsation de partie imaginaire négative est exponentiellement croissante en temps. Le schéma admet alors des solutions non physiques et est instable. Il nous faut donc garantir que toutes les pulsations sont réelles, ce qui se fait en imposant la condition :

$$(4.2.1) \quad \frac{\Delta t}{2h} \sqrt{\lambda_i^*(kh)} \leq 1,$$

pour tout  $k$  de  $[0, \frac{2\pi}{h}]$  et pour tout  $i = 1..N_{\Phi}$ . Pour le schéma Leap-Frog, les valeurs propres ne dépendent pas de  $\Delta t$  et cette condition devient

$$\frac{\Delta t}{h} \leq \frac{2}{\sqrt{\lambda_{LF,\max}^*}} \text{ avec } \lambda_{LF,\max}^* = \max_{k \in [0, \frac{2\pi}{h}]} \max_{i=1..N_{\Phi}} \lambda_i^*(kh).$$

Pour calculer une approximation de  $\lambda_i^*(kh)$ , nous avons calculé le déterminant de  $K_{LF}^*(kh)$  à l'aide de Maple, puis nous avons calculé ses racines pour 1000 valeurs de  $k$  réparties uniformément sur l'intervalle  $[0; \frac{2\pi}{h}]$ . La condition CFL est ensuite obtenue à partir de la valeur propre de module maximal. Nous poserons par la suite  $\gamma = \frac{\Delta t}{h}$  et nous noterons  $\gamma_{\max}$  la valeur maximale de  $\gamma$  garantissant la stabilité des schémas, i.e., le coefficient CFL des schémas.



On rappelle qu'on a la relation

$$K_{ADER}^* = - \left( \mathcal{B} + \frac{\Delta t^2}{24h^2} \mathcal{B}^3 \right)^2 = -\mathcal{B}^2 - \frac{\Delta t^2}{12h^2} \mathcal{B}^4 - \frac{\Delta t^4}{576h^2} \mathcal{B}^6.$$

Comme  $K_{LF}^* = -\mathcal{B}^2$ , on remarque que la condition (4.2.1) peut se réécrire

$$\frac{\Delta t}{2h} \sqrt{\lambda_{LF,i}^*(kh) - \frac{\Delta t^2}{12h^2} \lambda_{LF,i}^{*2}(kh) + \frac{\Delta t^4}{576h^2} \lambda_{LF,i}^{*3}(kh)} \leq 1,$$

pour tout  $k$  de  $[0; \frac{2\pi}{h}]$  et pour tout  $i = 1..N_\Phi$ . En posant  $X = \frac{\Delta t^2}{4h^2} \lambda_{LF,i}^*(kh)$ , on obtient

$$\frac{X^3}{36} - \frac{X^2}{3} + X - 1 \leq 0.$$

Ce polynôme admet deux racines complexes et une racine réelle de valeur approximative 8.1072. Pour que le schéma ADER d'ordre quatre soit stable, il faut donc que

$$\frac{\Delta t^2}{h^2} \lambda_{LF,i}^*(kh) \leq 8.1072, \quad \forall k \in \left[0; \frac{2\pi}{h}\right] \text{ et } \forall i = 1..N_\Phi.$$

Autrement dit,

$$\frac{\Delta t}{h} \leq 2.8473 \frac{2}{\sqrt{\lambda_{LF,\max}^*}}.$$

On retrouve ici le facteur 2.8473 déjà calculé par Lanteri et Fahs dans [32].

Le calcul de la condition de stabilité du schéma  $\nabla$  est plus compliqué car nous ne pouvons pas exprimer les valeurs propres associées à l'opérateur d'ordre quatre en fonction de celles de l'opérateur d'ordre deux. Nous avons donc procédé par dichotomie : on initialise l'algorithme en utilisant la CFL du Leap-Frog ; les bornes inférieures et supérieures de l'algorithme sont respectivement la moitié de la CFL du Leap-Frog et 1,5 fois la CFL du Leap-Frog. Ensuite, pour un pas de temps donné, on calcule les valeurs propres  $\lambda_i^*(kh)$  pour 1000 valeurs de  $k$  équiréparties dans  $[0; \frac{2\pi}{h}]$ . Si toutes ces valeurs propres respectent la condition (4.2.1), alors on augmente le pas de temps, sinon on le diminue.

Nous poserons par la suite  $\gamma = \frac{\Delta t}{h}$  et nous noterons  $\gamma_{\max}$  la valeur maximale de  $\gamma$  garantissant la stabilité des schémas, i.e., le coefficient CFL des schémas. Les trois conditions CFL sont présentées dans le tableau 4.1 pour différentes approximations de  $\mathcal{P}^0$  à  $\mathcal{P}^6$ . On remarque que la CFL du schéma  $\nabla$  est légèrement inférieure à celle du schéma Leap-Frog.

	$\mathcal{P}^0$	$\mathcal{P}^1$	$\mathcal{P}^2$	$\mathcal{P}^3$	$\mathcal{P}^4$	$\mathcal{P}^5$	$\mathcal{P}^6$
$\gamma_{LF,\max}$	2	0.5	0.247	0.15	0.01	0.0732	0.0551
$\gamma_{ADER,\max}$	5.6	1.4	0.693	0.421	0.284	0.204	0.154
$\gamma_{\nabla,\max}$	2	0.488	0.234	0.138	0.0921	0.0658	0.0492

TABLE 4.1 – Coefficients de stabilité du schéma Leap-Frog, ADER et  $\nabla$  pour des approximations en espace de  $\mathcal{P}^0$  à  $\mathcal{P}^6$ .

### 4.3 Analyse de dispersion des schémas

Maintenant que nous connaissons la condition CFL de chacun de nos schémas, nous pouvons passer à l'analyse de dispersion. Cette analyse consiste à étudier les variations des pulsations discrètes  $\omega_i(kh)$  et à les comparer à la pulsation  $\omega(k)$  de l'onde plane continue. La première difficulté réside dans l'existence de  $N_\Phi$  pulsations discrètes pour une seule pulsation continue. En fait, seule une pulsation représente la pulsation physique, les autres sont associées à des ondes planes discrètes parasites. La qualité du schéma dépendra donc **a)** de sa capacité à approcher au mieux la pulsation continue et **b)** de sa capacité à filtrer les ondes planes parasites dans un sens que nous préciserons par la suite. Pour pouvoir étudier les différents schémas, nous devons tout d'abord être capables de différencier la pulsation physique des pulsations parasites. Un premier critère de discrimination repose sur le fait que la vitesse de phase  $\omega(k)/k = c = 1$  est indépendante de  $k$ . Nous pouvons donc en déduire que la pulsation discrète doit vérifier cette propriété pour des maillages très raffinés, c'est-à-dire que  $\omega_i$  est tel que

$$(4.3.1) \quad \lim_{kh \rightarrow 0} \frac{\omega_i(kh)}{k} = 1.$$

D'après (4.1.9), on a

$$(4.3.2) \quad \frac{\omega_{hi}(kh)}{k} = \frac{2}{k\Delta t} \arcsin \left( \frac{\Delta t}{2h} \sqrt{\lambda_i^*(kh)} \right), \quad i = 1..N_\Phi.$$

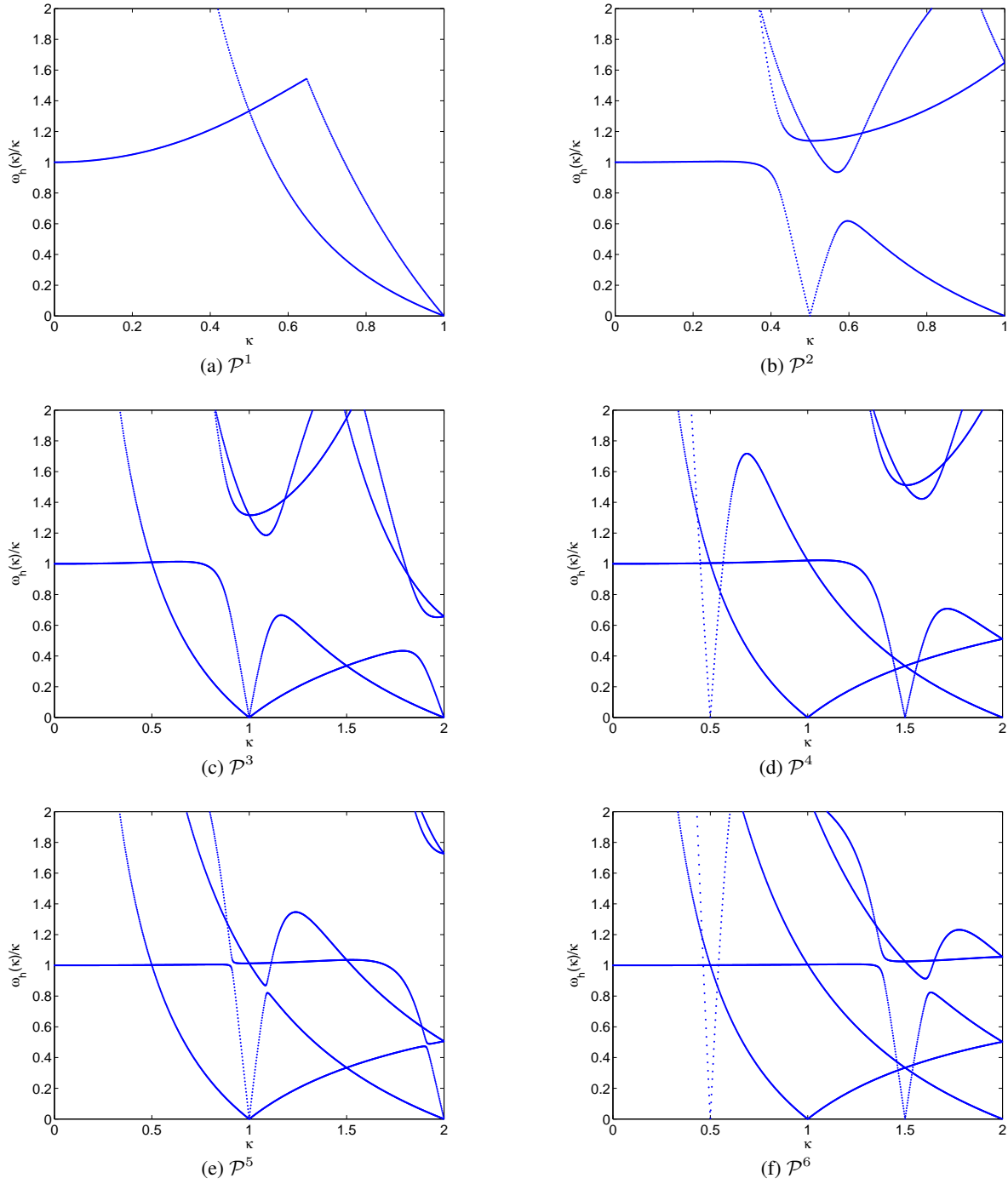
Cette quantité est appelée vitesse de phase de l'onde. En utilisant le paramètre  $\gamma = \Delta t/h$ , on peut la réécrire sous la forme

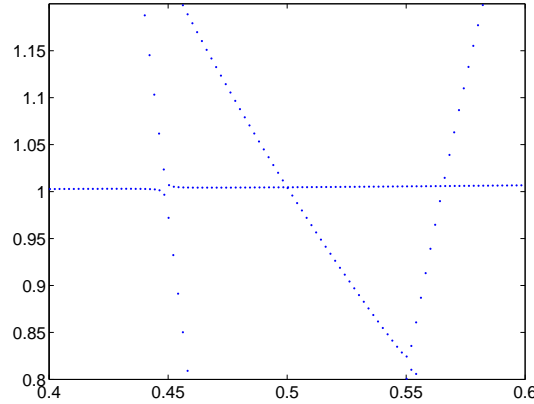
$$(4.3.3) \quad \frac{\omega_{hi}(kh)}{k} = \frac{2}{kh\gamma} \arcsin \left( \frac{\gamma}{2} \sqrt{\lambda_i^*(kh)} \right), \quad i = 1..N_\Phi.$$

Introduisons la quantité  $\kappa = \frac{kh}{2\pi}$  qui représente l'inverse du nombre d'éléments par longueur d'onde. Pour analyser la précision du schéma, nous allons exprimer  $\omega_{hi}$  en fonction de  $\kappa$  :

$$(4.3.4) \quad \frac{\omega_{hi}(2\pi\kappa)}{k} = \frac{1}{\pi\kappa\gamma} \arcsin \left( \frac{\gamma}{2} \sqrt{\lambda_i^*(2\pi\kappa)} \right), \quad i = 1..N_\Phi.$$

La définition (4.3.1) est pertinente à basse fréquence (ou pour des maillages très raffinés), mais elle ne suffit pas, à haute fréquence (ou pour des maillages grossiers), pour déterminer la pulsation physique. Pour illustrer ce point, nous représentons sur les figures 4.2.(a) à 4.2.(f) les valeurs de  $\omega_i(2\pi\kappa)/k$  en fonction de  $\kappa$  obtenues pour le schéma Leap-Frog de  $\mathcal{P}^1$  à  $\mathcal{P}^6$ . Pour les degrés élevés, la valeur propre qui approche le mieux la valeur propre physique à basse fréquence n'est pas la même que celle qui l'approche à haute fréquence. En effet, la valeur propre qui vérifie (4.3.1) décroît vers 0 lorsque  $\kappa$  augmente, et d'autres valeurs propres sont alors beaucoup plus proches de la condition  $\frac{\omega_i(\kappa)}{k} = 1$ . Ce phénomène est particulièrement visible pour des éléments  $\mathcal{P}^5$  au voisinage de  $\kappa = 1$  mais il est également présent pour des éléments  $\mathcal{P}^4$  et  $\mathcal{P}^6$ , à cause de croisements de valeurs propres indistinguables à l'oeil nu. Pour vérifier cela, nous avons fait un zoom autour du point  $\kappa = 0.5$  pour les éléments  $\mathcal{P}^4$  sur la figure 4.3. On remarque que la vitesse de phase associée à la pulsation définie par (4.3.1) décroît brutalement au voisinage de  $\kappa = 0.45$ . Cependant, une autre pulsation permet de bien approcher la vitesse de phase pour  $\kappa > 0.45$  (on observe des phénomènes similaires pour les éléments  $\mathcal{P}^6$ ). On voit donc que le choix d'utiliser (4.3.1) pour déterminer la pulsation physique n'est pas judicieux pour des maillages grossiers.

FIGURE 4.2 – Courbes de dispersion pour le schéma Leap-Frog pour des éléments  $\mathcal{P}^1$  à  $\mathcal{P}^6$

FIGURE 4.3 – Courbes de dispersion pour le schéma Leap-Frog pour des éléments  $\mathcal{P}^4$  (zoom)

Une solution alternative pourrait être de choisir, pour chaque valeur de  $\kappa$ , la pulsation associée à la vitesse de phase la plus proche de 1. L'inconvénient de cette solution est qu'elle ne permet pas réellement de discriminer les pulsations au niveau des croisements.

Nous avons donc privilégié une troisième approche basée sur l'étude des vecteurs propres associés à chaque pulsation. Nous choisissons la valeur propre dont le vecteur propre associé approche au mieux  $e^{2i\pi\kappa x}$ . Pour calculer ce vecteur propre, nous devons tenir compte de l'éventuelle multiplicité des valeurs propres pour certaines valeurs de  $kh$ . Nous introduisons donc  $\nu(i)$  l'ordre de multiplicité de la valeur propre  $\lambda_i$  et  $(V_i^j)_{j=1..\nu(i)}$  une famille linéairement indépendante de  $\nu(i)$  vecteurs propres associés à  $\lambda_i$ . L'idée est ici de trouver le sous-espace propre contenant la meilleure approximation de l'onde physique au sens des moindres carrés. Pour pouvoir approcher la fonction  $e^{2i\pi\kappa x}$ , nous interpolons chaque vecteur propre  $V_i^j$  par le vecteur  $v_i^j$  défini par  $v_i^j(\kappa, x) = \sum_{l=1}^{p+1} \phi_l(x)[V_i^j(\kappa)]_l$  et nous calculons une approximation de  $e^{2i\pi\kappa x}$  par la méthode des moindres carrés sur mille points équitablement répartis sur l'intervalle  $[0; 1]$ . Ce calcul nous permet d'exprimer l'approximation  $u$  sous la forme

$$u(x) = \sum_{i=1}^{N_\Phi} \sum_{j=1}^{\nu(i)} \beta_i^j v_i^j(\kappa, x).$$

Nous notons alors

$$I = \arg \min_{i=1..N_\Phi} \left\| e^{2i\pi\kappa x} - \sum_{j=1}^{\nu(i)} \beta_i^j v_i^j(\kappa, x) \right\|_{L^2([0,1])}.$$

Si le nombre d'éléments par longueur d'onde est suffisamment grand (autrement dit, si  $\kappa$  est suffisamment petit), nous aurons  $u(x) \approx \sum_{j=1}^{\nu(i)} \beta_I^j v_I^j(\kappa, x)$ . et  $\frac{1}{\pi\kappa\gamma} \arcsin\left(\frac{\gamma}{2}\sqrt{\lambda_I^*(2\pi\kappa)}\right) \approx 1$ .

Dans le tableau 4.2.(a), nous donnons les valeurs propres et leur multiplicité pour  $\kappa = 0.25$  et une approximation  $\mathcal{P}^5$  et nous représentons sur les figures 4.4.(a) et (b) les vecteurs propres associés. Nous voyons que le cinquième vecteur propre approche parfaitement la solution  $e^{2i\pi x/4}$ .

Intéressons-nous maintenant au cas  $\kappa = 1$  où une valeur propre est nulle et double. Nous donnons dans le tableau 4.2.(b) les valeurs propres et leur multiplicité pour  $\kappa = 1$  et une approximation  $\mathcal{P}^5$  et nous représentons sur les figures 4.5.(a) et (b) les vecteurs propres associés. La troisième valeur permet d'approcher relativement correctement la solution  $e^{2i\pi x}$ , mais nous remarquons la présence d'un mode parasite de vitesse nulle, car associé à la valeur propre nulle (en pointillés cyan), d'une forme similaire aux ondes parasites observées numériquement.

i	$\lambda_i$	$\nu(i)$
1	562	1
2	540	1
3	106	1
4	62.6	1
5	2.47	1
6	22.2	1

(a)  $\kappa = 0.25$  et  $\mathcal{P}^5$

i	$\lambda_i$	$\nu(i)$
1	0	2
2	380	2
3	39.7	2

(b)  $\kappa = 1$  et  $\mathcal{P}^5$

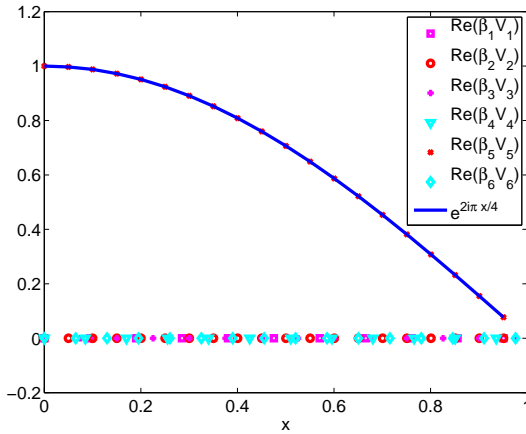
i	$\lambda_i$	$\nu(i)$
1	170	2
2	9.87	2
3	0	1

(c)  $\kappa = 0.5$  et  $\mathcal{P}^4$

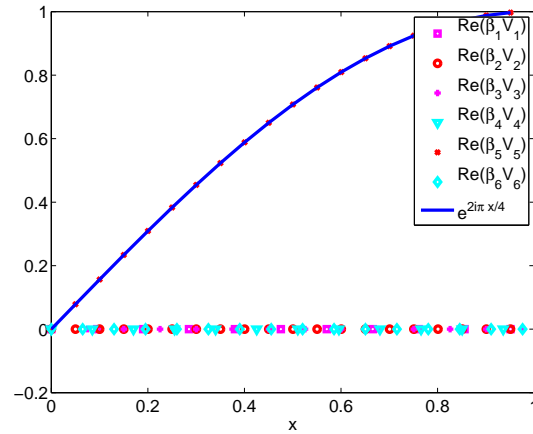
i	$\lambda_i$	$\nu(i)$
1	380	2
2	39,8	2
3	0	1

(d)  $\kappa = 1$  et  $\mathcal{P}^4$

TABLE 4.2 – Valeurs propres et leur multiplicité

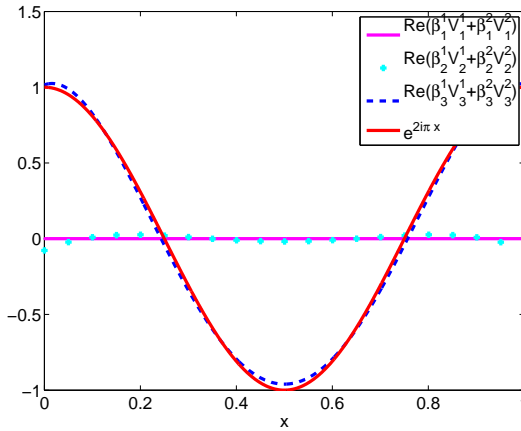


(a) Partie réelle

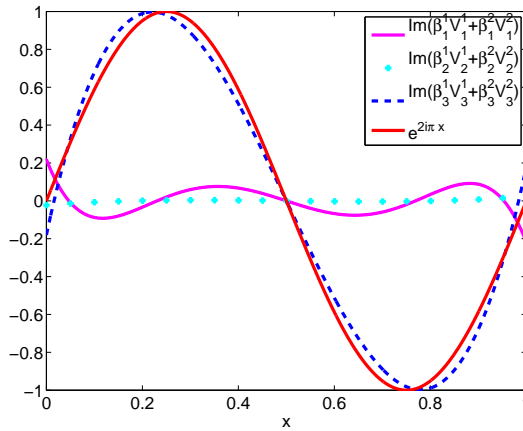


(b) Partie imaginaire

FIGURE 4.4 – Vecteurs propres pour  $\kappa = 0.25$  et  $\mathcal{P}^5$



(a) Partie réelle



(b) Partie imaginaire

FIGURE 4.5 – Vecteurs propres pour  $\kappa = 1$  et  $\mathcal{P}^5$

Pour les ordres pairs, nous considérons  $m = 4$  et les cas  $\kappa = 0.5$  et  $\kappa = 1$ , qui possèdent tous les deux des valeurs propres nulles (voir Tabs. 4.2.(c) et 4.2.(d) et Figs. 4.6 et 4.7). On observe un mode parasite de vitesse nulle (pointillés bleus) pour  $\kappa = 0.5$ , d'intensité plus faible que pour  $\kappa = 1$  et  $m = 5$ . Cependant,  $\kappa = 0.5$  correspond à deux éléments par longueur d'onde, ce qui veut dire que pour les ordres pairs, le mode parasite apparaît sur des maillages plus fins que pour les ordres impairs. Pour  $\kappa = 1$ , le mode parasite associé à la vitesse nulle est constant en espace (pointillés bleus).

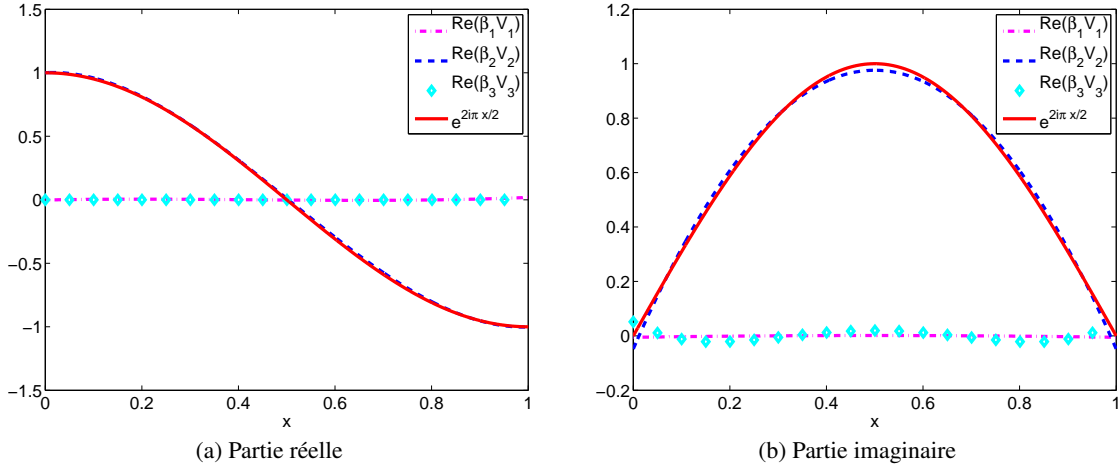


FIGURE 4.6 – Vecteurs propres pour  $\kappa = 0.5$  et  $\mathcal{P}^4$

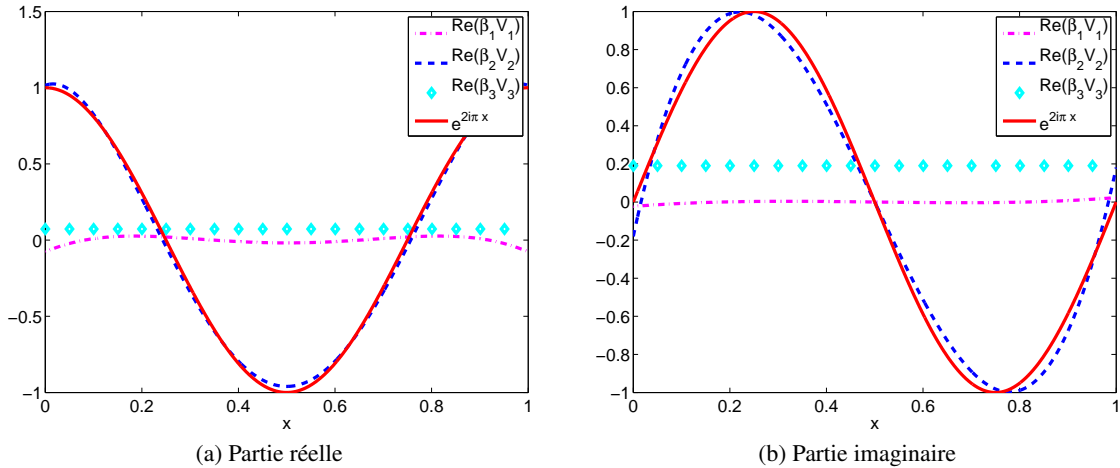
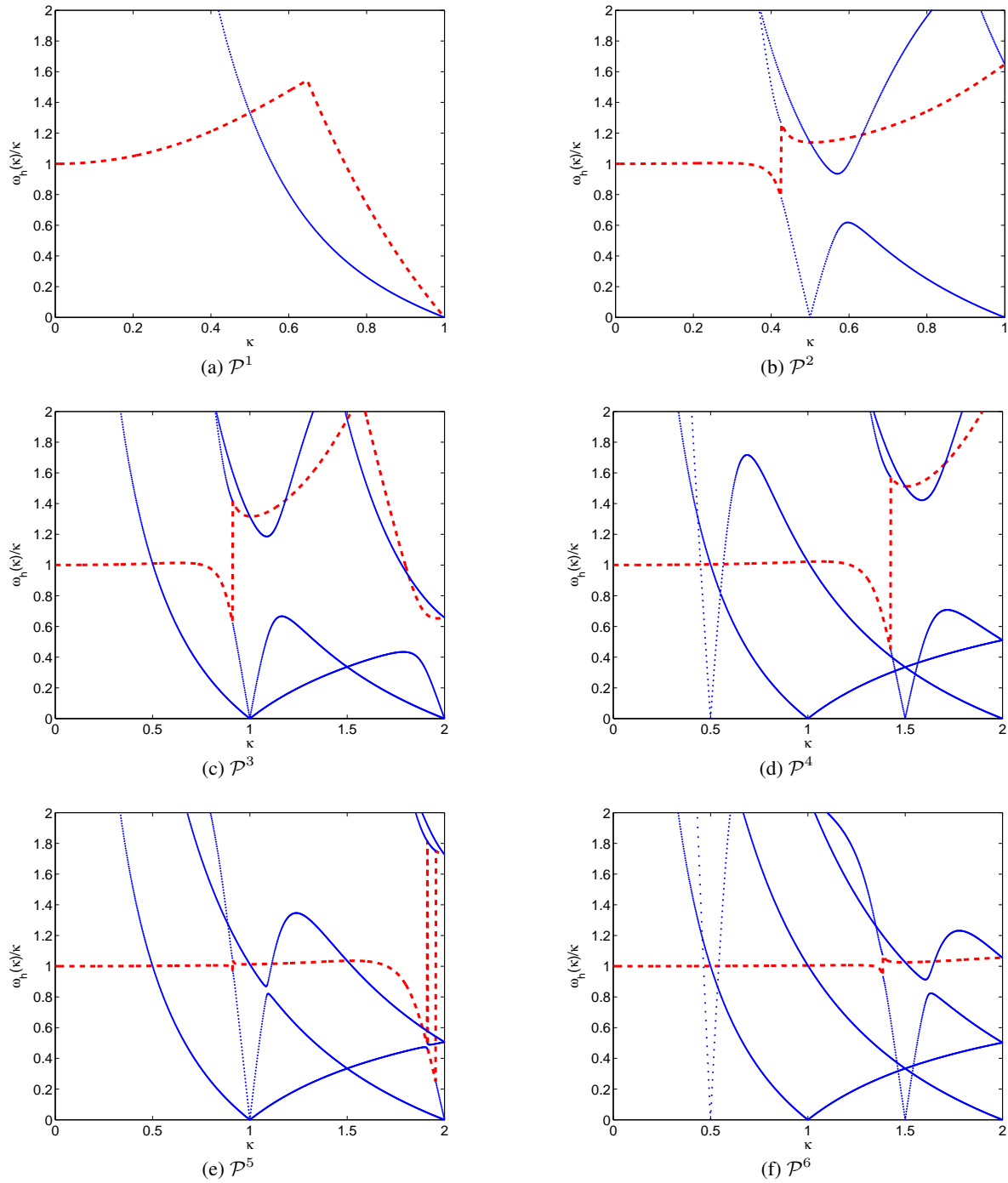


FIGURE 4.7 – Vecteurs propres pour  $\kappa = 1$  et  $\mathcal{P}^4$

Maintenant que nous avons un critère pour déterminer la pulsation physique, nous pouvons la représenter sur les courbes de dispersion. Sur les figures 4.8(a) à (f), nous représentons en pointillés rouges la valeur propre approchant le mieux la valeur propre physique pour des degrés d'approximation variant de  $\mathcal{P}^1$  à  $\mathcal{P}^6$ . Nous pouvons remarquer que cette valeur propre est discontinue au voisinage de deux points par longueur d'onde ( $\kappa \approx 0.5$ ) pour les éléments  $\mathcal{P}^3$  et  $\mathcal{P}^5$  et au voisinage de 0.66 point par longueur d'onde ( $\kappa \approx 1.5$ ) pour les éléments  $\mathcal{P}^4$  et  $\mathcal{P}^6$ . Pour localiser plus précisément ces discontinuités, nous représentons l'erreur  $\omega_{hi}(2\pi\kappa)/k - 1$  pour  $\mathcal{P}^1$  et  $\mathcal{P}^2$  sur la figure 4.9.(a) et de  $\mathcal{P}^3$

FIGURE 4.8 – Courbes de dispersion pour le schéma Leap-Frog pour des EF de  $\mathcal{P}^1$  à  $\mathcal{P}^6$

à  $\mathcal{P}^6$  sur la figure 4.9.(b). Cela nous permet de remarquer une (faible) discontinuité supplémentaire aux alentours de  $\kappa = 0.5$  (deux points par longueur d'onde) pour les éléments  $\mathcal{P}^4$ . Finalement, nous représentons ces erreurs en échelle  $\log - \log$  sur la figure 4.11.(a). Signalons qu'en plus d'une perte de précision sur la vitesse de phase, les discontinuités observées ont bien évidemment un effet dramatique sur la vitesse de groupe, qui est la dérivée de la pulsation par rapport à  $\kappa$ . Nous n'avons pas représenté ces erreurs car elles n'apportent rien de plus.

Nous pouvons également calculer l'erreur de projection de l'onde plane sur le sous-espace propre associé à la valeur propre physique. Sur la figure 4.10, nous représentons les variations de l'erreur relative  $L^2$  en fonction de  $\kappa$ . La figure 4.11.(b) représente cette même erreur relative en échelle  $\log - \log$ . Nous remarquons que les variations sont beaucoup moins fortes que pour l'erreur sur la vitesse de phase.

