



HAL
open science

Optimisation topologique en mécanique des fluides et analyse numérique de problèmes de propagation d'ondes en régime harmonique

Pierre-Henri Cocquet

► **To cite this version:**

Pierre-Henri Cocquet. Optimisation topologique en mécanique des fluides et analyse numérique de problèmes de propagation d'ondes en régime harmonique. Analyse numérique [math.NA]. Université de Pau et des Pays de l'Adour (UPPA), 2024. tel-04442272

HAL Id: tel-04442272

<https://hal.science/tel-04442272v1>

Submitted on 6 Feb 2024

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.



Habilitation à Diriger des Recherches

Université de Pau et des Pays de l'Adour

Spécialité : Mathématiques appliquées

Optimisation topologique en mécanique des fluides et analyse numérique de problèmes de propagation d'ondes en régime harmonique

par

Pierre-Henri Cocquet

soutenue le 25 janvier 2024 devant le jury composé de

Hélène Barucq	DR INRIA, Equipe Projet MAKUTU, Pau	examinatrice
Alain Bastide	Professeur, PIMENT, Université de la Réunion	examinateur
Franck Boyer	Professeur, IMT-UMR 5219, Université Toulouse III	rapporteur
Martin J. Gander	Professeur ordinaire, Section de mathématiques, Genève	examinateur
Christophe Geuzaine	Professeur ordinaire, Montefiore Institute, Liège	rapporteur
Florence Hubert	Professeur, I2M-UMR 7373, Université Aix Marseille	rapporteure
Laurent Pécastaing	Professeur, SIAME, UPPA	examinateur

Table des matières

I	Partie administrative	4
1	Informations professionnelles	5
2	Expériences professionnelles	6
3	Diplômes	7
4	Activités collectives et administratives	7
5	Distinctions	7
6	Projets en cours et achevés	8
7	Rayonnement, activités internationales	8
8	Activités d'enseignement (depuis 2018)	10
9	Domaines de recherche actuels	12
10	Liste des publications	12
11	Participation à des événements scientifiques	17
12	Encadrement de stages	19
13	Encadrement post-doctoral	19
14	Encadrement doctoral	19
II	Exposé des recherches réalisées au cours de la période postérieure à la thèse	21
1	Introduction	22
2	Optimisation topologique en mécanique des fluides	25
1	Optimisation de bathymétrie pour des ondes de petites amplitudes en eaux peu profondes	26
1.1	Existence d'une bathymétrie optimale	28
1.2	Discrétisation du problème d'optimisation	29

1.3	Résultats numériques	31
2	Optimisation topologique visant des solides optimisés avec des conductivités constantes par morceaux	33
2.1	Fonctions d'interpolation fluide-solide avec une conductivité constante par morceaux	37
2.2	Existence et unicité de solution faible et convergence d'une approximation éléments finis	38
2.3	Existence d'au moins une solution optimale et approximation éléments finis	40
2.4	Expériences numériques	42
3	Optimisation topologique avec conditions limites non-lisses en sortie . . .	45
3.1	Existence et unicité d'une solution faible	47
3.2	Convergence de la solution du problème régularisé	48
3.3	Conditions nécessaires d'optimalité du premier ordre	49
4	Pénalisation du modèle de Navier-Stokes-Darcy et application à des problèmes d'optimisation topologique visant des milieux poreux optimisés	53
4.1	Convergence du modèle de Navier-Stokes-Darcy pénalisé	54
4.2	Optimisation topologique visant des milieux poreux optimisés . .	57
5	Contributions à l'étude physique de problèmes d'optimisation topologique	59
5.1	Optimisation d'un <i>single pipe</i> et d'un <i>bend pipe</i>	61
5.2	Optimisation d'un canal vertical asymétriquement chauffé	63
3	Méthodes numériques pour des problèmes de propagation d'onde en régime harmonique	69
1	Présentation de l'équation de Helmholtz	71
2	Préconditionneur de type <i>Laplacien-shifté</i> pour l'équation de Helmholtz . .	73
2.1	Algorithme multigrille pour l'équation de Helmholtz avec shift complexe	74
2.2	Influence du shift sur la convergence d'un 2-grilles en 1d	76
2.3	Résultat de convergence pour un W-cycle	78
2.4	Extension du <i>Shifted Laplace Preconditioner : the expansion preconditioner</i>	80
3	Schémas différences finies avec correction de dispersion	85
3.1	Correction de dispersion en 1d	86
3.2	Correction de dispersion asymptotique pour un schéma à 9 points en 2d	88

3.3	Correction de dispersion asymptotique pour des schémas différences finies généraux	92
3.4	Résultats numériques : correction de dispersion et multigrille . . .	99

III Exposé sur les perspectives de recherche **101**

Première partie

Partie administrative

1 Informations professionnelles

Pierre-Henri Cocquet

Maître de Conférences

Université de Pau et des Pays de l'Adour



Né le 08 Décembre 1986 à Grande-Synthe (59)

Marié, 2 enfants

Nationalité Française

Portable : 06 52 20 54 31

E-mail : pierre-henri.cocquet@univ-pau.fr

Adresses professionnelles :

- IUT des Pays de l'Adour, Département MT2E, Avenue de l'Université, 64000 Pau.
 - Laboratoire des sciences pour l'ingénieur appliquées à la mécanique et au génie électrique (SIAME).
-

2 Expériences professionnelles

- 02/2024-07/2024 Délégation CNRS à l'Université Laval de Québec.
Accueilli au département de mathématiques et de statistique dans le Groupe Interdisciplinaire de Recherche en éléments Finis (GIREF).
- Depuis 09/2021 Maître de conférences en section 26 du CNU à l'Université de Pau et des Pays de l'Adour.
Membre du Laboratoire des sciences pour l'ingénieur appliquées à la mécanique et au génie électrique (SIAME).
Membre de IUT des Pays de l'Adour, Département MT2E.
- 09/2018-08/2021 Délégation croisée à l'Université de Pau et des Pays de l'adour.
Accueilli à l'IUT Génie Thermique et énergie (GTE).
Chercheur associé au Laboratoire des sciences pour l'ingénieur appliquées à la mécanique et au génie électrique (SIAME).
- 09/2014-08/2018 Maître de conférences en section 26 du CNU à l'Université de la Réunion.
Membre du Laboratoire de Physique et d'Ingénierie Mathématique pour l'énergie et l'Environnement (PIMENT).
Membre de l'école Supérieure d'Ingénieurs Réunion Océan Indien (ESIROI).
- 06/2013-08/2014 Post-Doctorat avec charges d'enseignements, Section de mathématique de l'Université de Genève.
- 10/2012-05/2013 Attaché temporaire d'enseignement et de recherche (ATER) à temps partiel, INSA de Toulouse.
- 09/2009-08/2012 Doctorant en mathématiques appliquées, Unité M2SN du département DTIM, ONERA, Toulouse.
- 09/2009-08/2012 Agent contractuel en charge d'enseignements, Université Toulouse III.

3 Diplômes

- 2012 **Qualifié** pour les fonctions de Maître de Conférences en 26ème section du CNU.
- 2012 **Doctorat** de Mathématiques Appliquées de l'Université Toulouse III, *étude mathématique et numérique de modèles homogénéisés de métamatériaux*. Soutenue le 7 décembre 2012. Mention : *Très Honorable*
Jury :
- A. Bendali, Professeur, Université Toulouse III, IMT-INSA. (**Directeur**)
 - V. Mouysset, Ingénieur de recherche ONERA, Toulouse. (**Directeur**)
 - A.S. Bonnet-Ben Dhia, Directrice de Recherche CNRS, UMA ENSTA. (**Rapporteur**)
 - E. Bonnetier, Professeur, Université Joseph Fourier, LJK. (**Rapporteur**)
 - B. Després Professeur, Université Paris 6, LJLL. (**Examineur**)
 - J.P. Raymond, Professeur, Université Toulouse III, IMT. (**Examineur**)
 - S. Tordeux, Maître de Conférences-HDR, Université de Pau et des Pays de l'Adour, Projet Magique 3D-LMA (**Examineur**).
- 2009 **Master 2** recherche en mathématiques appliquées, Université Toulouse III
Obtenu avec la mention *Très Bien*.
- 2008 **Master 1** de mathématiques fondamentales, Université Toulouse III
Obtenu avec la mention *Très Bien*.

4 Activités collectives et administratives

- Depuis 2022 **Membre élu** au conseil de l'IUT des Pays de l'Adour.
- Depuis 10/2021 **Membre élu** au conseil du Laboratoire SIAME.
- 09/2019-08/2022 **Directeur des études** 1ère année IUT GTE.
- 01/2019-08/2019 **Directeur des études** 2ème année IUT GTE.
- 09/2014-08/2018 Membre du **conseil du Laboratoire** PIMENT.
- 09/2014-09/2016 **Directeur des études** de l'ESIROI.
- 2011-2013 **Représentant des doctorants** de l'équipe Mathématiques pour l'Industrie et la Physique (MIP), Université Toulouse III.

5 Distinctions

- 09/2018-08/2022 Bénéficiaire de la Prime d'Encadrement Doctoral et de Recherche (**PEDR**).

6 Projets en cours et achevés

2019-2023	Coordinateur scientifique du projet ANR JCJC, ANR-19-CE40-0011, Optimisation topologique des transferts thermiques dans les fluides (Acronyme : O-TO-TT-FU). Collaborateurs : Alain Bastide et Delphine Ramalingom.
2022-2023	Projet de recherche contractuel avec EDF. Passive Ice Cooling (Acronyme : PIC). Collaborateurs : Yves Le Guer et Nicolas Chauchat.
2022-2025	Membre du projet ESR 2022 (projet région Nouvelle Aquitaine), Mixing Optimization for enhanced Microalgae Production in Open Ponds (Acronyme : MOMiPOP). Collaborateurs : Yves Le Guer et Nicolas Chauchat.
2024-2028	Partenaire du projet ANR PRC, Vers une nouvelle génération d'inversions conjointes InSAR - Gravité (Acronyme : DGMod). Collaborateurs : Valérie Cayol, Jean-Luc Froger, Lydie Gailler, Gilda Currenti, Olivier Bodart et Jonas Koko.

7 Rayonnement, activités internationales

Invitations dans des minisymposia :

- *Parallel solvers for Helmholtz problems* (DD XXVII, Prague, 2022).
- *On domain decomposition methods and preconditioners for PDE-constrained optimization problems* (DD XXVII, Prague, 2022).
- *Domain Decomposition and Multilevel Methods for Wave Propagation* (DD XXV, St. John's, Newfoundland, 2018).
- *Domain Decomposition for frequency domain wave problems* (DD XXIV, Svalbard, 2017).

Séjours en tant qu'invité :

- Organisation et participation au RESEARCH IN RESIDENCE intitulé "*Décomposition de domaine pour des problèmes de contrôle optimal*" au CIRM (Centre international de rencontres mathématiques) du 16/10/2023 au 20/10/2023. Participants : Alexandre Vieira et Martin J. Gander.
- Invitation au LAGA (Laboratoire Analyse, Géométrie et Applications) de l'Université Paris 13 du 27/03/2023 au 31/03/2023 par Martin J. Gander. Présentation au séminaire de modélisation et calcul scientifique.
- Invitation au CIRM (Centre international de rencontres mathématiques) du 03/10/2022 au 07/10/2022 par Martin J. Gander.
- Invitation à la section de Mathématiques de Genève du 30/05/2022 au 03/06/2022 par Martin J. Gander.
- Invitation à la section de Mathématiques de Genève du 12/12/2016 au 16/12/2016 par Martin J. Gander. Présentation au séminaire d'analyse numérique.

Rapporteur d'articles scientifiques : SIAM Journal on Scientific Computing, Journal of Scientific computing, Applied Mathematics Letter, Computers and Mathematics with Applications, Journal Of Computational Physics, Numerical Algorithms, International Journal of Computer Mathematics, Communications in Computational Physics.

8 Activités d'enseignement (depuis 2018)

2022-2023, BUT MTEE, Université de Pau et des Pays de l'Adour (216 h eq. TD)

Section :	1ère année BUT
Volume horaire :	104h TD.
Programme :	Mathématiques appliquées, Trigonométrie, Nombres complexes, fonctions.
Section :	2ème année BUT
Volume horaire :	70h TD.
Programme :	Suites, base de calcul financier, statistiques à une et deux variables, Intervalles de confiance.
Section :	2ème année BUT
Volume horaire :	42h TP.
Programme :	Transferts thermiques.

2021-2022, BUT MTEE, Université de Pau et des Pays de l'Adour (252 h eq. TD)

Section :	1ère année BUT
Volume horaire :	8h CM, 104h TD.
Programme :	Mathématiques appliquées, Trigonométrie, Nombres complexes, Equations différentielles, Intrégrales multiples, ...
Section :	2ème année DUT
Volume horaire :	80h TP.
Programme :	Transferts thermiques.
Section :	2ème année DUT
Volume horaire :	56h TP.
Programme :	Introduction à Scilab, Dichotomie, Méthode de Newton, Intégration approchée.

2019-2021, IUT GTE, Université de Pau et des Pays de l'Adour (236 h eq. TD)

Section :	1ère année IUT
Volume horaire :	8h CM, 96h TD.
Programme :	Mathématiques appliquées, Trigonométrie, Nombres complexes, Formules de Taylor, Equations différentielles, Intrégrales multiples, ...
Section :	2ème année IUT
Volume horaire :	80h TP.
Programme :	Transferts thermiques.
Section :	1ère année IUT
Volume horaire :	48h TP.
Programme :	Mécanique des fluides.

2018-2019, IUT GTE, Université de Pau et des Pays de l'Adour (225 h e.q. TD)

- Section : 1ère année IUT
Volume horaire : 96h TD.
Programme : Mathématiques appliquées, Trigonométrie, Nombres complexes, Formules de Taylor, Fonctions.
- Section : 2ème année IUT
Volume horaire : 2h CM et 12h TD.
Programme : Mathématiques appliquées, Calcul de longueur de courbes.
- Section : 2ème année IUT
Volume horaire : 6h CM, 12h TD et 64h TP.
Programme : Transferts thermiques, Conduction instationnaire, Rayonnement, Convection naturelle et forcée.
- Section : Licence professionnelle
Volume horaire : 10h CM et 14h TD.
Programme : Statique des fluides, Equation de Bernoulli, Échangeurs thermiques, Méthode DTLM, NUT.

9 Domaines de recherche actuels

- Méthodes numériques pour des problèmes de propagation d'onde en régime harmonique : *Equation de Helmholtz, Schémas différences finies avec correction de dispersion, méthode multigrille.*
- Optimisation topologique en mécanique des fluides : *Système de Navier-Stokes incompressible, Transferts de chaleur dans un fluide, Analyse mathématique (ex : existence, unicité, ...) et numérique (ex : convergence, ...).*

10 Liste des publications

Thèse

- [1] P.H. Cocquet, *étude mathématique et numérique de modèles homogénéisés de méta-matériaux*,
Thèse de l'université de Toulouse, 2012.

Articles parus dans des revues internationales à comité de lecture

- [18] A. Vieira, P.H. Cocquet, *The Boussinesq system with non-smooth boundary conditions : existence, relaxation and topology optimization*,
SIAM Journal on Control and Optimization, vol. 60, no 6, p. 3457-3484,
2022. (<https://doi.org/10.1137/21M146515>).
- [17] A. Vieira, A. Bastide, P.H. Cocquet, *Topology Optimization for Steady-State Anisothermal Flow Targeting Solids with Piecewise Constant Thermal Diffusivity*,
Applied Mathematics & Optimization, vol. 85, no 3, p. 1-32, 2022.
(<https://doi.org/10.1007/s00245-022-09828-5>).
- [16] A. Aldor, Y. Moguen, K. El Omari, C. Habchi, P.H. Cocquet, Y. Le Guer, *Heat transfer enhancement by chaotic advection in a novel sine-helical channel geometry*,
International Journal of Heat and Mass Transfer, vol. 193, p. 122870,
2022. (<https://doi.org/10.1016/j.ijheatmasstransfer.2022.122870>).
- [15] F.Z. Boudara, H. Bouzekri, Y. Benhammadi, P.H. Cocquet, M. Rivaletto, L. Pécastaing, A. Silvestre de Ferron, Y. Le Guer, *Accurate numerical modeling of convective heat transfer coefficient for a high power PLA-core toroidal electromagnetic coupler subject to cooling*,
International Journal of Thermal Sciences, vol. 170, p. 107125, 2021.
(<https://doi.org/10.1016/j.ijthermalsci.2021.107125>).
- [14] P.H. Cocquet, S. Rizzo, J. Salomon, *Optimization of bathymetry for long waves with small amplitude*,
SIAM Journal on Control and Optimization, vol. 59, no. 6, p. 4429-4456,
2021. (<https://doi.org/10.1137/20M1326337>).

- [13] P.H. Cocquet, M. J Gander, X. Xiang, *Closed Form Dispersion Corrections Including a Real Shifted WaveNumber for Finite Difference Discretizations of 2D Constant Coefficient Helmholtz Problems*, SIAM Journal on Scientific Computing, vol. 43, no. 1, p. A278-A308, 2021. (<https://doi.org/10.1137/19M1304350>).
- [12] P.H. Cocquet, M. Rakotobe, D. Ramalingom, A. Bastide, *Error analysis for the finite element approximation of the Darcy-Brinkman-Forchheimer model for porous media with mixed boundary conditions*, Journal of Computational and Applied Mathematics, vol. 381, p. 113008, 2021. (<https://doi.org/10.1016/j.cam.2020.113008>).
- [11] A. Bastide, P.H. Cocquet, D. Ramalingom and R. Maleck, *A multi-objective optimization problem in mixed and natural convection for a vertical channel asymmetrically heated*, Structural and Multidisciplinary Optimization, 60(5), 2001-2020, 2019 (<https://doi.org/10.1007/s00158-019-02306-7>).
- [10] A. Bastide, P.H. Cocquet et D. Ramalingom, *Penalization model for Navier-Stokes-Darcy equation with application to porosity-oriented topology optimization*, Mathematical Models and Methods in Applied Sciences (M3AS), 2018 (<https://doi.org/10.1142/S0218202518500409>).
- [9] A. Bastide, P.H. Cocquet et D. Ramalingom, *A new interpolation technique to deal with fluid-porous media interfaces for topology optimization of heat transfer*, Computers & Fluids, Volume 168, 144-158, 2018 (<https://doi.org/10.1016/j.compfluid.2018.04.005>).
- [8] D. Ramalingom, P.H. Cocquet et A. Bastide, *Numerical study of natural convection in asymmetrically heated channel considering thermal stratification and surface radiation*, Numerical Heat Transfer, Part A : Applications, 72(9), 681-696, 2017 (doi.org/10.1080/10407782.2017.1400337).
- [7] P.H. Cocquet et M.J. Gander, *How Large a Shift is Needed in the Shifted Helmholtz Preconditioner for its Effective Inversion by Multigrid?*, SIAM Journal on Scientific Computing, vol. 39-2, p. A385–A414, 2017 ([doi:10.1137/15M102085X](https://doi.org/10.1137/15M102085X)).
- [6] A. Bendali, P.H. Cocquet et S. Tordeux, *Approximation by Multipoles of the Multiple Acoustic Scattering by Small Obstacles in Three Dimensions and Application to the Foldy Theory of Isotropic Scattering*, Archive for Rational Mechanics and Analysis, vol. 219, no. 3, p. 1017–1059, 2016 ([doi 10.1007/s00205-015-0915-5](https://doi.org/10.1007/s00205-015-0915-5)).

- [5] K. Boutarene, P.H. Cocquet, *Scattering of a scalar time-harmonic wave by a penetrable obstacle with a thin layer*, European Journal of Applied Mathematics, vol. 27, no. 2, p. 264–310, 2016 (doi :10.1017/S095679251500056X).
- [4] P.H. Cocquet, P.A. Mazet et V. Mouysset, *On the existence and uniqueness of a solution for some frequency-dependent partial differential equations coming from the modelling of metamaterials*, SIAM Journal on Mathematical Analysis, vol. 44, no. 6, p. 3806–3833, 2012 (doi :10.1137/100810071).
- [3] P.H. Cocquet, A. Bendali et S. Tordeux, *Scattering of a scalar time-harmonic wave by N small spheres by the method of matched asymptotic expansions*, Numerical Analysis and Applications, vol. 5, no. 2, p. 116–123, 2012 (doi : 10.1134/S1995423912020036).
- [2] P.H. Cocquet, P.A. Mazet et V. Mouysset, *Sur l'existence et l'unicité de solutions pour des équations de Drude-Born-Fedorov homogénéisées en domaine borné et applications aux méta-matériaux*, Comptes-Rendus à l'académie des Sciences, Série I, Mathématiques, vol. 349, no. 1, p. 99–103, 2011.
- [1] P.H. Cocquet, P.A. Mazet et V. Mouysset, *Theoretical analysis of some homogenized metamaterials and application of PML to perform cloaking and back-scattering invisibility*, PIERS Online, vol. 6, no. 2, p. 189–194, 2010.

Publications dans des actes de congrès avec comité de lecture

- [11] P.H. Cocquet, A. Vieira, *Optimized Schwarz method for coupled direct-adjoint problems applied to parameter identification in advection-diffusion equation*,
Accepté pour publication dans les proceedings de DD 27.
- [10] D Marti, D Ramalingom, P.H. Cocquet, A Bastide, *Utilisation de fonctions objectifs locales pour l'optimisation topologique des échanges thermiques d'un canal vertical asymétriquement chauffé*,
Congrès annuel de la société française de thermique, 2021 (<https://doi.org/10.25855/SFT2021-043>).
- [9] P.H. Cocquet, M. J. Gander et X. Xiang, *Dispersion Correction for Helmholtz in 1D with piecewise constant wavenumber*,
Domain Decomposition Methods in Science and Engineering XXV, LNCSE, Springer-Verlag, 2019.
- [8] P.H. Cocquet et M.J. Gander, *Analysis of the shifted Helmholtz expansion preconditioner for the Helmholtz equation*,
Domain Decomposition Methods in Science and Engineering XXIV, LNCSE, Springer-Verlag, 2018 (doi : 10.1007/978-3-319-93873-8.17).
- [7] P.H. Cocquet, M. J. Gander et X. Xiang, *A finite difference method with optimized dispersion correction for the Helmholtz equation*,
Domain Decomposition Methods in Science and Engineering XXIV, LNCSE, Springer-Verlag, 2018 (doi : 10.1007/978-3-319-93873-8.18).
- [6] D. Ramalingom, P.H. Cocquet et A. Bastide, *Optimisation topologique des échanges thermiques dans un canal vertical asymétriquement chauffé*,
Publié dans les actes du 25^{ème} Congrès Français de Thermique, 2017.
- [5] G. Riviere, P.H. Cocquet, E. Fock, et A. Bastide, *Optimisation topologique d'une bouche de ventilation en régime turbulent*,
Conférence Francophone de l'International Building Performance Simulation Association (IBPSA), 2017.
- [4] A. Bendali, P.H. Cocquet et S. Tordeux, *Improvement of the Foldy-Lax model for three dimensional multiple scattering by small obstacles*,
The 12th International Conference on Mathematical and Numerical Aspects of Waves (WAVES 2015), Karlsruhe, Juillet 2015.
- [3] P.H. Cocquet et J. Michaud, *Error of an explicit Fuzzy Domain Decomposition Method (eFDDM) : What do matched asymptotic expansions teach us ?*,
Domain Decomposition Methods in Science and Engineering XXII, LNCSE, Springer-Verlag, 2016 (doi : 10.1007/978-3-319-18827-0.33).

- [2] P.H. Cocquet et M.J. Gander, *On the minimal shift in the shifted Laplacian preconditioner for multigrid to work*, Domain Decomposition Methods in Science and Engineering XXII, LNCSE, Springer-Verlag, 2016 (doi : 10.1007/978-3-319-18827-0_12).
- [1] P.H. Cocquet, P.A. Mazet et V. Mouysset, *Mathematical study of many example of homogenized metamaterials*, The 10th International Conference on Mathematical and Numerical Aspects of Waves (WAVES 2011), Vancouver, Juillet 2011.

Articles soumis

- [4] P.H. Cocquet, F.Z. Mihami, Y. Moguen, V. Roeber, *Serre-Green-Naghdi equations with optimized dispersion properties through a modified auxiliary elliptic equation*.
- [3] P.H. Cocquet, M.J. Gander, *Asymptotic dispersion correction in general finite difference schemes for Helmholtz problems*.
- [2] P.H. Cocquet, M.J. Gander, A. Vieira *Decompose-then-optimize or optimize-then-decompose : domain decomposition for optimal control problem*. Article en cours de re-soumission.
- [1] D. Marti, D. Ramalingom, P.H. Cocquet and A. Bastide, *Topology Optimization of Fluid Flow Using Local Thermal Objective Function Based on Field Synergy Principle*.

Rapports de recherche

- [3] P.H. Cocquet, P.A. Mazet et V. Mouysset, *PRF METAMAT : Rapport 1ère année*, RT 1/14549 DEMR, Février 2010.
- [2] P.H. Cocquet, P.A. Mazet et V. Mouysset, *PRF METAMAT : Rapport 2ème année*, RT 1/16746 DEMR, Mars 2011.
- [1] P.H. Cocquet, P.A. Mazet et V. Mouysset, *PRF METAMAT : Rapport 3ème année*, RT 1/18280 DEMR, Avril 2012.

11 Participation à des événements scientifiques

Communications dans des conférences internationales

- [11] *LiQuOFETI : a FETI-inspired method for elliptic quadratic optimal control problems* (Poster),
HYBRID RESEARCH SCHOOL Domain Decomposition for Optimal Control Problems (2644), CIRM, Septembre 2022, Marseille.
- [10] *Dispersion minimizing finite difference method for the two-dimensional Helmholtz equation*,
Participation au minisymposium *Parallel solvers for Helmholtz problems*.
27th International Domain Decomposition Conference, Juillet 2022, Prague, République Tchèque.
- [9] *Dispersion correction for finite difference scheme for the Helmholtz equation*,
Participation au minisymposium *Domain Decomposition and Multilevel Methods for Wave Propagation*.
International conference on Domain Decomposition Method (DD XXV),
Juillet 2018, St. John's, Newfoundland, Canada.
- [8] *What is the minimal shift to use in the Shifted Helmholtz Preconditioner for its effective inversion by multigrid?*,
Participation au minisymposium *Domain Decomposition for frequency domain wave problems*.
International conference on Domain Decomposition Method (DD XXIV),
Février 2017, Svalbard, Norvège.
- [7] *Improvement of the Foldy-Lax model for three dimensional multiple scattering by small obstacles*,
The 12th International Conference on Mathematical and Numerical Aspects of Waves (WAVES 2015),
Juillet 2015, Karlsruhe, Allemagne.
- [6] *Shifted Laplacian multigrid preconditionner for indefinite Helmholtz problems*,
Swiss Numerics Days, Avril 2014, Zürich, Suisse.
- [5] *Approximation numérique de modèles homogénéisés de méta-matériaux*,
Congrès SMAI 2013, Mai 2013, Seignosse Le Penon, France.
- [4] *Mathematical study of many example of homogenized metamaterials*,
The 10th International Conference on Mathematical and Numerical Aspects of Waves (WAVES 2011),
Juillet 2011, Vancouver, Canada.
- [3] *On the well-posedness and the numerical approximation of some problems coming from the modelling of metamaterials*,
The International Congress on Industrial and Applied Mathematics (ICIAM 2011),
Juillet 2011, Vancouver, Canada.

- [2] *Sur l'existence et de l'unicité de solutions pour les équations de Maxwell-Drude-Born-Fedorov en présence de méta-matériaux homogénéisés,*
Congrès SMAI 2011, Mai 2011, Guidel, France.
- [1] *Theoretical analysis of some homogenized metamaterials and application of PML to perform cloaking and back-scattering invisibility,*
Progress In Electromagnetics Research Symposium (PIERS 2010), Mars 2010, Xi'an, Chine.

Communications dans des Conférences nationales

- [6] *Préconditionneur multigrille pour l'équation de Helmholtz,*
Journées ONDES du Sud-Ouest (JO-SO, Poster), Février 2014, Toulouse, France.
- [5] *éléments pour l'étude et la simulation de milieux en présence de méta-matériaux,*
Colloque National Métamatériaux (Invitation), Septembre 2013, Paris, France.
- [4] *Un schéma numérique pour l'approximation de méta-matériaux acoustiques,*
Colloque EDP-Normandie (Invitation, Poster), Octobre 2013, Caen, France.
- [3] *Analyse asymptotique d'un problème de diffraction multiple et justification d'un modèle de Foldy-Lax,*
Journées ARAMIS (Invitation), Octobre 2013, Pau, France.
- [2] *Modélisation mathématique et homogénéisation non-périodique formelle de métamatériaux,*
Journées DYNAMO 2013 (Invitation), Mars 2013, Orléans, France.
- [1] *Réalisation de métamatériaux acoustiques par homogénéisation non périodique et développements asymptotiques raccordés,*
Colloque National Métamatériaux (Poster), Mars 2011, Paris, France.

12 Encadrement de stages

- [3] Léa Gosnik, *Etude de l'ébullition générée par la chute d'un cylindre chaud dans de l'eau au repos : modèle simplifié d'un écoulement pyroclastique*,
Stage de 2^{ème} année ENSGTI.
Co-encadrement à 30% avec Yves Le Guer (35%) et Jean-Pierre Bedecarrats (35%).
Durée : du 30/05/2022 au 02/09/2022.
- [2] Hajar Bouzekri, *Modélisation thermique d'un tore creux soumis à un flux de chaleur pariétal*,
Stage de 2^{ème} année ENSGTI
Co-encadrement à 50% avec Yves Le Guer (50%).
Durée : 2 mois (avril à juin 2020).
- [1] Aubert Lerouge, *Unsteady thermal transfer modeling during the quenching of a hot horizontal cylinder in water*,
Stage de 2^{ème} année ENSGTI.
Co-encadrement à 50% avec Yves Le Guer (50%).
Durée : 2 mois (avril à juin 2020).

13 Encadrement post-doctoral

- [1] Alexandre Vieira, *Optimisation topologique pour les transferts de chaleur dans les fluides*.
Contrat financé par le projet ANR-JCJC-O-TO-TT-FU (ANR-19-CE40-0011).
Durée : 2 ans et 4 mois (de janvier 2020 à avril 2022).
Encadrement à 100%
Actuellement Post-doctorant dans l'équipe INRIA CASTOR.

14 Encadrement doctoral

- [6] Imane Skifa, *Optimisation du mélange pour la culture intensifiée de microalgues en bassins ouverts*,
Co-direction à 50% avec Yves Le Guer (50%).
Thèse débutée à l'université de Pau et des Pays de l'Adour en juin 2023.
- [5] David Marti, *Caractérisation et optimisation d'écoulements de fluides confinés par actionneur plasma en convection naturelle*,
Co-encadrement à 35% avec Delphine Ramalingom (30%) et Alain Bastide (35%).
Thèse débutée à l'université de la Réunion en 2020.

- [4] Abbass Eldor, *Optimisation des performances d'un nouvel échangeur multifonctionnel à canaux sinus-hélicoïdaux*,
Co-encadrement à 30% avec Yves Le Guer (35%) et Charbel Habchi (35%).
Thèse débutée à l'université de Pau et des Pays de l'Adour en octobre 2019, Soutenance effectuée le 08 décembre 2022.
Actuellement en création d'entreprise au Liban.
- [3] Michaël Rakotobe, *Modélisation d'écoulement de fluide en milieu poreux hétérogène et application en optimisation topologique*,
Co-encadrement à 40% avec Laeticia Adelard (20%) et Alain Bastide (40%).
Thèse débutée à l'université de la Réunion en 2017, Soutenance effectuée le 13 octobre 2022.
Actuellement enseignant vacataire en mathématiques dans l'enseignement secondaire.
- [2] Delphine Ramalingom, *Optimisation topologique des transferts thermiques et massiques dans un canal asymétriquement chauffé*,
Co-encadrement à 50% avec Alain Bastide (50%).
Thèse débutée à l'université de la Réunion en 2013. Soutenance effectuée le 13/12/2017.
Actuellement Ingénieure en calcul scientifique à l'université de la Réunion (laboratoire PIMENT).
- [1] Garry Rivière, *Optimisation topologique d'écoulements turbulents et application à la ventilation des bâtiments*,
Co-encadrement à 33% avec Alain Bastide (33%) et Eric Fock (33%).
Thèse débutée à l'université de la Réunion en 2013. Soutenance effectuée le 01/03/2017.
Actuellement PRCE à l'université de la Réunion (département SBE).

Deuxième partie

Exposé des recherches réalisées au cours de la période postérieure à la thèse

Chapitre 1

Introduction

De 2009 à 2012, j'ai effectué ma thèse de doctorat à l'ONERA (centre de Toulouse) sous la direction de Vincent Mouysset, Pierre-Alain Mazet et Abderrahmane Bendali. Pendant ma thèse, j'ai travaillé sur l'étude mathématique et l'approximation numérique de modèles homogénéisés de métamatériaux qui sont des milieux dont certains paramètres physiques (ex : permittivité, perméabilité) dépendent de la fréquence et peuvent avoir des parties réelles négatives pour certaines fréquences.

Dans ma thèse, j'ai notamment obtenu des résultats d'existence et d'unicité pour des systèmes de propagation d'ondes en régime harmonique en présence de métamatériaux. Les modèles étudiés ont été : les équations de Maxwell en présence de milieux diélectriques, bi-anisotropiques ou chiraux, le système de l'acoustique et celui de l'élasticité linéaire. J'ai également adapté puis analysé la convergence d'un schéma éléments finis de type moindre carré pour ces systèmes.

Une seconde partie de ma thèse portait sur l'analyse asymptotique d'un problème de diffraction multiple, pour l'équation de Helmholtz, par des petits obstacles. J'ai utilisé pour cela la méthode des développements asymptotiques raccordés pour obtenir un développement asymptotique, à tout ordre, du champ diffracté par les petits objets. Ces résultats ont permis d'obtenir formellement un indice de réfraction équivalent pour le milieu contenant les objets proposant ainsi une méthode d'homogénéisation pour créer des métamatériaux. Ils ont également été utilisés pour justifier et améliorer le modèle de Foldy-Lax qui donne une approximation du champ diffracté par plusieurs petits obstacles. Les résultats mentionnés précédemment sont regroupés en partie dans mon manuscrit de thèse et dans les publications suivantes :

- Cocquet, P. H., Mazet, P. A., & Mouysset, V. (2012). *On the existence and uniqueness of a solution for some frequency-dependent partial differential equations coming from the modeling of metamaterials*. SIAM Journal on Mathematical Analysis, 44(6), 3806-3833.
- Bendali, A., Cocquet, P. H., & Tordeux, S. (2016). *Approximation by multipoles of the multiple acoustic scattering by small obstacles in three dimensions and application to the*

foldy theory of isotropic scattering. Archive for Rational Mechanics and Analysis, 219, 1017-1059.

Ce document présente une synthèse des travaux de recherche effectués après ma thèse de doctorat soutenue en décembre 2012. Il regroupe ainsi les apports obtenus pendant mon post-doctorat effectué à Genève de juin 2013 à juillet 2014 puis ceux réalisés depuis mon recrutement en tant que Maître de conférences à l'université de la Réunion de 2014 à 2018 puis à l'université de Pau et des Pays de l'Adour depuis 2018. Depuis mon post-doctorat, je travaille notamment avec Martin J. Gander sur l'analyse numérique de problème de propagation d'onde en régime harmonique. Après mon recrutement à l'université de la Réunion, j'ai également initié une collaboration avec un collègue physicien, Alain Bastide, sur l'optimisation topologique en mécanique des fluides. Ce mémoire sera ainsi composé de deux parties qui correspondent à des travaux dans ces deux domaines.

Le second chapitre concerne l'étude de problèmes d'optimisation topologique qui visent à déterminer la localisation d'un solide au sein d'un fluide permettant de minimiser une fonction coût donnée. Dans ce type de problème d'optimisation avec contraintes données par des équations aux dérivées partielles (EDP), la localisation du fluide est effectuée avec un terme de pénalisation qui s'annule dans les zones fluides du domaine de calcul et devient très grand (idéalement infini) dans les zones solides. De tels problèmes reviennent ainsi à minimiser une fonction coût en déterminant des paramètres dans une EDP.

On présentera tout d'abord des résultats portant sur un problème d'optimisation de bathymétrie pour des vagues de petites amplitudes en eaux peu profondes modélisées par une équation d'onde en régime harmonique. Bien que ce ne soit pas un problème d'optimisation topologique, cela rentre parfaitement dans la problématique d'identification de paramètres dans une EDP. On présentera ensuite des résultats en lien avec l'optimisation topologique des transferts thermiques dans les fluides. Dans ce cadre, les contraintes sont données par le système de Navier-Stokes sous l'approximation de Boussinesq couplé à une équation d'énergie. On donnera d'abord des résultats pour des écoulements stationnaires puis instationnaires avec prise en compte d'un écoulement retour en sortie via des conditions limites non-lisses. L'étude d'un problème d'optimisation topologique visant à localiser des milieux poreux modélisés par une loi de Darcy sera ensuite présenté. Pour chacun des cas précédemment mentionnés, on donnera essentiellement des théorèmes d'existence et d'unicité d'une solution faible à l'EDP considérée, des résultats d'existence d'au moins une solution optimale et des conditions nécessaires d'optimalité du premier ordre pour les problèmes d'optimisation ainsi que des résultats de convergence de la solution du problème d'optimisation discrétisé par éléments finis vers la solution du problème d'optimisation continu. Des simulations numériques permettant d'illustrer certains résultats théoriques seront également présentés.

Une seconde partie plus succincte donnera quelques résultats numériques obtenus en résolvant des problèmes d'optimisation topologique dans un cadre plus physique et applicatif. Pour les cas présentés, l'objectif est de minimiser les pertes de charge (l'énergie nécessaire au fluide pour se mouvoir) tout en maximisant les transferts thermiques.

Le troisième chapitre de ce mémoire concernera l'analyse numérique de problèmes de propagation d'ondes en régime harmonique. On présentera tout d'abord l'équation de Helmholtz et les principales difficultés associées à sa résolution numérique.

La seconde section présentera le travail réalisé sur un préconditionneur de type *Laplacien shifté* (*Shifted Laplace Preconditioner* ou encore SLP en anglais) utilisé pour accélérer la résolution de systèmes linéaires issus d'une discrétisation de l'équation de Helmholtz. On donnera des conditions permettant de choisir le *shift* pour lequel l'inversion du préconditionneur peut être réalisée avec une méthode multigrille. Ce chapitre se terminera avec la présentation d'une extension du SLP appelée *expansion preconditioner* ainsi qu'une étude, théorique et numérique, sur le choix du *shift* pour lequel ce dernier est un " bon " préconditionneur.

La troisième section présentera des schémas différences finies avec correction de dispersion. Ces derniers sont basés sur l'introduction, dans le schéma, d'un nombre d'onde dépendant du pas de maillage et de paramètres libres qui sont ensuite optimisés afin de réduire l'erreur de dispersion. L'avantage principal de l'approche proposée repose sur une expression explicite, pour des pas de maillage suffisamment petits, des paramètres minimisant l'erreur de dispersion. On montrera également l'effet de la réduction de la dispersion numérique sur l'erreur relative du schéma à travers plusieurs expérimentations numériques.

Pour terminer cette introduction, j'aimerais préciser que les deux parties de ce mémoire représentent les deux principaux axes de recherche sur lesquels je travaille depuis ma soutenance de thèse. Ainsi, bien que depuis mon arrivée à l'université de Pau et des Pays de l'Adour j'ai initié plusieurs collaborations au sein du laboratoire SIAME, les autres travaux ne relevant pas directement de ces deux thématiques ne seront pas détaillés dans ce document.

Chapitre 2

Optimisation topologique en mécanique des fluides

Cette partie présente des travaux effectués en lien avec l'optimisation topologique appliquée à la mécanique des fluides. Son contenu est basé sur les publications suivantes :

- [Chap2-A1] Ramalingom, D., Cocquet, P. H., & Bastide, A. (2018). *A new interpolation technique to deal with fluid-porous media interfaces for topology optimization of heat transfer*. Computers & Fluids, 168, 144-158.
- [Chap2-A2] Bastide, A., Cocquet, P. H., & Ramalingom, D. (2018). *Penalization model for Navier-Stokes-Darcy equations with application to porosity-oriented topology optimization*. Mathematical Models and Methods in Applied Sciences, 28(08), 1481-1512.
- [Chap2-A3] Ramalingom, D., Cocquet, P. H., Maleck, R., & Bastide, A. (2019). *A multi-objective optimization problem in mixed and natural convection for a vertical channel asymmetrically heated*. Structural and Multidisciplinary Optimization, 60(5), 2001-2020.
- [Chap2-A4] Cocquet, P. H., Riffo, S., & Salomon, J. (2021). *Optimization of bathymetry for long waves with small amplitude*. SIAM Journal on Control and Optimization, 59(6), 4429-4456.
- [Chap2-A5] Vieira, A., Bastide, A., & Cocquet, P. H. (2022). *Topology Optimization for Steady-State Anisothermal Flow Targeting Solids with Piecewise Constant Thermal Diffusivity*. Applied Mathematics & Optimization, 85(3), 1-32.
- [Chap2-A6] Vieira, A., & Cocquet, P. H. (2022). *The Boussinesq System with Nonsmooth Boundary Conditions : Existence, Relaxation, and Topology Optimization*. SIAM Journal on Control and Optimization, 60(6), 3457-3484.

1 Optimisation de bathymétrie pour des ondes de petites amplitudes en eaux peu profondes

Cette section présente des résultats issus de [Chap2-A4].

