

Université de Montréal

**Étude des discrétisations superconsistantes et
application à la résolution numérique
d'équations d'advection-diffusion**

par

François De l'Isle

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en mathématiques

28 décembre 2017

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**Étude des discrétisations superconsistantes et
application à la résolution numérique
d'équations d'advection-diffusion**

présenté par

François De l'Isle

a été évalué par un jury composé des personnes suivantes :

Anne Bourlioux

(président-rapporteur)

Robert G. Owens

(directeur de recherche)

Jean-Christophe Nave

(membre du jury)

Mémoire accepté le

30 mars 2018

SOMMAIRE

Dans ce mémoire, nous reprenons les travaux de Fatone, Funaro et al. [10] sur les discrétisations superconsistantes. Nous établissons d'abord une définition précise de la superconsistance et son lien avec la consistance. On élabore aussi une méthode explicite de construction de la discrétisation pour une vaste gamme d'opérateurs linéaires puis on applique la méthode sur l'opérateur d'advection-diffusion que l'on retrouve dans l'équation du même nom et dans les équations de Navier-Stokes. Une étude analytique de l'opérateur discret résultant est d'abord faite puis des tests numériques sont effectués et analysés. Les problèmes résolus sont choisis tels que le terme diffusif est négligeable devant le terme advectif ce qui crée typiquement une couche limite dans la solution. Pour ces problèmes, les méthodes par différence finie centrée classiques risquent de générer des oscillations non désirées proche des couches limites alors que les solutions données par la méthode superconsistante s'avèrent beaucoup plus stables. Dans ce travail, on se limite à la dimension 1 et 2 car les mêmes simulations en dimension 3 demanderaient trop de ressources de calcul.

Mots clés : Différences finies, Méthode numérique, Équation d'advection-diffusion, Navier-Stokes, Consistance, Superconsistance, Couche limite

SUMMARY

This thesis develops the work of Fatone, Funaro et al. [10] on superconsistent discretizations. We first give a precise definition of superconsistency and how it relates to consistency. We then suggest an explicit approach to constructing a superconsistent scheme from a linear operator and we apply the method to the advection-diffusion operator which we find, for example, in advective diffusive problems and the Navier Stokes equations. Both an analytical study and some numerical tests are provided. The choice of problems is made so that the diffusive term is small compared to the advective term. Solutions to such problems typically contain a boundary layer and classical centered finite difference methods may display some spurious oscillations in and near that layer. In contrast, solutions obtained with the superconsistent method are far more stable. In this work, only problems in one and two dimensions are treated since three dimensional problems would be too demanding on computational resources.

Keywords : Finite differences, Numerical method, Advection-diffusion equation, Navier-Stokes, Consistency, Superconsistency, Boundary layer

TABLE DES MATIÈRES

| | |
|---|------|
| Sommaire | v |
| Summary | vii |
| Liste des tableaux | xiii |
| Liste des figures | xvii |
| Introduction | 1 |
| Chapitre 1. Définitions et cadre d'étude | 5 |
| 1.1. Notation | 5 |
| 1.2. Discrétisation d'une équation différentielle en dimension 1 | 6 |
| 1.2.1. Consistance, convergence et stabilité | 7 |
| 1.2.2. Méthode de construction de L^h | 9 |
| 1.2.3. Superconsistance | 10 |
| 1.2.3.1. Exemple simple de superconsistance | 10 |
| 1.2.3.2. Note sur les travaux originaux | 12 |
| 1.3. Discrétisation d'une équation différentielle en dimension 2 | 13 |
| 1.3.1. Consistance, convergence et stabilité | 14 |
| 1.3.2. Méthode de construction de L^h | 14 |
| Chapitre 2. Application à des problèmes simples en une dimension | 17 |
| 2.1. Méthode de construction du schéma superconsistant | 17 |
| 2.2. Construction du schéma de dérivée première superconsistant .. | 18 |
| 2.2.1. Analyse de la consistance | 19 |
| 2.3. Construction du schéma de dérivée seconde superconsistant ... | 19 |
| 2.3.1. Analyse de la consistance | 20 |
| 2.4. Problème d'advection-diffusion en dimension 1 | 20 |

| | | |
|--------------------|--|-----------|
| 2.4.1. | Analyse de la consistance du schéma | 22 |
| 2.4.2. | Analyse de la stabilité du schéma..... | 23 |
| 2.4.3. | Exemple numérique..... | 24 |
| 2.4.3.1. | Calcul de l'ordre de convergence..... | 25 |
| 2.4.3.2. | Comparaison des résultats..... | 26 |
| Chapitre 3. | Application à un problème stationnaire d'advection-diffusion en dimension 2 | 31 |
| 3.1. | Dérivation de l'équation d'advection-diffusion..... | 31 |
| 3.1.1. | Théorème de la divergence | 31 |
| 3.1.2. | Équation de continuité..... | 32 |
| 3.1.3. | Équation d'advection-diffusion..... | 32 |
| 3.2. | Présentation du problème..... | 34 |
| 3.3. | Construction du schéma superconsistant | 34 |
| 3.4. | Analyse de la consistance du schéma..... | 36 |
| 3.5. | Deuxième approche (Fatone et al.)..... | 37 |
| 3.6. | Comparaison de la localisation des points de collocation | 39 |
| 3.7. | Analyse de la stabilité des schémas De l'Isle-Owens et Fatone et al..... | 41 |
| 3.7.1. | Stabilité dans le cas particulier $\beta_1 = 0$ (ou $\beta_2 = 0$) | 42 |
| 3.7.2. | Stabilité de la méthode de Fatone et al..... | 43 |
| 3.7.3. | Stabilité de la méthode De l'Isle-Owens..... | 45 |
| 3.8. | Algorithme et test numérique..... | 46 |
| 3.8.1. | Construction du problème de test..... | 46 |
| 3.8.2. | Résultats numériques d'un problème avec solution connue.. | 49 |
| 3.8.3. | Résultats numériques d'un problème avec solution exacte inconnue | 55 |
| Chapitre 4. | Application à un problème de Navier-Stokes ... | 63 |
| 4.1. | Postulats et principes généraux de mécanique des fluides..... | 63 |
| 4.1.1. | Mécanique des milieux continus..... | 63 |
| 4.1.2. | Viscosité et fluide newtonien..... | 64 |
| 4.1.3. | Fluide incompressible..... | 64 |

| | | |
|----------|---|-----|
| 4.2. | Dérivation des équations de Navier-Stokes | 65 |
| 4.2.1. | Représentation lagrangienne | 65 |
| 4.2.2. | Conservation de la masse | 65 |
| 4.2.3. | Conservation de la quantité de mouvement | 65 |
| 4.2.4. | Équations de Navier-Stokes pour un fluide newtonien incompressible | 66 |
| 4.2.5. | Nombre de Reynolds | 67 |
| 4.2.6. | Problème de cavité entraînée | 67 |
| 4.3. | Équations du problème | 68 |
| 4.3.1. | Méthodes de projection | 69 |
| 4.3.1.1. | Choix de θ | 71 |
| 4.4. | Discrétisation des équations | 73 |
| 4.5. | Analyse des résultats numériques | 76 |
| 4.5.1. | Présentation des résultats | 76 |
| 4.5.1.1. | Re=100 | 78 |
| 4.5.1.2. | Re=400 | 81 |
| 4.5.1.3. | Re=1000 | 83 |
| 4.5.1.4. | Re=3200 | 85 |
| 4.5.1.5. | Re=5000 | 87 |
| 4.5.1.6. | Ordre de convergence | 89 |
| 4.5.2. | Discussion des résultats | 90 |
| | Conclusion | 93 |
| | Annexe A. Développement en série de Taylor | A-i |
| | Annexe B. Inégalités discrètes | B-i |
| | Annexe C. Identités sommation | C-i |
| | Annexe D. Résultats numériques supplémentaires du chapitre 2 | D-i |
| | Annexe E. Résultats numériques supplémentaires du chapitre 3 | E-i |
| | Bibliographie | E-i |

LISTE DES TABLEAUX

| | | |
|-----|--|----|
| 2.1 | Comparaison des erreurs pour $\epsilon = 0.1$ avec $\beta = 1$ et $f = 1 \dots \dots$ | 26 |
| 2.2 | Comparaison des erreurs pour $\epsilon = 0.001$ avec $\beta = 1$ et $f = 1 \dots \dots$ | 27 |
| 2.3 | Comparaison des erreurs pour $\epsilon = 0.00001$ avec $\beta = 1$ et $f = 1 \dots$ | 27 |
| 2.4 | Comparaison des ordres de convergence des solutions discrètes pour différents ϵ avec $\beta = 1$ et $f = 1 \dots \dots \dots \dots \dots \dots$ | 28 |
| 3.1 | Comparaison des points de collocation pour $\epsilon = 1/100$, $\ \vec{\beta}\ = 1$, $h_x = h_y = 1/10 \dots \dots \dots \dots \dots \dots$ | 40 |
| 3.2 | Comparaison des points de collocation pour $\epsilon = 1/1000$, $\ \vec{\beta}\ = 1$, $h_x = h_y = 1/10 \dots \dots \dots \dots \dots \dots$ | 40 |
| 3.3 | Comparaison des points de collocation pour $\epsilon = 1/100$, $\ \vec{\beta}\ = 10$, $h_x = h_y = 1/10 \dots \dots \dots \dots \dots \dots$ | 40 |
| 3.4 | Comparaison des points de collocation pour $\epsilon = 1/100$, $\ \vec{\beta}\ = 10$, $h_x = h_y = 1/100 \dots \dots \dots \dots \dots \dots$ | 40 |
| 3.5 | Comparaison des erreurs $\ U - u^h\ $ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.1$, $\beta_1 = 1$, $\beta_2 = 0 \dots$ | 50 |
| 3.6 | Comparaison des erreurs $\ U - u^h\ $ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = 1$, $\beta_2 = 0 \dots$ | 51 |
| 3.7 | Comparaison des erreurs $\ U - u^h\ $ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 1e - 005$, $\beta_1 = 1$, $\beta_2 = 0$ | 51 |
| 3.8 | Ordre de convergence de l'erreur en fonction de $h = \frac{1}{n}$ (calculé comme la pente du graphique log-log) pour chacune des deux méthodes avec $\beta_1 = 1$, $\beta_2 = 0 \dots \dots \dots \dots \dots \dots$ | 51 |
| 3.9 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024) \dots \dots \dots \dots \dots \dots$ | 57 |

| | | |
|------|---|----|
| 3.10 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$ | 58 |
| 3.11 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$ | 59 |
| 3.12 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$ | 60 |
| 3.13 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$ | 61 |
| 3.14 | Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$ | 62 |
| 4.1 | Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode de De l'Isle-Owens pour $Re=100$ | 78 |
| 4.2 | Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode de Fatone et al. pour $Re=100$ | 78 |
| 4.3 | Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode centrée pour $Re=100$ | 79 |
| 4.4 | Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 100$, $n = 32$, $\Delta t = 0.00625$ | 79 |
| 4.5 | Valeur de $\ \vec{U}_G - \vec{U}\ _2 + \ \vec{V}_G - \vec{V}\ _2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 400$, $\Delta t = 0.00625$ | 81 |
| 4.6 | Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 400$, $n = 32$, $\Delta t = 0.00625$ | 82 |
| 4.7 | Valeur de $\ \vec{U}_G - \vec{U}\ _2 + \ \vec{V}_G - \vec{V}\ _2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 1000$, $\Delta t = 0.00625$... | 83 |
| 4.8 | Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 1000$, $n = 32$, $\Delta t = 0.00625$ | 83 |

| | | |
|------|--|----|
| 4.9 | Valeur de $\ \vec{U}_G - \vec{U}\ _2 + \ \vec{V}_G - \vec{V}\ _2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 3200$, $\Delta t = 0.00625$... | 85 |
| 4.10 | Localisation du centre des vortex Primaire-BL-BR-T (de haut en bas) et valeur de Φ en ces points pour $Re = 3200$, $n = 32$, $\Delta t = 0.00625$ | 85 |
| 4.11 | Valeur de $\ \vec{U}_G - \vec{U}\ _2 + \ \vec{V}_G - \vec{V}\ _2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 5000$, $\Delta t = 0.00625$... | 87 |
| 4.12 | Localisation du centre des vortex Primaire-BL-BR-T (de haut en bas) et valeur de Φ en ces points pour $Re = 5000$, $n = 32$, $\Delta t = 0.00625$ | 87 |
| 4.13 | Ordre de convergence de l'erreur en fonction de h | 89 |

LISTE DES FIGURES

| | | |
|-----|---|----|
| 2.1 | Graphique de la solution exacte (2.4.11) pour $\epsilon = 1/10$ (bleu), $\epsilon = 1/100$ (vert), $\epsilon = 1/1000$ (rouge)..... | 25 |
| 2.2 | Graphique à l'échelle log-log de l'erreur en fonction de n pour $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001, 0.00001$ (lecture de gauche à droite puis de haut en bas)..... | 29 |
| 2.3 | Comparaison entre les solutions discrètes et la solution exacte pour $\epsilon = 0.01$ et $n = 16, 32$ (de gauche à droite) | 30 |
| 2.4 | Comparaison entre les solutions discrètes et la solution exacte sur l'intervalle $x \in [0.75, 1]$ pour $\epsilon = 0.001$ et $n = 64, 128$ (de gauche à droite) | 30 |
| 2.5 | Comparaison entre les solutions discrètes et la solution exacte sur l'intervalle $x \in [0.9, 1]$ pour $\epsilon = 0.00001$ et $n = 64, 128$ (de gauche à droite) | 30 |
| 3.1 | Localisation "upwind" du point de collocation dans la méthode de Fatone et al..... | 38 |
| 3.2 | Graphique de la solution exacte (3.8.8) avec $\epsilon = 0.1$ (gauche) et $\epsilon = 0.01$ (droite)..... | 49 |
| 3.3 | Graphique log-log de l'évolution de l'erreur $\ U - u^h\ $ en fonction de $h = \frac{1}{n}$ pour les 2 méthodes avec $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001, 0.00001$ (de gauche à droite puis de haut en bas), $\beta_1 = 1$ et $\beta_2 = 0$ | 52 |
| 3.4 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 0.01, \beta_1 = 1, \beta_2 = 0$ | 53 |
| 3.5 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 0.001, \beta_1 = 1, \beta_2 = 0$ | 53 |

| | | |
|------|--|----|
| 3.6 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 1e - 4$, $\beta_1 = 1$, $\beta_2 = 0$ | 54 |
| 3.7 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 1e - 5$, $\beta_1 = 1$, $\beta_2 = 0$ | 54 |
| 3.8 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 64, 128$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$ | 57 |
| 3.9 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 64, 128$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$ | 58 |
| 3.10 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$ | 59 |
| 3.11 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$ | 60 |
| 3.12 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$ | 61 |
| 3.13 | Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$ | 62 |
| 4.1 | Domaine et conditions frontières pour le problème de cavité entraînée. | 68 |

| | | |
|------|--|----|
| 4.2 | Exemple de grille de représentation uniforme décalée. Les points verts et bleus constituent la grille de U^k et de V^k respectivement. | 73 |
| 4.3 | Solution minimale X de l'équation (4.4.2) en fonction de β_1 pour $Re = 100$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite) | 75 |
| 4.4 | Solution minimale X de l'équation (4.4.2) en fonction de β_1 pour $Re = 5000$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite) | 75 |
| 4.5 | Valeur de $X = \alpha\beta_1$ tel que α est la solution minimale de l'équation (4.4.4) et en fonction de β_1 et β_2 . Les autres paramètres sont fixés : $Re = 5000$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite) | 76 |
| 4.6 | Lignes de courant de la solution en régime permanent pour $Re=1000$ (gauche) et $Re=3200$ (droite). Ces images sont tirées de Yang et Fu [11] | 77 |
| 4.7 | Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 100$, $\Delta t = 0.00625$ et $n = 32$ | 79 |
| 4.8 | Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 100$, $n = 32$ et pour la méthode De l'Isle-Owens. La localisation des centres de référence (Ghia) est en bleu | 80 |
| 4.9 | Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 400$, $\Delta t = 0.00625$ et $n = 32$ | 81 |
| 4.10 | Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 400$, $n = 32$ et pour les méthodes De l'Isle Owens (haut),et centrée (bas). La localisation des centres de référence (Ghia) est en bleu | 82 |
| 4.11 | Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 1000$, $\Delta t = 0.00625$ et $n = 32$ | 83 |
| 4.12 | Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 1000$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu | 84 |

| | | |
|------|---|------|
| 4.13 | Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 3200$, $\Delta t = 0.00625$ et $n = 32$ | 85 |
| 4.14 | Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 3200$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu | 86 |
| 4.15 | comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 5000$, $\Delta t = 0.00625$ et $n = 32$ | 87 |
| 4.16 | Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 5000$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu | 88 |
| 4.17 | Graphique log-log de $e_U = \ \vec{U}_G - \vec{U}\ _2$ (gauche) et $e_V = \ \vec{V}_G - \vec{V}\ _2$ (droite) pour $Re = 100, 400, 1000$ (de haut en bas) et pour les 3 méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C) | 89 |
| D.1 | Graphique log-log de l'erreur en fonction de n pour $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001, 0.00001$ (lecture de gauche à droite puis de haut en bas). Le problème résolu est l'équation d'advection-diffusion (2.4.1) avec membre de droite $f = x$ et $u = 0$ sur la frontière. La solution exacte est donnée par (2.4.14) | D-i |
| E.1 | $\epsilon = 0.001$, $r = 1$, $n_x = 30$, $n_y = 30$ | E-i |
| E.2 | $\epsilon = 0.001$, $r = 5$, $n_x = 30$, $n_y = 30$ | E-i |
| E.3 | $\epsilon = 0.001$, $r = 5$, $n_x = 50$, $n_y = 50$ | E-ii |
| E.4 | $\epsilon = 0.001$, $r = 10$, $n_x = 30$, $n_y = 30$ | E-ii |
| E.5 | $\epsilon = 0.001$, $r = 10$, $n_x = 50$, $n_y = 50$ | E-ii |

INTRODUCTION

Pour n'en nommer que quelques-uns, l'étude d'une relation de proie-prédateur, le calcul de la diffusion de la chaleur dans une barre de fer ou la simulation d'une réaction chimique par un système de réaction-diffusion sont des problèmes pratiques de divers domaines de la science que l'on peut décrire par des modèles mathématiques. Ces modèles sont constitués d'équations différentielles (EDO) ou aux dérivées partielles (EDP) et on voudrait résoudre ces équations pour trouver les solutions de ces problèmes sous forme de fonctions.