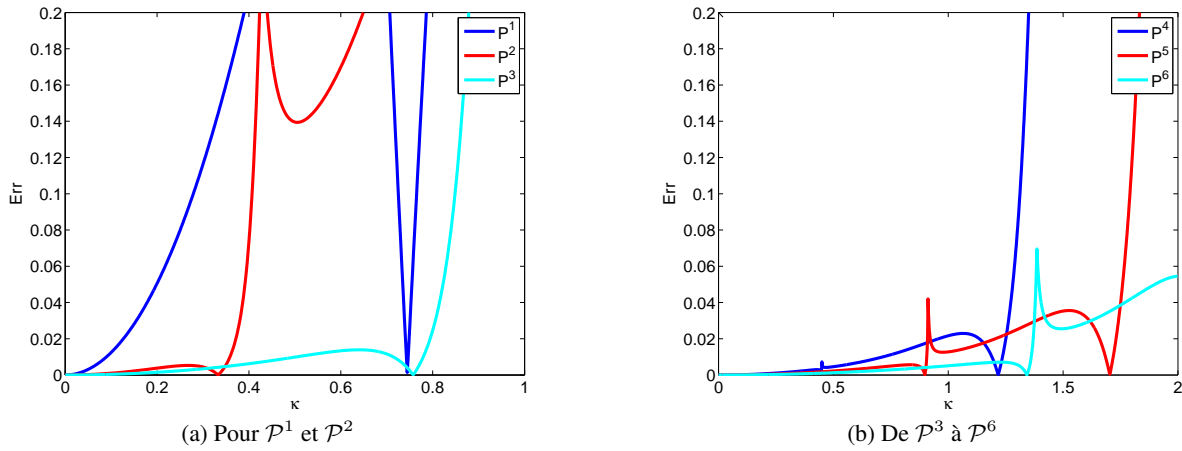


FIGURE 4.9 – Erreur sur le calcul de la valeur propre physique

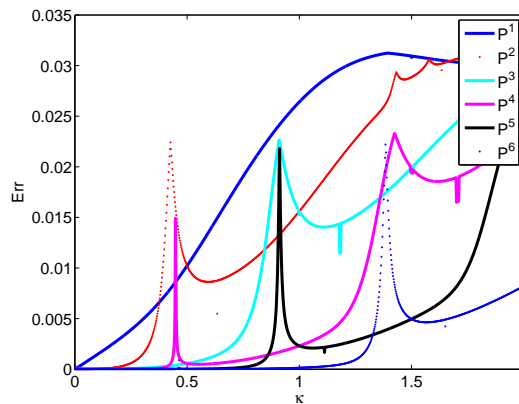


FIGURE 4.10 – Erreur sur le calcul du vecteur propre physique

À partir des courbes en échelle  $\log - \log$  des figures 4.11.(a) et 4.11.(b), nous pouvons déterminer l'ordre de convergence du schéma Leap-Frog pour le calcul de la vitesse de phase et de la projection de la solution exacte sur le sous-espace propre (Tabs. 4.3.(a) et (b)). Quel que soit l'ordre d'approximation en espace, la vitesse de phase converge à l'ordre 2, ce qui était attendu puisque le schéma Leap-Frog est d'ordre deux en temps. En ce qui concerne la projection de la solution exacte, qui dépend



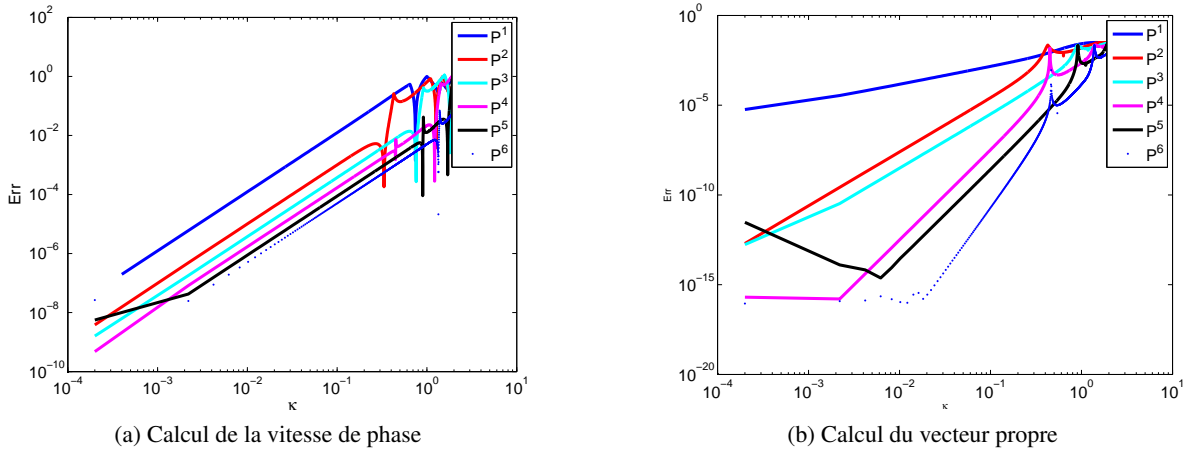


FIGURE 4.11 – Erreurs en échelle log – log

uniquement du schéma en espace, on constate que les schémas de degré  $2m$  et  $2m + 1$  convergent à l'ordre  $2m + 1$ , alors qu'on aurait pu s'attendre à une convergence d'ordre  $m + 1$  pour des polynômes de degré  $m$ .

Nous pouvons bien évidemment mener une étude similaire pour les schémas  $\nabla$  et ADER. Cependant, nous ne présentons pas toutes les courbes et nous nous limitons à indiquer les ordres de convergence dans les tableaux 4.3.(a) et (b). On remarque que la convergence sur la vitesse de phase est d'ordre  $\min(4, m + 1)$ , ce qui était attendu puisque les schémas sont d'ordre 4 en temps. Cependant, la convergence de la projection de la solution exacte, qui ne dépend que de la discrétisation en espace, reste sous-optimale pour  $m$  impair. Ce phénomène explique pourquoi nous n'avons pas pu obtenir la convergence optimale pour des éléments  $\mathcal{P}^3$  avec des conditions initiales. Le problème de la source ponctuelle s'explique lui par la présence de modes parasites pour certaines valeurs de  $kh$ .

$m$	Schéma LF	Schéma $\nabla$	Schéma ADER
1	2.01	2.07	1.86
2	1.99	5.40	4.32
3	2.00	3.65	3.97
4	2.00	4.00	4.01
5	2.00	4.00	4.00
6	2.00	3.98	4.00

(a) Ordre de convergence de la vitesse de phase

Schéma LF	Schéma $\nabla$	Schéma ADER
0.99	1.02	1.02
3.03	3.10	3.10
3.06	3.05	3.04
5.07	5.09	5.10
5.05	5.03	5.04
7.07	7.08	7.09

(b) Ordre de convergence du vecteur propre

TABLE 4.3 – Ordre de convergence des schémas Leap-Frog,  $\nabla$  et ADER

L'analyse que nous venons de mener montre que le schéma en espace laisse exister des modes parasites et que la convergence n'est pas optimale pour des degrés de discrétisation impairs. Il semble donc nécessaire de modifier le schéma en espace. Pour cela, nous proposons de pénaliser la formulation DG que nous avons choisie. En introduisant un paramètre de pénalisation, nous espérons que les modes parasites disparaissent, ce qui devrait se traduire par un ordre de convergence optimal.

#### 4.4 Schéma avec pénalisation

Nous allons étudier l'effet de l'introduction d'un paramètre de pénalisation dans la formulation variationnelle. Le schéma proposé par Ainsworth, Monk et Muniz [4] s'écrit

$$(4.4.1) \quad \left\{ \begin{array}{l} \sum_K \int_K \frac{\partial p(x,t)}{\partial t} q(x) + \sum_{\Gamma} \int_{\Gamma} \{ \{ \mathbf{v}(x,t) \} \} [[q(x)]] \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(x,t) \cdot \nabla q(x) \\ \quad + \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma}^{11} [[p(x,t)]] [[q(x)]] + \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma}^{12} \cdot [[\mathbf{v}(x,t)]] [[q(x)]] = 0 \\ \sum_K \int_K \frac{\partial \mathbf{v}(x,t)}{\partial t} \cdot \mathbf{w}(x) + \sum_{\Gamma} \int_{\Gamma} \{ \{ p(x,t) \} \} [[\mathbf{w}(x)]] \cdot \mathbf{n} - \sum_K \int_K p(x,t) \nabla \cdot \mathbf{w}(x) \\ \quad + \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma}^{21} \cdot [[\mathbf{w}(x,t)]] [[p(x)]] + \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma}^{22} [[\mathbf{w}(x,t)]] [[p(x)]] = 0 \end{array} \right.$$

où  $\alpha_{\Gamma}^{11}$  et  $\alpha_{\Gamma}^{22}$  sont des scalaires dont la valeur dépend de la face  $\Gamma$  et  $\alpha_{\Gamma}^{12}$  et  $\alpha_{\Gamma}^{21}$  sont des vecteurs dont la norme et la direction dépendent de  $\Gamma$ . On pourrait s'étonner de ne pas voir apparaître le vecteur normal  $\mathbf{n}$  au côté des termes de saut. Cela est dû au fait que nous considérons des produits de deux termes de saut et que  $\mathbf{n} \cdot \mathbf{n} = 1$ . Remarquons également que les quatre termes de pénalisation ne perturbent pas la consistance du schéma, car  $[[p]] = 0$  et  $[[\mathbf{v}]] \cdot \mathbf{n} = 0$  pour  $p$  et  $\mathbf{v}$  suffisamment régulières.

Pour déterminer plus précisément les paramètres  $\alpha_{\Gamma}^{ij}$ , nous allons étudier leur effet sur l'énergie définie dans la propriété 2.2.1 du chapitre 2. En posant  $q = p$  et  $\mathbf{w} = \mathbf{v}$ , on vérifie facilement que

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \left( \sum_K \int_K \frac{1}{\mu(\mathbf{x})} |p(\mathbf{x},t)|^2 + \sum_K \int_K \rho(\mathbf{x}) \|\mathbf{v}(\mathbf{x},t)\|^2 \right) = \\ & - \sum_{\Gamma} \int_{\Gamma} \left( \alpha_{\Gamma}^{11} |[p(x,t)]|^2 + \alpha_{\Gamma}^{22} |[[\mathbf{v}(x)]]|^2 + (\alpha_{\Gamma}^{12} + \alpha_{\Gamma}^{21}) \cdot [[\mathbf{v}(x,t)]] [[p(x)]] \right). \end{aligned}$$

Pour garantir la conservation de l'énergie

$$\sum_K \int_K \frac{1}{\mu(\mathbf{x})} |p(\mathbf{x},t)|^2 + \sum_K \int_K \rho(\mathbf{x}) \|\mathbf{v}(\mathbf{x},t)\|^2,$$

il faut donc choisir  $\alpha_{\Gamma}^{11} = \alpha_{\Gamma}^{22} = 0$  et  $\alpha_{\Gamma}^{12} = -\alpha_{\Gamma}^{21}$ . Pour simplifier les notations, nous poserons donc par la suite  $\alpha_{\Gamma} = \alpha_{\Gamma}^{12}$ , et nous considèrerons la formulation variationnelle

$$(4.4.2) \quad \left\{ \begin{array}{l} \sum_K \int_K \frac{\partial p(x,t)}{\partial t} q(x) + \sum_{\Gamma} \int_{\Gamma} \{ \{ \mathbf{v}(x,t) \} \} [[q(x)]] \cdot \mathbf{n} - \sum_K \int_K \mathbf{v}(x,t) \cdot \nabla q(x) \\ \quad + \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma} \cdot [[\mathbf{v}(x,t)]] [[q(x)]] = 0 \\ \sum_K \int_K \frac{\partial \mathbf{v}(x,t)}{\partial t} \cdot \mathbf{w}(x) + \sum_{\Gamma} \int_{\Gamma} \{ \{ p(x,t) \} \} [[\mathbf{w}(x)]] \cdot \mathbf{n} - \sum_K \int_K p(x,t) \nabla \cdot \mathbf{w}(x) \\ \quad - \sum_{\Gamma} \int_{\Gamma} \alpha_{\Gamma} \cdot [[\mathbf{w}(x,t)]] [[p(x)]] = 0 \end{array} \right.$$

Il reste maintenant à déterminer la norme et l'orientation du vecteur  $\alpha_{\Gamma}$ . Ainsworth *et al* [4] ont considéré des maillages cartésiens en dimension deux, et ils ont choisi  $\alpha_{\Gamma}$  orthogonal aux arêtes du

maillage, dirigé vers la droite pour les arêtes verticales, et vers le haut pour les arêtes horizontales (voir Fig. 4.12). Comme nous le verrons par la suite, le choix de la direction est beaucoup plus compliqué pour des maillages triangulaires. Pour le moment, nous nous concentrons sur le cas 1D, où  $\alpha_\Gamma$  est un scalaire. Le maillage étant uniforme, nous supposons que ce paramètre est constant et nous posons  $\alpha_\Gamma = \alpha$ . Pour la suite de ce travail, nous considérerons  $\alpha = 0.5$  indépendamment de l'ordre de discrétisation en espace, ce qui nous a semblé plausible lors des tests numériques.

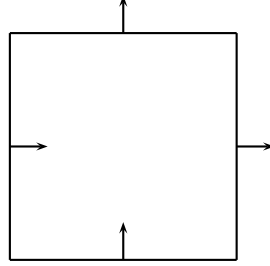


FIGURE 4.12 – Orientation du vecteur  $\alpha_\Gamma$  pour des maillages cartésiens.

Nous allons maintenant refaire l'analyse de dispersion en tenant compte du paramètre de pénalisation. Pour cela, il nous suffit de remplacer respectivement les matrices  $K$ ,  $K^W$  et  $K^E$  par les matrices  $K_\alpha$ ,  $K_\alpha^W$  et  $K_\alpha^E$  définies par

$$K_{\alpha ij} = K_{ij} - \alpha \left( \hat{\phi}_j(1)\hat{\phi}_i(1) - \alpha\hat{\phi}_j(0)\hat{\phi}_i(0) \right),$$

$$K_\alpha^W{}_{ij} = K_{ij}^W + \alpha\hat{\phi}_j(0)\hat{\phi}_i(1) \text{ et } K_\alpha^E{}_{ij} = K_{ij}^E + \alpha\hat{\phi}_j(1)\hat{\phi}_i(0).$$

Pour les schémas Leap-Frog et  $\nabla$ , le schéma totalement discrétisé se réécrit alors élément par élément sous la forme

$$(4.4.3) \quad \begin{cases} hM \frac{V_I^{n+\frac{1}{2}} - V_I^{n-\frac{1}{2}}}{\Delta t} = K_\alpha P_I^n + K_\alpha^W P_{I-1}^n + K_\alpha^E P_{I+1}^n, \\ hM \frac{P_I^{n+1} - P_I^n}{\Delta t} = K_{-\alpha} V_I^{n+\frac{1}{2}} + K_{-\alpha}^W V_{I-1}^{n+\frac{1}{2}} + K_{-\alpha}^E V_{I+1}^{n+\frac{1}{2}}, \end{cases}$$

Le schéma ADER d'ordre 4 s'écrit sous la forme

$$(4.4.4) \quad \begin{cases} hM \frac{V_I^{n+\frac{1}{2}} - V_I^{n-\frac{1}{2}}}{\Delta t} = K_\alpha P_I^n + K_\alpha^W P_{I-1}^n + K_\alpha^E P_{I+1}^n \\ \quad + \frac{\Delta t^2}{24} \left( K_\alpha \tilde{P}_I^n + K_\alpha^W \tilde{P}_{I-1}^n + K_\alpha^E \tilde{P}_{I+1}^n \right), \\ hM \frac{P_I^{n+1} - P_I^n}{\Delta t} = K_{-\alpha} V_I^{n+\frac{1}{2}} + K_{-\alpha}^W V_{I-1}^{n+\frac{1}{2}} + K_{-\alpha}^E V_{I+1}^{n+\frac{1}{2}} \\ \quad + \frac{\Delta t^2}{24} \left( K_{-\alpha} \tilde{V}_I^{n+\frac{1}{2}} + K_{-\alpha}^W \tilde{V}_{I-1}^{n+\frac{1}{2}} + K_{-\alpha}^E \tilde{V}_{I+1}^{n+\frac{1}{2}} \right), \end{cases}$$

où  $M$  et  $K$  sont les matrices définies pour le schéma Leap-Frog et le vecteur  $\tilde{P}_I^n$  est défini pour tout  $I$  de  $\mathbb{Z}$  par

$$(4.4.5) \quad \begin{cases} hM \tilde{V}_I^{n+\frac{1}{2}} = K_\alpha P_I^n + K_\alpha^W P_{I-1}^n + K_\alpha^E P_{I+1}^n \\ hM \tilde{P}_I^n = K_{-\alpha} \tilde{V}_I^{n+\frac{1}{2}} + K_{-\alpha}^W \tilde{V}_{I-1}^{n+\frac{1}{2}} + K_{-\alpha}^E \tilde{V}_{I+1}^{n+\frac{1}{2}}. \end{cases}$$

Le vecteur  $\tilde{V}_I^{n+\frac{1}{2}}$  est défini pour tout  $I \in \mathbb{Z}$  par

$$(4.4.6) \quad \begin{cases} hM\tilde{P}_I^{n+1} = K_{-\alpha}V_I^{n+\frac{1}{2}} + K_{-\alpha}^W V_{I-1}^{n+\frac{1}{2}} + K_{-\alpha}^E V_{I+1}^{n+\frac{1}{2}}, \\ hM\tilde{V}_I^{n+\frac{1}{2}} = K_{\alpha}\tilde{P}_I^n + K_{\alpha}^W \tilde{P}_{I-1}^n + K_{\alpha}^E \tilde{P}_{I+1}^n. \end{cases}$$

En posant

$$P_{I+1}^{n+1} = e^{i(kh+\omega_h\Delta t)} P_I^n \text{ et } V_{I+1}^{n+\frac{1}{2}} = e^{i(kh+\omega_h\Delta t)} V_I^{n-\frac{1}{2}},$$

l'équation (4.4.3) devient, pour les schémas Leap-Frog et  $\nabla$

$$(4.4.7) \quad \begin{cases} \frac{2i}{h\Delta t} \sin\left(\frac{\omega_h\Delta t}{2}\right) e^{i\frac{\omega_h\Delta t}{2}} M V_I^{n+\frac{1}{2}} = \left(K_{\alpha} + K_{\alpha}^W e^{-ikh} + K_{\alpha}^E e^{ikh}\right) P_I^n, \\ \frac{2i}{h\Delta t} \sin\left(\frac{\omega_h\Delta t}{2}\right) e^{-i\frac{\omega_h\Delta t}{2}} M P_I^n = \left(K_{-\alpha} + K_{-\alpha}^W e^{-ikh} + K_{-\alpha}^E e^{ikh}\right) V_I^{n+\frac{1}{2}}. \end{cases}$$

Ce système peut se réécrire uniquement en fonction de  $P_I^n$  sous la forme

$$(4.4.8) \quad \sin^2\left(\frac{\omega_h\Delta t}{2}\right) P_I^n = \frac{\Delta t^2}{4h^2} K_{\alpha}^*(kh) P_I^n,$$

avec

$$K^*(kh) = -M^{-1} \left( K_{-\alpha} + K_{-\alpha}^W e^{-ikh} + K_{-\alpha}^E e^{ikh} \right) M^{-1} \left( K_{\alpha} + K_{\alpha}^W e^{-ikh} + K_{\alpha}^E e^{ikh} \right).$$

Comme pour le cas sans pénalisation, on note  $K_{\alpha LF}^*(kh)$  et  $K_{\alpha \nabla}^*(kh)$  les matrices associées respectivement au schéma Leap-Frog et au schéma  $\nabla$ . Par la suite nous utilisons la notation

$$\mathcal{K}_{\alpha} = K_{\alpha} + K_{\alpha}^W e^{-ikh} + K_{\alpha}^E e^{ikh}.$$

Pour le schéma ADER, on obtient le système

$$(4.4.9) \quad \sin^2\left(\frac{\omega_h\Delta t}{2}\right) P_I^n = \frac{\Delta t^2}{4h^2} K_{\alpha ADER}^* P_I^n,$$

avec

$$(4.4.10) \quad K_{\alpha ADER}^* = - \begin{pmatrix} M^{-1}\mathcal{K}_{-\alpha} + \frac{\Delta t^2}{24h^2} M^{-1}\mathcal{K}_{-\alpha} M^{-1}\mathcal{K}_{\alpha} M^{-1}\mathcal{K}_{-\alpha} \\ M^{-1}\mathcal{K}_{\alpha} + \frac{\Delta t^2}{24h^2} M^{-1}\mathcal{K}_{\alpha} M^{-1}\mathcal{K}_{-\alpha} M^{-1}\mathcal{K}_{\alpha} \end{pmatrix}.$$

Nous allons maintenant montrer que les matrices  $K_{\alpha}^*$  sont diagonalisables dans  $\mathbb{R}$  et que leurs valeurs propres sont réelles. Pour cela, nous utiliserons les mêmes arguments que ceux utilisés pour le cas  $\alpha = 0$ .

**Theorème 4.4.1.** *La matrice  $\mathcal{K}_{-\alpha} M^{-1} \mathcal{K}_{\alpha}$  est Hermitienne.*

**Démonstration :** Il suffit de remarquer que  $\mathcal{K}_{\alpha}^T = -\overline{\mathcal{K}_{-\alpha}}$ , de telle sorte que

$$\left(\mathcal{K}_{-\alpha} M^{-1} \mathcal{K}_{\alpha}\right)^T = \mathcal{K}_{\alpha}^T M^{-1} \mathcal{K}_{-\alpha}^T = \overline{\mathcal{K}_{-\alpha} M^{-1} \mathcal{K}_{\alpha}}$$

■

On en déduit ensuite que  $M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}$  est diagonalisable et que ses valeurs propres sont réelles en utilisant le lemme suivant, dont la démonstration est identique à celle du lemme 4.1.1

**Lemme 4.4.1.** *Si  $M$  est une matrice symétrique définie positive et  $N$  une matrice Hermitienne,  $M^{-1}N$  est diagonalisable et toutes ses valeurs propres sont réelles.*

**Corollaire 4.4.1.** *Toutes les valeurs propres de  $K_{\alpha_{LF}}^*$  et  $K_{\alpha_{\nabla}}^*$  sont réelles.*

Nous avons un résultat similaire pour  $K_{\alpha_{ADER}}^*$

**Theorème 4.4.2.** *Toutes les valeurs propres de  $K_{\alpha_{ADER}}^*$  sont réelles.*

**Démonstration :** On réécrit tout d'abord  $K_{\alpha_{ADER}}^*$  sous la forme

$$(4.4.11) \quad \begin{aligned} K_{\alpha_{ADER}}^* = & - M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha} \\ & - \frac{\Delta t^2}{24h^2}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha} \\ & - \frac{\Delta t^2}{24h^2}M^{-1}\mathcal{K}_{\alpha}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}M^{-1}\mathcal{K}_{-\alpha} \\ & - \frac{\Delta t^4}{576h^4}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}. \end{aligned}$$

On observe donc que la matrice  $K_{\alpha_{ADER}}^*$  s'exprime comme des puissances de la matrice diagonalisable  $M^{-1}\mathcal{K}_{-\alpha}M^{-1}\mathcal{K}_{\alpha}$  dont les valeurs propres sont réelles. On en déduit donc que  $K_{\alpha_{ADER}}^*$  est diagonalisable et que ses valeurs propres sont réelles.

■

Il reste encore à établir la positivité des valeurs propres. Nous n'avons pas réussi à démontrer cette propriété mais les tests numériques que nous avons réalisés pour de nombreuses valeurs de  $\alpha$  et de  $kh$  nous ont permis d'établir la conjecture suivante.

**Conjecture 4.4.1.** *Toutes les valeurs propres de  $K_{\alpha_{LF}}^*$ ,  $K_{\alpha_{\nabla}}^*$  et  $K_{\alpha_{ADER}}^*$  sont positives.*

#### 4.4.1 Condition nécessaire de stabilité des schémas avec pénalisation

En procédant comme pour le cas sans pénalisation, nous pouvons maintenant déterminer les conditions de stabilité des trois schémas avec pénalisation. Dans le tableau 4.4, nous donnons le coefficient de stabilité obtenu avec  $\alpha = 0.5$  et nous rappelons celui obtenu sans pénalisation. Le ratio entre ces deux coefficients est d'environ 2/3, ce qui veut dire que le schéma avec pénalisation est *a priori* plus coûteux en temps de calcul.

#### 4.4.2 Analyse de dispersion des schémas avec pénalisation

Nous nous concentrons maintenant sur le schéma Leap-Frog et nous représentons l'évolution des différentes vitesses de phase en fonction de  $\kappa$  pour  $m = 1..6$  sur les figures 4.13 et 4.14. Les courbes de gauche sont celles obtenues avec  $\alpha = 0.5$ , les courbes de droite sont celles que nous avons obtenues sans pénalisation. Nous remarquons que l'introduction du paramètre de pénalisation

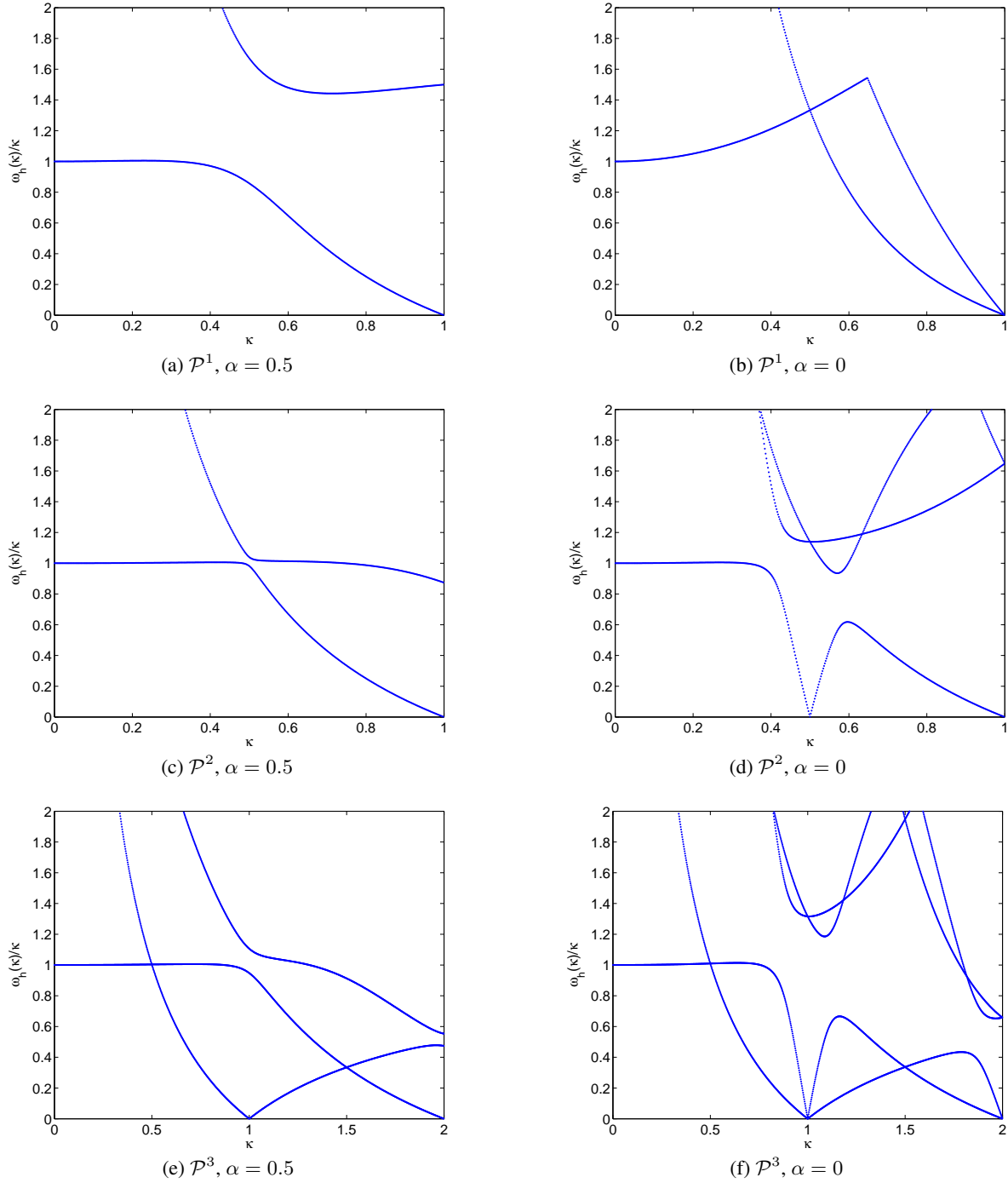


FIGURE 4.13 – Courbes de dispersion pour le schéma Leap-Frog avec pénalisation  $\alpha = 0.5$  pour des éléments finis de  $\mathcal{P}^1$  à  $\mathcal{P}^3$

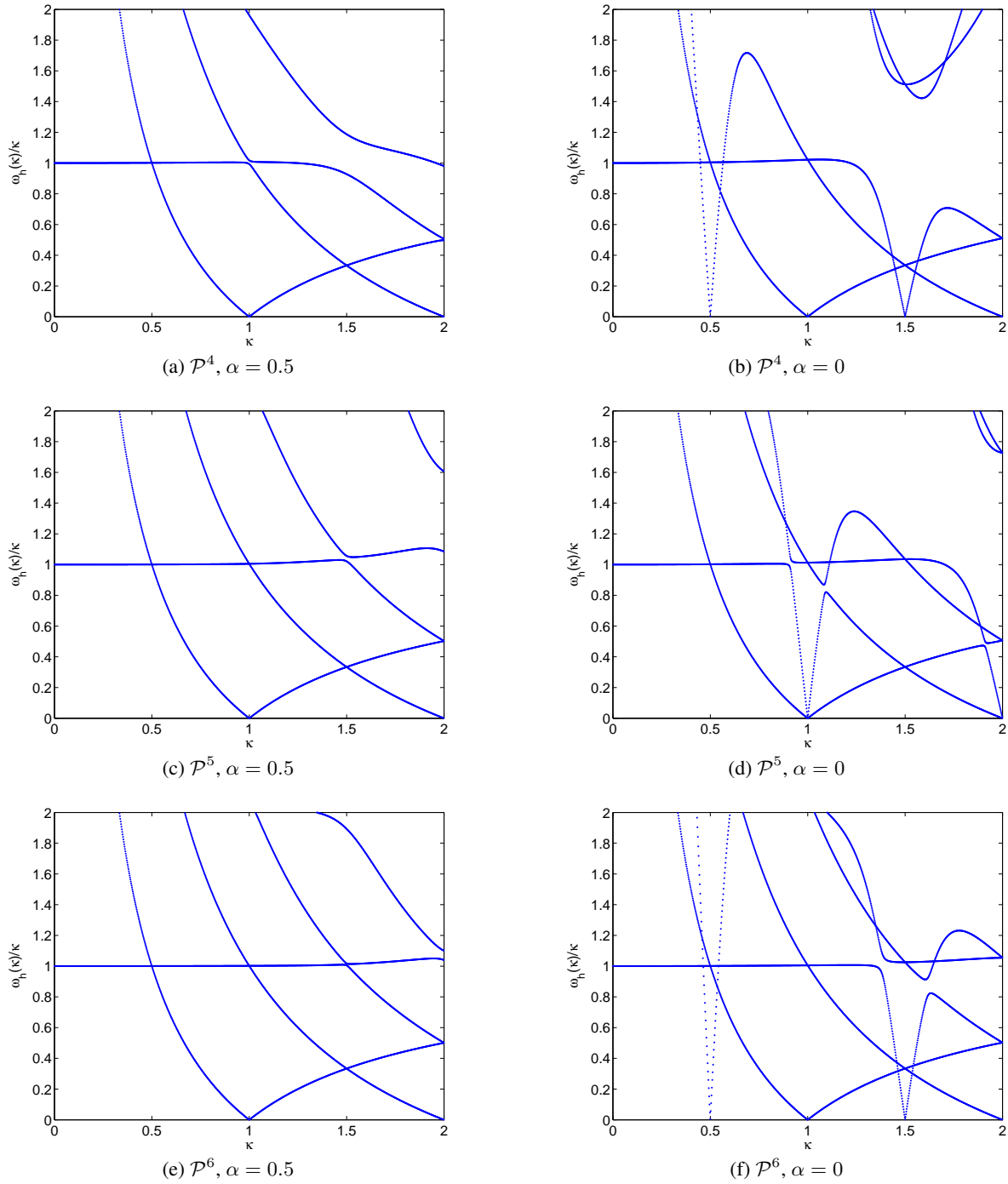


FIGURE 4.14 – Courbes de dispersion pour le schéma Leap-Frog avec pénalisation  $\alpha = 0.5$  pour des éléments finis de  $\mathcal{P}^4$  à  $\mathcal{P}^6$

	$\mathcal{P}^1$	$\mathcal{P}^2$	$\mathcal{P}^3$	$\mathcal{P}^4$	$\mathcal{P}^5$	$\mathcal{P}^6$
$\alpha = 0$	0.5	0.247	0.15	0.01	0.0732	0.0552
$\alpha = 0.5$	0.333	0.164	0.0955	0.0619	0.0451	0.0335

TABLE 4.4 – Coefficients de stabilité du schéma Leap-Frog, avec  $\alpha = 0$  et  $\alpha = 0.5$  pour des approximations en espace de  $\mathcal{P}^0$  à  $\mathcal{P}^6$ .

supprime une valeur propre nulle pour  $\kappa = 0.5$  pour les ordres pairs et pour  $\kappa = 1$  pour les ordres impairs.

Intéressons-nous maintenant à l'effet du paramètre de pénalisation sur les vecteurs propres. Nous donnons dans le tableau 4.5.(a) les valeurs propres et leur multiplicité pour  $\kappa = 0.25$  et  $\mathcal{P}^5$  et nous représentons sur les figures 4.15 les vecteurs propres associés. Nous voyons que le cinquième vecteur propre approche à nouveau parfaitement la solution  $e^{2i\pi x/4}$ .

i	$\lambda_i$	$\nu(i)$
1	1820	1
2	546	1
3	138	1
4	63	1
5	2.47	1
6	22.2	1

(a)  $\kappa = 0.25$  et  $\mathcal{P}^5$

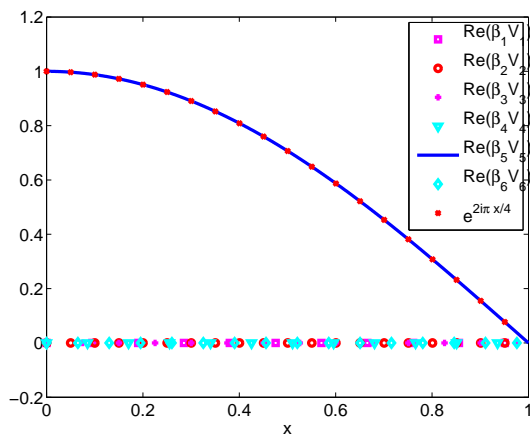
i	$\lambda_i$	$\nu(i)$
1	1970	1
2	380	1
3	180	1
4	0	1
5	398	1
6	395	1

(b)  $\kappa = 1$  et  $\mathcal{P}^5$

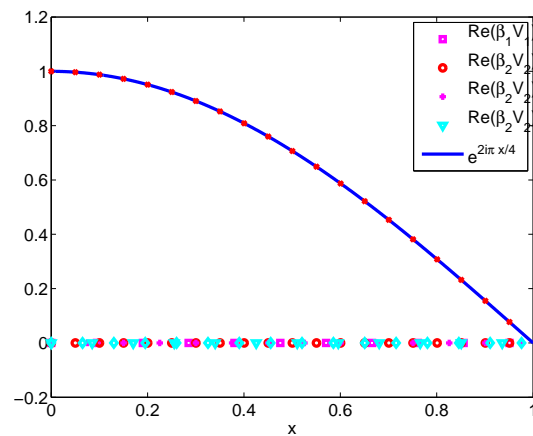
i	$\lambda_i$	$\nu(i)$
1	1020	1
2	145	1
3	0	1
4	40.4	1
5	38.5	1

(c)  $\kappa = 1$  et  $\mathcal{P}^4$

TABLE 4.5 – Valeurs propres et leur multiplicité pour un coefficient de pénalisation  $\alpha = 0.5$



(a) Partie réelle



(b) Partie imaginaire

FIGURE 4.15 – Vecteurs propres pour  $\kappa = 0.25$  et  $\mathcal{P}^5$  pour un coefficient de pénalisation  $\alpha = 0.5$

Contrairement au cas où  $\alpha = 0$ , il n'y a plus qu'une valeur propre nulle pour  $\kappa = 1$ . Nous donnons dans le tableau 4.5.(b) les valeurs propres et leur multiplicité pour  $\kappa = 1$  et  $\mathcal{P}^5$  et nous représentons sur les figures 4.16.(a) et (b) les vecteurs propres associés. La cinquième valeur permet d'approcher



relativement correctement la partie réelle de la solution  $e^{2i\pi x}$ , alors que la sixième permet d’approcher la partie imaginaire. Ce comportement s’explique par le fait que ces deux valeurs sont très proches. Ainsi, la somme des deux vecteurs propres nous donne une bonne approximation des parties réelles et imaginaires. En ce qui concerne la valeur propre nulle, le vecteur qui lui est associé (triangles cyans) est quasiment nul. Cette valeur propre ne génère donc pas de modes parasites pour  $\alpha = 0.5$ , contrairement à ce nous avons observé avec le schéma sans pénalisation.