Les objectifs de [Chap2-A4] étaient de traiter un problème d'identification de bathymétrie indépendante du temps en résolvant un problème d'optimisation où les contraintes sont données par une équation de Helmholtz avec des conditions limites de Robin. Le modèle a été obtenu en simplifiant tout d'abord le système de Navier-Stokes incompressible à surface libre en supposant que H/L était assez petit (H est une hauteur relative et L une longueur caractéristique horizontale). Cette analyse asymptotique permet d'obtenir le modèle de Saint-Venant. En supposant ensuite que les vagues ont des petites amplitudes par rapport à H , l'équation de Saint-Venant devient une équation d'ondes dans laquelle la bathymétrie intervient dans la partie spatiale de l'opérateur. En supposant enfin que la bathymétrie est une perturbation à support compacte d'un fond plat et en considérant des solutions harmoniques en temps, on peut réécrire l'équation d'ondes comme une équation de Helmholtz où la bathymétrie intervient dans la partie principale de l'opérateur. Concernant la bathymétrie à optimiser, nous avons suivi [37] et supposé qu'elle est peu régulière voire même discontinue.

Notre problème d'identification de bathymétrie revient à déterminer des coefficients dans une EDP afin de minimiser une fonction coût. Ce type de problème a déjà été étudié par exemple dans [6, 14] mais ces travaux traitent de problème d'optimisation où les contraintes sont données par des EDP elliptiques réelles et coercives. Ainsi ces résultats ne s'appliquent pas à l'équation de Helmholtz qui est non-coercive et complexe. Un travail connexe peut être trouvé dans [23] dans lequel le paramètre d'optimisation est une permittivité et apparaît donc comme un opérateur de multiplication dans l'équation de Helmholtz. Toutefois, dans [Chap2-A4], la bathymétrie étant supposée non-régulière et intervenant dans la partie principale de l'équation de Helmholtz, les résultats précédemment mentionnés ne s'appliquent pas.

Dans le cas d'eaux peu profondes et des vagues de petites amplitudes, on peut montrer que l'élévation du niveau d'eau η satisfait l'équation d'onde ci-dessous :

$$\frac{\partial^2 \eta}{\partial t^2} - \operatorname{div}(gz_b \nabla \eta) = 0, \quad (2.1)$$

où $g = 9.81 \text{ m/s}^2$ et $-z_b(x)$ est la bathymétrie avec $z_b > 0$. Dans notre étude, nous avons simplifié d'avantage (2.1) en considérant des solutions harmoniques en temps, soit

$$\eta(x, t) = \operatorname{Re}\{\psi_{\text{tot}}(x)e^{-i\omega t}\},$$

et une bathymétrie de la forme

$$z_b(x) := z_0 + \delta z_b(x),$$

où z_0 est une constante représentant un fond plat et δz_b est une perturbation à support compact dans Ω . Le domaine $\Omega \subset \mathbb{R}^2$ est un ouvert borné de frontière Lipchitzienne. En considérant alors δz_b comme un obstacle, nous avons réécrit (2.1) comme un problème de diffraction d'une onde plane $\psi_0(x) = e^{ik_0 x \cdot d}$ sur l'obstacle caractérisé par la perturbation de bathymétrie δz_b . L'onde totale ψ_{tot} vérifie alors

$$\begin{cases} \operatorname{div}((1+q)\nabla\psi_{\text{tot}}) + k_0^2\psi_{\text{tot}} = 0 & \text{dans } \Omega, \\ \nabla(\psi_{\text{tot}} - \psi_0) \cdot \mathbf{n} - ik_0(\psi_{\text{tot}} - \psi_0) = 0 & \text{sur } \partial\Omega, \end{cases} \quad (2.2)$$

où $q(x) = \frac{\delta z_b(x)}{z_0}$, $k_0 = \frac{\omega}{\sqrt{gz_0}}$ et \mathbf{n} est la normale unitaire sortante à $\partial\Omega$. La condition limite intervenant dans (2.2) représente une approximation de la condition de radiation de Sommerfield et assure que les ondes diffractées sortent du domaine.

En décomposant l'onde totale suivant $\psi_{\text{tot}} = \psi_0 + \psi_{\text{sc}}$ où ψ_{sc} est l'onde diffractée, l'équation (2.2) devient

$$\begin{cases} \operatorname{div}((1+q)\nabla\psi_{\text{sc}}) + k_0^2\psi_{\text{sc}} = -\operatorname{div}(q\nabla\psi_0) & \text{dans } \Omega, \\ \nabla\psi_{\text{sc}} \cdot \mathbf{n} - ik_0\psi_{\text{sc}} = 0 & \text{sur } \partial\Omega. \end{cases} \quad (2.3)$$

Pour certains choix de ψ_0 et de g , (2.2) et (2.3) sont équivalents à

$$\begin{cases} -\operatorname{div}((1+q)\nabla\psi) - k_0^2\psi = \operatorname{div}(q\nabla\psi_0) & \text{in } \Omega, \\ (1+q)\nabla\psi \cdot \mathbf{n} - ik_0\psi = g - q\nabla\psi_0 \cdot \mathbf{n} & \text{on } \partial\Omega, \end{cases} \quad (2.4)$$

dont la formulation variationnelle est

$$a(q; \psi, \phi) = b(q; \phi), \quad \forall \phi \in H^1(\Omega), \quad (2.5)$$

avec

$$\begin{aligned} a(q; \psi, \phi) &= \int_{\Omega} ((1+q)\nabla\psi \cdot \nabla\bar{\phi} - k_0^2\psi\bar{\phi}) \, dx - ik_0 \int_{\partial\Omega} \psi\bar{\phi} \, d\sigma, \\ b(q; \phi) &= - \int_{\Omega} q\nabla\psi_0 \cdot \nabla\bar{\phi} \, dx + \langle g, \bar{\phi} \rangle_{H^{-1/2}, H^{1/2}}. \end{aligned} \quad (2.6)$$

En supposant que $q \in L^\infty(\Omega)$ et $1 + q(x) \geq \alpha > 0$, on peut démontrer que le problème variationnel (2.5) admet une unique solution (voir [22, Theorems 2.1, 2.4]).

L'objectif principal de [Chap2-A3] est d'étudier, pour une fonction coût \mathcal{J} , le problème d'optimisation de bathymétrie ci-dessous

$$\begin{aligned} &\min \mathcal{J}(q, \psi), \\ &\text{avec } (q, \psi) \in U_{\text{ad}} \times H^1(\Omega) \text{ et } \psi \text{ vérifie (2.5)}. \end{aligned} \quad (2.7)$$

L'ensemble U_{ad} des bathymétries admissibles est un sous-ensemble de $BV(\Omega)$ correspondant aux fonctions à variations bornées sur Ω . La motivation principale pour utiliser des telles fonctions comme paramètres d'optimisation est que cet ensemble contient, en particulier, des fonctions discontinues.

On donne maintenant quelques propriétés de $BV(\Omega)$ et on réfère à [3] pour plus d'informations. L'espace $BV(\Omega)$ est l'ensemble des fonctions $q \in L^1(\Omega)$ dont le gradient au sens des distributions Dq est une mesure de Radon bornée à valeur dans \mathbb{R}^2 . C'est un espace de Banach pour la norme

$$\|q\|_{BV(\Omega)} = \|q\|_{L^1(\Omega)} + |Dq|(\Omega),$$

où $|Dq|(\Omega)$ est la variation de q définie par [3, p. 119, Definition 3.4]

$$|Dq|(\Omega) = \sup \left\{ \int_{\Omega} q \operatorname{div}(\varphi) \, dx \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2) \text{ et } \|\varphi\|_{L^\infty(\Omega)} \leq 1 \right\}. \quad (2.8)$$

Dans la suite, la convergence faible* dans $BV(\Omega)$ sera utilisée dans certains résultats. Cette dernière est indiquée par

$$q_n \rightharpoonup q \text{ faible* dans } BV(\Omega),$$

et signifie que

$$q_n \rightarrow q \text{ dans } L^1(\Omega) \text{ et } \lim_{n \rightarrow +\infty} \int_{\Omega} \psi \cdot dDq_n = \int_{\Omega} \psi \cdot dDq, \forall \psi \in \mathcal{C}^0(\Omega, \mathbb{R}^2).$$

L'ensemble des bathymétries admissibles $U_{\alpha, \Lambda}$ est alors donné par

$$U_{\Lambda} = \{q \in BV(\Omega) \mid \alpha - 1 \leq q(x) \leq \Lambda \text{ p.p. } x \in \Omega\},$$

et certains résultats présentés ci-dessous font également intervenir l'ensemble admissible suivant

$$U_{\Lambda, \kappa} = \{q \in U_{\Lambda} \text{ tel que } |Dq|(\Omega) \leq \kappa\},$$

où $\Lambda \geq \max\{\alpha - 1, 0\}$.

1.1 Existence d'une bathymétrie optimale

Tout d'abord, on présente un résultat de régularité sur l'application qui, à tout $q \in U_{\alpha, \Lambda}$, associe la solution du problème variationnel (2.5).

Théorème 2.1 *Pour tout $q \in U_{\alpha, \Lambda}$, on note $\psi(q)$ la solution de (2.5). Alors l'application*

$$q \in (U_{\alpha, \Lambda}, \text{faible*}) \mapsto \psi(q) \in (H^1(\Omega), \text{fort}),$$

est continue.

Le Théorème 2.1 permet ensuite de montrer l'existence d'au moins une solution au problème d'optimisation (2.52).

Théorème 2.2 *On suppose que $U_{\alpha, \Lambda} = U_{\Lambda}$ et que la fonction coût $(q, \psi) \in U_{\Lambda} \mapsto \mathcal{J}(q, \psi) \in \mathbb{R}$ vérifie :*

(A1) Il existe $\beta > 0$ et \mathcal{J}_0 tels que

$$\mathcal{J}(q, \psi) = \mathcal{J}_0(q, \psi) + \beta |Dq|(\Omega).$$

(A2) $\forall (q, \psi) \in \mathbf{U}_\Lambda \times H^1(\Omega)$, $\mathcal{J}_0(q, \psi) \geq m > -\infty$.

(A3) $(q, \psi) \mapsto \mathcal{J}_0(q, \psi)$ est semi-continue inférieurement pour la topologie (faible*, faible) de $BV(\Omega) \times H^1(\Omega)$.

Alors le problème d'optimisation (2.7) admet au moins une solution optimale dans $\mathbf{U}_\Lambda \times H^1(\Omega)$.

Dans le cas où $\mathbf{U}_{\text{ad}} = \mathbf{U}_{\Lambda, \kappa}$, on a un résultat d'existence similaire.

Théorème 2.3 On suppose que $\mathbf{U}_{\text{ad}} = \mathbf{U}_{\Lambda, \kappa}$ et que la fonction coût $(q, \psi) \in \mathbf{U}_{\Lambda, \kappa} \mapsto \mathcal{J}(q, \psi) \in \mathbb{R}$ vérifie (A2) – (A3) données dans le Théorème 2.2 avec $\beta = 0$. Alors le problème d'optimisation (2.7) admet au moins une solution optimale.

Notons que pour $\mathbf{U}_{\text{ad}} = \mathbf{U}_\Lambda$ (voir Théorème (2.2)), la fonction coût doit contenir un terme de pénalisation de la forme $\beta |Dq|(\Omega)$ pour obtenir l'existence d'une solution optimale. Ce dernier permet en effet de contrôler la norme $BV(\Omega)$ d'une suite minimisante et ainsi d'extraire une sous-suite faiblement convergente. Dans le cas où $\mathbf{U}_{\text{ad}} = \mathbf{U}_{\Lambda, \kappa}$, la norme $BV(\Omega)$ de la suite minimisante est naturellement contrôlée car

$$\forall q \in \mathbf{U}_{\Lambda, \kappa} : \|q\|_{BV(\Omega)} \leq 2 \max(\Lambda, \kappa, |\alpha - 1|),$$

et on peut ainsi prendre $\beta = 0$.

1.2 Discrétisation du problème d'optimisation

Cette section étudie la convergence d'une discrétisation éléments finis du problème d'optimisation (2.7). On considère une famille quasi-uniforme de triangulations $\{\mathcal{T}_h\}_{h>0}$ de Ω et l'espace éléments finis \mathbb{P}_1 associé

$$\mathcal{V}_h = \{\phi_h \in \mathcal{C}(\overline{\Omega}) \mid \phi_h|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_h\}.$$

L'utilisation d'éléments finis continus est motivée par le résultat de régularité ci-dessous qui assure que toute solution faible de (2.5) est continue sur Ω .

Théorème 2.4 Soit $g \in L^2(\partial\Omega)$ et $q \in L^\infty(\Omega)$ tel que $1 + q(x) \geq \alpha > 0$ pour presque tout $x \in \Omega$. On note $\psi(q)$ la solution de (2.5). Alors $\psi(q) \in C^0(\Omega)$.

L'ensemble des bathymétries admissibles du problème d'optimisation discrétisé est défini par

$$\mathbf{U}_h = \mathbf{U}_{\text{ad}} \cap \mathcal{K}_h,$$

où \mathcal{K}_h est un espace éléments finis qui dépendra du choix de U_{ad} . La discrétisation du problème d'optimisation (2.7) est alors

$$\text{Trouver } q_h^* \in U_h \text{ tel que } \tilde{\mathcal{J}}(q_h^*) \leq \tilde{\mathcal{J}}(q_h), \forall q_h \in U_h, \quad (2.9)$$

où $\tilde{\mathcal{J}}(q_h) = \mathcal{J}(q_h, \psi_h(q_h))$ est la fonctionnelle réduite et $\psi_h = \psi_h(q_h) \in \mathcal{V}_h$ est la solution du problème de Helmholtz discrétisé

$$a(q_h; \psi_h, \phi_h) = b(q_h; \phi_h), \forall \phi_h \in \mathcal{V}_h. \quad (2.10)$$

La convergence de ψ_h vers ψ est précisée ci-dessous.

Théorème 2.5 *Soit $(q_h)_h \subset U_h$ une suite qui converge faible* vers q dans $BV(\Omega)$. Soit $(\psi_h(q_h))_h$ la suite des solutions du problème variationnel discret (2.10). Alors, lorsque h tend vers 0, $\psi_h(q_h)$ converge fortement dans $H^1(\Omega)$ vers $\psi(q)$ satisfaisant (2.5).*

Le Théorème 2.5 permet de démontrer la convergence des solutions optimales du problème d'optimisation discrétisé (2.9) vers celle du problème continu (2.7). Toutefois, selon l'espace de bathymétries admissibles considéré, l'espace admissible discret U_h est différent. Pour $U_{ad} = U_\wedge$, on a le résultat suivant :

Théorème 2.6 *Considérons l'ensemble admissible $U_{ad} = U_\wedge$ et l'ensemble U_h suivant*

$$U_h = U_\wedge \cap \mathcal{K}_{h,1},$$

$$\mathcal{K}_{h,1} = \{q_h \in L^\infty(\Omega) \mid q_h|_T \in \mathbb{P}_1(T), \forall T \in \mathcal{T}_h\}.$$

Supposons que les hypothèses (A1) – (A2) – (A3) du Théorème 2.2 soient satisfaites et que la fonction coût $\mathcal{J}_0 : (q, \psi) \in U_\wedge \times H^1(\Omega) \mapsto \mathcal{J}_0(q, \psi) \in \mathbb{R}$ est continue pour la topologie (faible, forte) de $BV(\Omega) \times H^1(\Omega)$. Soit $(q_h^*, \psi_h(q_h^*)) \in U_{\wedge, h} \times \mathcal{V}_h$ une paire optimale de (2.9).*

Alors la suite $(q_h^)_h \subset U_\wedge$ est bornée et il existe une sous-suite (la même notation est utilisée) et $q^* \in U_\wedge$ tels que $q_h^* \rightharpoonup q^*$ faible* dans $BV(\Omega)$, $\psi(q_h^*) \rightarrow \psi(q^*)$ fort dans $H^1(\Omega)$ et*

$$\tilde{\mathcal{J}}(q^*) \leq \tilde{\mathcal{J}}(q), \forall q \in U_\wedge.$$

Ainsi, tout point d'accumulation de $(q_h^, \psi_h(q_h^*))$ est une solution du problème d'optimisation (2.7).*

Dans le cas où $U_{ad} = U_{\wedge, \kappa}$, nous avons un résultat similaire au Théorème 2.6 en modifiant l'ensemble U_h .

Théorème 2.7 *Considérons l'ensemble admissible $U_{ad} = U_{\wedge, \kappa}$ et l'ensemble U_h suivant*

$$U_h = U_{\wedge, \kappa} \cap \mathcal{K}_{h,0},$$

$$\mathcal{K}_{h,0} = \{q_h \in L^\infty(\Omega) \mid q_h|_T \in \mathbb{P}_0(T), \forall T \in \mathcal{T}_h\}.$$

Supposons que $\beta = 0$ et que les hypothèses (A2) – (A3) du Théorème 2.2 sont satisfaites. On suppose également que $\mathcal{J} : (q, \psi) \in \mathcal{U}_\Lambda \times H^1(\Omega) \mapsto \mathcal{J}(q, \psi) \in \mathbb{R}$ est continue pour la topologie (faible*, forte) de $BV(\Omega) \times H^1(\Omega)$. Soit $(q_h^*, \psi_h(q_h^*)) \in \mathcal{U}_h \times \mathcal{V}_h$ une paire optimale de (2.9). Alors les conclusions du Théorème 2.6 sont également valables.

L'ensemble \mathcal{U}_h du Théorème 2.6 est composé d'éléments finis \mathbb{P}_1 tandis que, dans le Théorème 2.7, \mathcal{U}_h est composé d'éléments \mathbb{P}_0 . Cette différence provient essentiellement des hypothèses assurant l'existence d'une solution optimale qui imposent d'avoir $\beta > 0$ lorsque $\mathcal{U}_{\text{ad}} = \mathcal{U}_\Lambda$ (voir Théorème 2.2). En effet, pour démontrer le Théorème 2.6, il faut passer à la limite dans l'inégalité (2.9). Ce passage à la limite nécessite d'avoir un ensemble \mathcal{U}_h dense dans \mathcal{U}_Λ pour la topologie forte de $BV(\Omega)$ ce qui est le cas avec des éléments finis \mathbb{P}_1 mais pas avec des éléments \mathbb{P}_0 . Lorsque $\mathcal{U}_{\text{ad}} = \mathcal{U}_{\Lambda, \kappa}$, le passage à la limite dans l'inégalité (2.9) peut s'effectuer pour un ensemble \mathcal{U}_h dense dans $\mathcal{U}_{\Lambda, \kappa}$ pour la topologie faible* de $BV(\Omega)$ permettant donc de discrétiser les bathymétries avec des éléments \mathbb{P}_0 .

1.3 Résultats numériques

On présente maintenant des résultats numériques pour lesquels les hypothèses des Théorèmes 2.6 et 2.7 assurant la convergence des optimums discrets vers les optimums continus ne sont pas vérifiées. Pour cela, les bathymétries sont approximées par des éléments de $\mathcal{K}_{h,0}$ mais la contrainte $|Dq_h|(\Omega) \leq \kappa$ n'est pas imposée et ainsi le Théorème 2.7 ne s'applique plus.

Pour le cas test présenté ici, on considère une onde plane $\psi_0(x) = e^{ik_0 x \cdot \mathbf{d}}$ avec $\mathbf{d} = (0, 1)^\top$ et

$$k_0 = \frac{\omega_0}{\sqrt{gz_0}}, \quad \omega_0 = \frac{2\pi}{T_0}, \quad T_0 = 20, \quad g = 9.81, \quad z_0 = 3.$$

Le domaine de calcul est $\Omega = [0, L]^2$ où $L = \frac{10\pi}{k_0} \approx 542.5$. Le domaine Ω est discrétisé avec un maillage structuré composé de 8192 triangles et le pas de maillage est $\Delta x = \Delta y = 8.476472$.

La résolution du problème d'optimisation est effectuée avec une méthode de type *subspace trust-region* telle que celle décrite dans [17] et [18]. Chaque itération nécessite la résolution d'un système linéaire qui est réalisée avec un gradient conjugué. On impose également une contrainte L^∞ sur $q_h : 1 \geq q_h \geq -0.9$.

Pour définir le cas test, on considère la bathymétrie

$$q_{\text{ref}}(x, y) = e^{-\tau \left((x - \frac{L}{4})^2 + (y - \frac{L}{4})^2 \right)} + e^{-\tau \left((x - \frac{3L}{4})^2 + (y - \frac{3L}{4})^2 \right)},$$

avec $\tau = 10^{-3}$. L'objectif est de reconstruire q_{ref} sur le domaine $\Omega_q = [\frac{L}{8}, \frac{3L}{8}]^2 \cup [\frac{5L}{8}, \frac{7L}{8}]^2$, en minimisant la fonction coût

$$\mathcal{J}(q, \psi_{\text{tot}}) = \frac{\omega_0^2}{2} \int_{\Omega_0} |\psi_{\text{tot}}(x, y) - \psi_{\text{ref}}(x, y)|^2 dx dy,$$

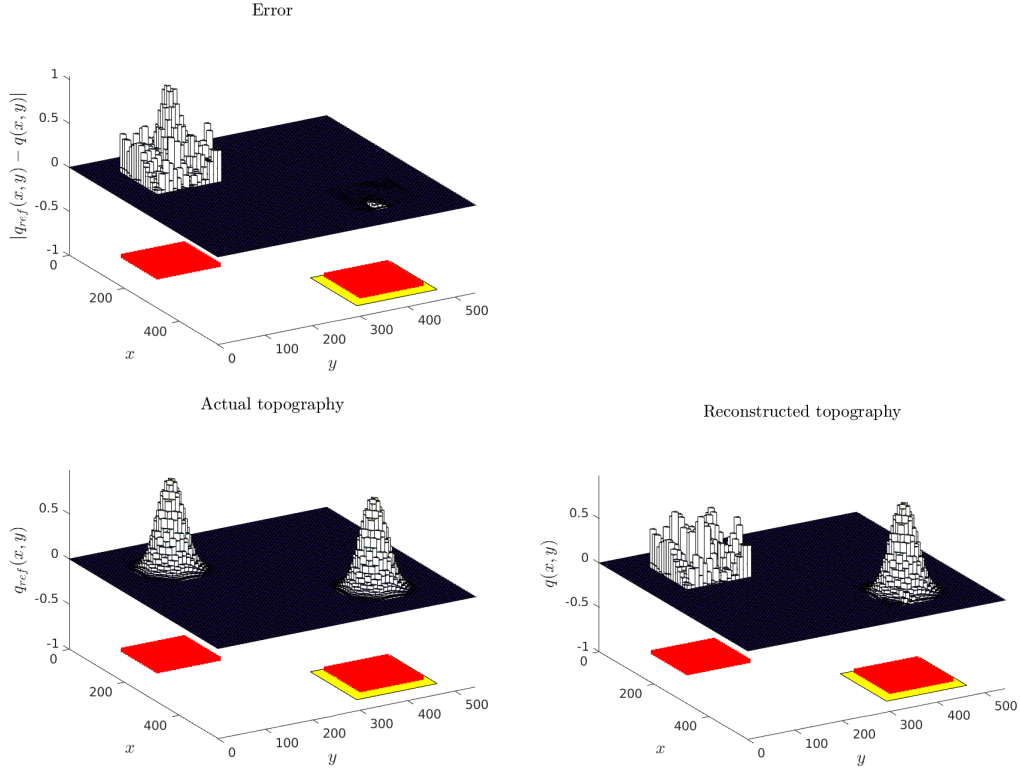


FIGURE 2.1 – (En haut) Erreur entre q_{ref} et la solution numérique de (2.9); (En bas à droite) Représentation de q_{ref} ; (En bas à gauche) solution numérique de (2.9). Les zones jaunes représentent Ω_0 et les zones rouges correspondent à Ω_q , le support de q_{ref} .

où ψ_{ref} est la solution de (2.10) avec $q_h = q_{\text{ref}}$ et $\Omega_0 = [\frac{3L}{4} - \delta, \frac{3L}{4} + \delta]^2$, $\delta = \frac{L}{6}$.

Les résultats sont donnés dans la Figure 2.1. On peut voir notamment que la bathymétrie n'est pas reconstruite dans les zones qui n'appartiennent pas au domaine d'observation Ω_0 . Au contraire, la bathymétrie est très bien retrouvée dans $\Omega_0 \cap \Omega_q$ ce qui correspond aux zones appartenant au support de q_{ref} .

Dans cet expérience numérique, les hypothèses du Théorème 2.7 ne sont pas satisfaites car on n'impose pas $|Dq_h|(\Omega) \leq \kappa$. La Figure 2.2 représente $|Dq_h^*(\Omega)|$ pour différentes valeurs de h . On peut y observer que $|Dq_h^*(\Omega)| = \mathcal{O}(h^{-s})$ pour un certain $s > 0$. Ces résultats numériques montrent l'importance d'imposer une borne supérieure sur $|Dq_h|$ (soit dans l'espace admissible U_{ad} ou bien dans la fonction coût) pour démontrer l'existence de bathymétries optimales (Théorèmes 2.2 et 2.3) et la convergence des optimum discrets vers les optimum continus (Théorèmes 2.6 et 2.7).

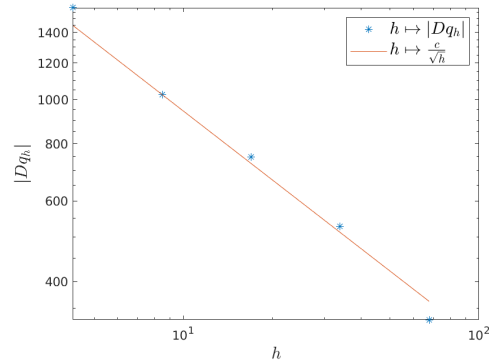


FIGURE 2.2 – Valeur de $|Dq_h^*|(\Omega)$ pour différents h .

2 Optimisation topologique visant des solides optimisés avec des conductivités constantes par morceaux

Cette section présente des résultats issus de [Chap2-A5].

Un problème d'optimisation topologique (voir par exemple [1]) consiste à déterminer la localisation d'un solide au sein d'un fluide pour minimiser une fonction coût. La résolution d'un tel problème nécessite donc d'utiliser la localisation et les caractéristiques physiques du matériau recherché comme paramètres d'optimisation. Pour sa simplicité d'utilisation, le modèle de pénalisation pour les équations de Navier-Stokes introduit dans [2] est majoritairement adopté. Ce dernier permet de localiser le solide avec l'aide d'un terme de friction, faisant intervenir une fonction indicatrice du solide, qui définit la zone fluide du domaine lorsqu'il est nul et la zone solide lorsqu'il tend vers l'infini. La localisation du solide étant réalisée avec un paramètre discret, le paramètre de contrôle appartient à un ensemble qui n'est ni convexe ni fermé pour la topologie faible* de L^∞ . Ainsi, l'étude mathématique et numérique de ce problème d'optimisation est, comme indiqué dans [20], difficile.

Afin de passer outre ces difficultés, T. Borrvall et J. Petersson [7] ont considéré une variable d'optimisation α définie sur tout le domaine de calcul et pouvant prendre des valeurs continues entre 0 et 1. Ils ont ensuite régularisé l'indicatrice intervenant dans le terme de pénalisation avec ce qu'on appellera une fonction d'interpolation fluide-solide permettant ainsi de localiser les zones fluides et solides en fonction des valeurs de α . Un problème d'optimisation topologique revient alors à déterminer le paramètre α qui minimise la fonction coût considérée.

L'étude mathématique (par exemple : existence, unicité de solution faible ; existence de solution au problème d'optimisation, ...) de problèmes d'optimisation topologique a été réalisée pour un écoulement de Stokes stationnaire dans [7, 20] et pour le modèle de Navier-Stokes stationnaire dans [21]. Les travaux précédemment mentionnés considèrent des conditions limites de Dirichlet homogènes ce qui permet notamment de simplifier la dérivation d'estimations a priori pour la solution du système de Navier-Stokes incom-

pressible [8, 34]. Un des apports principaux de [Chap2-A5] est l'étude d'un problème d'optimisation topologique pour des écoulements anisothermes modélisés par le système de Navier-Stokes, sous l'approximation de Boussinesq, couplé à une équation d'énergie avec des conditions limites mixtes.

Un autre apport de [Chap2-A5] a été de proposer une fonction d'interpolation fluide-solide permettant de localiser des solides ayant des conductivités thermiques constantes par morceaux. Par rapport à d'autres approches proposées pour obtenir de tels solides optimisés (voir par exemple [28, 38]), notre fonction d'interpolation est régulière et n'utilise qu'une seule variable d'optimisation additionnelle pour interpoler les différentes constantes des conductivités ciblées.

On décrit maintenant la procédure pour obtenir le modèle utilisé comme contrainte dans un problème d'optimisation topologique. Considérons pour cela $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$ tel que $\Omega = \Omega_f \cup \Omega_s$ où le fluide occupe la zone Ω_f et le solide la zone Ω_s . Le système de Navier-Stokes incompressible sous l'approximation de Boussinesq sous sa forme sans dimension est :

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 & \text{dans } \Omega_f, \\ (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - \text{Re}^{-1} \Delta \mathbf{u} - \text{Ri} \theta \mathbf{e}_y &= 0 & \text{dans } \Omega_f, \\ \nabla \cdot (\mathbf{u} \theta) - \nabla \cdot (\text{Re}^{-1} \text{Pr}^{-1} k(x) \nabla \theta) &= 0 & \text{dans } \Omega, \\ \mathbf{u} &= 0 & \text{dans } \Omega_s, \end{aligned} \tag{2.11}$$

où $\text{Re} = (V_0 L_0) / \nu$ est le nombre de Reynolds, $\text{Ri} = (g \beta L_0 \delta T) / V_0^2$ le nombre de Richardson, $\text{Pr} = \rho_0 \nu c_p / k_f$ le nombre de Prandtl. La fonction $x \in \Omega \mapsto k(x)$ est définie par

$$k(x) = \frac{1}{k_f} (k_f \mathbf{1}_{\Omega_s}(x) + k_s (1 - \mathbf{1}_{\Omega_s}(x))),$$

et représente le ratio des conductivités thermiques. Les quantités de référence V_0, L_0, ρ_0 sont respectivement une vitesse, une longueur et la masse volumique du fluide à une température de référence T_0 . De plus, β est le coefficient de dilatation thermique du fluide, ν sa viscosité cinématique, $\delta T = -\phi / k_f$ avec ϕ un flux thermique sur une paroi et k_f la conductivité du fluide et enfin c_p est la capacité calorifique du fluide à pression constante.

En suivant [2], le solide peut être localisé avec un terme de pénalisation de la forme $\eta^{-1} \mathbf{1}_{\Omega_s} \mathbf{u}$ qui est ajouté dans le système (2.11). Ce terme permet, en faisant tendre η vers 0, de récupérer que $\mathbf{u}|_{\Omega_s} = 0$ et $\mathbf{u}|_{\partial \Omega_s} = 0$ correspondant à une condition de non-glissement sur les parois solides (voir [2, Lemma 4.4 et Theorem 4.1]). La version pénalisée du système (2.11) est alors

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 & \text{dans } \Omega, \\ (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - \text{Re}^{-1} \Delta \mathbf{u} - \text{Ri} \theta \mathbf{e}_y + \frac{1}{\eta} \mathbf{1}_{\Omega_s}(x) \mathbf{u} &= 0 & \text{dans } \Omega, \\ \nabla \cdot (\mathbf{u} \theta) - \nabla \cdot (\text{Re}^{-1} \text{Pr}^{-1} k(x) \nabla \theta) &= 0 & \text{dans } \Omega. \end{aligned} \tag{2.12}$$

Remarque 2.8 Notons que k_τ représente le ratio des conductivités du solide par celle du fluide car nous travaillons avec la forme adimensionnée des équations. Par abus de langage, on parlera uniquement de conductivité pour les valeurs prises par k_τ .

La dernière étape pour définir le modèle utilisé comme contrainte dans le problème d'optimisation topologique consiste à introduire une fonction $\alpha : x \in \Omega \mapsto \mathbb{R}^+$ et à régulariser l'indicatrice du solide avec une fonction $h_\tau(\alpha(x))$ vérifiant

$$\begin{aligned} h_\tau(s) &\xrightarrow{\tau \rightarrow +\infty} 0 \text{ pour } s < \alpha_0, \\ h_\tau(s) &\xrightarrow{\tau \rightarrow +\infty} \alpha_{\max} \text{ pour } s \geq \alpha_0, \end{aligned}$$

où $\alpha_{\max} = \eta^{-1}$ et la convergence ci-dessus est ponctuelle. Les zones fluide et solide dans Ω peuvent alors être déterminées (de façon approchée) par

$$\Omega_f := \{x \in \Omega \mid \alpha(x) < \alpha_0\}, \quad \Omega_s := \{x \in \Omega \mid \alpha(x) \geq \alpha_0\}, \quad (2.13)$$

et nous pouvons désormais utiliser $\alpha : x \in \Omega \mapsto \alpha(x) \in [0, \alpha_{\max}]$ comme paramètre d'optimisation. Dans [Chap2-A5], nous avons proposé une fonction d'interpolation fluide-solide permettant d'obtenir des solides optimisés ayant des conductivités thermiques constantes par morceaux (voir Section 2.1). La variable d'optimisation gérant les différentes constantes est notée φ et on note ainsi $\xi = (\alpha, \varphi)$ les variables d'optimisation.

Le système utilisé comme contrainte dans le problème d'optimisation topologique étudié dans [Chap2-A5] est enfin

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \quad \text{dans } \Omega, \\ \nabla p - A\Delta \mathbf{u} - (\mathbf{u} \cdot \nabla) \mathbf{u} + h_\tau(\alpha) \mathbf{u} - B\theta \mathbf{e}_y &= 0 \quad \text{dans } \Omega, \\ -\nabla \cdot (Ck_\tau(\alpha, \varphi) \nabla \theta) + \nabla \cdot (\mathbf{u}\theta) &= 0 \quad \text{dans } \Omega, \end{aligned} \quad (2.14)$$

où A, B, C sont les constantes physiques ci-dessous

$$A = \frac{1}{\text{Re}}, \quad B = \text{Ri}, \quad C = \frac{1}{\text{Re Pr}}.$$

Les conditions aux limites considérées sont mixtes et sont définies sur $\partial\Omega = \Gamma$ décomposé selon $\Gamma = \Gamma_w \cup \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$ tels que $|\Gamma_w| > 0$, $|\Gamma_{\text{in}}| > 0$, $|\Gamma_{\text{out}}| > 0$ et $\overline{\Gamma_{\text{in}}} \cap \overline{\Gamma_{\text{out}}} = \emptyset$. La partie Γ_w concerne des conditions de non-glissement pour la vitesse, Γ_{in} représente l'entrée du domaine et Γ_{out} sa sortie. Les conditions limites correspondantes sont

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_{\text{in}}, & \theta &= 0, & \text{sur } \Gamma_{\text{in}}, \\ \mathbf{u} &= 0, & k_\tau \partial_n \theta &= \phi, & \text{sur } \Gamma_w, \\ A \partial_n \mathbf{u} - \mathbf{n}p &= 0, & \partial_n \theta &= 0, & \text{sur } \Gamma_{\text{out}}, \end{aligned} \quad (2.15)$$

où $\mathbf{u}_{\text{in}} \in H_{00}^{1/2}(\Gamma_{\text{in}})^d$ est une vitesse en entrée¹, ϕ est un flux thermique, \mathbf{n} est la normale unitaire sortante et ∂_n désigne la dérivée normale.

Pour définir la formulation variationnelle de (2.14)-(2.15), on introduit les ensembles suivants

$$\begin{aligned} \mathcal{X}_1^u &= \left\{ \mathbf{v} \in H^1(\Omega)^d \mid \mathbf{v}|_{\Gamma_w} = 0 \right\}, \\ \mathcal{X}^u &= \left\{ \mathbf{v} \in H^1(\Omega)^d \mid \mathbf{v}|_{(\Gamma_w \cup \Gamma_{\text{in}})} = 0 \right\}, \\ \mathcal{X}^\theta &= \left\{ \theta \in H^1(\Omega) \mid \theta|_{\Gamma_{\text{in}}} = 0 \right\}. \end{aligned}$$

1. La notation $H_{00}^{1/2}(\Gamma_{\text{in}})$ désigne l'ensemble des $\psi \in H^{1/2}(\Gamma_{\text{in}})$ tels que l'extension de ψ par zéro à Γ est dans $H^{1/2}(\Gamma)$.

Pour traiter la condition limite de Dirichlet inhomogène sur Γ_{in} , on utilise le [16, Lemma 16] qui assure l'existence d'un $\mathbf{V} \in H^1(\Omega)^d$ tel que $\nabla \cdot \mathbf{V} = 0$, $\mathbf{V}|_{\Gamma_w} = 0$, $\mathbf{V}|_{\Gamma_{\text{in}}} = \mathbf{u}_{\text{in}}$ et $\|\mathbf{V}\|_{H^1(\Omega)^d} \leq M_{\mathbf{V}} \|\mathbf{u}_{\text{in}}\|_{H_0^{1/2}(\Gamma_{\text{in}})^d}$ pour $M_{\mathbf{V}} > 0$.

En introduisant

$$\mathbf{u} = \mathbf{w} + \mathbf{V},$$

où $\mathbf{w} \in X^u$, la formulation variationnelle de (2.14)-(2.15) est

Trouver $(\mathbf{w}, \theta, p) \in X^u \times X^\theta \times L^2(\Omega)$ tels que :

$$\begin{cases} \mathbf{a}(\xi; \mathbf{w}, \mathbf{v}_1) + \mathbf{b}(\mathbf{v}_1, p) = \langle \mathbf{G}(\xi; \mathbf{w}) + \mathbf{F}(\theta), \mathbf{v}_1 \rangle, & \forall \mathbf{v}_1 \in X^u, \\ \tilde{\mathbf{a}}(\xi; \theta, v_2) + \tilde{\mathbf{c}}(\theta, \mathbf{w} + \mathbf{V}, v_2) = \langle \tilde{\mathbf{G}}, v_2 \rangle, & \forall v_2 \in X^\theta, \\ \mathbf{b}(\mathbf{w}, q) = 0, & \forall q \in L^2(\Omega), \end{cases} \quad (2.16)$$

où les différentes formes multilinéaires sont définies par

$$\begin{aligned} \mathbf{a}(\xi; \mathbf{u}, \mathbf{v}_1) &= \int_{\Omega} [A \nabla \mathbf{u} : \nabla \mathbf{v}_1 + h_\tau(\xi) \mathbf{u} \cdot \mathbf{v}_1] \, dx, \\ \mathbf{b}(\mathbf{u}, q) &= - \int_{\Omega} q (\nabla \cdot \mathbf{u}) \, dx, \quad \mathbf{c}(\mathbf{u}, \mathbf{v}_1, \mathbf{w}) = \int_{\Omega} (\mathbf{u} \cdot \nabla) \mathbf{v}_1 \cdot \mathbf{w} \, dx, \quad f(\theta, \mathbf{v}_1) = - \int_{\Omega} B \theta \mathbf{e}_y \cdot \mathbf{v}_1 \, dx, \\ \tilde{\mathbf{a}}(\xi; \theta, v_2) &= \int_{\Omega} C k_\tau(\xi) \nabla \theta \cdot \nabla v_2 \, dx, \quad \tilde{\mathbf{c}}(\theta, \mathbf{u}, v_2) = \int_{\Omega} (\nabla \theta \cdot \mathbf{u}) v_2 \, dx, \end{aligned}$$

et les termes sources sont

$$\begin{aligned} \langle \mathbf{G}(\xi; \mathbf{w}) + \mathbf{F}(\theta), \mathbf{v} \rangle &= -c(\mathbf{w} + \mathbf{V}, \mathbf{w} + \mathbf{V}, \mathbf{v}) - f(\theta, \mathbf{v}) - \mathbf{a}(\xi; \mathbf{V}, \mathbf{v}), \\ \langle \tilde{\mathbf{G}}, v_2 \rangle &= \int_{\Gamma_w} C \phi v_2 \, d\sigma. \end{aligned}$$

Le problème d'optimisation considéré est enfin donné par

$$\begin{aligned} \min \mathcal{J}(\xi, \mathbf{w}, \theta, p) \\ \text{où } \begin{cases} (\mathbf{w}, \theta, p) \text{ est solution de (2.16) paramétrée par } \xi, \\ \xi \in \mathcal{U}_{\text{ad}}, \end{cases} \end{aligned} \quad (2.17)$$

où \mathcal{J} est une fonction coût. De plus, on recherche des ξ admissibles dans l'ensemble \mathcal{U}_{ad} donné ci-dessous :

$$\xi \in \mathcal{U}_{\text{ad}} = \{\xi \in \text{BV}(\Omega)^m : 0 \leq \xi(x) \leq \xi_{\text{max}} \text{ p.p. dans } \Omega, |\text{D}\xi|(\Omega) \leq \kappa\},$$

où les inégalités sont comprises élément par élément et $\xi_{\text{max}} \in \mathbb{R}^m$, $\xi_{\text{max}} > 0$ et $\kappa > 0$.

Dans la suite, on présente tout d'abord la fonction d'interpolation fluide-solide permettant de prendre en compte des solides avec des conductivités constantes par morceaux (voir Section 2.1). Nous donnerons ensuite un résultat d'existence et l'unicité d'une solution faible au problème variationnel (2.16) ainsi que la convergence d'une approximation éléments finis avec des éléments de Taylor-Hood. L'existence d'au moins une solution au problème d'optimisation (2.17) et la convergence d'un optimum discret vers un optimum continu sera ensuite présentée. Après avoir dérivé des conditions nécessaires d'optimalité du premier ordre, cette partie terminera avec des simulations numériques illustrant certains résultats théoriques.