Des techniques de résolution analytique ont été développées pour des cas particuliers d'EDOs et d'EDPs et sur des domaines simples [9], mais en général ces techniques échouent pour des problèmes plus généraux ou lorsque la géométrie du domaine est compliquée. Si la résolution exacte n'est pas envisageable, on peut cependant faire une étude analytique des modèles afin par exemple de déterminer si une solution existe, si elle est unique et si elle dépend de façon continue aux données du problème. Ces analyses ne donnent en général aucune information sur la forme de la solution et on doit donc avoir recours à des résolutions numériques.

À partir des équations analytiques, on construit alors un modèle discret dont les solutions sont des approximations de la valeur réelle de la solution en différents points du domaine. Les 3 méthodes numériques les plus couramment utilisées sont les méthodes de différences finies (FDM), les méthodes de volumes finis (FVM) et les méthodes d'éléments finis (FEM). Ce mémoire traitera de la première méthode et d'une nouvelle approche pour construire les modèles discrets.

Les FDM utilisent la théorie des développements limités en série de Taylor pour approximer les dérivées qui composent les EDOs et les EDPs. Prenons par exemple $u(x)$, une fonction infiniment dérivable sur l'intervalle $(x, x + h)$. L'expansion suivante de $u(x + h)$ autour de x :

$$u(x + h) = u(x) + hu'(x) + \frac{h^2}{2}u''(\xi) \text{ avec } \xi \in (x, x + h) \quad (0.0.1)$$

peut être réécrite de la façon suivante :

$$\frac{u(x+h) - u(x)}{h} = u'(x) + \frac{h}{2}u''(\xi) \text{ avec } \xi \in (x, x+h) \quad (0.0.2)$$

En supposant que le domaine de la fonction est subdivisé en plusieurs points équidistants séparés de $h > 0 : x_0, x_0 + h, x_0 + 2h \dots$, alors de l'équation (0.0.2) découle une formule qui approxime $u'(x_i)$ en utilisant les 2 points x_i et $x_i + h$.

$$\frac{u(x_i+h) - u(x_i)}{h} = u'(x_i) + \frac{h}{2}u''(\xi_i) \text{ avec } \xi_i \in (x_i, x_i+h) \quad (0.0.3)$$

De (0.0.3) on remarque que lorsque h diminue, c'est à dire lorsqu'on utilise plus de points sur le domaine et que ceux ci se rapprochent localement, le terme supplémentaire au membre de droite tend vers 0 et donc

$$\frac{u(x_i+h) - u(x_i)}{h} \rightarrow u'(x_i) \text{ lorsque } h \rightarrow 0$$

Cette propriété selon laquelle l'équation discrétisée se "rapproche" de l'équation exacte lorsqu'on raffine le maillage est la notion de consistance et la vitesse à laquelle le terme d'erreur disparaît est l'ordre de consistance. Dans l'équation (0.0.3), la consistance est d'ordre 1 car le terme d'erreur $\frac{h}{2}u''(\xi_i)$ tend vers 0 à la même vitesse que h . Dans le schéma suivant qui utilise 3 points, la consistance est d'ordre 2 car le terme d'erreur tend vers 0 à la vitesse de h^2 :

$$\frac{-3u(x_i) + 4u(x_i+h) - u(x_i+2h)}{2h} = u'(x_i) - \frac{h^2}{3}u''(\xi_i) \text{ avec } \xi_i \in (x_i, x_i+2h) \quad (0.0.4)$$

Avoir un meilleur ordre de consistance implique qu'un raffinement du maillage aura un plus grand effet sur la précision de l'approximation de l'équation. Par exemple, dans le cas de la discrétisation d'ordre 1 (0.0.3), diviser h par 2 aura pour effet de diviser le terme dominant de l'erreur par 2 alors que pour la discrétisation d'ordre 2 (0.0.4), le terme dominant sera divisé par 4.

Les discrétisations qui utilisent plus de points auront en général un meilleur ordre de consistance, mais en contrepartie elles demanderont plus de temps de calcul et l'implémentation des conditions frontières sera moins aisée.

Les articles [10, 12, 13] présentent une méthode pour augmenter l'ordre de consistance d'un schéma et introduit ainsi les schémas superconsistants. L'essence de la méthode réside dans le choix du point de collocation, c'est à dire le point où est évaluée l'équation différentielle. On pourra vérifier par exemple qu'en modifiant le point où est évaluée la dérivée dans (0.0.3) par $x_i + h/2$, on obtient un schéma d'ordre 2 :

$$\frac{u(x_i + h) - u(x_i)}{h} = u'(x_i + h/2) + \frac{h^2}{24}u'''(\xi_i) \text{ avec } \xi_i \in (x_i, x_i + h) \quad (0.0.5)$$

Pour des opérateurs simples, le schéma superconsistant peut coïncider avec des schémas connus et ainsi dans l'exemple précédent on reconnaît la discrétisation par différences centrées. L'intérêt de la méthode devient plus apparent lorsque l'opérateur contient plusieurs ordres de dérivée car les schémas résultants diffèrent de ceux que l'on retrouve couramment dans la littérature. En plus d'améliorer la consistance, on verra que la modification du point de collocation permet aussi de diminuer considérablement la présence d'oscillations non désirées dans les solutions numériques ce qui indique une meilleure stabilité des schémas. Ce mémoire est donc une extension des travaux de Fatone et al. [10] et a pour but d'approfondir la théorie des méthodes superconsistantes qui consistent à optimiser la localisation du point de collocation en vue d'améliorer la qualité d'un schéma numérique.

Pour étudier la méthode, nous construirons explicitement des schémas superconsistants à partir d'équations différentielles et aux dérivées partielles. Les discrétisations obtenues, en plus d'être analysées sur le plan théorique, seront aussi implémentées dans le logiciel de calcul MATLAB afin de générer les solutions numériques de quelques problèmes choisis.

À priori, le choix des équations qui serviront à tester la méthode est arbitraire, mais nous avons décidé d'utiliser le modèle d'advection-diffusion dont la résolution numérique peut s'avérer délicate dû à la présence d'une ou plusieurs couches limites.

En une dimension, un exemple d'une telle équation est :

$$-\epsilon u_{xx} + u_x = 1 \quad (0.0.6)$$

Pour des conditions frontières de Dirichlet homogènes, une discontinuité se crée proche de la frontière lorsque $\epsilon \rightarrow 0$. Pour des valeurs de ϵ petites mais non nulles, on observe alors une couche limite où la solution change très rapidement. Typiquement, si la méthode numérique utilisée n'est pas adaptée, la solution trouvée risque de présenter des oscillations non désirées proche de cette couche limite et ce, même pour des maillages très fins [16].

On pourra constater que la méthode de discrétisation superconsistante est naturellement performante lorsqu'on l'applique à la résolution d'équations de ce type car les solutions présentent très peu d'oscillations.

Dans le premier chapitre, on détaille le cadre mathématique général dans lequel nous développons la théorie. On donne les définitions des propriétés de

consistance, stabilité et convergence qui sont les trois propriétés principales que l'on cherche à obtenir et améliorer lorsqu'on construit un schéma numérique. On définit ensuite précisément la propriété de superconsistance et le lien avec la consistance. La théorie est développée d'abord en dimension 1 puis est adaptée en dimension 2.

Dans le deuxième chapitre on commence par décrire une technique générale pour construire un schéma superconsistant à partir d'une équation différentielle (une dimension). On illustre ensuite cette technique en l'appliquant sur des opérateurs simples de dérivée première et de dérivée seconde puis on termine par une étude d'un exemple numérique où on utilise la méthode pour résoudre un problème de la forme (0.0.6) avec conditions aux bords. Nous comparons ces solutions avec la solution exacte ainsi qu'à celles obtenues par d'autres méthodes classiques.

Aux deux derniers chapitres, on applique la méthode en 2 dimensions ce qui permet de donner une interprétation physique plus intéressante aux modèles. Nous testons la méthode superconsistante sur le problème d'advection-diffusion à nombre de Peclet élevé (chapitre 3) et sur les équations de Navier-Stokes pour un fluide visqueux incompressible et à nombre de Reynolds élevé (chapitre 4). Ces deux problèmes sont régis par des équations présentant un terme diffusif négligeable par rapport au terme advectif et donc d'une forme analogue à (0.0.6) en deux dimensions. Dans chaque cas, nous donnons le raisonnement physique permettant de poser les équations du modèle et nous générons plusieurs solutions numériques que nous comparons avec celles obtenues par d'autres méthodes. Encore une fois, nous verrons que la méthode superconsistante est stable malgré la présence d'une couche limite. L'analyse théorique de la discrétisation ne sera faite qu'au chapitre 3.

Chapitre 1

DÉFINITIONS ET CADRE D'ÉTUDE

1.1. NOTATION

| | |
|--|--|
| \mathcal{X}, \mathcal{Y} | : espaces vectoriels de fonctions |
| $\mathcal{V} \subset \mathbb{R}^n, \mathcal{W} \subset \mathbb{R}^m$ | : espaces vectoriels de vecteurs colonnes |
| $\vec{x} \in \mathcal{V}$ | : grille de représentation des solutions |
| $\vec{z} \in \mathcal{W}$ | : grille de collocation de l'EDO ou EDP |
| z_i | : point de collocation en 1 dimension |
| (τ_i, μ_j) | : point de collocation en 2 dimensions |
| $L : \mathcal{X} \rightarrow \mathcal{Y}$ | : opérateur différentiel linéaire |
| $L^h : \mathcal{V} \rightarrow \mathcal{W}$ | : matrice représentant la discrétisation de L |
| $f \in \mathcal{Y}$ | : membre de droite de l'EDO ou de l'EDP |
| $u, v \in \mathcal{X}$ | : solutions exacte de l'EDO ou de l'EDP |
| $U, V \in \mathcal{V}$ | : approximations des solutions u et v respectivement sur \vec{x} |
| $u^h \in \mathcal{V}$ | : projection de u sur \vec{x} |
| $f^h \in \mathcal{W}$ | : projection de f sur \vec{z} |

1.2. DISCRÉTISATION D'UNE ÉQUATION DIFFÉRENTIELLE EN DIMENSION 1

Soit \mathcal{X}, \mathcal{Y} deux espaces vectoriels, $L : \mathcal{X} \rightarrow \mathcal{Y}$ un opérateur différentiel linéaire inversible et $u \in \mathcal{X}$. Soit $\Omega = [a, b] \subset \mathbb{R}$ et $u : \Omega \rightarrow \mathbb{R}$, on considère l'équation générale

$$Lu = f \tag{1.2.1}$$

où $f : I = [c, d] \rightarrow \mathbb{R}$ est une fonction donnée. La solution exacte u est souvent très difficile à trouver, alors on veut calculer une valeur approchée de la solution en différents points de l'espace.

Soit $n \in \mathbb{N}$, on construit une subdivision de Ω à $n + 1$ points :

$$a = x_0 < x_1 < \dots < x_n = b$$

C'est la **grille de représentation** et on la note $\vec{x} = \{x_i\}_{i=0,1,\dots,n}$. On appelle le **pas** h la distance entre 2 points et, sauf indications contraires, on le choisira constant c.-à-d. $h = x_{i+1} - x_i \quad \forall i = 0, \dots, n - 1$.

On construit aussi une subdivision de I à $m + 1$ points.

$$c = z_0 < z_1 < \dots < z_m = d$$

C'est la **grille de collocation** et on la note $\vec{z} = \{z_i\}_{i=0,1,\dots,m}$

On note u^h la restriction de u sur la grille de représentation et f^h la restriction de f sur la grille de collocation. Soit $u_i := u(x_i)$ et $f_i := f(z_i)$, on a donc $u^h = (u_0, \dots, u_n)$ et $f^h = (f_0, \dots, f_m)$.

L'indice h en exposant indique que l'objet mathématique est discret et dépend du pas h .

Soit $L^h : \mathcal{V} \subset \mathbb{R}^{n+1} \rightarrow \mathcal{W} \subset \mathbb{R}^{m+1}$ une discrétisation de L , on considère le système linéaire discret associé à (1.2.1) :

$$L^h U = f^h \tag{1.2.2}$$

L'équation (1.2.2) est aussi appelée le **schéma numérique** associé à (1.2.1). L^h est une matrice de dimension $(m+1) \times (n+1)$ et le vecteur $U = (U_0, U_1, \dots, U_n) \in \mathcal{V}$ est la solution du système discret et est une approximation de u^h .

Il est important de distinguer les vecteurs $(Lu)^h$ et $L^h u^h$. Le premier résulte de l'évaluation de Lu sur \vec{z} et vaut f^h par (1.2.1). Par (1.2.2) on a donc $(Lu)^h = L^h U$. Le deuxième provient de l'application de l'opérateur discrétisé sur u^h et est une combinaison linéaire des éléments u_i de u^h qui approxime f^h .

Ainsi $U_i \approx u_i$ et $(L^h u^h)_i \approx f_i$ et un bon schéma doit minimiser l'erreur sur ces deux approximations.

1.2.1. Consistance, convergence et stabilité

Définition 1.2.1. Soit u la solution exacte de (1.2.1) et U la solution exacte de (1.2.2) le **résidu** au point z_i est

$$r_i = (L^h u^h)_i - (Lu)(z_i) = (L^h u^h)_i - f_i$$

et l'**erreur d'approximation** au point x_i est

$$e_i = U_i - u_i$$

Les normes du vecteur résidu et du vecteur erreur sont alors respectivement

$$r = \|L^h u^h - f^h\|_{\mathcal{W}} \text{ et } e = \|U - u^h\|_{\mathcal{V}}$$

où $\|\cdot\|_{\mathcal{W}} : \mathcal{W} \rightarrow [0, \infty]$, $\|\cdot\|_{\mathcal{V}} : \mathcal{V} \rightarrow [0, \infty]$ sont des normes discrètes

On appellera par la suite r le résidu global et e l'erreur globale.

Remarque 1.2.1. Dans ce travail, sauf indication contraire on utilisera la norme 2 pour les vecteurs dans \mathbb{R}^n

$$\|U\|_2 = \sqrt{\sum_{i=1}^n U_i^2}$$

et on la notera simplement $\|\cdot\|$. La norme induite d'une matrice $A \in \mathbb{R}^{m \times n}$ est alors

$$\|A\| = \max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|}$$

On donne maintenant les définitions des 3 propriétés principales d'un schéma numérique soit la consistance, la convergence et la stabilité.

Définition 1.2.2. Soit u la solution du problème continu (1.2.1) et U la solution du problème discret (1.2.2) avec $h > 0$ le pas du maillage, le schéma (1.2.2) est **consistant** si

$$r \rightarrow 0 \quad \text{lorsque} \quad h \rightarrow 0$$

où r est le résidu global. Il est consistant d'ordre k si $r = O(h^k)$ où nous utilisons la notation grand ordre de Landau c.-à-d.

$$r = O(h^k) \implies \exists C, H > 0 \text{ tel que } r \leq Ch^k \text{ lorsque } h \leq H$$

Définition 1.2.3. Soit u la solution du problème continu (1.2.1) et U la solution du problème discret (1.2.2) avec $h > 0$ le pas du maillage, le schéma (1.2.2) est

convergent si

$$e \rightarrow 0 \quad \text{lorsque} \quad h \rightarrow 0$$

où e est l'erreur globale. Il est convergent d'ordre k si $e = O(h^k)$.

Définition 1.2.4. Soit L^h la matrice de (1.2.2), la discrétisation (1.2.2) est **stable** s'il existe une constante $C > 0$ indépendante de h telle que

$$\|(L^h)^{-1}\| \leq C \quad \forall h$$

Ces 3 propriétés sont liées par la proposition suivante

Proposition 1.2.1. (Lax) *Un schéma numérique consistant et stable est convergent.*

DÉMONSTRATION. Soit e et r l'erreur et le résidu global respectivement

$$e = \|U - u^h\| = \|(L^h)^{-1}f^h - (L^h)^{-1}L^h u^h\| \leq \|(L^h)^{-1}\| \|f^h - L^h u^h\| \leq Cr$$

où la dernière inégalité n'est vraie que pour un schéma stable. \square

On remarque de la dernière inégalité que l'ordre de convergence est au moins aussi élevé que l'ordre de consistance. En pratique, puisqu'on ne connaît généralement pas la solution exacte du problème, il est difficile d'étudier théoriquement l'ordre de convergence de la solution discrète. On tente plutôt d'étudier et d'améliorer la consistance de la discrétisation. Par la proposition précédente, si le schéma est stable, on est assuré d'améliorer du même coup la convergence. Pour étudier la stabilité d'un schéma, on se servira du résultat suivant.