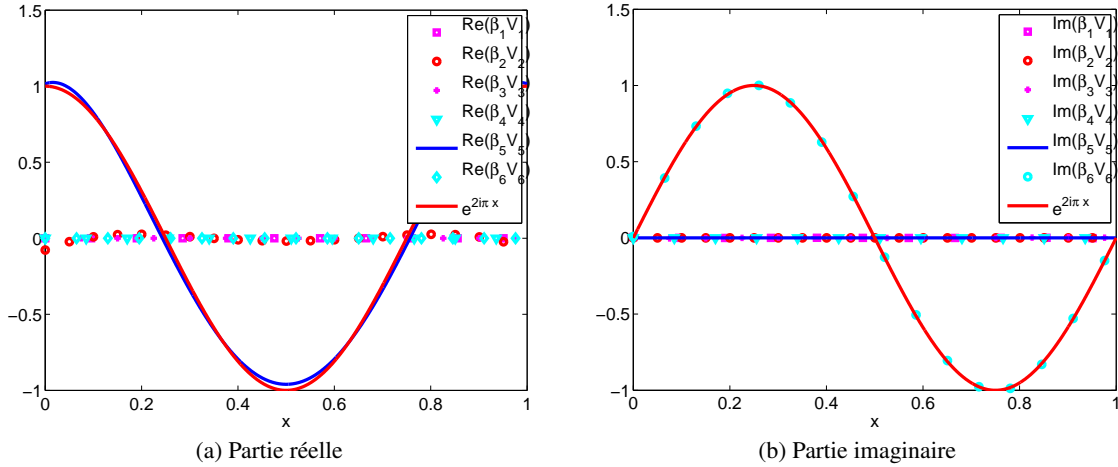


FIGURE 4.16 – Vecteurs propres pour  $\kappa = 1$  et  $\mathcal{P}^5$  pour un coefficient de pénalisation  $\alpha = 0.5$ . L’ajout du terme de pénalisation a permis de supprimer l’effet des modes parasites visibles sur la figure 4.5

Pour les ordres pairs, nous considérons  $m = 4$  et le cas  $\kappa = 1$ , puisqu’il n’y a plus de valeur propre nulle en  $\kappa = 0.5$ . Les valeurs propres sont représentées dans le tableau 4.5.(c) et les vecteurs propres sur la figure 4.17. Comme pour le cas précédent, le mode parasite de vitesse nulle (pointillés cyans) est d’amplitude négligeable. Cependant, aucun vecteur propre ne permet d’approcher correctement la solution. Ceci est à nouveau dû au fait que les deux valeurs propres à 38.5 et 40.4 sont relativement proches. Nous avons représenté la somme de ces deux vecteurs propres sur la figure 4.18 pour montrer que cette somme permettait d’approcher correctement la solution.

Sur les figures 4.19 et 4.20 nous représentons en pointillés rouges la valeur propre approchant au mieux la valeur propre physique. Les figures de gauche représentent les résultats obtenus pour un paramètre de pénalisation  $\alpha = 0.5$ , celles de droite sont celles que nous avons obtenues sans le terme de pénalisation.

Nous remarquons que l’ajout du paramètre de pénalisation permet d’obtenir une meilleure approximation de la vitesse de phase pour un faible nombre d’éléments par longueur d’onde. Cette observation sera d’ailleurs confirmée lors de l’étude de la convergence du schéma.

Sur les figures 4.21.(a) et (b), nous représentons les variations de l’erreur sur la vitesse de phase en fonction de  $\kappa$ , pour  $m = 1.6$  et sur la figure 4.22, les variations de l’erreur sur le vecteur propre physique en fonction de  $\kappa$ . Nous représentons ces erreurs en échelle  $\log - \log$  sur les figures 4.23.(a) et (b) et nous donnons la pente obtenue par régression linéaire dans le tableau 4.6.

Le paramètre de pénalisation permet donc d’obtenir l’ordre de convergence optimal pour le calcul des vecteurs propres quel que soit  $m$ , alors qu’avec le schéma sans pénalisation, on perd un ordre de convergence pour  $m$  impair. Pour illustrer ce point, nous comparons les courbes de convergence des valeurs propres et vecteurs propres pour  $m = 1.6$  sur les figures 4.24 et 4.25.

Pour finir, nous donnons dans les tableaux 4.7 et 4.8 le nombre minimum d’éléments par longueur

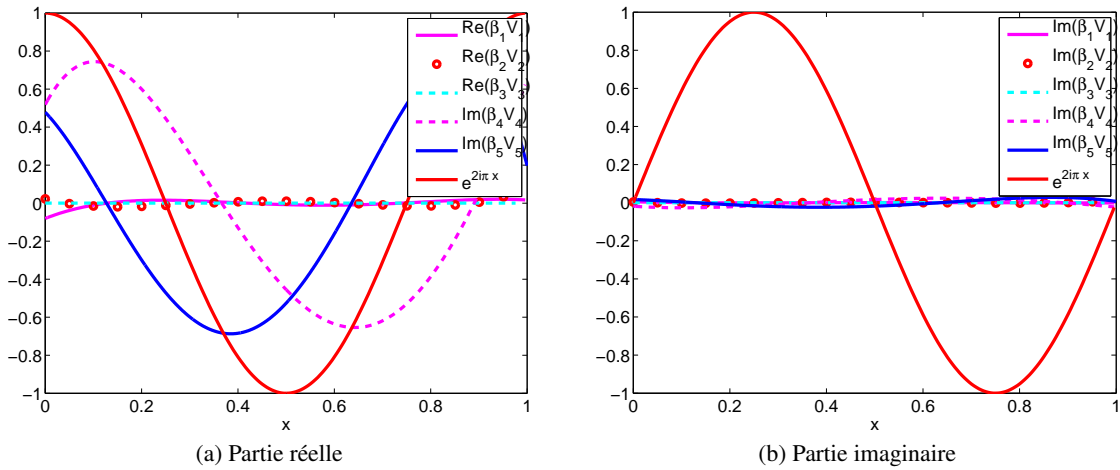


FIGURE 4.17 – Vecteurs propres pour  $\kappa = 1$  et  $\mathcal{P}^4$  pour un coefficient de pénalisation  $\alpha = 0.5$

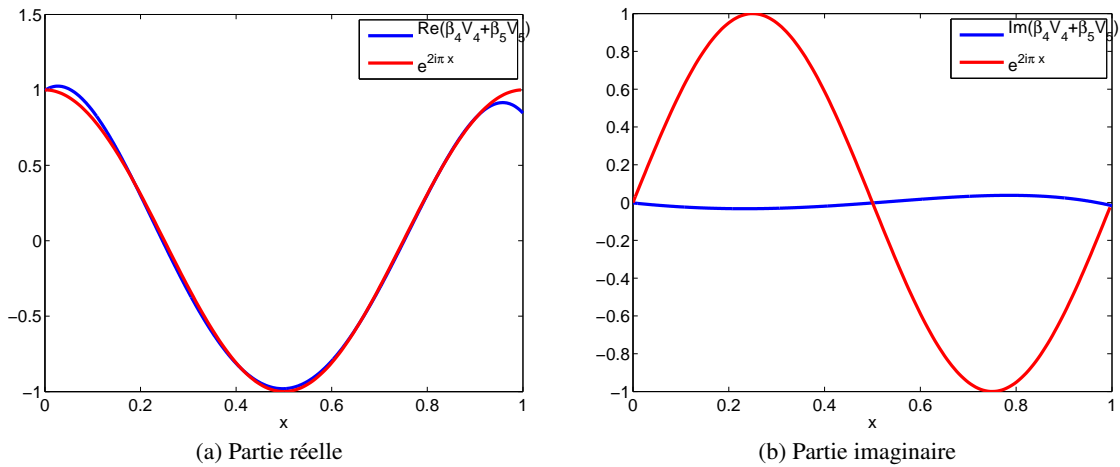


FIGURE 4.18 – Vecteurs propres pour  $\kappa = 1$  et  $\mathcal{P}^4$  pour un coefficient de pénalisation  $\alpha = 0.5$

$m$	Vitesse de phase	Vecteur propre
1	1.99	2.01
2	1.99	3.00
3	2.00	3.99
4	2.00	4.99
5	2.00	5.96
6	2.00	7.01

(a) Schéma Leap-Frog

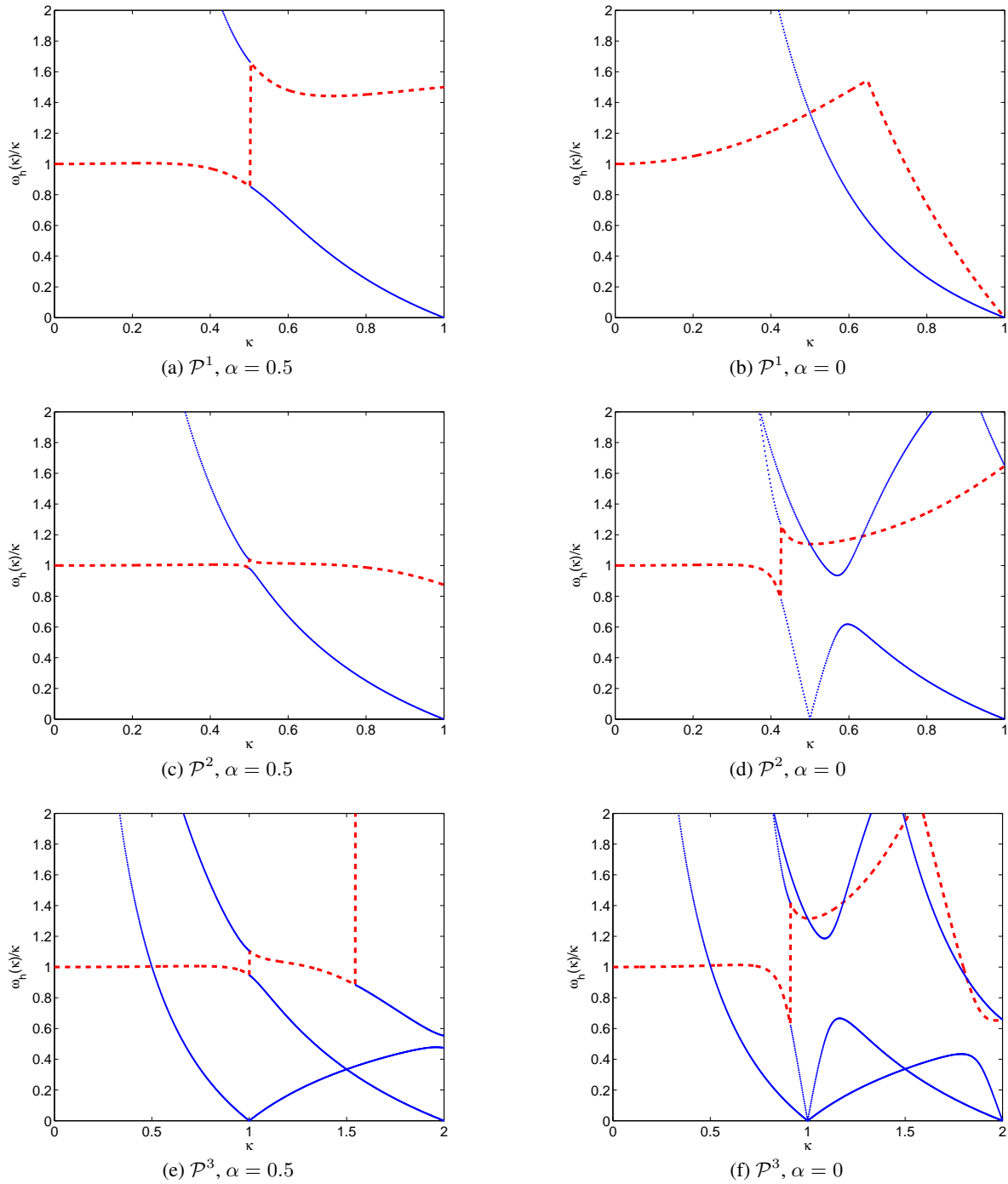
Vitesse de phase	Vecteur propre
4.1	2.01
6.1	2.88
4.3	4.05
3.99	4.91
3.97	6.04
3.83	6.90

(b) Schéma  $\nabla$

Vitesse de phase	Vecteur propre
4.05	2.01
4.12	3.00
4.01	4.00
4.00	5.00
4.02	5.96
3.99	7.02

(c) Schéma ADER

TABLE 4.6 – Ordre de convergence des schémas Leap-Frog,  $\nabla$  et ADER avec pénalisation

FIGURE 4.19 – Courbes de dispersion pour le schéma Leap-Frog pour des éléments  $\mathcal{P}^1$  à  $\mathcal{P}^3$

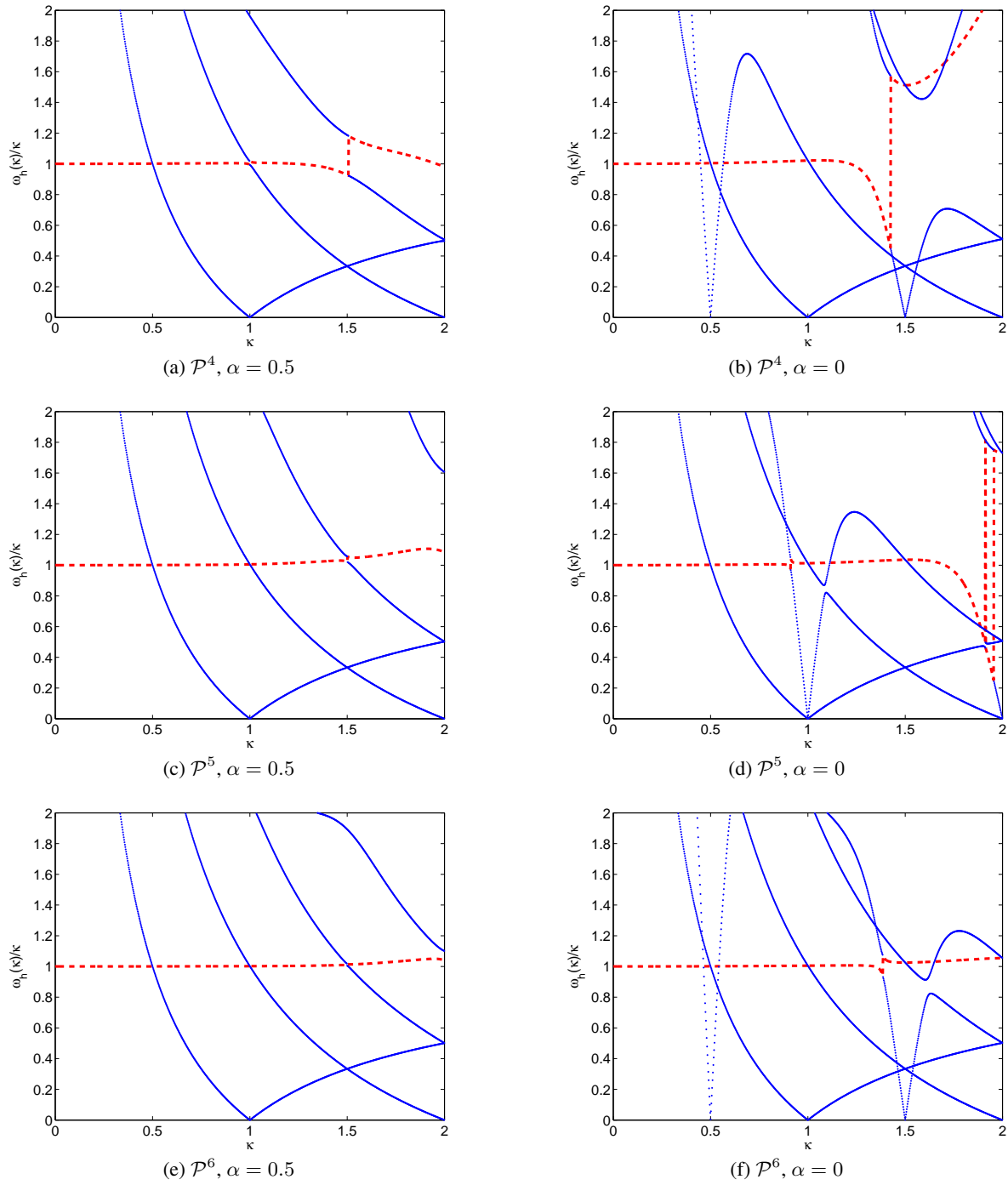


FIGURE 4.20 – Courbes de dispersion pour le schéma Leap-Frog pour des éléments  $\mathcal{P}^4$  à  $\mathcal{P}^6$

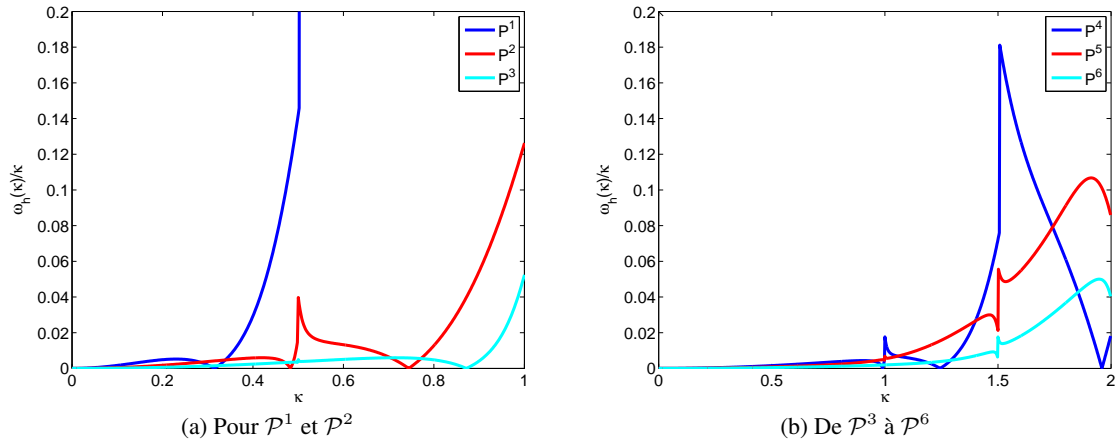


FIGURE 4.21 – Erreur sur le calcul de la valeur propre physique

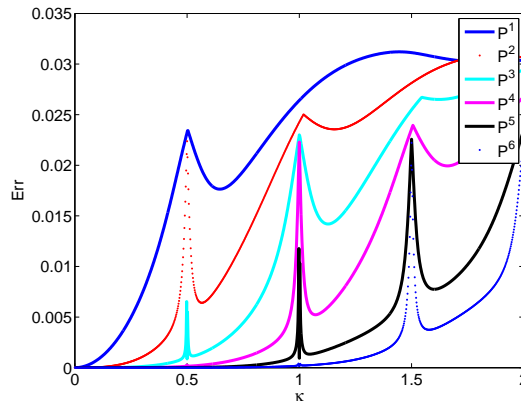


FIGURE 4.22 – Erreur sur le calcul du vecteur propre physique

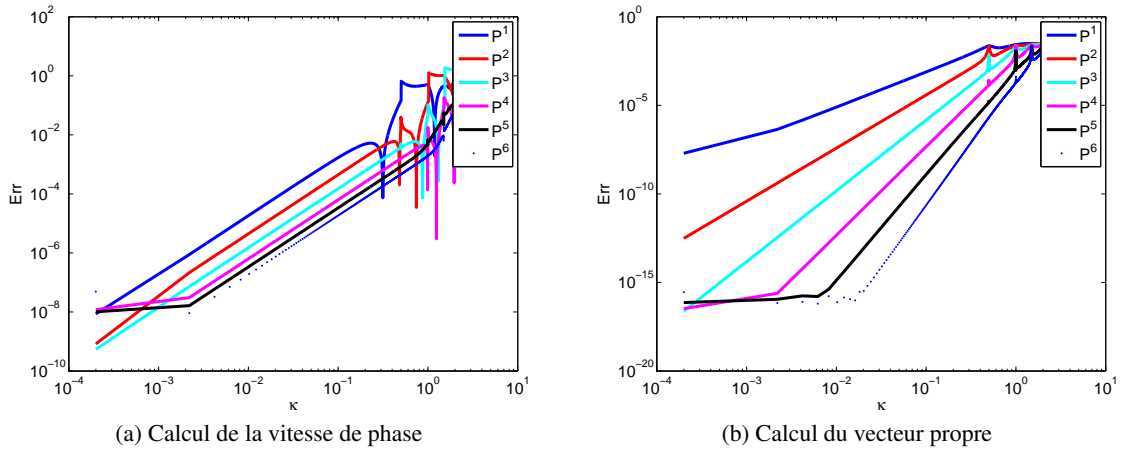


FIGURE 4.23 – Erreurs en échelle log-log

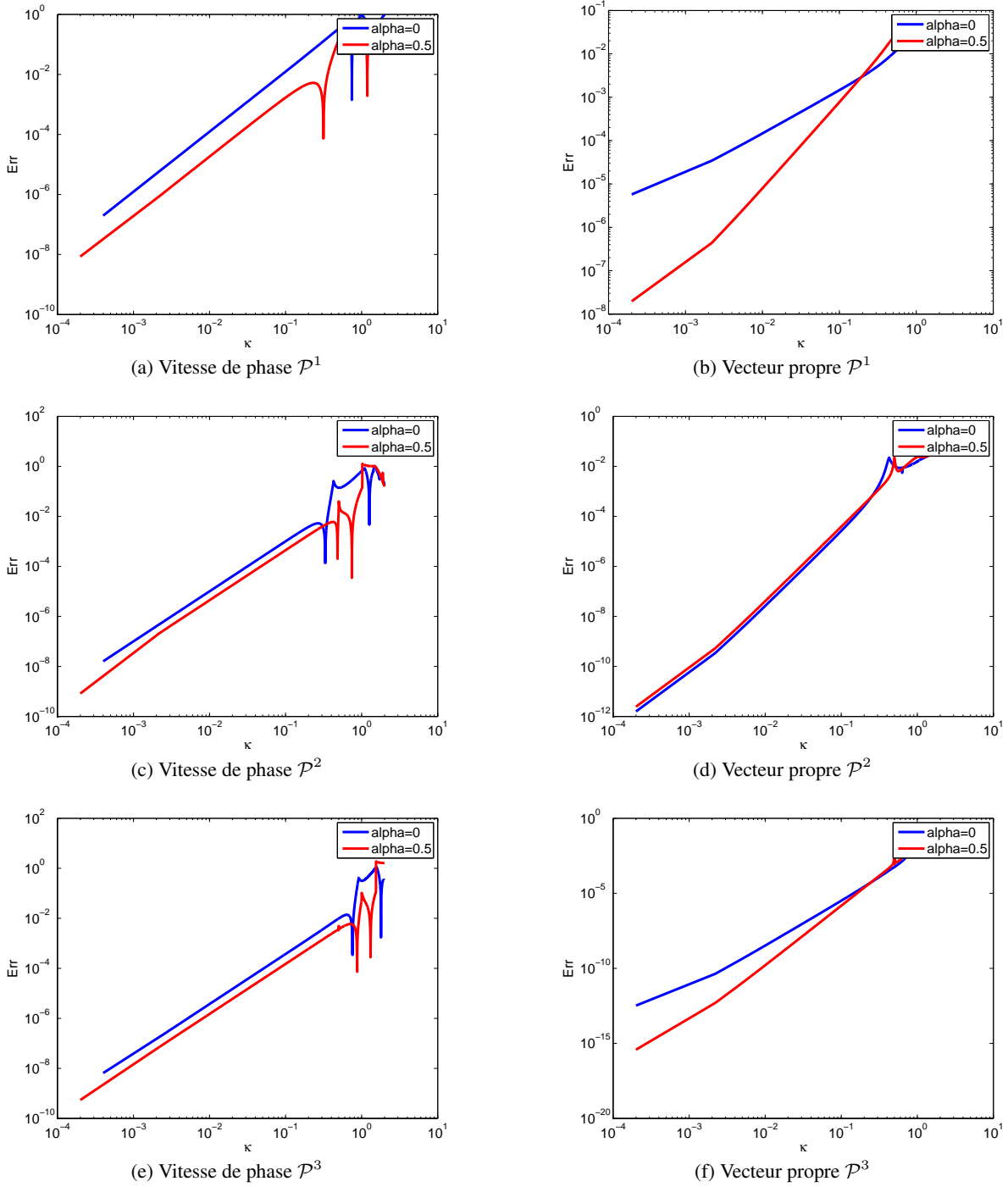


FIGURE 4.24 – Comparaison des courbes de convergence des schémas avec et sans pénalisation pour  $m = 1..3$

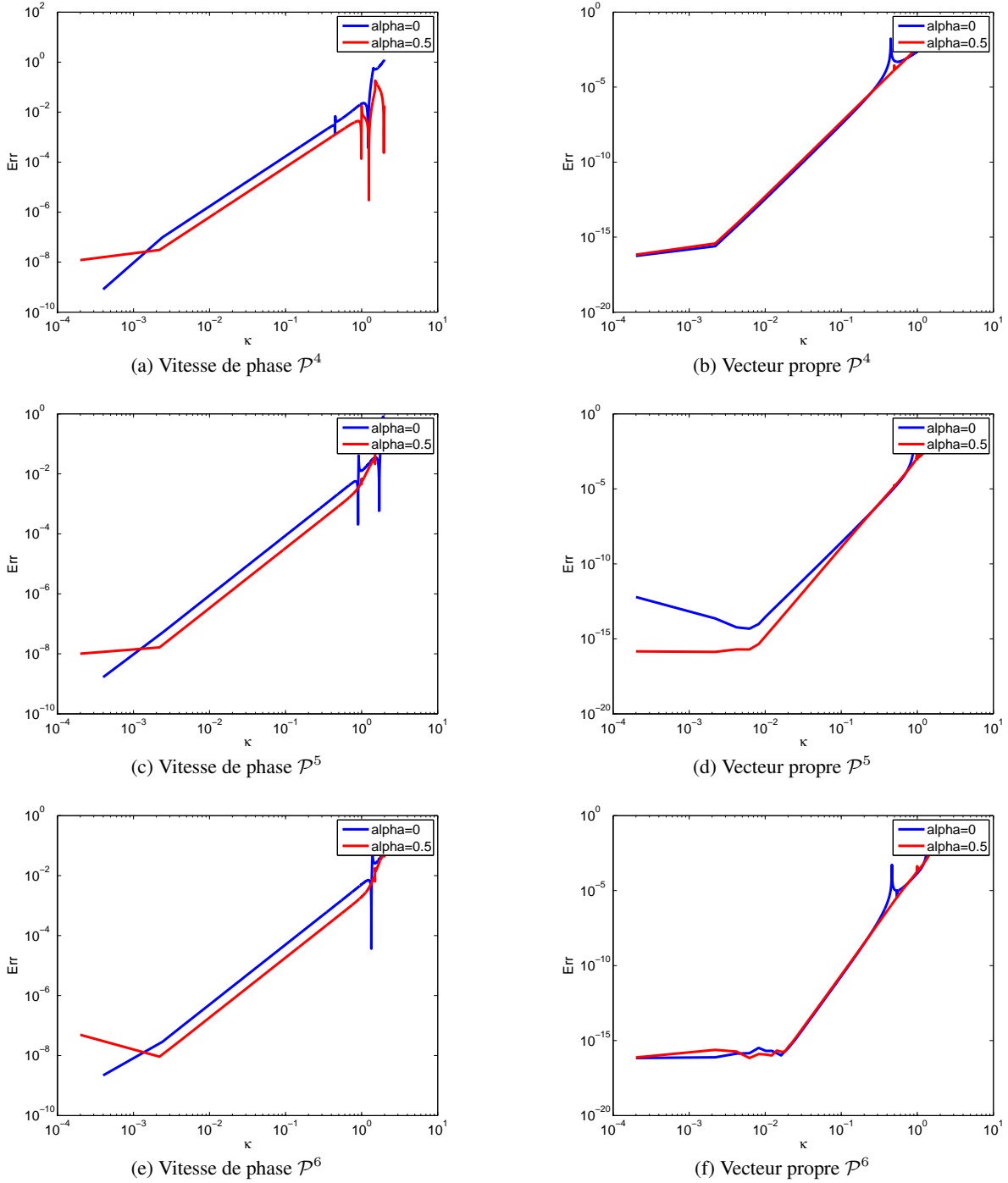


FIGURE 4.25 – Comparaison des courbes de convergence des schémas avec et sans pénalisation pour  $m = 4..6$

d'onde nécessaire pour obtenir une précision donnée sur le calcul de la vitesse de phase et du vecteur propre.

## 4.5 Conclusion

Ce chapitre avait pour objectif de comprendre pourquoi le schéma de discrétisation en espace que nous utilisons ne permettait pas d'obtenir la convergence optimale et générait des modes parasites. Pour cela, nous avons réalisé une analyse de dispersion qui nous a conduit à modifier la formulation DG. Nous avons introduit un paramètre de pénalisation en étendant à des maillages triangulaires la méthode proposée par Ainsworth, Monk et Muniz [4] pour des maillages cartésiens. Nous avons observé que les modes parasites étaient éliminés et que l'ordre de convergence était restauré. Toutefois, le paramètre affecte la condition CFL du schéma d'environ 30% et augmente donc les temps de calcul. Nous avons mené une analyse détaillée du nouveau schéma dans le cas 1D et les résultats doivent être confirmés à 2D avant d'envisager une mise en œuvre dans la plateforme DIVA. Dans le chapitre suivant, nous allons donc étudier l'effet de ce paramètre dans les configurations 2D du chapitre 3.