2.1 Fonctions d'interpolation fluide-solide avec une conductivité constante par morceaux

On commence par définir la fonction d'interpolation fluide-solide h_τ intervenant dans (2.14). Elle est basée sur la régularisation de la fonction de Heaviside ci-dessous :

$$\tilde{h}_\tau(y, y_0, a, b) = a + (b - a) \left(\frac{1}{1 + \exp(-\tau(y - y_0))} - \frac{1}{1 + \exp(\tau y_0)} \right), \quad (2.18)$$

où $y \in [0, y_{\max}]$ dont on peut vérifier directement la convergence ponctuelle suivante

$$\lim_{\tau \rightarrow +\infty} \tilde{h}_\tau(y, y_0, a, b) = \begin{cases} a & \text{if } y < y_0, \\ (a + b)/2 & \text{if } y = y_0, \\ b & \text{if } y > y_0. \end{cases} \quad (2.19)$$

On peut alors définir la fonction h_τ par

$$h_\tau(\alpha) = \tilde{h}_\tau(\alpha, \alpha_0, 0, \alpha_{\max}), \quad (2.20)$$

pour un $\alpha_0 \in (0, \alpha_{\max})$ fixé, une constante α_{\max} assez grande pour pénaliser la vitesse du fluide dans le solide qui sera représenté par les zones dans lesquelles $\alpha(x) \geq \alpha_0$ (voir également (2.13)).

On présente maintenant la fonction d'interpolation permettant d'obtenir des solides optimisés ayant des conductivités thermiques constantes par morceaux. Pour $j = 1, \dots, N$, on note k_j les différentes valeurs de conductivités ciblées et φ la variable d'optimisation permettant de les interpoler. On définit maintenant k_τ comme suit

$$k_\tau(\alpha, \varphi) = \frac{1}{k_f} (k_f + \tilde{h}_\tau(\alpha, \alpha_0, 0, 1)(\chi_\tau(\varphi) - k_f)), \quad (2.21)$$

où k_f est la conductivité du fluide et χ_τ va permettre d'interpoler les différents k_j en fonction des valeurs prises par φ . Les zones fluides de Ω pouvant être déterminées par les $x \in \Omega$ tels que $\alpha(x) < \alpha_0$, la propriété (2.19) entraîne

$$\lim_{\tau \rightarrow +\infty} k_\tau(\alpha(x), \varphi(x)) = 1.$$

Les zones solides étant déterminées par les $x \in \Omega$ tels que $\alpha(x) \geq \alpha_0$, on obtient

$$\lim_{\tau \rightarrow +\infty} k_\tau(\alpha(x), \varphi(x)) = (k_f + \lim_{\tau \rightarrow +\infty} (\chi_\tau(\varphi(x)) - k_f))/k_f = \lim_{\tau \rightarrow +\infty} \chi_\tau(\varphi(x))/k_f.$$

On peut ainsi définir χ_τ comme suit

$$\chi_\tau(\varphi) = k_1 + \sum_{j=1}^{N-1} \tilde{h}_\tau(\varphi, \varphi_j, 0, a_j), \quad (2.22)$$

où $\varphi_j \in [0, \varphi_{\max}]$ et $\varphi_i < \varphi_j$ pour tout $i < j$, $i, j \in \{1, \dots, N-1\}$. Les constantes a_j sont ensuite déterminées en imposant que

$$\lim_{\tau \rightarrow +\infty} \chi_\tau(\varphi) = k_j \text{ pour } \varphi_{j-1} < \varphi < \varphi_j, \quad 2 \leq j \leq N. \quad (2.23)$$

En utilisant (2.19), les conditions (2.23) donnent le système linéaire triangulaire

$$\sum_{j=1}^i a_j = k_{i+1} - k_1, i = 1, 2, \dots, N - 1, \quad (2.24)$$

qui permet de déterminer les a_j et ainsi de définir complètement la fonction χ_τ . La Figure 2.3 représente les deux fonctions d'interpolation fluide-solide décrite ci-dessus.

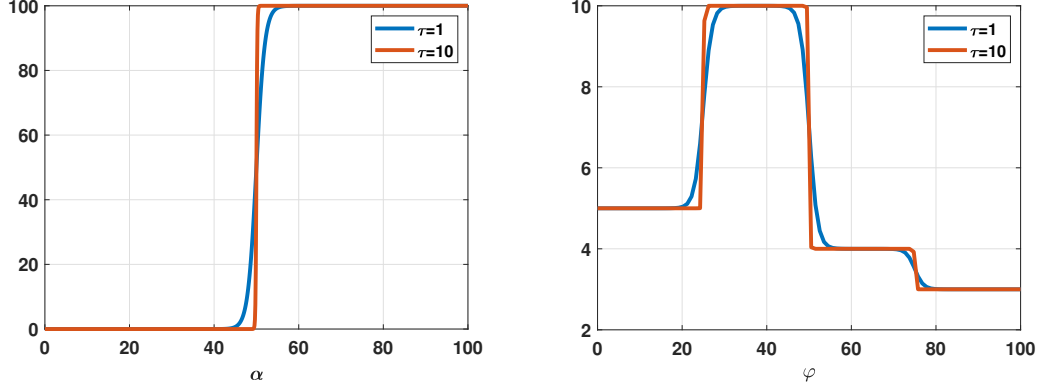


FIGURE 2.3 – (Gauche) $\tilde{h}_\tau(\xi, 50, 0, 100)$ pour différentes valeurs de τ . (Droite) χ_τ avec $\phi_{\max} = 100$ et $k_j \in \{5, 10, 4, 3\}$ pour différentes valeurs de τ .

2.2 Existence et unicité de solution faible et convergence d'une approximation éléments finis

L'existence et l'unicité d'une solution pour (2.16) est donnée dans le résultat ci-dessous :

Théorème 2.9 *On suppose que les fonctions k_τ et h_τ sont bornées et positives sur leur ensemble respectif de définition et qu'il existe $k_{\min} > 0$ tel que, pour tout $\xi \in \mathcal{U}_{\text{ad}}$, on ait $k_{\min} \leq k_\tau(\xi)$. Alors, si B , $\|u_{in}\|_{H_{00}^{1/2}(\Gamma_{in})}$ et $\|\Phi\|_{L^2(\Gamma)}$ sont suffisamment petit, il existe une unique solution faible de (2.16).*

La démonstration du Théorème 2.9 est basée sur un point fixe, ce qui impose d'avoir des termes sources assez petits pour fonctionner. Les bornes supérieures sur les termes sources dépendent de k_{\min} , de la constante intervenant dans la condition inf-sup vérifiée par la forme bilinéaire $b(\cdot, \cdot)$ et d'autres constantes intervenant dans l'inégalité de Poincaré ou dans des inclusions continues entre espaces de Sobolev. Un point important reste que si les termes sources sont assez petits alors la solution faible de (2.16) vérifie

$$\|\mathbf{w}\|_{X^u} + \|\theta\|_{X^\theta} + \|\mathbf{p}\|_{L^2(\Omega)} \leq R,$$

où $R > 0$ ne dépend pas de τ ni du paramètre d'optimisation ξ . Les normes ci-dessus sont définies par $\|\cdot\|_{X^u} = \sqrt{a(\xi; \cdot, \cdot)}$ et $\|\cdot\|_{X^\theta} = \|\cdot\|_{H^1(\Omega)}$. En supposant que la zone solide

Ω_s est connue et en prenant $h_\tau = \frac{1}{\tau}1_{\Omega_s}$, le champ de vitesse associé à (2.16) satisfait $\|\sqrt{h_\tau}\mathbf{w}\|_{L^2(\Omega)} = \frac{1}{\sqrt{\tau}}\|\mathbf{w}\|_{L^2(\Omega_s)} \leq R$ ce qui implique

$$\|\mathbf{w}\|_{L^2(\Omega_s)} \leq R\sqrt{\tau}.$$

De plus,

$$\|\mathbf{w}\|_{L^2(\partial\Omega_s)} \leq C\sqrt{\|\mathbf{w}\|_{L^2(\Omega_s)}\|\mathbf{w}\|_{H^1(\Omega)}} \leq C\sqrt{R\tau}^{1/4}\sqrt{\|\mathbf{w}\|_{X^u}} \leq CR\tau^{1/4},$$

et on retrouve ainsi le fait que, lorsque $\tau \rightarrow 0$, le champ de vitesse associé au problème pénalisé tend vers 0 dans le solide avec la condition de non-glissement en paroi. On a également les mêmes ordres de convergence en τ que ceux obtenus dans le cas d'une condition de Dirichlet homogène sur $\partial\Omega$ et sans prise en compte des transferts de chaleur dans le fluide (voir [2, Corollary 4.1, Lemma 4.4]).

La discrétisation de (2.16) est réalisée avec une triangulation quasi-uniforme de Ω de sorte que $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$. On note h_K une taille caractéristique de l'élément K et $h = \sup_{K \in \mathcal{T}_h} h_K$. Les éléments finis de Taylor-Hood [33] sont ensuite utilisés pour approximer les solutions de (2.16). On cherche donc des approximations $(\mathbf{w}_h, \theta_h, p_h) \in X_h^u \times X_h^\theta \times M_h$ où

$$\begin{aligned} X_h^u &= \{\mathbf{v}_h \in X^u \mid \forall K \in \mathcal{T}_h, \mathbf{v}_h|_K \in \mathbb{P}_2(K)\}, \\ X_h^\theta &= \{\psi_h \in X^\theta \mid \forall K \in \mathcal{T}_h, \psi_h|_K \in \mathbb{P}_2(K)\}, \\ M_h &= \{q_h \in \mathcal{C}^0(\bar{\Omega}) \mid \forall K \in \mathcal{T}_h, q_h|_K \in \mathbb{P}_1(K)\}. \end{aligned}$$

Pour discrétiser également les paramètres d'optimisation ξ , on utilise l'ensemble suivant

$$U_h = U_{ad} \cap \mathcal{K}_h \text{ avec } \mathcal{K}_h = \{\xi_h \in L^\infty(\Omega)^2 \mid \xi_h|_K \in \mathbb{P}_0(T)^2, \forall K \in \mathcal{T}_h\},$$

qui correspond à des fonctions constantes sur chaque maille.

Pour $\xi_h \in \mathcal{K}_h$, la discrétisation de (2.16) est alors

Trouver $(\mathbf{w}_h, \theta_h, p_h) \in X_h^u \times X_h^\theta \times M_h$ tel que :

$$\begin{cases} a(\xi_h; \mathbf{w}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = \langle G(\xi_h; \mathbf{w}_h) + F(\theta_h), \mathbf{v}_h \rangle, & \forall \mathbf{v}_h \in X_h^u, \\ \tilde{a}(\xi_h; \theta_h, \psi_h) + \tilde{c}(\theta_h, \mathbf{w}_h + \mathbf{V}, \psi_h) = \langle \tilde{G}, \psi_h \rangle, & \forall \psi_h \in X_h^\theta, \\ b(\mathbf{w}_h, q_h) = 0, & \forall q_h \in M_h. \end{cases} \quad (2.25)$$

Le résultat de convergence ci-dessous a été démontré.

Théorème 2.10 *On suppose que B , $\|\mathbf{u}_{in}\|_{H_{00}^{1/2}(\Gamma_{in})}$ et $\|\Phi\|_{L^2(\Gamma)}$ sont assez petits pour assurer l'existence et l'unicité d'une solution faible pour (2.16) paramétrée par ξ . Soit $\xi_h \in U_h$ une suite qui converge faible* vers ξ dans $BV(\Omega)^2$. Alors, pour h assez petit, il existe une unique solution $(\mathbf{w}_h, \theta_h, p_h)$ de (2.25) et on a de plus*

$$\lim_{h \rightarrow 0} \left(\|\mathbf{w}_h - \mathbf{w}\|_{X^u} + \|\theta_h - \theta\|_{X^\theta} + \|p_h - p\|_{L^2(\Omega)} \right) = 0.$$

L'existence et l'unicité d'une solution au problème (2.25) sont obtenues comme pour (2.16) ce qui nécessite d'avoir des termes sources assez petits. La restriction sur le pas de maillage provient de la validité de la condition inf-sup discrète. Sous ces hypothèses, l'unique solution de (2.25) vérifie

$$\|\mathbf{w}_h\|_{X^u} + \|\theta_h\|_{X^\theta} + \|\mathbf{p}_h\|_{L^2(\Omega)} \leq R,$$

où R est indépendant de h . Nous pouvons alors extraire de la suite $(\mathbf{w}_h, \theta_h, \mathbf{p}_h)_h$ une sous-suite faiblement convergente et la démonstration du Théorème 2.10 est réalisée, en particulier, en passant à la limite dans la formulation variationnelle (2.25).

2.3 Existence d'au moins une solution optimale et approximation éléments finis

Pour chaque $\xi \in \mathcal{U}_{\text{ad}}$, le Théorème 2.9 donne l'existence et l'unicité de $(\mathbf{w}(\xi), \theta(\xi), \mathbf{p}(\xi))$ satisfaisant (2.16). On peut de plus démontrer que l'application

$$\xi \in (\mathcal{U}_{\text{ad}}, \text{faible}^*) \mapsto (\mathbf{w}(\xi), \theta(\xi), \mathbf{p}(\xi)) \in (X^u \times X^\theta \times L^2(\Omega), \text{forte})$$

est continue. En utilisant cette propriété et des suites minimisantes, on peut montrer l'existence d'au moins une solution au problème d'optimisation (2.17).

Théorème 2.11 *On suppose que*

(A1) $\inf_{\mathcal{U}_{\text{ad}} \times X^u \times X^\theta \times L^2(\Omega)} \mathcal{J} > -\infty.$

(A2) \mathcal{J} est semi-continue inférieurement pour la topologie (faible*, faible, faible, faible) de $BV(\Omega)^2 \times X^u \times X^\theta \times L^2(\Omega).$

Alors le problème d'optimisation (2.17) admet au moins une solution dans $\mathcal{U}_{\text{ad}} \times X^u \times X^\theta \times L^2(\Omega).$

La discrétisation éléments finis du problème d'optimisation (2.17) est

$$\begin{aligned} & \min \mathcal{J}(\xi_h, \mathbf{w}_h, \theta_h, \mathbf{p}_h) \\ & \text{tels que } \begin{cases} (\mathbf{w}_h, \theta_h, \mathbf{p}_h) \text{ vérifient (2.25) paramétré par } \xi_h, \\ \xi_h \in \mathcal{U}_h = \mathcal{U}_{\text{ad}} \cap \mathcal{K}_h. \end{cases} \end{aligned} \quad (2.26)$$

En utilisant des techniques similaires à celles utilisées pour la démonstration du Théorème 2.6, nous avons obtenu le résultat de convergence suivant.

Théorème 2.12 *On suppose que (A1)-(A2) du Théorème 2.11 sont satisfaites et que \mathcal{J} est continue pour la topologie (faible*, forte, forte, forte) de $BV(\Omega)^m \times X^u \times X^\theta \times L^2(\Omega).$ Soit $(\xi_h^*, \mathbf{w}_h, \theta_h, \mathbf{p}_h) \in \mathcal{U}_h \times X_h^u \times X_h^\theta \times M_h$ une solution optimale de (2.26). Alors tout point d'accumulation de la suite $(\xi_h, \mathbf{w}_h, \theta_h, \mathbf{p}_h)_h$ est une solution optimale de (2.17).*

Des conditions nécessaires d'optimalité du premier ordre pour (2.17) peuvent être obtenues en appliquant le [26, Theorem 1.48]. Ce résultat requiert la différentiabilité de la fonctionnelle réduite $\mathcal{J}_1 : \xi \in \mathcal{U}_{\text{ad}} \mapsto \mathcal{J}(\xi, \mathbf{w}(\xi), \theta(\xi), \mathbf{p}(\xi)) \in \mathbb{R}$ ainsi que l'inversibilité, pour tout $\xi \in \mathcal{U}_{\text{ad}}$, de l'application $\partial_{(\mathbf{w}, \theta, \mathbf{p})} e(\xi; \mathbf{w}(\xi), \theta(\xi), \mathbf{p}(\xi))$ où e est l'opérateur à valeur dans le dual de $X^u \times X^\theta \times L^2(\Omega)$ pour lequel le problème variationnel (2.16) est équivalent à

$$\langle e(\xi; \mathbf{w}, \theta, \mathbf{p}), (\mathbf{v}_1, \mathbf{v}_2, \mathbf{q}) \rangle_{(X^u)' \times (X^\theta)' \times L^2(\Omega), X^u \times X^\theta \times L^2(\Omega)} = 0, \forall (\mathbf{v}_1, \mathbf{v}_2, \mathbf{q}) \in X^u \times X^\theta \times L^2(\Omega).$$

L'étude de l'inversibilité de l'application $\partial_{(\mathbf{w}, \theta, \mathbf{p})} e(\xi; \mathbf{w}(\xi), \theta(\xi), \mathbf{p}(\xi))$ revient à étudier l'existence et l'unicité de la solution au système linéarisé de (2.16). Pour des termes sources $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3) \in (X^u)' \times (X^\theta)' \times L^2(\Omega)$, cela est équivalent à montrer qu'il existe un unique $(\mathbf{v}, \zeta, \rho)$ satisfaisant le problème suivant

$$\begin{aligned} & \text{Trouver } (\mathbf{v}, \zeta, \rho) \in X^u \times X^\theta \times L^2(\Omega) \text{ tels que} \\ & \left\{ \begin{array}{ll} \alpha(\xi; \mathbf{v}, \mathbf{v}_1) + \mathbf{b}(\mathbf{v}_1, \rho) + \mathbf{c}'(\mathbf{w} + \mathbf{V}, \mathbf{v}, \mathbf{v}_1) + \mathbf{f}(\zeta, \mathbf{v}_1) = \langle \mathbf{f}_1, \mathbf{v}_1 \rangle_{(X^u)', X^u}, & \forall \mathbf{v}_1 \in X^u \\ \tilde{\alpha}(\xi; \zeta, \mathbf{v}_2) + \tilde{\mathbf{c}}(\zeta, \mathbf{w} + \mathbf{V}, \mathbf{v}_2) + \tilde{\mathbf{c}}(\theta, \mathbf{v}, \mathbf{v}_2) = \langle \mathbf{f}_2, \mathbf{v}_2 \rangle_{(X^\theta)', X^\theta} & \forall \mathbf{v}_2 \in X^\theta \\ \mathbf{b}(\mathbf{v}, \mathbf{q}) = (\mathbf{f}_3, \mathbf{q})_{L^2(\Omega)}, & \forall \mathbf{q} \in L^2(\Omega), \end{array} \right. \end{aligned} \quad (2.27)$$

avec

$$\mathbf{c}'(\mathbf{w}, \mathbf{v}, \mathbf{v}) = \int_{\Omega} ((\mathbf{w} \cdot \nabla) \mathbf{v} \cdot \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{w} \cdot \mathbf{v}) \, dx.$$

Nous avons obtenu le résultat suivant :

Théorème 2.13 *On suppose que \mathbf{u}_{in} et $(\mathbf{w}, \theta, \mathbf{p})$ sont assez petits. Alors (2.27) admet une unique solution qui est de plus continue par rapport aux données $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$.*

En combinant le Théorème 2.13 avec [26, Theorem 1.48], un optimum ξ^* vérifie les conditions nécessaires d'optimalité ci-dessous :

$$\langle \nabla_{\xi} \mathcal{J}_1(\xi^*), \xi - \xi^* \rangle_{(L^\infty(\Omega))', L^\infty(\Omega)} \geq 0, \forall \xi \in \mathcal{U}_{\text{ad}},$$

où le gradient de la fonctionnelle réduite \mathcal{J}_1 est défini à partir du modèle adjoint. Pour calculer ce dernier, on suppose que la fonction coût est donnée par

$$\mathcal{J}(\xi, \mathbf{u}, \mathbf{p}, \theta) = \int_{\Omega} J_{\Omega}(\xi, \mathbf{u}, \theta, \mathbf{p}) + \int_{\Gamma} J_{\Gamma}(\xi, \mathbf{u}, \theta, \mathbf{p}),$$

où $J_{\Omega} : \mathcal{U}_{\text{ad}} \times X_1^u \times X^\theta \times L^2(\Omega) \rightarrow L^1(\Omega)$, $J_{\Gamma} : \mathcal{U}_{\text{ad}} \times X_1^u \times X^\theta \times L^2(\Omega) \rightarrow L^1(\Gamma)$ sont Fréchet-différentiables. Le système adjoint associé à (2.17) est alors

$$\left\{ \begin{array}{l} \nabla \lambda^{\mathbf{p}} - \mathbf{h}_{\tau}(\alpha) \lambda^{\mathbf{u}} + \theta \nabla \lambda^{\theta} + \mathbf{A} \Delta \lambda^{\mathbf{u}} + \nabla \lambda^{\mathbf{u}} \cdot \mathbf{u} - (\lambda^{\mathbf{u}} \cdot \nabla) \mathbf{u} = -\frac{\partial J_{\Omega}}{\partial \mathbf{u}}, \\ \mathbf{B} \lambda^{\mathbf{u}} \cdot \mathbf{e}_y + \mathbf{u} \cdot \nabla \lambda^{\theta} + \nabla \cdot (\mathbf{C} \mathbf{k}_{\tau}(\xi) \nabla \lambda^{\theta}) = -\frac{\partial J_{\Omega}}{\partial \theta}, \\ \nabla \cdot \lambda^{\mathbf{u}} = -\frac{\partial J_{\Omega}}{\partial \mathbf{p}}, \end{array} \right. \quad (2.28)$$

avec les conditions limites

$$\begin{aligned}
\text{sur } \Gamma_w : \quad & \lambda^u \cdot \mathbf{t} = 0, \lambda^u \cdot \mathbf{n} = \frac{\partial J_\Gamma}{\partial p}, \lambda^\theta (\mathbf{u} \cdot \mathbf{n}) + Ck_\tau(\xi) \partial_n \lambda^\theta = \frac{\partial J_\Gamma}{\partial \theta}, \\
\text{sur } \Gamma_{in} : \quad & \lambda^u \cdot \mathbf{t} = 0, \lambda^u \cdot \mathbf{n} = \frac{\partial J_\Gamma}{\partial p}, \lambda^\theta = 0, \partial_n \lambda^p = 0, \\
\text{sur } \Gamma_{out} : \quad & \lambda^\theta (\mathbf{u} \cdot \mathbf{n}) + Ck_\tau(\xi) \partial_n \lambda^\theta = \frac{\partial J_\Gamma}{\partial \theta}, \\
& \lambda^p \mathbf{n} + \lambda^\theta \theta \mathbf{n} + A \partial_n \lambda^u + (\mathbf{u} \cdot \mathbf{n}) \lambda^u = \frac{\partial J_\Gamma}{\partial \mathbf{u}},
\end{aligned} \tag{2.29}$$

où \mathbf{t} est un vecteur unitaire tangent. Le gradient de la fonction coût est alors

$$\begin{aligned}
\nabla_\xi \mathcal{J} &= \frac{\partial J_\Omega}{\partial \xi} - \frac{\partial h_\tau}{\partial \alpha}(\xi) \mathbf{u} \cdot \lambda^u - C \frac{\partial k_\tau}{\partial \xi}(\xi) \nabla \theta \cdot \nabla \lambda^\theta & \text{sur } \Omega, \\
\nabla_\xi \mathcal{J} &= \frac{\partial J_\Gamma}{\partial \xi} & \text{sur } \Gamma_w,
\end{aligned} \tag{2.30}$$

avec $\mathbf{u} = \mathbf{w} + \mathbf{V}$ où (\mathbf{w}, θ, p) est la solution faible du problème direct (2.16).

2.4 Expériences numériques

On illustre maintenant certains résultats théoriques et les effets de la fonction d'interpolation permettant d'obtenir des solides optimisés avec des paramètres physiques constants par morceaux donnés dans la Section 2.1.

Le code a été réalisé avec Python et FEniCS/DOLFIN [27]. On utilise l'approche *optimize-then-discretize* et la résolution du problème d'optimisation (2.17) est effectuée avec l'algorithme L-BFGS-B [11]. La résolution du problème direct (2.16) a été réalisée avec une méthode de Newton dont l'initialisation est calculée avec un point fixe. Le domaine $\Omega = [0, L] \times [0, H]$ est représenté dans la Figure 2.4 et la fonction coût vise à

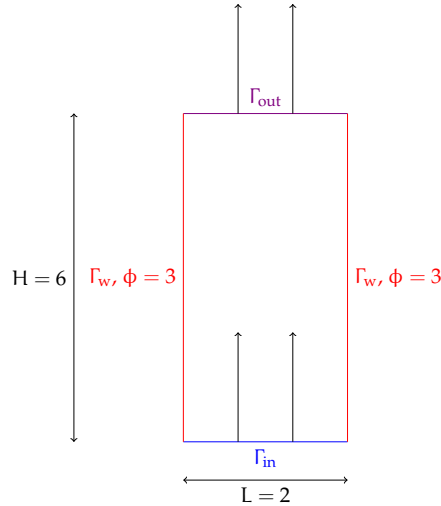


FIGURE 2.4 – Présentation du domaine de calcul Ω et des conditions limites

maximiser la température en sortie du canal. La vitesse à l'entrée du canal est le profil parabolique

$$\forall x \in \Gamma_{\text{in}} : \mathbf{u}_{\text{in}}(x) = 1.8x(2-x)\mathbf{e}_y.$$

Le problème d'optimisation résolu numériquement est ainsi

$$\begin{aligned} & \min - \int_{\Gamma_{\text{out}}} \theta \\ \text{tels que } & \begin{cases} \mathbf{u} = \mathbf{w} + \mathbf{V} \text{ où } (\mathbf{w}, \theta, p) \text{ satisfait (2.16),} \\ (\alpha, \varphi) \in \mathcal{U}_{\text{ad}}. \end{cases} \end{aligned}$$

Les résultats numériques ont été obtenus avec les paramètres ci-dessous :

$$Re = 100, Ri = 1.8, \tau = 30, \alpha_{\text{max}} = 10^8, \alpha_0 = 1, k_j \in \{5, 6, 8, 11, 13, 16\}.$$

La taille caractéristique du maillage est paramétrée dans FEniCS par n qui indique le nombre de cellule dans une direction. Ainsi, plus n est grand, plus le maillage est fin. La Figure 2.5 montre les solides optimisés pour différentes valeurs de n . On peut observer sur la Figure 2.5 que les formes des solides optimisés sont assez proches pour les différents pas de maillages. Toutefois, les champs de conductivités optimisés présentent davantage de différences. Néanmoins, ces résultats ne sont pas incompatibles avec le Théorème 2.12 qui assure la convergence d'une sous-suite de minimiseurs discrets vers un minimum de la fonction coût. En effet, pour chaque maillage considéré, les valeurs de la fonction coût après optimisation sont très proches de -1.5 .

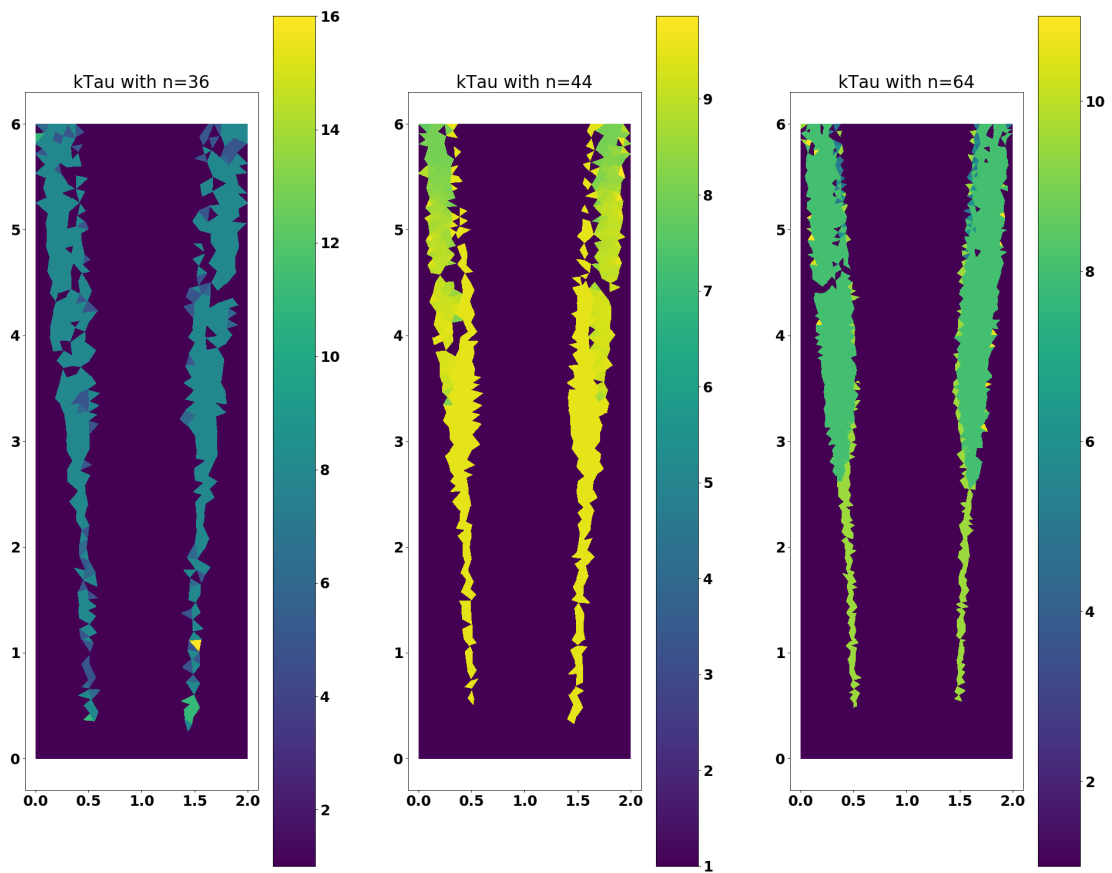


FIGURE 2.5 – Evolution du k_τ optimisé pour différentes valeurs de n . Les zones de fluide correspondent aux zones bleues dans lesquelles $k_\tau = 1$.

3 Optimisation topologique avec conditions limites non-lisses en sortie

Cette section présente des résultats issus de [Chap2-A6].

L'objectif de [Chap2-A6] est d'étudier un problème d'optimisation topologique où les contraintes sont données par le système de Navier-Stokes incompressible, instationnaire et sous l'approximation de Boussinesq, avec un terme de pénalisation pour la vitesse permettant de localiser un solide, en prenant en compte des conditions non-lisses en sortie. Ces conditions limites sont appelées *Directional-do-nothing* (DDN) et permettent de modéliser le cas d'un écoulement retour en sortie d'un domaine de calcul [10, 4, 12, 8]. Plus précisément, les conditions limites DDN se ramènent à des conditions de sortie usuelles ($\text{Re}^{-1} \partial_n \mathbf{u} - n\mathbf{p} = 0$, appelée aussi conditions limites *Do-Nothing*) lorsque l'écoulement est sortant et font intervenir le carré de la vitesse pour un écoulement rentrant dans le domaine. Les conditions DDN se rapprochent de plus des conditions locales et globales de type Bernoulli étudiées dans [9, 19] mais ont l'avantage de s'inclure plus naturellement dans une formulation variationnelle du système de Navier-Stokes incompressible.

Le cas stationnaire du modèle de Navier-Stokes avec des conditions limites DDN en sortie a été étudié dans [5] et le système de Boussinesq dans [12]. L'existence et l'unicité d'une solution faible pour le système de Navier-Stokes instationnaire avec des conditions DDN en sortie ont été obtenus dans [10, 8]. Le système de Boussinesq instationnaire a également été traité dans [5, p. 16, Theorem 3.2] mais la stratégie de démonstration est basée sur un point fixe ce qui permet de démontrer l'existence et l'unicité pour des termes sources et des paramètres physiques assez petits. Dans [Chap2-A6], nous avons obtenu l'existence (et l'unicité en 2D) de solution pour le modèle de Boussinesq instationnaire avec conditions DDN en sortie en s'affranchissant de l'hypothèse des petits termes sources et petits paramètres physiques. Nous avons ensuite étudié un problème d'optimisation topologique utilisant ce modèle comme contrainte.

Le modèle sans dimension considéré dans [Chap2-A6] est donné par

$$\begin{aligned}
 \nabla \cdot \mathbf{u} &= 0, & \text{dans } \Omega \\
 \partial_t \theta + \nabla \cdot (\mathbf{u}\theta) - \nabla \cdot (Ck(\alpha)\nabla\theta) &= 0, & \text{dans } \Omega \\
 \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} - A\Delta\mathbf{u} + \nabla p - B\theta\mathbf{e}_y + h(\alpha)\mathbf{u} &= \mathbf{f}, & \text{dans } \Omega \\
 \mathbf{u}(0, \mathbf{x}) &= \mathbf{u}_0(\alpha; \mathbf{x}) & \mathbf{x} \in \Omega, \\
 \theta(0, \mathbf{x}) &= \theta_0(\alpha; \mathbf{x}) & \mathbf{x} \in \Omega,
 \end{aligned} \tag{2.31}$$

où \mathbf{u} est la vitesse du fluide, p sa pression, θ la température et $\mathbf{u}_0(\alpha), \theta_0(\alpha)$ sont des conditions initiales. On a de plus $A = \text{Re}^{-1}$ avec Re le nombre de Reynolds, $B = \text{Ri}$ est le nombre de Richardson et $C = (\text{Re Pr})^{-1}$ où Pr est le nombre de Prandtl, $\mathbf{e}_y = (0, 1)$ en 2d et $\mathbf{e}_y = (0, 1, 0)$ en 3d. La fonction $\alpha : \mathbf{x} \in \Omega \mapsto \alpha(\mathbf{x}) \in \mathbb{R}^+$ est le paramètre d'optimisation permettant de discriminer les zones fluides et solides dans le domaine Ω et $h(\alpha), k(\alpha)$ sont définies à partir des fonctions d'interpolation fluide/solide telles que celles données

dans la Section 2.1.

Pour les conditions aux limites, on considère tout d'abord une fonction $\beta : (t, x) \in [0, T] \times \Gamma_{\text{out}} \mapsto \beta(t, x) \in \mathbb{R}$ et on note

$$\forall s \in \mathbb{R} : s^+ = \text{pos}(s) = \max(0, s), s^- = \text{neg}(s) = \max(0, -s), s = s^+ - s^-.$$

En supposant que $\partial\Omega = \Gamma$ est Lipschitzienne et peut être décomposée en 3 parties disjointes $\Gamma = \Gamma_w \cup \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$, les conditions limites sont données par

$$\begin{aligned} \underline{\text{sur}} \Gamma_{\text{in}} : \quad & \mathbf{u} = \mathbf{u}_{\text{in}}, \theta = 0, \\ \underline{\text{sur}} \Gamma_w : \quad & \mathbf{u} = 0, \text{Ck}\partial_n\theta = \phi, \\ \underline{\text{sur}} \Gamma_{\text{out}} : \quad & A\partial_n\mathbf{u} - \mathbf{np} = A\partial_n\mathbf{u}^{\text{ref}} - \mathbf{np}^{\text{ref}} - \frac{1}{2}(\mathbf{u} \cdot \mathbf{n})^-(\mathbf{u} - \mathbf{u}^{\text{ref}}), \\ & \text{Ck}\partial_n\theta + \beta(\mathbf{u} \cdot \mathbf{n})^-\theta = 0, \end{aligned} \tag{2.32}$$

où $\phi \in L^2(0, T; L^2(\Gamma_w))$, $\mathbf{f} \in L^2(0, T; (H^1(\Omega))')$, $\mathbf{u}_{\text{in}} \in L^2(0, T; H_{00}^{1/2}(\Gamma_{\text{in}}))$, $\partial_n = \mathbf{n} \cdot \nabla$ et $(\mathbf{u}^{\text{ref}}, \mathbf{p}^{\text{ref}})$ est une donnée. Les conditions de type *Directional-Do-Nothing* sont celles imposées sur la sortie Γ_{out} et ont été en partie introduites dans [10, 4, 12, 8].

Dans [Chap2-A6], nous avons étudié le problème d'optimisation topologique ci-dessous

$$\begin{aligned} & \min \mathcal{J}(\alpha, \mathbf{u}, \theta) \\ \text{tel que } & \begin{cases} (\mathbf{u}, \theta) \text{ vérifie (2.31) – (2.32) paramétré par } \alpha, \\ \alpha \in \mathcal{U}_{\text{ad}}, \end{cases} \end{aligned} \tag{2.33}$$

où \mathcal{J} est une fonction coût. L'ensemble \mathcal{U}_{ad} peut être défini comme aux Sections 1 et 2 mais pourrait également être remplacé par tout ensemble dont l'inclusion dans $L^p(\Omega)$ pour un certain $1 \leq p < +\infty$ est compacte.

Une des difficultés concernant l'étude du problème d'optimisation (2.33) provient du caractère non-lisse (ici : non-dérivable) des conditions aux limites. En particulier, cela rend délicat la dérivation de conditions nécessaires d'optimalité du premier ordre. Pour contourner cette difficulté, nous avons régularisé les conditions limites (voir par exemple [32]) en remplaçant la fonction neg par une suite neg_ε de classe C^1 définie par exemple dans [32] par

$$\text{neg}_\varepsilon(s) = \begin{cases} \text{neg}(s) & \text{si } |s| \geq \frac{\varepsilon}{2}, \\ \frac{1}{2} \left(\frac{1}{2} - \frac{s}{\varepsilon} \right)^3 \left(\frac{3\varepsilon}{2} + s \right) & \text{si } |s| < \frac{\varepsilon}{2}. \end{cases} \tag{2.34}$$

On introduit également la notation $\text{pos}_\varepsilon(s) = s + \text{neg}_\varepsilon(s)$ pour la régularisation de la partie positive de $s \in \mathbb{R}$. Parmi les propriétés de la fonction $\text{neg}_\varepsilon(\cdot)$, on peut noter que la convergence de $\text{neg}_\varepsilon(\cdot)$ vers neg est uniforme sur \mathbb{R} lorsque $\varepsilon \rightarrow 0$.

Pour définir la formulation faible de (2.31)-(2.32), on introduit

$$V^u = \{\mathbf{u} \in H^1(\Omega)^d \mid \nabla \cdot \mathbf{u} = 0, \mathbf{u}|_{\Gamma_{\text{in}} \cup \Gamma_w} = 0\},$$

et on note H^u la fermeture de V^u dans $(L^2(\Omega))^d$. De même, on note

$$V^\theta = \{\theta \in H^1(\Omega) \mid \theta|_{\Gamma_{\text{in}}} = 0\},$$

et $H^0 = L^2(\Omega)$. La formulation faible de (2.31)-(2.32) où la fonction neg est remplacée neg_ε est alors donnée par

$$\begin{aligned} & \int_{\Omega} \partial_t \theta_\varepsilon \varphi - \int_{\Omega} \theta_\varepsilon \mathbf{u}_\varepsilon \cdot \nabla \varphi + \int_{\Gamma_{\text{out}}} ((\mathbf{u}_\varepsilon \cdot \mathbf{n}) + \beta \text{neg}_\varepsilon (\mathbf{u}_\varepsilon \cdot \mathbf{n})) \theta_\varepsilon \varphi \\ & + \int_{\Omega} Ck \nabla \theta_\varepsilon \cdot \nabla \varphi = \int_{\Gamma_w} \phi \varphi, \end{aligned} \quad (2.35)$$

$$\begin{aligned} & \int_{\Omega} \partial_t \mathbf{u}_\varepsilon \cdot \Psi + \frac{1}{2} \{ ((\mathbf{u}_\varepsilon \cdot \nabla) \mathbf{u}_\varepsilon) \cdot \Psi - ((\mathbf{u}_\varepsilon \cdot \nabla) \Psi) \cdot \mathbf{u}_\varepsilon \} + A \nabla \mathbf{u}_\varepsilon : \nabla \Psi \\ & + \int_{\Omega} h(\alpha) \mathbf{u}_\varepsilon \cdot \Psi - B \theta_\varepsilon \cdot \mathbf{e}_y \cdot \Psi + \frac{1}{2} \int_{\Gamma_{\text{out}}} \text{pos}_\varepsilon (\mathbf{u}_\varepsilon \cdot \mathbf{n}) (\mathbf{u}_\varepsilon \cdot \Psi) \\ & = \int_{\Omega} \mathbf{f} \cdot \Psi + \int_{\Gamma_{\text{out}}} (A \partial_n \mathbf{u}^{\text{ref}} - \mathbf{n} p^{\text{ref}}) \cdot \Psi + \frac{1}{2} \int_{\Gamma_{\text{out}}} \text{neg}_\varepsilon (\mathbf{u}_\varepsilon \cdot \mathbf{n}) (\mathbf{u}^{\text{ref}} \cdot \Psi), \end{aligned} \quad (2.36)$$

pour tout $(\Psi, \varphi) \in V^u \times V^\theta$. On appelle notamment solution faible de (2.31)-(2.32), avec les conditions limites régularisées, tout $(\mathbf{u}_\varepsilon, \theta_\varepsilon) \in L^2(0, T; V^u) \times L^2(0, T; V^\theta)$ tel que $(\partial_t \mathbf{u}_\varepsilon, \partial_t \theta_\varepsilon) \in (V^u)' \times (V^\theta)'$ vérifiant (2.35)-(2.36) au sens des distributions dans $\mathcal{D}'(]0, T[)$.

Le problème d'optimisation avec conditions limites régularisées est enfin

$$\begin{aligned} & \min \mathcal{J}(\alpha_\varepsilon, \mathbf{u}_\varepsilon, \theta_\varepsilon) \\ & \text{où } \begin{cases} (\mathbf{u}_\varepsilon, \theta_\varepsilon) \text{ vérifient (2.35) - (2.36) paramétré par } \alpha_\varepsilon, \\ \alpha_\varepsilon \in \mathcal{U}_{\text{ad}}. \end{cases} \end{aligned} \quad (2.37)$$

Dans la suite, nous donnerons un théorème d'existence de solution faible pour (2.35)-(2.36) ainsi qu'un résultat de convergence lorsque $\varepsilon \rightarrow 0$. On dérivera ensuite des conditions nécessaires d'optimalité du premier ordre pour le problème d'optimisation (2.37) qui permettront, après un passage à la limite $\varepsilon \rightarrow 0$, de dériver des conditions d'optimalité pour le problème d'optimisation (2.33) avec des conditions limites non-lisses en sortie.