Proposition 1.2.2. *Un schéma est stable si et seulement si pour tout membre de droite f il existe une constante C_f indépendante de h telle que*

$$\|U\| \leq C_f \|f^h\|$$

DÉMONSTRATION. (\implies) : On a $U = (L^h)^{-1}f^h$ donc $\|U\| = \|(L^h)^{-1}f^h\| \leq \|(L^h)^{-1}\| \|f^h\| \leq C \|f^h\|$

(\impliedby) : On a $U = (L^h)^{-1}f^h$ donc $\|(L^h)^{-1}f^h\| \leq C_f \|f^h\|$. La dernière inégalité est vrai pour tout membre de droite f^h en particulier pour celui qui maximise $\|(L^h)^{-1}f^h\|$ et le résultat découle de la définition de la norme matricielle. \square

1.2.2. Méthode de construction de L^h

Soit L un opérateur différentiel d'ordre s , supposons que l'on cherche la valeur de $Lu(z_i)$ pour un certain $z_i \in \vec{z}$. On commence par choisir un ensemble de $k + 1$ points de \vec{x} avec $k \geq s$.

$$X_{0_i} < X_{1_i} \cdots < X_{k_i}$$

On construit P_i le polynôme d'interpolation locale de degré k tel que $P_i(X_{j_i}) = u(X_{j_i})$, $j = 0, \dots, k$. L'indice i indique que pour chaque point de collocation z_i le polynôme d'interpolation peut utiliser un ensemble de points différents. On va désormais omettre les indices i sur les points X par soucis de clarté. On peut écrire P_i dans la base des polynômes de Lagrange comme

$$P_i(x) = \sum_{j=0}^k l_j(x)u(X_j)$$

où

$$l_j(x) = \prod_{z=0, z \neq j}^k \frac{x - X_z}{X_j - X_z} = \frac{1}{h^k} \prod_{z=0, z \neq j}^k \frac{x - X_z}{j - z} \quad (1.2.3)$$

Ensuite on évalue l'action de L sur ce polynôme au point z_i

$$(L^h u^h)_i := LP_i(z_i) = \sum_{j=0}^k L(l_j(z_i))u(X_j) = \sum_{j=0}^k c_j^i u(X_j)$$

où on a posé $c_j^i = L(l_j(z_i))$

La contrainte $k \geq s$ assure que le degré du polynôme est au moins aussi élevé que l'ordre de l'opérateur. En faisant cette construction pour chaque point z_i de la grille de collocation, on obtient les coefficients c_j^i de la combinaison linéaire qui donne $LP_i(z_i)$. Ces coefficients forment la matrice L^h . De (1.2.3), on peut vérifier que la s -ième dérivée de $l_j(x)$ satisfait $l_j^{(s)}(z_i) \in O(h^{-s})$. C'est le terme dominant dans $L(l_j(z_i))$ lorsque $h \rightarrow 0$. On a donc la relation suivante

$$c_j^i \in O(h^{-s})$$

Définition 1.2.5. Soit $k \in \mathbb{N}$, on définit $\mathbb{P}_k[x]$ l'espace des polynômes de degré au plus k dans la variable x . On allègera l'écriture en posant $\mathbb{P}_k := \mathbb{P}_k[x]$

Proposition 1.2.3. Soit L un opérateur d'ordre s et L^h sa discrétisation à $k + 1$ points construite à partir du polynôme d'interpolation de degré k et avec $k \geq s$, alors

$$L^h p^h = (Lp)^h \quad \forall p \in \mathbb{P}_k$$

DÉMONSTRATION. La preuve est directe par construction de L^h car il existe un unique polynôme de degré au plus k passant par $k + 1$ points donc le polynôme d'interpolation P_i est précisément égal au polynôme interpolé p . \square

Avec ces derniers résultats, nous pouvons définir la propriété de superconsistance.

1.2.3. Superconsistance

Définition 1.2.6. Soit L un opérateur linéaire d'ordre s , L^h sa discrétisation locale à $k + 1$ points avec $k \geq s$ et soit Q l'espace des polynômes p tels que

$$(Lp)^h - L^h p^h = 0 \quad \text{sur } \vec{z} \quad (1.2.4)$$

le schéma L^h est dit **superconsistant** si $Q = \mathbb{P}_{k+1}$.

De la Proposition 1.2.3, $Q = \mathbb{P}_k$ pour tout choix de \vec{z} . La superconsistance est donc définie comme une augmentation de la dimension de l'espace Q des fonctions v tel que le résidu est nul sur \vec{z} . On donne un exemple pour illustrer cette propriété.

1.2.3.1. Exemple simple de superconsistance

Soit $u : [0, 1] \rightarrow \mathbb{R}$ une fonction donnée et $L = \frac{d}{dx}$ l'opérateur de dérivée première. On choisit un maillage uniforme de taille $n \in \mathbb{N}$ avec $h = 1/n$ de sorte que $x_i = ih$, $i = 0, 1, \dots, n$ et

$$u_i := u(x_i)$$

On choisit en premier lieu $z_i = x_i$, $i = 0, \dots, n - 1$. On considère L_a^h la discrétisation de L par différence avant définie ainsi :

$$(L_a^h u^h)_i := \frac{u_{i+1} - u_i}{h} \approx (Lu)(x_i)$$

On peut vérifier facilement en explicitant les polynômes généraux $p_1(x) = a_1x + b_1$ et $p_2(x) = a_2x^2 + b_2x + c_2$ de degré 1 et 2, respectivement, que

$$Lp_1(x_i) = a_1 = (L_a^h p_1^h)_i$$

mais que

$$Lp_2(x_i) = 2a_2x_i + b_2 \neq a_2(x_i + x_{i+1}) + b_2 = (L_a^h p_2^h)_i$$

Appelons L_c^h la discrétisation de L par différence centrée où le point de collocation est cette fois le point milieu entre deux points consécutifs du maillage de représentation \vec{x} . Prenons pour cela $z_i = x_i + \frac{h}{2}$, $i = 0, 1, \dots, n - 1$. On a

$$(L_c^h u^h)_i := \frac{u_{i+1} - u_i}{h} \approx (Lu) \left(x_i + \frac{h}{2} \right) = (Lu) \left(\frac{x_i + x_{i+1}}{2} \right)$$

On peut vérifier encore une fois que

$$Lp_1(x_i + h/2) = a_1 = (L_c^h p_1^h)_i$$

mais cette fois on a aussi

$$Lp_2(x_i + h/2) = 2a_2 \frac{x_i + x_{i+1}}{2} + b_2 = a_2(x_{i+1} + x_i) + b_2 = (L_c^h p_2^h)_i$$

Ainsi, pour l'opérateur par différence avant, $Q = \mathbb{P}_1$ soit un espace de dimension 2, alors que $Q = \mathbb{P}_2$, un espace de dimension 3, pour l'opérateur par différence centrée. Comme on a utilisé 2 points dans la construction de L_a^h et de L_c^h , par la Définition 1.2.6, seul l'opérateur centré est superconsistant.

La proposition suivante fait le lien entre superconsistance et la consistance.

Proposition 1.2.4. *Soit $L = a_0(x) + a_1(x) \frac{d}{dx} + \dots + a_s(x) \frac{d^s}{dx^s}$ un opérateur différentiel linéaire d'ordre s et soit L^h sa discrétisation par différences finies à $k + 1$ points, $k \geq s$, construite à partir du polynôme d'interpolation de degré k , alors la discrétisation est consistante d'ordre au moins $k + 1 - s$. De plus si $(Lp)^h = L^h p^h \forall p \in \mathbb{P}_{k+1}$ (superconsistance) alors la discrétisation est consistante d'ordre au moins $k + 2 - s$.*

DÉMONSTRATION. Soit u une fonction lisse (on assume la régularité nécessaire de u pour que les dérivées aient un sens), en suivant la construction présentée en début de section on a

$$Lu(z_i) \approx (L^h u^h)_i = \sum_{l=0}^k c_l u(X_l) \quad \text{où } c_j \in O(h^{-s})$$

Soit $T_j(x) = \frac{(x - z_i)^j}{j!} u^{(j)}(z_i)$, alors le développement de Taylor de u autour de z_i est

$$u(x) = T_0(x) + T_1(x) + T_2(x) + \dots$$

En chaque point de la grille de collocation on obtient

$$(L^h u^h)_i = \sum_{l=0}^k c_l u(X_l) = \sum_{l=0}^k c_l \left(\sum_{j=0}^{\infty} T_j(X_l) \right) = \sum_{j=0}^{\infty} \left(\sum_{l=0}^k c_l T_j(X_l) \right) = \sum_{j=0}^{\infty} (L^h T_j^h)_i$$

Par la Proposition 1.2.3, $(L^h T_j^h)_i = LT_j(z_i)$, $j \leq k$.

De plus, par définition des $T_j(x)$ on a

$$LT_j(z_i) = \begin{cases} a_j(z_i)u^{(j)}(z_i) & j \leq s \\ 0 & j > s \end{cases}$$

ce qui implique par définition de L que

$$\sum_{j=0}^k (L^h T_j^h)_i = \sum_{j=0}^k LT_j(z_i) = Lu(z_i)$$

On conclut que

$$(L^h u^h)_i = Lu(z_i) + \sum_{j=k+1}^{\infty} (L^h T_j^h)_i$$

Comme la distance $(X_l - z_j)$ est proportionnelle à h on a $T_j(X_l) \in O(h^j)$ $l = 0, 1, \dots, k$ et donc $(L^h T_j^h)_i \in O(h^{j-s})$ et l'ordre de consistance correspond à l'ordre du premier terme de la somme.

$$(L^h u^h)_i = Lu(z_i) + O(h^{k+1-s})$$

Si par hypothèse on avait superconsistance alors $(L^h T_{k+1}^h)_i = LT_{k+1}(z_i) = 0$ car $s \leq k$ et on obtient

$$(L^h u^h)_i = Lu(z_i) + \sum_{j=k+2}^{\infty} (L^h T_j^h)_i = Lu(z_i) + O(h^{k+2-s})$$

□

1.2.3.2. Note sur les travaux originaux

À notre connaissance, la propriété de superconsistance a été introduite par D. Funaro et L. Fatone entre autres dans les articles [10, 12, 13] qui sont le point de départ de notre recherche. La définition donnée dans ces travaux est la suivante :

Soit Q l'espace des fonctions v tel que le résidu $L^h v^h - (Lv)^h$ est nul sur \vec{z} , la discrétisation L^h est **superconsistante** si la dimension de Q est plus grande que le nombre de points ayant servi à discrétiser localement L .

Aucune information n'est cependant donnée sur la nature de Q ou de L^h dans la définition. Dans ce cadre général, il était difficile d'élaborer une théorie autour de la superconsistance, son implémentation et son impact sur la convergence des solutions. En particulier, il n'était pas clair si une discrétisation superconsistante était nécessairement consistante. Il suffit de considérer par exemple le phénomène de Runge [7] qui se présente lorsqu'on essaye d'interpoler certaines fonctions lisses

par des polynômes de degrés élevés définis sur tout l'intervalle. L'erreur d'interpolation grandit lorsqu'on raffine trop le maillage et donc un opérateur L^h construit à partir de ces polynômes ne serait pas nécessairement consistant. Nous avons décidé de développer la théorie dans le cas spécifique où Q est un espace de polynômes et L^h est un opérateur construit à partir d'un polynôme d'interpolation local. Ces choix ont permis d'obtenir la Proposition 1.2.4 et ainsi de justifier l'étymologie du terme "superconsistance". Après rétrospection, il aurait aussi semblé plus logique de définir la superconsistance directement à partir de l'ordre de consistance plutôt qu'à partir de la taille de l'espace Q . Nous avons cependant gardé la Définition 1.2.6 car elle respecte la façon dont la superconsistance a été introduite par ses auteurs et car une méthode explicite de construction des schémas en découle naturellement comme nous le verrons au chapitre 2.

La prochaine section est dédiée à l'adaptation de la théorie superconsistante en dimension 2.

1.3. DISCRÉTISATION D'UNE ÉQUATION DIFFÉRENTIELLE EN DIMENSION 2

La théorie en dimension 1 était relativement simple à développer, mais l'ajout d'une deuxième variable complique la notion de dimension de l'espace Q et du même fait la définition de la superconsistance.

Soit $u : \Omega \rightarrow \mathbb{R}$ et $f : I \rightarrow \mathbb{R}$ on considère l'équation

$$Lu = f$$

On pose cette fois $\Omega = [a_x, b_x] \times [a_y, b_y] \subset \mathbb{R}^2$ et $I = [c_x, d_x] \times [c_y, d_y] \subset \mathbb{R}^2$.

Aussi, on s'intéresse aux opérateurs L linéaires d'ordre w dans chaque variable de la forme

$$L = \sum_{p=0}^w \sum_{q=0}^w a_{p,q} \frac{\partial^{p+q}}{\partial x^p \partial y^q} \quad (1.3.1)$$

Nos résultats restent vrai pour des opérateurs plus généraux mais les démonstrations sont plus lourdes et les exemples des prochains chapitres se limitent à des opérateurs de la forme (1.3.1). Soit $n_x, n_y, m_x, m_y \in \mathbb{N}$, la grille de représentation et la grille de collocation sont respectivement

$$\vec{x} = \{(x_i, y_j)\}_{\substack{i=0,1,\dots,n_x \\ j=0,1,\dots,n_y}}$$

$$\vec{z} = \{(\tau_i, \mu_j)\}_{\substack{i=0,1,\dots,m_x \\ j=0,1,\dots,m_y}}$$

1.3.1. Consistance, convergence et stabilité

On peut linéariser la numérotation des indices et garder la même notation que pour la dimension 1. Le point (x_i, y_j) , devient alors $x_{i+j \cdot n_x}$. On combine les pas de discrétisation dans les 2 directions en posant $h = \max(h_x, h_y)$. Les Définitions 1.2.1 à 1.2.4 restent alors inchangées.

1.3.2. Méthode de construction de L^h

Soit \vec{x} la grille de représentation de u , et L un opérateur différentiel d'ordre w dans chaque variable de la forme (1.3.1). Pour trouver $(Lu)(\tau_i, \mu_j)$ on commence par choisir un ensemble de $k_x + 1$, $k_x \geq w$, points dans la direction x et $k_y + 1$, $k_y \geq w$, points dans la direction y . Encore par soucis de simplifier les preuves, on se restreint à $k_x = k_y = k$.

$$X_{0_i} < X_{1_i} < \dots < X_{k_i}$$

$$Y_{0_j} < Y_{1_j} < \dots < Y_{k_j}$$

Comme à la Section 1.2.2 les indices i et j indiquent que pour chaque point de collocation (τ_i, μ_j) on a possiblement des ensembles de points d'interpolation différents. On omettra ces indices par soucis d'alléger la notation.

On construit $P_{i,j}$ le polynôme d'interpolation local de degré k dans chaque variable tel que $P_{i,j}(X_r, Y_s) = u_{r,s} := u(X_r, Y_s)$, $r, s = 0, 1, \dots, k$. On peut l'écrire dans la base des polynômes de Lagrange comme

$$P_{i,j}(x, y) = \sum_{r=0}^k \sum_{s=0}^k l_r(x) q_s(y) u_{r,s}$$

avec

$$l_r(x) = \prod_{\substack{j=0 \\ j \neq r}}^k \frac{x - X_j}{X_r - X_j} \quad q_s(y) = \prod_{\substack{j=0 \\ j \neq s}}^k \frac{y - Y_j}{Y_s - Y_j}$$

Ensuite on évalue $LP_{i,j}$ au point de collocation.

$$(L^h u^h)_{i,j} := LP_{i,j}(\tau_i, \mu_j) = \sum_{r,s=0,\dots,k} L(l_r(\tau_i) q_s(\mu_j)) u_{r,s} = \sum_{r,s=0,\dots,k} c_{r,s} u_{r,s}$$

Par la forme des polynômes de Lagrange et de L on peut vérifier que les coefficients $c_{r,s} = L(l_r(\tau_i) q_s(\mu_j))$ satisfont

$$c_{r,s} \in O(h^{-w}) \quad \text{lorsque } h \rightarrow 0$$

Certains coefficients peuvent aussi être nuls si \vec{z} coïncide avec \vec{x} .

En dimension 1, si w est l'ordre de l'opérateur alors le polynôme d'interpolation utilisé dans la construction de L^h a $k + 1$ points avec $k \geq w$. Aussi, on a au moins $Q = \mathbb{P}_k$ et pour avoir superconsistance on exige que $Q = \mathbb{P}_{k+1}$ où Q est l'espace des fonctions telles que $(Lu)^h = L^h u^h$ sur \vec{z} . En dimension 2, pour lier adéquatement l'ordre de L , avec le nombre de points utilisés dans l'interpolation et pour définir Q on utilisera les espaces suivants :

Définition 1.3.1. Soit $k_x, k_y \in \mathbb{N}$, on définit $\mathbb{P}_{k_x, k_y}[x, y]$ l'espace des polynômes de degré au plus k_x dans la variable x et au plus k_y dans la variable y . On allègera l'écriture en posant $\mathbb{P}_{k_x, k_y} := \mathbb{P}_{k_x, k_y}[x, y]$

On note au passage que \mathbb{P}_{k_x, k_y} est de dimension $(k_x + 1)(k_y + 1)$. Ce choix permet d'avoir le résultat suivant, analogue à la Proposition 1.2.3

Proposition 1.3.1. Soit L un opérateur d'ordre w en x et en y de la forme (1.3.1). Soit L^h sa discrétisation à $k + 1$ points en x et $k + 1$ points en y , avec $k \geq w$, construite à partir du polynôme d'interpolation alors

$$L^h p^h = (Lp)^h \quad \forall p \in \mathbb{P}_{k, k}$$

DÉMONSTRATION. La preuve est immédiate comme en dimension 1 puisque L^h est construit comme tel. \square

On a donc $Q = \mathbb{P}_{k, k}$. Pour augmenter la dimension de Q on a le choix d'augmenter le degré des polynômes dans l'une ou l'autre des variables. Pour obtenir un résultat analogue à la Proposition 1.2.4 on ajoutera à la base de Q un polynôme dans $\mathbb{P}_{k+1, 0}$ et un dans $\mathbb{P}_{0, k+1}$. On obtient la définition suivante de la superconsistance en dimension 2.