		10%	1%	0.1%	0.01%	$10^{-3}\%$	$10^{-4}\%$	$10^{-5}\%$
$\mathcal{P}^1$	$\alpha = 0$	3.6	11	35	110	330	1000	2500
	$\alpha = 0.5$	2.1	2.8	13	42	135	415	1250
$\mathcal{P}^2$	$\alpha = 0$	2.5	2.8	10	32	100	310	1000
	$\alpha = 0.5$	1	2	6.7	21	66	210	625
$\mathcal{P}^3$	$\alpha = 0$	1.2	2	6	20	60	190	550
	$\alpha = 0.5$	1	1.1	3.9	12	28	120	380
$\mathcal{P}^4$	$\alpha = 0$	0.76	2.2	4	13	40	120	380
	$\alpha = 0.5$	0.7	1	2.5	8	25	78	250
$\mathcal{P}^5$	$\alpha = 0$	0.5	1.1	3	9	30	90	300
	$\alpha = 0.5$	0.5	0.9	1.8	5.8	18	57	180
$\mathcal{P}^6$	$\alpha = 0$	<0.5	0.73	2.24	7	22	70	220
	$\alpha = 0.5$	<0.5	0.66	1.3	4.3	14	43	135

TABLE 4.7 – Nombre d'éléments par longueur d'onde nécessaire pour obtenir une précision donnée sur la vitesse de phase avec et sans pénalisation



		10%	1%	0.1%	0.01%	10 <sup>-3</sup> %	10 <sup>-4</sup> %	10 <sup>-5</sup> %
$\mathcal{P}^1$	$\alpha = 0$	<0.5	2	14	140	1250	>5000	>5000
	$\alpha = 0.5$	<0.5	2.9	9	27	86	260	830
$\mathcal{P}^2$	$\alpha = 0$	<0.5	2.5	3.5	6.5	14	29	63
	$\alpha = 0.5$	<0.5	2.1	3.4	7.3	15	34	73
$\mathcal{P}^3$	$\alpha = 0$	<0.5	1.2	1.7	3.4	7	15	31
	$\alpha = 0.5$	<0.5	2	2	3.5	6.2	11	20
$\mathcal{P}^4$	$\alpha = 0$	<0.5	2.2	2.3	2.6	3.5	5.2	8
	$\alpha = 0.5$	<0.5	1	1.3	2.1	3.4	5.4	8.6
$\mathcal{P}^5$	$\alpha = 0$	<0.5	1.1	1.2	1.4	2.1	3.2	4.9
	$\alpha = 0.5$	<0.5	1	1	1.4	2.1	3.2	4.8
$\mathcal{P}^6$	$\alpha = 0$	<0.5	0.7	0.8	2.2	2.3	2.6	3.2
	$\alpha = 0.5$	<0.5	0.7	0.8	1.1	1.6	2.2	3.1

TABLE 4.8 – Nombre d'éléments par longueur d'onde nécessaire pour obtenir une précision donnée sur le vecteur propre avec et sans pénalisation

		10%	1%	0.1%	0.01%	$10^{-3}\%$	$10^{-4}\%$	$10^{-5}\%$
$\mathcal{P}^1$	$\alpha = 0$	3.4	10	32	100	310	1000	2500
	$\alpha = 0.5$	3	5	9	16	28	50	90
$\mathcal{P}^2$	$\alpha = 0$	2.5	3	4	6	9	13	19
	$\alpha = 0.5$	1.3	2	3	6	9	13	20
$\mathcal{P}^3$	$\alpha = 0$	1.2	1.4	1.6	3	4	6	9
	$\alpha = 0.5$	0.8	1	2.1	2.9	4	5	7
$\mathcal{P}^4$	$\alpha = 0$	0.76	2.2	2.2	2.4	2.8	3	5
	$\alpha = 0.5$	0.7	1.1	1.3	1.6	2	2.6	3.4
$\mathcal{P}^5$	$\alpha = 0$	0.5	1.1	1.2	1.3	1.5	1.9	3
	$\alpha = 0.5$	0.7	0.8	0.9	1.1	1.5	2	2.6
$\mathcal{P}^6$	$\alpha = 0$	<0.5	0.73	0.8	0.9	2.1	2.2	2.6
	$\alpha = 0.5$	<0.5	0.5	0.7	1	1.2	1.4	1.7

TABLE 4.9 – Nombre d'éléments par longueur d'onde nécessaire pour obtenir une précision donnée sur la vitesse de phase, schéma  $\nabla$  avec et sans pénalisation

		10%	1%	0.1%	0.01%	$10^{-3}\%$	$10^{-4}\%$	$10^{-5}\%$
$\mathcal{P}^1$	$\alpha = 0$	3	9	30	100	300	800	2500
	$\alpha = 0.5$	2.2	4	6	11	19	34	60
$\mathcal{P}^2$	$\alpha = 0$	2.5	3	4	6	9	13	18
	$\alpha = 0.5$	1.1	2	2.6	4	5	8	12
$\mathcal{P}^3$	$\alpha = 0$	1.2	1.4	1.6	3	4	6	9
	$\alpha = 0.5$	0.7	1.2	2	2.1	2.7	4	5
$\mathcal{P}^4$	$\alpha = 0$	0.76	0.9	2.2	2.4	2.8	3	4
	$\alpha = 0.5$	0.7	1	1.1	1.3	2	2.1	2.8
$\mathcal{P}^5$	$\alpha = 0$	0.5	1.1	1.2	1.3	1.5	1.9	3
	$\alpha = 0.5$	<0.5	0.8	1.1	1.3	1.7	2.2	2.8
$\mathcal{P}^6$	$\alpha = 0$	<0.5	0.73	0.8	0.9	2.2	2.2	2.6
	$\alpha = 0.5$	<0.5	0.7	0.8	1	1.3	1.6	2

TABLE 4.10 – Nombre d'éléments par longueur d'onde nécessaire pour obtenir une précision donnée sur la vitesse de phase, schéma *ADEP* avec et sans pénalisation

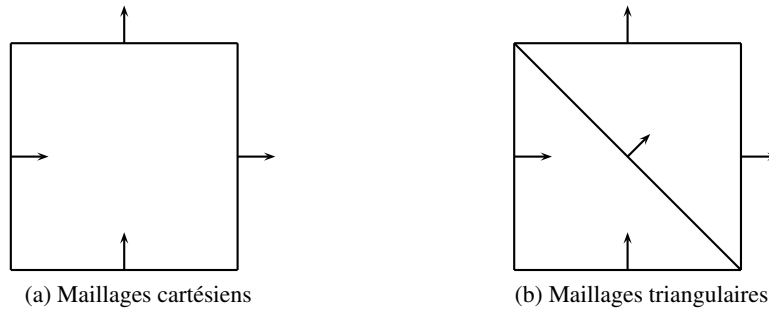
## Chapitre 5

# Étude de l'effet du terme de pénalisation en dimension deux

Nous allons maintenant reprendre les expériences 2D menées au chapitre 3, en ajoutant, cette fois, le terme de pénalisation. Ces expériences ont pour but de montrer l'intérêt de cette pénalisation dans une configuration plus réaliste alors que le cas 1D nous a servi à exhiber les modes parasites. Nous montrons que cette pénalisation permet non seulement de restaurer l'ordre de convergence des schémas mais aussi supprimer les ondes parasites (sections 5.1 et 5.2). Nous comparons également les performances des schémas écrits pour la formulation d'ordre un aux performances des schémas construits pour la formulation d'ordre deux. Nous finissons ce chapitre en illustrant (section 5.5) deux des difficultés que nous avons rencontrées en portant nos algorithmes dans la plate-forme DIVA. La première survient, semble-t-il, parce que les calculs sont réalisés en simple précision. Cette configuration est justifiée par les énormes coûts de calcul qui sont nécessaires pour effectuer des calculs en double précision. La seconde est posée par le choix du paramètre de pénalisation qui n'est pas évident dans le cas d'un maillage triangulaire non structuré. Il semblerait que la formulation soit très sensible à ce choix.

### 5.1 Cas test académique avec condition initiale

Dans l'analyse de dispersion 1D, le paramètre de pénalisation  $\alpha_\Gamma$  était un scalaire et pour passer au cas 2D, nous devons maintenant considérer un vecteur. A l'origine, Ainsworth, Monk et Muniz [4] ont considéré des maillages cartésiens et ont orienté le vecteur  $\alpha_\Gamma$  orthogonalement à chacune des arêtes (voir Fig. 5.1.(a)). Pour les arêtes horizontales, le vecteur était dirigé vers le haut et pour les arêtes verticales il était dirigé vers la droite. Nous avons suivi la même procédure, mais nous devons en plus définir le vecteur sur les arêtes correspondant à l'hypoténuse des triangles rectangles du maillage. Nous l'avons choisi dans un premier temps orthogonal à l'arête et dirigé vers le haut et vers la droite (voir Fig. 5.1.(b)). Dans ce cas, que nous appellerons Configuration 1, tous les vecteurs sont orientés vers le haut et la droite du domaine. Nous testerons par la suite une deuxième configuration, en orientant le vecteur  $\alpha_\Gamma$  sur l'hypoténuse vers le bas et vers la gauche. Dans cette configuration, tous les vecteurs sont dirigés vers l'intérieur du premier triangle et vers l'extérieur du second.

FIGURE 5.1 – Orientation du vecteur  $\alpha_T$ .

### 5.1.1 Cas test académique avec condition initiale

#### 5.1.1.1 Le schéma DG-Flux avec pénalisation

Nous reprenons ici l'expérience définie à la section 3.3.1 et nous définissons le vecteur de pénalisation de la configuration 1. Nous avons vu que le paramètre de pénalisation détériore la condition CFL et nous avons donc divisé par deux le pas de temps de chaque expérience pour garantir la stabilité des solutions. La table 5.1 est constituée de trois tableaux : (a) pour le schéma LF, (b) pour le schéma  $\nabla$ , (c) pour le schéma ADER. Chaque tableau contient les erreurs relatives  $L^2$  et les valeurs du pas de temps correspondantes pour un facteur de raffinement  $\delta$  donné ainsi que l'évolution de l'ordre de convergence entre deux raffinements. Tout cela est obtenu pour une approximation de type  $\mathcal{P}^3$ . Les tables 5.2 et 5.3 représentent les mêmes quantités pour des approximations de type  $\mathcal{P}^4$  et  $\mathcal{P}^5$  respectivement.

$\delta$	dt	Erreur	CV
1	9.6e-04	3.9e-02	
2	4.8e-04	9.7e-03	2.0
4	2.4e-04	2.4e-03	2.0
8	1.2e-04	6.1e-04	2.0
16	6.0e-05	1.5e-04	2.0

(a) LF

dt	Erreur	CV
9.6e-04	8.2e-05	
4.8e-04	5.2e-06	4.0
2.4e-04	3.2e-07	4.0
1.2e-04	2.0e-08	4.0
6.0e-05	1.3e-09	4.0

(b)  $\nabla$

dt	Erreur	CV
2.7e-03	8.6e-05	
1.3e-03	5.4e-06	4.0
6.7e-04	3.4e-07	4.0
3.4e-04	2.1e-08	4.0
1.7e-04	1.3e-09	4.0

(c) ADER

TABLE 5.1 – Erreur et taux de convergence en fonction du taux de raffinement  $\delta$  avec pénalisation et pour des éléments finis  $\mathcal{P}^3$ .

Nous constatons que la convergence est d'ordre quatre, quel que soit le degré des éléments finis utilisés. Nous représentons l'erreur en fonction des coûts de calcul pour des éléments  $\mathcal{P}_3$  (resp.  $\mathcal{P}_4$  et  $\mathcal{P}_5$ ) sur la figure 5.2.(a) (resp. 5.3.(a) et 5.4.(a)). On peut observer qu'à chaque fois, pour un nombre d'opérations fixé, les erreurs du schéma  $\nabla$  sont inférieures aux erreurs des deux autres schémas. La convergence des schémas est illustrée sur les figures 5.2.(b), 5.3.(b) et 5.4.(b). En abscisse, on représente les valeurs du pas de discrétisation  $h$ . On peut observer que l'ordre de 4 de convergence est bien atteint, grâce au paramètre de pénalisation.

$\delta$	dt	Erreur	CV
1	6.6e-04	1.8e-02	
2	3.3e-04	4.6e-03	2.0
4	1.7e-04	1.1e-03	2.0
8	8.3e-05	2.4e-04	2.3
16	4.1e-05	5.9e-05	2.0

(a) LF

dt	Erreur	CV
6.6e-04	2.3e-06	
3.3e-04	7.1e-08	5.0
1.7e-04	2.3e-09	4.9
8.3e-05	7.1e-11	5.0
4.1e-05	5.6e-12	3.7

(b)  $\nabla$

dt	Erreur	CV
1.8e-03	3.1e-06	
9.2e-04	1.4e-07	4.5
4.6e-04	7.4e-09	4.2
2.3e-04	4.4e-10	4.1
1.2e-04	1.4e-11	4.9

(c) ADER

TABLE 5.2 – Erreur et taux de convergence en fonction du taux de raffinement  $\delta$  avec pénalisation et pour des éléments finis  $\mathcal{P}^4$ .

$\delta$	dt	Erreur	CV
1	4.8e-04	9.7e-03	
2	2.4e-04	2.4e-03	2.0
4	1.2e-04	6.1e-04	2.0
8	6.0e-05	1.5e-04	2.0
16	3.0e-05	1.2e-05	3.6

(a) LF

dt	Erreur	CV
4.8e-04	9.3e-08	
2.4e-04	3.6e-09	4.7
1.2e-04	2.2e-10	4.1
6.0e-05	1.4e-11	4.0
3.0e-05	3.0e-13	5.5

(b)  $\nabla$

dt	Erreur	CV
1.3e-03	5.0e-07	
6.7e-04	3.1e-08	4.0
3.4e-04	1.9e-09	4.0
1.7e-04	1.2e-10	4.0
8.4e-05	1.6e-12	6.2

(c) ADER

TABLE 5.3 – Erreur et taux de convergence en fonction du taux de raffinement  $\delta$  avec pénalisation et pour des éléments finis  $\mathcal{P}^5$ .

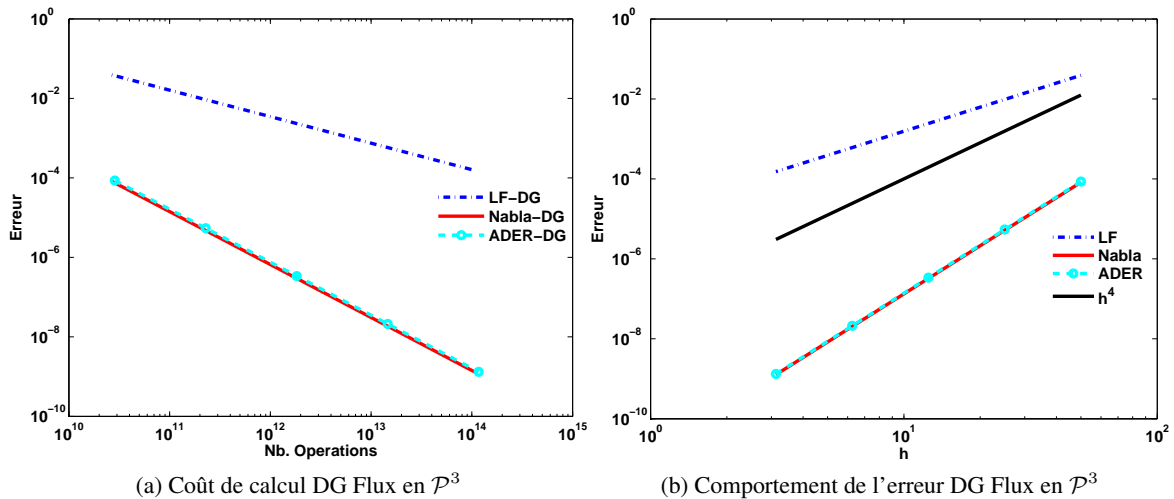


FIGURE 5.2 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degrés 3.

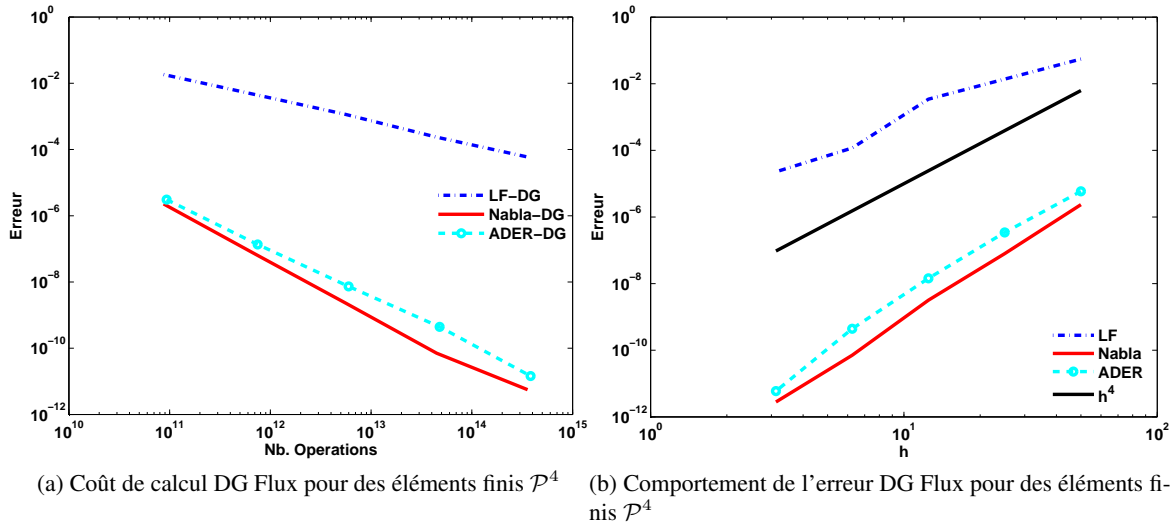


FIGURE 5.3 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 4.

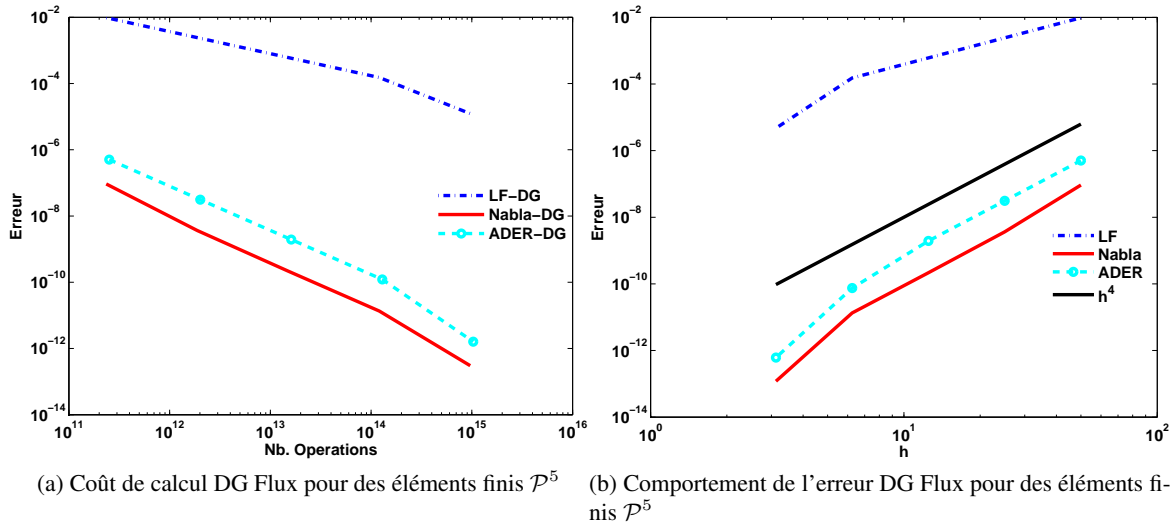


FIGURE 5.4 – Courbes d'erreur et coût de calcul, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  et pour des éléments finis de degré 5.

## 5.2 Cas de la source ponctuelle

### 5.2.1 Le schéma DG-Flux avec pénalisation

Nous étudions maintenant le cas de la source ponctuelle, et nous définissons le paramètre de pénalisation de la configuration 1. Les tableaux 5.4, 5.5 et 5.6 représentent les erreurs relatives  $L^2$  obtenues au récepteur. Comme nous l'espérons, l'ajout du terme de pénalisation permet d'obtenir la convergence des schémas. Cependant, cette convergence reste sous-optimale pour des éléments  $\mathcal{P}^3$ . Dans ce cas, le schéma LF converge uniquement à l'ordre un. Les schémas ADER et  $\nabla$  apportent une amélioration notable car ils convergent à l'ordre 3 pour des maillages grossiers et à l'ordre 4 pour des maillages fins. Pour des éléments  $\mathcal{P}^4$  et  $\mathcal{P}^5$ , nous obtenons bien la convergence attendue, quel que soit le schéma.

$\delta$	dt	Erreur	CV
1	9.6e-04	3.5e-02	
2	4.8e-04	7.0e-02	-1.0
4	2.4e-04	3.6e-02	1.0
8	1.2e-04	1.8e-02	1.0
16	6.0e-05	9.2e-03	1.0

(a) LF

dt	Erreur	CV
9.6e-04	1.3e-01	
4.8e-04	1.2e-02	3.5
2.4e-04	1.0e-03	3.5
1.2e-04	7.0e-05	3.9
6.0e-05	4.5e-06	4.0

(b)  $\nabla$

dt	Erreur	CV
2.7e-03	1.4e-01	
1.3e-03	1.2e-02	3.6
6.7e-04	1.0e-03	3.6
3.4e-04	6.7e-05	3.9
1.7e-04	4.3e-06	4.0

(c) ADER

TABLE 5.4 – Erreur et ordre de convergence en fonction du pas de maillage  $h$  pour des éléments fins  $\mathcal{P}^3$  avec pénalisation. Les schémas convergent mais nous n'obtenons pas la convergence d'ordre quatre.

$\delta$	dt	Erreur	CV
1	6.6e-04	1.4e-02	
2	3.3e-04	2.1e-03	2.7
4	1.7e-04	5.2e-04	2.0
8	8.3e-05	1.3e-04	2.0
16	4.1e-05	3.3e-05	2.0

(a) LF

dt	Erreur	CV
6.6e-04	5.3e-02	
3.3e-04	1.4e-03	5.3
1.7e-04	2.3e-05	5.9
8.3e-05	3.6e-07	6.0
4.1e-05	5.7e-09	6.0

(b)  $\nabla$

dt	Erreur	CV
1.8e-03	5.1e-02	
9.2e-04	1.4e-03	5.2
4.6e-04	2.6e-05	5.8
2.3e-04	4.2e-07	5.9
1.2e-04	7.1e-09	5.9

(c) ADER

TABLE 5.5 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments fins  $\mathcal{P}^4$  avec pénalisation. La convergence est bien d'ordre quatre (et même supérieure) pour les schémas  $\nabla$  et ADER.



$\delta$	dt	Erreur	CV
1	9.6e-04	4.4e-03	
2	2.4e-04	1.1e-03	2.0
4	1.2e-04	2.8e-04	2.0
8	6.0e-05	6.9e-05	2.0
16	3.0e-05	1.7e-05	2.0

(a) LF

dt	Erreur	CV
4.8e-04	3.5e-03	
2.4e-04	1.0e-04	5.1
1.2e-04	2.6e-06	5.3
6.0e-05	4.7e-08	5.8
3.0e-05	7.5e-10	6.0

(b)  $\nabla$

dt	Erreur	CV
1.3e-03	3.8e-03	
6.7e-04	1.1e-04	5.1
3.4e-04	2.6e-06	5.4
1.7e-04	4.5e-08	5.9
8.4e-05	9.0e-10	5.6

(c) ADER

TABLE 5.6 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^5$  avec pénalisation. La convergence est bien d'ordre quatre (et même supérieure) pour les schémas  $\nabla$  et ADER.

### 5.2.1.1 Le schéma IPDG

Après l'introduction du paramètre de pénalisation on doit diviser la CFL par deux et les coûts de calculs sont alors multipliés par deux. On se demande alors si le schéma DG-flux avec pénalisation est compétitif par rapport à d'autres schémas de discrétisation en espace. Nous avons donc également testé les schémas LF, ADER et  $\nabla$  basés sur une discrétisation en espace IPDG pour la formulation de l'équation des ondes du second ordre. Les erreurs obtenues sont présentées dans les tableaux 5.7, 5.8 et 5.9. Comme pour les schémas DG-Flux, la convergence n'est pas d'ordre quatre pour ADER-IPDG et  $\nabla$ -IPDG avec des éléments  $\mathcal{P}^3$  sur des maillages grossiers, mais elle s'en rapproche sur des maillages plus fins. Par contre, le schéma LF-IPDG converge bien à l'ordre deux alors que le schéma LF-DG flux ne convergeait qu'à l'ordre 1.

$\delta$	dt	Erreur	CV
1	1.4e-03	5.1e-02	
2	6.8e-04	9.4e-03	2.4
4	3.4e-04	2.2e-03	2.1
8	1.7e-04	5.5e-04	2.0
16	8.5e-05	1.4e-04	2.0

(a) LF

dt	Erreur	CV
1.4e-03	2.5e-02	
6.8e-04	3.0e-03	3.0
3.4e-04	3.7e-04	3.0
1.7e-04	2.9e-05	3.7
8.5e-05	1.9e-06	3.9

(b)  $\nabla$

dt	Erreur	CV
2.3e-03	2.9e-02	
1.2e-03	3.6e-03	3.0
5.8e-04	3.8e-04	3.2
2.9e-04	2.8e-05	3.8
1.5e-04	1.8e-06	3.9

(c) ADER

TABLE 5.7 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^3$ , schéma IPDG.

Ces différents tableaux ne font qu'illustrer les ordres de convergence des différentes méthodes. Pour un maillage donné, chaque schéma a un pas de temps qui lui est propre et donne une précision différente. Comme pour les schémas DG-Flux, nous allons déterminer les coûts de calcul d'une itération des schémas IPDG. Pour les schémas LF et  $\nabla$ , IPDG ne requiert qu'une multiplication par la matrice de raideur qui est de taille  $N_\phi^2 N \times N_\phi^2 N$  et contient  $4N \times N_\phi^2$  éléments non nuls. Rappelons que le nombre de degrés de liberté par élément  $N_\phi$  est le même pour les schémas IPDG et DG-Flux.

$\delta$	dt	Erreur	CV
1	9.6e-04	2.4e-02	
2	4.8e-04	4.4e-03	2.5
4	2.4e-04	1.1e-03	2.0
8	1.2e-04	2.8e-04	2.0
16	6.0e-05	6.9e-05	2.0

(a) LF

dt	Erreur	CV
9.6e-04	1.6e-02	
4.8e-04	1.3e-04	7.0
2.4e-04	3.2e-06	5.3
1.2e-04	6.1e-08	5.7
6.0e-05	8.9e-10	6.1

(b)  $\nabla$

dt	Erreur	CV
1.7e-03	1.7e-02	
8.3e-04	2.5e-04	6.1
4.1e-04	2.5e-06	6.6
2.1e-04	8.0e-08	5.0
1.0e-04	4.1e-09	4.3

(c) ADER

TABLE 5.8 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^4$ , schéma IPDG.

$\delta$	dt	Erreur	CV
1	1.4e-03	9.6e-03	
2	3.6e-04	2.4e-03	2.0
4	1.8e-04	6.1e-04	2.0
8	8.9e-05	1.5e-04	2.0
16	4.4e-05	3.8e-05	2.0

(a) LF

dt	Erreur	CV
7.1e-04	2.4e-03	
3.6e-04	2.4e-05	6.7
1.8e-04	3.2e-07	6.2
8.9e-05	1.1e-08	4.8
4.4e-05	2.4e-10	5.5

(b)  $\nabla$

dt	Erreur	CV
1.2e-03	1.1e-03	
6.1e-04	1.3e-05	6.4
3.0e-04	6.5e-07	4.3
1.5e-04	2.1e-08	5.0
7.6e-05	1.1e-09	4.3

(c) ADER

TABLE 5.9 – Erreur et taux de convergence en fonction du pas de maillage  $h$  pour des éléments finis  $\mathcal{P}^5$ , schéma IPDG.

Le coût d'une itération IPDG est donc  $4NN_{\Phi}^2$  alors que celui d'une itération DG-Flux est de  $16NN_{\Phi}^2$ , donc quatre fois plus cher.

En ce qui concerne les schémas ADER, les pas de temps et la précision sont sensiblement les mêmes pour IPDG et DG-Flux. Par contre, en appliquant une formulation IPDG, nous ne faisons que deux multiplications par la matrice de raideur, le coût d'une itération IPDG-ADER est donc de  $8NN_{\Phi}^2$  contre  $48NN_{\Phi}^2$ .

Nous comparons les coûts de calcul des différentes méthodes sur la figure 5.5. Nous remarquons

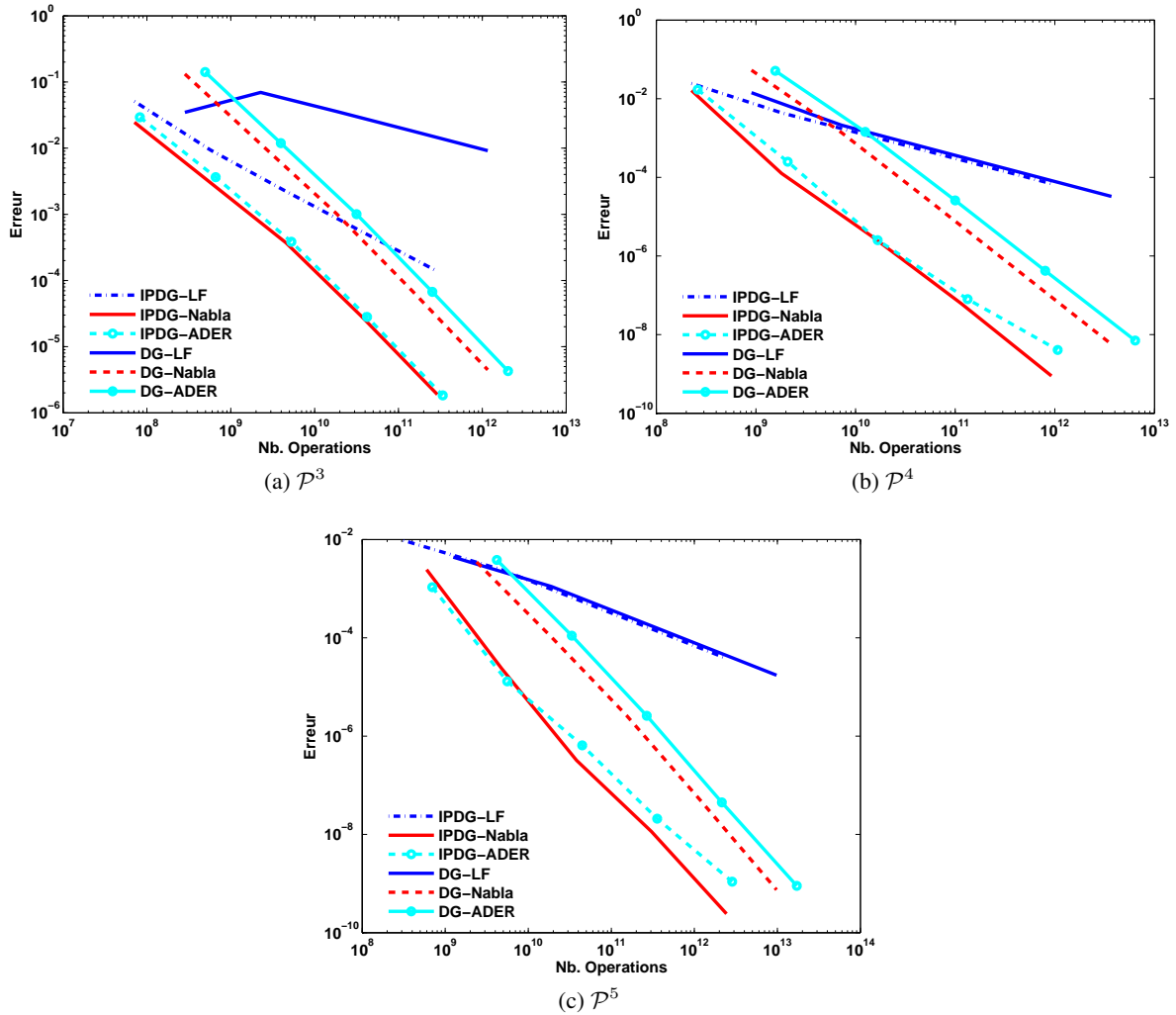


FIGURE 5.5 – Comparaisons des coûts de calcul entre les méthode DG-Flux et IPDG.

que :

- (i) Les schémas IDPG-LF et DG-Flux-LF ont des coûts de calcul similaires lorsqu'on les utilise avec des éléments  $\mathcal{P}^4$  ou  $\mathcal{P}^5$ . Avec des éléments  $\mathcal{P}^3$ , le schéma IPDG-LF montre de bien meilleures performances, avec des coûts de calcul 100 à 1000 moins élevés.
- (ii) En ce qui concerne les schémas d'ordre élevé en temps, IPDG-ADER et IPDG- $\nabla$  présentent des performances similaires et bien supérieures à celles des schémas DG-Flux-ADER et DG-Flux- $\nabla$ . On observe un facteur 10 entre les coûts de calcul des schémas de type IPDG et ceux des

schémas de type DG-Flux.

L'ajout du paramètre de pénalisation permet donc d'obtenir une meilleure précision, mais au prix d'une augmentation des coûts de calcul, de telle sorte que les schémas IPDG semblent nettement plus avantageux. Nous pouvons également rajouter que les schémas DG-Flux nécessitent de stocker plus d'inconnues que les schémas IPDG. Dans un code parallèle, il faut également communiquer les valeurs des inconnues à chaque demi pas de temps en DG-Flux, alors qu'il suffit de le faire à chaque pas de temps en IPDG. La formulation IPDG semble donc plus intéressante à tout point de vue.

### 5.3 Validation sur benchmark : le modèle Marmousi

Nous reprenons maintenant le cas-test Marmousi pour illustrer l'effet du paramètre de pénalisation. Les sismogrammes obtenus sont représentés sur les figures 5.6, 5.7 et 5.8.

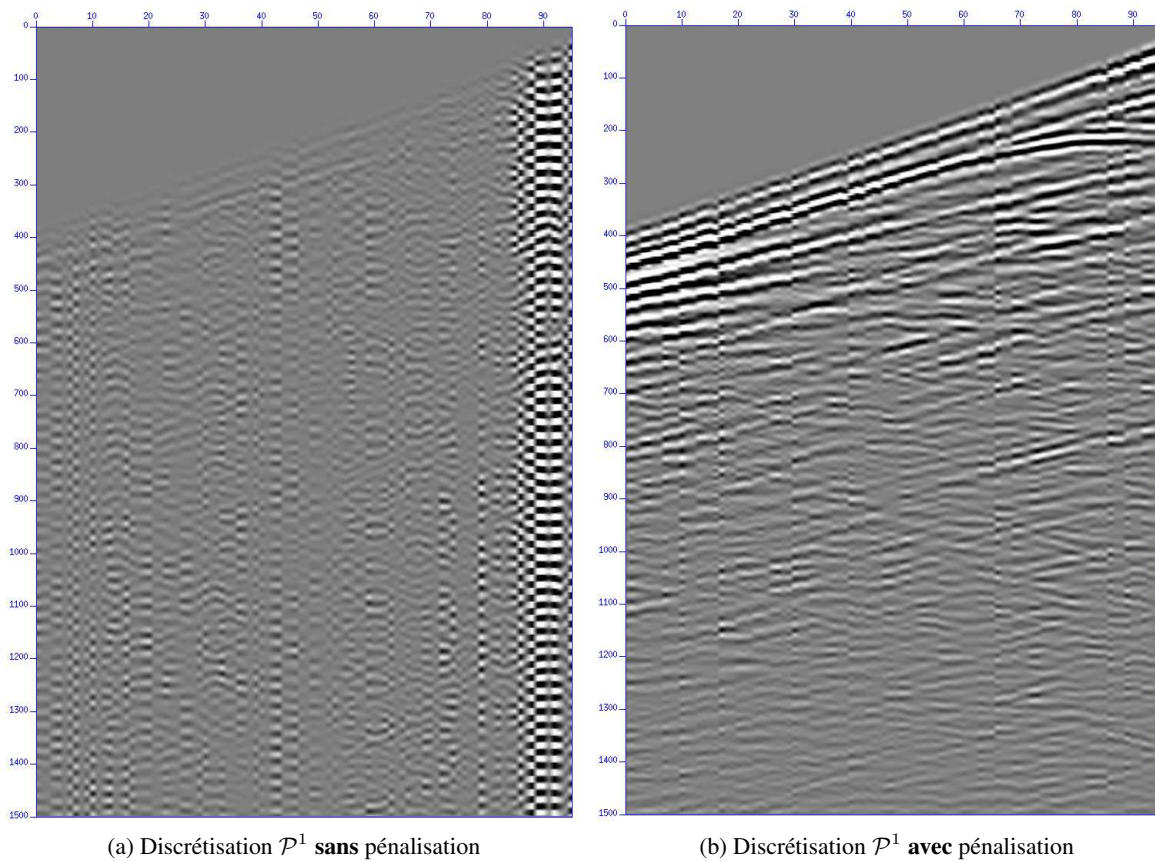


FIGURE 5.6 – Comparaison des sismogrammes obtenus sur le cas test Marmousi avec et sans pénalisation pour des éléments  $\mathcal{P}^1$ .