3.1 Existence et unicité d'une solution faible

L'existence de solutions faibles pour (2.35)-(2.36) est obtenue sous les hypothèses ci-dessous :

Hypothèses 1

- Le terme source vérifie $\mathbf{f} \in L^2(0, T; (H^1(\Omega))')$.
- $(\mathbf{u}^{\text{ref}}, p^{\text{ref}})$ sont tels que

$$\begin{cases} \mathbf{u}^{\text{ref}} \in L^r(0, T; (H^1(\Omega))^d) \cap L^\infty(0, T; (L^2(\Omega))^d) \\ \text{où } r = 2 \text{ si } d = 2 \text{ et } r = 4 \text{ si } d = 3, \\ \nabla \cdot \mathbf{u}^{\text{ref}} = 0, \\ \partial_t \mathbf{u}^{\text{ref}} \in L^2(0, T; (L^2(\Omega))^d), \\ \mathbf{u}^{\text{ref}} = 0 \text{ sur } \Gamma_w \\ \mathbf{u}^{\text{ref}} = \mathbf{u}_{\text{in}} \text{ sur } \Gamma_{\text{in}}. \end{cases}$$

et $A \partial_n \mathbf{u}^{ref} - \mathbf{p}^{ref} \mathbf{n} \in L^2(0, T; H^{-\frac{1}{2}}(\partial\Omega))$.

- Il existe k_{min} tel que $k(x) \geq k_{min} > 0$ et $h(x) \geq 0$ pour presque tout $x \in \Omega$.
- La condition initiale $\mathbf{u}_0 : \mathcal{U}_{ad} \mapsto \mathbf{V}^u$ (resp. $\theta_0 : \mathcal{U}_{ad} \mapsto \mathbf{V}^\theta$) est Fréchet-différentiable. De plus, pour tout $\alpha \in \mathcal{U}_{ad}$, $\mathbf{u}_0(\alpha)|_{\Gamma_{in}} = \mathbf{u}_{in}(0)$, $\mathbf{u}_0(\alpha)|_{\Gamma_w} = 0$ et $\theta_0(\alpha)|_{\Gamma_{in}} = 0$.
- $\beta \in L^\infty(0, T; L^\infty(\Gamma_{out}))$ vérifie $\beta(t, x) \geq \frac{1}{2}$ pour presque tout $(t, x) \in [0, T] \times \Gamma_{out}$.

En notant $\mathbf{u}_\varepsilon = \mathbf{v}_\varepsilon + \mathbf{u}^{ref}$ et en prenant $\varphi = \theta_\varepsilon$, $\Psi = \mathbf{v}_\varepsilon$ dans (2.35)-(2.36), on obtient les estimations a priori suivantes

$$\sup_{[0, T]} \|\theta_\varepsilon\|_{L^2(\Omega)} + \int_0^T \|\nabla \theta_\varepsilon\|_{L^2(\Omega)}^2 \leq c^\theta, \quad (2.38)$$

$$\sup_{[0, T]} \|\mathbf{v}_\varepsilon\|_{L^2(\Omega)} + \int_0^T \|\nabla \mathbf{v}_\varepsilon\|_{L^2(\Omega)}^2 \leq c^v, \quad (2.39)$$

où les constantes positives c^θ, c^v sont indépendantes de ε et de α . L'existence de solutions peut alors être obtenue avec la méthode de Faedo-Galerkin (voir par exemple [8, 10]) combinée à des arguments de compacité.

Théorème 2.14 *Supposons que les hypothèses 1 soient vérifiées. Alors pour tout $(\mathbf{v}_0, \theta_0) \in H^u \times H^\theta$ et tout $T > 0$, il existe $\mathbf{v}_\varepsilon \in L^\infty(0, T; H^u) \cap L^2(0, T; \mathbf{V}^u)$, $\theta_\varepsilon \in L^\infty(0, T; H^\theta) \cap L^2(0, T; \mathbf{V}^\theta)$ vérifiant (2.35)-(2.36). De plus, en notant $\mathbf{u}_0 = \mathbf{v}_0 + \mathbf{u}^{ref}(0)$ et $\mathbf{u}_\varepsilon = \mathbf{v}_\varepsilon + \mathbf{u}^{ref}$, on a*

$$\left(\int_\Omega \mathbf{u}_\varepsilon \cdot \Psi \right) (0) dx = \int_\Omega \mathbf{u}_0 \cdot \Psi dx, \quad \left(\int_\Omega \theta_\varepsilon \varphi \right) (0) dx = \int_\Omega \theta_0 \varphi dx,$$

pour tout $(\Psi, \varphi) \in \mathbf{V}^u \times \mathbf{V}^\theta$. Enfin, $\mathbf{v}'_\varepsilon \in L^{\frac{4}{3}}(0, T; (\mathbf{V}^u)')$ et $\theta'_\varepsilon \in L^{\frac{4}{3}}(0, T; (\mathbf{V}^\theta)')$.

Le Théorème 2.14 assure l'existence d'au moins une solution faible pour (2.35)-(2.36). L'unicité peut également être démontrée en dimension $d = 2$ (voir par exemple [8, 10]).

3.2 Convergence de la solution du problème régularisé

En utilisant que les estimations a priori (2.38) sont indépendantes de ε , on peut obtenir un résultat de convergence pour \mathbf{u}_ε et θ_ε lorsque $\varepsilon \rightarrow 0$.

Théorème 2.15 *Considérons $(\alpha_\varepsilon)_\varepsilon \subset \mathcal{U}_{ad}$ et $\alpha \in \mathcal{U}_{ad}$ tels que α_ε converge faible* vers α dans $BV(\Omega)$. Soit $(\mathbf{v}_\varepsilon, \theta_\varepsilon)$ une (suite de) solutions de (2.35)-(2.36) paramétrée par α_ε et soit $\mathbf{u}_\varepsilon = \mathbf{v}_\varepsilon + \mathbf{u}^{ref}$. Alors il existe $(\mathbf{v}, \theta) \in L^\infty(0, T; H^u) \cap L^2(0, T; \mathbf{V}^u) \times L^\infty(0, T; H^\theta) \cap L^2(0, T; \mathbf{V}^\theta)$ tel que, en notant $\mathbf{u} = \mathbf{v} + \mathbf{u}^{ref}$, on a une sous-suite de $(\mathbf{v}_\varepsilon, \theta_\varepsilon)$ (la même notation est utilisée) telle que les convergences ci-dessous sont vérifiées :*

- $\mathbf{u}_\varepsilon \xrightarrow{*} \mathbf{u}$ dans $L^\infty(0, T; H^u)$
- $\theta_\varepsilon \xrightarrow{*} \theta$ dans $L^\infty(0, T; H^\theta)$,

- $\mathbf{u}_\varepsilon \rightharpoonup \mathbf{u}$ dans $L^2(0, T; \mathbf{V}^u)$ et dans $L^2(0, T; (L^6(\Omega))^d)$,
- $\theta_\varepsilon \rightharpoonup \theta$ dans $L^2(0, T; \mathbf{V}^\theta)$ et dans $L^2(0, T; L^6(\Omega))$,
- $\mathbf{u}_\varepsilon \rightharpoonup \mathbf{u}$ dans $L^4(0, T; (L^2(\Gamma))^d)$
- $\theta_\varepsilon \rightharpoonup \theta$ in $L^4(0, T; L^2(\Gamma))$,
- $\mathbf{u}_\varepsilon \rightarrow \mathbf{u}$ dans $L^2(0, T; (L^2(\Omega))^d)$
- $\theta_\varepsilon \rightarrow \theta$ dans $L^2(0, T; L^2(\Omega))$,
- $\mathbf{u}_\varepsilon \rightarrow \mathbf{u}$ dans $L^2(0, T; (L^2(\Gamma))^d)$
- $\theta_\varepsilon \rightarrow \theta$ dans $L^2(0, T; L^2(\Gamma))$,
- $\partial_t \mathbf{u}_\varepsilon \rightharpoonup \partial_t \mathbf{u}$ dans $L^{\frac{4}{3}}(0, T; (\mathbf{V}^u)')$
- $\partial_t \theta_\varepsilon \rightharpoonup \partial_t \theta$ dans $L^{\frac{4}{3}}(0, T; (\mathbf{V}^\theta)')$.

De plus (\mathbf{v}, θ) est une solution faible de (2.31)-(2.32) paramétré par α .

Le Théorème 2.15 assure l'existence d'une sous-suite de solutions faibles du problème régularisé vers une solution du problème initial (2.31)-(2.32). Des résultats plus précis ont été obtenus en dimension $d = 2$ car la dérivée en temps de $\mathbf{u}_\varepsilon, \theta_\varepsilon$ est alors suffisamment régulière pour utiliser des arguments basés sur le Lemme de Gronwall ce qui permet notamment de démontrer l'unicité de la solution faible.

Théorème 2.16 *On suppose que $d = 2$ et que les hypothèses du Théorème 2.15 sont satisfaites. Alors, pour $\varepsilon \rightarrow 0$, on a $\mathbf{u}_\varepsilon \rightarrow \mathbf{u}$ fortement dans $L^\infty(0, T; L^2(\Omega)^2)$ et $\theta_\varepsilon \rightarrow \theta$ fortement dans $L^\infty(0, T; L^2(\Omega))$. De plus, $\nabla \mathbf{u}_\varepsilon \rightarrow \nabla \mathbf{u}$ fortement dans $L^2(0, T; L^2(\Omega)^2)$ et $\nabla \theta_\varepsilon \rightarrow \nabla \theta$ fortement dans $L^2(0, T; L^2(\Omega))$.*

Les Théorèmes 2.15 et 2.16 donnent la convergence, pour $\varepsilon \rightarrow 0$, des solutions faibles de (2.35)-(2.36) vers celles de (2.31)-(2.32). En dimension 3, nous avons la convergence d'une sous-suite seulement tandis qu'en dimension $d = 2$, l'unicité de la solution faible permet de montrer la convergence de l'intégralité de la suite.

3.3 Conditions nécessaires d'optimalité du premier ordre

Pour simplifier la présentation des résultats, on se restreint dans cette partie au cas de la dimension $d = 2$.

Considérons les ensembles suivants :

$$\mathbb{X}^u = L^2(0, T; \mathbf{V}^u) \cap L^\infty(0, T; \mathbf{H}^u) \text{ et } \mathbb{X}^\theta = L^2(0, T; \mathbf{V}^\theta) \cap L^\infty(0, T; \mathbf{H}^\theta),$$

qui sont munis des normes

$$\|\mathbf{u}\|_{\mathbb{X}^u} = \max\{\|\mathbf{u}\|_{L^2(0, T; \mathbf{V}^u)}, \|\mathbf{u}\|_{L^\infty(0, T; \mathbf{H}^u)}\}, \|\theta\|_{\mathbb{X}^\theta} = \max\{\|\theta\|_{L^2(0, T; \mathbf{V}^\theta)}, \|\theta\|_{L^\infty(0, T; L^2(\Omega))}\}.$$

En adaptant la démonstration du Théorème 2.16, on peut démontrer la continuité de l'application

$$\alpha \in (\mathcal{U}_{\text{ad}}, \text{faible*}) \mapsto (\mathbf{u}_\varepsilon(\alpha), \theta(\alpha)) \in (\mathbb{X}^u \times \mathbb{X}^\theta, \text{fort}),$$

où $(\mathbf{u}_\varepsilon(\alpha), \theta(\alpha))$ est la solution faible de (2.31)-(2.32) paramétré par α . En supposant alors que la fonction coût \mathcal{J} est semi-continue inférieurement pour la topologie (faible*, forte, forte) de $\mathcal{U}_{ad} \times L^2(0, T; V^u) \times L^2(0, T; V^\theta)$, on peut alors utiliser des suites minimisantes pour démontrer l'existence d'au moins une solution aux problèmes d'optimisation (2.37) et (2.33).

Pour définir les conditions nécessaires d'optimalité du premier ordre du problème d'optimisation régularisé, on ré-écrit (2.35)-(2.36) sous la forme équivalente

$$\langle e(\alpha_\varepsilon; \mathbf{u}_\varepsilon, \theta_\varepsilon), \mathbf{y} \rangle_{\mathcal{Y}, \mathcal{Y}} = 0, \quad \forall \mathbf{y} \in \mathcal{Y}$$

où $e : \mathcal{U}_{ad} \times W^u(0, T) \times W^\theta(0, T) \rightarrow \mathcal{Y}'$ avec $\mathcal{Y} := \mathbb{X}^u \times \mathbb{X}^\theta \times H^u \times H^\theta$ et

$$W^u(0, T) = \{\mathbf{u} \in \mathbb{X}^u; \partial_t \mathbf{u} \in (\mathbb{X}^u)'\},$$

$$W^\theta(0, T) = \{\theta \in \mathbb{X}^\theta; \partial_t \theta \in (\mathbb{X}^\theta)'\}.$$

Les deux premières composantes de $e(\alpha_\varepsilon; \mathbf{u}_\varepsilon, \theta_\varepsilon)$ donnent les équations de (2.35)-(2.36) et les deux dernières permettent d'imposer les conditions initiales.

Les conditions nécessaires d'optimalité du premier ordre pour le problème d'optimisation régularisé (2.37) sont obtenues en appliquant [26, Theorem 1.48] qui nécessite d'étudier l'inversibilité de l'opérateur $\partial_{(\mathbf{u}_\varepsilon, \theta_\varepsilon)} e(\alpha_\varepsilon; \cdot, \cdot)$. Ces dernières sont données dans le résultat suivant.

Théorème 2.17 *Soit α_ε^* une solution optimale de (2.37) et $(\mathbf{u}_\varepsilon^*, \theta_\varepsilon^*)$ la solution de (2.35)-(2.36) avec $\alpha = \alpha_\varepsilon^*$. Alors il existe des variables adjointes $(\lambda_\varepsilon^u, \lambda_\varepsilon^\theta) \in \mathbb{X}^u \times \mathbb{X}^\theta$ telles que, en notant $(\lambda_\varepsilon^{u_0}, \lambda_\varepsilon^{\theta_0}) = (\lambda_\varepsilon^u(0, \cdot), \lambda_\varepsilon^\theta(0, \cdot))$ et $\Lambda_\varepsilon = (\lambda_\varepsilon^u, \lambda_\varepsilon^\theta, \lambda_\varepsilon^{u_0}, \lambda_\varepsilon^{\theta_0})$:*

$$\begin{aligned} e(\alpha_\varepsilon^*, \mathbf{u}_\varepsilon^*, \theta_\varepsilon^*) &= 0, \\ (\partial_{(\mathbf{u}_\varepsilon, \theta_\varepsilon)} e(\alpha_\varepsilon^*; \mathbf{u}_\varepsilon^*, \theta_\varepsilon^*))^* \Lambda_\varepsilon &= -\partial_{(\mathbf{u}_\varepsilon, \theta_\varepsilon)} \mathcal{J}(\alpha_\varepsilon^*; \mathbf{u}_\varepsilon^*, \theta_\varepsilon^*), \\ \langle \partial_{\alpha_\varepsilon} \mathcal{J}(\alpha_\varepsilon^*; \mathbf{u}_\varepsilon^*, \theta_\varepsilon^*) + (\partial_{\alpha_\varepsilon} e(\alpha_\varepsilon^*; \mathbf{u}_\varepsilon^*, \theta_\varepsilon^*))^* \Lambda_\varepsilon, \alpha - \alpha_\varepsilon^* \rangle_{\mathcal{U}'_{ad}, \mathcal{U}_{ad}} &\geq 0, \quad \forall \alpha \in \mathcal{U}_{ad}, \\ \alpha_\varepsilon &\in \mathcal{U}_{ad}. \end{aligned} \tag{2.40}$$

Ci-dessus, $(\lambda_\varepsilon^u, \lambda_\varepsilon^\theta)$ est la solution faible du problème adjoint

$$(\partial_{(\mathbf{u}_\varepsilon, \theta_\varepsilon)} e(\alpha_\varepsilon; \mathbf{u}_\varepsilon, \theta_\varepsilon))^* \Lambda_\varepsilon = G, \tag{2.41}$$

pour $\mathbf{G} = -\partial_{(\mathbf{u}_\varepsilon, \theta_\varepsilon)} \mathcal{J}(\alpha_\varepsilon; \mathbf{u}_\varepsilon, \theta_\varepsilon) \in W^u(0, T)' \times W^\theta(0, T)'$ dont la formulation forte est

$$\begin{aligned}
-\partial_t \lambda_\varepsilon^u - A \Delta \lambda_\varepsilon^u + h(\alpha_\varepsilon) \lambda_\varepsilon^u + (\nabla \mathbf{u}_\varepsilon)^\top \lambda_\varepsilon^u - (\mathbf{u}_\varepsilon \cdot \nabla) \lambda_\varepsilon^u - \theta_\varepsilon \nabla \lambda_\varepsilon^\theta &= g_1 \\
\nabla \cdot \lambda_\varepsilon^u &= 0, \\
-\partial_t \lambda_\varepsilon^\theta + B \lambda_\varepsilon^u \cdot \mathbf{e}_y - \nabla \cdot (\text{Ck}(\alpha_\varepsilon) \nabla \lambda_\varepsilon^\theta) - \nabla \cdot (\mathbf{u}_\varepsilon \lambda_\varepsilon^\theta) &= g_2 \\
\lambda_\varepsilon^u|_{\Gamma_w \cup \Gamma_{in}} &= 0, \\
\lambda_\varepsilon^\theta|_{\Gamma_{in}} &= 0, \\
\partial_n \lambda_\varepsilon^\theta|_{\Gamma_w} &= 0,
\end{aligned} \tag{2.42a}$$

$$\begin{aligned}
A \partial_n \lambda_\varepsilon^u|_{\Gamma_{out}} - \frac{1}{2} (\text{pos}_\varepsilon(\mathbf{u}_\varepsilon \cdot \mathbf{n}) + (\mathbf{u}_\varepsilon \cdot \mathbf{n})) \lambda_\varepsilon^u - (1 + \beta \mu_\varepsilon) \theta_\varepsilon \lambda_\varepsilon^\theta \mathbf{n} \\
- \frac{1}{2} \mu_\varepsilon \left((\mathbf{u}_\varepsilon - \mathbf{u}^{ref}) \cdot \lambda_\varepsilon^u \right) \mathbf{n} &= 0, \\
\text{Ck}(\alpha_\varepsilon) \partial_n \lambda_\varepsilon^\theta + \beta \lambda_\varepsilon^\theta \text{neg}_\varepsilon(\mathbf{u}_\varepsilon \cdot \mathbf{n})|_{\Gamma_{out}} &= 0 \\
\lambda_\varepsilon^u(T) &= 0, \\
\lambda_\varepsilon^\theta(T) &= 0,
\end{aligned}$$

$$\mu_\varepsilon = \text{neg}'_\varepsilon(\mathbf{u}_\varepsilon \cdot \mathbf{n}). \tag{2.42b}$$

Pour obtenir des conditions nécessaires d'optimalité pour le problème d'optimisation initial (2.33), nous allons passer à la limite $\varepsilon \rightarrow 0$ dans (2.40) correspondant aux conditions d'optimalité du problème d'optimisation régularisé. La difficulté principale concerne le terme μ_ε défini par (2.42b) qui fait intervenir la fonction $\text{neg}'_\varepsilon(\cdot)$. Pour passer tout de même à la limite dans cette relation, on introduit la fonction de Heaviside multivaluée $H : \mathbb{R} \rightarrow [0, 1]$ suivante

$$H(s) = \begin{cases} \{0\} & \text{si } s < 0, \\ \{1\} & \text{si } s > 0, \\ [0, 1] & \text{si } s = 0. \end{cases} \tag{2.43}$$

On donne maintenant un résultat de convergence pour $(\lambda_\varepsilon^u, \lambda_\varepsilon^\theta)$ lorsque $\varepsilon \rightarrow 0$.

Lemme 2.18 Soient $(\alpha_\varepsilon)_\varepsilon \subset \mathcal{U}_{ad}$ et $\alpha \in \mathcal{U}_{ad}$ tels que α_ε converge vers α faible* dans $BV(\Omega)$. Considérons $(\lambda_\varepsilon^u, \lambda_\varepsilon^\theta)$ une solution faible de (2.42) paramétrée par α_ε . Alors, il existe une sous-suite de $(\lambda_\varepsilon^u, \lambda_\varepsilon^\theta)$ et $\lambda^u \in L^\infty(0, T; H^u) \cap L^2(0, T; V^u)$, $\lambda^\theta \in L^\infty(0, T; H^\theta) \cap L^2(0, T; V^\theta)$ tels que

- $\lambda_\varepsilon^u \rightarrow \lambda^u$ dans $L^\infty(0, T; (L^2(\Omega))^2)$ et $\lambda_\varepsilon^\theta \rightarrow \lambda^\theta$ dans $L^\infty(0, T; L^2(\Omega))$,
- $\lambda_\varepsilon^u \xrightarrow{\varepsilon \rightarrow 0} \lambda^u$ dans $L^2(0, T; (H^1(\Omega))^2)$ et $\lambda_\varepsilon^\theta \xrightarrow{\varepsilon \rightarrow 0} \lambda^\theta$ dans $L^2(0, T; (H^1(\Omega)))$,
- $\lambda_\varepsilon^u \xrightarrow{\varepsilon \rightarrow 0} \lambda^u$ dans $L^2(0, T; (L^2(\Gamma))^2)$ et $\lambda_\varepsilon^\theta \xrightarrow{\varepsilon \rightarrow 0} \lambda^\theta$ dans $L^2(0, T; (L^2(\Gamma)))$.

De plus, il existe $\mu \in L^\infty([0, T] \times \Gamma_{out})$ donné par

$$-\mu \in H(-\mathbf{u} \cdot \mathbf{n}) \text{ p.p. sur } \Gamma_{out},$$

tel que $(\lambda^u, \lambda^\theta)$ est une solution faible de (2.42a) paramétrée par α et μ , où on a remplacé $\text{neg}_\varepsilon(\cdot)$ (resp. $\text{pos}_\varepsilon(\cdot)$) par $\text{neg}(\cdot)$ (resp. $\text{pos}(\cdot)$).

En utilisant le Lemme 2.18, nous pouvons enfin donner des conditions nécessaires d'optimalité du premier ordre pour (2.33) correspondant au problème d'optimisation avec les conditions limites non-lisses en sortie.

Théorème 2.19 *Soit α^* une solution optimale de (2.33) et (\mathbf{u}^*, θ^*) la solution faible de (2.31)-(2.32) paramétré par α^* . Alors il existe $\mu \in L^\infty([0, T] \times \Gamma_{out})$ vérifiant $-\mu \in H(-\mathbf{u} \cdot \mathbf{n})$ p.p. sur Γ_{out} et des variables adjointes $(\lambda^{\mathbf{u}}, \lambda^\theta) \in \mathbb{X}^{\mathbf{u}} \times \mathbb{X}^\theta$ satisfaisant (2.42a) tels que*

$$\langle \partial_\alpha \mathcal{J}(\alpha^*; \mathbf{u}^*, \theta^*) + (\partial_\alpha e(\alpha^*; \mathbf{u}^*, \theta^*))' \Lambda, \alpha - \alpha^* \rangle_{\mathcal{U}'_{ad}, \mathcal{U}_{ad}} \geq 0, \forall \alpha \in \mathcal{U}_{ad},$$

où on a noté $\Lambda = (\lambda^{\mathbf{u}}, \lambda^\theta, \lambda^{\mathbf{u}_0}, \lambda^{\theta_0})$ où $(\lambda^{\mathbf{u}_0}, \lambda^{\theta_0}) = (\lambda^{\mathbf{u}}(0, \cdot), \lambda^\theta(0, \cdot))$.

4 Pénalisation du modèle de Navier-Stokes-Darcy et application à des problèmes d'optimisation topologique visant des milieux poreux optimisés

Cette section présente des résultats issus de [Chap2-A2].

La majeure partie des problèmes d'optimisation topologique visent à localiser des solides dans lesquels la vitesse du fluide s'annule. La motivation principale de [Chap2-A2] était de proposer une alternative en déterminant des milieux poreux optimisés modélisés par une loi de Darcy ce qui permet de ralentir l'écoulement sans le bloquer complètement. Les interactions fluide-poreux sont modélisées par le système de Navier-Stokes incompressible dans les zones fluides couplé à la loi de Darcy dans le poreux avec des conditions de transmission spécifiques à l'interface fluide/poreux. Nous avons tout d'abord introduit un paramètre α permettant de réécrire ce problème de transmission comme un modèle valable sur tout le domaine de calcul et correspondant au système de Navier-Stokes lorsque $\alpha = 0$ et à la loi de Darcy pour $\alpha > 0$. Le modèle proposé est en fait une perturbation singulière du problème de transmission de Navier-Stokes-Darcy dont la limite (en un certain sens décrit ci-dessous) est bien le modèle initial.

On décrit ci-dessous le problème de transmission de Navier-Stokes-Darcy et sa réécriture comme une EDP valable sur tout le domaine. Soit Ω un ouvert borné de \mathbb{R}^d pour $d = 3$. Le modèle de Navier-Stokes-Darcy pénalisé donné ci-dessous s'écrit sans soucis pour $d = 2$ et les résultats donnés dans cette section seront également valables. La frontière $\partial\Omega = \overline{\Gamma}_1 \cup \overline{\Gamma}_2$ ($\overline{\Gamma}_1 \cap \overline{\Gamma}_2 = \emptyset$) est supposée Lipschitzienne et de normale unitaire sortante \mathbf{n} . On suppose que Ω se décompose comme

$$\Omega = \Omega_p \cup \Omega_f,$$

où l'indice f représente la partie fluide et p la partie poreuse du domaine. Le modèle de Navier-Stokes pour la partie fluide couplé au modèle de Darcy pour la partie poreuse est donné par

$$\left\{ \begin{array}{ll} -\nu\Delta\mathbf{u}_f + \frac{1}{\rho_f}\nabla p_f + (\mathbf{u}_f \cdot \nabla)\mathbf{u}_f = \mathbf{f} & \text{dans } \Omega_f, \\ \operatorname{div} \mathbf{u}_f = 0 & \text{dans } \Omega_f \\ \alpha\mathbf{u}_p + \frac{1}{\rho_f}\nabla p_p = \mathbf{f} & \text{dans } \Omega_p, \\ \operatorname{div} \mathbf{u}_p = 0 & \text{dans } \Omega_p, \end{array} \right. \quad (2.44)$$

où $\mathbf{f} \in L^2(\Omega)^3$ est un terme source et les conditions limites sont données plus bas. Dans (2.44), ν est la viscosité cinématique du fluide et ρ_f sa masse volumique supposée ici constante et égale à 1. Les conditions à l'interface $\Gamma_{p,f} := (\partial\Omega_f \cap \partial\Omega_p) \setminus \Gamma$ considérées sont les suivantes (voir par exemple [15, 13])

$$\begin{aligned} \mathbf{u}_f \cdot \mathbf{n}_{pf} &= \mathbf{u}_p \cdot \mathbf{n}_{pf}, \\ \left(\nu\nabla\mathbf{u}_f - p_f\mathbb{I}_3 - \frac{1}{2}|\mathbf{u}_f|^2\mathbb{I}_3 \right) \mathbf{n}_{pf} &= -p_p\mathbf{n}_{pf}, \end{aligned} \quad (2.45)$$

où \mathbf{n}_{pf} est une normale unitaire fixée de l'interface.

L'idée principale de [Chap2-A2] est de réécrire le problème de transmission (2.44) comme une seule équation valable sur tout Ω . En écrivant le terme non-linéaire comme suit :

$$(\mathbf{u} \cdot \nabla)\mathbf{u} = \frac{1}{2}\nabla|\mathbf{u}|^2 - \mathbf{u} \times (\nabla \times \mathbf{u}),$$

nous pouvons introduire le problème de Navier-Stokes-Darcy pénalisé

$$\left\{ \begin{array}{ll} -\operatorname{div}(\nu e^{-\tau\alpha}\nabla\mathbf{u} - p\mathbb{I}_3 - \frac{1}{2}e^{-2\tau\alpha}|\mathbf{u}|^2\mathbb{I}_3) - e^{-2\tau\alpha}\mathbf{u} \times (\nabla \times \mathbf{u}) + \alpha\mathbf{u} = \mathbf{f} & \text{dans } \Omega, \\ \operatorname{div}(\mathbf{u}) = 0 & \text{dans } \Omega, \\ \mathbf{u} = 0 & \text{sur } \Gamma_1, \\ \nu e^{-\tau\alpha}\partial_{\mathbf{n}}\mathbf{u} - (p + \frac{1}{2}e^{-2\tau\alpha}|\mathbf{u}|^2)\mathbf{n} = 0 & \text{sur } \Gamma_2, \end{array} \right. \quad (2.46)$$

où div est, selon le cas, la divergence d'un champ de tenseur ou de vecteur et $\alpha \in L^\infty(\Omega)$. Notons que l'on retrouve les équations de Navier-Stokes dans le fluide ($\alpha = 0$) et, en passant formellement à la limite $\tau \rightarrow +\infty$ dans (2.46), on obtient le modèle de Darcy

$$\begin{aligned} \operatorname{div} \mathbf{u} &= 0, \\ \alpha\mathbf{u} + \nabla p &= \mathbf{f} \text{ dans } \Omega_p, \end{aligned}$$

dans le milieu poreux ($\alpha > 0$). Ainsi, le modèle (2.46) peut être considéré comme un modèle de pénalisation pour le problème de transmission fluide-poreux (2.44).

Dans la suite, nous donnerons un résultat d'existence et d'unicité pour (2.46) ainsi qu'un théorème de convergence du problème de Navier-Stokes-Darcy pénalisé vers le problème de transmission initial. Une illustration numérique d'un problème d'optimisation topologique utilisant (2.46) comme contrainte sera enfin donnée.

4.1 Convergence du modèle de Navier-Stokes-Darcy pénalisé

Soit \mathbb{X} le sous-espace de $H^1(\Omega)^3$ défini par

$$\mathbb{X} := \left\{ \psi \in H^1(\Omega)^3 \mid \psi = 0 \text{ sur } \Gamma_1 \right\}.$$

La formulation variationnelle de (2.46) est alors

Trouver $(\mathbf{u}, p) \in \mathbb{X} \times L^2(\Omega)$ tels que

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) + N(\mathbf{u}, \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) &= \int_{\Omega} \mathbf{f}(x) \cdot \mathbf{v}(x) dx + \int_{\Gamma_2} \varphi \cdot \mathbf{v} d\sigma, \quad \forall \mathbf{v} \in \mathbb{X}, \\ b(\mathbf{u}, q) &= 0, \quad \forall q \in L^2(\Omega), \end{aligned} \quad (2.47)$$

où les formes multi-linéaires sont données par

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \nu \sum_{j=1}^d \int_{\Omega} e^{-\tau\alpha} \nabla \mathbf{u}_j \cdot \nabla \mathbf{v}_j dx + \int_{\Omega} \alpha \mathbf{u} \cdot \mathbf{v} dx, \\ N(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= - \int_{\Omega} \frac{e^{-2\tau\alpha}}{2} \mathbf{u} \cdot \mathbf{v} \operatorname{div}(\mathbf{w}) + e^{-2\tau\alpha} \mathbf{u} \times (\nabla \times \mathbf{v}) \cdot \mathbf{w} dx, \\ b(\mathbf{v}, p) &= - \int_{\Omega} \operatorname{div} \mathbf{v} p dx. \end{aligned} \quad (2.48)$$

L'existence d'une solution faible pour τ fixé est donnée dans le résultat suivant.

Théorème 2.20 *On suppose que les hypothèses suivantes sont vérifiées :*

1. $\alpha \in L^\infty(\Omega)$ et $\alpha(x) \geq 0$ pour presque tout $x \in \Omega$.
2. $\Omega = \Omega_p \cup \Omega_f$ avec $\Omega_f = \{x \in \Omega \mid \alpha(x) = 0\}$ vérifiant $\overline{\Omega_f} \cap \Gamma_1 \neq \emptyset$.
3. Ω_p et Ω_f sont des ouverts bornés de \mathbb{R}^3 de frontière Lipschitz.
4. Il existe $\alpha_{\min} > 0$ tel que

$$\text{p.p. } x \in \Omega_p : \alpha_{\min} \leq \alpha(x).$$

Alors, pour tout $M > 0$ et tout τ tel que $0 < \tau \leq M$, il existe au moins une solution $(\mathbf{u}_\tau, p_\tau)$ de (2.47).

Le Théorème 2.20 donne l'existence d'une solution faible pour le problème de Navier-Stokes-Darcy pénalisé (2.46). L'unicité peut être également obtenue lorsque ν est suffisamment grand ou bien que les termes sources sont assez petits (voir par exemple Théorème 2.9).

Pour étudier la convergence lorsque $\tau \rightarrow +\infty$ des solutions de (2.46), il faut commencer par réécrire (2.46) comme un problème de transmission. En utilisant que $\Omega = \Omega_f \cup \Omega_p$ et en omettant la dépendance en τ pour alléger les expressions, le problème (2.46) s'écrit sous la forme suivante :

$$\left\{ \begin{array}{l} -\operatorname{div}(\nu e^{-\tau\alpha} \nabla \mathbf{u}_p - p_{\tau,p} \mathbb{I}_3 - \frac{1}{2} e^{-2\tau\alpha} |\mathbf{u}_p|^2 \mathbb{I}_3) \\ \quad - e^{-2\tau\alpha} \mathbf{u}_p \times (\nabla \times \mathbf{u}_p) + \alpha \mathbf{u}_p = \mathbf{f} \text{ dans } \Omega_p, \\ \operatorname{div}(\mathbf{u}_p) = 0 \text{ dans } \Omega_p, \\ -\operatorname{div}(\nu \nabla \mathbf{u}_f - p_f \mathbb{I}_3 - \frac{1}{2} |\mathbf{u}_f|^2 \mathbb{I}_3) - \mathbf{u}_f \times (\nabla \times \mathbf{u}_f) = \mathbf{f} \text{ dans } \Omega_f, \\ \operatorname{div}(\mathbf{u}_f) = 0 \text{ dans } \Omega_f, \\ \mathbf{u}_f = 0 \text{ sur } \Gamma_1 \cap \partial\Omega_f, \\ \mathbf{u}_p = 0 \text{ sur } \Gamma_1 \cap \partial\Omega_p, \\ \nu \partial_n \mathbf{u}_f - (p_f + \frac{1}{2} |\mathbf{u}_f|^2) \mathbf{n} = 0 \text{ sur } \Gamma_2 \cap \partial\Omega_f, \\ e^{-\tau\alpha} \nu \partial_n \mathbf{u}_p - (p_p + \frac{1}{2} e^{-2\tau\alpha} |\mathbf{u}_p|^2) \mathbf{n} = 0 \text{ sur } \Gamma_2 \cap \partial\Omega_p, \\ \mathbf{u}_f \cdot \mathbf{n}_{pf} - \mathbf{u}_p \cdot \mathbf{n}_{pf} = 0 \text{ sur } \Gamma_{p,f}, \\ \mathcal{T}_\tau[(\mathbf{u}_p, p_p), (\mathbf{u}_f, p_f)] = 0 \text{ sur } \Gamma_{p,f}, \end{array} \right. \quad (2.49)$$

Le Théorème 2.20 assure l'existence de $(\mathbf{u}_{s,\tau}, p_{\tau,p}, \mathbf{u}_{f,\tau}, p_{\tau,f}) \in H^1(\Omega_p)^3 \times L^2(\Omega_p) \times H^1(\Omega_f)^3 \times L^2(\Omega_f)^3$ qui est une solution faible du problème de transmission (2.49)-(2.45) pour τ fixé. Pour pouvoir passer à la limite $\tau \rightarrow +\infty$, nous avons tout d'abord dérivé, pour la solution de (2.49)-(2.45), des estimations dans lesquelles la dépendance en τ est explicite.

Théorème 2.21 *Supposons que les hypothèses du Théorème 2.20 sont valides et que $\alpha|_{\Omega_p} \in \text{Lip}(\Omega_p)$. Alors la solution de (2.49) satisfait les estimations ci-dessous*

$$\begin{aligned} \|\mathbf{u}_{f,\tau}\|_{L^2(\Omega_f)^3} + \|\mathbf{u}_{p,\tau}\|_{L^2(\Omega_p)^3} &\lesssim \|\mathbf{f}\|_{L^2(\Omega)^3}, \\ \|\nabla \mathbf{u}_{f,\tau}\|_{L^2(\Omega_f)^3} &\lesssim \|\mathbf{f}\|_{L^2(\Omega)^3}, \\ \|e^{-\tau\alpha} \nabla \mathbf{u}_{p,\tau}\|_{L^2(\Omega_p)^3} &\lesssim e^{-\tau\alpha_{\min}/2} \|\mathbf{f}\|_{L^2(\Omega)^3}, \\ \|\mathbf{p}_{p,\tau}\|_{L^2(\Omega_p)} + \|\mathbf{p}_{f,\tau}\|_{L^2(\Omega_f)} &\lesssim C, \end{aligned}$$

où $C > 0$ et la notation \lesssim indique que les constantes non-mentionnées dans les inégalités sont indépendantes de τ .

Le Théorème 2.21 entraîne que la suite $((\mathbf{u}_{p,\tau}, \mathbf{p}_{p,\tau}, \mathbf{u}_{f,\tau}, \mathbf{p}_{f,\tau}))_\tau$ est bornée, uniformément en τ , dans $L^2(\Omega_p)^3 \times L^2(\Omega_p) \times H^1(\Omega_f)^3 \times L^2(\Omega_f)$. Il existe donc $(\mathbf{u}_p, \mathbf{p}_p, \mathbf{u}_f, \mathbf{p}_f) \in L^2(\Omega_p)^3 \times L^2(\Omega_p) \times H^1(\Omega_f) \times L^2(\Omega_f)$ et une sous-suite telle que

$$\begin{aligned} \mathbf{u}_{s,\tau} &\rightarrow \mathbf{u}_p \text{ faible dans } L^2(\Omega_p)^3, \\ \mathbf{u}_{f,\tau} &\rightarrow \mathbf{u}_f \text{ faible dans } H^1(\Omega_f)^3 \text{ et fort dans } L^2(\Omega_f)^3, \\ \mathbf{p}_{p,\tau} &\rightarrow \mathbf{p}_p \text{ faible dans } L^2(\Omega_p) \text{ et } \mathbf{p}_{f,\tau} \rightarrow \mathbf{p}_f \text{ faible dans } L^2(\Omega_f), \\ e^{-\tau\alpha} \nabla \mathbf{u}_{s,\tau} &\rightarrow 0 \text{ fort dans } L^2(\Omega_p). \end{aligned} \tag{2.50}$$

Les convergences précédentes permettent de passer à la limite dans la formulation variationnelle du problème de Navier-Stokes-Darcy pénalisé (2.49).

Théorème 2.22 *Les distributions $(\mathbf{u}_p, \mathbf{p}_p, \mathbf{u}_f, \mathbf{p}_f) \in L^2(\Omega_p)^3 \times L^2(\Omega_f) \times H^1(\Omega_f)^3 \times L^2(\Omega_f)$ vérifie le problème limite ci-dessous*

$$\left\{ \begin{array}{ll} \nabla \mathbf{p}_p + \alpha \mathbf{u}_p = \mathbf{f} & \text{dans } \mathcal{D}'(\Omega_p, \mathbb{R}^3), \\ \operatorname{div}(\mathbf{u}_p) = 0 & \text{dans } \mathcal{D}'(\Omega_p), \\ -\operatorname{div}(\nu \nabla \mathbf{u}_f - \mathbf{p}_f \mathbb{I}_3) + (\mathbf{u}_f \cdot \nabla) \mathbf{u}_f = \mathbf{f} & \text{dans } \mathcal{D}'(\Omega_f, \mathbb{R}^3), \\ \operatorname{div}(\mathbf{u}_f) = 0 & \text{dans } \mathcal{D}'(\Omega_f), \\ \mathbf{u}_f = 0 & \text{dans } H^{1/2}(\Gamma_1 \cap \partial\Omega_f, \mathbb{R}^3), \\ \mathbf{n} \cdot \mathbf{u}_p = 0 & \text{dans } H^{-1/2}(\Gamma_1 \cap \partial\Omega_p) \\ \mathbf{u}_f \cdot \mathbf{n}_{pf} = \mathbf{u}_p \cdot \mathbf{n}_{pf} & \text{dans } H^{-1/2}(\Gamma_{p,f}), \\ \nu \partial_{\mathbf{n}} \mathbf{u}_f - (\mathbf{p}_f + \frac{1}{2} |\mathbf{u}_f|^2) \mathbf{n} = 0 & \text{dans } H^{-1/2}(\Gamma_2 \cap \partial\Omega_f, \mathbb{R}^3), \\ -\mathbf{p}_p = 0 & \text{dans } H^{1/2}(\Gamma_2 \cap \partial\Omega_p), \\ (\nu \nabla \mathbf{u}_f - \mathbf{p}_f \mathbb{I}_3 - \frac{1}{2} |\mathbf{u}_f|^2 \mathbb{I}_3) \mathbf{n}_{pf} = -\mathbf{p}_p \mathbf{n}_{pf} & \text{dans } (W^{1-1/q, q}(\Gamma_{p,f}, \mathbb{R}^3))', \end{array} \right. \tag{2.51}$$

où $q = 2p/(2-p)$ pour tout $p \in [1, 3/2]$.