Définition 1.3.2. Soit L un opérateur linéaire d'ordre w en x et en y de la forme (1.3.1), L^h sa discrétisation locale à $k + 1$ points en x et en y , avec $k \geq w$, construite à partir du polynôme d'interpolation. Soit Q l'espace des polynômes p tels que

$$(Lp)^h - L^h p^h = 0 \quad \text{sur } \vec{z} \tag{1.3.2}$$

le schéma L^h est dit **superconsistant** si

$$Q = \mathbb{P}_{k, k} \cup \mathbb{P}_{k+1, 0} \cup \mathbb{P}_{0, k+1}$$

Proposition 1.3.2. Soit L un opérateur différentiel linéaire d'ordre w en x et en y de la forme (1.3.1), L^h une discrétisation par différence finie à $k + 1$ points

en x en y , avec $k \geq w$, construite à partir du polynôme d'interpolation, alors la discrétisation est consistante d'ordre au moins $k+1-w$. De plus si $(Lp)^h = L^h p^h \forall p \in \mathbb{P}_{k,k} \cup \mathbb{P}_{k+1,0} \cup \mathbb{P}_{0,k+1}$ alors la discrétisation est consistante d'ordre au moins $k+2-w$.

DÉMONSTRATION. Soit u une fonction lisse, par construction de L^h on a localement

$$Lu(\tau_i, \mu_j) \approx (L^h u^h)_{i,j} = \sum_{p=0}^k \sum_{q=0}^k c_{p,q} u(X_p, Y_q)$$

où $c_{p,q} \in O(h^{-w})$.

En utilisant le développement en série de Taylor comme pour la preuve en dimension 1 et en posant

$$T_{r,s}(x, y) = \frac{(x - \tau_i)^r}{r!} \frac{(y - \mu_j)^s}{s!} \frac{\partial^{r+s} u}{\partial x^r \partial y^s}(\tau_i, \mu_j)$$

on trouve

$$(L^h u^h)_{i,j} = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} (L^h T_{r,s}^h)_{i,j}$$

Par la Proposition 1.3.1, $(L^h T_{r,s}^h)_{i,j} = LT_{r,s}(\tau_i, \mu_j)$ pour $r, s \leq k$ et le calcul explicite des $LT_{r,s}(\tau_i, \mu_j)$ mène à

$$(L^h u^h)_{i,j} = Lu(\tau_i, \mu_j) + \left(\sum_{r=k+1}^{\infty} \sum_{s=0}^k + \sum_{r=0}^k \sum_{s=k+1}^{\infty} + \sum_{r=k+1}^{\infty} \sum_{s=k+1}^{\infty} \right) (L^h T_{r,s}^h)_{i,j}$$

Comme $T_{r,s}(X_p, Y_q) \in O(h^{r+s})$ et $c_{p,q} \in O(h^{-w})$, alors $(L^h T_{r,s}^h)_{i,j} \in O(h^{r+s-w})$ et ainsi en regardant les premiers termes du résidu au membre de droite de l'équation précédente on trouve

$$(L^h u^h)_{i,j} = Lu(\tau_i, \mu_j) + O(h^{k+1-w})$$

Si on avait superconsistance alors $(L^h T_{k+1,0}^h)_{i,j} = LT_{k+1,0}(\tau_i, \mu_j) = 0$ et similairement $(L^h T_{0,k+1}^h)_{i,j} = 0$ donc on aurait

$$\begin{aligned} (L^h u^h)_{i,j} &= Lu(\tau_i, \mu_j) + \left(\sum_{r=k+2}^{\infty} \sum_{s=0}^k + \sum_{r=0}^k \sum_{s=k+2}^{\infty} + \sum_{r=k+1}^{\infty} \sum_{s=k+1}^{\infty} \right) (L^h T_{r,s}^h)_{i,j} \\ &= Lu(\tau_i, \mu_j) + O(h^{k+2-w}) \end{aligned}$$

□

Pour explorer d'avantage les propriétés d'un schéma superconsistant, il faut étudier des exemples spécifiques et c'est l'objet des prochains chapitres.

Chapitre 2

APPLICATION À DES PROBLÈMES SIMPLES EN UNE DIMENSION

2.1. MÉTHODE DE CONSTRUCTION DU SCHÉMA SUPERCONSISTANT

Soit \vec{x} la grille de représentation à $n + 1$ points en une dimension et \vec{z} la grille de collocation à $m + 1$ points. Jusqu'ici, la grille de collocation n'a jamais été fixée et c'est cette liberté de choix de \vec{z} qui permet la construction d'un schéma superconsistant comme on a pu le voir dans l'exemple 1.2.3.1. Soit L un opérateur d'ordre s et L^h sa discrétisation à $k + 1$ points $k \geq s$ construite à partir du polynôme d'interpolation. Soit $\chi_0, \chi_1, \dots, \chi_m \in \mathbb{P}_{k+1}$, pour avoir superconsistance au sens de la Définition 1.2.6, il suffit d'exiger que

$$L\chi_i(z_i) = (L^h\chi_i^h)_i \quad i = 0, 1, \dots, m \quad (2.1.1)$$

Soit Q l'espace des polynômes p tels que

$$(Lp)^h - L^hp^h = 0 \quad \text{sur } \vec{z}, \quad (2.1.2)$$

on a déjà $Q = \mathbb{P}_k$ par la Proposition 1.2.3. Soit $i = 0, 1, \dots, m$, tout polynôme $p_{k+1} \in \mathbb{P}_{k+1}$ peut s'écrire $p_{k+1} = \alpha_i\chi_i + p_{i,k}$ où $p_{i,k} \in \mathbb{P}_k$. Ainsi si (2.1.1) est vérifiée on a superconsistance car pour $i = 0, 1, \dots, m$:

$$\begin{aligned} Lp_{k+1}(z_i) &= L(\alpha_i\chi_i + p_{i,k})(z_i) \\ &= \alpha_i L\chi_i(z_i) + Lp_{i,k}(z_i) \\ &= \alpha_i (L^h\chi_i^h)_i + (L^hp_{i,k}^h)_i \\ &= (L^h(\alpha_i\chi_i^h + p_{i,k}^h))_i \\ &= (L^hp_{k+1}^h)_i \end{aligned} \quad (2.1.3)$$

et donc $Q = \mathbb{P}_{k+1}$

Soit $i \in \{0, 1, \dots, m\}$, pour trouver χ_i et z_i qui satisfont (2.1.1), une astuce est de choisir χ_i qui s'annule sur les points $X_{0_i}, X_{1_i}, \dots, X_{k_i}$ de \vec{x} qui ont servi à discrétiser localement l'opérateur (voir Section 1.2.2). Comme $(L^h \chi_i^h)_i$ est une combinaison linéaire des points $\chi_i(X_{j_i})$, on a nécessairement $(L^h \chi_i^h)_i = 0$. Il suffira donc de résoudre $(L\chi_i)(z_i) = 0$ pour z_i .

Pour bien comprendre la méthode de construction du schéma et l'impact du choix des points de collocation sur la consistance, on commence par étudier en détail la discrétisation des opérateurs dérivée première et dérivée seconde.

2.2. CONSTRUCTION DU SCHÉMA DE DÉRIVÉE PREMIÈRE SUPERCONSISTANT

Dans l'exemple de la Section 1.2.3.1, on a vu que le schéma avant pour la dérivée première n'était pas superconsistant car l'espace Q n'incluait que les polynômes de degré 1. On pouvait obtenir superconsistance en modifiant la grille de collocation \vec{z} pour obtenir un schéma centré. On détaille comment arriver rigoureusement à ce choix de \vec{z} .

Soit $L = \frac{d}{dx}$. On construit un maillage uniforme à $n + 1$ points de l'intervalle $[0, 1]$ avec $x_i = ih$, $i = 0, \dots, n$ et $h = 1/n$. Soit $u : [0, 1] \rightarrow \mathbb{R}$ une fonction donnée et $P_i(x)$ son polynôme d'interpolation local de degré 1 dans l'intervalle $[x_i, x_{i+1}]$

$$P_i(x) = \frac{1}{h}(x - x_i)u_{i+1} - \frac{1}{h}(x - x_{i+1})u_i$$

où $u_i := u(x_i)$.

L'opérateur discrétisé se construit en appliquant L sur chaque P_i et en évaluant le résultat au point de collocation z_i . On obtient l'approximation de $u'(z_i)$ suivante

$$(L^h u^h)_i = \frac{u_{i+1} - u_i}{h}$$

Pour tout choix z_i , cette discrétisation est exacte pour $p \in \mathbb{P}_1$ par (1.2.3). On remarque que pour $z_i = x_i$ on obtient le schéma de différence avant et que pour $z_i = x_{i+1}$ on obtient le schéma de différence arrière.

Pour avoir superconsistance on choisit les points z_i de sorte que $Q = \mathbb{P}_2$. On applique la méthode présentée à la section précédente.

On pose $\chi_i(x) = (x - x_i)(x - x_{i+1})$ et on résout $L\chi_i(z_i) = 0$.

$$L\chi_i(z_i) = 2z_i - 2x_i - h = 0 \iff z_i = x_i + h/2$$

On obtient la discrétisation superconsistante présentée dans l'exemple 1.2.3.1

$$(L^h u^h)_i = \frac{u_{i+1} - u_i}{h} \approx u'(x_i + h/2)$$

2.2.1. Analyse de la consistance

Par un développement en série de Taylor autour du point z_i on a

$$u(x) = u(z_i) + (x - z_i)u'(z_i) + \frac{(x - z_i)^2}{2}u''(z_i) + \frac{(x - z_i)^3}{6}u'''(z_i) + O(h^4)$$

et donc en posant $x = x_{i+1} = x_i + h$ on obtient

$$u_{i+1} = u(z_i) + (x_i - z_i + h)u'(z_i) + \frac{(x_i - z_i + h)^2}{2}u''(z_i) + \frac{(x_i - z_i + h)^3}{6}u'''(z_i) + O(h^4)$$

D'une manière analogue, poser $x = x_i$ donne

$$u_i = u(z_i) + (x_i - z_i)u'(z_i) + \frac{(x_i - z_i)^2}{2}u''(z_i) + \frac{(x_i - z_i)^3}{6}u'''(z_i) + O(h^4)$$

On soustrait u_i de u_{i+1} et on divise par h pour obtenir

$$\frac{u_{i+1} - u_i}{h} = u'(z_i) + \left(\frac{h}{2} + x_i - z_i\right)u''(z_i) + \frac{h^2 + 3h(x_i - z_i) + 3(x_i - z_i)^2}{6}u'''(z_i) + O(h^3)$$

On a ici un résultat qui illustre la Proposition 1.2.4. On remarque que pour tout choix $z_i \neq x_i + h/2$ le schéma est consistant d'ordre 1, mais pour $z_i = x_i + h/2$ on a consistance d'ordre 2. La superconsistance implique bien une consistance d'ordre plus élevé.

2.3. CONSTRUCTION DU SCHÉMA DE DÉRIVÉE SECONDE SUPER-CONSISTANT

Soit $L = \frac{d^2}{dx^2}$ et la même grille $x_i = ih$ que précédemment. Soit $u : [0, 1] \rightarrow \mathbb{R}$ une fonction donnée. Son polynôme d'interpolation local de degré 2 est

$$P_i(x) = \frac{1}{2h^2}(x - x_i)(x - x_{i+1})u_{i-1} - \frac{1}{h^2}(x - x_{i-1})(x - x_{i+1})u_i + \frac{1}{2h^2}(x - x_{i-1})(x - x_i)u_{i+1}$$

On applique L à P_i au point z_i et on obtient l'approximation de $u''(z_i)$ suivante

$$(L^h u^h)_i = \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2}$$

La discrétisation ci dessus est exacte pour $p \in \mathbb{P}_2$. On pose maintenant

$$\chi_i(x) = (x - x_{i-1})(x - x_i)(x - x_{i+1})$$

et on calcule

$$L\chi_i(z_i) = \chi_i''(z_i) = 6z_i - 6ih = 0 \iff z_i = ih = x_i$$

Le choix naturel $\vec{x} = \vec{z}$ donne ici une discrétisation superconsistante telle que $Q = \mathbb{P}_3$

2.3.1. Analyse de la consistance

Un développement en série de Taylor de u autour de z_i permet de montrer que

$$(L^h u^h)_i - Lu(z_i) = (x_i - z_i)u'''(z_i) + \frac{1}{12}(h^2 + 6(z_i - x_i)^2)u^{(4)}(z_i) + O(h^3)$$

Pour $z_i \neq x_i$ le schéma est consistant d'ordre h et pour $z_i = x_i$ la consistance est d'ordre h^2 .

Remarque 2.3.1. *Dans un tel exemple, la théorie superconsistante est plus intéressante si la grille n'est pas uniforme. On peut vérifier que si $(x_{i+1} - x_i) \neq (x_i - x_{i-1})$ alors les points de collocation idéaux ne seraient pas centrés par rapport aux points de la grille \vec{x} et l'on aurait*

$$z_i = \frac{1}{3}(x_{i-1} + x_i + x_{i+1})$$

Utiliser des schémas superconsistants avec des maillages non uniformes est une idée que nous n'avons pas suffisamment étudiée au cours de notre recherche pour pouvoir l'inclure dans ce travail mais elle mérite d'être explorée.

Dans le prochain exemple, on veut étudier la construction d'une discrétisation superconsistante lorsque l'opérateur combine une dérivée première et une dérivée seconde. Nous utilisons pour cela l'équation d'advection-diffusion que nous présenterons plus en détail au chapitre 3. On observe cette fois que même sur une grille uniforme, le point de collocation superconsistant n'est pas centré et dépend du rapport entre le coefficient du terme diffusif et celui du terme advectif. Aussi, dans les deux précédents exemples, la solution de $L\chi = 0$ est unique ce qui n'est pas le cas lorsque L contient plusieurs ordres de dérivée.

2.4. PROBLÈME D'ADVECTION-DIFFUSION EN DIMENSION 1

Soit $\Omega =]0, 1[\in \mathbb{R}$, $\partial\Omega = \{0, 1\}$ et $L = -\epsilon \frac{d^2}{dx^2} + \beta \frac{d}{dx}$ avec $\epsilon > 0$ et $\beta \in \mathbb{R}$ on considère le problème

$$\begin{cases} Lu = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (2.4.1)$$

On construit un maillage uniforme à $n + 1$ points

$$0 = x_0 < x_1 < \dots < x_n = 1$$

où $h = 1/n$ et $x_i = ih$. On reprend le polynôme d'interpolation de degré 2 de l'exemple précédent et on y applique le nouvel opérateur L . Après évaluation au point z_i on obtient l'approximation de $Lu(z_i)$ suivante

$$(L^h u^h)_i = (\beta(z_i - x_i) - \epsilon) \left(\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \right) + \beta \left(\frac{u_{i+1} - u_{i-1}}{2h} \right) \quad (2.4.2)$$

Soit $\chi_i(x) = (x - x_{i-1})(x - x_i)(x - x_{i+1})$, pour avoir un schéma superconsistant le point de collocation z_i doit satisfaire $L\chi(z_i) = 0$. Les solutions satisfont donc

$$L\chi(z_i) = 6\epsilon(x_i - z_i) + 3\beta(x_i - z_i)^2 - \beta h^2 = 0 \quad (2.4.3)$$

et sont

$$\begin{aligned} z_i &= x_i + \frac{\epsilon}{\beta} + \sqrt{\frac{\epsilon^2}{\beta^2} + \frac{h^2}{3}} \\ z_i &= x_i + \frac{\epsilon}{\beta} - \sqrt{\frac{\epsilon^2}{\beta^2} + \frac{h^2}{3}} \end{aligned} \quad (2.4.4)$$

Dépendamment du signe de β , une de ces deux solutions n'est pas bornée lorsque $\left| \frac{\epsilon}{\beta} \right| \rightarrow \infty$. La bonne solution est celle qui correspond à un décalage "upwind" c.-a.-d. $(x_i - z_i)\beta \geq 0$. Elle peut s'écrire

$$z_i = x_i - \frac{\epsilon}{\beta} \left(\sqrt{1 + \frac{\beta^2 h^2}{3\epsilon^2}} - 1 \right) \quad (2.4.5)$$

et ce peut importe le signe de β . Dans ce cas l'équation différentielle discrétisée s'écrit

$$\begin{aligned} (L^h u^h)_i &= -\epsilon \sqrt{1 + \frac{\beta^2 h^2}{3\epsilon^2}} \left(\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \right) + \beta \left(\frac{u_{i+1} - u_{i-1}}{2h} \right) \\ &\approx Lu \left(x_i - \frac{\epsilon}{\beta} \left(\sqrt{1 + \frac{\beta^2 h^2}{3\epsilon^2}} - 1 \right) \right) \end{aligned} \quad (2.4.6)$$