### 5.4 Application à l'imagerie sismique

Dans le chapitre précédent, nous avons utilisé, à des fins de validation, nos algorithmes pour la résolution du problème direct (phase de modelling). L'objectif final de cette thèse est de contribuer à la

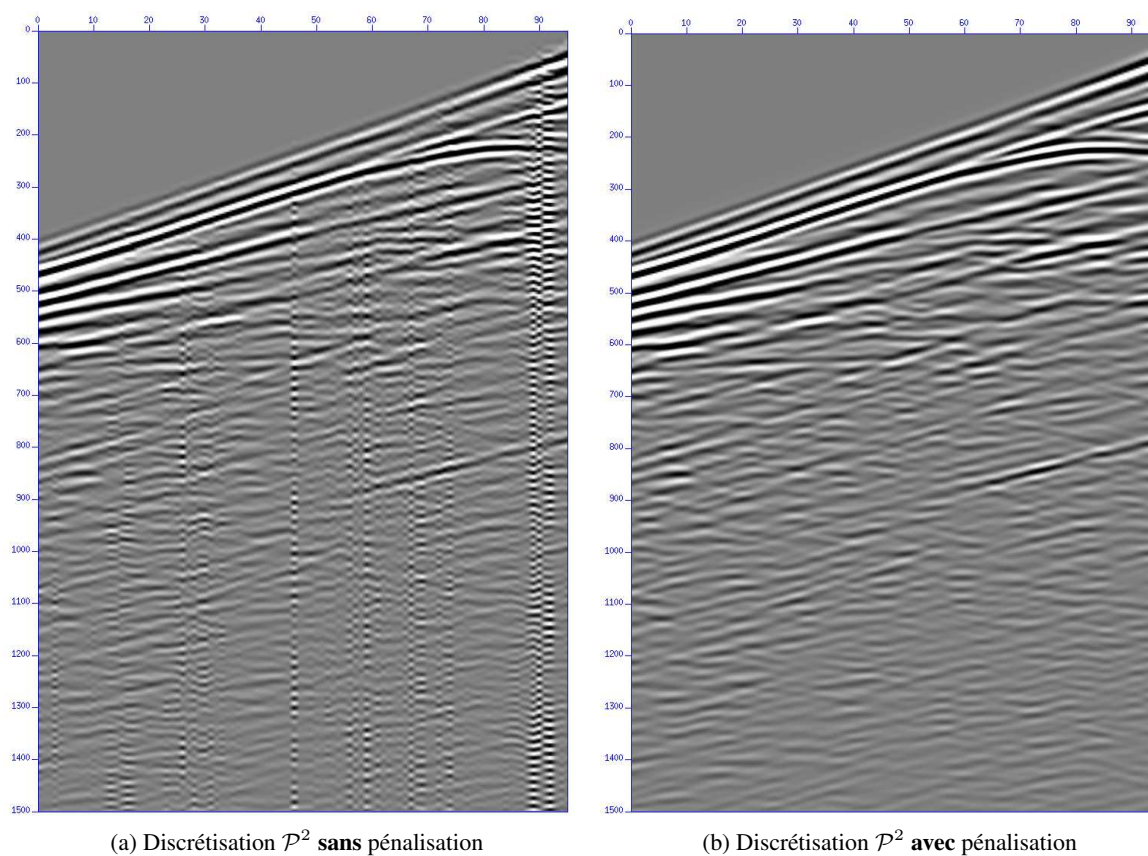


FIGURE 5.7 – Comparaison des sismogrammes obtenus sur le cas test Marmousi avec et sans pénalisation pour des éléments  $\mathcal{P}^2$ .

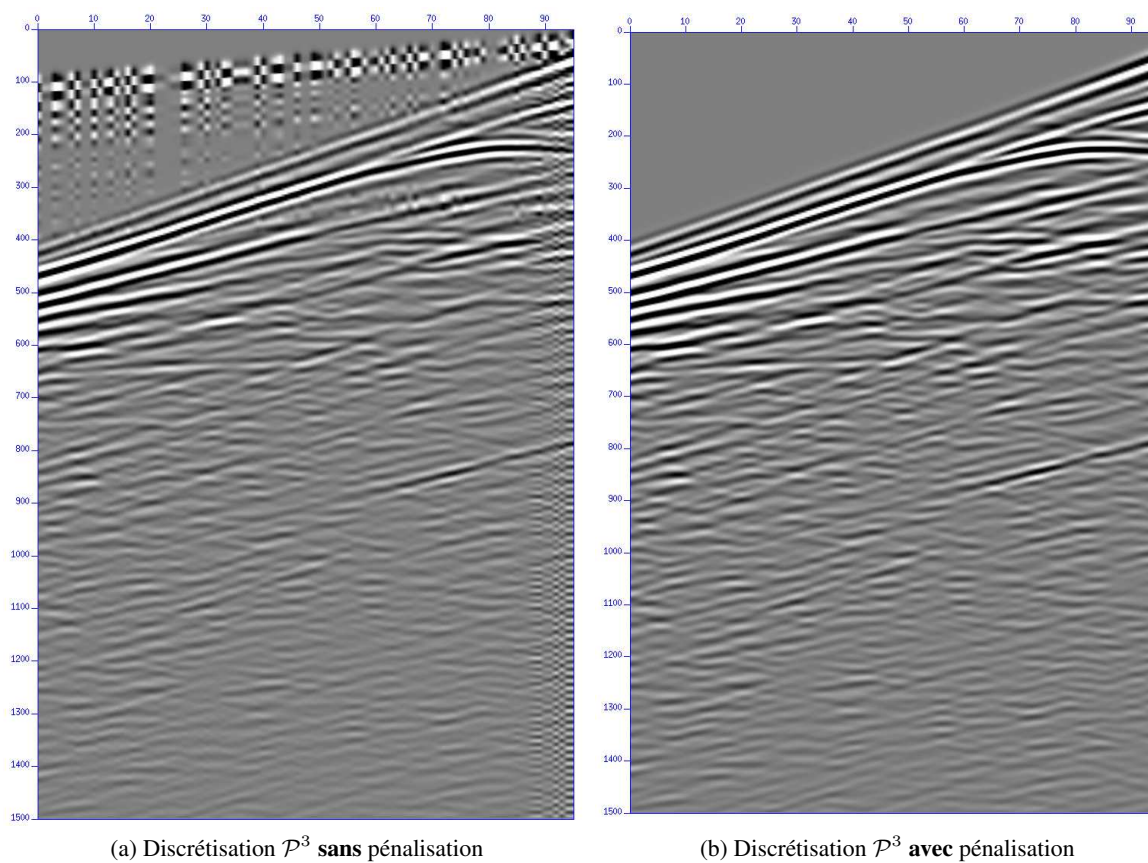
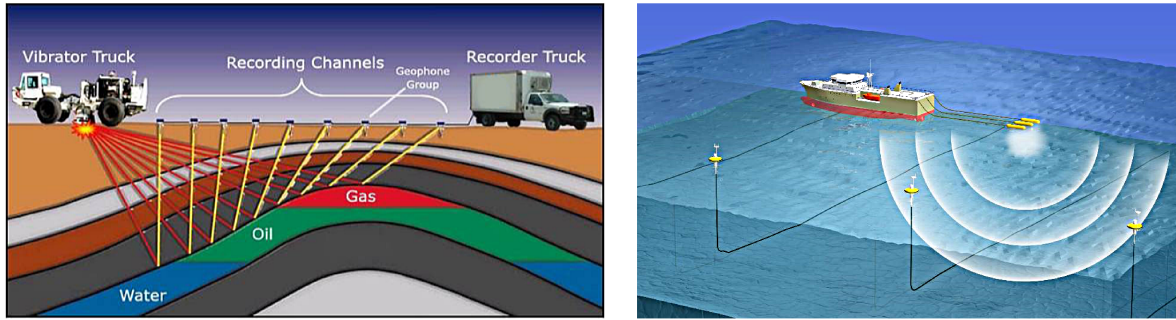


FIGURE 5.8 – Comparaison des sismogrammes obtenus sur le cas test Marmousi avec et sans pénalisation pour des éléments  $\mathcal{P}^3$ .



Acquisition Terrestre

Acquisition Marine

FIGURE 5.9 – Campagnes d’acquisitions sismiques Terrestre et Marine.

production d’images du sous-sol. Pour produire ces images, nous utilisons la Reverse Time Migration (RTM) qui repose sur la résolution d’équations d’ondes complètes. Nous avons donc implémenté et intégré au sein de la plateforme Total DIVA (Depth Imaging Velocity Analysis) les trois schémas LF,  $\nabla$  et ADER. Dans toute la suite, nous ne considérons que des éléments- $\mathcal{P}^3$  pour la discrétisation en espace et les schémas en temps d’ordre 4  $\nabla$ , ADER et d’ordre 2 LF. Nous commençons par rappeler le principe de la RTM.

#### 5.4.1 Présentation de la Reverse Time Migration

Nous présentons de façon brève le principe de la Reverse Time Migration (RTM) et pour de plus amples informations sur cette méthode, le lecteur pourra, par exemple, consulter l’ouvrage de Baysal *et al.* [10].

La migration par équation d’onde ou RTM se décompose en trois étapes, la seconde pouvant être itérée plusieurs fois :

1. Récupération du modèle de vitesse initial grâce à la campagne d’acquisition (voir figure 5.9).

- |  |  |
|--|--|
| <ol style="list-style-type: none"> <li>2. On résout l’équation <b>deux fois</b> pour chacune des sources,</li> </ol> | <ol style="list-style-type: none"> <li>(a) <b>Propager</b> les sources dans le modèle de vitesse</li> <li>(b) <b>Rétro-propager</b> le champ d’ondes enregistré aux récepteurs (cette phase se fonde sur la réversibilité en temps de l’équation des ondes)</li> </ol> |
|--|--|

3. Cross-corrélation de (a) et (b) : Condition d’imagerie

4. On compare au modèle initial et si il existe des points de discordance, retour à (2) après modification du modèle de vitesse.

#### 5.4.2 Condition d’imagerie et stratégie de stockage

Le principe d’imagerie que nous avons utilisé suit la méthodologie de Claerbout [17]. Il utilise une condition d’imagerie de la forme

$$(5.4.1) \quad I(\mathbf{x}) = \int_0^T \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{d}(\mathbf{x}, t) dt$$

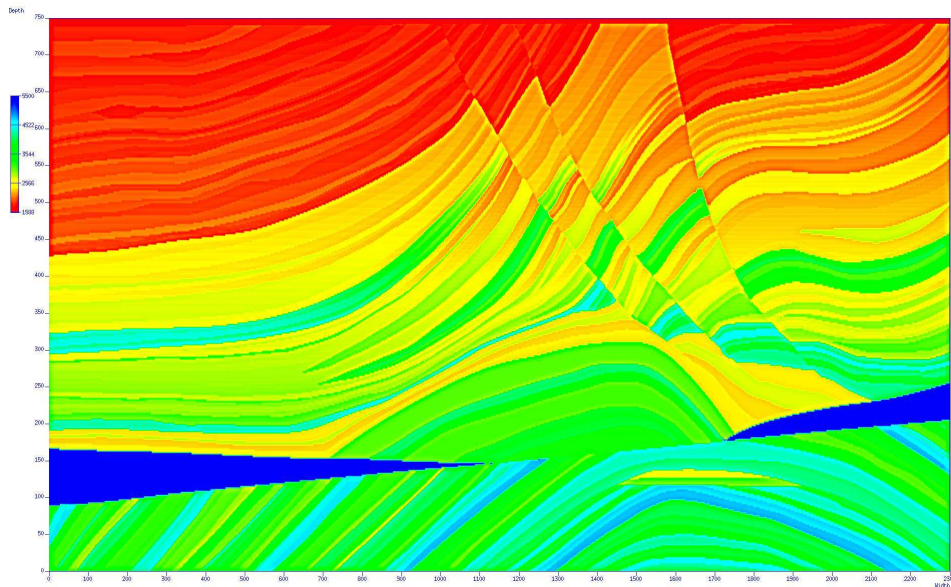
où  $\mathbf{u}(\mathbf{x}, t)$  désigne le champ d'ondes propagé par les sources et  $\mathbf{d}(\mathbf{x}, t)$  le champ d'ondes rétro-propagé par les récepteurs. La quantité  $T$  désigne le temps total de simulation. Afin d'obtenir des images plus réalistes, nous allons appliquer à l'image deux filtres. Le premier est appelé principe de Claerbout normalisé ou illumination, i.e.

$$(5.4.2) \quad I(\mathbf{x}) = \frac{1}{\|\mathbf{u}(\mathbf{x}, \cdot)\|_{L^2(0,T)}^2} \int_0^T \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{d}(\mathbf{x}, t) dt$$

L'autre condition que nous avons utilisée est un filtre passe-haut développé au sein de l'utilitaire "CWP/SU :Seismic Unix" ([http://www.seismicunix.com/w/Seismic\\_Unix](http://www.seismicunix.com/w/Seismic_Unix)). Nous avons également combiné les deux filtres précédents afin d'obtenir une image plus nette (en limitant les effets basse fréquence).

### 5.4.3 Caractéristiques du modèle

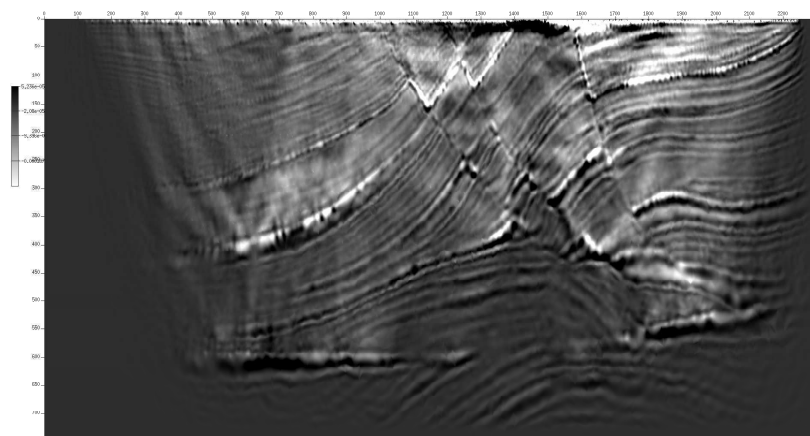
Nous avons de nouveau utilisé le modèle Marmousi présenté au chapitre 3. Nous rappelons qu'il contient 158 couches horizontales. Le domaine est large de 9.5 km et profond de 3 km. Le modèle de vitesse contient 751 points de données dans la direction Z et 2301 dans la direction X. Les vitesses varient de  $1500 \text{ m.s}^{-1}$  à  $5500 \text{ m.s}^{-1}$ . Nous considérons 240 sources de 12 Hz qui sont espacées de 25m. Nous propageons et enregistrons sur 3 sec. Nous considérons un maillage de 21469 éléments



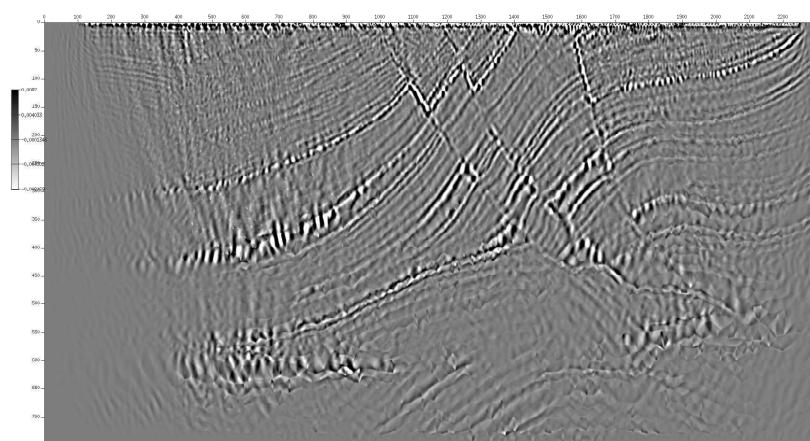
de type  $\mathcal{P}^3$ . Nous avons utilisé 8 noeuds de calcul de la plateforme Plafrim (<https://plafrim.bordeaux.inria.fr>). Chaque noeud est composé de deux Quad Core Intel Xeon E5420, avec une fréquence de 2.5GHz et dispose de 32Go de RAM (64 noeuds de calcul). Les noeuds sont reliés entre eux par un lien Infiniband DDR de 20Gb/s.

Si on applique le schéma LF, le temps de calcul est de 46080 s. En appliquant le schéma  $\nabla$ , les calculs durent 48960 s et 34560 s si on utilise le schéma ADER. Les résultats obtenus avec le schéma LF sont présentés sur la figure 5.10. Les résultats obtenus avec le schéma ADER sont présentés sur la figure 5.11 et les résultats obtenus avec le schéma  $\nabla$  sont présentés sur la figure 5.12.

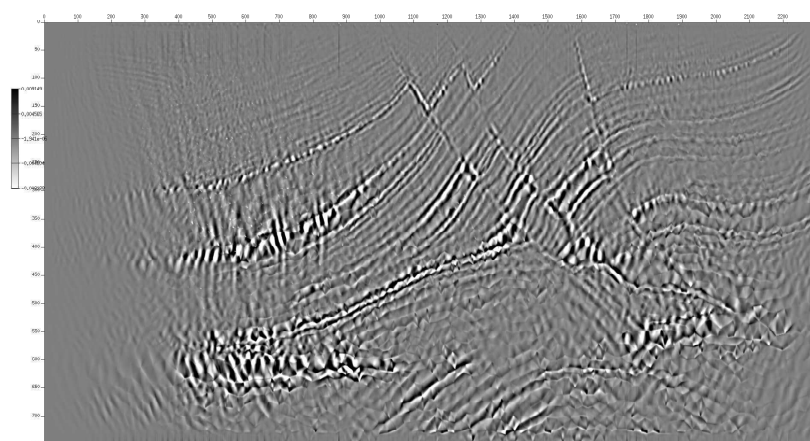




(a) Image RTM brute

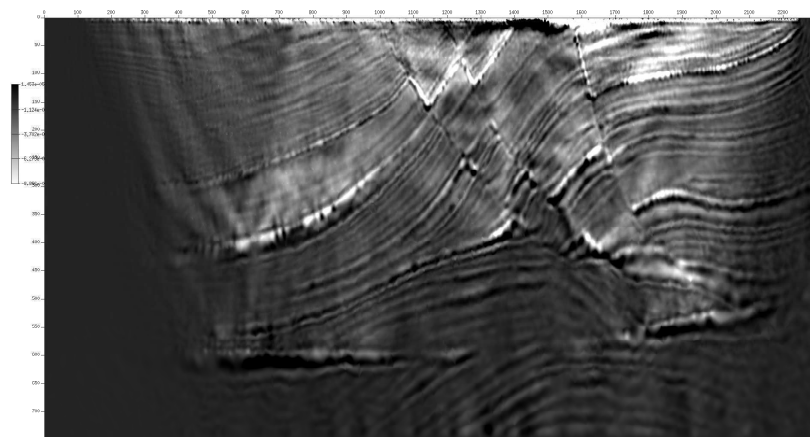


(b) Image RTM avec illumination

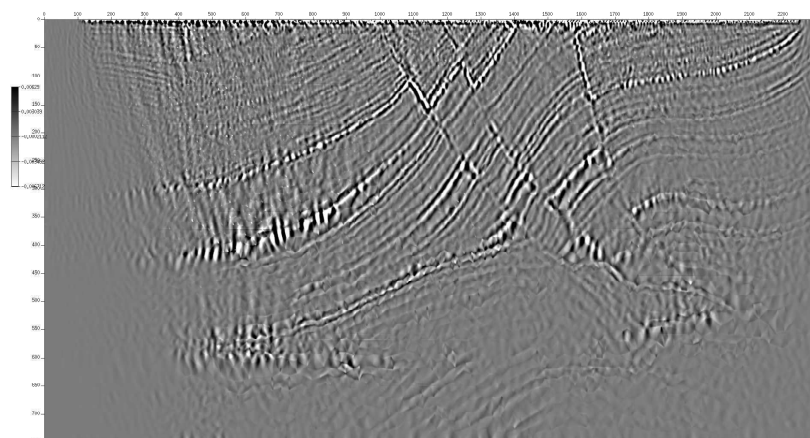


(c) Image RTM avec illumination et filtre passe-haut

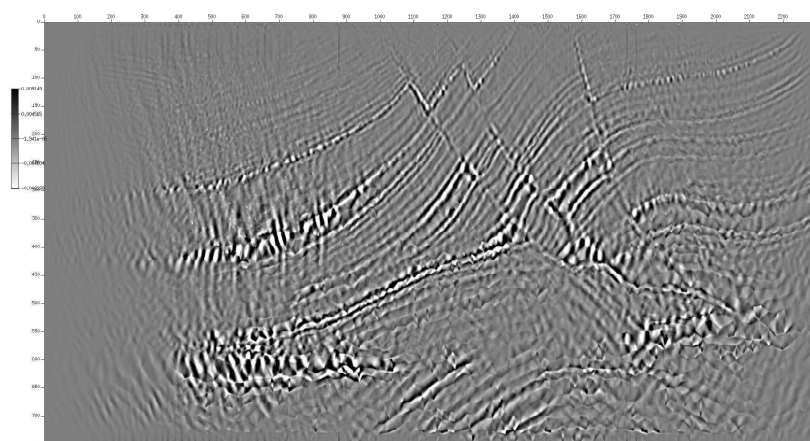
FIGURE 5.10 – Résultats RTM pour le schéma LF avec image sans traitement (a), avec illumination (b), avec illumination et filtre passe-haut (c).



(a) Image RTM brute

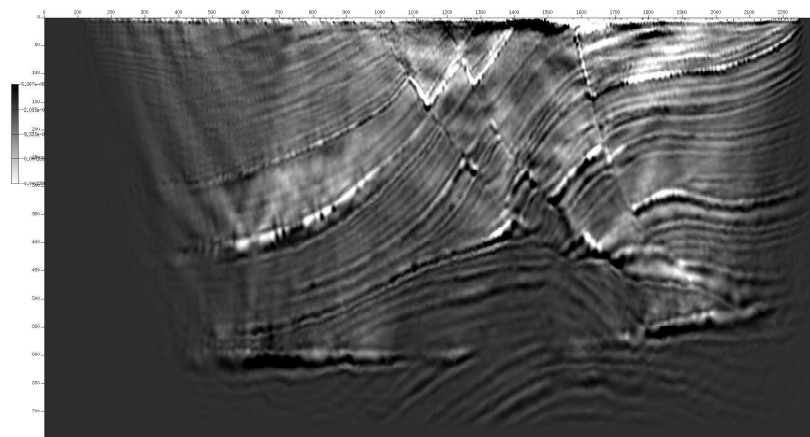


(b) Image RTM avec illumination

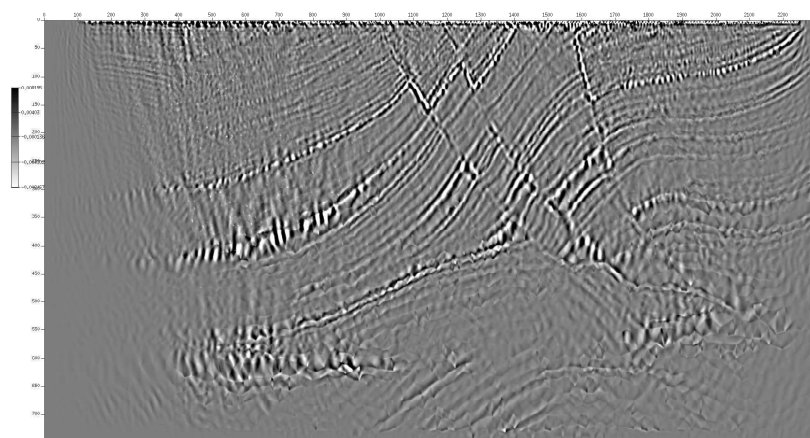


(c) Image RTM avec illumination et filtre passe-haut

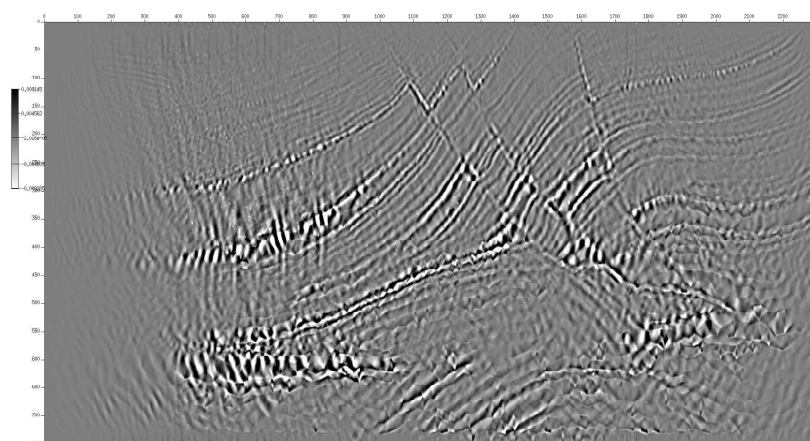
FIGURE 5.11 – Résultats RTM pour le schéma ADER avec image sans traitement (a), avec illumination (b), avec illumination et filtre passe-haut (c).



(a) Image RTM brute



(b) Image RTM avec illumination



(c) Image RTM avec illumination et filtre passe-haut

FIGURE 5.12 – Résultats RTM pour le schéma  $\nabla$  avec image sans traitement (a), avec illumination (b), avec illumination et filtre passe-haut (c).

Les résultats obtenus avec les trois schémas sont très similaires, ce qui est assez décevant car nous espérions une amélioration des images en utilisant des schémas d'ordre élevé en temps. On peut expliquer cela par le fait que le code DIVA est écrit en simple précision. Le gain de précision apporté par les schémas d'ordre élevé n'est donc pas pris en compte. Il faudrait donc récrire le code en double précision mais on augmenterait alors considérablement les coûts de calcul.

Se pose aussi la question de l'orientation du vecteur pénalisation sur un maillage non structuré. Pour les maillages structurés, nous pouvions définir simplement cette orientation à partir d'un élément de référence. Dans le cas des maillages non structurés, nous avons choisi de prendre un vecteur orthogonal à chaque arête et dirigé vers le haut ou vers la droite. Cependant, il n'y a pas de réelle justification à ce choix qui influence peut-être le niveau de précision du schéma.

Nous étudions ces deux questions à la section suivante.

## 5.5 Étude de l'effet d'autres facteurs sur la précision du code DIVA

Nous avons vu dans les sections précédentes que, sans paramètre de pénalisation, nos schémas ne convergeaient plus avec une source ponctuelle. Ce dernier était le facteur principal expliquant les problèmes de précision de nos schémas. Nous allons maintenant étudier l'effet de deux autres facteurs qui surviennent lors de la mise en œuvre industrielle des algorithmes.

### 5.5.1 Effet de la simple précision

Nous avons dans un premier temps refait les expériences avec une source ponctuelle dans le code DIVA, avec des éléments  $\mathcal{P}^3$  et des maillages réguliers. Les erreurs obtenues sont présentées dans le tableau 5.10 et les courbes de convergence sont sur la figure 5.13.

$\delta$	dt	Erreur	CV
1	5.86e-04	2.58e-01	
2	2.93e-04	3.06e-02	3.0
4	1.47e-04	1.16e-02	1.4
8	7.33e-05	5.69e-03	1.0

(a)  $\mathcal{P}^3$ , DG Flux Centrés DIVA, LF

dt	Erreur	CV
5.27e-04	2.50e-01	
2.64e-04	2.78e-02	3.1
1.32e-04	1.20e-02	1.2
6.59e-05	1.31e-02	-0.13

(b)  $\mathcal{P}^3$ , DG Flux Centrés DIVA,  $\nabla$

dt	Erreur	CV
1.56e-03	3.14e-01	
7.80e-04	6.62e-02	2.2
3.90e-04	3.00e-02	1.1
1.95e-04	1.50e-02	1.0

(c)  $\mathcal{P}^3$ , DG Flux Centrés DIVA, ADER

TABLE 5.10 – Erreur et taux de convergence en fonction du pas de maillage  $h$  avec pénalisation et pour des éléments finis  $\mathcal{P}^3$ . La simple précision n'offre pas les résultats espérés.

Il est évident que les schémas ne convergent pas avec l'ordre attendu. À partir d'une précision de  $10^{-2}$ , ils ne convergent même plus. Pensant à un bug, nous avons utilisé notre code prototype en simple précision. Les résultats obtenus sont présentés sur les figures 5.14 (pour les schémas DG-Flux) et 5.15 (pour les schémas IPDG).

Comme nous pouvions nous y attendre, les schémas ne convergent plus à partir d'une certaine précision, de l'ordre de  $10^{-4}$  à  $10^{-5}$ . Cette précision est cependant meilleure que celle de DIVA. Cela peut s'expliquer par le fait que notre code prototype n'utilise pas réellement de maillage, car il a été construit exclusivement pour des maillages structurés. Le calcul des fonctions de transformation entre

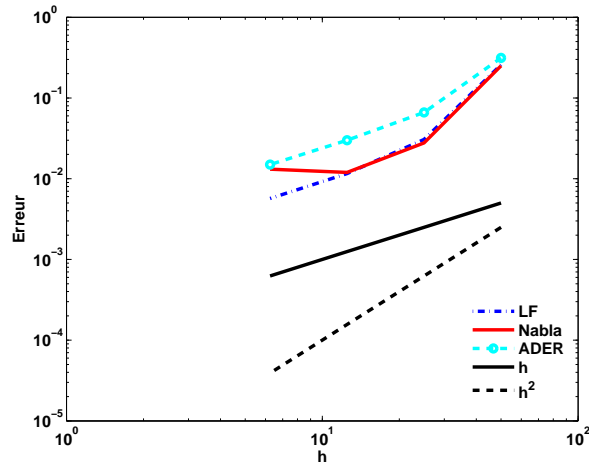
(a) Comportement de l'erreur DIVA simple précision en  $\mathcal{P}^3$ 

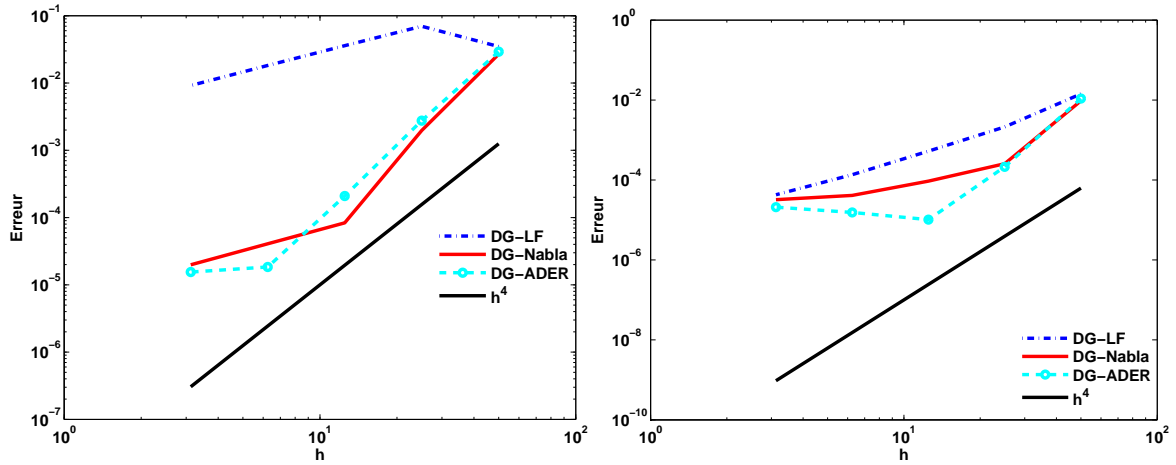
FIGURE 5.13 – Courbes d'erreur, en échelle loglog, pour les schémas LF, ADER et  $\nabla$  avec pénalisation et pour des éléments fins de degré 3 en simple précision.

chaque maille et la maille de référence est donc fait de manière exacte, alors que ce calcul est fait en simple précision dans DIVA, même lorsque le maillage est structuré.

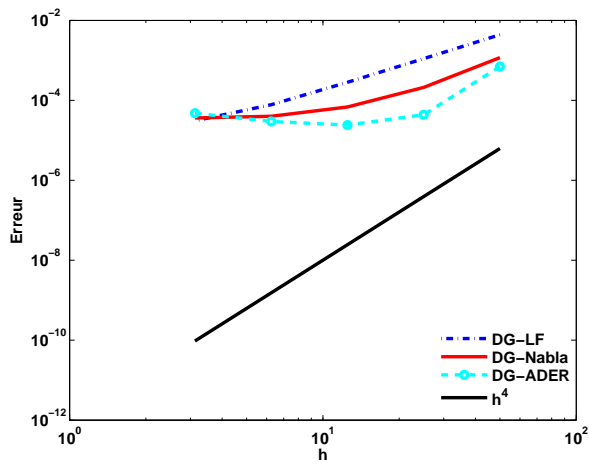
### 5.5.2 Effet du choix du vecteur de pénalisation

Dans le cas de maillages structurés, il est facile de définir l'orientation du vecteur de pénalisation, car il n'y a que trois types d'arêtes. C'est plus compliqué pour les maillages non structurés car les arêtes peuvent suivre toutes les directions. En suivant ce que nous avons fait pour les maillages structurés, nous avons décidé d'orienter les vecteurs  $\alpha_T$  orthogonalement aux arêtes et de les diriger vers le haut et vers la droite (voir figure 5.16.(a)). Nous n'avons pas de réelle justification théorique, mais nous allons montrer que, même sur des maillages structurés, un autre choix peut perturber la convergence. Nous avons donc décidé d'orienter le vecteur de pénalisation sur l'hypoténuse des triangles vers le bas et vers la gauche, en suivant la configuration 2(voir figure 5.16.(b)). Les résultats sont présentés dans les tableaux 5.11, 5.12 et 5.13. La convergence du schéma  $\nabla$  est d'ordre inférieur à quatre sur les maillages grossiers, mais cela est compensé par une super-convergence sur les maillages fins. Le schéma ADER converge bien à l'ordre quatre. Pour des éléments  $\mathcal{P}^3$ , le schéma LF ne converge qu'à l'ordre 1.