Le Théorème 2.22 permet d'affirmer que la solution faible du problème pénalisé (2.49) converge vers une solution faible du problème de transmission fluide-poreux modélisé par le système de Navier-Stokes pour la zone fluide et la loi de Darcy pour le milieu poreux avec les conditions de transmission (2.45).

4.2 Optimisation topologique visant des milieux poreux optimisés

Le problème d'optimisation topologique visant un milieu poreux optimisé peut être formulé comme un problème d'optimisation où α est le paramètre d'optimisation et les contraintes sont données par le modèle de Navier-Stokes-Darcy pénalisé (2.47). On a ainsi

$$\min_{(\mathbf{u}, \mathbf{p}, \alpha) \in \mathbb{X} \times L^2(\Omega) \times \mathcal{U}} \mathcal{J}(\alpha; \mathbf{u}, \mathbf{p}) \text{ où } (\mathbf{u}, \mathbf{p}) \text{ satisfait (2.47) et } \alpha \in \mathcal{U}_{\text{ad}}, \quad (2.52)$$

où \mathcal{U}_{ad} est l'ensemble des porosités admissibles. Dans [Chap2-A2], nous avons démontré l'existence d'au moins une solution optimale pour (2.52) et des conditions nécessaires d'optimalité du premier ordre dans le cas où la localisation du poreux est fixée dans $\Omega_p \subset \Omega$ et \mathcal{U}_{ad} est un sous-ensemble d'un espace de Sobolev permettant d'assurer que α est Lipschitzienne dans Ω_p . Notons que ce type de régularité a été obtenu en considérant un ensemble admissible $\mathcal{U}_{\text{ad}} \subset W^{k,2}(\Omega_p)$ pour k assez grand.

Pour illustrer numériquement l'approche, on considère le problème d'optimisation (2.52) avec $\Omega = [0, 1] \times [0, 2]$ et

$$\mathcal{J}(\alpha) = \frac{1}{2} \int_{\Omega} |\mathbf{u} - \mathbf{u}_d|^2 dx, \quad \mathbf{u}_d = \mathbf{e}_x,$$

où, pour simplifier, on considère le modèle de Stokes-Darcy pénalisé qui est donné par

$$\left\{ \begin{array}{ll} -\text{div}(\nu \mathbf{e}^{-\tau\alpha} \nabla \mathbf{u} - p \mathbb{I}_3) + \alpha \mathbf{u} = 0 & \text{dans } \Omega, \\ \text{div}(\mathbf{u}) = 0 & \text{dans } \Omega, \\ \mathbf{u} = \mathbf{u}_0 & \text{sur } \{0\} \times [0, 1], \\ \mathbf{u} = 0 & \text{sur } [0, 2] \times (\{0\} \cup \{1\}), \\ \nu \mathbf{e}^{-\tau\alpha} \partial_{\mathbf{n}} \mathbf{u} - p \mathbf{n} = 0 & \text{sur } \{2\} \times [0, 1]. \end{array} \right. \quad (2.53)$$

Notons que le Théorème 2.22 est valable pour le modèle de Darcy-Stokes pénalisé (2.53). Ainsi, lorsque $\tau \rightarrow +\infty$, la solution faible de (2.53) converge vers la solution faible du problème de transmission fluide-poreux où le fluide est modélisé par le système de Stokes et le milieu poreux par la loi de Darcy. De plus, les conditions d'interface sont

$$\begin{aligned} \mathbf{u}_f \cdot \mathbf{n}_{pf} &= \mathbf{u}_p \cdot \mathbf{n}_{pf}, \\ (\nu \nabla \mathbf{u}_f - p_f \mathbb{I}_3) \mathbf{n}_{pf} &= -p_p \mathbf{n}_{pf}, \end{aligned}$$

et sont similaires à celles données dans (2.45) sans le terme non-linéaire.

Le problème d'optimisation visant un poreux optimisé devient donc

$$\min_{(\mathbf{u}, \mathbf{p}, \alpha) \in \mathbb{X} \times L^2(\Omega) \times \mathcal{U}} \mathcal{J}(\alpha; \mathbf{u}, \mathbf{p}) \text{ où } (\mathbf{u}, \mathbf{p}) \text{ vérifie (2.53) et } \alpha \in \mathcal{U}_{\text{ad}}. \quad (2.54)$$

La résolution numérique de (2.54) est réalisée avec un algorithme de descente où le gradient est calculé à partir du modèle adjoint qui est donné par

$$\left\{ \begin{array}{ll} -\text{div}(\nu \mathbf{e}^{-\tau\alpha} \nabla \mathbf{v} - q \mathbb{I}_3) + \alpha \mathbf{v} = -(\mathbf{u} - \mathbf{u}_d) & \text{dans } \Omega, \\ \text{div}(\mathbf{v}) = 0 & \text{dans } \Omega, \\ \mathbf{v} = 0 & \text{sur } \{0\} \times [0, 1], \\ \mathbf{v} = 0 & \text{sur } [0, 2] \times (\{0\} \cup \{1\}), \\ \nu \mathbf{e}^{-\tau\alpha} \partial_{\mathbf{n}} \mathbf{v} - q \mathbf{n} = 0 & \text{sur } \{2\} \times [0, 1]. \end{array} \right. \quad (2.55)$$

L'algorithme utilisé est donné ci-dessous :

Algorithme 1 :

1. $\alpha = 0 \in \mathbb{P}^0$.
2. Calculer $(\mathbf{u}_h, p_h) \in \mathbb{P}^2 \times \mathbb{P}^1$ solution de (2.53).
3. Calculer $(\mathbf{v}_h, q_h) \in \mathbb{P}^2 \times \mathbb{P}^1$ solution du problème adjoint (2.55).
4. Mise à jour de la porosité :

$$\beta = \max \left\{ 0, \alpha - \lambda \Pi_h \left(-\nu \tau e^{-\tau \alpha} \sum_{j=1}^d (\nabla \mathbf{u}_{j,h} \cdot \nabla \mathbf{v}_{j,h}) + \mathbf{u}_h \cdot \mathbf{v}_h \right) \right\},$$

$$\alpha = \max \{ \alpha_{\min}, \beta \} \chi_{(\beta > 0)}$$

où $\Pi_h : L^2(\Omega) \rightarrow \mathbb{P}^0$ est le projecteur éléments finis et le max est réalisé sur chaque cellule du maillage. De plus, $\chi_{(\beta > 0)}$ est défini sur chaque cellule et renvoie 0 si $\beta = 0$ et la valeur de β si $\beta > 0$.

5. Retour à l'étape 2.

Dans **Algorithme 1**, on cherche, sur chaque cellule du maillage, une valeur pour α qui est soit nulle (zone fluide) soit supérieure à α_{\min} dans le milieu poreux.

Les simulations ci-dessous ont été réalisées avec FREE-FEM++ [24] en utilisant les paramètres suivants

$$\alpha_{\min} = 10, \nu = 1, \tau = 10, n = 100, \lambda = 10, n_{\text{iter}} = 1500, \varepsilon = 0.000001.$$

Nous avons ainsi effectué 1500 itérations de l'**Algorithme 1** car cela à suffit à minimiser la fonctionnelle qui est passée de $\mathcal{J} = 0.1801$ avant optimisation à $\mathcal{J} = 0.0043$ après ajout du milieu poreux optimisé.

La Figure 2.7 montre les composantes du champ de vitesse $\mathbf{u} = u_x \mathbf{e}_x + u_y \mathbf{e}_y$ avant optimisation et la Figure 2.8 celui après optimisation.

Le poreux optimisé est représenté dans la Figure 2.9. On peut voir plus précisément la localisation du poreux optimisé sur la Figure 2.9 qui représente $\exp(-\tau\alpha)$. Comme, lorsque $\alpha > 0$, le terme $\exp(-\tau\alpha)$ est numériquement nul, on a donc le modèle de Stokes dans les zones bleues et celui de Darcy dans les zones rouges. Ces résultats numériques montrent que le modèle de Stokes-Darcy pénalisé peut être utilisé comme contrainte dans un problème d'optimisation topologique visant un milieu poreux optimisé.

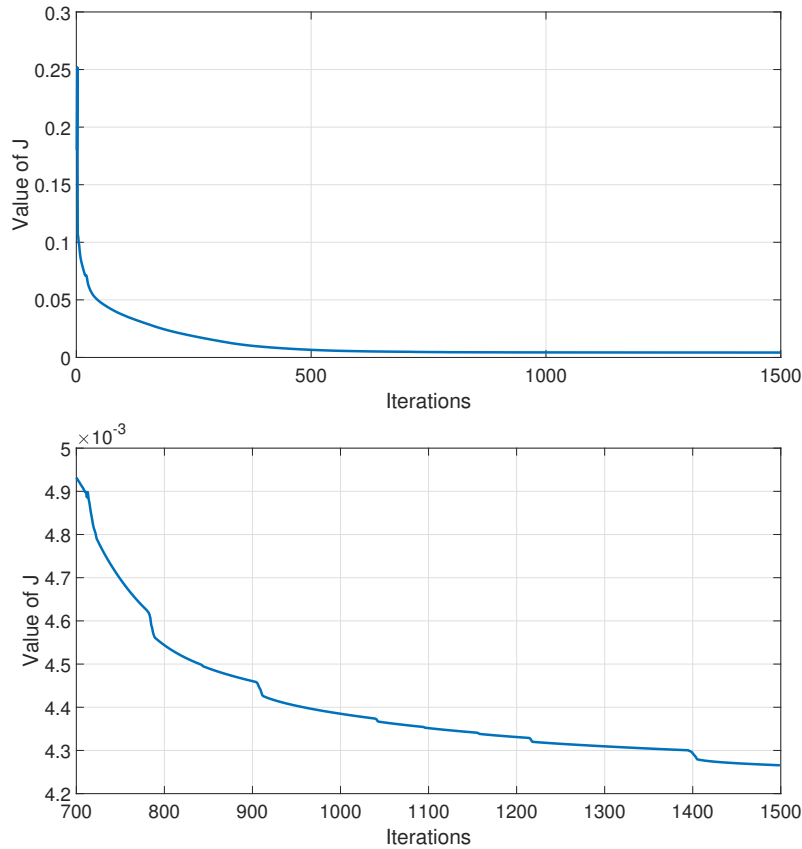


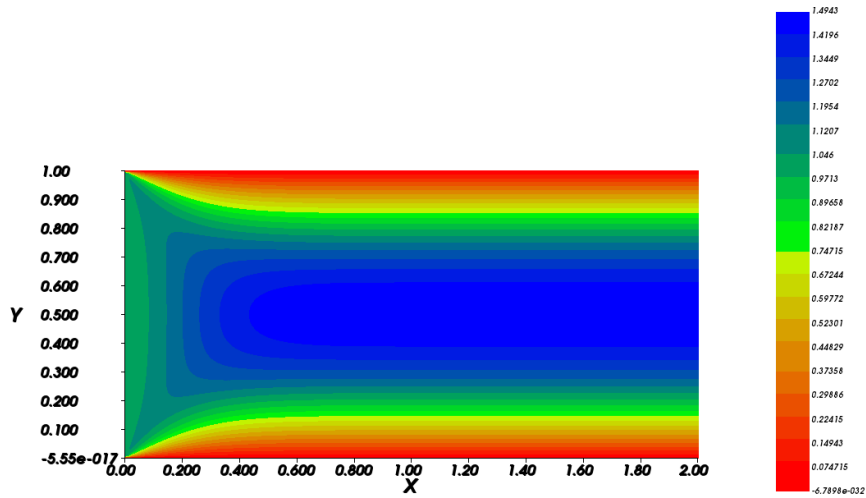
FIGURE 2.6 – Évolution de la fonction coût au cours des itérations de l’Algorithme 1. (En haut) Valeur de \mathcal{J} au cours des itérations. (En bas) Zoom sur les valeur de \mathcal{J} à partir de l’itération 700.

5 Contributions à l’étude physique de problèmes d’optimisation topologique

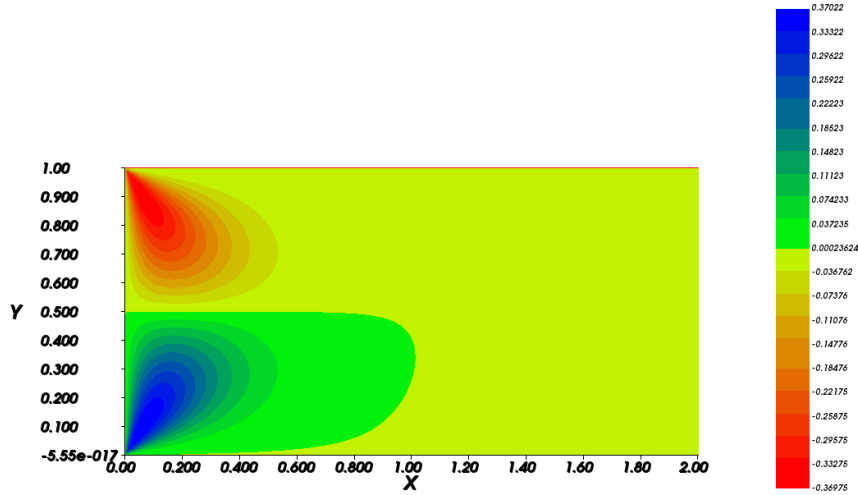
Cette partie présente succinctement des résultats issus de [Chap2-A1] et [Chap2-A3] qui sont également issus de la thèse de Delphine Ramalingom [31].

Ces deux papiers traitent d’optimisation topologique pour des écoulements avec prise en compte de la convection forcée. Le modèle considéré est le système de Navier-Stokes incompressible sous l’approximation de Boussinesq donné par (2.14)-(2.15).

Pour chaque résultat présenté ci-dessous, la résolution numérique du problème d’optimisation est basée sur l’approche *optimiser-puis-discrétiser*. On utilise un algorithme de descente de gradient où le gradient (2.30) est calculé avec le modèle adjoint (2.28)-(2.29). Les systèmes primal (2.14)-(2.15) et adjoint (2.28)-(2.29) sont tous deux discrétisés avec des volumes finis co-localisés, la résolution est réalisée avec l’algorithme SIMPLE et l’implémentation avec OPENFOAM [25].



u1 without alpha



u2 without alpha

FIGURE 2.7 – Champ de vitesse avant optimisation. (En haut) u_x . (En bas) u_y .

La mise à jour du paramètre d'optimisation α à l'itération k est réalisée par

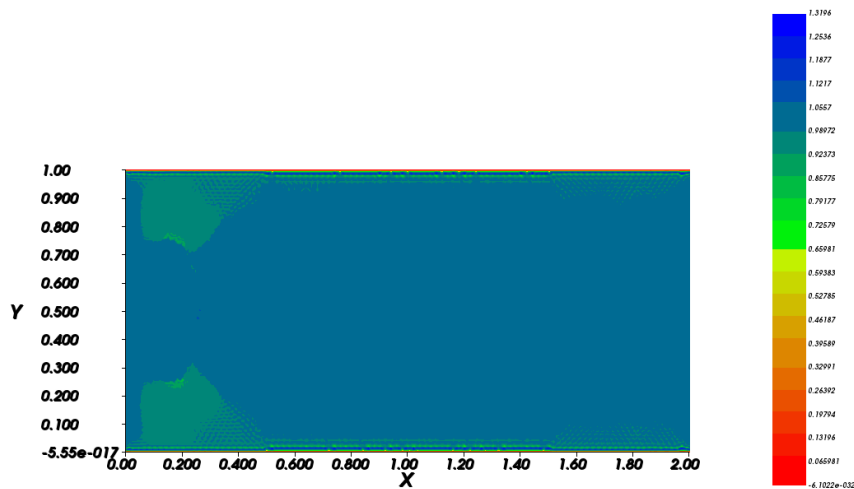
$$\alpha_{k+1} = \max(0, \min\{0, \alpha_k + \omega_k d_k\}).$$

Ci-dessus, la direction d_k est

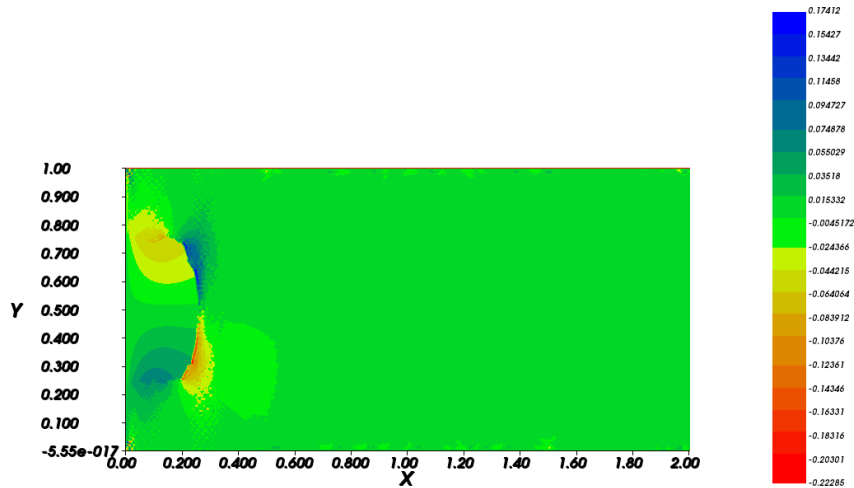
$$d_k = -\nabla_{\alpha} \mathcal{J}(\alpha_k) + \beta_k^{\text{PR}} d_{k-1},$$

où β_k^{PR} est la direction de descente associée à la méthode de Polack-Ribiere

$$\beta_k^{\text{PR}} = \frac{\nabla \mathcal{J}_k^T (\nabla \mathcal{J}_k - \nabla \mathcal{J}_{k-1})}{\nabla \mathcal{J}_{k-1}^T \nabla \mathcal{J}_{k-1}}, \quad \nabla \mathcal{J}_{k-1} = \nabla_{\alpha} \mathcal{J}(\alpha_{k-1}).$$



u1



u2

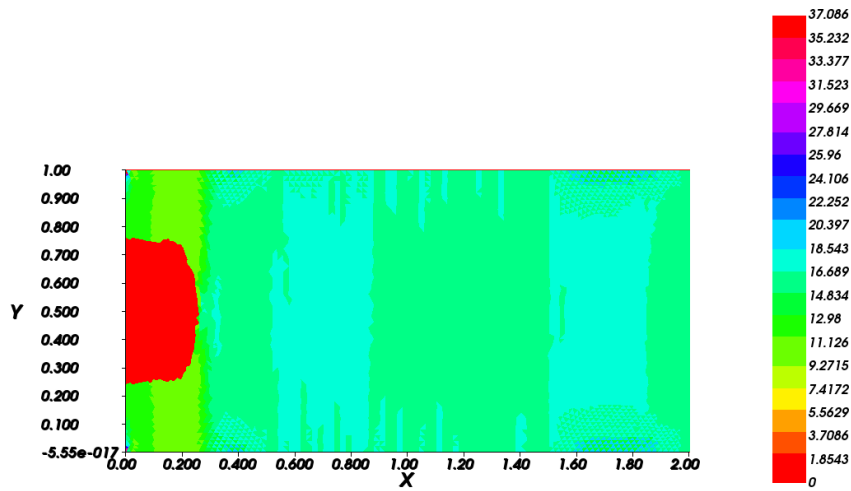
FIGURE 2.8 – Champ de vitesse après optimisation. (En haut) u_x . (En bas) u_y .

Une fois la direction d_k calculée, on met à jour le pas ω_k avec une recherche linéaire puis on itère jusqu'à avoir $\|\nabla_{\alpha} \mathcal{J}(\alpha_{k+1})\|_2 \leq \varepsilon$ pour une tolérance fixée ($\varepsilon = 10^{-7}$ dans les exemples ci-dessous).

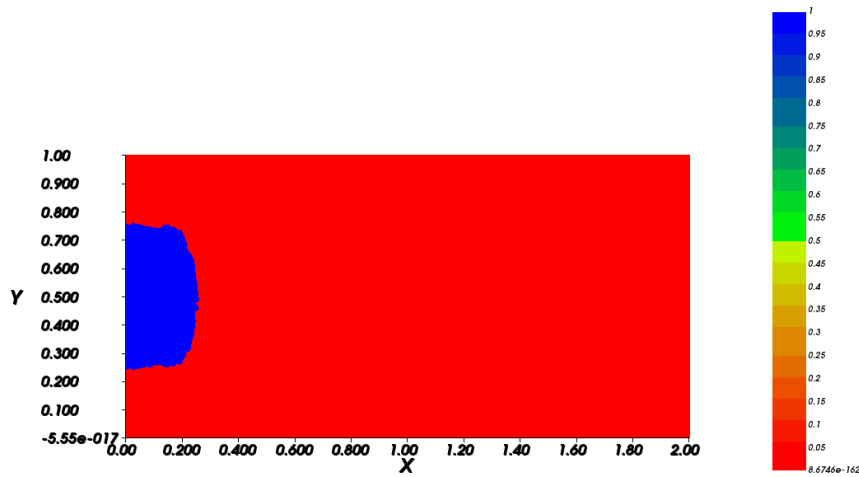
On présente maintenant quelques résultats numériques obtenus à partir de l'algorithme brièvement décrit ci-dessus.

5.1 Optimisation d'un *single pipe* et d'un *bend pipe*

Un des apports principaux de [Chap2-A1] a été de proposer d'utiliser (2.18) comme fonction d'interpolation fluide-solide. Cette fonction d'interpolation permet notamment



alpha P0



exp(-tau*alpha) P0

FIGURE 2.9 – α après 1500 itérations de l’Algorithme 1. (En haut) α . (En bas) $\exp(-\tau\alpha)$ où le fluide est représenté par la zone bleue et le poreux par la zone rouge.

d’éviter les motifs en damier dans les solides optimisés obtenus numériquement sans avoir à recourir à des techniques de filtrage ou de régularisation communément utilisées dans la littérature (voir par ex. [30, 36]). Nous avons ensuite optimisé la forme d’une *single pipe* et d’un *bend pipe* (voir Figure 2.10) en cherchant à minimiser les pertes de charge tout en maximisant les transferts thermiques. Ces deux géométries ont été considérées pour valider le choix de la fonction d’interpolation fluide-solide car elles correspondent à des cas classiques de la littérature. Les pertes de charge sont définies par

$$f_1(\mathbf{u}, p) = \int_{\Gamma} -\mathbf{n} \cdot \mathbf{u} \left(p + \frac{1}{2} |\mathbf{u}|^2 \right) d\sigma, \quad (2.56)$$

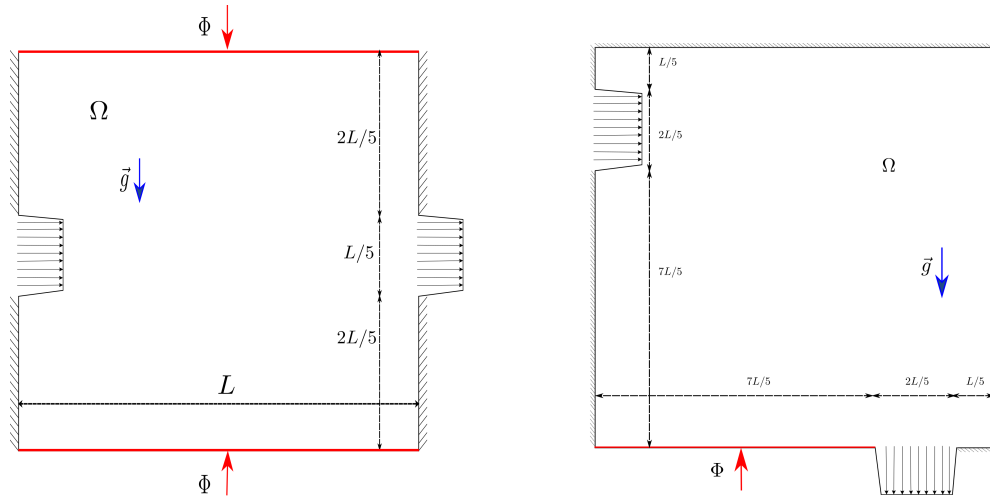


FIGURE 2.10 – Géométries étudiées dans [Chap2-A1], (Droite) le *single pipe*. (Gauche) le *bend pipe*.

et les transferts thermiques par

$$f_2(\mathbf{u}, \theta) = \int_{\Gamma} \mathbf{n} \cdot \mathbf{u} \theta \, d\sigma. \quad (2.57)$$

La fonction coût permettant de minimiser les pertes de charge tout en maximisant les transferts est définie par $\mathcal{J} = f_1 - f_2$ et le problème d'optimisation est ainsi :

$$\begin{aligned} \min \quad & - \int_{\Gamma} \left(p + \frac{1}{2} |\mathbf{u}|^2 + \theta \right) \mathbf{u} \cdot \mathbf{n} \, d\Gamma, \\ \text{où} \quad & (\mathbf{u}, p, \theta) \text{ vérifie (2.14) - (2.15) avec } \Phi = -1/k_{\tau}, \mathbf{u}_{\text{in}} = \mathbf{e}_x. \end{aligned} \quad (2.58)$$

Concernant les paramètres physiques, nous avons utilisé

$$\text{Re} = 500, \text{ Ri} = 2.8, \text{ Pr} = 0.71, \frac{k_s}{k_f} = \frac{(9.88 \times 10^{-5} \text{ m}^2 \cdot \text{s}^{-1})}{(2.25 \times 10^{-5} \text{ m}^2 \cdot \text{s}^{-1})},$$

où le ratio des conductivités correspond à la conductivité thermique de l'aluminium sur celle de l'air à 25°C. Les paramètres de la fonction d'interpolation fluide-solide sont fixés à $\alpha_{\text{max}} = 200$, $\alpha_0 = 50$ et $\tau = 2$.

Certains résultats issus de [Chap2-A1] sont présentés dans la Figure 2.11. Pour le *single pipe*, la valeur de la fonction coût sans optim était $\mathcal{J} = 0.015$ et de $\mathcal{J}^* = 6.966 \times 10^{-3}$ après optimisation. Pour le *bend pipe*, on a $\mathcal{J} = 0.008$ sans optimisation et $\mathcal{J}^* = 5.067 \times 10^{-3}$. Ainsi, pour ces deux cas, nous avons bien minimisé la fonction coût et les formes obtenus permettent de diminuer les pertes de charge tout en augmentant les transferts thermiques. Des résultats similaires ont été obtenus pour différentes valeurs des paramètres α_0 et τ .

5.2 Optimisation d'un canal vertical asymétriquement chauffé

L'objectif de [Chap2-A3] est de résoudre un problème d'optimisation topologique en convection forcée pour minimiser les pertes de charge tout en maximisant les transferts

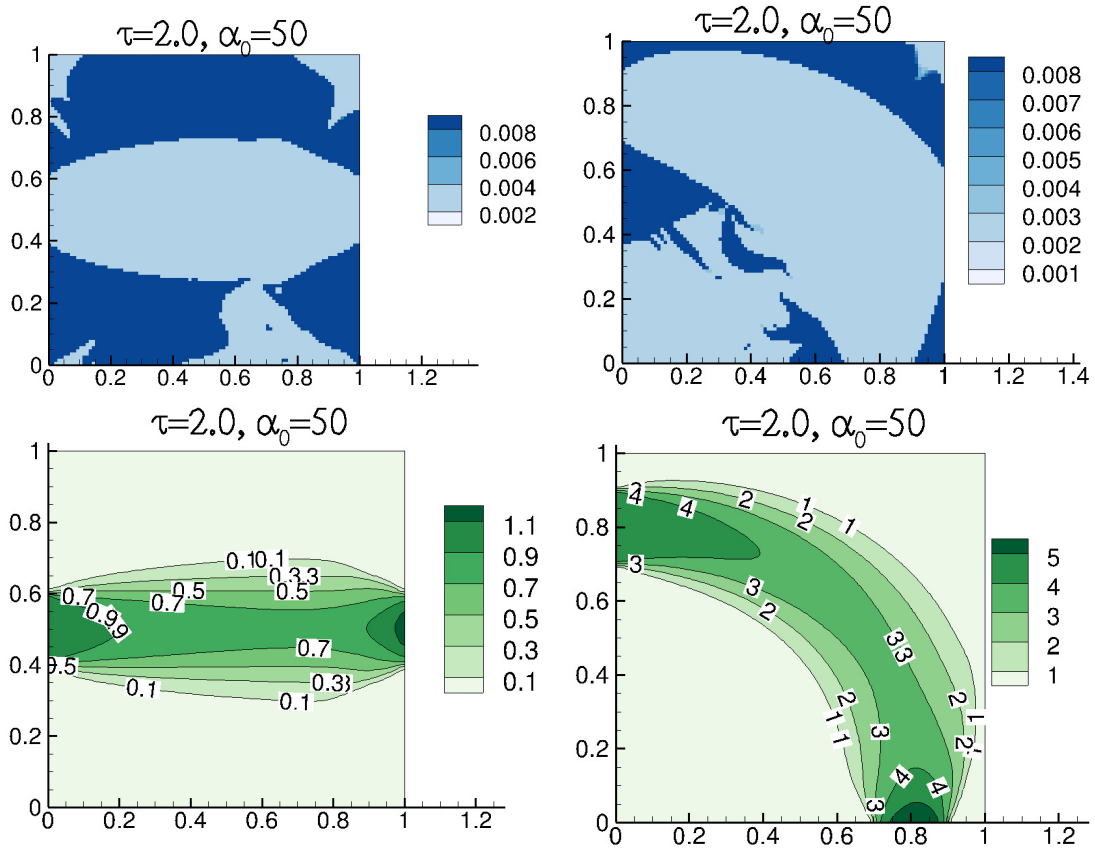


FIGURE 2.11 – (Haut) Valeur de $k_f k_\tau(\alpha)$ pour le α optimisé. (Bas) Module du champ de vitesse et lignes de courant. (Gauche) Résultats pour le *single pipe*. (Droite) Résultats pour le *bend pipe*. Le bleu foncé correspond aux zones solides du domaine et le bleu clair aux zones fluides.

thermiques dans un canal vertical asymétriquement chauffé (voir Figure 2.12). Cette géométrie permet par exemple de modéliser un mur Trombe [35] ou encore des cheminées solaires [19].

Un des premiers apports de [Chap2-A3] a été de montrer que la fonction coût (2.56) définissant les pertes de charge ne correspond pas dans certains cas à l'énergie nécessaire au fluide pour se mouvoir. Pour toute solution (\mathbf{u}, p, θ) de (2.14)-(2.15), on a en effet

$$\begin{aligned}
 f_1(\mathbf{u}, p) &= \int_{\Omega} \text{Re}^{-1} |\nabla \mathbf{u}|^2 + h_\tau(\gamma) |\mathbf{u}|^2 - \text{Ri}(\mathbf{u} \cdot \mathbf{e}_y) \theta \, dx \\
 &\quad - \text{Re}^{-1} \int_{\Gamma_{\text{out}}} (\mathbf{u} \cdot \mathbf{n}) p \, d\sigma - \text{Re}^{-1} \int_{\Gamma_{\text{in}}} \partial_{\mathbf{n}} \mathbf{u} \cdot \mathbf{u}_{\text{in}} \, d\sigma.
 \end{aligned}$$

Dans les simulations Volumes Finis réalisées dans [Chap2-A3], les conditions limites en sortie imposent en particulier une pression nulle $p = 0$ et les conditions limites en entrée imposent $\mathbf{u}_{\text{in}} = u_i \mathbf{e}_y$. Ainsi, dans le cas où $\text{Ri} = 0$ et la vitesse du fluide est constante au travers de l'entrée (dans ce cas $\partial_{\mathbf{n}} \mathbf{u} \cdot \mathbf{e}_y = 0$), la fonctionnelle f_1 devient

$$f_1(\mathbf{u}, p) = \int_{\Omega} \text{Re}^{-1} |\nabla \mathbf{u}|^2 + h_\tau(\gamma) |\mathbf{u}|^2 \, dx,$$

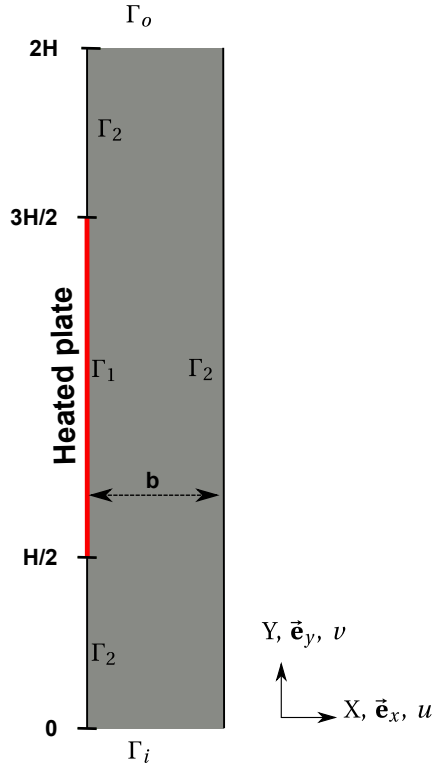


FIGURE 2.12 – Canal vertical asymétriquement chauffé

correspondant à l'énergie nécessaire au fluide pour se mouvoir (voir par ex. [29] pour des considérations similaires).

Les applications visées dans [Chap2-A3] utilisent toutefois des grands Ri , par exemple $Ri \geq 100$, ce qui nous a conduit à proposer une autre fonction coût pour réduire les pertes de charge et augmenter les transferts thermiques dans le canal. Ces fonctionnelles sont basées sur une approche systémique appliquée au canal vu comme un système entrée/sortie de vitesses moyennes et de températures. La version systémique des pertes de charge (ou encore de la puissance mécanique) est alors

$$\mathcal{J}_1(\mathbf{u}, p) = -\frac{1}{|\Gamma_{in}|} \int_{\Gamma_{in}} p_t \, d\sigma \int_{\Gamma_{in}} \mathbf{u} \cdot \mathbf{n} \, d\sigma - \frac{1}{|\Gamma_{out}|} \int_{\Gamma_{out}} p_t \, d\sigma \int_{\Gamma_{out}} \mathbf{u} \cdot \mathbf{n} \, d\sigma, \quad (2.59)$$

où $p_t = p + |\mathbf{u}|^2/2$ est la pression totale. La version systémique des transferts thermiques (2.57) est

$$\mathcal{J}_2(\mathbf{u}, \theta) = \frac{1}{|\Gamma_{in}|} \int_{\Gamma_{in}} \theta \, d\sigma \int_{\Gamma_{in}} \mathbf{u} \cdot \mathbf{n} \, d\sigma + \frac{1}{|\Gamma_{out}|} \int_{\Gamma_{out}} \theta \, d\sigma \int_{\Gamma_{out}} \mathbf{u} \cdot \mathbf{n} \, d\sigma. \quad (2.60)$$

La fonction coût considérée dans [Chap2-A3] vise à minimiser les pertes de charge données par (2.59) tout en maximisant les transferts thermiques avec (2.60). La résolution de ce problème d'optimisation multi-objectifs a été effectuée en mettant chaque fonctionnelle à la même échelle en considérant la fonction coût suivante

$$\hat{\mathcal{J}} = \omega \hat{\mathcal{J}}_1 - (1 - \omega) \hat{\mathcal{J}}_2, \quad \omega \in [0, 1],$$

où

$$\hat{f} = \frac{f - f_{\min}}{f_{\max} - f_{\min}}$$

et $\mathcal{J}_{i,\max}$ est la valeur de la fonction coût sans optimisation tandis que $\mathcal{J}_{i,\min}$ correspond à la valeur obtenue après optimisation en minimisant uniquement la fonctionnelle \mathcal{J}_i . On a ainsi résolu le problème d'optimisation suivant :

$$\begin{aligned} \min \quad & \hat{\mathcal{J}}(\mathbf{u}, p, \theta) \\ \text{où} \quad & (\mathbf{u}, p, \theta) \text{ vérifient (2.14) - (2.15) avec } \Phi = -1/k_\tau, \mathbf{u}_{\text{in}} = 6.1 \times (1 - \chi) \mathbf{e}_y. \end{aligned} \quad (2.61)$$

Les paramètres de la fonction d'interpolation sont

$$\alpha_0 = 20, \tau = 0.6 \text{ et } \alpha_{\max} = 2 \times 10^5,$$

et les paramètres physiques

$$\text{Re} = 400, \text{ Ri} = 100, 200, 400, \text{ Pr} = 0.71, \frac{k_s}{k_f} = 3.$$

Le paramètre d'optimisation α est supposé constant par morceaux sur le maillage et on impose des zones de fluide près de l'entrée et de la paroi chauffée. Ainsi :

$$\alpha = 0 \text{ sur } [0, 1] \times [0, 1] \cap [0, 0.10] \times [2.5, 7.5].$$

Les écoulements sans optimisation sont donnés dans la Figure 2.13 où les zones bleues correspondent à un écoulement retour au niveau de la sortie. Les écoulements après optimisation sont donnés dans la Figure 2.14 où les zones noires représentent les zones solides. On peut voir que les zones solides sont placées dans l'écoulement retour permettant ainsi de le réduire. La Figure 2.15 représente chaque fonction coût au fil des itérations de l'algorithme de descente présenté en début de section. On peut remarquer que \mathcal{J}_1 est bien minimisé, \mathcal{J}_2 est maximisé et ainsi les formes obtenus permettent bien de réduire les pertes de charge tout en augmentant les transferts thermiques.

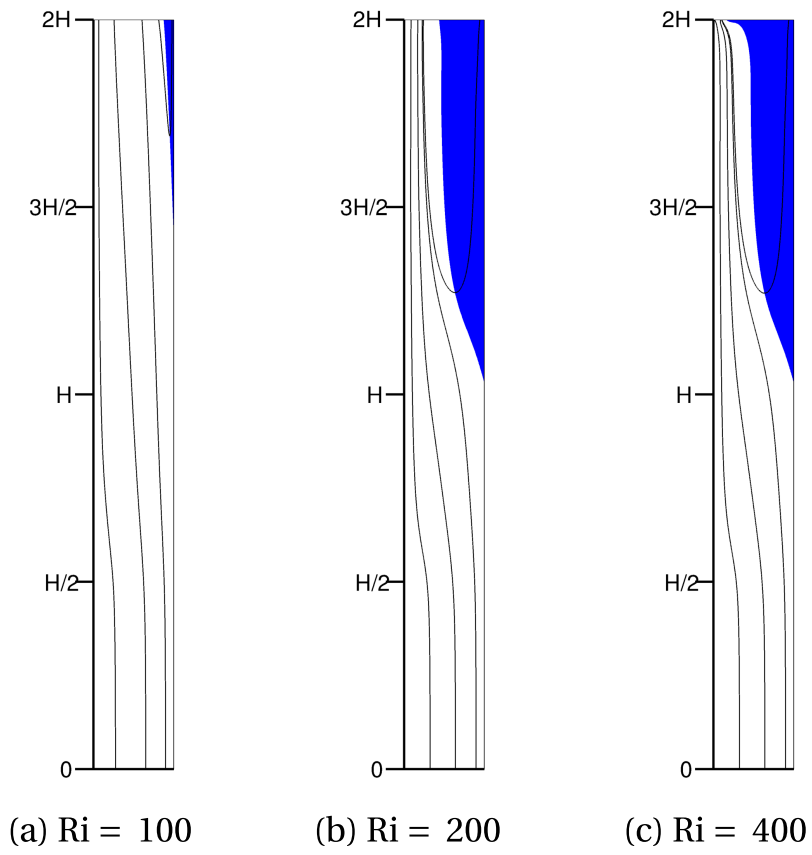


FIGURE 2.13 – Lignes de courant du champ de vitesse sans optimisation. La zone bleue correspond à l'écoulement retour.

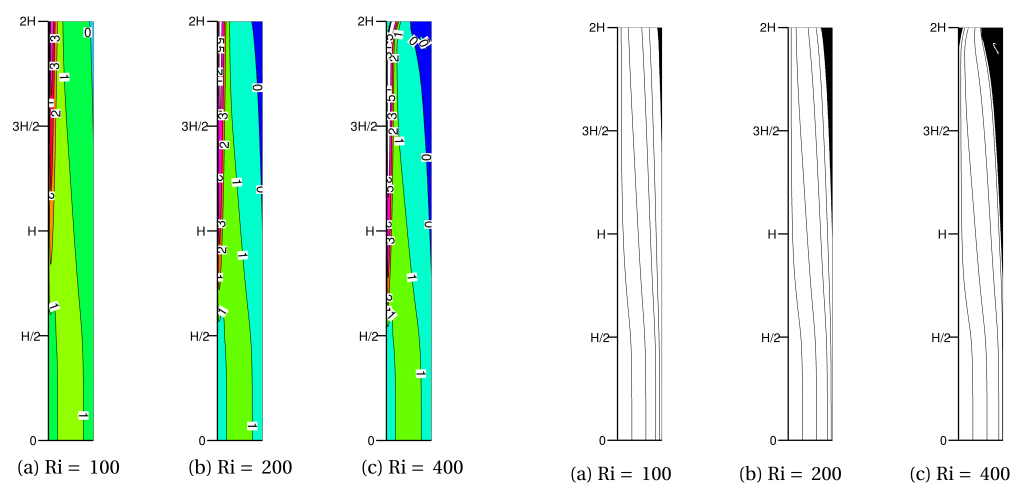


FIGURE 2.14 – Champs de vitesses après optimisation. (Gauche) Composante verticale de la vitesse. (Droite) Lignes de courant.