On vérifie que z_i est dans l'intervalle $[x_{i-1}, x_{i+1}]$. Soit $y = \left| \frac{\epsilon}{\beta} \right|$ on calcule la distance entre x_i et z_i

$$|z_i - x_i| = y \left(\sqrt{1 + \frac{h^2}{3y^2}} - 1 \right) = \sqrt{y^2 + \frac{h^2}{3}} - y$$

Comme $y > 0$ on obtient

$$\frac{d}{dy} \left(\sqrt{y^2 + \frac{h^2}{3}} - y \right) = \frac{1}{\sqrt{1 + \frac{h^2}{3y^2}}} - 1 < 0$$

Donc la distance $|z_i - x_i|$ décroît lorsque le rapport $\left| \frac{\epsilon}{\beta} \right|$ augmente. La distance maximale est atteinte lorsque ce rapport $\left| \frac{\epsilon}{\beta} \right|$ tend vers 0 auquel cas $|z_i - x_i| \rightarrow \frac{h}{\sqrt{3}}$. On obtient

$$0 < |z_i - x_i| < \frac{h}{\sqrt{3}} < h$$

2.4.1. Analyse de la consistance du schéma

En substituant u dans l'opérateur (2.4.2) par son développement en série de Taylor autour de z_i on trouve

$$\begin{aligned} (L^h u^h)_i - Lu(z_i) &= \left(-\epsilon(x_i - z_i) - \frac{\beta(x_i - z_i)^2}{2} + \frac{\beta h^2}{6} \right) u'''(z_i) \\ &+ \left(-\frac{\beta}{3}(x_i - z_i)^3 - \frac{\epsilon}{2}(x_i - z_i)^2 + \frac{\beta h^2}{12}(x_i - z_i) - \frac{\epsilon h^2}{12} \right) u''''(z_i) \\ &+ O(h^3) \end{aligned} \tag{2.4.7}$$

Comme $|x_i - z_i| < h$, l'opérateur est consistant d'ordre h . De plus, on remarque que les solutions de (2.4.3) annulent le premier terme et augmentent ainsi la consistance à h^2 . En analysant (2.4.7), on remarque qu'on peut éliminer de 3 façons différentes le terme d'ordre h sans annuler complètement le terme devant $u''''(z_i)$. On conclut que les 3 solutions suivantes assurent aussi une consistance d'ordre h^2

$$z_i \in \left\{ x_i, x_i + \frac{2\epsilon}{\beta}, x_i - \frac{\beta h^2}{6\epsilon} \right\} \tag{2.4.8}$$

Cependant on en élimine immédiatement deux puisque seule la première est bornée pour toute valeur de ϵ et β . Pour montrer l'avantage de la solution superconsistante (2.4.5) par rapport à la solution centrée $z_i = x_i$, il faut faire une étude de la stabilité.

Remarque 2.4.1. *Une autre façon d'avoir une consistance d'ordre h^2 sans éliminer complètement le coefficient de u'''' serait de choisir z_i tel que $z_i \rightarrow x_i$ quadratiquement lorsque $h \rightarrow 0$. On verra dans la prochaine section que certains choix de χ peuvent mener à ce résultat.*

2.4.2. Analyse de la stabilité du schéma

On est face à un problème qui présente un terme diffusif négligeable devant le terme advectif. Il est donc de rigueur de s'assurer que la solution est bornée lorsque h est fixé et $\epsilon \rightarrow 0$. On multiplie $f_i = (L^h u^h)_i$ par u_i en utilisant (2.4.2). On somme sur $i = 1, \dots, n-1$ avec les conditions frontières $u_0 = u_n = 0$. Puisque la différence $z_i - x_i$ est constante pour chaque i on trouve

$$\begin{aligned}
\sum_{i=1}^{n-1} f(z_i)u_i &= (\beta(z_i - x_i) - \epsilon) \sum_{i=1}^{n-1} \frac{u_{i-1}u_i - 2u_i^2 + u_{i+1}u_i}{h^2} + \beta \sum_{i=1}^{n-1} \frac{u_{i+1}u_i - u_{i-1}u_i}{2h} \\
&= (\beta(z_i - x_i) - \epsilon) \left(\sum_{i=1}^{n-1} \frac{u_{i-1}u_i}{h^2} - \sum_{i=1}^{n-1} \frac{2u_i^2}{h^2} + \sum_{i=1}^{n-1} \frac{u_{i+1}u_i}{h^2} \right) \\
&\quad + \beta \sum_{i=1}^{n-1} \frac{u_{i+1}u_i}{2h} - \beta \sum_{i=1}^{n-1} \frac{u_{i-1}u_i}{2h} \\
&= (\beta(z_i - x_i) - \epsilon) \left(\sum_{i=0}^{n-2} \frac{u_i u_{i+1}}{h^2} - \sum_{i=1}^{n-1} \frac{u_i^2}{h^2} - \sum_{i=0}^{n-2} \frac{u_{i+1}^2}{h^2} + \sum_{i=1}^{n-1} \frac{u_{i+1}u_i}{h^2} \right) \\
&= (\beta(z_i - x_i) - \epsilon) \left(\sum_{i=0}^{n-1} \frac{u_i u_{i+1}}{h^2} - \sum_{i=0}^{n-1} \frac{u_i^2}{h^2} - \sum_{i=0}^{n-1} \frac{u_{i+1}^2}{h^2} + \sum_{i=0}^{n-1} \frac{u_{i+1}u_i}{h^2} \right) \\
&= (\epsilon - \beta(z_i - x_i)) \sum_{i=0}^{n-1} \frac{(u_i - u_{i+1})^2}{h^2}
\end{aligned} \tag{2.4.9}$$

En imposant $(\epsilon - \beta(z_i - x_i)) \geq 0$ (choix upwind), et en utilisant les inégalités de Cauchy-Schwartz B.0.1 et de Poincaré B.0.2 on trouve

$$\|u\| \|f\| \geq \sum_{i=1}^{n-1} f(z_i)u_i = (\epsilon - \beta(z_i - x_i)) \sum_{i=0}^{n-1} \frac{(u_i - u_{i+1})^2}{h^2} \geq (\epsilon - \beta(z_i - x_i)) \|u\|^2$$

Avec le choix superconsistant (2.4.5) on obtient

$$\|u\| \leq \frac{1}{\epsilon - \beta(z_i - x_i)} \|f\| = \frac{1}{\sqrt{\epsilon^2 + \frac{\beta^2 h^2}{3}}} \|f\|$$

Si $h \rightarrow 0$, $\|u\|$ reste borné par $\frac{1}{\epsilon} \|f\|$

Avec le choix $z_i = x_i$ on a directement

$$\|u\| \leq \frac{1}{\epsilon} \|f\|$$

Dans les deux cas, on a la stabilité donnée par la Définition 1.2.4. Cependant, avec le choix centré la solution n'est pas borné lorsque $\epsilon \rightarrow 0$. En l'occurrence, en

prenant $\beta = 1$, $f = 1$ et $z_i = x_i$ on peut vérifier [12] que la solution numérique exacte au problème discret est

$$u_i = hi - \frac{\kappa^i - 1}{\kappa^n - 1} \quad \text{où} \quad \kappa = \frac{2\epsilon + h}{2\epsilon - h}$$

Si $h > 2\epsilon$ alors $\kappa < 0$ et la solution a un comportement oscillatoire dont l'amplitude est d'autant plus grande que ϵ est petit. On est donc contraint d'utiliser un maillage de l'ordre de ϵ ce qui peut être numériquement impraticable.

Avec le choix superconsistant, on a la borne suivante $\forall \epsilon$

$$\|u\| \leq \frac{\sqrt{3}}{|\beta|h} \|f\|$$

Dans ce problème simple d'advection-diffusion en dimension 1, le point de collocation superconsistant assure à la fois stabilité et consistance d'ordre h^2 pour un paramètre ϵ très petit.

2.4.3. Exemple numérique

On résout numériquement le problème (2.4.1) en comparant la discrétisation superconsistante (2.4.6), la discrétisation upwind (on pose $z_i = x_i - h/2$ dans (2.4.2)) et la discrétisation centrée (on pose $z_i = x_i$ dans (2.4.2)).

On choisit $\beta = 1$, $f = 1$ dans (2.4.1) et on s'intéresse à l'erreur

$$e(h) = \frac{1}{\sqrt{n}} \|U - u^h\|_2 = \sqrt{\frac{1}{n} \sum_{i=1}^n (U_i - u(x_i))^2} \quad (2.4.10)$$

où U est la solution numérique de $L^h U = f^h$ et u^h est la solution exacte évaluée sur la grille de représentation.

Avec les conditions Dirichlet homogènes on peut vérifier que la solution exacte est

$$u(x) = x - \frac{1 - \exp\left(\frac{x}{\epsilon}\right)}{1 - \exp\left(\frac{1}{\epsilon}\right)} \quad (2.4.11)$$

Pour $\epsilon \rightarrow 0$, la solution tend vers la fonction discontinue

$$\lim_{\epsilon \rightarrow 0} u(x) = \begin{cases} x & \text{si } 0 \leq x < 1 \\ 0 & \text{si } x = 1 \end{cases}$$

Pour des valeurs de ϵ très petites, une couche limite se crée à la frontière $x = 1$. On peut approximer la largeur de cette couche en calculant le point de maximum de la fonction

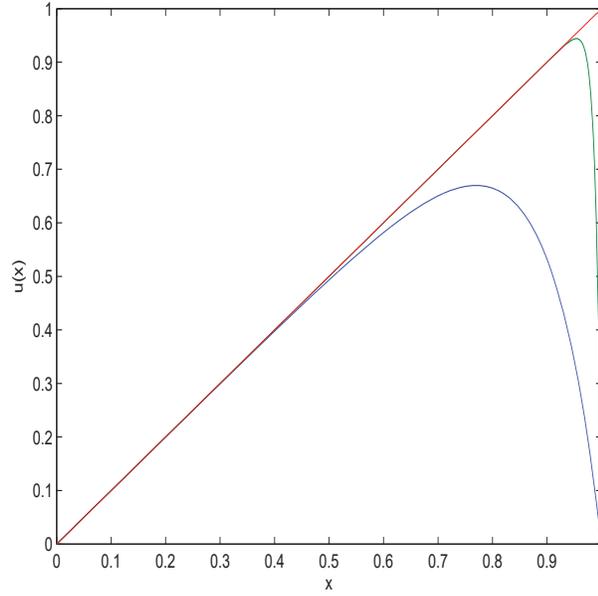


FIGURE 2.1. Graphique de la solution exacte (2.4.11) pour $\epsilon = 1/10$ (bleu), $\epsilon = 1/100$ (vert), $\epsilon = 1/1000$ (rouge)

$$u'(x_{max}) = 0 \iff x_{max} = \epsilon \log \left(\epsilon \left(\exp \left(\frac{1}{\epsilon} \right) - 1 \right) \right) \quad (2.4.12)$$

Pour $\epsilon \ll 1$ on peut estimer la largeur de la couche par

$$|1 - x_{max}| = \left| 1 - \epsilon \log \epsilon - \epsilon \log \left(\exp \left(\frac{1}{\epsilon} \right) - 1 \right) \right| \approx |\epsilon \log \epsilon| \quad (2.4.13)$$

La dernière approximation utilise $\exp(1/\epsilon) - 1 \approx \exp(1/\epsilon)$

2.4.3.1. Calcul de l'ordre de convergence

On rappelle que selon la Proposition 1.2.1, comme le schéma est stable, $e \leq Cr$ pour une certaine constante $C > 0$ et donc l'ordre de convergence est au moins aussi élevé que l'ordre de consistance. Comme on connaît la solution exacte dans notre exemple, on pourra calculer directement l'erreur et vérifier qu'elle diminue d'ordre h^2 . Si entre chaque calcul on raffine le maillage en divisant le pas h par 2 on pourra calculer l'ordre de convergence de l'erreur simplement en calculant le rapport de 2 erreurs consécutives. En effet, supposons que l'erreur $e(h)$ est asymptotiquement proportionnelle à h^α , alors il existe une constante A telle que pour h suffisamment petit on a

$$e(h) \sim Ah^\alpha$$

Ainsi

$$\frac{e(h)}{e(h/2)} \sim 2^\alpha$$

et, en prenant le log en base 2, on obtient l'ordre de convergence. Cette technique oblige l'utilisation de grilles exponentiellement de plus en plus grandes. En une dimension c'est facilement applicable mais pas en dimension 2. Une autre technique est de prendre le log de l'erreur car alors on a

$$\log(e(h)) \sim \log(A) + \alpha \log(h)$$

et α peut être calculé comme la pente de la droite du graphique du log de l'erreur en fonction du log de h . Cette pente peut être calculée simplement par la fonction `polyfit` de `MATLAB`. En pratique on obtient rarement une droite parfaite à cause en outre des erreurs d'arrondi machine et il faut choisir un intervalle de points adéquat, souvent en éliminant les extrémités où le maillage est trop grossier ou trop fin.

2.4.3.2. Comparaison des résultats

Dans les Tableaux 2.1 à 2.3 on donne l'erreur (2.4.10) en fonction de $h = 1/n$ pour les 3 méthodes de discrétisation et pour différentes valeurs de ϵ . Les courbes log-log des résultats sont montrées dans la Figure 2.2 et les ordres sont calculés et indiqués dans le Tableau 2.4. Certaines solutions U calculées sont présentées visuellement dans les Figures 2.3 à 2.5.

| n | superconsistant | upwind | centré |
|-----|-----------------|---------|---------------|
| 8 | 0.0175 | 0.0732 | 0.0226 |
| 16 | 0.00495 | 0.0422 | 0.0053 |
| 32 | 0.00127 | 0.0227 | 0.0013 |
| 64 | 0.000321 | 0.0118 | 0.000322 |
| 128 | $8.03e - 005$ | 0.00603 | $8.04e - 005$ |

TABLEAU 2.1. Comparaison des erreurs pour $\epsilon = 0.1$ avec $\beta = 1$ et $f = 1$

On peut voir dans la Figure 2.2 que les courbes log-log de l'erreur ne sont pas complètement droite. Pour un maillage trop fin, l'erreur d'arrondi machine domine l'erreur de troncature et on perd la convergence. Ce phénomène apparaît autour de $n = 10^5$ dans les trois premiers graphiques de la Figure 2.2 où on voit que l'erreur augmente avec n .

On a aussi des problèmes avec n trop petit. En l'occurrence, l'ordre de convergence n'est pas atteint tant que le maillage n'est pas assez fin par rapport à ϵ . De la Figure 2.1, on voit que la solution exacte décroît en s'approchant de la

| n | superconsistant | upwind | centré |
|------|-----------------|---------|---------------|
| 8 | 0.0255 | 0.00281 | 5.2 |
| 16 | 0.018 | 0.00394 | 1.15 |
| 32 | 0.0129 | 0.00548 | 0.334 |
| 64 | 0.00938 | 0.00753 | 0.152 |
| 128 | 0.00739 | 0.0101 | 0.065 |
| 256 | 0.00607 | 0.0118 | 0.0225 |
| 512 | 0.00338 | 0.0098 | 0.00581 |
| 1024 | 0.00114 | 0.00613 | 0.00134 |
| 2048 | 0.000307 | 0.00341 | 0.00032 |
| 4096 | $7.81e - 005$ | 0.00181 | $7.89e - 005$ |

TABLEAU 2.2. Comparaison des erreurs pour $\epsilon = 0.001$ avec $\beta = 1$ et $f = 1$

| n | superconsistant | upwind | centré |
|---------|-----------------|---------------|---------------|
| 8 | 0.0254 | $2.83e - 005$ | 552 |
| 16 | 0.018 | $4e - 005$ | 138 |
| 32 | 0.0127 | $5.66e - 005$ | 34.2 |
| 64 | 0.009 | $7.99e - 005$ | 8.3 |
| 128 | 0.00636 | 0.000113 | 1.89 |
| 256 | 0.0045 | 0.00016 | 0.474 |
| 512 | 0.00318 | 0.000225 | 0.216 |
| 1024 | 0.00225 | 0.000317 | 0.107 |
| 2048 | 0.0016 | 0.000444 | 0.0524 |
| 4096 | 0.00114 | 0.000615 | 0.0251 |
| 8192 | 0.00085 | 0.000839 | 0.0114 |
| 16384 | 0.000697 | 0.00109 | 0.0046 |
| 32768 | 0.000526 | 0.00115 | 0.00143 |
| 65536 | 0.00024 | 0.000848 | 0.000345 |
| 131072 | $7.25e - 005$ | 0.000502 | $8.01e - 005$ |
| 262144 | $1.89e - 005$ | 0.000273 | $1.94e - 005$ |
| 524288 | $4.78e - 006$ | 0.000143 | $4.81e - 006$ |
| 1048576 | $1.2e - 006$ | $7.34e - 005$ | $1.2e - 006$ |

TABLEAU 2.3. Comparaison des erreurs pour $\epsilon = 0.00001$ avec $\beta = 1$ et $f = 1$

frontière $x = 1$ et que ce changement dans le signe de la dérivée survient de plus en plus proche de la frontière lorsque ϵ diminue. Pour un maillage trop grossier par rapport à ϵ , ce changement brusque dans l'allure de la courbe n'est pas calculé adéquatement dans les solutions numériques donc la couche limite n'est en quelque sorte pas "capturée" par le maillage. Ainsi, dans les quatre derniers graphique de la Figure 2.2, on peut constater que l'ordre de convergence devient

| ϵ | superconsistant | upwind | centré |
|------------|-----------------|--------|--------|
| 1 | 2.00 | 0.99 | 2.00 |
| 0.1 | 2.00 | 0.99 | 2.00 |
| 0.01 | 2.00 | 0.98 | 2.00 |
| 0.001 | 2.00 | 0.99 | 2.00 |
| 0.0001 | 2.00 | 0.97 | 2.00 |
| $1e - 005$ | 1.97 | 0.93 | 2.02 |

TABLEAU 2.4. Comparaison des ordres de convergence des solutions discrètes pour différents ϵ avec $\beta = 1$ et $f = 1$

quadratique lorsque h est comparable à la largeur de la couche limite $\epsilon \log \epsilon$ calculée plus haut pour des petites valeurs de ϵ . Par exemple, pour $\epsilon = 10^{-4}$, on a $|\epsilon \log \epsilon| \approx 9.21 \times 10^{-4}$ ce qui correspond sur le graphique correspondant de la Figure 2.2 à peu près à l'endroit où la pente de la courbe devient -2 . On peut aussi voir dans les graphiques des Figures 2.3, 2.4 et 2.5 que les solutions numériques décroissent trop tôt pour n petit et que le changement dans l'allure de la courbe est plus doux que pour la solution exacte. C'est un phénomène de viscosité artificielle lié à la multiplication du terme diffusif par $\sqrt{1 + \frac{\beta^2 h^2}{3\epsilon^2}}$ dans l'équation discrète (2.4.6). L'augmentation du terme diffusif lisse la solution et améliore la stabilité mais en contrepartie on perd en précision pour h trop grand par rapport à la couche limite.