Ainsi, un simple changement dans l'orientation du vecteur de pénalisation peut considérablement changer le comportement des schémas numériques, même sur des maillages structurés. Dans le cas de maillages non structurés, où il n'est pas évident de définir une stratégie claire pour l'orientation de ce vecteur, cela peut très certainement affecter la précision des solutions. C'est un nouvel argument en faveur des méthodes IPDG qui sont basées sur des pénalisations scalaires, indépendantes de l'orientation des arêtes.

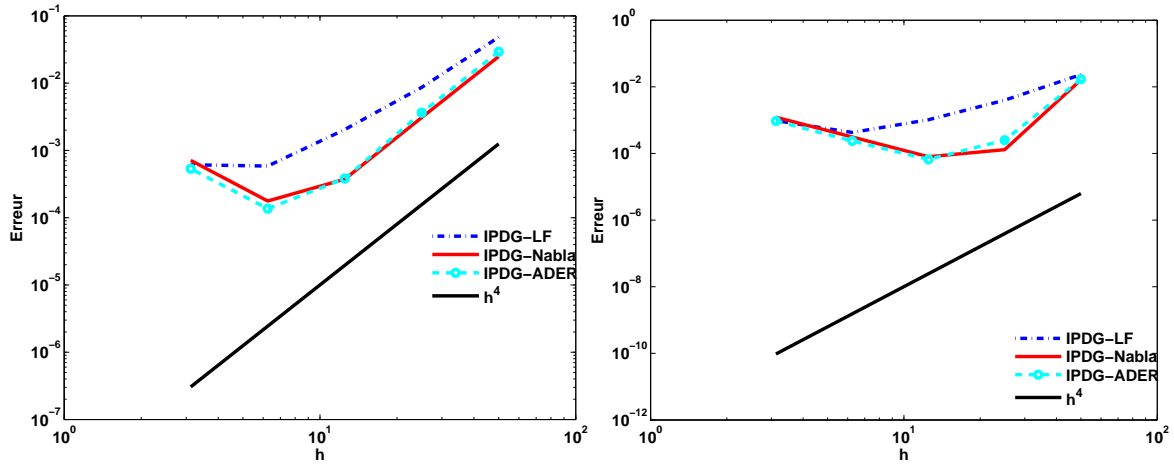


(a) Comportement de l'erreur en simple précision, DG Flux, pour des éléments  $\mathcal{P}^3$ , (b) Comportement de l'erreur en simple précision, DG Flux, pour des éléments  $\mathcal{P}^4$

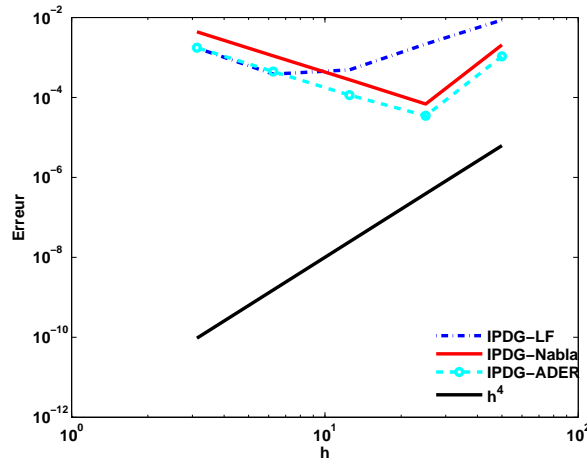


(c) Comportement de l'erreur en simple précision, DG Flux, pour des éléments  $\mathcal{P}^5$

FIGURE 5.14 – Courbes d'erreur, en échelle loglog, pour les schémas DG-LF, DG-ADER et DG- $\nabla$  et pour des éléments finis de degrés 3,4 et 5 en simple précision.



(a) Comportement de l'erreur en simple précision, IPDG, pour des éléments  $\mathcal{P}^3$ , (b) Comportement de l'erreur en simple précision, IPDG, pour des éléments  $\mathcal{P}^4$



(c) Comportement de l'erreur en simple précision, IPDG, pour des éléments  $\mathcal{P}^5$

FIGURE 5.15 – Courbes d'erreur, en échelle loglog, pour les schémas IPDG-LF, IPDG-ADER et IPDG- $\nabla$  et pour des éléments de degrés 3,4 et 5 en simple précision.

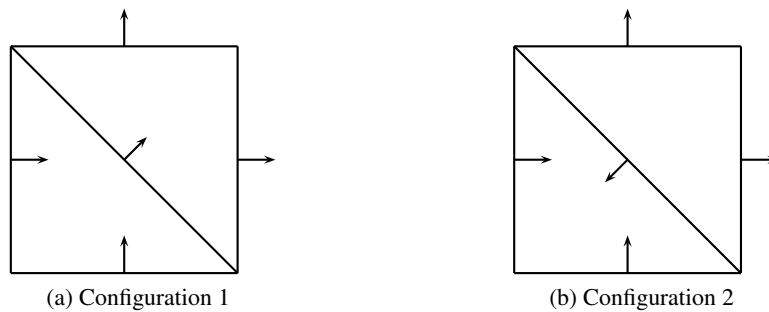


FIGURE 5.16 – Orientation du vecteur  $\alpha_\Gamma$ .

$\delta$	dt	Erreur	CV
1	9.6e-04	5.5e-02	
2	4.8e-04	6.9e-02	-0.3
4	2.4e-04	3.6e-02	1.0
8	1.2e-04	1.8e-02	1.0
16	6.0e-05	9.2e-03	1.0

(a) LF

dt	Erreur	CV
9.6e-04	5.4e-01	
4.8e-04	4.6e-02	3.5
2.4e-04	9.0e-03	2.4
1.2e-04	8.3e-03	0.1
6.0e-05	2.1e-04	5.3

(b)  $\nabla$

dt	Erreur	CV
2.7e-03	5.6e-01	
1.3e-03	4.8e-02	3.5
6.7e-04	3.3e-03	3.9
3.4e-04	2.1e-04	4.0
1.7e-04	1.3e-05	4.0

(c) ADER

TABLE 5.11 – Erreur et taux de convergence en fonction du pas de maillage  $h$  avec pénalisation et pour des éléments  $\mathcal{P}^3$ . Le schéma LF ne converge qu'à l'ordre 1 et la convergence du schéma  $\nabla$  est très irrégulière.

$\delta$	dt	Erreur	CV
1	6.6e-04	1.8e-02	
2	3.3e-04	2.2e-03	3.1
4	1.7e-04	5.2e-04	2.0
8	8.3e-05	1.3e-04	2.0
16	4.1e-05	3.3e-05	2.0

(a) LF

dt	Erreur	CV
6.6e-04	5.6e-02	
3.3e-04	1.1e-02	2.4
1.7e-04	3.4e-03	1.7
8.3e-05	2.8e-04	3.6
4.1e-05	2.0e-06	7.1

(b)  $\nabla$

dt	Erreur	CV
1.8e-03	4.5e-02	
9.2e-04	6.6e-03	2.8
4.6e-04	5.6e-04	3.6
2.3e-04	3.8e-05	3.9
1.2e-04	2.4e-06	4.0

(c) ADER

TABLE 5.12 – Erreur et taux de convergence en fonction du pas de maillage  $h$  avec pénalisation et pour des éléments  $\mathcal{P}^4$ . La convergence du schéma  $\nabla$  est très irrégulière.

$\delta$	dt	Erreur	CV
1	9.6e-04	4.5e-03	
2	2.4e-04	1.1e-03	2.0
4	1.2e-04	2.8e-04	2.0
8	6.0e-05	6.9e-05	2.0
16	3.0e-05	1.7e-05	2.0

(a) LF

dt	Erreur	CV
4.8e-04	4.2e-02	
2.4e-04	7.9e-04	5.7
1.2e-04	1.7e-04	2.2
6.0e-05	2.1e-06	6.4
3.0e-05	1.2e-08	7.5

(b)  $\nabla$

dt	Erreur	CV
1.3e-03	2.8e-02	
6.7e-04	7.9e-04	5.1
3.4e-04	1.4e-05	5.8
1.7e-04	2.3e-07	5.9
8.4e-05	3.7e-09	6.0

(c) ADER

TABLE 5.13 – Erreur et taux de convergence en fonction du pas de maillage  $h$  avec pénalisation et pour des éléments  $\mathcal{P}^5$ . La convergence du schéma  $\nabla$  est très irrégulière.



## 5.6 Conclusion

Ce chapitre avait pour objectif d'illustrer les effets du paramètre de pénalisation sur la convergence de la méthode DG-Flux. Dans le cas où on résout un problème aux limites avec des conditions initiales régulières, on observe qu'en ajoutant une pénalisation dans la formulation, l'ordre de convergence est optimal. Dans le cas d'une source ponctuelle, la formulation DG-Flux pénalisée donne de bons résultats puisque tous les artefacts sont éliminés. Cependant, le niveau de précision des calculs n'égale pas celui qui est atteint avec une formulation IPDG. C'est certainement lié au choix du paramètre de pénalisation. Dans la formulation IPDG, le paramètre est scalaire et la théorie permet de définir une règle claire qui fixe des valeurs admissibles en fonction du degré d'approximation. Pour pénaliser la formulation DG-Flux, on utilise un paramètre vectoriel que l'on fixe à partir de sa norme et d'un vecteur directeur. Le choix du paramètre de pénalisation est donc plus délicat à réaliser et il n'est pas évident de dire s'il existe un choix optimal de paramètre pour un maillage quelconque. On peut aussi se demander si l'approximation de la source n'empêche pas d'atteindre le niveau de convergence optimal. Nous n'avons pas réussi à obtenir des estimations a priori qui allaient dans ce sens. Toutefois, nous n'étions pas vraiment convaincus que chercher dans cette direction était une bonne idée car nous avons appliqué à la source exactement la même méthode d'approximation pour les formulations DG-Flux et IPDG. La formulation IPDG semble donc plus performante que la formulation DG-Flux. Néanmoins, le schéma DG-flux, basé sur la formulation du premier ordre de l'équation des ondes, a l'avantage de donner un accès direct aux gradients de champs de vecteurs. On peut ainsi utiliser des conditions d'imagerie plus sensibles aux variations du milieu élastique sans générer de calculs supplémentaires. Les résultats obtenus dans ce chapitre montrent que pour une fréquence donnée, les schémas d'ordre élevé permettent d'utiliser des maillages plus grossiers pour obtenir une précision équivalente à celle des schémas d'ordre 2. Pour un maillage donné, on peut ainsi considérer des fréquences plus hautes et obtenir des images plus précises. Toutefois, on observe que pour une précision de l'ordre de 1%, les schémas d'ordre élevé ne sont pas plus efficaces que les schémas d'ordre deux. Ils ne s'avèrent plus performants que pour un niveau de précision de l'ordre de 0.1%. Les schémas d'ordre élevé ne sont donc pas intéressants pour la RTM qui privilégie temps de calcul sur précision. Ils peuvent cependant être efficaces pour d'autres applications industrielles, comme par exemple les expériences de modélisation. Ces expériences ont pour but de déterminer différents paramètres, comme la position des capteurs et des sources, avant de lancer les campagnes d'acquisition. Dans ce contexte, la précision prime sur les temps de calcul et elle doit être supérieure à celle requise par la RTM. En conclusion, il nous semble plus intéressant d'utiliser des schémas IPDG couplés à un schéma d'ordre deux en temps pour des expériences de RTM acoustiques et des schémas IPDG couplés à des schémas  $\nabla$  pour des expériences de modélisation acoustique ou élastodynamique. En ce qui concerne la RTM élastique, s'il est évident qu'il faut utiliser des schémas en temps d'ordre deux, il faudra mener des tests numériques supplémentaires pour déterminer la méthode de discrétisation en espace la plus adaptée. Ces tests nécessitent une mise en œuvre préalable des conditions d'imagerie spécifiques à l'élastique dans DIVA, ce qui dépasse le cadre de cette thèse.

## Chapitre 6

# Discrétisation en temps avec des pas de temps locaux

L'objet de ce chapitre est de construire et analyser des schémas de discrétisation en temps qui sont adaptés à la formulation du premier ordre décrite au chapitre 2 et qui acceptent des pas de temps variables dans le domaine. L'utilisation de pas de temps locaux est d'un intérêt évident pour l'imagerie sismique. En effet, si les milieux à explorer sont parfois fortement hétérogènes, ils ne le sont pas partout et les maillages utilisés pour reproduire le milieu sont amenés à être composés de mailles de taille très variable. Une gestion optimisée des calculs numériques à réaliser est donc favorisée en adaptant le pas de temps à la taille des mailles recouvrant le domaine de calcul. En élaborant des techniques d'intégration en temps locales, on peut donc gagner considérablement sur les temps de calcul mais aussi, dans une perspective de RTM, sur l'occupation de la mémoire en stockant un nombre bien maîtrisé d'instantanés de propagation.

De nombreux travaux ont été consacrés à la construction de schémas à pas de temps locaux. Les applications visées sont diverses et assez nombreuses pour expliquer que dans ce manuscrit, nous nous limitons à ne citer que les travaux relatifs aux équations d'ondes. Pour le cas  $1D$  de l'équation des ondes, on peut citer [19, 20] qui a été suivi par [49] où on délivre une analyse d'erreur détaillée. La technique a ensuite été étendue aux équations de Maxwell en dimensions 2 et 3 dans [21] puis aux équations de l'élastodynamique dans [11] et au cas du couplage acoustique-élastodynamique dans [27]. Le couplage entre les zones maillées finement et grossièrement se fait, soit directement dans la formulation variationnelle en prenant en compte des conditions de transmission dans les termes de bord, soit par l'introduction d'un multiplicateur de Lagrange sur l'interface entre les deux milieux. Le premier traitement du couplage nécessite cependant de considérer deux formulations différentes dans les deux milieux. Par exemple, pour l'équation acoustique, on devra considérer la formulation en pression dans un des milieux et la formulation en vitesse ou en déplacement dans le deuxième milieu. On ne pourra donc pas traiter le cas d'un couplage entre trois milieux différents, car l'équation des ondes dans le troisième milieu devrait être à la fois formulée en pression pour assurer le couplage avec le deuxième milieu et en déplacement ou en vitesse pour assurer le couplage avec le deuxième milieu. Concernant la discrétisation en temps, il existe également deux possibilités : **a)** le schéma est complètement explicite, mais dans ce cas, la CFL peut ne pas être optimale, c'est-à-dire que le pas de temps peut être plus petit que les pas de temps des deux milieux pris séparément ; **b)** le schéma est explicite sauf sur l'interface où on doit résoudre un système linéaire liant toutes les inconnues de l'interface. Dans ce dernier cas, l'optimalité de la CFL est garantie par la conservation d'une énergie discrète. Le surcoût lié à l'inversion du système linéaire est faible par rapport au coût de calcul global

car ce système est d'une dimension inférieure à la dimension du problème, mais cette inversion peut néanmoins pénaliser les performances du code dans un cadre de calcul parallèle.

Un autre inconvénient de cette méthode est qu'elle se limite à la construction de schémas d'ordre deux en temps, et son extension à des schémas d'ordre plus élevés n'est pas évidente à réaliser.

Une autre technique, proposée par Piperno [57], est basée sur la théorie des intégrateurs symplectiques, qui est utilisée pour la discrétisation des systèmes Hamiltoniens. Dans ce travail, l'auteur développe deux algorithmes différents. Le premier, dit "explicite-implicite", utilise un schéma explicite sur le maillage grossier et implicite sur le maillage fin. Cette technique est très intéressante lorsque le nombre de petites mailles est négligeable par rapport au nombre de grandes mailles, car le coût de la résolution du système linéaire lié au schéma implicite est très faible. Dans le cas contraire, Piperno propose un second algorithme utilisant des schémas explicites dans les deux milieux. Cet algorithme garantit la conservation d'une énergie discrète et, contrairement à la méthode de Collino *et al*, la CFL est optimale sans avoir à résoudre de système linéaire à l'interface entre maillages fin et grossier. Cependant, les schémas en temps proposés sont également uniquement d'ordre deux et leur extension à des ordres supérieurs est toujours un problème ouvert.

Pour construire des schémas explicites d'ordre élevé, on peut se tourner vers les schémas ADER [50, 30], proposés par Dumbser, Käser *et al*. Ces schémas sont complètement explicites et la CFL est optimale. L'idée de la méthode consiste à effectuer d'abord la discrétisation en temps élément par élément, avec un pas de temps et un ordre de discrétisation variables sur chaque élément, puis à remplacer les opérateurs d'ordre élevé en espace par des opérateurs d'ordre deux *via* l'introduction d'inconnues auxiliaires. Dans le cas où le pas temps et l'ordre de discrétisation sont les mêmes sur tous les éléments, on peut montrer que cette méthode est équivalente à la technique de l'équation modifiée classique dont nous avons déjà parlé. Cependant, lorsque l'ordre et le pas de temps varient, il n'y a plus de conservation de l'énergie discrète et le schéma devient dissipatif, ce qui pénalise les expériences en temps long.

Ce chapitre renferme les travaux que nous avons réalisés dans l'optique de construire des schémas à pas de temps locaux d'ordre arbitrairement élevé. Pour respecter les contraintes dictées par l'application RTM, nous ne voulons pas introduire de multiplicateur de Lagrange ni résoudre un système linéaire sur l'interface entre les milieux fin et grossier. De plus, notre approche doit mener à un système discret pour lequel il existe une énergie conservée. Nous avons choisi de partir des travaux réalisés dans [26] pour les équations d'ondes du second ordre qui ont été étendus par Caroline Baldasari dans sa thèse [8] et dans [9] pour que les ordres de discrétisation soient aussi locaux. L'extension à des schémas en temps de type Adams-Bashworth et Runge-Kutta a été récemment proposée par Grote et Mitkova dans [37]. Dans un premier temps, (section 6.1), nous rappelons le principe de la technique pour la formulation du second ordre de l'équation des ondes, et nous montrons qu'il ne peut pas directement être appliqué à la formulation du premier ordre. Nous nous intéresserons alors (section 6.2) à des schémas à pas de temps locaux de type Adams-Bashforth. Puis, dans la section 6.3, nous proposons un nouveau schéma à pas de temps local pour la formulation du premier ordre.

## 6.1 Pas de temps local pour une équation hyperbolique d'ordre deux en temps

Dans cette section, nous rappelons le principe de la méthode de pas de temps local pour un système d'équations aux dérivées partielles d'ordre deux en temps comme par exemple l'équation des ondes

acoustiques. Après discrétisation en espace, le système s'écrit sous la forme :

$$(6.1.1) \quad \begin{cases} \mathcal{M} \frac{d^2 \mathbf{P}}{dt^2}(t) + \mathcal{K} \mathbf{P}(t) = \mathbf{F}(t), & t \in ]0, T[, \\ \mathbf{P}(0) = \mathbf{U}_0 \text{ et } \frac{d\mathbf{U}}{dt}(0) = \mathbf{U}_1, \end{cases}$$

où  $\mathbf{P}(t)$  est un vecteur dont les composantes sont les valeurs discrètes de la solution continue approchée en espace. Les vecteurs  $\mathbf{F}(t)$ ,  $\mathbf{U}_0$  et  $\mathbf{U}_1$  décrivent les approximations en espace du terme source et des conditions initiales. Les matrices  $\mathcal{M}$  et  $\mathcal{K}$  sont appelées respectivement matrice de masse et matrice de raideur. Notons que le système ne fait pas intervenir de terme d'amortissement qui se justifierait par exemple par l'emploi de conditions de bord absorbant.

Notre objectif étant de construire des schémas en temps explicites, nous sommes sous la contrainte d'avoir une matrice de masse idéalement diagonale ou facilement inversible, ce qui est le cas quand on choisit une méthode DG. Supposons maintenant qu'on ait sélectionné un schéma en temps  $S$ , par exemple, pour citer le plus commun, le schéma Leap-Frog. Dans tous les cas, la théorie garantit la stabilité si la condition CFL du schéma est respectée, c'est-à-dire si le pas de temps  $\Delta\tau$  vérifie  $\Delta\tau \leq \alpha_S h$  où  $\alpha_S$  est la constante CFL fixée par l'approximation en espace et en temps et  $h$  désigne la taille minimale du pas de discrétisation en espace. La théorie impose donc un choix global de  $\Delta\tau$  dicté par la valeur de  $h$ . Si le domaine de propagation est décrit par un maillage structuré, cela ne pose aucun problème. Mais dès qu'on s'intéresse à des cas réalistes, les milieux sont hétérogènes et peuvent inclure des effets de topographie. Il est donc plus efficace d'utiliser des maillages non structurés qui vont être constitués de petites mailles dans certaines zones et de plus ou moins grandes mailles dans d'autres. Respecter globalement la condition CFL va alors générer un nombre d'itérations en temps supérieur au nombre requis si la condition CFL était satisfaite localement. Et l'intérêt de réduire les temps de calcul n'est pas la seule motivation pour appliquer localement la condition CFL. En effet, lorsque le pas de temps est beaucoup plus petit que le pas de temps dicté par la théorie, on peut observer un phénomène de pollution causé par de la dispersion numérique.

Dans la suite, nous allons donc supposer que le maillage est constitué de zones maillées plus ou moins finement. On va noter  $h$  le pas de discrétisation correspondant à la maille la plus fine et on va supposer que les mailles les plus larges sont obtenues en multipliant  $h$  par un facteur  $m$  où  $m$  est un entier supérieur ou égal à 2 (voir figure 6.1). Pour simplifier la mise en oeuvre, nous envisageons d'utiliser le même schéma en temps partout. Ainsi, si on conserve la notation  $\Delta\tau$  pour le pas de temps lié à  $h$ , le pas de temps relatif au maillage grossier est noté  $\Delta t$  et on a la relation  $\Delta t = m\Delta\tau$ . La construction d'un schéma à pas de temps locaux va nécessiter d'intégrer en temps le système en utilisant *a priori* comme pas de temps  $\Delta t$  et lorsque cela est nécessaire, on effectuera des calculs sur un intervalle de temps  $\Delta t$  avec pour pas d'intégration  $\Delta\tau$ . C'est pourquoi nous allons commencer par décrire deux façons d'intégrer en temps le système pour montrer qu'en s'y prenant bien, on peut sélectionner le pas de temps qui sera soit  $\Delta t$  soit  $\Delta\tau$ . Dans la suite, pour simplifier l'exposé, nous noterons  $\mathcal{A} = \mathcal{M}^{-1}\mathcal{K}$  et nous ferons abstraction du terme source.

On découpe l'intervalle  $[0, T]$  en  $N$  intervalles de taille  $\Delta t$  et on introduit, à chaque instant  $t_n = n\Delta t$ , la variable auxiliaire  $\tilde{\mathbf{P}}_n(\tau) = \frac{\mathbf{P}(n\Delta t - \tau) + \mathbf{P}(n\Delta t + \tau)}{2}$ , définie pour  $\tau \in [-\Delta t, \Delta t]$ . Cette fonction

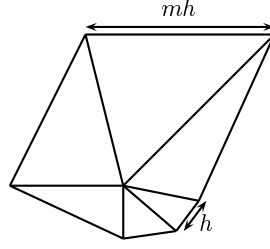


FIGURE 6.1 – Un exemple de maillage contenant des cellules fines et des cellules grossières.

est paire et vérifie,

$$(6.1.2) \quad \begin{cases} \frac{d^2 \tilde{\mathbf{P}}_n(\tau) + \mathcal{A} \tilde{\mathbf{P}}_n(\tau) = 0, & \tau \in ]0, \Delta t[, \\ \tilde{\mathbf{P}}_n(0) = \mathbf{P}(t_n) \text{ et } \frac{d\tilde{\mathbf{P}}_n}{d\tau}(0) = 0 \end{cases}$$

Remarquons d'abord que si nous disposions de la solution exacte de (6.1.2), nous pourrions calculer  $U$  à l'instant  $t_{n+1}$  en fonction de  $\mathbf{P}(t_{n-1})$  et  $\mathbf{P}(t_n)$  car on a la relation  $\mathbf{P}(t_{n+1}) = 2\tilde{\mathbf{P}}(\Delta t) - \mathbf{P}(t_{n-1})$ . En pratique, on ne sait calculer qu'une approximation de  $\tilde{\mathbf{P}}$  qui diffère, selon la technique retenue pour intégrer le système (6.1.2). Nous allons en décrire deux en commençant par la méthode la plus directe qui consiste à intégrer l'équation en temps en appliquant un schéma Leap-Frog de pas de temps  $\Delta\tau$ . L'algorithme de calcul qui en découle se décline de la façon suivante :

1. Initialiser  $\tilde{\mathbf{P}}_n^{1/m} = \tilde{\mathbf{P}}_n^0 - \frac{\Delta\tau^2}{2} \mathcal{A} \tilde{\mathbf{P}}_n^0 = \mathbf{P}^n - \frac{\Delta\tau^2}{2} \mathcal{A} \mathbf{P}_n$
2. Calculer  $\tilde{\mathbf{P}}_n^{(i+1)/m} = 2\tilde{\mathbf{P}}_n^i - \tilde{\mathbf{P}}_n^{(i-1)/m} - \Delta\tau^2 \mathcal{A} \tilde{\mathbf{P}}_n^i$ , pour  $i = 1..p-1$
3. En déduire  $\mathbf{P}^{n+1} = 2\tilde{\mathbf{P}}^1 - \mathbf{P}^{n-1}$ .

On peut remarquer que la parité de  $\tilde{\mathbf{P}}$  implique que  $\tilde{\mathbf{P}}^{1/m} = \tilde{\mathbf{P}}^{-1/m}$ . Cela signifie donc que l'initialisation décrite au point 1. correspond exactement à une itération du schéma Leap-Frog sur un pas de temps.

On vérifie facilement que ce schéma est équivalent au schéma Leap-Frog de pas de temps  $\Delta\tau$  appliqué directement à  $\mathbf{P}$  :

$$\frac{\mathbf{P}^{n+1} - 2\mathbf{P}^n + \mathbf{P}^{n-1}}{\Delta\tau^2} = -\mathcal{A}\mathbf{P}^2.$$

Supposons maintenant qu'avant de discrétiser le schéma, on approche le terme  $\mathcal{A}\tilde{\mathbf{P}}_n(\tau)$  par un développement de Taylor d'ordre 1 :

$$\mathcal{A}\tilde{\mathbf{P}}_n(\tau) = \mathcal{A}\tilde{\mathbf{P}}_n(0) + \tau \mathcal{A}\tilde{\mathbf{P}}_n'(0) + O(\Delta t^2).$$

Comme  $\tilde{\mathbf{P}}_n'(0) = 0$  grâce à la parité de  $\tilde{\mathbf{P}}_n$ , on a, en utilisant  $\tilde{\mathbf{P}}_n(0) = \mathbf{P}^n$ ,

$$\mathcal{A}\tilde{\mathbf{P}}_n(\tau) = \mathcal{A}\mathbf{P}^n + O(\Delta t^2).$$

Le système (6.1.2) est alors approché par :

$$(6.1.3) \quad \begin{cases} \frac{d^2 \tilde{\mathbf{P}}_n(\tau) + \mathcal{A}\mathbf{P}^n = 0, & \tau \in ]0, \Delta t[, \\ \tilde{\mathbf{P}}_n(0) = \mathbf{P}^n \text{ et } \frac{d\tilde{\mathbf{P}}_n}{d\tau}(0) = 0 \end{cases}$$

On résout ensuite ce système en appliquant un schéma Leap-Frog de pas de temps  $\Delta\tau$ . L'algorithme s'écrit alors

1.  $\tilde{\mathbf{P}}_n^{1/m} = \tilde{\mathbf{P}}_n^0 - \frac{\Delta\tau}{2}\mathcal{A}\mathbf{P}^n$ ,
2.  $\tilde{\mathbf{P}}_n^{(i+1)/m} = 2\tilde{\mathbf{P}}_n^i - \tilde{\mathbf{P}}_n^{(i-1)/m} - \Delta\tau^2\mathcal{A}\mathbf{P}^n$ , pour  $i = 1..p-1$
3.  $\mathbf{P}^{n+1} = 2\tilde{\mathbf{P}}^1 - \mathbf{P}^{n-1}$ .

Dans ce cas, on vérifie, après quelques calculs, que cet algorithme est équivalent à un schéma du type équation modifiée d'ordre deux et de pas de temps  $\Delta t$  :

$$\frac{\mathbf{P}^{n+1} - 2\mathbf{P}^n + \mathbf{P}^{n-1}}{\Delta t^2} = -\mathcal{A}\mathbf{P}^n.$$

En résumé, en modifiant l'approximation de  $\tilde{\mathbf{P}}$ , nous obtenons deux méthodes d'intégration en temps basées sur un schéma Leap-frog. La première intègre directement (6.1.2) et le pas de discrétisation est  $\Delta\tau$ . La seconde repose sur une version approchée de (6.1.2) qui consiste à modifier son second membre via une formule de Taylor et on se ramène ainsi à intégrer sur le pas de temps  $\Delta t$ . L'idée de la méthode à pas de temps local est alors de combiner ces deux schémas pour obtenir un schéma à pas de temps  $\Delta\tau$  sur les petites mailles et  $\Delta t$  sur les grandes mailles.

Nous supposons maintenant, sans perte de généralité, que les grandes mailles contiennent  $N_g$  degrés de liberté numérotés de 1 à  $N_g$  et que les petites mailles contiennent  $N_f$  degrés de liberté numérotés de  $N_g + 1$  à  $N_g + N_f$ . On note alors  $\mathbf{P}_g$  et  $\mathbf{P}_f$  les vecteurs de longueur  $N_g + N_f$ . Ils sont définis par :

$$\mathbf{P}_g = \begin{bmatrix} P_1, \\ \dots, \\ P_{N_f}, \\ 0, \\ \dots, \\ 0 \end{bmatrix} \quad \text{et} \quad \mathbf{P}_f = \begin{bmatrix} 0, \\ \dots, \\ 0, \\ P_{N_f+1}, \\ \dots, \\ P_{N_f+N_g} \end{bmatrix}.$$

On note  $\mathcal{P}$  la matrice de projection sur la grille fine. Elle est diagonale avec des coefficients nuls ou égaux à 1 sur la diagonale, répartis de telle sorte que  $\mathcal{P}\mathbf{P} = \mathbf{P}_f$  et  $(\mathcal{I} - \mathcal{P})\mathbf{P} = \mathbf{P}_g$ . On réécrit alors le système (6.1.2) sous la forme

$$(6.1.4) \quad \begin{cases} \frac{d^2\tilde{\mathbf{P}}_n}{d\tau^2}(\tau) + \mathcal{A}(\mathcal{I} - \mathcal{P})\tilde{\mathbf{P}}_n(\tau) + \mathcal{A}\mathcal{P}\tilde{\mathbf{P}}_n(\tau) = 0, \tau \in ]0, \Delta t[, \\ \tilde{\mathbf{P}}_n(0) = \mathbf{P}^n \text{ et } \frac{d\tilde{\mathbf{P}}_n}{d\tau}(0) = 0, \end{cases}$$

Nous avons vu qu'approcher le terme  $\mathcal{A}\tilde{\mathbf{P}}_n$  par un développement de Taylor nous permettait d'obtenir un schéma de pas de temps  $\Delta t$  alors que sans approximation, nous obtenions un schéma de pas de temps  $\Delta\tau$ . L'idée est donc de remplacer uniquement  $\mathcal{A}(\mathcal{I} - \mathcal{P})\tilde{\mathbf{P}}_n(\tau)$  par un développement de Taylor. On obtient ainsi le système

$$(6.1.5) \quad \begin{cases} \frac{d^2\tilde{\mathbf{P}}_n}{d\tau^2}(\tau) + \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{P}^n + \mathcal{A}\mathcal{P}\tilde{\mathbf{P}}_n(\tau) = 0, \tau \in ]0, \Delta t[, \\ \tilde{\mathbf{P}}_n(0) = \mathbf{P}^n \text{ et } \frac{d\tilde{\mathbf{P}}_n}{d\tau}(0) = 0. \end{cases}$$

Il ne reste plus qu'à discrétiser ce système par un schéma Leap-Frog de pas de temps  $\Delta\tau$  ; on obtient :

1.  $\tilde{\mathbf{P}}_n^{1/m} = \tilde{\mathbf{P}}_n^0 - \frac{\Delta\tau}{2} (\mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{P}^n + \mathcal{A}\mathcal{P}\mathbf{P}^n)$
2. pour  $i$  de 1 à  $m - 1$  ;

$$\tilde{\mathbf{P}}_n^{(i+1)/m} = 2\tilde{\mathbf{P}}_n^i - \tilde{\mathbf{P}}_n^{(i-1)/m} - \Delta\tau^2 \left( \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{P}^n + \mathcal{A}\mathcal{P}\tilde{\mathbf{P}}_n^i \right) ;$$

3.  $\mathbf{P}^{n+1} = 2\tilde{\mathbf{P}}_n^1 - \mathbf{P}^{n-1}$ .

Afin d'évaluer les performances d'une méthode d'intégration, il est important de pouvoir compter le nombre d'opérations nécessaires. Pour cela, il est utile de réécrire l'algorithme sous la forme

1. On pose  $v_1 = \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{P}^n$  et  $v_2 = \mathcal{A}\mathcal{P}\mathbf{P}^n$  ;
2.  $\tilde{\mathbf{P}}_n^{1/m} = \tilde{\mathbf{P}}_n^0 - \frac{\Delta\tau}{2} (v_1 + v_2)$  ;
3. Pour  $i$  de 1 à  $m - 1$ ,
  - (a) On pose  $v_2 = v_1 + \mathcal{A}\mathcal{P}\tilde{\mathbf{P}}_n^i$
  - (b)  $\tilde{\mathbf{P}}_n^{(i+1)/m} = 2\tilde{\mathbf{P}}_n^i - \tilde{\mathbf{P}}_n^{(i-1)/m} - \Delta\tau^2 v_2$  ;
4.  $\mathbf{P}^{n+1} = 2\tilde{\mathbf{P}}_n^1 - \mathbf{P}^{n-1}$ .

Pour un pas de temps  $\Delta t$ , il faut donc effectuer une multiplication par  $\mathcal{A}(\mathcal{I} - \mathcal{P})$  (étape 1) et  $m$  multiplications par  $\mathcal{A}\mathcal{P}$  (étapes 2 et 3.a), ce qui correspond bien au nombre d'opérations nécessaires pour un schéma Leap-Frog de pas de temps  $\Delta t$  sur les grandes mailles et de pas de temps  $\Delta\tau$  sur les petites mailles. Nous renvoyons à [26] pour l'étude de la stabilité de ce schéma. Les auteurs y montrent en particulier la conservation d'une énergie et la stabilité sous une condition CFL de type  $\Delta t \leq \alpha_{LF} h$  et  $\Delta\tau \leq \alpha_{LF} h/m$ . Le nouveau schéma permet donc d'utiliser deux pas de temps optimaux, le premier respectant la CFL du maillage grossier et le second la CFL du maillage fin.