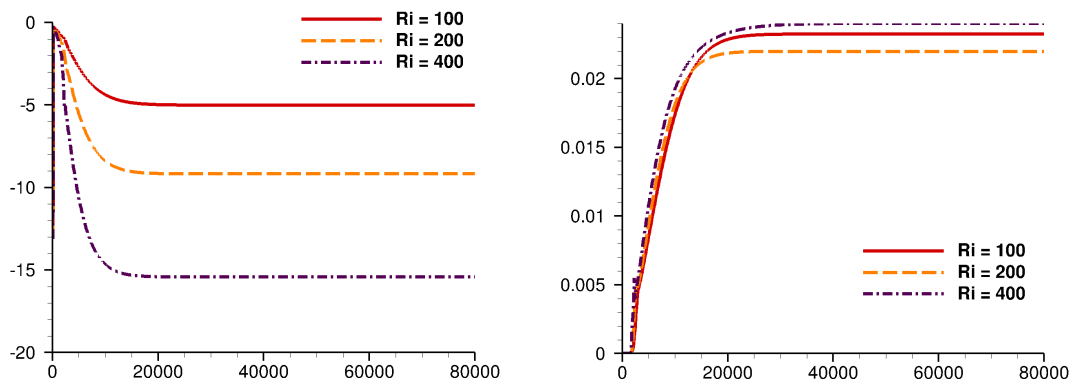


FIGURE 2.15 – (Gauche) J_1 . (Droite) J_2 .

Chapitre 3

Méthodes numériques pour des problèmes de propagation d'onde en régime harmonique

Cette seconde partie présente les travaux effectués en lien avec la résolution numérique de problèmes de propagation d'onde en régime harmonique. Son contenu est basé sur les publications suivantes :

- [Chap3-A1] Cocquet, P. H., & Gander, M. J. (2016). *On the minimal shift in the shifted Laplacian preconditioner for multigrid to work*. In Domain Decomposition Methods in Science and Engineering XXII (pp. 137-145). Springer, Cham.
- [Chap3-A2] Cocquet, P. H., & Gander, M. J. (2017). *How large a shift is needed in the shifted Helmholtz preconditioner for its effective inversion by multigrid?* SIAM Journal on Scientific Computing, 39(2), A438-A478.
- [Chap3-A3] Cocquet, P. H., & Gander, M. J. (2018). *Analysis of the shifted Helmholtz expansion preconditioner for the Helmholtz equation*. In Domain Decomposition Methods in Science and Engineering XXIV, LNCSE, Springer-Verlag.
- [Chap3-A4] Cocquet, P. H., Gander, M. J., & Xiang, X. (2018). *A finite difference method with optimized dispersion correction for the Helmholtz equation*. Domain Decomposition Methods in Science and Engineering XXIV, LNCSE, Springer-Verlag.
- [Chap3-A5] Cocquet, P. H., Gander, M. J., & Xiang, X. (2019). *Dispersion correction for Helmholtz in 1D with piecewise constant wavenumber*. Domain Decomposition Methods in Science and Engineering XXV, LNCSE, Springer-Verlag.
- [Chap3-A6] Cocquet, P. H., Gander, M. J., & Xiang, X. (2021). *Closed form dispersion corrections including a real shifted wavenumber for finite difference discretizations of 2d constant coefficient Helmholtz problems*. SIAM Journal on Scientific Computing, 43(1), A278-A308.

[Chap3-A7] Cocquet, P. H., & Gander, M. J. (2022). *Asymptotic dispersion correction in general finite difference schemes for Helmholtz problems*. Soumis à SIAM Journal on Scientific Computing, Première révision envoyée.

1 Présentation de l'équation de Helmholtz

L'équation de Helmholtz peut être dérivée à partir de l'équation des ondes

$$\frac{1}{c^2} \partial_{tt} v - \Delta v = S(t, x), \quad (3.1)$$

où c est la célérité de l'onde dans le milieu et S un terme source. On suppose que la source est périodique en temps de période $T = 2\pi/\omega$, où ω est une pulsation, et vérifie

$$S(t, x) = e^{i\omega t} f(x),$$

où f est aussi un terme source. On peut alors chercher des solutions de (3.1) sous la forme

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

et dans ce cas u vérifie l'équation de Helmholtz

$$-\Delta u(x) - k^2 u(x) = f(x), \quad (3.2)$$

où $k = \omega/c$ est appelé le nombre d'onde. L'équation (3.2) est un problème modèle modélisant la propagation d'onde en régime harmonique. Hormis la dérivation effectuée ci-dessus, il peut également intervenir par exemple dans les situations suivantes (voir [64, Chapitre 1] ou [62, Chapitre 1]) :

1. En électromagnétisme où

$$k = \omega \sqrt{\epsilon \mu},$$

avec ϵ la permittivité du milieu considéré et μ sa perméabilité.

2. En élasticité linéaire où

$$k = \omega \frac{\rho}{\lambda + 2\mu},$$

avec ρ la masse volumique et λ, μ les coefficients de Lamé du milieu considéré.

3. En mécanique des fluides à partir d'une linéarisation des équations de Saint-Venant qui modélisent la propagation de vagues en eaux peu profondes. Dans ce cas

$$k = \frac{\omega}{\sqrt{gH}},$$

où g est la constante d'accélération et H une hauteur d'eau caractéristique.

Soit maintenant $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) un domaine dont la normale extérieure est notée \mathbf{n} . On suppose que sa frontière $\partial\Omega$ est Lipschitzienne et peut être décomposée comme $\partial\Omega = \overline{\Gamma_D \cup \Gamma_N \cup \Gamma_R}$ où Γ_D, Γ_N et Γ_R sont mutuellement disjoints. L'équation de Helmholtz étudiée dans mes travaux de recherche correspond alors au problème aux limites suivant :

$$\begin{cases} -\Delta u(x) - k^2 u(x) = f(x), & x \in \Omega, \\ u|_{\partial\Omega} = 0, & x \in \Gamma_D, \\ \partial_{\mathbf{n}} u = 0, & x \in \Gamma_N, \\ \partial_{\mathbf{n}} u - iku = g, & x \in \Gamma_R, \end{cases} \quad (3.3)$$

où $f \in L^2(\Omega)$ et $g \in L^2(\Gamma_R)$ sont des termes sources. L'existence et l'unicité d'une solution satisfaisant (3.3) ont été démontrées, par exemple, dans [58, 61, 62]. Toutefois, la résolution numérique de (3.3) pour des grands nombres d'onde est délicate du fait des difficultés suivantes [49, 53] :

- La matrice H_k associée à une discrétisation (ex : différences finies, éléments finis, ...) de (3.3) n'est pas définie positive. De plus, pour $\Gamma_R \neq \emptyset$, bien que H_k soit symétrique ($H_k^T = H_k$), elle n'est pas Hermitienne ($H_k^* \neq H_k$). Ces défauts rendent la résolution des systèmes linéaires issus d'une discrétisation de (3.3) difficile avec des méthodes itératives.
- Les solutions de (3.3) oscillent avec une période proportionnelle à k^{-1} . Pour un maillage dont le pas est noté h , cela entraîne que h doit vérifier $h \sim k^{-1}$ pour approximer précisément les solutions de (3.3). Si N désigne le nombre de degré de liberté du schéma numérique considéré et d la dimension spatiale, alors $N \sim h^{-d}$ et la contrainte précédente entraîne que $N \sim k^d$ indiquant des systèmes linéaires de grande taille.
- Le point précédent entraîne que la taille du maillage h doit vérifier $kh \ll 1$ ou encore que le nombre de points par longueur d'onde $G = 2\pi/(kh)$ soit suffisamment grand. Toutefois, bien qu'avoir un nombre de point par longueur d'onde fixé assure que l'erreur entre u et son interpolant soit bornée uniformément en k , l'erreur entre u et la solution numérique reste croissante en k même pour kh assez petit. Ce phénomène est appelé *effet de pollution* [39, 65, 67] et peut être évité en considérant des maillages encore plus fins (par exemple $h \sim k^{-3/2}$ pour des éléments finis \mathbb{P}_1) engendrant ainsi des systèmes linéaires de très grande taille.

Les travaux de recherche présentés dans la suite de ce mémoire portent tout d'abord sur l'analyse d'un préconditionneur de type *Laplacien shifté* (voir Section 2) utilisé pour accélérer la résolution des systèmes linéaires issus d'une discrétisation de (3.3) puis sur des discrétisations différences finies avec correction de dispersion (voir Section 3) dont l'objectif est de réduire l'effet de pollution.

2 Préconditionneur de type *Laplacien-shifté* pour l'équation de Helmholtz

Dans cette partie, nous allons présenter majoritairement des résultats obtenus dans [Chap3-A2] qui portent sur le preconditionneur appelé *Shifted Laplace Preconditioner (SLP)*. Les résultats issus de [Chap3-A1,Chap3-A3] seront également mentionnés.

La résolution numérique du système linéaire issu d'une discrétisation de l'équation de Helmholtz (3.3) est principalement réalisée avec des méthodes de Krylov (ex : GMRES ou BICGStab) en raison de leur robustesse. Toutefois, sans un bon preconditionneur, leur convergence est lente et se dégrade lorsque k augmente [53].

En notant H_k la matrice issue d'une discrétisation (différences finies ou éléments finis) de l'équation de Helmholtz (3.3), le SLP est défini par $H_{k_s}^{-1}$ avec $k_s \in \mathbb{C}$ donné par

$$k_s^2 = k^2 + i\varepsilon,$$

où ε est appelé le *shift*.

Le SLP a été introduit en 2004 dans [51] et, étant issu d'une discrétisation de l'équation de Helmholtz, il hérite de certaines propriétés du problème continu et peut notamment être utilisé pour des milieux dont le nombre d'onde k est inhomogène. Pour ces raisons, ce preconditionneur a été ensuite intensément utilisé et étudié [52, 41, 70, 60].

Le choix du shift est déterminant pour les propriétés du SLP. En effet, pour un petit shift la matrice H_{k_s} est "proche" de H_k et ainsi le système preconditionné $H_{k_s}^{-1}H_k$ sera "proche" de la matrice identité indiquant alors que H_{k_s} est un bon preconditionneur. Toutefois, dans ce cas, l'inversion de H_{k_s} est difficile car cette dernière hérite des propriétés de H_k . Pour un shift assez grand, l'inversion de H_{k_s} peut être effectuée avec une méthode multigrille [52] qui a l'avantage, pour des EDP elliptiques et coercives, de converger rapidement avec un facteur de convergence indépendant du pas du maillage [59].

Plus précisément, il a été démontré dans [55, Theorem 1.4] que, pour $\varepsilon \sim k$, le SLP était un bon preconditionneur pour H_k dans le sens où le nombre d'itération de GMRES nécessaire pour résoudre le système preconditionné $H_{k_s}^{-1}H_k$ était indépendant du nombre d'onde. D'un autre côté, pour $\varepsilon \sim k^2$, l'inversion du SLP peut être réalisée avec des méthodes multigrilles [52, 70]. Il y a ainsi un compromis entre le fait d'avoir un bon preconditionneur ($\varepsilon \sim k$) et avoir une méthode itérative efficace pour calculer le SLP ($\varepsilon \sim k^2$).

Au vu de l'importance du choix du shift, nous avons étudié à la fois théoriquement et numériquement son influence sur la convergence d'une méthode multigrille. Dans la suite, on commencera par présenter l'algorithme multigrille. On donnera ensuite des résultats dans le cas monodimensionnel pour lequel nous avons un résultat de divergence, lorsque $0 < \varepsilon < k^2$, pour un algorithme à deux niveaux (appelé aussi algorithme ou méthode 2-grilles) utilisant un lisseur de Jacobi (voir Section 2.2). Ensuite, en utilisant un autre lisseur, nous avons obtenu un résultat théorique assurant la convergence d'un

multigrille pour tout $0 < \varepsilon \lesssim k^2$ si un nombre suffisamment grand de pas de lissage est effectué (voir Section 2.3). Enfin, la section 2.4 présentera une analyse d'une extension du SLP appelée *expansion preconditioner*.

2.1 Algorithme multigrille pour l'équation de Helmholtz avec shift complexe

Soit Ω un polygone convexe de \mathbb{R}^d avec $d = 1, 2, 3$. L'équation de Helmholtz avec shift complexe et conditions de Dirichlet homogènes est donnée par :

$$\begin{cases} -\Delta u(x) - (k^2 + i\varepsilon)u(x) = f(x), & x \in \Omega, \\ u|_{\partial\Omega} = 0, \end{cases} \quad (3.4)$$

où $f \in L^2(\Omega)$ est un terme source. La formulation variationnelle de (3.4) est

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ tel que :} \\ a(u, v) := \int_{\Omega} \nabla u \cdot \overline{\nabla v} - (k^2 + i\varepsilon)u\bar{v} dx = \int_{\Omega} f\bar{v} dx, \quad \forall v \in H_0^1(\Omega). \end{cases} \quad (3.5)$$

Considérons maintenant $\{\mathcal{T}_{h_l}\}$ une famille de triangulations régulières de Ω constituées de simplexes de taille caractéristique h_l . L'indice l permettra de définir des relations d'inclusion entre différents maillages de tailles caractéristiques différentes. On note \mathcal{V}_l l'espace éléments finis \mathbb{P}_1 correspondant :

$$\mathcal{V}_l = \{v \in \mathcal{C}(\overline{\Omega}) \mid v|_T \in \mathbb{P}_1 \text{ pour tous } T \in \mathcal{T}_l, v|_{\partial\Omega} = 0\}.$$

Le problème discret est ainsi donné par

$$\begin{cases} \text{Trouver } u_l \in \mathcal{V}_l \text{ tel que :} \\ a(u_l, v_l) = \int_{\Omega} f\bar{v}_l dx, \quad \forall v_l \in \mathcal{V}_l. \end{cases} \quad (3.6)$$

En notant $(\phi_j^l)_{j=1}^{N_l}$ la base nodale de \mathcal{V}_l , nous pouvons définir l'isomorphisme ci-dessous

$$F_l : \mathbb{C}^{N_l} \longrightarrow \mathcal{V}_l, F_l z = \sum_{j=1}^{N_l} z_j \phi_j^l, \quad (3.7)$$

et le problème (3.6) est alors équivalent au système linéaire

$$A_l z_l = b_l, \text{ avec } (A_l)_{i,j} = a(\phi_i^l, \phi_j^l) \text{ et } (b_l)_j = \int_{\Omega} f \phi_j dx, \quad (3.8)$$

et $u_l = F_l z_l$.

Nous présentons maintenant la méthode multigrille [59, 66]. Pour ce faire, la triangulation au niveau l est obtenue, par raffinement de maillage, à partir d'une triangulation plus grossière au niveau $l - 1$. Les espaces éléments finis associés vérifient alors

$$\mathcal{V}_{l-1} \subset \mathcal{V}_l,$$

où les pas de maillage satisfont $h_{l-1} \leq ch_l$ pour $c \geq 2$. Les méthodes multigrilles utilisent des opérateurs de restriction et de prolongation pour communiquer entre les différents niveaux (grilles fine et grossière). Comme $\mathcal{V}_{l-1} \subset \mathcal{V}_l$, l'opérateur identité $I_l : \mathcal{V}_{l-1} \rightarrow \mathcal{V}_l$ est bien défini et

$$\forall u_{l-1} \in \mathcal{V}_{l-1}, (I_l u_{l-1})(x) = \sum_{j=1}^{N_l} u_{l-1}(x_j^l) \phi_j^l(x),$$

où x_j^l sont les noeuds du maillage fin. L'opérateur I_l est l'opérateur de prolongation partant de la grille grossière vers la grille fine et sa représentation matricielle est

$$P_l : \mathbb{C}^{N_{l-1}} \mapsto \mathbb{C}^{N_l}, P_l := F_l^{-1} F_{l-1}. \quad (3.9)$$

Plus précisément, on peut calculer les éléments de sa matrice comme suit

$$\forall \mu \in \{1, \dots, N_{l-1}\}, (I_l \phi_\mu^{l-1})(x) = \sum_{j=1}^{N_l} \phi_\mu^{l-1}(x_j^l) \phi_j^l(x).$$

En suivant [59] l'opérateur de restriction canonique $R_l : \mathbb{C}^{N_l} \mapsto \mathbb{C}^{N_{l-1}}$ est la transposée de l'opérateur de prolongation. Ainsi

$$R_l := P_l^T. \quad (3.10)$$

L'algorithme multigrille [66, 59] utilisé pour résoudre (3.8) est enfin donné par

$$\left\{ \begin{array}{l} \text{function } z_l = \text{MGM}_l(z_l, \mathbf{b}_l) \\ \text{if } l = 0 \text{ then } z_0 = A_0^{-1} \mathbf{b}_0 \text{ else} \\ \quad z_l = S_l^{\nu_1}(z_l, \mathbf{b}_l); \quad \% \text{ pre - smoothing} \\ \quad \mathbf{d}_{l-1} = R_l(\mathbf{b}_l - A_l z_l); \\ \quad \mathbf{e}_{l-1}^0 = \mathbf{0}; \\ \quad \text{for } j = 1 \text{ to } \tau \text{ do} \\ \quad \quad \mathbf{e}_{l-1}^j = \text{MGM}_{l-1}(\mathbf{e}_{l-1}^{j-1}, \mathbf{d}_{l-1}); \\ \quad \text{end} \\ \quad z_l = z_l + P_l \mathbf{e}_{l-1}^\tau; \\ \quad z_l = S_l^{\nu_2}(z_l, \mathbf{b}_l); \quad \% \text{ post - smoothing} \\ \text{end} \end{array} \right. \quad (3.11)$$

Dans (3.11), S_l est appelé le lisseur et correspond à la matrice d'itération d'une méthode itérative (ex : Richardson, Jacobi, ...) utilisée pour résoudre un système linéaire $A_l z_l = \mathbf{b}_l$. La notation $S_l^\nu(z, \mathbf{b})$ indique que ν itérations de cette méthode ont été effectuées. De plus, ν_1 est le nombre de pré-pas de lissage et ν_2 le nombre de post-pas de lissage. Enfin, le "V-cycle" classique est obtenu pour $\tau = 1$ et le "W-cycle" pour $\tau = 2$.

En utilisant [66, Theorem 7.1 p.22] (ou [59] Lemma 7.14), la matrice d'itération de l'algorithme (3.11) est donnée par la relation de récurrence

$$\begin{aligned} C_{\text{MG},0} &= 0, \\ C_{\text{MG},l} &= S_l^{\nu_2} \left(I - P_l \left(I - C_{\text{MG},l-1}^\tau \right) A_{l-1}^{-1} R_l A_l \right) S_l^{\nu_1}. \end{aligned} \quad (3.12)$$

Notons que $C_{MG,l} = C_{MG,l}(\nu_2, \nu_1)$ et

$$\rho(C_{MG,l}(\nu_2, \nu_1)) = \rho(C_{MG,l}(0, \nu_1 + \nu_2)),$$

permettant ainsi de n'étudier que le cas $\nu = \nu_1 + \nu_2$. Dans la suite, la convergence de l'algorithme multigrille sera étudiée en calculant le rayon spectral de $C_{MG,l}$.

2.2 Influence du shift sur la convergence d'un 2-grilles en 1d

Cette partie présente des résultats issus de [Chap3-A2].

Considérons un maillage uniforme de $\Omega = (0, 1)$ avec N_l points intérieurs dont le pas de maillage est $h_l = 1/(N_l + 1)$. La base de \mathcal{V}_l est alors donnée par les fonctions "chapeaux"

$$\phi_j^l(x) = \max\left(0, 1 - \frac{|x - jh_l|}{h_l}\right), \quad j = 1, \dots, N_l.$$

La matrice associée au système (3.6) est la matrice tri-diagonale

$$A_l = \text{tridiag}\left(-\frac{1}{h_l} - \frac{h_l}{6}(k^2 + i\varepsilon), \frac{2}{h_l} - \frac{2h_l}{3}(k^2 + i\varepsilon), -\frac{1}{h_l} - \frac{h_l}{6}(k^2 + i\varepsilon)\right).$$

Le maillage grossier est défini par le maillage uniforme avec $N_{l-1} = (N_l - 1)/2$ points intérieurs et son pas de maillage est $h_{l-1} = 1/(N_{l-1} + 1) = 2h_l$. Le lisseur utilisé est une méthode de Jacobi

$$\mathbf{u}_{m+1} = \mathbf{u}_m + \omega D^{-1}(\mathbf{f} - A_l \mathbf{u}_m),$$

où $D = \text{diag}(A_l)$ et ω est donné par

$$\omega = \frac{12 - 4h^2(k^2 + i\varepsilon)}{18 - 3h^2(k^2 + i\varepsilon)}. \quad (3.13)$$

Notons que la matrice d'itération est alors donnée par $S = I_d - \omega D^{-1}A_l$. La matrice d'itération d'une méthode à deux niveaux, appelé également algorithme 2-grilles, est notée T et son expression est donnée dans (3.12)

$$T = S_l^{\gamma_2} \left(I_d - P_l A_{l-1}^{-1} R_l A_l \right) S_l^{\gamma_1},$$

où les matrices des opérateurs de restriction R_l et de prolongation P_l sont définies respectivement par (3.10) et (3.9). En utilisant une diagonalisation par bloc de la matrice d'itération T , nous avons obtenu des résultats de convergence pour la méthode à deux niveaux en fonction des valeurs du shift ε .

Théorème 3.1 *Soit $\nu = \nu_1 + \nu_2$ le nombre total de pas de lissage. On suppose que $\varepsilon = Ck^{2-\delta}$ pour $0 < \delta \leq 2$. Alors, pour tout k et h tels que*

$$kh = \sqrt{6} + o(1) \text{ pour } h \rightarrow 0,$$

on a

$$\rho(T) \geq \left(\frac{2 \cdot 6^{\delta/2}}{3 C h^{\delta}} \right)^{\nu} + o\left(\frac{1}{h^{\delta \nu}} \right),$$

et ainsi l'algorithme 2-grilles diverge.

Pour $\varepsilon = Ck^2$, si $C > 0$ est une constante suffisamment grande alors l'algorithme 2-grilles converge.

Afin d'illustrer numériquement le Théorème 3.1, considérons une grille uniforme de $\Omega = (0, 1)$ avec $N_1 = 2^{10} - 1$ points intérieurs ainsi que les 3 nombres d'ondes ci-dessous

$$k_1 = \frac{\sqrt{6}}{h_1} = 2508.3, \quad k_2 = \frac{1}{2h_1} = 512, \quad k_3 = h_1^{-2/3} = 101.6.$$

Le nombre d'onde k_1 vérifie $k_1 h_1 = \sqrt{6}$ ce qui est proche des hypothèses du Théorème 3.1 pour lesquelles l'algorithme 2-grilles diverge, k_2 vérifie $k_2 h_1 = 0.5$ correspondant à considérer environ 12 points par longueur d'onde et k_3 est choisi de sorte qu'il n'y ait pas d'effet de pollution car, dans ce cas, $k^3 h^2$ est assez petit [39].

ε	1	\sqrt{k}	k	$k^{3/2}$	k^2
$k = k_1, \nu = 1$	$4.1943 \cdot 10^6$	$8.3748 \cdot 10^4$	$1.6722 \cdot 10^3$	33.3906	0.7454
$\nu = 2$	$1.7592 \cdot 10^{13}$	$7.0137 \cdot 10^9$	$2.7962 \cdot 10^6$	$1.1149 \cdot 10^3$	0.5556
$\nu = 4$	$3.0949 \cdot 10^{26}$	$4.9191 \cdot 10^{19}$	$7.8188 \cdot 10^{12}$	$1.2433 \cdot 10^6$	0.3221
$k = k_2, \nu = 1$	40.1940	40.1460	26.9411	1.6269	0.3623
$\nu = 2$	29.4072	29.3721	19.7105	1.1755	0.1623
$\nu = 4$	31.8974	31.8593	21.3795	1.2722	0.1141
$k = k_3, \nu = 1$	0.5093	0.5089	0.4732	0.3454	0.3344
$\nu = 2$	0.1626	0.1625	0.1605	0.1235	0.1138
$\nu = 4$	0.1202	0.1200	0.1052	0.0633	0.0633

TABLEAU 3.1 – Rayon spectral de T pour $k = k_1, k_2, k_3$.

Le rayon spectral de T est donné dans le Tableau 3.1 pour $\nu = 1, 2, 4$. Comme annoncé dans le Théorème 3.1, l'algorithme 2-grilles diverge lorsque kh est proche de $\sqrt{6}$ sauf si $\varepsilon = k^2$. La situation est similaire pour $k = k_2$. Notons finalement que l'algorithme 2-grilles converge pour chaque ε considéré pour $k = k_3$ correspondant au nombre d'onde pour lequel il n'y a pas d'effet de pollution. Toutefois, cette observation n'implique pas qu'une méthode multigrille convergera dans la mesure où, pour certains niveaux l , des combinaisons de k, h entraînant la divergence (ex : $k = k_1, k = k_2$) pourraient apparaître.

Remarque 3.2 Des résultats similaires au Théorème 3.1 ont été également obtenus pour une discrétisation différences finies dans [Chap3-A1].

2.3 Résultat de convergence pour un W-cycle

Cette partie présente des résultats issus de [Chap3-A2].

L'objectif de cette section est de présenter une extension du Théorème 3.1 en dimensions supérieures. Notons que la démonstration du Théorème 3.1 a été réalisée en diagonalisant la matrice d'itération d'un 2-grilles. Toutefois, ce type de démonstration est difficile à étendre en dimension supérieure notamment pour des domaines Ω non-rectangulaires et des maillages non-uniformes.

L'idée principale de la démonstration de la convergence d'un W-cycle est de borner directement la norme Euclidienne de sa matrice d'itération (3.12) en faisant apparaître une relation de récurrence entre deux niveaux successifs. Un résultat général de convergence d'un W-cycle est donné ci-dessous.

Théorème 3.3 (Theorem 7.20,[66]) *On considère un W-cycle, soit $\tau = 2$. Supposons qu'il existe des constantes positives C_A, C_S et une fonction positive décroissante $\nu \mapsto g(\nu)$ vérifiant $\lim_{\nu \rightarrow +\infty} g(\nu) = 0$, telles que¹*

$$\begin{aligned} \|A_l^{-1} - P_l A_{l-1}^{-1} R_l\|_2 &\leq C_A \|A_l\|_2^{-1}, & \text{(The approximation property)} \\ \|A_l S^\nu\|_2 &\leq g(\nu) \|A_l\|_2, \nu \geq 1, & \text{(The smoothing property)} \\ \|S^\nu\|_2 &\leq C_S, \nu \geq 1. & \text{(The multigrid contraction number)} \end{aligned}$$

Alors pour tout $\xi \in (0, 1)$ il existe un ν_ξ tel que quelque soit $\nu \geq \nu_\xi$, on ait

$$\|C_{MG,l}\|_2 \leq \xi.$$

Nous avons ensuite démontré que les hypothèses du Théorème 3.3 étaient satisfaites pour une discrétisation \mathbb{P}_1 de l'équation (3.4).

Quelques notations sont tout d'abord nécessaires avant de donner ce résultat. Notons $C_{\text{cont}}, C_{\text{coer}}$ les constantes de continuité et de coercivité de la forme bilinéaire a définie par (3.5). On a

$$|a(\mathbf{u}, \mathbf{v})| \leq C_{\text{cont}} \|\nabla \mathbf{u}\|_{L^2(\Omega)} \|\nabla \mathbf{v}\|_{L^2(\Omega)}, \quad |a(\mathbf{u}, \mathbf{u})| \geq C_{\text{coer}} \|\nabla \mathbf{u}\|_{L^2(\Omega)}^2.$$

Le domaine Ω étant convexe, les résultats usuels de régularité elliptique assurent que, si $f \in L^2(\Omega)$, alors la solution de (3.4) est dans $H^2(\Omega)$ et on note C_{H^2} , la constante positive pour laquelle

$$\|\mathbf{u}\|_{H^2(\Omega)} \leq C_{H^2} \|f\|_{L^2(\Omega)}.$$

La dépendance en k, ε de ces trois constantes étant importante pour étudier la convergence de la méthode multigrille en fonction de ces paramètres, elle est maintenant donnée explicitement.

1. On gardera la dénomination anglaise des inégalités ci-après.

Théorème 3.4 Les constantes C_{cont} , C_{coer} et C_{H^2} sont données par

$$\begin{aligned} C_{\text{cont}} &= (1 + C_{\text{P}}^2 |k^2 + i\varepsilon|), \\ C_{\text{coer}} &= \min\{\Re\gamma, \varepsilon\Im\gamma - k^2\Re\gamma\}, \\ C_{\text{H}^2} &= C(\Omega) \left(1 + \frac{|k^2 + i\varepsilon|}{\varepsilon} + \frac{C_{\text{P}}^2}{C_{\text{coer}}} \right), \end{aligned}$$

où C_{P} est la constante de l'inégalité de Poincaré ($\forall \mathbf{u} \in \text{H}_0^1(\Omega) : \|\mathbf{u}\|_{\text{L}^2(\Omega)} \leq C_{\text{P}} \|\nabla \mathbf{u}\|_{\text{L}^2(\Omega)}$), $\gamma \in \mathbb{C}$ vérifie $|\gamma| = 1$, $0 < \Re\gamma \leq \varepsilon/|k^2 + i\varepsilon|$ et $\Im\gamma > 0$ et $C(\Omega) > 0$ ne dépend que du domaine Ω .

Nous pouvons maintenant donner le résultat de convergence pour l'algorithme multigrille appliqué au problème de Helmholtz avec shift complexe.

Théorème 3.5 Ci-dessous $C(\Omega) > 0$ désigne une constante ne dépendant que du domaine Ω . On suppose que Ω est convexe. Alors, il existe $C(\Omega) > 0$ telle que :

- The approximation property :

$$\|A_1^{-1} - P_1 A_{1-1}^{-1} R_1\|_2 \leq C(\Omega) C_{\text{reg}} C_{\text{cont}} \|A_1\|_2^{-1},$$

avec

$$C_{\text{reg}} = \left(\frac{C_{\text{cont}}}{C_{\text{coer}}} \right)^2 C_{\text{cont}} C_{\text{H}^2}^2 \|f\|_{\text{L}^2(\Omega)},$$

- The multigrid contraction number : Pour tout $\nu \geq 1$, on a

$$\|S^\nu\|_2 \leq C(\Omega) \sqrt{\frac{C_{\text{cont}}}{C_{\text{coer}}}}, \quad (3.14)$$

où $S = I_d - \frac{\omega}{2} \gamma |D|^{-1} A_1$ est un lisseur de Jacobi modifié avec D la partie diagonale de A_1 .

- The smoothing property : Pour tout $\omega \in \left(0, \frac{2C(\Omega)C_{\text{coer}}^2}{C_{\text{cont}}^2}\right)$, on a

$$\|A_1 S^\nu\|_2 \leq 2C(\Omega) \sqrt{\frac{2}{\pi\omega^2\nu}} \sqrt{\frac{C_{\text{cont}}}{C_{\text{coer}}}} \|A_1\|_2. \quad (3.15)$$

Remarque 3.6 Des résultats similaires aux Théorèmes 3.4 et 3.5 ont été obtenus pour des conditions limites de Robin sur $\partial\Omega$ (voir [Section 5, Chap3-A2]).

Le Théorème 3.5 combiné au Théorème 3.3 montre que l'algorithme multigrille (3.11) appliqué à (3.4) avec le lisseur de Jacobi modifié S converge pour tout k et tout ε si un nombre suffisamment grand de pas de lissage est effectué. Le facteur de convergence est de plus décroissant en fonction du nombre de pas de lissage ν . Toutefois, le nombre de pas de lissage minimal ν_{min} pour lequel on ait

$$\forall \nu \geq \nu_{\text{min}} : \|C_{\text{MG},l}\|_2 < 1,$$

dépend de k et de ε . De même, le paramètre ω intervenant dans le lisseur de Jacobi modifié dépend de k , ε et du nombre complexe γ introduit dans le Théorème 3.4. En exprimant les constantes de continuité C_{cont} et de coercivité C_{coer} lorsque $\varepsilon < O(k^2)$, on a

$$\omega = \left(\frac{\Re\gamma}{1 + |k^2 + i\varepsilon|} \right)^2 \text{ où } \gamma = \frac{\varepsilon + ik^2}{|\varepsilon + ik^2|} \text{ et } \varepsilon < \text{cste} \times k^2.$$

Des résultats numériques calculant le rayon spectral de la matrice d'itération d'un 2-grilles sont donnés dans le Tableau 3.2. On peut voir que $\rho(T)$ est très proche de 1 pour tous les

ν	ε	$k = 10$	$k = 20$	$k = 40$	$k = 80$	$k = 160$
1	1	0.99999999975240	1.000000000000025	1.000000000000046	1.000000000001082	1.000000000000053
3	1	0.99999999925725	0.99999999999980	1.000000000000049	1.000000000001164	1.000000000000047
1	k	0.999999975841890	0.99999999801596	0.99999999998355	1.000000000000026	1.000000000000046
3	k	0.999999927525641	0.99999999404773	0.99999999994992	0.99999999999972	1.000000000000064
1	$k^{3/2}$	0.999999382351815	0.99999984191422	0.99999999600494	0.99999999983235	0.99999999999554
3	$k^{3/2}$	0.999998147056534	0.99999952574267	0.99999998801486	0.99999999949603	0.99999999998608

TABLEAU 3.2 – Rayon spectral de la matrice d'itération T d'un 2-grilles en 2d avec conditions de Neumann homogène. $\Omega = (0, 1)^2$, le maillage fin est composé de 49×49 points, le maillage grossier est composé de 25×25 points.

cas considérés. Cela s'explique par le fait que le paramètre $\omega = O(k^{-4})$ et donc devient très petit numériquement entraînant que le lisseur S est très proche de $S \approx I_d$. Ainsi, la matrice T devient très proche de l'opérateur grille grossière $T \approx (I - P_1 A_{1-1}^{-1} R_1 A_1)$ qui est une projection. Pour certaines valeurs des paramètres considérés, le rayon spectral de T est bien une fonction décroissante du nombre de pas de lissage entraînant la convergence de cet algorithme si le nombre de pas de lissage devient (très!) grand. Toutefois, pour certains jeux de paramètres, $\rho(T) > 1$ et donc ce algorithme 2-grilles diverge.

Pour $\varepsilon = \text{cste} \times k^2$, le paramètre ω devient

$$\omega = \left(\frac{-(\Re\gamma)k^2 + (\Im\gamma)\varepsilon}{1 + |k^2 + i\varepsilon|} \right)^2 \text{ où } \gamma = \frac{1}{2} + i\frac{\sqrt{3}}{2} \text{ et } \varepsilon = \text{cste} \times k^2.$$

Dans ce cas, lorsque $\varepsilon = \text{cste} \times k^2$ pour une constante suffisamment grande, le ratio $C_{\text{cont}}/C_{\text{coer}}$ devient indépendant de k et le paramètre ω n'est pas numériquement 0 lorsque k augmente (voir [Theorem 4.7, Chap3-A2]). Le Théorème 3.4 assurant la convergence du multigrille avec le lisseur de Jacobi modifié sera donc valable également numériquement contrairement au cas $\varepsilon < \text{cste} \times k^2$. Le rayon spectral de la matrice d'itération pour un 2-grilles et un W -cycle à 3-niveaux sont donnés respectivement dans les Tableaux 3.3 et 3.4. On peut voir que le rayon spectral d'un 2-grilles et d'un W -cycle à 3-niveaux est strictement plus petit que 1 et décroît lorsque le nombre de pas de lissage augmente. Cela est conforme avec les conclusions du Théorème 3.5.

2.4 Extension du Shifted Laplace Preconditioner : the expansion preconditioner

Cette partie présente des résultats issus de [Chap3-A3].

k	10	20	40	80	160
kh	0.1000	0.2000	0.4000	0.8000	1.6000
$\nu = 1$	0.9764	0.9757	0.9736	0.9648	0.9306
$\nu = 3$	0.9309	0.9289	0.9228	0.8982	0.8059
$\nu = 5$	0.8875	0.8843	0.8747	0.8361	0.6980

TABLEAU 3.3 – Rayon spectral de la matrice d’itération d’un 2-grilles en 2d avec condition de Neumann homogène et $\varepsilon = 1.5k^2$. $\Omega = (0, 1)^2$, le maillage fin est composé de $101 \times 101 = 10201$ points, le maillage grossier est composé de 51×51 points.

k	10	20	40	80	160
kh	0.1000	0.2000	0.4000	0.8000	1.6000
$\nu = 1$	0.9764	0.9757	0.9736	0.9648	0.9306
$\nu = 3$	0.9309	0.9289	0.9229	0.8982	0.8058
$\nu = 5$	0.8876	0.8844	0.8748	0.8362	0.6978

TABLEAU 3.4 – Rayon spectral de la matrice d’itération d’un W-cycle à 3-niveaux en 2d avec conditions de Neumann homogène et $\varepsilon = 1.5k^2$. $\Omega = (0, 1)^2$, le maillage fin est composé de $101 \times 101 = 10201$ points, le maillage grossier est composé de 51×51 points.

On présente maintenant des résultats concernant une extension du *Shifted Laplace Preconditioner* (SLP) appelée *Expansion preconditioner* qui a été introduite dans [46, 47]. En notant H_k la matrice issue d’une discrétisation de l’équation de Helmholtz, il est possible de représenter H_k^{-1} par une série faisant intervenir uniquement $H_{k+i\varepsilon}^{-1}$ et le *Expansion preconditioner* est ainsi défini en tronquant cette série à l’ordre voulu. On retrouve en particulier le SLP usuel en ne conservant que le premier terme de cette série.

Considérons l’équation de Helmholtz avec shift complexe et conditions de Robin sur $\partial\Omega$

$$\begin{cases} -\Delta u(x) - (k^2 + i\varepsilon)u(x) = f(x), & x \in \Omega, \\ \partial_n u - i\eta u = 0, & \text{sur } \partial\Omega, \end{cases} \quad (3.16)$$

où \mathbf{n} est la normale unitaire sortante de $\partial\Omega$ et $\eta > 0$ est un paramètre. La formulation variationnelle de (3.16) est

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ tel que pour tout } v \in H^1(\Omega) : \\ a_\varepsilon(u, v) := \int_\Omega \nabla u \cdot \overline{\nabla v} - (k^2 + i\varepsilon)u\bar{v} dx - i\eta \int_{\partial\Omega} u\bar{v} d\sigma = \int_\Omega f\bar{v} dx. \end{cases} \quad (3.17)$$

En considérant une triangulation uniforme de Ω et des éléments finis \mathbb{P}_1 , le problème discret est

$$\begin{cases} \text{Trouver } u_h \in \mathcal{V}_h \text{ tel que :} \\ a_\varepsilon(u_h, v_h) = \int_\Omega f\bar{v}_h dx, \quad \forall v_h \in \mathcal{V}_h. \end{cases} \quad (3.18)$$

Le problème (3.18) est équivalent au système linéaire suivant

$$A_\varepsilon \mathbf{w}_h = \mathbf{b}_h$$

où

$$A_\varepsilon = K - (k^2 + i\varepsilon)M - i\eta N.$$

Ci-dessus, K est la matrice de rigidité, M la matrice de masse et N la matrice de masse associée à $\partial\Omega$. Si ϕ_j , pour $j = 1, \dots, N$ sont les fonctions de base de \mathcal{V}_h , ces matrices sont

$$K = \left(\int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx \right)_{i,j}, \quad M = \left(\int_{\Omega} \phi_i \phi_j \, dx \right)_{i,j}, \quad N = \left(\int_{\partial\Omega} \phi_i \phi_j \, dx \right)_{i,j}.$$

On note A_0 la matrice associée à une discrétisation \mathbb{P}_1 de l'équation de Helmholtz sans shift ($\varepsilon = 0$).

La première idée du *expansion preconditioner* est d'exprimer A_0^{-1} uniquement à partir d'inverse de A_ε . En utilisant que $A_0 = A_\varepsilon + i\varepsilon M$ et une série de Neumann, on obtient une expression pour l'inverse de A_0 faisant intervenir uniquement des inverses A_ε .

Théorème 3.7 *L'inverse de A_0 s'exprime comme suit*

$$A_0^{-1} = \left(\sum_{j \geq 0} (-i\varepsilon)^j (A_\varepsilon^{-1} M)^j \right) A_\varepsilon^{-1},$$

où la série converge pour la norme matricielle subordonnée à la norme $\|\mathbf{x}\|_M = \sqrt{\langle M\mathbf{x}, \mathbf{x} \rangle}$.

En suivant alors [46], le *expansion preconditioner* d'ordre $n \in \mathbb{N}^+$ est défini en tronquant la série de Neumann donnée dans le Théorème 3.7,

$$EX(n) = \left(\sum_{j=0}^{n-1} (-i\varepsilon)^j (A_\varepsilon^{-1} M)^{j+1} \right) M^{-1} = \left(\sum_{j=0}^{n-1} (-i\varepsilon)^j (A_\varepsilon^{-1} M)^j \right) A_\varepsilon^{-1}. \quad (3.19)$$

Le système de Helmholtz préconditionné est ainsi donné par

$$EX(n)A_0 \mathbf{w}_1 = EX(n) \mathbf{b}_1. \quad (3.20)$$

Notons que, lorsque $n = 1$, $EX(1)$ est exactement le SLP. Pour étudier la convergence de GMRES utilisé sur le système préconditionné avec le *expansion preconditioner* (3.20), nous utilisons les estimations d'Elman [55, Theorem 1.8]. Pour un système $B\mathbf{w} = \mathbf{b}$ résolu avec GMRES, on note \mathbf{r}_m le résidu de GMRES après m itérations. En supposant que $\|I_d - B\|_2 \leq \sigma < 1$, on a

$$\frac{\|\mathbf{r}_m\|_2}{\|\mathbf{r}_0\|_2} \leq \left(\frac{2\sqrt{\sigma}}{(1+\sigma)^2} \right)^m. \quad (3.21)$$

Il reste donc à calculer σ dans le cas où GMRES est utilisé pour résoudre (3.20) donc pour $B = EX(n)A_0$. Le résultat ci-dessous a été obtenu.