Les ordres de convergence ont donc été calculés en éliminant les maillages trop grossiers ou trop fins et on voit dans le Tableau 2.4 qu'ils correspondent aux ordres théoriques. La convergence pour la méthode centrée et pour la méthode superconsistante est quadratique alors qu'elle est linéaire pour la méthode upwind.

Le fait que h doit être suffisamment petit par rapport à ϵ pour atteindre un ordre de convergence quadratique peut sembler contraignant d'autant plus que pour un tel h la méthode superconsistante n'apporte pas d'amélioration visible par rapport à la méthode centrée. Cependant si on s'intéresse à la magnitude de l'erreur ainsi qu'à la stabilité de la méthode, la discrétisation superconsistante est nettement meilleure que celle centrée. On réfère aux Figures 2.3 et 2.4 dans lesquelles on note les fortes oscillations de la méthode centrée.

Il semble que pour ϵ petit la méthode upwind performe mieux sur les maillages grossiers (voir les trois derniers graphiques de la Figure 2.2). Cependant la convergence n'est pas encore atteinte donc le choix du problème avec $f = 1$ est certainement en cause et on ne peut rien généraliser.

Dans [12] sont présentés d'autres résultats numériques pour un problème avec un membre de droite $f = x$ non constant et tel que la solution exacte est donnée

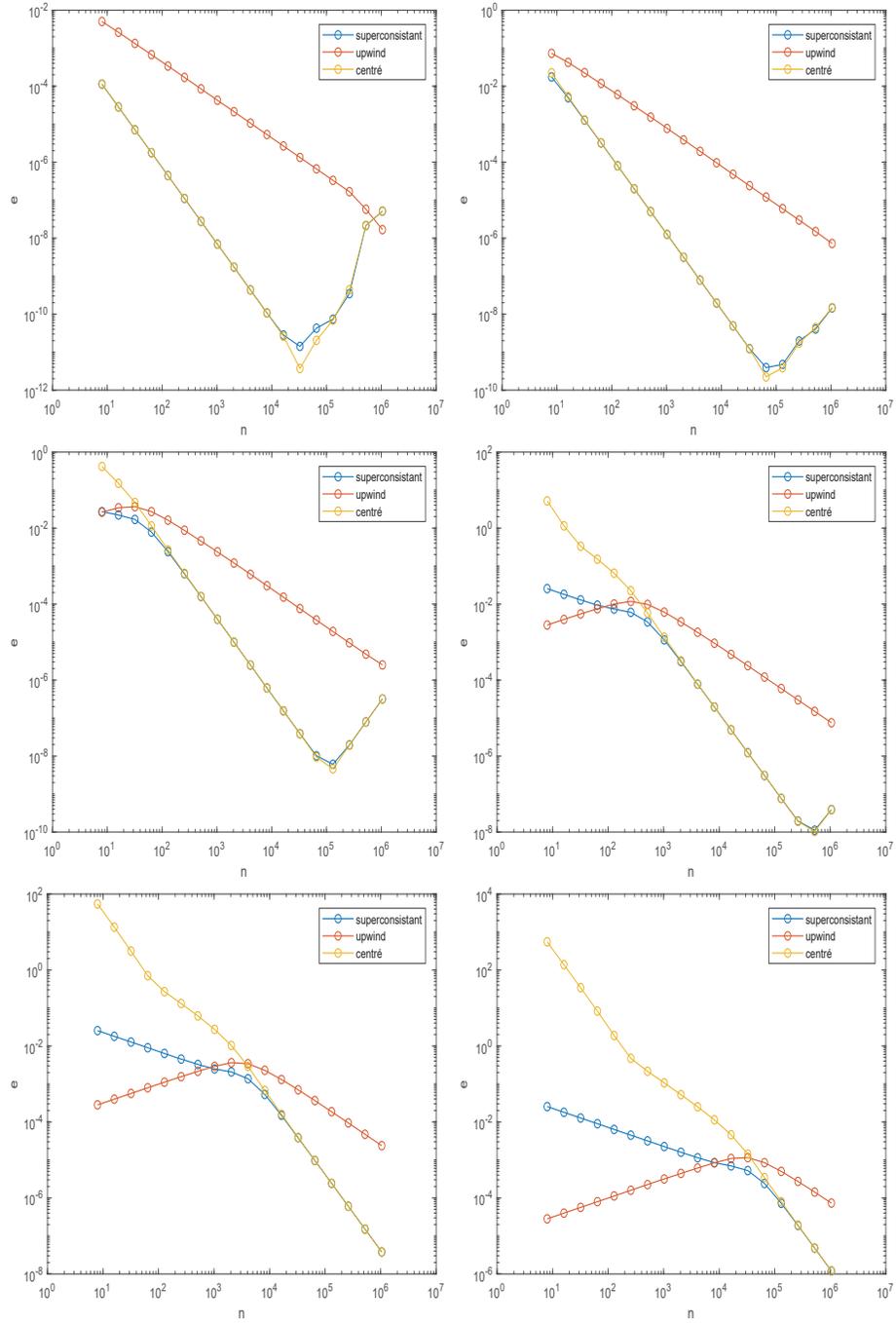


FIGURE 2.2. Graphique à l'échelle log-log de l'erreur en fonction de n pour $\epsilon = 1, 0.1, 0.01, 0.0001, 0.00001$ (lecture de gauche à droite puis de haut en bas)

par.

$$u(x) = \epsilon x + \frac{x^2}{2} - \left(\frac{1}{2} + \epsilon\right) \frac{1 - \exp\left(\frac{x}{\epsilon}\right)}{1 - \exp\left(\frac{1}{\epsilon}\right)} \quad (2.4.14)$$

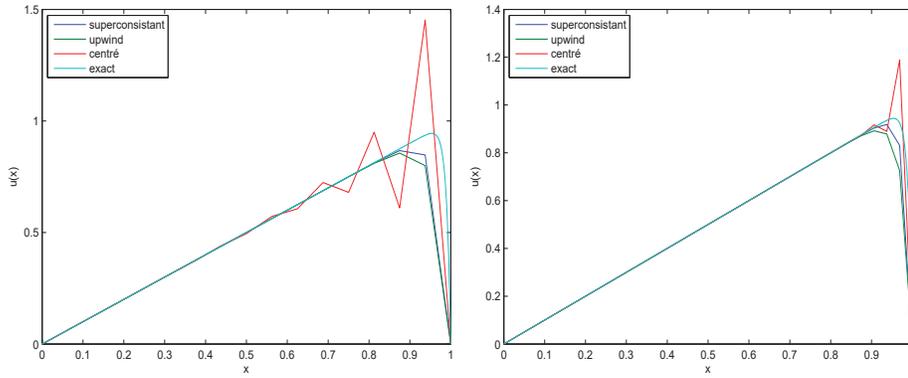


FIGURE 2.3. Comparaison entre les solutions discrètes et la solution exacte pour $\epsilon = 0.01$ et $n = 16, 32$ (de gauche à droite)

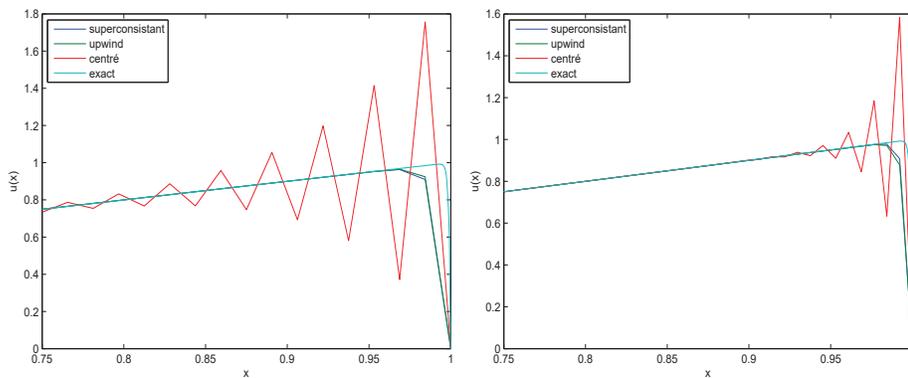


FIGURE 2.4. Comparaison entre les solutions discrètes et la solution exacte sur l'intervalle $x \in [0.75, 1]$ pour $\epsilon = 0.001$ et $n = 64, 128$ (de gauche à droite)

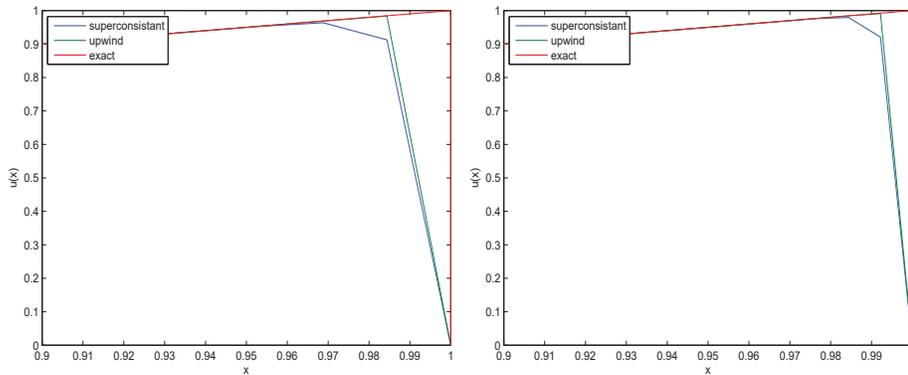


FIGURE 2.5. Comparaison entre les solutions discrètes et la solution exacte sur l'intervalle $x \in [0.9, 1]$ pour $\epsilon = 0.00001$ et $n = 64, 128$ (de gauche à droite)

La Figure D.1 en annexe montre les courbes log-log de l'erreur pour ce problème. Les mêmes conclusions peuvent être tirées de ces graphiques et on voit que la méthode superconsistante est la plus performante des trois méthodes.

Chapitre 3

APPLICATION À UN PROBLÈME STATIONNAIRE D'ADVECTION-DIFFUSION EN DIMENSION 2

Dans la première section on explicite sommairement le développement menant à l'obtention de l'équation d'advection-diffusion. On peut retrouver ces résultats entre autres dans les ouvrages [4, 19].

3.1. DÉRIVATION DE L'ÉQUATION D'ADVECTION-DIFFUSION

3.1.1. Théorème de la divergence

Le Théorème de la divergence est un résultat mathématique qui lie la divergence d'un champ de vecteurs à l'intérieur d'un domaine et le flux de ce champ sur la frontière du domaine. Il est valide en dimension $n \in \mathbb{N}$ mais dans l'étude des fluides on ne s'intéresse qu'à $n = 3$. On renvoi à [19] pour les détails de la preuve.

Théorème 3.1.1. (*Divergence*)

Soit V un sous ensemble compact de \mathbb{R}^3 dont la frontière S est lisse et soit \vec{F} un champ de vecteurs de classe C^1 dans V alors

$$\int_V \nabla \cdot \vec{F} dV = \int_S \vec{F} \cdot \hat{n} dS \quad (3.1.1)$$

où \hat{n} est le vecteur unitaire normal à S et orienté vers l'extérieur de V .

3.1.2. Équation de continuité

Le principe physique de conservation de la masse est décrit en hydrodynamique par l'équation de continuité. Soit V un volume fixe délimité par une surface S , en suivant le raisonnement présenté dans [4] on a

$$\frac{d}{dt} \int_V \rho dV = - \int_S \rho \vec{v} \cdot \hat{n} dS \quad (3.1.2)$$

où le membre de gauche est le taux de variation de la masse dans V et le membre de droite est le flux de masse traversant la frontière S . \vec{n} est le vecteur unitaire normal à S de sorte que seules les contributions normales du courant de masse $\rho \vec{v}$ influent sur la masse. Il est orienté vers l'extérieur du volume donc le signe négatif indique qu'un flux sortant entraîne une diminution de la masse. On applique le Théorème de la divergence (3.1.1) au membre de droite et on trouve

$$\frac{d}{dt} \int_V \rho dV = - \int_V \nabla \cdot (\rho \vec{v}) dV \quad (3.1.3)$$

Pour un volume fini, l'intégrale et la dérivée commutent donc

$$\int_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) \right) dV = 0 \quad (3.1.4)$$

et comme V est arbitraire l'intégrande doit être nulle ce qui donne l'**équation de continuité**

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \quad (3.1.5)$$

3.1.3. Équation d'advection-diffusion

L'équation d'advection-diffusion décrit des phénomènes où il y a un déplacement d'une entité au sein d'un système dû à un transport par un champ de vecteurs, l'advection, ainsi qu'à un processus de diffusion. Pour illustrer cela dans le cadre de l'hydrodynamique, on s'intéresse à une certaine concentration c (par exemple de sel) dans un fluide de densité ρ qui se déplace à vitesse \vec{v} . Dans cet exemple, un changement dans la concentration c est dû au champ de vecteurs \vec{v} ainsi qu'au phénomène de diffusion inhérent du sel dans le fluide.

On s'intéressera à la situation où le fluide est en régime permanent c'est à dire que le champ de vitesse \vec{v} est indépendant de t .

Soit ϕ le courant de masse de l'entité, il est donné par la somme du courant dû à l'écoulement et du courant dû à la diffusion et on l'écrit :

$$\phi = \rho c \vec{v} - \rho D \nabla c \quad (3.1.6)$$

Dans le premier terme, ρc est simplement la masse de l'entité par unité de volume de fluide donc $\rho c \vec{v}$ est la masse de l'entité traversant une unité de surface par unité de temps.

Le deuxième terme découle directement de la loi de Fick [2] qui stipule que la diffusion se fait des régions de haute concentration vers les régions de basse concentration d'où le signe négatif du terme $-\nabla c$. Finalement $D > 0$ est le coefficient de diffusivité en unité de surface par unité de temps.

Soit V un volume fixe dans le fluide de frontière S , la masse de l'entité est donnée par

$$M = \int_V \rho c dV \quad (3.1.7)$$

On pose maintenant que le changement de masse M dans le volume fixe V est égal au flux de masse ϕ sur la frontière S c.-à-d.

$$\frac{d}{dt} \int_V \rho c dV = - \int_S (\rho c \vec{v} - \rho D \nabla c) \cdot \hat{n} dS \quad (3.1.8)$$

Comme dans l'équation (3.1.2), \hat{n} est le vecteur unitaire normal à S et orienté vers l'extérieur. V est un volume fini indépendant de t donc les opérateurs de dérivation et d'intégration commutent au membre de gauche. Après application du Théorème de la divergence (3.1.1) au membre de droite on obtient

$$\int_V \left(\frac{\partial}{\partial t} (\rho c) + \nabla \cdot (\rho c \vec{v} - \rho D \nabla c) \right) dV = 0 \quad (3.1.9)$$

et comme V est arbitraire, l'intégrande doit être nulle. Après quelques manipulations on arrive à

$$c \frac{\partial \rho}{\partial t} + c \nabla \cdot (\rho \vec{v}) + \rho \frac{dc}{dt} + \rho \vec{v} \cdot \nabla c - \nabla \cdot (\rho D \nabla c) = 0 \quad (3.1.10)$$

et par l'équation de continuité (3.1.5) les deux premiers termes disparaissent. Il reste

$$\rho \frac{\partial c}{\partial t} + \rho \vec{v} \cdot \nabla c - \nabla \cdot (\rho D \nabla c) = 0 \quad (3.1.11)$$

C'est l'équation générale d'**advection-diffusion** pour un fluide en régime permanent. Supposons maintenant que la concentration de l'entité s'est stabilisée dans le fluide c.-à-d. $\frac{\partial c}{\partial t} = 0$. Supposons de plus que D et ρ sont constants dans tout le fluide et qu'on ajoute un terme f au membre de droite pour introduire une source (création de l'entité) ou un puit (élimination de l'entité), l'équation (3.1.11) se réécrit

$$\vec{v} \cdot \nabla c - D \nabla^2 c = f \quad (3.1.12)$$

où on a utilisé la relation suivante entre la divergence du gradient et le laplacien $\nabla \cdot (\nabla) = \nabla^2$. On obtient l'équation d'advection-diffusion stationnaire que nous

avons utilisée en dimension 1 à la dernière section du chapitre 2 et que nous étudierons maintenant en dimension 2.