L'inconvénient de cette méthode est qu'elle repose fortement sur la parité de la fonction  $\tilde{\mathbf{P}}_n$ , qui n'est garantie qu'en l'absence de terme de dissipation dans l'équation des ondes. Il s'agit là d'une réelle contrainte car il est fréquent d'introduire des conditions de bord absorbant qui injectent de la dissipation dans les équations. C'est pourquoi nous nous sommes intéressés à des méthodes ne nécessitant pas l'introduction d'une fonction auxiliaire paire.

## 6.2 Utilisation de la méthode d'Adams-Bashforth

Nous rappelons dans cette section la méthode de pas de temps local proposée par Grote et Mitkova [37] qui repose sur les schémas d'Adams-Bashforth. Cette méthode a l'avantage d'être adaptée aux équations hyperboliques d'ordre un en temps, et elle est directement utilisable pour les équations d'ondes en formulation du premier ordre.

### 6.2.1 Présentation de la méthode

La méthode d'Adams-Bashforth (AB) permet d'obtenir des schémas de discrétisation en temps explicites pour des équations différentielles ordinaires de la forme

$$(6.2.1) \quad \begin{cases} \frac{d\mathbf{Y}}{dt}(t) = f(\mathbf{Y}(t), t), \\ \mathbf{Y}(0) = \mathbf{Y}_0. \end{cases}$$

Les schémas semi-discrétisés en espace que nous avons présentés précédemment peuvent se réécrire de cette manière en posant

$$f(\mathbf{Y}(t), t) = \mathcal{A}\mathbf{Y}(t) = \begin{bmatrix} 0 & \mathcal{A}_1 \\ \mathcal{A}_2 & 0 \end{bmatrix} \mathbf{Y}(t), \quad \mathbf{Y}(t) = \begin{bmatrix} \mathbf{P}(t) \\ \mathbf{V}(t) \end{bmatrix} \text{ et } \mathbf{Y}_0 = \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{V}_0 \end{bmatrix}.$$

Nous supposons maintenant que nous disposons d'une approximation  $\tilde{\mathbf{Y}}(t)$  de  $\mathbf{Y}(t)$  d'ordre  $s$  jusqu'au temps  $t_n = n\Delta t$  avec  $n \geq s$ . Nous approchons ensuite  $f(\tilde{\mathbf{Y}}(t), t)$  sur l'intervalle  $[t_n, t_{n+1}]$  par son polynôme d'interpolation  $f_n(t)$  de degré  $s - 1$  aux instants  $t_{n-s+1}, t_{n-s+2}, \dots, t_n$  :

$$f(\tilde{\mathbf{Y}}(t), t) \approx f_n(t) = \sum_{j=1}^s \beta_j \left( \frac{t - t_n}{\Delta t} \right) f(\tilde{\mathbf{Y}}(t_{n-s+j}), t_{n-s+j}),$$

avec

$$\beta_j(\tau) = \frac{(-1)^{s-j}}{(j-1)!(s-j)!} \prod_{\substack{i=1 \\ i \neq j}}^s (\tau + i - s).$$

On obtient finalement l'équation différentielle ordinaire (EDO),

$$(6.2.2) \quad \begin{cases} \frac{d\tilde{\mathbf{Y}}}{dt}(t) = f_n(t), \\ \tilde{\mathbf{Y}}(t_n) = \mathbf{Y}^n, \end{cases}$$

pour  $t \in ]t_n, t_{n+1}[$ . Le second membre de cette équation étant polynomial, la solution peut être calculée exactement en utilisant la relation

$$\tilde{\mathbf{Y}}(t_{n+1}) - \tilde{\mathbf{Y}}(t_n) = \Delta t \int_{t_n}^{t_{n+1}} \frac{d\tilde{\mathbf{Y}}}{dt}(t) dt = \Delta t \int_{t_n}^{t_{n+1}} f_n(t) dt.$$

On obtient finalement

$$\tilde{\mathbf{Y}}(t_{n+1}) - \tilde{\mathbf{Y}}(t_n) = \Delta t \sum_{j=1}^s \gamma_j f(\tilde{\mathbf{Y}}(t_{n-s+j}), t_{n-s+j}), \text{ avec } \gamma_j = \int_0^1 \beta_j(t) dt.$$

Le schéma d'Adams-Bashforth d'ordre  $s$  pour l'équation des ondes semi-discrétisée en espace s'écrit donc, en faisant l'abus de notation  $\tilde{\mathbf{Y}}^n = \mathbf{Y}^n$ ,

$$\mathbf{Y}^{n+1} - \mathbf{Y}^n = \Delta t \sum_{j=1}^s \gamma_j \mathcal{A} \mathbf{Y}^{n-s+j}.$$

En particulier, le schéma d'Adams-Bashforth d'ordre 4 (AB4) que nous considérerons par la suite s'écrit

$$(6.2.3) \quad \mathbf{Y}^{n+1} = \mathbf{Y}^n + \frac{\Delta t}{24} (55\mathcal{A}\mathbf{Y}^n - 59\mathcal{A}\mathbf{Y}^{n-1} + 37\mathcal{A}\mathbf{Y}^{n-2} - 9\mathcal{A}\mathbf{Y}^{n-3})$$

Pour initialiser ces schémas, qui nécessitent la connaissance de la solution sur plusieurs pas de temps, on utilisera une méthode de type Runge-Kutta (RK) du même ordre que la méthode AB. Nous renvoyons à [44] pour plus de détails sur les propriétés des schémas d'Adams-Bashforth.



## 6.2.2 Méthode d'Adams-Bashforth avec pas de temps local

Dans cette section, nous proposons de réinterpréter la méthode à pas de temps local pour les schémas d'Adams-Bashforth qui a été proposée dans [37]. Avant de considérer les méthodes à pas de temps local proprement dites, nous allons montrer deux façons différentes de construire des schémas AB d'ordre 4 à pas temps global  $\Delta\tau = \Delta t/m$ . Nous supposons donc que nous disposons d'une approximation  $\tilde{\mathbf{Y}}(t)$  de  $\mathbf{Y}(t)$  d'ordre  $s$  jusqu'au temps  $t_{n,m} = n\Delta t + m\Delta\tau$  avec  $n \geq s$  et  $m \leq p$ . La différence entre les deux schémas que nous allons présenter réside dans l'approximation de  $f(\tilde{\mathbf{Y}}(t), t)$ .

Dans le premier cas, nous choisissons d'utiliser le polynôme d'interpolation de degré  $s - 1$  aux instants  $t_{n,m-s+1}, t_{n,m-s+2}, \dots, t_{n,m}$  :

$$f(\tilde{\mathbf{Y}}(t), t) \approx f_{n,m}^1(t) = \sum_{j=1}^s \beta_j \left( \frac{t - t_{n,m}}{\Delta\tau} \right) f(\tilde{\mathbf{Y}}(t_{n,m-s+j}), t_{n,m-s+j}).$$

et nous résolvons l'équation

$$(6.2.4) \quad \begin{cases} \frac{d\tilde{\mathbf{Y}}}{dt}(t) = f_{n,m}^1(t), \\ \tilde{\mathbf{Y}}(t_{n,m}) = \mathbf{Y}^{n,m}, \end{cases}$$

pour  $t \in ]t_{n,m}, t_{n,m+1}[$ . En opérant comme précédemment, on obtient,

$$\tilde{\mathbf{Y}}(t_{n,m+1}) - \tilde{\mathbf{Y}}(t_{n,m}) = \Delta\tau \sum_{j=1}^s \gamma_j f(\tilde{\mathbf{Y}}(t_{n,m-s+j}), t_{n,m-s+j}).$$

On obtient donc le schéma AB d'ordre 4 classique, de pas de temps  $\Delta\tau$ .

La seconde méthode est obtenue en deux temps. Nous interpolons tout d'abord  $f$  par son polynôme d'interpolation  $f_n(t)$  de degré  $s - 1$  aux instants  $t_{n-s+1}, t_{n-s+2}, \dots, t_n$  pour obtenir l'équation différentielle

$$(6.2.5) \quad \begin{cases} \frac{d\tilde{\mathbf{Y}}}{dt}(t) = f_n(t), \\ \tilde{\mathbf{Y}}(t_{n,m}) = \tilde{\mathbf{Y}}^{n,m}, \end{cases}$$

pour  $t \in ]t_{n,m}, t_{n,m+1}[$ . Dans un second temps, nous remplaçons  $f_n(t)$  par son polynôme d'interpolation de degré  $s - 1$  aux instants  $t_{n,m-s+1}, t_{n,m-s+2}, \dots, t_{n,m}$  :

$$f_n(t) = f_{n,m}^2(t) = \sum_{j=1}^s \beta_j \left( \frac{t - t_{n,m}}{\Delta\tau} \right) f_n(t_{n,m-s+j}).$$

En utilisant l'expression de  $f_n$ , on obtient

$$f_{n,m}^2(t) = \sum_{j=1}^s \beta_j \left( \frac{t - t_{n,m}}{\Delta\tau} \right) \left( \sum_{k=1}^s \beta_k \left( \frac{t_{n,m-s+j} - t_n}{\Delta t} \right) f(\tilde{\mathbf{Y}}(t_{n-s+k}), t_{n-s+k}) \right),$$

soit

$$f_{n,m}^2(t) = \sum_{k=1}^s \beta_{m,k} \left( \frac{t - t_{n,m}}{\Delta\tau} \right) f(\tilde{\mathbf{Y}}(t_{n-s+k}), t_{n-s+k}),$$

avec

$$\beta_{m,k}(\xi) = \sum_{j=1}^s \beta_j(\xi) \beta_k \left( \frac{t_{n,m-s+j} - t_n}{\Delta t} \right) = \sum_{j=1}^s \beta_j(\xi) \beta_k \left( \frac{m-s+j}{p} \right).$$

On peut noter que  $f_{n,m}^2$  n'est qu'une réécriture de  $f_n$  sous une autre forme. En effet  $f_n$  étant un polynôme de degré  $s$ , il est son propre polynôme d'interpolation de degré  $s$ .

L'équation différentielle à résoudre s'écrit finalement

$$(6.2.6) \quad \begin{cases} \frac{d\tilde{\mathbf{Y}}}{dt}(t) = f_{n,m}^2(t), \\ \tilde{\mathbf{Y}}(t_{n,m}) = \mathbf{Y}^{n,m}, \end{cases}$$

pour  $t \in ]t_{n,m}, t_{n,m+1}[$ . En résolvant cette équation, on obtient le schéma numérique

$$\tilde{\mathbf{Y}}(t_{n,m+1}) - \tilde{\mathbf{Y}}(t_{n,m}) = \Delta\tau \sum_{k=1}^s \gamma_{m,k} f(\tilde{\mathbf{Y}}(t_{n-s+k}), t_{n-s+k}),$$

avec

$$\gamma_{j,k} = \sum_{j=1}^s \beta_k \left( \frac{m-s+j}{p} \right) \int_0^1 \beta_j(t) dt.$$

Comme nous l'avons dit,  $f_{n,m}^2$  est égal à  $f_n$ . Les deux équations (6.2.2) et 6.2.6) sont donc équivalentes et admettent la même solution. Ainsi, la solution donnée par cette seconde méthode est la même que celle donnée par le schéma AB d'ordre 4 classique de pas de temps global  $\Delta t$ .

En résumé, nous avons présenté deux schémas de pas de temps global  $\Delta\tau$ . L'un est équivalent au schéma AB d'ordre 4 de pas de temps  $\Delta\tau$ , l'autre est équivalent au schéma AB d'ordre 4 de pas de temps  $\Delta t$ . Nous allons maintenant montrer comment combiner ces deux schémas pour obtenir un schéma à pas de temps local. Nous supposons que  $f(\mathbf{Y}(t), t)$  peut se décomposer en deux parties  $f(\mathbf{Y}(t), t) = f^{[g]}(\mathbf{Y}(t), t) + f^{[f]}(\mathbf{Y}(t), t)$ , où  $f^{[g]}(\mathbf{Y}(t), t)$  est une fonction oscillant rapidement avec  $t$  et  $f^{[f]}(\mathbf{Y}(t), t)$  une fonction oscillant encore plus rapidement. L'idée consiste maintenant à approcher  $f^{[g]}(\mathbf{Y}(t), t)$  par son interpolant de degré  $s$  aux instants  $t_{n-s+1}, t_{n-s+2}, \dots, t_n$  et  $f^{[f]}(\mathbf{Y}(t), t)$  par son interpolant de degré  $s$  aux instants  $t_{n,m-s+1}, t_{n,m-s+2}, \dots, t_{n,m}$ . Plus précisément, on écrit :

$$f^{[g]}(\tilde{\mathbf{Y}}(t), t) \approx f_{n,m}^{[g]}(t) = \sum_{k=1}^s \beta_{m,k} \left( \frac{t - t_{n,m}}{\Delta\tau} \right) f^{[g]}(\tilde{\mathbf{Y}}(t_{n-s+k}), t_{n-s+k}),$$

et

$$f^{[f]}(\tilde{\mathbf{Y}}(t), t) \approx f_{n,m}^{[f]}(t) = \sum_{j=1}^s \beta_j \left( \frac{t - t_{n,m}}{\Delta\tau} \right) f^{[f]}(\tilde{\mathbf{Y}}(t_{n,m-s+j}), t_{n,m-s+j}),$$

ce qui nous conduit à résoudre l'équation

$$(6.2.7) \quad \begin{cases} \frac{d\tilde{\mathbf{Y}}}{dt}(t) = f_{n,m}^{[g]}(t) + f_{n,m}^{[f]}(t), \\ \tilde{\mathbf{Y}}(t_{n,m}) = \mathbf{Y}^{n,m}. \end{cases}$$

La solution de cette équation vérifie

$$\tilde{\mathbf{Y}}(t_{n,m+1}) - \tilde{\mathbf{Y}}(t_{n,m}) = \Delta\tau \sum_{j=1}^s \gamma_{m,j} f^{[g]}(\tilde{\mathbf{Y}}(t_{n-s+j}), t_{n-s+j}) + \Delta\tau \sum_{j=1}^s \gamma_j f^{[f]}(\tilde{\mathbf{Y}}(t_{n,m-s+j}), t_{n,m-s+j}).$$

Il nous reste maintenant à exprimer  $f^{[g]}$  et  $f^{[f]}$  dans le cas particulier de l'équation des ondes semi-discrétisée. Nous avons  $f(\mathbf{Y}(t), t) = \mathcal{A}\mathbf{Y}(t)$ . Comme pour le schéma de Störmer-Verlet, nous décomposons le vecteur  $\mathbf{Y}$  sur les petites mailles et les grandes mailles en utilisant la matrice de projection  $\mathcal{P} : \mathbf{Y} = (\mathcal{I} - \mathcal{P})\mathbf{Y} + \mathcal{P}\mathbf{Y}$ .  $f(\mathbf{Y}(t), t)$  devient alors

$$f(\mathbf{Y}(t), t) = \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{Y}(t) + \mathcal{A}\mathcal{P}\mathbf{Y}(t) = f^{[g]}(\mathbf{Y}(t), t) + f^{[f]}(\mathbf{Y}(t), t).$$

Le schéma AB d'ordre 4 à pas de temps local pour l'équation des ondes semi-discrétisée s'écrit donc

$$\mathbf{Y}^{n,m+1} = \mathbf{Y}^{n,m} + \Delta\tau \sum_{j=1}^s \gamma_{m,j} \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{Y}^{n-s+j} + \Delta\tau \sum_{j=1}^s \gamma_j \mathcal{A}\mathcal{P}\mathbf{Y}^{n,m-s+j}.$$

En guise de conclusion à cette section, nous présentons ci-dessous l'algorithme permettant de calculer  $\mathbf{Y}^{n+1,0}$  à partir de  $\mathbf{Y}^{n,0}$ . Nous supposons qu'au début de cet algorithme, nous connaissons  $\mathbf{Y} = \mathbf{Y}^{n,0}$ ,  $\mathbf{W}_j^{[g]} = \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{Y}^{n-s+j}$  et  $\mathbf{W}_j^{[f]} = \mathcal{A}\mathcal{P}\mathbf{Y}^{n,m-s+j}$  pour  $j = 1 \dots s$ .

- Pour  $m = 1, \dots, p$ 
  - $\mathbf{Y} := \mathbf{Y} + \frac{\Delta t}{p} \sum_{l=1}^s \gamma_{m,l} \mathbf{W}_l^{[g]} + \frac{\Delta t}{p} \mathcal{A}\mathcal{P} \sum_{l=1}^s \gamma_l \mathbf{W}_l^{[f]}$ .
  - $\mathbf{W}_i^{[f]} = \mathbf{W}_{i+1}^{[f]}$ ,  $i = 1, \dots, s-1$
  - $\mathbf{W}_s^{[f]} = \mathcal{A}\mathcal{P}\mathbf{Y}$
  - Fin Pour
  - $\mathbf{W}_i^{[g]} = \mathbf{W}_{i+1}^{[g]}$ ,  $i = 1, \dots, s-1$
  - $\mathbf{W}_s^{[g]} = \mathcal{A}(\mathcal{I} - \mathcal{P})\mathbf{Y}$

Comme nous l'avons vu, cette technique nécessite de stocker la solution sur plusieurs pas de temps, ce qui augmente considérablement les coûts de calculs de la technique Reverse Time Migration. En particulier, la plate-forme DIVA utilise l'algorithme de Griewank, qui optimise les coûts de calcul et de stockage par l'utilisation de "checkpoints". Ces checkpoints sont des sauvegardes de la solution à plusieurs instants qui sont judicieusement choisis et permettent de relancer l'expérience de propagation à un instant donné sans avoir à repartir de  $t = 0$ . Avec un schéma Leap-Frog appliqué à la formulation d'ordre un, il suffit de stocker la solution à l'instant  $n$ . Avec un schéma AB global, il faut en plus stocker la solution aux instants  $n-1$ ,  $n-2$  et  $n-3$ . En utilisant la technique de pas de temps local, il faut encore stocker la solution en  $n-1/m$ ,  $n-2/m$  et  $n-3/m$ , ce qui multiplie par six la quantité d'inconnues à stocker si on compare avec le schéma Leap-Frog. Cette solution n'est donc pas envisageable pour des problèmes d'imagerie sismique. Nous avons donc proposé une nouvelle méthode, ne nécessitant pas de stocker la solution aux instants précédents, et reposant sur le schéma de Störmer-Verlet qui est une généralisation du schéma Leap-Frog.

### 6.3 Utilisation du schéma de Störmer-Verlet

Les deux méthodes que nous avons présentées précédemment nécessitent de connaître toute la solution à l'instant  $t_n$ . Nous ne pouvons donc pas les adapter facilement aux schémas Leap-frog pour les formulations du premier ordre. En effet, dans ce cas, une partie de la solution (par exemple la pression) est calculée aux instants  $t_n$  et l'autre (par exemple la vitesse) est déterminée aux instants  $t_{n+\frac{1}{2}}$ . C'est pour cela que nous nous sommes tournés vers le schéma de Störmer-Verlet (SV) qui est équivalent au schéma Leap-Frog tout en permettant de calculer les deux inconnues aux instants  $t_n$ .

### 6.3.1 Présentation de la méthode

Nous rappelons ci-dessous l'expression du schéma semi-discrétisé que nous considérons :

$$(6.3.1) \quad \begin{cases} \mathcal{M}_p \frac{d\mathbf{P}}{dt} = -\mathcal{K}^T \mathbf{V}, \\ \mathcal{M}_v \frac{d\mathbf{V}}{dt} = \mathcal{K} \mathbf{P}. \end{cases}$$

Les matrices de masse  $\mathcal{M}_p$  et  $\mathcal{M}_v$  étant diagonales par blocs, nous pouvons simplifier le système après inversion. On obtient :

$$(6.3.2) \quad \begin{cases} \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V}, \\ \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P} \end{cases}$$

avec  $\mathcal{A}_v = -\mathcal{M}_p^{-1} \mathcal{K}^T$  et  $\mathcal{A}_p = \mathcal{M}_v^{-1} \mathcal{K}$ .

Le schéma de Störmer-Verlet de pas de temps  $\Delta t$  s'écrit

$$(6.3.3) \quad \begin{cases} \frac{\mathbf{V}^{n+\frac{1}{2}} - \mathbf{V}^n}{\Delta t/2} = \mathcal{A}_p \mathbf{P}^n, \\ \frac{\mathbf{P}^{n+1} - \mathbf{P}^n}{\Delta t} = \mathcal{A}_v \mathbf{V}^{n+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n+1} - \mathbf{V}^{n+\frac{1}{2}}}{\Delta t/2} = \mathcal{A}_p \mathbf{P}^{n+1}. \end{cases}$$

Dans les équations (6.3.3),  $\mathbf{P}^n$  désigne le vecteur pression calculé à l'instant  $t_n = n\Delta t$  et  $\mathbf{V}^n$  désigne le vecteur vitesse calculé lui aussi à l'instant  $t_n$ . On peut montrer que ce schéma est équivalent au schéma Leap-Frog classique. On s'en convainc immédiatement pour le calcul de la pression qui s'exprime directement à l'instant  $t_{n+1}$  en fonction de l'instant  $t_n$ . Pour le calcul de la vitesse, on est convaincu en additionnant les lignes 1 et 3 de (6.3.3).

Comme à la section précédente, nous allons construire deux schémas à pas de temps global  $\Delta\tau$ , le premier étant équivalent à un schéma SV à pas de temps global  $\Delta\tau$ , le second à un schéma SV à pas de temps global  $\Delta t$ .

Le premier schéma est obtenu très simplement en écrivant l'équation différentielle ordinaire vérifiée par  $\mathbf{P}$  et  $\mathbf{V}$  sur un pas de temps  $\Delta t$ , puis en discrétisant cette équation par un schéma SV à pas de temps  $\Delta\tau$ . On obtient alors successivement

$$(6.3.4) \quad \begin{cases} \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V}(t), & \text{pour } t \in ]t_n, t_{n+1}[ \\ \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}(t), & \text{pour } t \in ]t_n, t_{n+1}[ \\ \mathbf{P}(t_n) = \mathbf{P}^n, \quad \mathbf{V}(t_n) = \mathbf{V}^n \end{cases}$$

et

$$(6.3.5) \quad \left\{ \begin{array}{l} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,i+1}. \end{array} \right.$$

pour  $i = 0..m-1$ , où  $\mathbf{P}^{n,i}$  et  $\mathbf{V}^{n,i}$  désignent les approximations respectives de  $\mathbf{P}$  et  $\mathbf{V}$  au temps  $t_{n,i} = n\Delta t + i\Delta\tau$ . On vérifie facilement que ce schéma est équivalent au schéma de Störmer-Verlet de pas de temps global  $\Delta\tau$ .

Le second schéma est plus compliqué à obtenir. On écrit à nouveau l'équation différentielle ordinaire vérifiée par  $\mathbf{P}$  et  $\mathbf{V}$  sur un pas de temps  $\Delta t$ , puis on approche les quantités  $\mathbf{V}(t)$  et  $\mathbf{P}(t)$  du second membre par des constantes. L'idée la plus simple est d'approcher  $\mathbf{V}(t)$  par  $\mathbf{V}(t_{n+1/2})$  si  $t \in ]t_n, t_{n+1}[$ , et  $\mathbf{P}(t)$  par  $\mathbf{P}(t_n)$  si  $t \in ]t_n, t_{n+1/2}[$  et par  $\mathbf{P}(t_{n+1})$  si  $t \in ]t_{n+1/2}, t_{n+1}[$ . Le système à résoudre s'écrit alors

$$(6.3.6) \quad \left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}^n, \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \mathbf{V}(t_n) = \mathbf{V}^n \end{array} \right. \\ \left\{ \begin{array}{l} \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V}^{n+1/2}, \quad \text{pour } t \in ]t_n, t_{n+1}[ \\ \mathbf{P}(t_n) = \mathbf{P}^n \end{array} \right. \\ \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}^{n+1}, \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \mathbf{V}(t_{n+1/2}) = \mathbf{V}^{n+1/2} \end{array} \right. \end{array} \right.$$

L'inconvénient de cette formulation est qu'elle ne permet plus de calculer  $\mathbf{V}$  et  $\mathbf{P}$  alternativement. Il faut en effet connaître  $\mathbf{V}^{n+1/2}$  pour résoudre la deuxième équation et  $\mathbf{P}^{n+1}$  pour résoudre la troisième. Si nous voulons utiliser un schéma de pas de temps  $\Delta t/m$ , nous devons donc d'abord utiliser  $m/2$  fois le premier système, sans actualiser  $\mathbf{P}$ , puis  $m$  fois le deuxième système, sans actualiser  $\mathbf{V}$  et enfin  $m/2$  fois le troisième système, sans actualiser  $\mathbf{P}$ . De plus, si  $m$  est impair, nous ne pourrions jamais calculer  $\mathbf{V}^{n+1/2}$  : nous n'aurons accès qu'à  $\mathbf{V}^{n+(m-1)/2m}$  et  $\mathbf{V}^{n+(m+1)/2m}$ .

Nous choisissons donc d'effectuer deux approximations supplémentaires, à l'aide de développements de Taylor :

$$\mathbf{V}(t_{n+1/2}) \approx \mathbf{V}(t_n) + \frac{\Delta t}{2} \frac{d\mathbf{V}}{dt}(t_n) = \mathbf{V}(t_n) + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}(t_n)$$

et

$$\mathbf{P}(t_{n+1}) \approx \mathbf{P}(t_n) + \Delta t \frac{d\mathbf{P}}{dt}(t_n) = \mathbf{P}(t_n) + \Delta t \mathcal{A}_v \mathbf{V}(t_{n+1/2}) \approx \mathbf{P}(t_n) + \Delta t \mathcal{A}_v \left( \mathbf{V}(t_n) + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}(t_n) \right).$$

Le nouveau système peut alors s'écrire sous la forme

$$(6.3.7) \quad \left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}^n, \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^n \right), \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \mathbf{V}(t_n) = \mathbf{V}^n, \quad \mathbf{P}(t_n) = \mathbf{P}^n \end{array} \right. \\ \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \left( \mathbf{P}^n + \Delta t \mathcal{A}_v \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^n \right) \right), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^n \right), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \mathbf{V}(t_{n+1/2}) = \mathbf{V}^{n+1/2}, \quad \mathbf{P}(t_{n+1/2}) = \mathbf{P}^{n+1/2} \end{array} \right. \end{array} \right.$$

Nous choisissons maintenant de discrétiser chaque système par un schéma de Störmer-Verlet de pas de temps  $\Delta\tau$ . Nous allons donc effectuer  $m$  pas de Störmer-Verlet, la première moitié sera consacrée à la discrétisation du premier système, de  $]t_n, t_{n+1/2}[$ ; la seconde à la discrétisation du second système, de  $]t_{n+1/2}, t_{n+1}[$ . Bien évidemment, dans le cas où  $m$  est impair, nous devons faire un traitement particulier pour le pas de temps du milieu. Il nous faut donc distinguer deux cas :

— **Si  $m$  est pair**, les  $m/2$  premiers pas de temps sont consacrés au premier système, les  $m/2$  suivants au second. L'algorithme s'écrit :

$$— \mathbf{W}_1 = \mathbf{V}^n + \frac{\Delta\tau}{2} \mathcal{A}_v \mathbf{P}^n, \quad \mathbf{W}_2 = \mathbf{P}^n + \Delta t \mathcal{A}_v \left( \mathbf{V}^n + \frac{\Delta\tau}{2} \mathcal{A}_p \mathbf{P}^n \right);$$

$$— \text{pour } i = 0, \dots, m/2 - 1 : \left\{ \begin{array}{l} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,0}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{W}_1, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,0}. \end{array} \right.$$

$$— \text{pour } i = m/2, \dots, m - 1 : \left\{ \begin{array}{l} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{W}_2, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{W}_1, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{W}_2. \end{array} \right.$$

— **Si  $m$  est impair**. Les  $(m-1)/2$  premiers pas de temps sont consacrés au premier système, les  $(m-1)/2$  derniers au second et le pas de temps  $(m+1)/2$  assure la transition entre les deux systèmes. L'algorithme écrit :

$$— \mathbf{W}_1 = \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^n, \quad \mathbf{W}_2 = \mathbf{P}^n + \Delta t \mathcal{A}_p \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^n \right);$$

$$\text{— pour } m = 0, \dots, (l-1)/2 - 1 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,0}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{W}_1, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,0}. \end{cases}$$

$$\text{— pour } m = (l-1)/2 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{P}^{n,0}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{W}_1, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{W}_2. \end{cases}$$

$$\text{— pour } m = (l-1)/2 + 1, \dots, l-1 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{W}_2, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v \mathbf{W}_1, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p \mathbf{W}_2. \end{cases}$$

On vérifie, après quelques calculs, que ce schéma est bien équivalent à un schéma SV de pas de temps  $\Delta t$ , que  $m$  soit pair ou impair.

Il nous reste maintenant à combiner les deux schémas précédents, comme dans le cas de la formulation du second ordre, pour obtenir un schéma à pas de temps local. Pour cela, on considère tout d'abord le système

$$(6.3.8) \quad \left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}(t), \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V}(t), \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \mathbf{V}(t_n) = \mathbf{V}^n, \quad \mathbf{P}(t_n) = \mathbf{P}^n \end{array} \right. \\ \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p \mathbf{P}(t), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v \mathbf{V}(t), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \mathbf{V}(t_{n+1/2}) = \mathbf{V}^{n+1/2}, \quad \mathbf{P}(t_{n+1/2}) = \mathbf{P}^{n+1/2} \end{array} \right. \end{array} \right.$$

On décompose ensuite les inconnues  $\mathbf{P}$  et  $\mathbf{V}$  dans les seconds membres sous la forme

$$(6.3.9) \quad \begin{cases} \mathbf{V}(t) = (\mathcal{I} - \mathcal{P})\mathbf{V}(t) + \mathcal{P}\mathbf{V}(t), \\ \mathbf{P}(t) = (\mathcal{I} - \mathcal{P})\mathbf{P}(t) + \mathcal{P}\mathbf{P}(t), \end{cases}$$

où  $\mathcal{P}$  est la matrice de projection définie à la section précédente. L'étape suivante consiste à utiliser des approximations pour les inconnues des grandes mailles. En suivant la même démarche

que pour le schéma à pas de temps global, nous choisissons d'abord d'approcher  $(\mathcal{I} - \mathcal{P})\mathbf{V}(t)$  par  $(\mathcal{I} - \mathcal{P})\mathbf{V}(t_{n+1/2})$  pour  $t \in ]t_n, t_{n+1}[$ , et  $(\mathcal{I} - \mathcal{P})\mathbf{P}(t)$  par  $(\mathcal{I} - \mathcal{P})\mathbf{P}(t_n)$  pour  $t \in ]t_n, t_{n+1/2}[$  et par  $(\mathcal{I} - \mathcal{P})\mathbf{P}(t_{n+1})$  pour  $t \in ]t_{n+1/2}, t_{n+1}[$ . Dans un second temps, nous utilisons les approximations

$$(\mathcal{I} - \mathcal{P})\mathbf{V}(t_{n+1/2}) \approx (\mathcal{I} - \mathcal{P}) \left[ \mathbf{V}(t_n) + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}(t_n) \right]$$

et

$$(\mathcal{I} - \mathcal{P})\mathbf{P}(t_{n+1}) \approx (\mathcal{I} - \mathcal{P}) \left[ \mathbf{P}(t_n) + \Delta t \mathcal{A}_v \left( \mathbf{V}(t_n) + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}(t_n) \right) \right].$$

Le système d'équations devient alors

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^n + \mathcal{A}_p\mathbf{P}(t), \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v(\mathcal{I} - \mathcal{P}) \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^n \right) + \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{V}(t), \quad \text{pour } t \in ]t_n, t_{n+1/2}[ \\ \mathbf{V}(t_n) = \mathbf{V}^n, \quad \mathbf{P}(t_n) = \mathbf{P}^n \end{array} \right. \\ \left\{ \begin{array}{l} \frac{d\mathbf{V}}{dt} = \mathcal{A}_p(\mathcal{I} - \mathcal{P}) \left( \mathbf{P}^n + \Delta t \mathcal{A}_p \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^n \right) \right) + \mathcal{A}_p \mathcal{P} \mathbf{P}(t), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \frac{d\mathbf{P}}{dt} = \mathcal{A}_v(\mathcal{I} - \mathcal{P}) \left( \mathbf{V}^n + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^n \right) + \mathcal{A}_v \mathcal{P} \mathbf{V}(t), \quad \text{pour } t \in ]t_{n+1/2}, t_{n+1}[ \\ \mathbf{V}(t_{n+1/2}) = \mathbf{V}^{n+1/2}, \quad \mathbf{P}(t_{n+1/2}) = \mathbf{P}^{n+1/2} \end{array} \right. \end{array} \right.$$

L'algorithme de pas de temps local est finalement obtenu en discrétisant ce système par un schéma de Störmer-Verlet de pas temps  $\Delta\tau$ . Pour écrire cet algorithme, nous distinguons à nouveau deux cas.