Théorème 3.8 Pour tous $\varepsilon > 0$, $\eta > 0$, $h > 0$ et $n \in \mathbb{N}^+$, on a les estimations suivantes

$$\begin{aligned} \mathcal{N}(\mathbb{I}_d - \text{EX}(1)A_0) &\leq \varepsilon \mathcal{N}(A_\varepsilon^{-1}M), \\ \forall n \geq 1, \mathcal{N}(\mathbb{I}_d - \text{EX}(n)A_0) &\leq \frac{1 + \varepsilon \mathcal{N}(A_\varepsilon^{-1}M)}{1 - \varepsilon \mathcal{N}(A_\varepsilon^{-1}M)} (\varepsilon \mathcal{N}(A_\varepsilon^{-1}M))^n, \end{aligned}$$

où $\mathcal{N}(B)$ est n'importe quelle norme matricielle ou bien $\rho(B)$.

Le Théorème 3.8 et (3.19) entraînent que la convergence de GMRES appliqué au système préconditionné (3.20) est majoritairement impactée par le terme $\varepsilon \mathcal{N}(A_\varepsilon^{-1}M)$. En particulier, si ce dernier vérifie $\varepsilon \mathcal{N}(A_\varepsilon^{-1}M) < 1$ alors la convergence de GMRES pourra être accélérée en augmentant le nombre n de termes inclus dans le *expansion preconditioner*.

Théorème 3.9 Supposons que les hypothèses (H1) ou bien (H2) soient vérifiées :

$$(H1) \eta \sim k \text{ et } k^3 h^2 \leq C_0 \text{ pour } C_0 \text{ assez petit.}$$

$$(H2) \eta \lesssim k, k \geq k_0 \text{ pour un certain } k_0 > 0 \text{ et } kh \sqrt{|k^2 - \varepsilon|} \leq C_0 \text{ avec } C_0 \text{ assez petit.}$$

Alors il existe $C_2 > 0$ qui dépend uniquement de Ω tel que pour tout $\varepsilon > 0$ avec $\varepsilon C_2 < k$, on a

$$\forall n \geq 1, \mathcal{N}(\mathbb{I}_d - \text{EX}(n)A_0) \leq \left(\frac{C_2 \varepsilon}{k} \right)^n \frac{k + C_2 \varepsilon}{k - C_2 \varepsilon},$$

où $\mathcal{N}(\cdot) = \rho(\cdot)$ si (H1) est satisfaite et $\mathcal{N}(\cdot) = \|\cdot\|_2$ si (H2) est vérifiée.

Concernant les hypothèses du Théorème 3.9 : (H1) permet de ne plus avoir d'erreur de pollution (voir [50]) tandis que (H2) est une contrainte provenant de [55, Lemma 3.5] qui assure la quasi-optimalité (l'optimalité pour des maillages assez fins) de l'approximation éléments finis. Le Théorème 3.9 est une extension directe du [55, Theorem 1.4] au cas du *expansion preconditioner*. Il entraîne que, lorsque $\varepsilon \lesssim k$, la vitesse de convergence de GMRES pour résoudre le système préconditionné (3.20) devient indépendante du nombre d'onde et peut être accélérée en augmentant le nombre de terme dans l'*expansion preconditioner*.

Nous présentons maintenant des résultats numériques donnant le nombre d'itération de GMRES pour résoudre (3.20) lorsque $\varepsilon = k, k^{3/2}, k^2$ et $n = 1, 2, 3$. En effet, bien que le Théorème 3.9 assure la convergence de GMRES, indépendamment de k , le calcul de $\text{EX}(n)$ nécessite des inversions de A_ε qui peuvent être difficiles à réaliser (voir Section 2.3).

On choisit $\Omega = (0, 1)^2$, $\eta = k$ et $f = 1$. On discrétise (3.18) avec des éléments \mathbb{P}_1 . Le pas de maillage h est choisit de sorte que $kh^{3/2} = 1$ et ainsi les hypothèses (H1) du Théorème 3.9 sont vérifiées. Le Tableau 3.5 donne le nombre d'itération de GMRES appliqué au système préconditionné (3.20) permettant d'atteindre un résidu relatif de $1e - 6$.

Notons que pour $n = 1, 2, 3$, le nombre d'itérations est constant pour $\varepsilon = k$ et décroît très légèrement lorsque n augmente ce qui est conforme aux résultats du Théorème 3.9. Lorsque $\varepsilon = k^{3/2}$, le nombre d'itération semble constant pour $n = 2, 3$ et augmente

	n = 1			n = 2			n = 3		
ϵ	k	$k^{3/2}$	k^2	k	$k^{3/2}$	k^2	k	$k^{3/2}$	k^2
k = 5	5	6	8	5	6	9	5	7	9
k = 10	6	8	13	5	8	12	5	7	13
k = 20	6	11	24	5	8	21	4	7	20

TABLEAU 3.5 – Nombre d’itérations de GMRES pour le problème de Helmholtz préconditionné avec le *expansion preconditioner*.

linéairement avec k pour $\epsilon = k^2$. Ces résultats numériques semblent indiquer qu’un nombre assez grand de termes dans $EX(n)$ permet de définir un bon préconditionneur pour des shifts plus petits que $\epsilon = k^2$. Ce progrès reste toutefois assez coûteux puisqu’il est nécessaire d’inverser plusieurs fois la matrice A_ϵ pour calculer $EX(n)$.

3 Schémas différences finies avec correction de dispersion

Les ondes planes sont les fonctions $e^{ikx \cdot \theta}$ avec $\theta \in \mathcal{S}^{d-1}$ et sont des solutions de l'équation de Helmholtz homogène

$$-\Delta u(x) - k^2 u(x) = 0.$$

L'erreur de dispersion est le fait que les ondes planes discrètes $e^{ik_d x \cdot \theta}$ satisfaisant l'équation de Helmholtz homogène discrétisée sont différentes des ondes planes continues. En effet, le nombre d'onde discret k_d dépend de θ et du pas de maillage h et on a généralement

$$k_d(\theta, h) \neq k.$$

L'erreur de dispersion est en partie responsable de l'effet de pollution [39, 65, 67] qui indique qu'avoir un nombre de points par longueur d'onde assez grand (ou de manière équivalente avoir kh assez petit) ne permet pas d'avoir une erreur relative bornée uniformément en (kh) .

Les schémas avec correction de dispersion visent à réduire l'erreur de dispersion définie par $|k_d(\theta, h) - k|$. Ces derniers sont principalement développés pour les différences finies en considérant des schémas d'approximation contenant des paramètres libres qui sont ensuite optimisés numériquement afin de minimiser l'erreur de dispersion [42, 43, 48, 72, 40, 68].

Dans cette section, on présente des résultats obtenus dans les publications [Chap3-A4, Chap3-A5, Chap3-A6, Chap3-A7]. Tous ces travaux portent sur la construction de schémas différences finies avec une correction de dispersion effectuée sans utiliser une optimisation numérique. Hormis le fait que nous avons des formules explicites pour les paramètres libres intervenant dans les schémas, nos résultats sont également basés sur l'introduction d'un nombre d'onde *shifté*, noté \hat{k} , dans le schéma. Ce \hat{k} permet d'introduire un paramètre supplémentaire afin de réduire davantage l'erreur de dispersion et permettra également de réduire l'erreur de dispersion pour des schémas usuels (ex : schéma différences finies à 5-points) qui ne contiennent pas naturellement de paramètres. Notre approche s'appelle *correction de dispersion asymptotique* car les formules explicites sont obtenues en minimisant l'erreur de dispersion lorsque $h \rightarrow 0$ (ou $G \rightarrow +\infty$ où $G = 2\pi/(kh)$ est le nombre de points par longueur d'onde).

Dans la suite de ce chapitre, les résultats suivants seront présentés en détail :

1. On montrera dans la Section 3.1 comment déterminer, pour un problème monodimensionnel, le nombre d'onde shifté \hat{k} pour lequel un schéma standard à 3-points ne présente plus d'erreur de dispersion.
2. Dans la Section 3.2, un schéma différences finies à 9-points avec des paramètres libres sera considéré. Les paramètres et le shift \hat{k} seront déterminés explicitement afin de réduire l'erreur de dispersion pour un nombre de points par longueur d'onde assez grand.

3. La Section 3.3 présentera une généralisation de l'approche décrite dans la Section 3.2 pour des schémas différences finis généraux.
4. Enfin, dans la Section 3.4, des schémas différences finies avec correction de dispersion seront utilisés dans un algorithme multigrille et des résultats numériques de convergence seront présentés

3.1 Correction de dispersion en 1d

Cette section présente des résultats issus de [Chap3-A5,Chap3-A7].

L'équation de Helmholtz dans $\Omega = (0, 1)$ avec conditions de Dirichlet homogène est donnée par

$$\begin{cases} -u''(x) - k^2 u(x) = f(x) & \text{dans } (0, 1), \\ u(x) = 0, & x \in \{0, 1\}, \end{cases} \quad (3.22)$$

où f est un terme source. On suppose que $k^2 \notin \pi\mathbb{N}$ de sorte que (3.22) admette une unique solution.

On considère une grille uniforme $\{x_j\}_{j=1}^n = \{j/(n+1)\}_{j=1}^n$ de $(0, 1)$ composée de $n \in \mathbb{N}^+$ points intérieurs dont le pas de maillage est $h = 1/(n+1)$. En utilisant un schéma à 3-points pour la dérivée seconde, la discrétisation de (3.22) est

$$-\frac{1}{h^2} (u_{i-1} - 2u_i + u_{i+1}) - \tilde{k}^2 u_i = f(x_i), \quad i = 1 \cdots n, \quad (3.23)$$

où u_i est une approximation de $u(x_i)$. Le nombre d'onde \tilde{k} est $\tilde{k} = k$ sans correction de dispersion ou $\tilde{k} = \hat{k}$ lorsque la correction de dispersion est utilisée. En supposant que $f = 0$, en négligeant les bords pour se ramener à un domaine infini et en insérant $u_j := e^{ik_d x_j}$ dans (3.23), le nombre d'onde discret vérifie

$$\cos(k_d h) = 1 - \frac{\tilde{k}^2 h^2}{2}.$$

Notons que $k_d \neq k$ ce qui correspond à l'erreur de dispersion. En cherchant maintenant \hat{k} tel que $k_d = k$, on obtient

$$\hat{k} = \sqrt{\frac{2}{h^2} (1 - \cos(kh))},$$

ce qui correspond au nombre d'onde shifté obtenu dans [53]. Pour ce choix de \hat{k} , le nombre d'onde discret associé au schéma (3.23) vérifie $k_d = k$ et ainsi ce schéma n'a pas d'erreur de dispersion.

On étudie l'influence de la correction de dispersion, c'est à dire du choix $\tilde{k} = \hat{k}$ dans le schéma (3.23), sur l'erreur dans le résultat suivant.

Théorème 3.10 *Supposons que le terme source f vérifie*

$$\|f\|_{L^\infty(0,1)} \lesssim 1, \quad \|f''\|_{L^\infty(0,1)} \lesssim k^2,$$

où la notation \lesssim signifie que les constantes n'apparaissant pas sont indépendantes de k et h . On note $\mathbf{e}_{\tilde{k}} := (u_i - u(x_i))_{i=1}^n$ l'erreur entre la solution de (3.22) et son approximation vérifiant (3.23) pour $\tilde{k} \in \{k, \hat{k}\}$.

- Sans correction de dispersion : Si $\tilde{k} = k$, $kh < 2$ et $\frac{4}{h^2} \sin\left(\frac{j\pi h}{2}\right)^2 - k^2 \neq 0$ pour $j = 1, \dots, n$ (cette hypothèse assure que le problème discret est inversible) alors l'erreur vérifie

$$\|\mathbf{e}_k\|_\infty \lesssim \frac{k^2 h^3}{|\sin(\theta)| |\sin(\theta/h)|} \left(1 + k \left(1 + \frac{1}{|\sin(k)|} \right) \right),$$

où θ vérifie $\cos(\theta) = 1 - (kh)^2/2$.

- Avec correction de dispersion : Si $\tilde{k} = \hat{k}$, $kh \notin \pi\mathbb{N}$ et $k \notin \pi\mathbb{N}$, alors l'erreur vérifie

$$\|\mathbf{e}_{\hat{k}}\|_\infty \lesssim \frac{kh}{|\sin(kh)|} \frac{1}{|\sin(k)|} \left(\frac{(kh)^2}{k} + (kh)^2 \left(1 + \frac{1}{|\sin(k)|} \right) \right).$$

Ainsi, pour le schéma avec shift $\tilde{k} = \hat{k}$, l'erreur décroît comme un $O(G^{-2})$ pour tout nombre d'onde.

Dans le cas sans correction de dispersion ($\tilde{k} = k$), un calcul montre que la borne supérieure de l'erreur donnée par le Théorème 3.10 est bornée par $((kh)^2 + k(kh)^4)$. Le terme $k(kh)^4$ est caractéristique de l'effet de pollution et montre que garder kh petit n'est pas suffisant pour empêcher l'erreur d'augmenter avec le nombre d'onde tout en gardant kh assez petit.

Lorsque la correction de dispersion est utilisée ($\tilde{k} = \hat{k}$), la borne supérieure de l'erreur ne dépend que de kh indiquant que l'effet de pollution a été supprimé.

Pour illustrer numériquement ces résultats, on choisit

$$f = \sin(kx),$$

et dans ce cas, la solution exacte de (3.22) est

$$u_{\text{ex}}(x) = -\frac{x \cos(kx)}{2k} + \left(\frac{\cos(k)}{2 \sin(k)k} \right) \sin(kx).$$

La figure 3.1 montre l'erreur relative avec et sans correction de dispersion pour $k = 200, 2000$ pour un nombre de points par longueur d'onde $G = 2\pi/(kh)$ variant entre 230 et 2. Ces résultats numériques confirment ceux obtenus dans le Théorème 3.10 et montrent l'absence d'effet de pollution lorsqu'on utilise la correction de dispersion.

Remarque 3.11 Dans [Chap3-A5], nous avons étendu la correction de dispersion décrite ci-dessus pour le schéma à 3-points pour des milieux dont le nombre d'onde est constant par morceaux. Les résultats numériques obtenus montrent que la correction de dispersion diminue grandement l'erreur ce qui permet de considérer des maillages plus grossiers que ceux utilisés pour avoir une précision similaire sans correction.

Remarque 3.12 Dans [Chap3-A7], nous avons étendu la construction du nombre d'onde shifté \hat{k} à des discrétisations différences finies générales permettant, en 1d, de supprimer l'erreur de dispersion pour différents schémas.

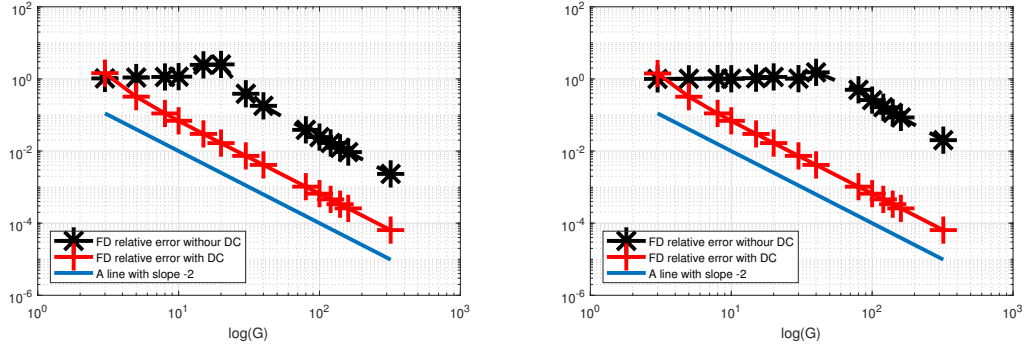


FIGURE 3.1 – Représentation en échelle log-log des erreurs relatives en fonction de G pour $k = 200$ (à gauche) et $k = 2000$ (à droite).

3.2 Correction de dispersion asymptotique pour un schéma à 9 points en 2d

Cette section présente des résultats issus de [Chap3-A6].

Les résultats présentés dans la Section 3.1 montrent qu'il est possible de supprimer totalement l'erreur de dispersion, et ainsi l'effet de pollution, en 1d en modifiant le schéma avec l'introduction d'un nombre d'onde shifté \hat{k} . En dimension supérieure, bien qu'il ne soit pas possible de supprimer l'erreur de dispersion, il est néanmoins possible de la réduire.

On présente tout d'abord un résultat assurant que l'erreur de dispersion est du même ordre que l'erreur d'approximation du schéma différences finies sur des ondes planes. Soit donc une discrétisation de l'opérateur de Helmholtz

$$\mathcal{H}u = -(\Delta + k^2)u,$$

qui est donnée par

$$(\mathcal{H}(k, h)u)_{i,j} = (-\Delta_h u)_{i,j} - k^2(\mathcal{M}_h u)_{i,j},$$

où h est le pas de maillage, supposé uniforme, et l'indice i, j indique que l'approximation est calculée au point (x_i, y_j) . Pour $\xi = (\xi_1, \xi_2)$ et $\mathbf{x} = (x, y)$, les symboles continu σ et discret σ_d sont

$$\sigma(k, \xi) = |\xi|^2 - k^2, \quad \sigma_d(k, h, \xi) = (e^{-i\mathbf{x} \cdot \xi})_{i,j} (\mathcal{H}(k, h)e^{i\mathbf{x} \cdot \xi})_{i,j}. \quad (3.24)$$

Pour un maillage uniforme, σ_d ne dépend pas de i, j . Les relations de dispersion continue et discrète sont alors définies par

$$\mathcal{D}_c := \{\xi \in \mathbb{R}^2 \mid |\xi|^2 - k^2 = 0\}, \quad \mathcal{D}_d := \{\xi \in \mathbb{R}^2 \mid \sigma_d(k, h, \xi) = 0\}, \quad (3.25)$$

et, pour $\theta \in \mathbb{S}^1$, le nombre d'onde discret $k_d = k_d(k, h, \theta)$ satisfait

$$\sigma_d(k, h, k_d(k, h, \theta)\theta) = 0. \quad (3.26)$$

En fonction des situations, il peut être commode de travailler avec le pas de maillage h ou le nombre de point par longueur d'onde G . On peut aussi définir le nombre d'onde discret comme une fonction de k, θ, G à partir de la relation

$$\sigma_d \left(k, \frac{2\pi}{kG}, k_d \theta \right) = 0. \quad (3.27)$$

L'erreur de dispersion est estimée dans le résultat ci-dessous.

Théorème 3.13 *On suppose que le symbole discret est une fonction régulière en chacune de ses variables pour G assez grand et que*

$$\nabla_{\xi} \sigma_d \left(k, \frac{2\pi}{kG}, \xi \right) = c_2 \xi + N(k, G, \xi),$$

où c_2 est une constante et $\|N(k, G, \xi)\| \leq \tilde{c}(\xi)G^{-1}$ où la fonction $\xi \in \mathbb{R}^2 \mapsto \tilde{c}(\xi) > 0$ est continue. Alors les deux assertions suivantes sont équivalentes :

i) *Le schéma différences finies approxime les ondes planes à l'ordre $q \in \mathbb{N}^*$:*

$$\forall \theta \in \mathcal{S}^1, (\mathcal{H}(k, h)e^{ikx \cdot \theta})_{i,j} = O_{G \rightarrow \infty}(G^{-q})k^2.$$

ii) *Le nombre d'onde discret admet le développement suivant*

$$k_d = k \left(1 + O_{G \rightarrow +\infty}(G^{-q}) \right).$$

Si de plus la relation de dispersion discrète admet une représentation polaire alors ii) est également équivalent à

iii) *La distance entre les relations de dispersion continue et discrète vérifie :*

$$\text{dist}(\mathcal{D}_d, \mathcal{D}_c) = kO_{G \rightarrow \infty}(G^{-q}).$$

Les hypothèses du Théorème 3.13 sont aisément vérifiables pour des schémas différences finies usuels. Ce résultat assure que l'erreur de dispersion ainsi que la distance entre les relations de dispersion discrète et continue sont du même ordre que l'erreur d'approximation du schéma sur les ondes planes. Dans la suite, nous allons effectuer une correction de dispersion pour un schéma différences finies basé sur un stencil à 9-points contenant des paramètres libres. Après avoir introduit le nombre d'onde shifté \hat{k} dans ce schéma, la correction de dispersion sera réalisée en minimisant, lorsque $G \rightarrow +\infty$, la distance entre les deux relations de dispersion permettant d'obtenir des formules explicites pour les paramètres optimisés et pour \hat{k} .

Pour une grille uniforme de pas de maillage $h > 0$, le schéma différences finies à 9-points est donné par

$$\begin{aligned} (\mathcal{H}_h v)_{i,j} &:= \left(\frac{4a}{h^2} - k^2 b \right) v(x_i, y_j) \\ &+ \left(\frac{1-2a}{h^2} - \frac{k^2 c}{4} \right) (v(x_{i-1}, y_j) + v(x_{i+1}, y_j) + v(x_i, y_{j-1}) + v(x_i, y_{j+1})) \\ &- \left(\frac{1-a}{h^2} + k^2 \frac{1-b-c}{4} \right) (v(x_{i-1}, y_{j-1}) + v(x_{i+1}, y_{j-1}) + v(x_{i-1}, y_{j+1}) + v(x_{i+1}, y_{j+1})), \end{aligned} \quad (3.28)$$

où a, b, c sont des constantes positives. En choisissant ces constantes comme suit

$$a = \frac{5}{6}, \quad c = \frac{b}{4} + \frac{b}{2} - \frac{5}{12} = 0, \quad (3.29)$$

le schéma (3.28) appliqué à toute onde plane vérifie

$$\begin{aligned} \left(\mathcal{H}_h e^{ikx \cdot e(\theta)} \right)_{i,j} &= -\frac{e^{ik(x_i, y_j) \cdot e(\theta)}}{720} h^4 k^6 \left(3 + (90c - 16)(\cos(\theta)^4 - \cos(\theta)^2) \right) \\ &+ k^2 O((kh)^6), \end{aligned}$$

et le Théorème 3.13 assure donc que $k_d = k(1 + O_{G \rightarrow +\infty}(G^{-4}))$.

On introduit maintenant le nombre d'onde shifté \hat{k} dans le schéma à 9-points. La relation (3.29) permettant d'exprimer a et b en fonction de c et il reste désormais deux paramètres \hat{k}, \hat{c} dans le schéma suivant :

$$\begin{aligned} \left(\mathcal{H}_h^{\text{asympt}} v \right)_{i,j} &:= \left(\frac{4a}{h^2} - \hat{k}^2 b \right) v(x_i, y_j) \\ &+ \left(\frac{1-2a}{h^2} - \frac{\hat{k}^2 \hat{c}}{4} \right) (v(x_{i-1}, y_j) + v(x_{i+1}, y_j) + v(x_i, y_{j-1}) + v(x_i, y_{j+1})) \\ &- \left(\frac{1-a}{h^2} + \hat{k}^2 \frac{1-b-\hat{c}}{4} \right) (v(x_{i-1}, y_{j-1}) + v(x_{i+1}, y_{j-1}) + v(x_{i-1}, y_{j+1}) + v(x_{i+1}, y_{j+1})). \end{aligned} \quad (3.30)$$

En minimisant, pour $G \rightarrow +\infty$, la distance entre la relation de dispersion continue et celle de $\mathcal{H}_h^{\text{asympt}}$ on a obtenu des expressions explicites pour \hat{c} et \hat{k} réalisant le minimum.

Théorème 3.14 Soient (a, b, \hat{c}) vérifiant 3.29 et \hat{k}, \hat{c} définis par

$$\begin{aligned} \hat{k} &= k - k \frac{\pi^4}{30} G^{-4} - k \frac{\pi^6}{192} G^{-6}, \\ \hat{c} &= \frac{8}{45} - \frac{\pi^2}{54} G^{-2}. \end{aligned}$$

Alors le nombre d'onde discret \hat{k}_d associé au schéma différences finies $\mathcal{H}_h^{\text{asympt}}$ vérifie

$$\hat{k}_d = k + k O_{G \rightarrow \infty}(G^{-6}).$$

Le Théorème 3.14 combiné au Théorème 3.13 assure que le schéma $\mathcal{H}_h^{\text{asympt}}$ est, avec l'aide du nombre d'onde shifté \hat{k} , d'ordre 6 sur les ondes planes tandis qu'il est d'ordre 4 sans nombre d'onde shifté.

Il est important de noter que l'erreur de dispersion a été minimisée uniquement lorsque $G \rightarrow +\infty$. Cette méthode est ainsi appelée *correction de dispersion asymptotique* et son principal avantage est d'obtenir des formules explicites pour les paramètres et le shift \hat{k} .

Pour qualifier la validité de cette approche, on considère le schéma $\mathcal{H}_h^{\text{opt}}$ défini par (3.30) où les paramètres (\hat{k}, \hat{c}) sont maintenant optimisés numériquement en minimisant la distance entre la relation de dispersion continue et celle du schéma $\mathcal{H}_h^{\text{opt}}$. La minimisation

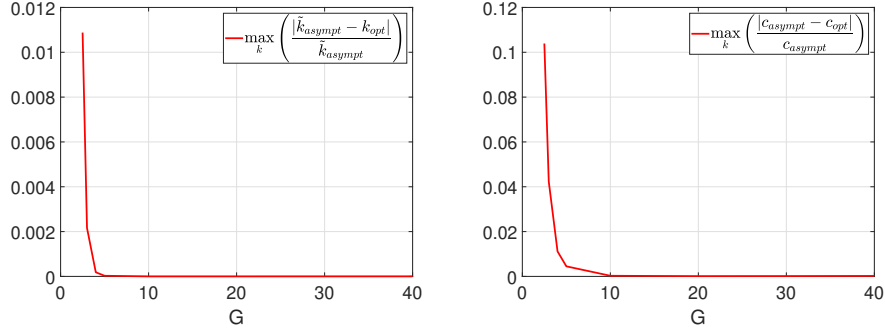


FIGURE 3.2 – (Gauche) Erreur relative entre k_{opt} et k_{asympt} du Théorème 3.14. (Droite) Erreur relative entre c^{opt} et c^{asympt} du Théorème 3.14. Pour chaque G , les erreurs sont calculées en prenant l’erreur relative maximale obtenue pour chaque $k \in \{20, 40, 80, 100, 160, 180, 200, 250, 300\}$.

a été effectuée avec la méthode de Nelder-Mead. On note $(k_{\text{opt}}, c_{\text{opt}})$ les paramètres ainsi calculés et $(k_{\text{asympt}}, c_{\text{asympt}})$ ceux obtenus explicitement dans le Théorème 3.14. Les erreurs relatives entre les paramètres asymptotiques et les optimaux sont données dans la Figure 3.2. On peut ainsi voir que les formules asymptotiques sont très proches des formules obtenues à partir d’une optimisation numérique pour $G \geq 10$ justifiant ainsi l’approche de correction de dispersion asymptotique si le nombre de points par longueur d’onde est suffisamment grand.

Pour terminer cette section, on étudie la précision des schémas $\mathcal{H}_h^{\text{asympt}}$ et $\mathcal{H}_h^{\text{opt}}$, pour résoudre le problème de Helmholtz suivant :

$$\begin{cases} \Delta u(\theta, \mathbf{x}) + k^2 u(\theta, \mathbf{x}) = 0, & \text{dans } \Omega := (-1, +1) \times (-1, +1), \\ u(\theta, \mathbf{x}) = \exp(ik\mathbf{x} \cdot \boldsymbol{\theta}), & \text{sur } \partial\Omega, \end{cases} \quad (3.31)$$

où $\boldsymbol{\theta} = (\cos(s), \sin(s))$. Notons que

$$u(\theta, \mathbf{x}) = \exp(ik\mathbf{x} \cdot \boldsymbol{\theta})$$

est l’unique solution de (3.31) dès lors que k^2 n’est pas valeur propre du Laplacien avec une condition de Dirichlet homogène sur $\partial\Omega$.

Les Théorèmes 3.14 et 3.13 entraînent que $\mathcal{H}_h^{\text{asympt}}$ est d’ordre 6 sur les ondes planes et nous allons donc comparer les performances de ce schéma avec d’autres schémas d’ordre 6. On considère pour cela le schéma issu [69], noté $\mathcal{H}^{\text{Sutmann}}$ et celui de [72, Eq. (2.15)], noté \mathcal{H}^{Wu} . Ce dernier est basé sur le même schéma différences finies que $\mathcal{H}^{\text{Sutmann}}$ mais une minimisation numérique de l’erreur de dispersion a été effectuée pour déterminer certains paramètres.

On considère un maillage cartésien uniforme de Ω dont les points $\mathbf{x}_{i,j} = (x_i, y_j) \in \Omega$ satisfont $|x_i - x_{i\pm 1}| = h$ et $|y_j - y_{j\pm 1}| = h$. On calcule alors une solution numérique pour (3.31) approximée par les 4 schémas différences finies pour les angles

$$\boldsymbol{\theta}_l = (\cos(s_l), \sin(s_l)) \text{ avec } s_l \in \left\{ \frac{2\pi}{N}l, l = 0, \dots, N \right\}.$$

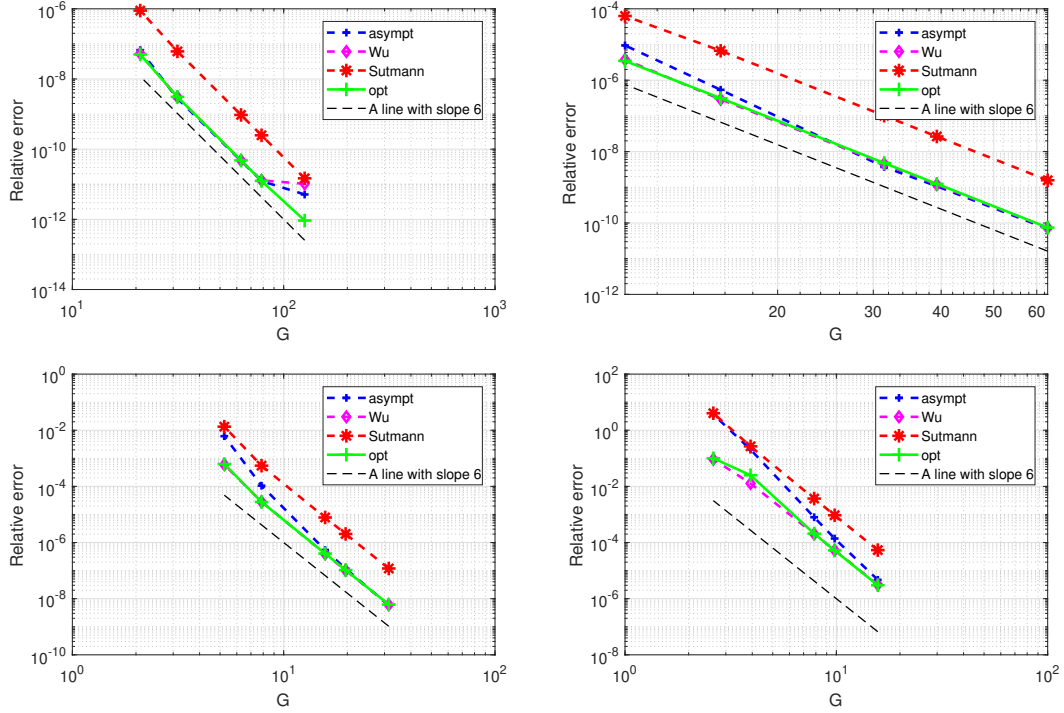


FIGURE 3.3 – (haut, gauche) Erreurs relatives pour $k = 10$, (haut, droite) $k = 20$, (bas, gauche) $k = 40$ et (bas, droite) $k = 80$.

En notant $\mathbf{u}_h(\theta) := (\mathbf{u}_{h,i,j}(\theta))_{i,j}$ la solution numérique, on calcule la moyenne des erreurs pour chaque angle considéré

$$\text{err}(h) = \frac{1}{N+1} \sum_{l=0}^N \frac{\|\mathbf{u}(\theta_l, \mathbf{x}_{i,j}) - \mathbf{u}_h(\theta_l)\|_2}{\|\mathbf{u}_h(\theta_l)\|_2}, \quad \mathbf{x}_{i,j} = (x_i, y_j), \quad \|\mathbf{v}\|_2 = \sqrt{\sum_{i,j} |v_{i,j}|^2}.$$

Les résultats obtenus sont visibles dans la Figure 3.3 pour $N = 20$, les pas de maillages $h \in \{0.005, 0.008, 0.01, 0.02, 0.03\}$ et les nombres d'ondes $k \in \{10, 20, 40, 80\}$. Tout d'abord, le schéma $\mathcal{H}^{\text{Sutmann}}$ étant le seul sans correction de dispersion, il est le moins précis des 4. Les schémas \mathcal{H}^{Wu} et \mathcal{H}^{opt} ont une précision très similaire. De même pour $\mathcal{H}_h^{\text{asympt}}$ et $\mathcal{H}_h^{\text{opt}}$ dès lors que le nombre de point par longueur d'onde est suffisamment grand soit pour $k = 10, 20$.

3.3 Correction de dispersion asymptotique pour des schémas différences finies généraux

Cette section présente des résultats issus de [Chap3-A7].

La correction de dispersion présentée dans le Théorème 3.14 a été effectuée en minimisant la distance entre les relations de dispersion continue et discrète. Malheureusement, cette technique est difficile à étendre pour des schémas différences finis généraux ou bien en dimensions supérieures.

Dans cette section, on présente des résultats issus de [Chap3-A7] qui permettent de déterminer pour un schéma différences finies général le *shift asymptotique optimal* \widehat{k} minimisant l'erreur de dispersion pour des pas de maillages h assez petits. La méthode présentée est donc une méthode de correction de dispersion dite asymptotique car l'optimisation est réalisée pour des maillages assez fins.

Pour une grille uniforme de pas de maillage h , on considère une discrétisation différences finies générales de $\mathcal{H} = -(\Delta + k^2)$ donnée par

$$(\mathcal{H}_h u)_i = -(\Delta_h u)_i - k^2(M_h u)_i,$$

où l'indice i indique que l'approximation est calculée au point x_i , $(\Delta_h u)_i$ est une approximation du Laplacien de u au point x_i et $(M_h u)_i$ correspond à une approximation de $u(x_i)$. On rappelle que le symbole discret est défini par

$$\sigma_d(k, h, \xi) = \left(e^{-ix \cdot \xi} \right)_i \left(\mathcal{H}_h e^{ix \cdot \xi} \right)_i,$$

et le nombre d'onde discret est donné pour tout $\theta \in \mathcal{S}^{d-1}$ par $k_d := k_d(k, h, \theta)$ qui vérifie

$$\sigma_d(k, h, k_d \theta) = 0.$$

On introduit maintenant le nombre d'onde shifté \widehat{k} dans le schéma différences finies

$$(\mathcal{H}_h u)_i = -(\Delta_h u)_i - \widehat{k}^2(M_h u)_i. \quad (3.32)$$

Le symbole discret associé au schéma (3.32) vérifie ainsi $\sigma_d(\widehat{k}, h, k\theta)$. Dans la suite, nous allons montrer comment déterminer explicitement \widehat{k} pour minimiser l'erreur de dispersion pour h assez petit. On suppose pour cela que le symbole discret satisfait les hypothèses ci-dessous :

(H1) Pour $h \rightarrow 0$, le symbole discret vérifie

$$\sigma_d(k, h, k\theta) = h^p \mathcal{E}(k, k\theta) + O(h^{p+1}),$$

où \mathcal{E} est une fonction régulière.

(H2) Pour k fixé, la suite de fonctions $(\nabla_\xi \sigma_d(k, h, \cdot))_h$ converge uniformément, pour $h \rightarrow 0$, vers $\nabla_\xi \sigma_c(k, \cdot)$ dans un voisinage compact de $\xi = k\theta$ pour tout $\theta \in \mathcal{S}^{d-1}$.

(H3) Pour ξ fixé, la suite de fonctions $(\partial_k \sigma_d(\cdot, h, \xi))_h$ converge uniformément, pour $h \rightarrow 0$, vers $\partial_k \sigma_c(\cdot, \xi)$ dans un voisinage compact du nombre d'onde k .

L'hypothèse (H1) est vérifiée pour tout schéma d'ordre p sur les ondes planes. En effet, dans ce cas, on a

$$(\mathcal{H}_h u_\theta)_i = h^p \mathcal{E}(k, k\theta) u_\theta(x_i) + O(h^{p+1}),$$

pour toute onde plane $u_\theta(x) = e^{ikx \cdot \theta}$ où $\theta \in \mathcal{S}^{d-1}$. Ainsi, (H1) peut être vérifiée directement en calculant le développement de Taylor, pour $h \rightarrow 0$, du symbole discret en

$\xi = k\theta$. De plus, les hypothèses (H2)-(H3) peuvent également être vérifiées en calculant le développement de Taylor des dérivées de σ_d par rapport à (k, ξ) .

En cherchant le nombre d'onde shifté sous la forme

$$\widehat{k} = k + k_p h^p,$$

on a démontré que le nombre d'onde discret associé au schéma (3.32) admettait le développement

$$k_d(k, h, \theta) = k - \frac{h^p}{2k} (-2kk_p + \mathcal{E}(k, k\theta)) + O(h^{p+1}).$$

On peut ainsi minimiser l'erreur de dispersion lorsque $h \rightarrow 0$ en déterminant k_p comme suit

$$k_p^{\text{asy}} := \arg \min_{k_p} \left(\max_{\theta \in \mathcal{S}^{d-1}} |-2kk_p + \mathcal{E}(k, k\theta)| \right), \quad (3.33)$$

où k_p^{asy} est appelé le *shift asymptotique optimal*. Ce dernier peut être calculé explicitement à partir des valeurs minimale et maximale de la fonction $\theta \mapsto \mathcal{E}(k, k\theta)$.

Théorème 3.15 *On suppose qu'il existe $\theta_{\min}, \theta_{\max} \in \mathcal{S}^{d-1}$ tels que*

$$\mathcal{E}_{\min} := \mathcal{E}(k, k\theta_{\min}) \leq \mathcal{E}(k, k\theta) \leq \mathcal{E}(k, k\theta_{\max}) =: \mathcal{E}_{\max},$$

où les bornes supérieure et inférieure peuvent dépendre de k . Alors le shift asymptotique optimal k_p^{asy} défini par (3.33) est unique et

$$k_p^{\text{asy}} = \frac{1}{4k} (\mathcal{E}_{\max} + \mathcal{E}_{\min}).$$

L'erreur de dispersion relative admet de plus le développement suivant

$$\max_{\theta \in \mathcal{S}^{d-1}} \left| \frac{k_d(\widehat{k}^{\text{asy}}, h, \theta) - k}{k} \right| = \frac{h^p}{2k^2} \left| \frac{\mathcal{E}_{\max} - \mathcal{E}_{\min}}{2} \right| + O(h^{p+1}),$$

où $\widehat{k}^{\text{asy}} = k + h^p k_p^{\text{asy}}$.

Le Théorème 3.15 permet de calculer le shift asymptotique optimal pour tout schéma différences finies vérifiant (H1)-(H2)-(H3) et 2 exemples seront présentés dans la suite. Afin de quantifier l'effet du shift sur la réduction de l'erreur de dispersion, on introduit le facteur de réduction

$$R_f := \frac{\max_{\theta} |\mathcal{E}(k, k\theta)|}{\max_{\theta} |-2kk_p^{\text{asy}} + \mathcal{E}(k, k\theta)|} = 2 \frac{\max\{|\mathcal{E}_{\max}|, |\mathcal{E}_{\min}|\}}{|\mathcal{E}_{\max} - \mathcal{E}_{\min}|}, \quad (3.34)$$

correspondant au ratio des développements, avec et sans shift, des nombres d'onde discrets tronqués à l'ordre p . Dans les cas présentés ci-dessous, on a toujours $R_f \geq 2$.

Application au schéma à 5 points

Le schéma différences finies à 5 points pour l'équation de Helmholtz est

$$\left(\mathcal{H}_h^{5pt} \mathbf{u}\right)_{i,j} = \frac{-\mathbf{u}_{i+1,j} - \mathbf{u}_{i-1,j} + 4\mathbf{u}_{i,j} - \mathbf{u}_{i,j+1} - \mathbf{u}_{i,j-1}}{h^2} - k^2 \mathbf{u}_{i,j},$$

et son symbole discret est donc

$$\sigma_d^{5pt}(k, h, \xi) = \frac{4 - 2(\cos(h\xi_1) + \cos(h\xi_2))}{h^2} - k^2.$$

En effectuant un développement de Taylor lorsque $h \rightarrow 0$, on peut vérifier aisément que \mathcal{H}_h^{5pt} satisfait les hypothèses (H1)-(H2)-(H3) avec $p = 2$ et

$$\varepsilon(k, k\theta) = -\frac{k^4}{12} \left(\cos(s)^4 + \sin(s)^4 \right) = -\frac{k^4}{12} \left(2 \cos(s)^4 - 2 \cos(s)^2 + 1 \right),$$

où on a utilisé que $\theta \in \mathcal{S}^1$ pouvait s'exprimer comme $\theta = (\cos(s), \sin(s))$ avec $s \in [0, 2\pi]$. En appliquant le Théorème 3.15 sur \mathcal{H}_h^{5pt} , on obtient une formule explicite pour le shift asymptotique optimal et le facteur de réduction associés au schéma à 5 points.

Théorème 3.16 *Pour le schéma \mathcal{H}_h^{5pt} , on a*

$$k_2^{asy} = -\frac{k^3}{32}, \quad R_f = 4.$$

On étudie maintenant numériquement l'influence du shift asymptotique optimal. Considérons pour cela le cas test suivant

$$\begin{cases} -\Delta u - k^2 u = 0, & \text{dans } \Omega = (0, 1)^2, \\ u = f, & \text{sur } \{0\} \times [0, 1] \cup \{1\} \times [0, 1], \\ \partial_n u + iku = g, & \text{sur } [0, 1] \times \{0\} \cup [0, 1] \times \{1\}, \end{cases} \quad (3.35)$$

où f, g sont calculés de sorte que $u(x, y) = \sin(k(x + y) \sqrt{2}/2)$ soit la solution exacte. La condition de Robin est discrétisée avec un point fantôme pour conserver un schéma d'ordre 2. On calcule les erreurs suivantes

$$\text{err}_\infty(\tilde{k}) = \frac{\|\mathbf{u} - \mathbf{u}_h(\tilde{k})\|}{\|\mathbf{u}\|_\infty},$$

où $\mathbf{u}_h(\tilde{k})$ est la solution numérique sans shift (dans ce cas $\tilde{k} = k$) ou avec shift (dans ce cas $\tilde{k} = \hat{k} = k - h^2 k^3 / 32$, voir Théorème 3.16).