Le **nombre de Peclet** est un nombre sans dimension qui représente le rapport entre le transfert de l'entité par diffusion et celui par advection. Si v_0 et L sont respectivement les échelles caractéristiques de vitesse et de taille du domaine, on a :

$$Pe = \frac{Lv_0}{D}$$

Pour l'équation adimensionnée, on peut poser $v_0 = L = 1$ auquel cas $Pe = 1/D$. On s'intéresse au problème lorsque le nombre de Peclet est élevé donc lorsque le coefficient de diffusion est faible.

Pour respecter la nomenclature des chapitres précédents et symboliser la différence d'échelle entre v et D on réécrit (3.1.12) de la façon suivante

$$\vec{\beta} \cdot \nabla u - \epsilon \nabla^2 u = f \quad (3.1.13)$$

avec $\epsilon > 0$.

3.2. PRÉSENTATION DU PROBLÈME

Soit $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$ on considère le problème

$$\begin{cases} Lu = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases}$$

avec L l'opérateur d'advection-diffusion provenant de l'équation (3.1.13) en 2 dimensions :

$$L = -\epsilon \frac{\partial^2}{\partial x^2} - \epsilon \frac{\partial^2}{\partial y^2} + \beta_1 \frac{\partial}{\partial x} + \beta_2 \frac{\partial}{\partial y} \quad (3.2.1)$$

où $\epsilon > 0$ est le coefficient de diffusion et $\vec{\beta}$ le champ de vitesse que l'on choisira constant sur le domaine. ϵ sera choisi très petit par rapport à $\|\vec{\beta}\|$ dans les simulations. Finalement, on construit une grille de représentation uniforme avec $h_x = 1/n_x$ et $h_y = 1/n_y$

$$\begin{aligned} x_i &= ih_x & i &= 0, 1, \dots, n_x \\ y_j &= jh_y & j &= 0, 1, \dots, n_y \end{aligned}$$

3.3. CONSTRUCTION DU SCHÉMA SUPERCONSISTANT

Le polynôme d'interpolation local $p_{i,j}$ utilise les 9 points $(x_k, y_m), k = i - 1, i, i + 1$ et $m = j - 1, j, j + 1$.

$$\begin{aligned}
p_{i,j}(x,y) &= \frac{1}{4h_x^2 h_y^2} (x-x_{i+1})(x-x_i)(y-y_{j+1})(y-y_j)u_{i-1,j-1} \\
&- \frac{1}{2h_x^2 h_y^2} (x-x_{i+1})(x-x_{i-1})(y-y_{j+1})(y-y_j)u_{i,j-1} \\
&+ \frac{1}{4h_x^2 h_y^2} (x-x_{i-1})(x-x_i)(y-y_{j+1})(y-y_j)u_{i+1,j-1} \\
&- \frac{1}{2h_x^2 h_y^2} (x-x_{i+1})(x-x_i)(y-y_{j+1})(y-y_{j-1})u_{i-1,j} \\
&+ \frac{1}{h_x^2 h_y^2} (x-x_{i+1})(x-x_{i-1})(y-y_{j+1})(y-y_{j-1})u_{i,j} \\
&- \frac{1}{2h_x^2 h_y^2} (x-x_{i-1})(x-x_i)(y-y_{j+1})(y-y_{j-1})u_{i+1,j} \\
&+ \frac{1}{4h_x^2 h_y^2} (x-x_{i+1})(x-x_i)(y-y_{j-1})(y-y_j)u_{i-1,j+1} \\
&- \frac{1}{2h_x^2 h_y^2} (x-x_{i+1})(x-x_{i-1})(y-y_{j-1})(y-y_j)u_{i,j+1} \\
&+ \frac{1}{4h_x^2 h_y^2} (x-x_{i-1})(x-x_i)(y-y_{j-1})(y-y_j)u_{i+1,j+1}
\end{aligned}$$

On évalue $Lp_{i,j}(\tau_i, \mu_j)$. On pose $X = x_i - \tau_i$ et $Y = y_j - \mu_j$. Après réorganisation des termes on trouve

$$\begin{aligned}
(L^h u^h)_{i,j} &= \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h_y^2} \left(-\epsilon - \beta_2 Y + \frac{\epsilon}{2h_x^2} (X^2 + Y^2) + \frac{XY}{2h_x^2} (\beta_1 Y + \beta_2 X) \right) \\
&+ \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h_x^2} \left(-\epsilon - \beta_1 X + \frac{\epsilon}{2h_y^2} (X^2 + Y^2) + \frac{XY}{2h_y^2} (\beta_1 Y + \beta_2 X) \right) \\
&+ \frac{u_{i+1,j} - u_{i-1,j}}{2h_x} \left(\beta_1 - \frac{2\epsilon X}{h_y^2} - \frac{\beta_1 Y^2}{h_y^2} - \frac{2\beta_2 XY}{h_y^2} \right) \\
&+ \frac{u_{i,j+1} - u_{i,j-1}}{2h_y} \left(\beta_2 - \frac{2\epsilon Y}{h_x^2} - \frac{\beta_2 X^2}{h_x^2} - \frac{2\beta_1 XY}{h_x^2} \right) \\
&+ \frac{u_{i-1,j-1} - u_{i-1,j} + u_{i-1,j+1} - u_{i,j-1} - u_{i,j+1} + u_{i+1,j-1} - u_{i+1,j} + u_{i+1,j+1}}{2h_x^2 h_y^2} \times \\
&\quad (-\epsilon(X^2 + Y^2) - XY(\beta_1 Y + \beta_2 X)) \\
&+ \frac{u_{i-1,j-1} + u_{i-1,j+1} - u_{i+1,j-1} - u_{i+1,j+1}}{4h_x^2 h_y^2} (-2\epsilon X h_x - \beta_1 h_x Y^2 - 2\beta_2 XY h_x) \\
&+ \frac{u_{i-1,j-1} - u_{i-1,j+1} + u_{i+1,j-1} - u_{i+1,j+1}}{4h_x^2 h_y^2} (-2\epsilon Y h_y - \beta_2 h_y X^2 - 2\beta_1 XY h_y) \\
&+ \frac{u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1} + u_{i+1,j+1}}{4h_x^2 h_y^2} (-h_x h_y (\beta_1 Y + \beta_2 X)) \\
&\approx Lu(\tau_i, \mu_j) \\
&= f(\tau_i, \mu_j)
\end{aligned} \tag{3.3.1}$$

où on reconnaît entre autre les dérivées premières et secondes par différences centrées.

Soit $g(x) := (x - x_{i-1})(x - x_i)(x - x_{i+1})$ et $h(y) := (y - y_{j-1})(y - y_j)(y - y_{j+1})$ on pose maintenant

$$\chi(x, y) := g(x) + h(y) \quad (3.3.2)$$

En résolvant $L\chi = 0$, on obtient les points de collocation tel que $Lg = Lh = 0$. Par la Proposition (1.3.2) on aura un schéma superconsistant puisque tout polynôme dans $\mathbb{P}_{3,0}$ peut s'écrire comme une combinaison de $g(x)$ et d'un polynôme dans $\mathbb{P}_{2,2}$ et qu'un polynôme dans $\mathbb{P}_{0,3}$ peut s'écrire comme une combinaison de $h(y)$ et d'un polynôme dans $\mathbb{P}_{2,2}$. On cherche les couples (x, y) tels que

$$L\chi(x, y) = Lg + Lh = -\epsilon g''(x) + \beta_1 g'(x) - \epsilon h''(y) + \beta_2 h'(y) = 0 \quad (3.3.3)$$

On obtient dans chaque variable la même équation que celle provenant du problème en une dimension (2.4.3). Après avoir posé $X = x_i - \tau_i$ et $Y = y_j - \mu_j$ on cherche donc les solutions de

$$\begin{aligned} 6\epsilon X + 3\beta_1 X^2 - \beta_1 h_x^2 &= 0 \\ 6\epsilon Y + 3\beta_2 Y^2 - \beta_2 h_y^2 &= 0 \end{aligned} \quad (3.3.4)$$

Elles s'écrivent

$$\begin{aligned} X &= \frac{\epsilon}{\beta_1} \left(\sqrt{1 + \frac{\beta_1^2 h_x^2}{3\epsilon^2}} - 1 \right) \\ Y &= \frac{\epsilon}{\beta_2} \left(\sqrt{1 + \frac{\beta_2^2 h_y^2}{3\epsilon^2}} - 1 \right) \end{aligned} \quad (3.3.5)$$

On peut montrer comme nous l'avons fait pour le problème en une dimension que

$$\begin{aligned} 0 &< |X| < \frac{h_x}{\sqrt{3}} \\ 0 &< |Y| < \frac{h_y}{\sqrt{3}} \end{aligned}$$

3.4. ANALYSE DE LA CONSISTANCE DU SCHÉMA

On considère l'expansion de Taylor de u autour de (τ_i, μ_j)

$$u(x, y) = u(\tau_i, \mu_j) + (x - \tau_i) \frac{\partial}{\partial x} u(\tau_i, \mu_j) + (y - \mu_j) \frac{\partial}{\partial y} u(\tau_i, \mu_j) + \frac{(x - \tau_i)^2}{2} \frac{\partial^2}{\partial x^2} u(\tau_i, \mu_j)$$

$$\begin{aligned}
& + (x - \tau_i)(y - \mu_j) \frac{\partial^2}{\partial x \partial y} u(\tau_i, \mu_j) + \frac{(y - \mu_j)^2}{2} \frac{\partial^2}{\partial y^2} u(\tau_i, \mu_j) + \frac{(x - \tau_i)^3}{6} \frac{\partial^3}{\partial x^3} u(\tau_i, \mu_j) \\
& + \frac{(x - \tau_i)^2(y - \mu_j)}{2} \frac{\partial^3}{\partial x^2 \partial y} u(\tau_i, \mu_j) + \frac{(x - \tau_i)(y - \mu_j)^2}{2} \frac{\partial^3}{\partial x \partial y^2} u(\tau_i, \mu_j) + \frac{(y - \mu_j)^3}{6} \frac{\partial^3}{\partial y^3} u(\tau_i, \mu_j) \\
& + \frac{(x - \tau_i)^4}{24} \frac{\partial^4}{\partial x^4} u(\tau_i, \mu_j) + \frac{(x - \tau_i)^3(y - \mu_j)}{6} \frac{\partial^4}{\partial x^3 \partial y} u(\tau_i, \mu_j) + \frac{(x - \tau_i)^2(y - \mu_j)^2}{4} \frac{\partial^4}{\partial x^2 \partial y^2} u(\tau_i, \mu_j) \\
& + \frac{(x - \tau_i)(y - \mu_j)^3}{6} \frac{\partial^4}{\partial x \partial y^3} u(\tau_i, \mu_j) + \frac{(y - \mu_j)^4}{24} \frac{\partial^4}{\partial y^4} u(\tau_i, \mu_j) + O((x - \tau_i)^5 + (y - \mu_j)^5)
\end{aligned}$$

En remplaçant les termes $u_{i,j}, u_{i,j-1}, \dots$ dans (3.3.1) par leur expansion de Taylor et en soustrayant $Lu(\tau_i, \mu_j)$ on trouve le résidu suivant.

$$\begin{aligned}
(L^h u^h)_{i,j} - Lu(\tau_i, \mu_j) &= \frac{1}{6}(\beta_1 h_x^2 - 3\beta_1 X^2 - 6\epsilon X) u_{xxx}(\tau_i, \mu_j) \\
&+ \frac{1}{6}(\beta_2 h_y^2 - 3\beta_2 Y^2 - 6\epsilon Y) u_{yyy}(\tau_i, \mu_j) \\
&+ \frac{1}{12}(h_x^2(X\beta_1 - \epsilon) - 2X^2(2X\beta_1 + 3\epsilon)) u_{xxx}(\tau_i, \mu_j) \quad (3.4.1) \\
&+ \frac{1}{12}(h_y^2(Y\beta_2 - \epsilon) - 2Y^2(2Y\beta_2 + 3\epsilon)) u_{yyy}(\tau_i, \mu_j) \\
&+ O(h^3)
\end{aligned}$$

où $h = \max(h_x, h_y)$. Posons

$$c_{xxx} = \frac{1}{6}(\beta_1 h_x^2 - 3\beta_1 X^2 - 6\epsilon X) \quad (3.4.2)$$

$$c_{yyy} = \frac{1}{6}(\beta_2 h_y^2 - 3\beta_2 Y^2 - 6\epsilon Y) \quad (3.4.3)$$

Ce sont les seuls coefficient d'ordre h dans le résidu donc le choix (3.3.5) permet d'augmenter l'ordre du schéma à h^2 puisque alors $c_{xxx} = c_{yyy} = 0$.

Comme pour le cas unidimensionnel, le choix $X = Y = 0$ assure aussi une consistance d'ordre 2 mais on verra que ce schéma n'est pas stable pour des petites valeurs de ϵ .

3.5. DEUXIÈME APPROCHE (FATONE ET AL.)

Notre choix de χ donné en (3.3.2) est motivé par notre définition de superconsistance et par la Proposition (1.3.2). Dans les travaux originaux [10, 12], l'annulation des coefficients d'ordre h , c_{xxx} et c_{yyy} , n'est pas le critère qui motive le choix de χ et Fatone et al. présentent une autre approche basée sur un argument de stabilité upwind. Soit $g(x)$ et $h(y)$ comme dans (3.3.2), ils posent

$$\chi(x, y) = g(x)h(y) = (x - x_{i-1})(x - x_i)(x - x_{i+1})(y - y_{j-1})(y - y_j)(y - y_{j+1}) \quad (3.5.1)$$

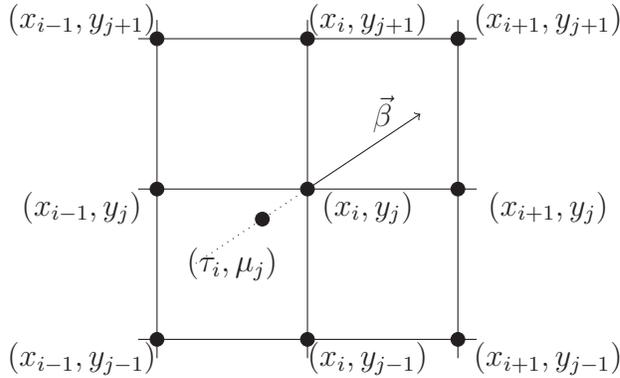


FIGURE 3.1. Localisation "upwind" du point de collocation dans la méthode de Fatone et al.

On obtient l'équation

$$L\chi(\tau_i, \mu_j) = (\beta_1(3X^2 - h_x^2) + 6\epsilon X)(-Y^3 + h^2 Y) + (\beta_2(3Y^2 - h_y^2) + 6\epsilon Y)(-X^3 + h^2 X) = 0$$

Ils choisissent alors de poser

$$\begin{aligned} X &= \alpha\beta_1 \\ Y &= \alpha\beta_2 \end{aligned} \tag{3.5.2}$$

et de résoudre l'équation résultante pour α .

Le but est de positionner (τ_i, μ_j) sur la droite passant par (x_i, y_j) dans la direction de $\vec{\beta}$ et selon un décalage upwind comme illustré sur la figure 3.1. Ils cherchent donc la plus petite valeur $\alpha > 0$ qui satisfait

$$-6\beta_1^2\beta_2^2\alpha^4 - 6(\beta_1^2 + \beta_2^2)\epsilon\alpha^3 + 4(h_y^2\beta_1^2 + h_x^2\beta_2^2)\alpha^2 + 6\epsilon(h_x^2 + h_y^2)\alpha - 2h_x^2h_y^2 = 0$$

La discrétisation résultante n'est pas superconsistante au sens de notre Définition 1.3.2. Cependant, l'espace Q tel que L et L^h coïncident sur la grille de collocation est plus grand que $\mathbb{P}_{2,2}$ car on a rajouté $\chi = g(x)h(y)$ à sa base vectorielle. Aussi, on verra que cette discrétisation est aussi d'ordre 2 donc on voudrait pouvoir dire qu'elle est aussi superconsistante. Pour cela, il faudrait idéalement généraliser leur approche à d'autres types d'opérateurs et démontrer qu'elle donne toujours un schéma d'ordre de consistance plus élevé. Pour un futur travail, il serait donc intéressant d'élargir notre définition de la superconsistance pour inclure un schéma comme celui de Fatone et al.

Remarque 3.5.1. Dans les cas particulier $\beta_1 = 0$ ou $\beta_2 = 0$, le problème devient unidimensionnel et les solutions (3.5.2) sont les même que (3.3.5). Aussi, si $\beta_1 =$

β_2 , alors $X = Y$ et on peut vérifier facilement que la méthode de Fatone et al. [10, 12] et celle décrite dans la Section 3.3 coïncident.