— **Si  $m$  est pair**, les  $m/2$  premiers pas de temps sont consacrés au premier système, les  $m/2$  suivants au second. L'algorithme s'écrit :

$$— \mathbf{W}_1 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^{n,0} \right), \mathbf{W}_2 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{P}^{n,0} + \Delta t \mathcal{A}_v \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^{n,0} \right) \right);$$

$$— \text{pour } i = 0, \dots, m/2 - 1 : \left\{ \begin{array}{l} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p \mathcal{P} \mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v \mathcal{P} \mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p \mathcal{P} \mathbf{P}^{n,i+1}. \end{array} \right.$$

$$— \text{pour } i = m/2, \dots, m - 1 : \left\{ \begin{array}{l} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p \mathcal{P} \mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v \mathcal{P} \mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p \mathcal{P} \mathbf{P}^{n,i+1}. \end{array} \right.$$

— **Si  $m$  est impair**. Les  $(i-1)/2$  premiers pas de temps sont consacrés au premier système, les  $(i-1)/2$  derniers au second et l'itération  $(i+1)/2$  assure la transition entre les deux systèmes. L'algorithme s'écrit :



$$\begin{aligned}
& - \mathbf{W}_1 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^{n,0} \right), \mathbf{W}_2 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{P}^{n,0} + \Delta t \mathcal{A}_p \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^{n,0} \right) \right); \\
& - \text{pour } i = 0, \dots, (m-1)/2 - 1 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} & = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v\mathcal{P}\mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}. \end{cases} \\
& - \text{pour } i = (m-1)/2 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} & = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v\mathcal{P}\mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}. \end{cases} \\
& - \text{pour } i = (m-1)/2 + 1, \dots, m-1 : \begin{cases} \frac{\mathbf{V}^{n,i+\frac{1}{2}} - \mathbf{V}^{n,i}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i}, \\ \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} & = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v\mathcal{P}\mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,m+\frac{1}{2}}}{\Delta\tau/2} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}. \end{cases}
\end{aligned}$$

Comme pour le schéma de Störmer-Verlet global, nous pouvons réduire le nombre d'opérations à certains endroits de l'algorithme, pour se ramener à un schéma de type Leap-Frog. Nous le montrons dans le cas pair, le cas impair étant similaire et nous nous concentrons sur la première phase de l'algorithme (pour  $i = 0, \dots, (m-1)/2 - 1$ ). À l'instant  $i + 1$ , la dernière étape de la boucle s'écrit

$$\frac{\mathbf{V}^{n,i+1} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1},$$

et à l'instant  $i + 1$ , la première étape de la boucle s'écrit

$$\frac{\mathbf{V}^{n,i+\frac{3}{2}} - \mathbf{V}^{n,i+1}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}.$$

En additionnant ces deux équations, nous obtenons

$$\frac{\mathbf{V}^{n,i+\frac{3}{2}} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}.$$

Nous pouvons procéder de la même manière pour l'étape deux, et l'algorithme devient

$$\begin{aligned}
& - \mathbf{W}_1 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_v \mathbf{P}^{n,0} \right), \mathbf{W}_2 = (\mathcal{I} - \mathcal{P}) \left( \mathbf{P}^{n,0} + \Delta t \mathcal{A}_v \left( \mathbf{V}^{n,0} + \frac{\Delta t}{2} \mathcal{A}_p \mathbf{P}^{n,0} \right) \right); \\
& - \frac{\mathbf{V}^{n,\frac{1}{2}} - \mathbf{V}^{n,0}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,0} = \mathcal{A}_p\mathbf{P}^{n,0}
\end{aligned}$$

$$\begin{aligned}
& \text{— pour } i = 1, \dots, m/2 - 1 : \begin{cases} \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} & = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v\mathcal{P}\mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+\frac{3}{2}} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}. \end{cases} \\
& \text{— } \frac{\mathbf{V}^{n,m/2} - \mathbf{V}^{n,m/2-\frac{1}{2}}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{P}^{n,0} + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,m/2} \\
& \text{— } \frac{\mathbf{V}^{n,m/2+\frac{1}{2}} - \mathbf{V}^{n,m/2}}{\Delta\tau/2} = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,m/2} \\
& \text{— pour } i = m/2 + 1, \dots, m - 1 : \begin{cases} \frac{\mathbf{P}^{n,i+1} - \mathbf{P}^{n,i}}{\Delta\tau} & = \mathcal{A}_v(\mathcal{I} - \mathcal{P})\mathbf{W}_1 + \mathcal{A}_v\mathcal{P}\mathbf{V}^{n,i+\frac{1}{2}}, \\ \frac{\mathbf{V}^{n,i+\frac{3}{2}} - \mathbf{V}^{n,i+\frac{1}{2}}}{\Delta\tau} & = \mathcal{A}_p(\mathcal{I} - \mathcal{P})\mathbf{W}_2 + \mathcal{A}_p\mathcal{P}\mathbf{P}^{n,i+1}. \end{cases}
\end{aligned}$$

Remarquons que nous pouvons encore économiser une opération en regroupant les calculs de  $\mathbf{V}^{n,m/2}$  et de  $\mathbf{V}^{n,m/2+\frac{1}{2}}$  entre les deux phases de l'algorithme.

Nous allons maintenant étudier le comportement de cet algorithme sur des expériences numériques en dimension 1.

## 6.4 Résultats numériques

Dans cette section, nous testons numériquement la stabilité des schémas à pas de temps locaux SV et AB. Le domaine de calcul est un segment de longueur 6m divisé en trois parties de 2m. La partie centrale est maillée avec des mailles de longueur  $h = 0.1\text{m}$ . Les deux autres parties sont maillées avec des mailles de longueur  $2h = 0.2\text{m}$  (voir figure 6.2). Nous imposons des conditions de bord périodiques et  $c = 1\text{m}\cdot\text{s}^{-1}$ . Les conditions initiales sont

$$p_0(x) = (x - x_0)e^{-\frac{4(x-x_0)^2}{\pi^2 r_0^2}}, \quad v_0(x) = (x - x_0)e^{-\frac{4(x-x_0)^2}{\pi^2 r_0^2}}.$$

Ainsi, la solution exacte est donnée par

$$p_{ex}(x, t) = \sum_{i=0}^{+\infty} p_0(x - 6i - t).$$

Nous choisissons  $x_0 = 2$  et  $r_0 = 2$ .

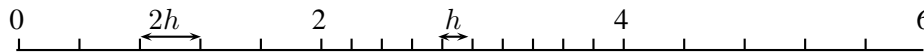


FIGURE 6.2 – Le domaine de calcul

### 6.4.1 Le schéma SV

Nous testons dans un premier temps le schéma SV à pas de temps local. Pour fixer le pas de temps, nous utilisons la condition CFL du schéma global est  $\Delta t \leq 0.03h$  et nous choisissons  $m = 2$ .

Nous représentons sur la figure 6.3.(a) l'évolution de l'erreur  $L^2$  en espace et en temps après 40s d'expérience, ce qui laisse le temps à l'onde d'effectuer environ 7 fois le tour du domaine. À titre de comparaison, nous représentons l'erreur obtenue avec les schéma SV à pas de temps global sur la figure 6.3.(b). La solution à pas de temps local semble évoluer normalement et l'erreur est meilleure que la solution à pas de temps global. Cependant, si nous regardons l'erreur en temps plus long (après environ 15 tours du domaine, la solution du schéma à pas de temps local est clairement instable (voir Figs. 6.3.(c) et 6.3.(d)).

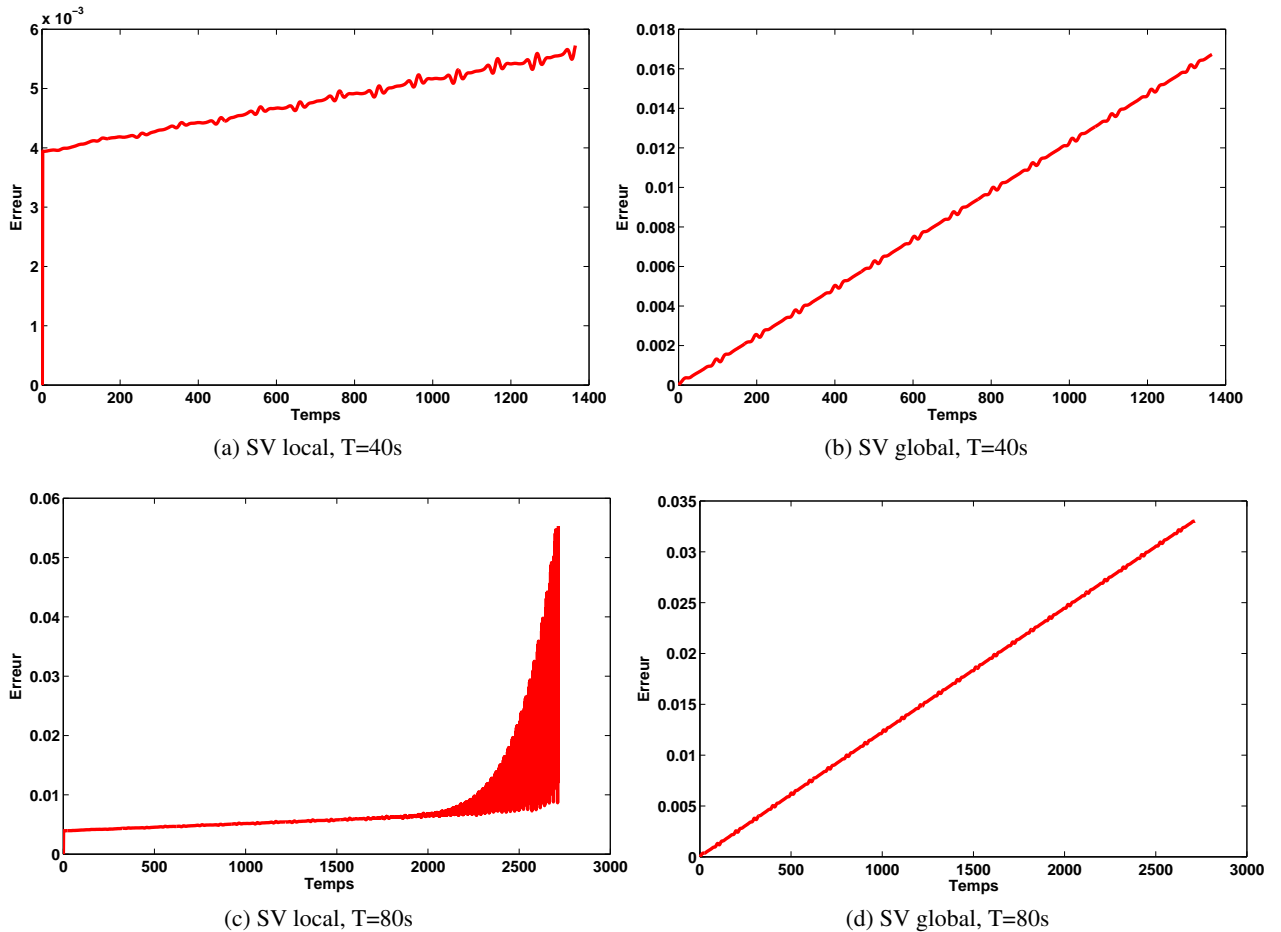


FIGURE 6.3 – Erreur au cours du temps pour les schémas SV à pas de temps local et global.

Nous avons testé notre schéma avec des pas de temps plus faibles, mais nous n'avons pas pu supprimer l'instabilité. Pour comprendre ce phénomène, il faut calculer numériquement la CFL du schéma. Pour cela, nous remarquons que notre schéma peut se mettre sous la forme  $\mathbf{Y}^{n+1} = \mathbf{Y}^n + \Delta t \mathcal{A} \mathbf{Y}^n = (\mathcal{I} + \Delta t \mathcal{A}) \mathbf{Y}^n$ , avec  $\mathbf{Y} = \begin{bmatrix} \mathbf{P}^n \\ \mathbf{P}^n \end{bmatrix}$ . Il n'est pas nécessaire de déterminer l'expression de la matrice  $\mathcal{A}$  analytiquement. En effet, l'algorithme de pas de temps local prend en entrée un vecteur  $\mathbf{Y}^n$  et renvoie en sortie le vecteur  $\mathbf{Y}^{n+1} = (\mathcal{I} + \Delta t \mathcal{A}) \mathbf{Y}^n$ . Si nous remplaçons le vecteur  $\mathbf{Y}^n$  par la matrice identité, nous récupérons donc en sortie la matrice  $(\mathcal{I} + \Delta t \mathcal{A}) \mathbf{Y}^n$ . Pour déterminer la stabilité du schéma, il nous faut vérifier que les valeurs propres de cette matrice ont toutes un module plus petit que 1. Nous représentons sur la figure 6.4.(a) le module de la plus grande valeur propre en fonction de

$\Delta t$ . Nous représentons le résultat obtenu pour le schéma à pas de temps global sur la figure 6.4.(b) À première vue, les valeurs propres restent bien de module inférieur à 1 jusqu'à  $\Delta t = 0.03s$ , cependant, si on zoome autour de 1 (figure 6.4.(c) et 6.4.(d)), on s'aperçoit que la valeur propre maximale du schéma à pas de temps local est légèrement plus grande que 1, ce qui explique les explosions en temps long que nous avons observées.

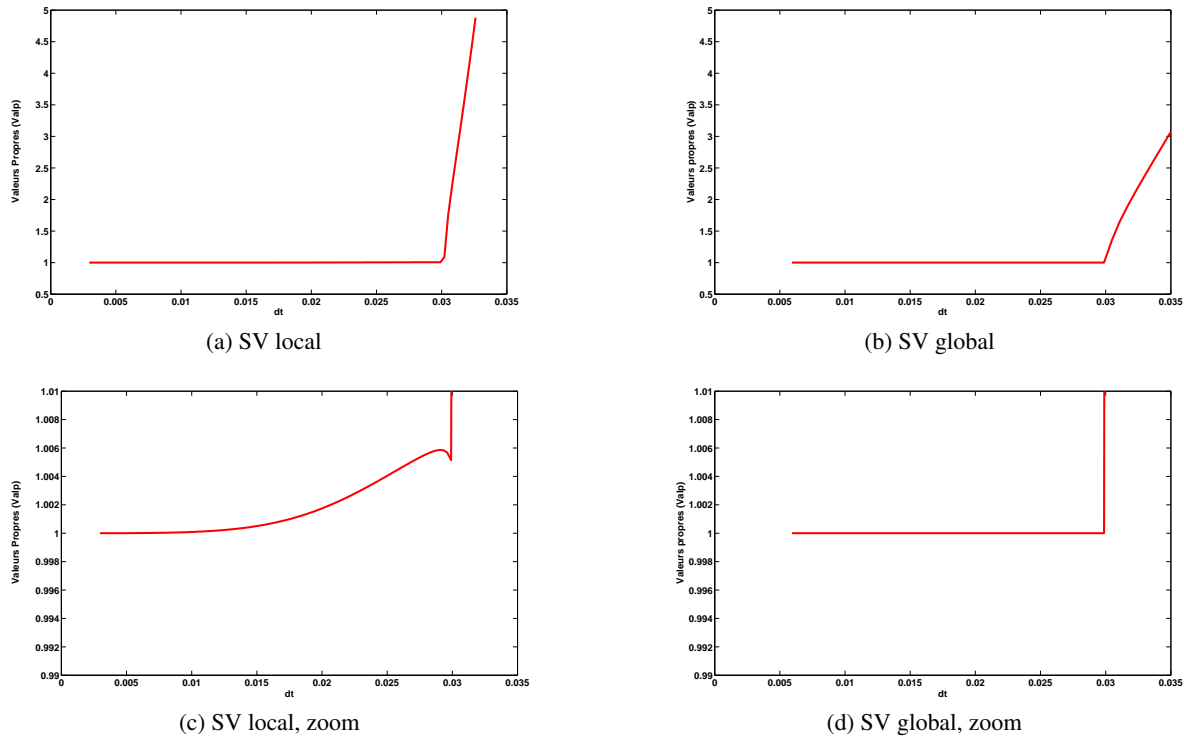


FIGURE 6.4 – Module de la plus grande valeur propre de  $(I + \Delta tA)$  en fonction de  $\Delta t$

## 6.4.2 Le schéma AB

Nous avons mené des expériences similaires pour le schéma AB et nous n'avons pas observé d'instabilités. Les erreurs obtenues pour les schémas à pas de temps local et global sont présentées sur la figure 6.5. Pour étudier la stabilité des schémas d'Adams-Bashworth, nous avons procédé comme à la section précédente, en remarquant que ce schéma pouvait se mettre sous la forme  $\mathbf{W}^{n+1} = (\mathcal{I} + \Delta tA)\mathbf{Y}^n$ , avec  $\mathbf{W}^n = [\mathbf{Y}^n, \mathbf{Y}^{n-1}, \mathbf{Y}^{n-2}, \mathbf{Y}^{n-1/m}, \mathbf{Y}^{n-2/m}, \mathbf{Y}^{n-3/m}]^T$ . Il suffit alors de remplacer successivement chaque vecteur par la matrice identité dans l'algorithme à pas de temps local pour construire la matrice  $\mathcal{I} + \Delta tA$ . Nous représentons sur la figure 6.6 l'évolution de la plus grande des valeurs propres de cette matrice en fonction de  $\Delta t$ . Le module reste bien inférieur à 1 pour tout  $\Delta t \leq 0.065$ , ce qui est environ cinq fois plus faible que pour le schéma SV, et représente un surcoût considérable. Nous avons également testé ces schémas en introduisant des conditions absorbantes et en inversant les régions fines et grossières (i.e. en mettant les régions grossières à l'extérieur), nous n'avons pas non plus observé d'instabilités. Néanmoins, en plus des problèmes liés à l'augmentation des coûts de calculs dans un cadre RTM, la condition CFL des schémas, cinq fois plus petite que celle du Leap-Frog, ne permet pas d'envisager une utilisation dans un code d'imagerie sismique.

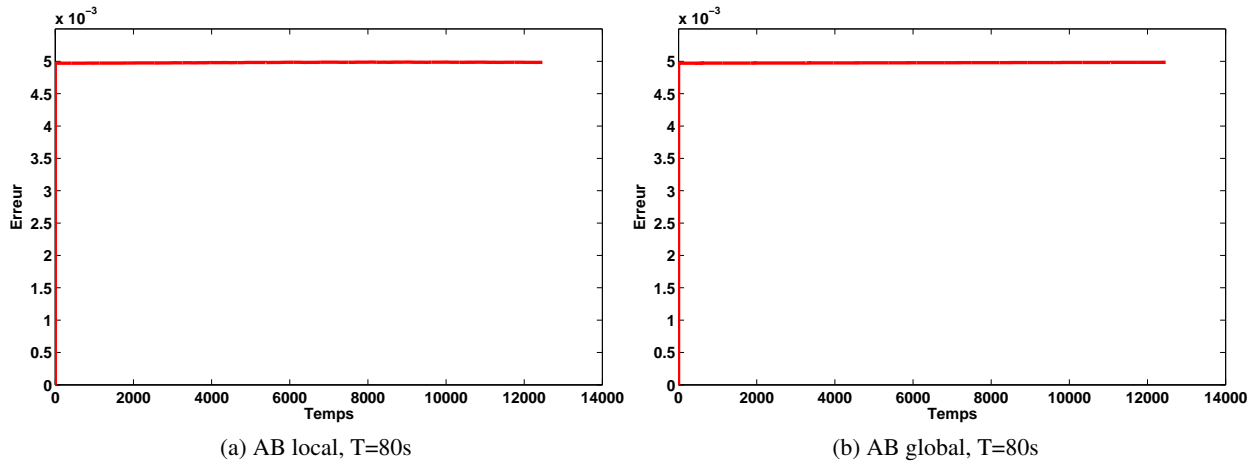
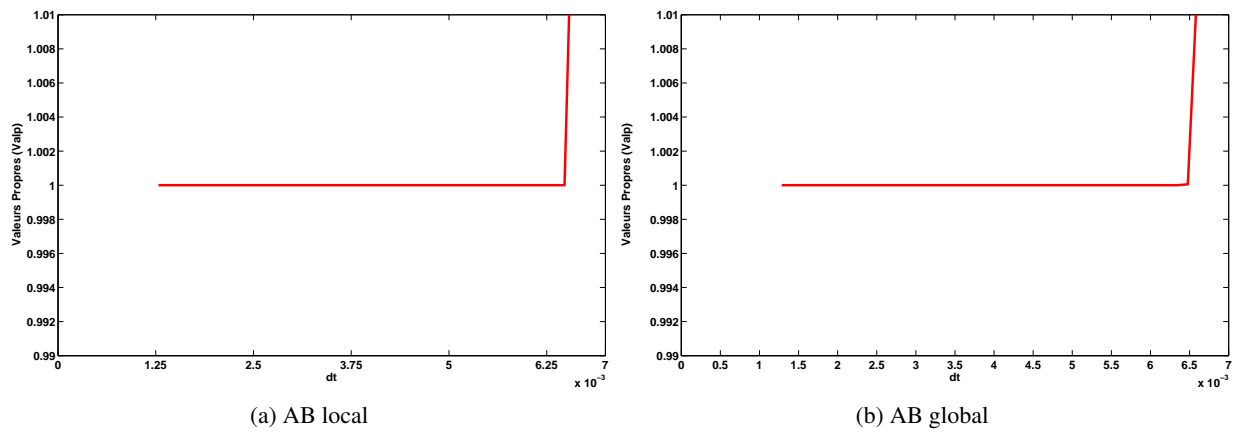


FIGURE 6.5 – Erreur au cours du temps pour les schémas AB à pas de temps local et global.

FIGURE 6.6 – Module de la plus grande valeur propre de  $(I + \Delta t A)$  en fonction de  $\Delta t$

## 6.5 Conclusion

Le schéma à pas de temps local d'Adams-Bashforth est stable et permet d'atteindre les conditions de stabilité optimales à la fois sur les maillages fins et grossiers. Cependant, sa mise en œuvre pour des applications géophysiques est irréaliste car c'est un schéma multipas qui nécessite le stockage de nombreuses inconnues auxiliaires. Nous avons donc proposé une technique de pas de temps local basée sur le schéma de Störmer-Verlet et permettant de limiter le stockage des inconnues. Ce schéma est efficace en temps court, mais se révèle malheureusement inconditionnellement instable en temps long. La mise au point d'une technique efficace de pas de temps local pour l'imagerie sismique reste donc un problème ouvert. Il faudrait probablement se tourner plutôt vers des techniques combinant des schémas explicites sur les maillages grossiers et implicites sur les maillages fins.



# Bibliographie

- [1] C. AGUT : *Schémas numériques d'ordre élevé en temps et en espace pour l'équation des ondes*. Thèse, Université de Pau et des Pays de l'Adour, décembre 2011.
- [2] C. AGUT et J. DIAZ : New high order schemes based on the modified equation technique for solving the wave equation. Rapport de recherche RR-7331, INRIA, juillet 2010.
- [3] C. AGUT et J. DIAZ : Stability analysis of the interior penalty discontinuous Galerkin method for the wave equation. *ESAIM : Mathematical Modelling and Numerical Analysis*, 2012. To appear.
- [4] M. AINSWORTH, P. MONK et W. MUNIZ : Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *Journal of Scientific Computing*, 27(1-3):5–40, 2006.
- [5] D. ARNOLD : An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, Vol. 19, No. 4 : pp. 742-760, 1982.
- [6] D. ARNOLD, F. BREZZI, B. COCKBURN et L. MARINI : Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779 (electronic), 2001/02.
- [7] C. BALDASSARI, H. BARUCQ, H. CALANDRA, B. DENEL et J. DIAZ : High-order discontinuous Galerkin method for the reverse time migration. *Communications in Computational Physics*, 11(2):660–673, janvier 2012.
- [8] C. BALDASSARI : *Modélisation et simulation numérique pour la migration terrestre par équation d'ondes*. Thèse, Université de Pau et des Pays de l'Adour, décembre 2009.
- [9] C. BALDASSARI, H. BARUCQ, H. CALANDRA et J. DIAZ : Numerical performances of a hybrid local-time stepping strategy applied to the reverse time migration. *Geophysical Prospecting*, 59(5):907–919, 2011.
- [10] E. BAYSAL, D. D. KOSLOFF et J. W. SHERWOOD : Reverse time migration. *Geophysics*, 48(11):1514–1524, 1983.
- [11] E. BÉCACHE, P. JOLY et J. RODRÍGUEZ : Space-time mesh refinement for elastodynamics. Numerical results. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):355–366, 2005.
- [12] M. BERNACKI, S. LANTERI et S. PIPERNO : Time-domain parallel simulation of heterogeneous wave propagation on unstructured grids using explicit, nondiffusive, discontinuous Galerkin methods. *J. Comput. Acoust.*, 14(1):57–81, 2006.
- [13] E. CASAROTTI, M. STUPAZZINI, S. LEE, D. KOMATITSCH, A. PIERSANTI et J. TROMP : CUBIT and seismic wave propagation based upon the spectral-element method : An advanced unstructured mesher for complex 3D geological media. *In Proceedings of the 16th International Meshing Roundtable*, volume 5B.4, pages 579–597, 2008.



- [14] E. T. CHUNG et B. ENGQUIST : Optimal discontinuous Galerkin methods for wave propagation. *SIAM J. Numer. Anal.*, 44(5):2131–2158, 2006.
- [15] E. T. CHUNG et B. ENGQUIST : Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions. *SIAM J. Numer. Anal.*, 47(5):3820–3848, 2009.
- [16] P. G. CIARLET : *The finite element method for elliptic problems*. Access Online via Elsevier, 1978.
- [17] J. F. CLAERBOUT : Imaging the Earth’s interior. *Blackwell Scientific Publications, Oxford, London, Edinburgh*, 399 pp., 1985.
- [18] B. COCKBURN et C. SHU : The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, Vol. 35, No. 6 : pp. 2440-2463, 1998.
- [19] F. COLLINO, T. FOUQUET et P. JOLY : A conservative space-time mesh refinement method for the 1-D wave equation. I. Construction. *Numer. Math.*, 95(2):197–221, 2003.
- [20] F. COLLINO, T. FOUQUET et P. JOLY : A conservative space-time mesh refinement method for the 1-D wave equation. II. Analysis. *Numer. Math.*, 95(2):223–251, 2003.
- [21] F. COLLINO, T. FOUQUET et P. JOLY : Conservative space-time mesh refinement methods for the FDTD solution of Maxwell’s equations. *J. Comput. Phys.*, 211(1):9–35, 2006.
- [22] M. A. DABLAIN : The application of high order differencing for the scalar wave equation. *Geophysics*, 51(1):54–66, 1986.
- [23] M. A. DABLAIN : High order differencing for the scalar wave equation. *SEG Technical Program Expanded Abstracts*, 3(1):854–854, 1984.
- [24] S. DELCOURTE, L. FEZOUÏ et N. GLINSKY-OLIVIER : A high-order discontinuous Galerkin method for the seismic wave propagation. *ESAIM*, 2009.
- [25] J. DIAZ et M. J. GROTE : Energy conserving local time stepping for second order wave equations. *SIAM*, 31:1985–2014, 2009.
- [26] J. DIAZ et M. GROTE : Energy conserving explicit local time stepping for second-order wave equations. *SIAM J. Sci. Comput.*, 31(3):1985–2014, 2009.
- [27] J. DIAZ et P. JOLY : Robust high order non-conforming finite element formulation for time domain fluid-structure interaction. *Journal of Computational Acoustics*, 13(3):430–431, 2006.
- [28] J. DOUGLAS, Jr. et T. DUPONT : Interior penalty procedures for elliptic and parabolic Galerkin methods. *In Computing methods in applied sciences (Second Internat. Sympos., Versailles, 1975)*, pages 207–216. Lecture Notes in Phys., Vol. 58. Springer, Berlin, 1976.
- [29] M. DUMBSER et C.-D. MUNZ : ADER discontinuous Galerkin schemes for aeroacoustics. *Comptes Rendus Mécanique, Vol 333, No. 9, pp. 683-687*, 2004.
- [30] M. DUMBSER, M. KÄSER et E. F. TORO : An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - V. local time stepping and p-adaptivity. *Geophysical Journal International*, 171(2):695–717, 2007.
- [31] M. DUMBSER et C.-D. MUNZ : {ADER} discontinuous Galerkin schemes for aeroacoustics. *Comptes Rendus Mécanique*, 333(9):683 – 687, 2005.
- [32] H. FAHS et S. LANTERI : A high-order non-conforming discontinuous Galerkin method for time-domain electromagnetics. *Journal of Computational and Applied Mathematics*, 234(4):1088 – 1096, 2010. Proceedings of the Thirteenth International Congress on Computational and Applied Mathematics (ICCAM-2008), Ghent, Belgium, 7-11 July, 2008.

- [33] T. FEI et K. LARNER : Elimination of numerical dispersion in finite-difference modeling and migration by flux-corrected transport. *Geophysics*, 60(6):1830–1842, 1995.
- [34] L. FEZOUÏ, S. LANTERI, S. LOHRENGEL et S. PIPERNO : Convergence and stability of a discontinuous Galerkin time-domain method for the 3D heterogeneous Maxwell equations on unstructured meshes. *ESAIM : Mathematical Modelling and Numerical Analysis*, 39(06):1149–1176, 2005.
- [35] B. FORNBERG : High-order finite differences and the pseudospectral method on staggered grids. *SIAM J. Numer. Anal.*, 27(4):904–918, 1990.
- [36] A. GRIEWANK et A. WALTHER : Algorithm 799 : revolve : an implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Transactions on Mathematical Software (TOMS)*, 26(1):19–45, 2000.
- [37] M. J. GROTE et T. MITKOVA : High-order explicit local time-stepping methods for damped wave equations. *Journal of Computational and Applied Mathematics*, 239(0):270 – 289, 2013.
- [38] M. J. GROTE, A. SCHNEEBELI et D. SCHÖTZAU : Interior penalty discontinuous Galerkin method for Maxwell’s equations : Energy norm error estimates. *Journal of Computational and Applied Mathematics*, 204(2):375 – 386, 2007. Special Issue : The Seventh International Conference on Mathematical and Numerical Aspects of Waves (WAVES’05).
- [39] M. J. GROTE, A. SCHNEEBELI et D. SCHÖTZAU : Interior penalty discontinuous Galerkin method for Maxwell’s equations : optimal  $L^2$ -norm error estimates. *IMA J. Numer. Anal.*, 28(3): 440–468, 2008. submitted.
- [40] M. GROTE et T. MITKOVA : Explicit local time-stepping for Maxwell’s equations. *J. Comput. App. Math.*, 34, 3283–3302, 2010.
- [41] M. GROTE, A. SCHNEEBELI et D. SCHÖTZAU : Discontinuous Galerkin finite element method for the wave equation. *SIAM J. Numer. Anal.*, 44(6):2408–2431 (electronic), 2006.
- [42] M. GROTE et D. SCHÖTZAU : Optimal error estimates for the fully discrete interior penalty DG method for the wave equation. *J. Sc. Computing*, 40:257–272, 2009.
- [43] L. GUILLOT, Y. CAPDEVILLE et J.-J. MARIGO : 2-D non-periodic homogenization of the elastic wave equation : SH case. *Geophysical Journal International*, 182(3):1438–1454, 2010.
- [44] E. HAIRER, S. NØRSETT et G. WANNER : Solving ordinary differential equations I : Nonstiff problems. *Springer*, 2000.
- [45] O. HOLBERG : *Large-scale wave equation computations using high-order difference operators*. Thèse de doctorat, University Microfilms International, 1987.
- [46] H. IGEL, P. MORA et B. RIOLLET : Anisotropic wave propagation through finite-difference grids. *Geophysics*, 60(4):1203–1216, 1995.
- [47] L. JAMESON : High order schemes for resolving waves : Number of points per wavelength, Explicit local time-stepping for Maxwell’s equations. *J. Sci. Comput.*, Vol 15, pp. 417-439, 2000.
- [48] L. JAMESON : AMR vs high order schemes. *J. Sci. Comput.*, Vol 18, pp. 1-24, 2003.
- [49] P. JOLY et J. RODRÍGUEZ : An error analysis of conservative space-time mesh refinement methods for the one-dimensional wave equation. *SIAM J. Numer. Anal.*, 43(2):825–859 (electronic), 2005.
- [50] M. KÄSER et M. DUMBSER : An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - I. the two-dimensional isotropic case with external source terms. *Geophysical Journal International*, 166(2):855–877, 2006.

- [51] D. KOMATITSCH, C. BARNES et J. TROMP : Wave propagation near a fluid-solid interface : a spectral element approach. *Geophysics*, 65(2):623–631, 2000.
- [52] D. KOMATITSCH : *Méthodes spectrales et éléments spectraux pour l'équation de l'élastodynamique 2D et 3D en milieu hétérogène (Spectral and spectral-element methods for the 2D and 3D elastodynamics equations in heterogeneous media)*. Thèse de doctorat, Institut de Physique du Globe, Paris, France, May 1997. 187 pages.
- [53] Y. KONDDOHA, Y. HOSAKA et K. ISHII : Kernel optimum nearly analytical discretization algorithm applied to parabolic and hyperbolic equations. *Comput. Math. Appl*, 1994.
- [54] P. LAX et B. WENDROFF : Difference schemes for hyperbolic equations with high-order of accuracy. *Comm. Pure Appl. Math.*, Vol 17, pp. 381-398, 1964.
- [55] B. LOMBARD : Modélisation numérique de la propagation et de la diffraction d'ondes mécaniques. 2010.
- [56] H. MIZUTANI, R. J. GELLER et N. TAKEUCHI : Comparison of accuracy and efficiency of time-domain schemes for calculating synthetic seismograms. *Physics of the Earth and Planetary Interiors*, 119(1):75–97, 2000.
- [57] S. PIPERNO : Symplectic local time-stepping in non-dissipative DGTD methods applied to wave propagation problems. *M2AN Math. Model. Numer. Anal.*, 40(5):815–841, 2006.
- [58] A. SEI et W. SYMES : Error analysis of numerical schemes for the wave equation in heterogeneous media. *CRPC-TR93536*, 1993.
- [59] K. SHAHBAZI : An explicit expression for the penalty parameter of the interior penalty method. *Journal of Computational Physics*, 205(2):401 – 407, 2005.
- [60] G. R. SHUBIN et J. B. BELL : A modified equation approach to constructing fourth-order methods for acoustic wave propagation. *SIAM J. Sci. Statist. Comput.*, 8:135–151, 1987.
- [61] L. VERLET : Computer experiments on classical fluids. I. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.
- [62] J. VIRIEUX : P-SV wave propagation in heterogeneous media : Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901, 1986.
- [63] D. YANG, J. TENG, Z. ZHANG et E. LIU : Finite-difference modelling in two-dimensional anisotropic media using a flux-corrected transport technique. *Geophys. J. Int. (2002) 148*, 320-328, 2002.
- [64] D. YANG, J. TENG, Z. ZHANG et E. LIU : A nearly analytic discrete method for acoustic and elastic wave equations in anisotropic media. *Bulletin of the Seismological Society of America*, Vol. 93, No. 2, pp. 882–890, 2003.
- [65] D. YANG, J. PENG, M. LU et T. TERLAKY : Optimal nearly analytic discrete approximation to the scalar wave equation. *Bulletin of the Seismological Society of America*, 96(3):1114–1130, 2006.