Les erreurs relatives sont données dans la Figure 3.4 pour $k = 20, 40, 80, 160$ et des pas de maillages vérifiants $h = 1/(n + 1)$ pour $n = 2^j$ et $j = 4, \dots, 10$. Ces résultats montrent que, pour h assez petit, le shift asymptotique optimal réduit l'erreur relative. En calculant les ratios entre les erreurs avec et sans shift asymptotique optimal dans le Tableau 3.6, on peut voir que nous pouvons réduire l'erreur relative d'un facteur 2 pour un nombre de points par longueur d'onde assez grand (ou de manière équivalente : lorsque h est assez petit).

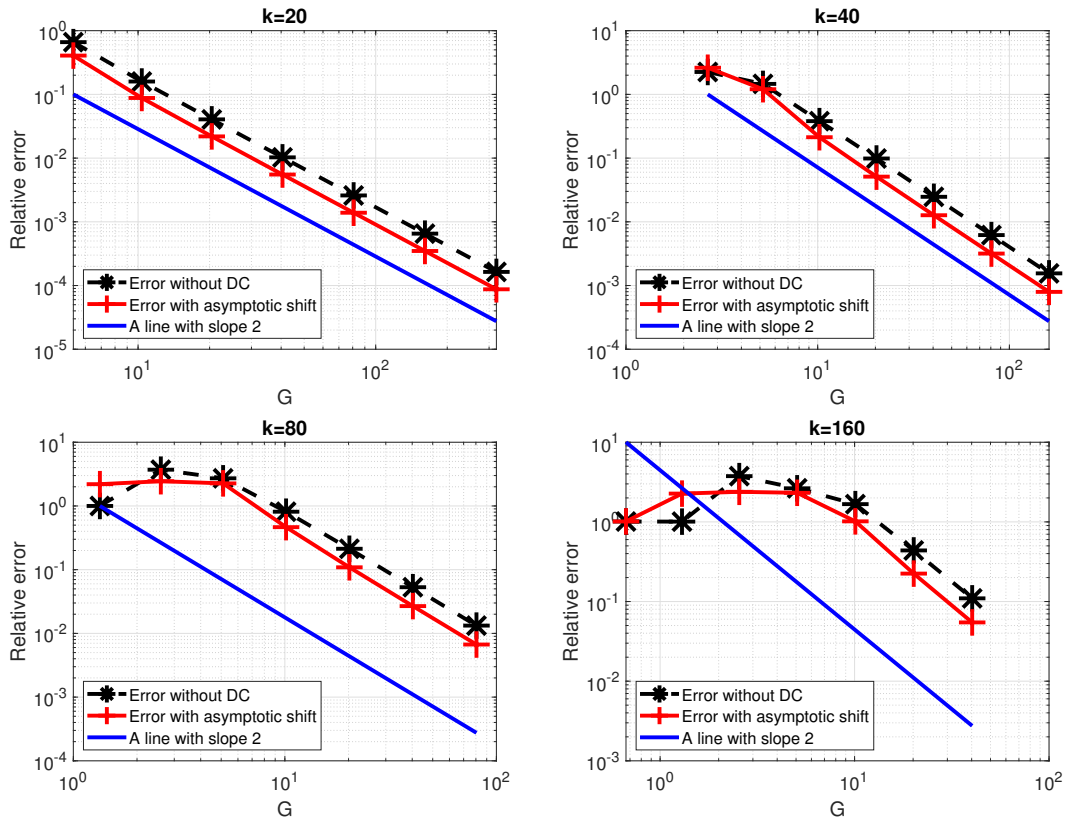


FIGURE 3.4 – Erreurs relatives avec et sans shift asymptotique optimal.

n	2 ⁴	2 ⁵	2 ⁶	2 ⁷	2 ⁸	2 ⁹	2 ¹⁰
k = 20	0.6153	0.5514	0.5406	0.5376	0.5353	0.5342	0.5335
k = 40	1.1662	0.8251	0.5594	0.5190	0.5136	0.5126	0.5123
k = 80	2.1981	0.6594	0.8327	0.5723	0.5118	0.5052	0.5045
k = 160	1.0015	2.2481	0.6354	0.8776	0.6078	0.5119	0.4988

TABLEAU 3.6 – Ratio $\text{err}_\infty(\hat{k}) / \text{err}_\infty(k)$.

Application à un schéma à 9 points

On applique maintenant le Théorème 3.15 au schéma à 9 points ci-dessous

$$\begin{aligned}
 (\mathcal{H}_h^{9\text{-pts}} v)_i &:= \left(\frac{4a}{h^2} - k_g^2 b \right) v(x_i, y_j) \\
 &+ \left(\frac{1-2a}{h^2} - \frac{k_g^2 c}{4} \right) (v(x_{i-1}, y_j) + v(x_{i+1}, y_j) + v(x_i, y_{j-1}) + v(x_i, y_{j+1})) \\
 &- \left(\frac{1-a}{h^2} + k_g^2 \frac{1-b-c}{4} \right) (v(x_{i-1}, y_{j-1}) + v(x_{i+1}, y_{j-1}) + v(x_{i-1}, y_{j+1}) + v(x_{i+1}, y_{j+1})),
 \end{aligned} \tag{3.36}$$

où a, b, c et k_g sont des constantes positives définies par

$$a = \frac{5}{6}, \quad b = \frac{5}{6} - \frac{c}{2}, \quad c = \frac{8}{45} + c_2 G^{-2}, \quad k_g = k - \frac{\pi^4 k}{30} G^{-4},$$

et c_2 est un paramètre libre. Ce schéma est similaire à (3.28) mais les constantes a, b, c ont été choisies pour que $\mathcal{H}_h^{9\text{-pts}}$ soit directement d'ordre 6. En écrivant $\theta = (\cos(s), \sin(s))$, le symbole discret de $\mathcal{H}_h^{9\text{-pts}}$ satisfait le développement

$$\begin{aligned}\sigma_d^{9\text{-pts}}(k, h, k\theta) &= -\frac{k^8 h^6}{6048\pi^2} (2 \cos(s)^8 \pi^2 - 4 \cos(s)^6 \pi^2 + 6 \cos(s)^4 \pi^2 \\ &\quad + 189c_2 \cos(s)^4 - 4 \cos(s)^2 \pi^2 - 189 \cos(s)^2 c_2 + \pi^2) + O(h^8) \\ &= h^6 \mathcal{E}(k, k\theta, c_2) + O(h^8),\end{aligned}$$

et on peut aisément vérifier que les hypothèses (H1)-(H2)-(H3) sont vérifiées avec $p = 6$ et $\mathcal{E}(k, k\theta, c_2)$ défini ci-dessus. En appliquant le Théorème 3.15 au schéma $\mathcal{H}_h^{9\text{-pts}}$, on a démontré le résultat suivant.

Théorème 3.17 *Soit c_2^* le paramètre minimisant l'erreur de dispersion pour h assez petit. Ce dernier est ainsi défini par*

$$c_2^* = \arg \min_{c_2} |\mathcal{E}(k, k\theta, c_2)|.$$

On a alors

$$c_2^* = -\frac{\pi^2}{54}, \quad k_6^{\text{asy}} = -\frac{k^7}{12288}, \quad \text{et} \quad R_f = 64.$$

Pour étudier numériquement l'effet du shift asymptotique optimal associé à $\mathcal{H}_h^{9\text{-pts}}$, on considère le cas test ci-dessous

$$\begin{cases} -\Delta u - k^2 u = 0, & \text{dans } \Omega = (-1, 1)^2, \\ u = f, & \text{sur } \partial\Omega, \end{cases} \quad (3.37)$$

où le terme source f est calculé de sorte que $u(x, y) = \sin(k(x + y) \sqrt{2}/2)$ soit la solution exacte. Sans shift asymptotique optimal, les constantes intervenant dans $\mathcal{H}_h^{9\text{-pts}}$ sont

$$a = \frac{5}{6}, \quad b = \frac{5}{6} - \frac{c}{2}, \quad c = \frac{8}{45} - \frac{\pi^2}{54} G^{-2}, \quad k_g = k \left(1 - \frac{\pi^4}{30} G^{-4} \right),$$

Lorsqu'on utilise le nombre d'onde shifté, ces constantes deviennent

$$\begin{aligned}c &= \frac{8}{45} - \frac{\pi^2}{54} \left(\frac{2\pi}{(k + h^6 k_6^{\text{asy}})h} \right)^{-2}, \\ k_g^{\text{asy}} &= (k + h^6 k_6^{\text{asy}}) \left(1 - \frac{\pi^4}{30} \left(\frac{2\pi}{(k + h^6 k_6^{\text{asy}})h} \right)^{-4} \right),\end{aligned} \quad (3.38)$$

où k_6^{asy} est donné dans le Théorème 3.17. Lorsque $h \rightarrow 0$, les constantes (3.38) vérifient

$$c^{\text{asy}} = \frac{8}{45} - \frac{\pi^2}{54} G^{-2} + O(h^7), \quad k_g^{\text{asy}} = k \left(1 - \frac{\pi^4}{30} G^{-4} - \frac{\pi^6}{192} G^{-6} \right) + O(h^7).$$

Le schéma $\mathcal{H}_h^{9\text{-pts}}$ étant d'ordre 6, on peut définir le schéma $\mathcal{H}_h^{9\text{-pts}}$ avec correction de dispersion en tronquant les développements à l'ordre 6 des constantes c^{asy} et k_g^{asy} . Le schéma ainsi obtenu est alors exactement celui du Théorème 3.14 issu de [Chap3-A6].

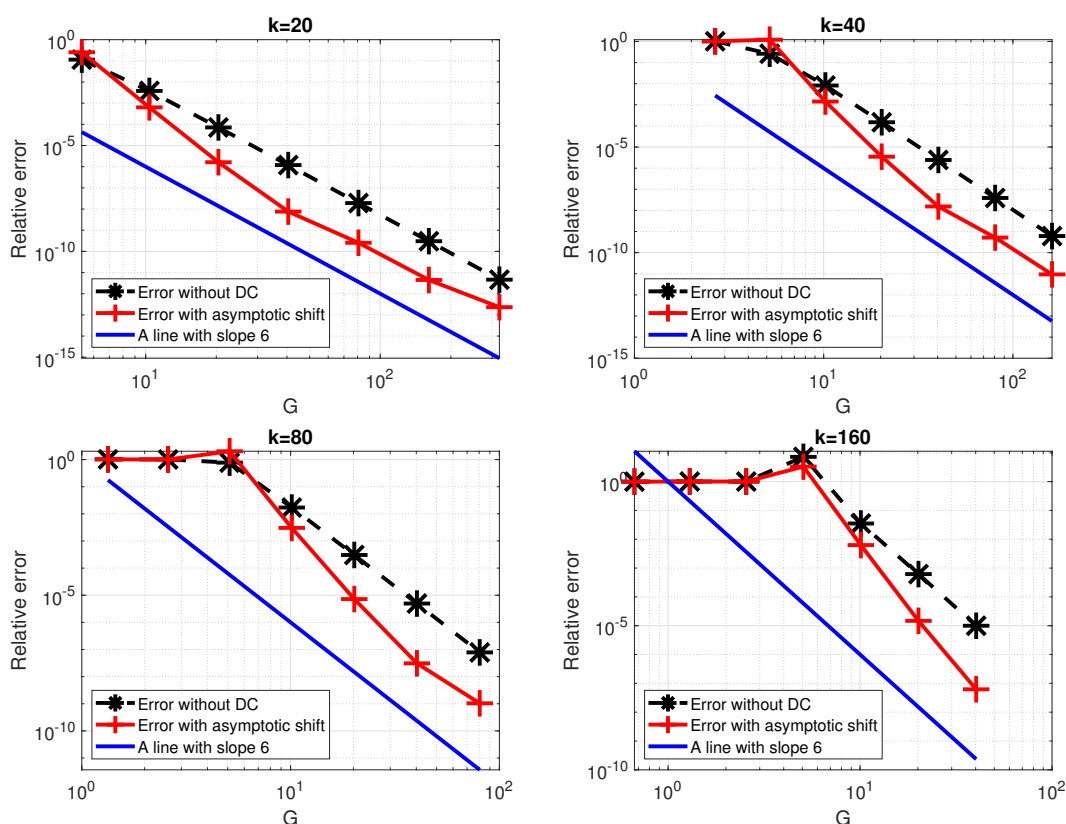


FIGURE 3.5 – Erreurs relatives avec et sans shift asymptotique optimal pour $\mathcal{H}_h^{9\text{-pts}}$.

Les erreurs relatives, avec et sans shift asymptotique optimal, sont données dans la Figure 3.5 pour $k = 20, 40, 80, 160$ et des pas de maillages vérifiant $h = 1/(n + 1)$ pour $n = 2^j$ et $j = 4, \dots, 10$. On peut voir que la correction de dispersion asymptotique réduit l'erreur relative dès que le pas de maillage est assez petit.

Les facteurs de réduction sont également donnés dans le Tableau 3.7. Les cas pour lesquels l'erreur relative avec correction de dispersion est plus grande que celle sans correction ($(k, n) = (20, 2^5)$ et $(k, n) = (80, 2^6)$) correspondent à des combinaisons de k, h proches d'une valeur propre. Pour les autres choix de (k, n) , l'utilisation du nombre d'onde shifté avec shift asymptotique optimal diminue l'erreur relative.

n	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}
k = 20	2.231	0.167	0.022	0.006	0.013	0.014	0.050
k = 40	0.994	4.647	0.174	0.023	0.006	0.013	0.015
k = 80	0.999	0.997	2.627	0.177	0.023	0.006	0.013
k = 160	0.999	0.999	0.998	0.458	0.178	0.023	0.006

TABLEAU 3.7 – Ratio des erreurs $\text{err}_\infty(\hat{k}) / \text{err}_\infty(k)$.

3.4 Résultats numériques : correction de dispersion et multigrille

Cette section présente des résultats issus de [Chap3-A6].

On présente dans cette section des résultats numériques étudiant l'influence de la correction de dispersion sur la convergence de l'algorithme multigrille (3.11). Considérons pour cela l'équation de Helmholtz sur $\Omega = (0, 1)^2$ avec conditions de Dirichlet homogènes. Cette dernière est discrétisée avec le schéma $\mathcal{H}_h^{9\text{-pts}}$ avec ou sans shift asymptotique optimal calculé dans le Théorème 3.17.

On utilise des opérateurs de restriction et de prolongation classiques donnés par

$$\begin{aligned} (R_l \mathbf{u})_{i,j} &:= \frac{1}{16} (4u_{2i,2j} + 2(u_{2i-1,2j} + u_{2i+1,2j} + u_{2i,2j-1} + u_{2i,2j+1}) \\ &\quad + u_{2i-1,2j-1} + u_{2i+1,2j-1} + u_{2i-1,2j+1} + u_{2i+1,2j+1}), \\ P_l &:= 4R_l^T. \end{aligned} \quad (3.39)$$

Pour le lisseur, nous utilisons celui dont la matrice d'itération est

$$S_l = I_d - \omega_l A_l^* A_l,$$

avec $\omega_l = \rho(A_l)^{-2}$. La matrice d'itération de ce lisseur est Hermitienne et son spectre vérifie $\sigma(S_l) \subset [0, 1]$. On a donc

$$\|S_l\|_2 = \rho(S_l) \leq 1,$$

ce qui implique que son *multigrid contraction number* (voir Théorème 3.3) satisfait $\|S_l^\nu\|_2 \leq 1$ pour tout $\nu \in \mathbb{N}^*$.

Pour chaque niveau l , une grille uniforme de $\Omega = (0, 1)^2$ composée de $n_l \times n_l$ points est considérée. Le pas de maillage est donné par

$$n_l = 2^l - 1, \quad h_l = \frac{1}{n_l + 1},$$

et ainsi les grilles sont emboîtées. Le même nombre de pas de lissage ν est utilisé à chaque niveau. Pour chaque simulation, le nombre d'onde k est choisi de sorte à avoir environ π points par longueur d'onde au niveau le plus grossier. Le niveau le plus fin est toujours composé d'une grille ayant $n_p \times n_p$ points avec $p = 6$. Les résultats pour des V-cycles ($\tau = 1$ dans Algorithme 3.11) sont représentés dans le Tableau 3.8. Le Tableau 3.9 présente des résultats similaires pour un W-cycle ($\tau = 2$ dans Algorithme 3.11).

On peut noter que le rayon spectral de la matrice d'itération de l'algorithme multigrille est une fonction décroissante du nombre de pas de lissage ν . Cette propriété est attendue au moins théoriquement pour un W-cycle au vu du Théorème 3.3. Ces résultats montrent que le shift asymptotique optimal réduit le nombre de pas de lissage nécessaires pour avoir un algorithme multigrille convergent.

Remarque 3.18 Dans [Chap3-A5], en 1d et pour des milieux dont le nombre d'onde est constant par morceaux, nous avons également montré numériquement que la correction de dispersion réduisait le nombre de pas de lissage pour lequel l'algorithme multigrille convergeait.

$L = 2, k = 64 \setminus v$	1	5	10	20
9-pts	35.0880	18.5364	13.3576	12.1106
asympt	22.4954	6.9003	1.6712	0.6875
$L = 3, k = 32$				
9-pts	13.6200	6.3527	5.6742	5.6195
asympt	25.8102	5.2241	0.8920	0.6922
$L = 4, k = 16$				
9-pts	7.3096	3.0443	2.2077	2.0340
asympt	25.2834	4.5225	0.8597	0.6956

TABLEAU 3.8 – Rayon spectral de la matrice d’itération d’un V-cycle ($\tau = 1$) et $p = 6$.

$L = 3, k = 32 \setminus v$	1	5	10	20
9-pts	176.9828	43.2154	33.1786	32.2884
asympt	573.3670	36.0669	1.8799	0.4950
$L = 4, k = 16$				
9-pts	2.4772×10^3	0.1379×10^3	0.0335×10^3	0.0210×10^3
asympt	3.1880×10^5	1.105×10^3	3.0124	0.2558

TABLEAU 3.9 – Rayon spectral de la matrice d’itération d’un W-cycle ($\tau = 2$) et $p = 6$.

Troisième partie

Exposé sur les perspectives de recherche

L'objectif de cette partie est de présenter quelques perspectives de recherche en lien avec les travaux détaillés dans la partie II. Ces perspectives pourront également donner lieu, si possible, à des encadrements de thèses ou de post-doctorats.

Les premières perspectives de recherche que je donne ci-dessous sont issues des résultats présentés dans le chapitre 2 qui portait sur l'étude de problèmes d'optimisation topologique en mécanique des fluides.

Décomposition de domaine pour des problèmes d'optimisation avec contraintes d'EDP :

Les résultats numériques issus de la résolution de problèmes d'optimisation topologique exposés dans le chapitre 2 ont été obtenus en utilisant des algorithmes de descente de gradient. Nous avons de plus adopté l'approche *Optimize-then-discretize* en calculant le gradient de la fonction coût à partir d'un modèle adjoint continu. Une perspective pertinente concerne donc l'accélération de la résolution numérique des problèmes d'optimisation topologique. On propose pour cela d'étudier des méthodes de décomposition de domaine (DDM) appliquées à des problèmes d'optimisation, avec contraintes données par des EDP. Ces dernières sont actuellement en plein développement (voir par exemple [44, 57, 44, 45, 56]). Toutefois, dans la mesure où un problème d'optimisation topologique implique des EDP non-linéaires dans lesquelles les paramètres d'optimisation interviennent également de manière non-linéaire, des avancées sont encore à effectuer dans ce domaine.

Une première piste serait d'appliquer directement des méthodes (DDM) sur les problèmes direct et adjoint pour accélérer le calcul du gradient de la fonction coût. Une idée pour aller au-delà d'appliquer directement DDM pour résoudre le problème direct puis l'adjoint serait de construire des méthodes DD sur le système couplant ces deux EDP. Nous avons commencé à étudier dans [71] ce type de technique, en se basant sur une méthode de Schwarz optimisée, pour un problème d'identification de paramètres dans une équation d'advection diffusion linéaire. Les résultats numériques obtenus sont plutôt encourageants et une analyse numérique plus poussée des performances de cette méthode constitue une perspective à court terme.

Pour aller au-delà des études numériques de [71], un second axe de travail visera l'analyse mathématique de la convergence de méthodes DDM pour des problèmes d'optimisation avec contraintes données par des EDP. Pour traiter des problèmes de difficultés croissantes, on pourra commencer par exemple par l'analyse d'une méthode de décomposition de domaine pour un problème d'optimisation linéaire quadratique où les contraintes seront données par une EDP elliptique du second ordre. Ces travaux permettront d'étendre ceux issus de [57] au cas sans recouvrement avec conditions de Robin sur l'interface. Dans la continuité, nous proposerons et analyserons des méthodes de décomposition de domaine sans recouvrement pour déterminer un terme source dans le système de Navier-Stokes incompressible minimisant un écart de vitesse à une cible. Les résultats de [73]

pourront être utilisés et étendus au système d'optimalité associé au problème d'optimisation considéré. Une dernière tâche pourra enfin porter sur l'extension des méthodes DDM précédemment développées à un problème d'optimisation topologique en mécanique des fluides visant à trouver une répartition de solide optimisée minimisant, par exemple, les pertes de charge.

Une dernière voie d'exploration possible consiste à effectuer une décomposition de domaine directement au niveau du problème d'optimisation. Cela peut se réaliser en décomposant par sous-domaine les EDP agissant comme contraintes puis à traiter ces nouvelles contraintes avec différentes techniques d'optimisation. Nous avons appelé cette approche "*Décomposer-puis-Optimiser*" par opposition aux approches plus classiques, appelées "*Optimiser-puis-Décomposer*", qui consistent à construire et analyser des méthodes DDM sur les conditions nécessaires d'optimalité du premier ordre du problème d'optimisation.

Avec cette approche, nous avons traité en détail le cas d'un problème d'optimisation linéaire quadratique où les conditions de continuité à l'interface fictive ont été traitées avec une méthode de Lagrangien augmenté. Un des avantages du "*Décomposer-puis-Optimiser*" est le grand nombre de variantes provenant des différents choix permettant de décomposer les contraintes par sous-domaine puis de les prendre en compte avec des techniques d'optimisation. Une perspective à long terme consistera à étudier théoriquement et numériquement les propriétés de certaines de ces variantes et, si possible, de faire le lien entre ces dernières et des méthodes de décomposition de domaine directement appliquées aux conditions d'optimalité du premier ordre.

Les perspectives données ci-dessous concernent des suites envisageables, à moyen et long terme, des résultats exposés dans le chapitre 3 qui portait sur l'analyse numérique de problèmes de propagation d'onde en régime harmonique.

Influence de la correction de dispersion sur l'algorithme multigrille : Les résultats numériques présentés dans la Section 3.4 montrent que la correction de dispersion associée à un algorithme multigrille permet d'obtenir une méthode itérative convergente dès lors qu'un nombre assez grand de pas de lissage est effectué. Le choix du lisseur étant déterminant pour obtenir ce type de résultat, une piste future serait d'étudier davantage l'effet du lisseur et de la correction de dispersion sur la convergence d'une méthode multigrille. Un résultat similaire a déjà été obtenu en 1d dans [54] où la correction de dispersion assure la convergence d'un algorithme multigrille avec un lisseur de type Jacobi à deux-pas. La démonstration utilisant une diagonalisation par bloc de la matrice d'itération de la méthode multigrille, une première piste naturelle concerne l'utilisation de cette diagonalisation en dimension supérieure. On pourra par exemple se baser sur l'analyse Fourier présentée dans [59, Chapitre 8] qui permet d'étudier la convergence d'un algorithme multigrille, avec un lisseur de type Jacobi, pour l'équation de Poisson. Un autre axe de développement envisageable concerne l'étude de l'influence de la correction de dispersion sur le shift (complexe) minimal pour lequel le préconditionneur de

type *Laplacien shifté* pourrait être calculer avec une méthode multigrille.

Correction de dispersion pour différents modèles et d'autres discrétisations : La correction de dispersion asymptotique proposée concerne actuellement uniquement les schémas différences finies et l'équation de Helmholtz. Cette dernière étant basée sur l'introduction d'un nombre d'onde shifté dans le schéma puis sur le développement du nombre d'onde discret, il n'y a aucune obstruction à l'appliquer à d'autres méthodes de discrétisation. Néanmoins, la principale difficulté reste le calcul du développement du nombre d'onde discret ce qui constituera une des principales pistes à étudier pour la suite.

L'extension de cette technique à d'autres modèles de propagation d'ondes en régime harmonique, telles que les équations de Maxwell ou le système de l'élasticité linéaire, constitue un autre axe de futurs développements. Ces modèles étant vectoriels, leurs relations de dispersion ne sont plus systématiquement une sphère ce qui constitue une différence majeure, et donc une difficulté supplémentaire, avec le modèle de Helmholtz qui est scalaire.

Correction de dispersion pour des équations d'ondes temporelles : En régime harmonique, la pulsation ω est fixée, et la correction de dispersion peut être effectuée en perturbant le nombre d'onde k dans le schéma puis en réduisant la distance entre les relations de dispersion continu et discrète ce qui revient à réduire l'écart les nombres d'onde continu k et discret k_d . En régime temporel, la difficulté supplémentaire est qu'il faut maintenant réduire la distance entre les relations de dispersion qui sont données par

$$\mathcal{D}_c = \{(\omega, k) \mid \sigma(\omega, k) = 0\}, \quad \mathcal{D}_d = \{(\omega_d, k_d) \mid \sigma_d(\omega_d, k_d, h, \delta t) = 0\},$$

où δt est le pas de temps et σ, σ_d sont respectivement les symboles continu et discret. En particulier, en régime harmonique, la méthode utilisée pour déterminer le développement de k_d lorsque $h \rightarrow 0$ utilise fortement le caractère scalaire de k_d . On ne peut donc pas utiliser cette méthode pour estimer la distance entre $(\omega, k) \in \mathcal{D}_c$ et $(\omega_d, k_d) \in \mathcal{D}_d$. Cette étape constituant le point de départ pour effectuer la correction de dispersion asymptotique pour le régime harmonique, de nouvelles pistes restent donc à explorer.

Une piste alternative à la technique précédemment décrite peut être trouvée dans [63] qui présente une méthode numérique pour réduire l'erreur de dispersion pour une discrétisation différences finies d'une équation d'onde. L'approche est basée sur l'ajout d'un terme faisant intervenir des dérivées spatiales d'ordre 4 dans l'équation des ondes. En discrétisant le modèle ainsi perturbé avec des différences finies (schéma à 3 points pour les dérivées spatiales et schéma explicite en temps), on obtient une approximation de la vitesse des ondes numériques d'ordre $(kh)^4$ tandis que l'ordre d'approximation est en $(kh)^2$ sans la perturbation. Ce gain de précision rend l'approche très intéressante ce qui donne lieu à des perspectives de recherche. L'extension de cette technique de correction de dispersion en dimensions supérieures n'ayant, à ma connaissance, pas été effectuée, cela constituera une première piste de développement. Une seconde perspective concernera l'analyse

d'erreur entre le modèle perturbé, qui est un problème de perturbation singulière, et l'équation d'onde initialement considérée.

Bibliographie

- [1] Alexandersen, J., & Andreasen, C. S. (2020). A review of topology optimisation for fluid-based problems. *Fluids*, 5(1), 29.
- [2] Angot, P., Bruneau, C. H., & Fabrie, P. (1999). A penalization method to take into account obstacles in incompressible viscous flows. *Numerische Mathematik*, 81(4), 497-520.
- [3] Ambrosio, L., Fusco, N., & Pallara, D. (2000). *Functions of bounded variation and free discontinuity problems*. Courier Corporation.
- [4] Arndt, D., Braack, M., & Lube, G. (2016). Finite elements for the Navier-Stokes problem with outflow condition. In *Numerical Mathematics and Advanced Applications ENUMATH 2015* (pp. 95-103). Springer International Publishing.
- [5] Arndt, R., Ceretani, A. N., & Rautenberg, C. N. (2020). On existence and uniqueness of solutions to a Boussinesq system with nonlinear and mixed boundary conditions. *Journal of Mathematical Analysis and Applications*, 490(1), 124201.
- [6] Banks, H. T., & Kunisch, K. (2012). *Estimation techniques for distributed parameter systems*. Springer Science & Business Media.
- [7] Borrvall, T., & Petersson, J. (2003). Topology optimization of fluids in Stokes flow. *International journal for numerical methods in fluids*, 41(1), 77-107.
- [8] Boyer, F., & Fabrie, P. (2012). *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models* (Vol. 183). Springer Science & Business Media.
- [9] Brangeon, B., Joubert, P., & Bastide, A. (2015). Influence of the dynamic boundary conditions on natural convection in an asymmetrically heated channel. *International Journal of Thermal Sciences*, 95, 64-72.
- [10] Bruneau, C. H., & Fabrie, P. (1996). New efficient boundary conditions for incompressible Navier-Stokes equations : a well-posedness result. *ESAIM : Mathematical Modelling and Numerical Analysis*, 30(7), 815-840.
- [11] Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing*, 16(5), 1190-1208.

- [12] Ceretani, A. N., & Rautenberg, C. N. (2019). The Boussinesq system with mixed non-smooth boundary conditions and do-nothing boundary flow. *Zeitschrift für angewandte Mathematik und Physik*, 70(1), 14.
- [13] Çeşmelioglu, A., & Riviere, B. (2008). Analysis of time-dependent Navier-Stokes flow coupled with Darcy flow, *J. Numer. Math.*, 16 (4), pp. 249-280.
- [14] Chen, Z., & Zou, J. (1999). An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. *SIAM Journal on Control and Optimization*, 37(3), 892-910.
- [15] Chidyagwai, P., & Rivière, B. (2009). On the solution of the coupled Navier-Stokes and Darcy equations. *Computer Methods in Applied Mechanics and Engineering*, 198(47-48), 3806-3820.
- [16] Cocquet, P. H., Rakotobe, M., Ramalingom, D., & Bastide, A. (2021). Error analysis for the finite element approximation of the Darcy-Brinkman-Forchheimer model for porous media with mixed boundary conditions. *Journal of Computational and Applied Mathematics*, 381, 113008.
- [17] Coleman, T. F., & Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization*, 6(2), 418-445.
- [18] Coleman, T. F., & Li, Y. (1994). On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Mathematical programming*, 67(1-3), 189-224.
- [19] Desrayaud, G., Chénier, E., Joulin, A., Bastide, A., Brangeon, B., Caltagirone, J. P., ... & Zoubir, A. (2013). Benchmark solutions for natural convection flows in vertical channels submitted to different open boundary conditions. *International journal of thermal sciences*, 72, 18-33.
- [20] Evgrafov, A. (2005). The limits of porous materials in the topology optimization of Stokes flows. *Applied Mathematics and Optimization*, 52(3), 263-277.
- [21] Garcke, H., Hinze, M., Kahle, C., & Lam, K. F. (2018). A phase field approach to shape optimization in Navier-Stokes flow with integral state constraints. *Advances in Computational Mathematics*, 44, 1345-1383.
- [22] Graham, I., & Sauter, S. (2020). Stability and finite element error analysis for the Helmholtz equation with variable coefficients. *Mathematics of Computation*, 89(321), 105-138.
- [23] Haslinger, J., & Mäkinen, R. A. (2015). On a topology optimization problem governed by two-dimensional Helmholtz equation. *Computational Optimization and Applications*, 62, 517-544.
- [24] Hecht, F. New development in FreeFem++. *J. Numer. Math.* 20 (2012), no. 3-4, 251-265. 65Y15.
- [25] Weller, Henry G., et al.. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics* 12.6 (1998) : 620-631.

- [26] Hinze, M., Pinnau, R., Ulbrich, M., & Ulbrich, S. (2008). Optimization with PDE constraints (Vol. 23). Springer Science & Business Media.
- [27] Logg, A., & Wells, G. N. (2010). DOLFIN : Automated finite element computing. *ACM Transactions on Mathematical Software (TOMS)*, 37(2), 1-28.
- [28] Long, K., Wang, X., & Gu, X. (2018). Multi-material topology optimization for the transient heat conduction problem using a sequential quadratic programming algorithm. *Engineering optimization*, 50(12), 2091-2107.
- [29] Lv, Y., & Liu, S. (2018). Topology optimization and heat dissipation performance analysis of a micro-channel heat sink. *Meccanica*, 53(15), 3693-3708.
- [30] Marck, Gilles, Maroun Nemer, and Jean-Luc Harion. "Topology optimization of heat and mass transfer problems : laminar flow." *Numerical Heat Transfer, Part B : Fundamentals* 63.6 (2013) : 508-539.
- [31] Ramalingom-Barbary, D. (2017). Optimisation topologique des transferts thermiques et massiques dans un canal asymétriquement chauffé (Doctoral dissertation, La Réunion).
- [32] Schiela, A., & Wachsmuth, D. (2013). Convergence analysis of smoothing methods for optimal control of stationary variational inequalities with control constraints. *ESAIM : Mathematical Modelling and Numerical Analysis*, 47(3), 771-787.
- [33] Taylor, C., & Hood, P. (1973). A numerical solution of the Navier-Stokes equations using the finite element technique. *Computers & Fluids*, 1(1), 73-100.
- [34] Temam, R. (2001). Navier-Stokes equations : theory and numerical analysis (Vol. 343). American Mathematical Soc..
- [35] Yedder, R. B., & Bilgen, E. (1991). Natural convection and conduction in Trombe wall systems. *International Journal of Heat and Mass Transfer*, 34(4-5), 1237-1248.
- [36] Wang, G., Wang, D., Liu, A., Dbouk, T., Peng, X., & Ali, A. (2023). Design and performance enhancement of thermal-fluid system based on topology optimization. *Applied Mathematical Modelling*, 116, 168-186.
- [37] Wintermeyer, N., Winters, A. R., Gassner, G. J., & Kopriva, D. A. (2017). An entropy stable nodal discontinuous Galerkin method for the two dimensional shallow water equations on unstructured curvilinear meshes with discontinuous bathymetry. *Journal of Computational Physics*, 340, 200-242.
- [38] Zuo, W., & Saitou, K. (2017). Multi-material topology optimization using ordered SIMP interpolation. *Structural and Multidisciplinary Optimization*, 55, 477-491.
- [39] Babuska, I. M., & Sauter, S. A. (1997). Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wavenumbers?. *SIAM Journal on numerical analysis*, 34(6), 2392-2423.
- [40] Feng, Q., Han, B., & Michelle, M. (2021). Sixth Order Compact Finite Difference Method for 2D Helmholtz Equations with Singular Sources and Reduced Pollution Effect. *arXiv preprint arXiv :2112.07154*.

- [41] Chen, H., Wu, H., & Xu, X. (2015). Multilevel preconditioner with stable coarse grid corrections for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 37(1), A221-A244.
- [42] Cheng, D., Tan, X., & Zeng, T. (2017). A dispersion minimizing finite difference scheme for the Helmholtz equation based on point-weighting. *Computers & Mathematics with Applications*, 73(11), 2345-2359.
- [43] Chen, Z., Cheng, D., & Wu, T. (2012). A dispersion minimizing finite difference scheme and preconditioned solver for the 3D Helmholtz equation. *Journal of Computational Physics*, 231(24), 8152-8175.
- [44] Ciaramella, G., Kwok, F., & Müller, G. (2023). A Nonlinear Optimized Schwarz Preconditioner for Elliptic Optimal Control Problems. In *Domain Decomposition Methods in Science and Engineering XXVI* (pp. 391-398). Cham : Springer International Publishing.
- [45] Ciaramella, G., Halpern, L., & Mechelli, L. (2022). Convergence analysis and optimization of a Robin Schwarz waveform relaxation method for periodic parabolic optimal control problems.
- [46] Cools, S., & Vanroose, W. (2015). Generalization of the complex shifted Laplacian : on the class of expansion preconditioners for Helmholtz problems. *ArXiv e-prints*.
- [47] Cools, S., & Vanroose, W. (2017). On the Optimality of Shifted Laplacian in a Class of Polynomial Preconditioners for the Helmholtz Equation. In *Modern Solvers for Helmholtz Problems* (pp. 53-81). Springer International Publishing.
- [48] Dastour, H., & Liao, W. (2021). A generalized optimal fourth-order finite difference scheme for a 2D Helmholtz equation with the perfectly matched layer boundary condition. *Journal of Computational and Applied Mathematics*, 394, 113544.
- [49] Diwan, G. C., Moiola, A., & Spence, E. A. (2019). Can coercive formulations lead to fast and accurate solution of the Helmholtz equation?. *Journal of Computational and Applied Mathematics*, 352, 110-131.
- [50] Du, Y., & Wu, H. (2015). Preasymptotic error analysis of higher order FEM and CIP-FEM for Helmholtz equation with high wave number. *SIAM Journal on Numerical Analysis*, 53(2), 782-804.
- [51] Erlangga, Y. A., Vuik, C., & Oosterlee, C. W. (2004). On a class of preconditioners for solving the Helmholtz equation. *Applied Numerical Mathematics*, 50(3-4).
- [52] Erlangga, Y. A., Oosterlee, C. W., & Vuik, C. (2006). A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM Journal on Scientific Computing*, 27(4), 1471-1492.
- [53] Ernst, O. G., & Gander, M. J. (2011). Why it is difficult to solve Helmholtz problems with classical iterative methods. *Numerical analysis of multiscale problems*, 325-363.

- [54] Ernst, O. G., & Gander, M. J. (2013). Multigrid methods for Helmholtz problems : a convergent scheme in 1D using standard components. *Direct and Inverse Problems in Wave Propagation and Applications*. De Gruyter, 135-186.
- [55] Gander, M. J., Graham, I. G., & Spence, E. A. (2015). Applying GMRES to the Helmholtz equation with shifted Laplacian preconditioning : what is the largest shift for which wavenumber-independent convergence is guaranteed ?. *Numerische Mathematik*, 131(3), 567-614.
- [56] Gander, M. J., & Lu, L. D. (2023). Dirichlet-Neumann and Neumann-Neumann Methods for Elliptic Control Problems. *arXiv preprint arXiv :2308.12764*.
- [57] Gong, W., Kwok, F., & Tan, Z. (2022). Convergence analysis of the Schwarz alternating method for unconstrained elliptic optimal control problems. *arXiv preprint arXiv :2201.00974*.
- [58] Graham, I., & Sauter, S. (2020). Stability and finite element error analysis for the Helmholtz equation with variable coefficients. *Mathematics of Computation*, 89(321), 105-138.
- [59] Hackbusch, W. (2013). *Multi-grid methods and applications (Vol. 4)*. Springer Science & Business Media.
- [60] Hannukainen, A. (2013). Field of values analysis of a two-level preconditioner for the Helmholtz equation. *SIAM Journal on Numerical Analysis*, 51(3), 1567-1584.
- [61] Hetmaniuk U., Stability estimates for a class of Helmholtz problems, *Communications in Mathematical Sciences*, 5 (3) (2007) 665-678.
- [62] Ihlenburg, F. (Ed.). (1998). *Finite element analysis of acoustic scattering*. New York, NY : Springer New York.
- [63] Krenk, S. (2001). Dispersion-corrected explicit integration of the wave equation. *Computer methods in applied mechanics and engineering*, 191(8-10), 975-987.
- [64] Martin, P. A. (2006). *Multiple scattering : interaction of time-harmonic waves with N obstacles (No. 107)*. Cambridge University Press.
- [65] Melenk, J. M., & Sauter, S. (2011). Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM Journal on Numerical Analysis*, 49(3), 1210-1243.
- [66] A. Reusken. Introduction to multigrid methods for elliptic boundary value problems, *Multiscale Simulation Methods in Molecular Sciences*, p. 467-506, 2009.
- [67] Spence, E. A. (2022). A simple proof that the hp-FEM does not suffer from the pollution effect for the constant-coefficient full-space Helmholtz equation. *arXiv preprint arXiv :2202.06939*.
- [68] Stolk, C. C. (2016). A dispersion minimizing scheme for the 3-D Helmholtz equation based on ray theory. *Journal of computational Physics*, 314, 618-646.

- [69] Sutmann, G. (2007). Compact finite difference schemes of sixth order for the Helmholtz equation, *Journal of Computational and Applied Mathematics*, 203, 15-31.
- [70] Umetani, N., MacLachlan, S. P., & Oosterlee, C. W. (2009). A multigrid-based shifted Laplacian preconditioner for a fourth-order Helmholtz discretization. *Numerical Linear Algebra with Applications*, 16(8), 603-626.
- [71] Vieira A., Cocquet P.H. (2023) Optimized Schwarz method for coupled direct-adjoint problems applied to parameter identification in advection-diffusion equation. Accepted in Proceedings of the conference Domain Decomposition Methods in Science and Engineering XXVII. Springer. Lien HAL : <https://hal.science/hal-03992586/document>
- [72] Wu, T., & Xu, R. (2018). An optimal compact sixth-order finite difference scheme for the Helmholtz equation. *Computers & Mathematics with Applications*, 75(7), 2520-2537.
- [73] Xu, X., Chow, C. O., & Lui, S. H. (2005). On nonoverlapping domain decomposition methods for the incompressible Navier-Stokes equations. *ESAIM : Mathematical Modelling and Numerical Analysis*, 39(6), 1251-1269.