3.6. COMPARAISON DE LA LOCALISATION DES POINTS DE COLLOCATION

Supposons que $\beta_1, \beta_2 > 0$. Les solutions (3.3.5) s'écrivent

$$X = \sqrt{\frac{\epsilon^2}{\beta_1^2} + \frac{h_x^2}{3}} - \frac{\epsilon}{\beta_1}$$

$$Y = \sqrt{\frac{\epsilon^2}{\beta_2^2} + \frac{h_y^2}{3}} - \frac{\epsilon}{\beta_2}$$

On remarque que $\epsilon \rightarrow 0 \implies (X, Y) \rightarrow \left(\frac{h_x}{\sqrt{3}}, \frac{h_y}{\sqrt{3}} \right)$ et ce peu importe l'orientation de $\vec{\beta}$. Cependant si β_1 est négligeable devant β_2 mais pas devant ϵ , le vecteur $\vec{\beta}$ est orienté presque verticalement on aura $X \approx 0$ comme solution de (3.5.2). À la lumière de cet exemple, pour observer des différences entre les deux méthodes, le paramètre $\vec{\beta}$ devra être orienté dans une direction proche de celle d'un axe et ϵ devra être très petit. On peut noter qu'il est possible de faire varier la précision de la grille dans l'un ou l'autre des axes. Par exemple on peut prendre $h_x = \frac{\alpha}{\beta_2}$, $h_y = \frac{\alpha}{\beta_1}$ pour un certain α . De cette façon X et Y dépendent à la fois de β_1 et de β_2 . Cette fois si β_1 est négligeable devant β_2 on aura $h_x \ll h_y$ et la position du point de collocation de la méthode superconsistante approchera celle de la méthode "upwind".

On fait maintenant quelques tests numériques pour comparer la localisation des points selon les 2 méthodes. Dans le Tableau 3.1 on compare les solutions (X, Y) de (3.3.5) avec celles de (3.5.2) pour quelques $\vec{\beta} = \|\vec{\beta}\|(\cos \theta, \sin \theta)$ dans le premier quadrant. On donne aussi la valeur de c_{xxx} et c_{yyy} définis par (3.4.2) et (3.4.3) respectivement, pour les points de (3.5.2). On rappelle que ces coefficients sont toujours nuls pour les solutions de (3.3.5). On observe que les points coïncident lorsque $\theta = \pi/4$ et qu'ils sont très différents en $\theta = \pi/16$

Dans le Tableau 3.2 on a diminué la valeur de ϵ et on observe comme attendu que les solutions de (3.3.5) varient peu lorsque l'orientation de $\vec{\beta}$ change.

Dans le Tableau 3.3 on a augmenté la valeur de $\|\vec{\beta}\|$ en gardant le rapport $\epsilon/\|\beta\|$ constant. On observe que les points de collocation ne changent pas. Cependant c_{xxx} et c_{yyy} dépendent linéairement de β_1 et β_2 respectivement et comme ϵ est négligeable on remarque que la norme de c_{xxx} et c_{yyy} a augmenté linéairement.

| θ | Solutions de (3.3.5) | | Solutions de (3.5.2) | | c_{xxx} | c_{yyy} |
|-------------------|----------------------|----------|----------------------|----------|-----------|-----------|
| | X | Y | X | Y | | |
| $\frac{\pi}{16}$ | 0.048432 | 0.025947 | 0.058810 | 0.011698 | -0.000649 | 0.000194 |
| $\frac{\pi}{8}$ | 0.047917 | 0.037242 | 0.056739 | 0.023502 | -0.000514 | 0.000297 |
| $\frac{3\pi}{16}$ | 0.046947 | 0.042476 | 0.052541 | 0.035107 | -0.000287 | 0.000232 |
| $\frac{\pi}{4}$ | 0.045300 | 0.045300 | 0.045300 | 0.045300 | 0. | 0. |

TABLEAU 3.1. Comparaison des points de collocation pour $\epsilon = 1/100$, $\|\vec{\beta}\| = 1$, $h_x = h_y = 1/10$

| θ | Solutions de (3.3.5) | | Solutions de (3.5.2) | | c_{xxx} | c_{yyy} |
|-------------------|----------------------|----------|----------------------|----------|-----------|-----------|
| | X | Y | X | Y | | |
| $\frac{\pi}{16}$ | 0.056724 | 0.052836 | 0.069193 | 0.013763 | -0.000782 | 0.000292 |
| $\frac{\pi}{8}$ | 0.056663 | 0.055181 | 0.067780 | 0.028075 | -0.000650 | 0.000459 |
| $\frac{3\pi}{16}$ | 0.056545 | 0.055963 | 0.064315 | 0.042974 | -0.000398 | 0.000370 |
| $\frac{\pi}{4}$ | 0.056338 | 0.056338 | 0.056338 | 0.056338 | 0. | 0. |

TABLEAU 3.2. Comparaison des points de collocation pour $\epsilon = 1/1000$, $\|\vec{\beta}\| = 1$, $h_x = h_y = 1/10$

| θ | Solutions de (3.3.5) | | Solutions de (3.5.2) | | c_{xxx} | c_{yyy} |
|-------------------|----------------------|----------|----------------------|----------|-----------|-----------|
| | X | Y | X | Y | | |
| $\frac{\pi}{16}$ | 0.056724 | 0.052836 | 0.069193 | 0.013763 | -0.007824 | 0.002929 |
| $\frac{\pi}{8}$ | 0.056663 | 0.055181 | 0.067780 | 0.028075 | -0.006502 | 0.004589 |
| $\frac{3\pi}{16}$ | 0.056545 | 0.055963 | 0.064315 | 0.042974 | -0.003982 | 0.003700 |
| $\frac{\pi}{4}$ | 0.056338 | 0.056338 | 0.056338 | 0.056338 | 0. | 0. |

TABLEAU 3.3. Comparaison des points de collocation pour $\epsilon = 1/100$, $\|\vec{\beta}\| = 10$, $h_x = h_y = 1/10$

| θ | Solutions de (3.3.5) | | Solutions de (3.5.2) | | c_{xxx} | c_{yyy} |
|-------------------|----------------------|----------|----------------------|----------|-----------|-----------|
| | X | Y | X | Y | | |
| $\frac{\pi}{16}$ | 0.004843 | 0.002595 | 0.005881 | 0.001170 | -0.000065 | 0.000019 |
| $\frac{\pi}{8}$ | 0.004792 | 0.003724 | 0.005674 | 0.002350 | -0.000051 | 0.000030 |
| $\frac{3\pi}{16}$ | 0.004695 | 0.004248 | 0.005254 | 0.003511 | -0.000029 | 0.000023 |
| $\frac{\pi}{4}$ | 0.004530 | 0.004530 | 0.004530 | 0.004530 | 0. | 0. |

TABLEAU 3.4. Comparaison des points de collocation pour $\epsilon = 1/100$, $\|\vec{\beta}\| = 10$, $h_x = h_y = 1/100$

Finalement dans le Tableau 3.4 on a diminué la valeur de h_x et h_y et on observe que la norme de c_{xxx} et c_{yyy} a diminué quadratiquement.

Ce dernier résultat est discuté dans [10] et on peut remarquer que X et Y définis par (3.5.2) tendent vers 0 à la vitesse de h^2 . Ainsi, même si les points de collocation n'ont pas été choisis dans le but d'annuler directement les termes d'ordre h dans l'expansion de Taylor, ils entraînent une convergence quadratique

de ces derniers vers 0 lorsqu'on raffine le maillage. Le schéma de Fatone et al. est donc aussi consistant d'ordre 2.

En choisissant différents X et Y dans (3.3.1) on obtient les 3 schémas suivants.

- . De l'Isle-Owens (DO) (3.3.5)
- . Fatone et al. (F) (3.5.2)
- . Centré (C) ($X = Y = 0$)

Dans la suite de ce travail, en plus de comparer la méthode superconsistante à la méthode classique centrée, on étudiera aussi les différences entre la méthode superconsistante de De l'Isle Owens (DO) et celle de Fatone et al. (F). Pour éviter toute confusion, lorsqu'on parlera de méthode superconsistante, on fera référence à (DO). La méthode (F) ne sera pas qualifiée de superconsistante dans ce travail car elle ne satisfait pas à notre définition.

3.7. ANALYSE DE LA STABILITÉ DES SCHÉMAS DE L'ISLE-OWENS ET FATONE ET AL.

On remplace $u_{i,j}$ par $U_{i,j}$ dans l'opérateur (3.3.1) pour obtenir le système discret

$$L^h U = f(\tau_i, \mu_j) \quad (3.7.1)$$

Pour faire une étude de la stabilité on utilisera l'idée de Fatone et al. présentée dans [10]. Comme le développement mathématique dans l'article contenait quelques erreurs il n'est pas inutile de refaire la preuve dans sa totalité ici.

On multiplie l'équation (3.7.1) par $U_{i,j}$ et on somme sur les i, j . En utilisant les conditions frontières $U_{0,j} = U_{n,j} = U_{i,0} = U_{i,m} = 0$ on a les égalités (C.0.1) à (C.0.5) et on obtient

$$\begin{aligned} & \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} - \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} \left(-\epsilon - \beta_2 Y + \frac{\epsilon}{2h_x^2} (X^2 + Y^2) + \frac{XY}{2h_x^2} (\beta_1 Y + \beta_2 X) \right) \\ & - \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \left(-\epsilon - \beta_1 X + \frac{\epsilon}{2h_y^2} (X^2 + Y^2) + \frac{XY}{2h_y^2} (\beta_1 Y + \beta_2 X) \right) \\ & + \frac{(U_{i,j} - U_{i+1,j})^2 - (U_{i,j} - U_{i+1,j+1})^2 - (U_{i+1,j} - U_{i,j+1})^2 + (U_{i,j} - U_{i,j+1})^2}{2h_x^2 h_y^2} \times \\ & (-\epsilon(X^2 + Y^2) - XY(\beta_1 Y + \beta_2 X)) \\ & + \frac{(U_{i+1,j} - U_{i,j+1})^2 - (U_{i,j} - U_{i+1,j+1})^2}{4h_x^2 h_y^2} (-h_x h_y (\beta_1 Y + \beta_2 X)) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} \left(\epsilon + \beta_2 Y - \frac{\epsilon}{h_x^2} (X^2 + Y^2) - \frac{XY}{h_x^2} (\beta_1 Y + \beta_2 X) \right) \\
&\quad + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \left(\epsilon + \beta_1 X - \frac{\epsilon}{h_y^2} (X^2 + Y^2) - \frac{XY}{h_y^2} (\beta_1 Y + \beta_2 X) \right) \\
&\quad + \frac{(U_{i+1,j} - U_{i,j+1})^2}{4h_x^2 h_y^2} (-h_x h_y (\beta_1 Y + \beta_2 X) + 2\epsilon (X^2 + Y^2) + 2XY (\beta_1 Y + \beta_2 X)) \\
&\quad + \frac{(U_{i,j} - U_{i+1,j+1})^2}{4h_x^2 h_y^2} (h_x h_y (\beta_1 Y + \beta_2 X) + 2\epsilon (X^2 + Y^2) + 2XY (\beta_1 Y + \beta_2 X)) \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} f(\tau_i, \mu_j)
\end{aligned}$$

On pose $W = \epsilon(X^2 + Y^2) + XY(\beta_1 Y + \beta_2 X)$ et en utilisant (C.0.6), on peut simplifier la dernière équation

$$\begin{aligned}
&\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} \left(\epsilon + \beta_2 Y - \frac{W}{h_x^2} \right) + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \left(\epsilon + \beta_1 X - \frac{W}{h_y^2} \right) \\
&\quad + \frac{\beta_1 Y + \beta_2 X}{2h_x h_y} ((U_{i,j} - U_{i+1,j})(U_{i+1,j} - U_{i+1,j+1}) + (U_{i,j} - U_{i,j+1})(U_{i,j} - U_{i+1,j})) \\
&\quad + \left(\frac{W}{2} \right) \left(\frac{(U_{i,j} - U_{i+1,j+1})^2}{h_x^2 h_y^2} + \frac{(U_{i+1,j} - U_{i,j+1})^2}{h_x^2 h_y^2} \right) \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} f(\tau_i, \mu_j)
\end{aligned} \tag{3.7.2}$$

À partir de ce point on étudie des cas spécifiques

3.7.1. Stabilité dans le cas particulier $\beta_1 = 0$ (ou $\beta_2 = 0$)

Considérons le cas $\beta_1 = 0$. On a alors $X = 0$ et $W = \epsilon Y^2$. L'équation (3.7.2) devient

$$\begin{aligned}
&\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} f(\tau_i, \mu_j) = \\
&\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} \left(\epsilon + \beta_2 Y - \frac{\epsilon Y^2}{h_x^2} \right) + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \left(\epsilon - \frac{\epsilon Y^2}{h_y^2} \right)
\end{aligned}$$

$$\begin{aligned}
& + \left(\frac{\epsilon Y^2}{2} \right) \left(\frac{(U_{i,j} - U_{i+1,j+1})^2}{h_x^2 h_y^2} + \frac{(U_{i+1,j} - U_{i,j+1})^2}{h_x^2 h_y^2} \right) \\
& \geq \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \left(\beta_2 Y + \epsilon \left(1 - \frac{Y^2}{h_x^2} \right) \right) \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2}
\end{aligned}$$

Pour un maillage tel que $h_x \neq h_y$ on a besoin de la contrainte supplémentaire $Y < h_x$. Ainsi $\left(1 - \frac{Y^2}{h_x^2} \right) > 0$ et comme on a toujours $\beta_2 Y > 0$, les inégalités de Cauchy-Schwartz (B.0.1) et de Poincaré (B.0.2) impliquent

$$\|U\| \leq \frac{1}{\left(\beta_2 Y + \epsilon \left(1 - \frac{Y^2}{h_x^2} \right) \right)} \|f\|$$

Pour $\beta_1 = 0$, les points de collocation des schémas de De l'Isle-Owens et de Fatone et al. coïncident et on rappelle que les solutions pour X et Y sont

$$\begin{aligned}
& X = 0 \\
& Y = \frac{\epsilon}{\beta_2} \left(\sqrt{1 + \frac{\beta_2^2 h_y^2}{3\epsilon^2}} - 1 \right)
\end{aligned}$$

La borne de $\|U\|$ est donc

$$\|U\| \leq \frac{1}{\left(\sqrt{\epsilon^2 + \frac{\beta_1^2 h_y^2}{3}} - \frac{\epsilon Y^2}{h_x^2} \right)} \|f\|$$

Lorsque $\epsilon \rightarrow 0$, la solution reste bornée.

$$\|U\| \leq \frac{\sqrt{3}}{|\beta_1| h_y} \|f\|$$

Aussi, en utilisant la règle de L'Hôpital on peut vérifier que $Y/h_y \rightarrow 0$ lorsque $h_y \rightarrow 0$. Donc si on raffine le maillage de façon à ce qu'il existe une constante C tel que $h_y < Ch_x$, la solution reste bornée lorsque $h_x, h_y \rightarrow 0$.

$$\|U\| \leq \frac{1}{\epsilon} \|f\|$$

Le cas $\beta_2 = 0$ est analogue avec $\|U\| \leq \frac{1}{\left(\beta_1 X + \epsilon \left(1 - \frac{X^2}{h_y^2} \right) \right)} \|f\|$

3.7.2. Stabilité de la méthode de Fatone et al.

On reprend (3.7.2). Avec $X = \alpha\beta_1$ et $Y = \alpha\beta_2$, en choisissant η tel que

$$0 < \eta \leq \min \left(\epsilon - \frac{W}{h_y^2}, \epsilon - \frac{W}{h_x^2} \right) \quad (3.7.3)$$

et en utilisant le fait que $W > 0$ on a

$$\begin{aligned} & \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} f(\tau_i, \mu_j) \geq \\ & \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \eta \left(\frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \right) \\ & \quad + \alpha \beta_2^2 \frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} + \alpha \beta_1^2 \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \\ & \quad + \frac{\alpha \beta_1 \beta_2}{h_x h_y} ((U_{i,j} - U_{i+1,j})(U_{i+1,j} - U_{i+1,j+1}) + (U_{i,j} - U_{i,j+1})(U_{i,j} - U_{i+1,j})) \\ & = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \eta \left(\frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \right) \\ & \quad + \frac{1}{2} \left(\sqrt{\alpha \beta_2^2} \frac{(U_{i+1,j} - U_{i+1,j+1})}{h_y} + \sqrt{\alpha \beta_1^2} \frac{(U_{i,j} - U_{i+1,j})}{h_x} \right)^2 \\ & \quad + \frac{1}{2} \left(\sqrt{\alpha \beta_2^2} \frac{(U_{i,j} - U_{i,j+1})}{h_y} + \sqrt{\alpha \beta_1^2} \frac{(U_{i,j} - U_{i+1,j})}{h_x} \right)^2 \\ & \geq \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \eta \left(\frac{(U_{i,j} - U_{i,j+1})^2}{h_y^2} + \frac{(U_{i,j} - U_{i+1,j})^2}{h_x^2} \right) \\ & \geq \eta \sum_{i=0}^n \sum_{j=0}^m U_{i,j}^2 \end{aligned}$$

La dernière inégalité utilise (B.0.3). Avec l'inégalité de Cauchy-Schwartz (B.0.1) on obtient

$$\|U\| \|f\| \geq \eta \|U\|^2 \implies \|U\| \leq \frac{1}{\eta} \|f\| \quad (3.7.4)$$

Malheureusement, la borne dépend de ϵ . Considérons h suffisamment petit pour obtenir un η qui respecte la contrainte (3.7.3), supposons $h = h_x = h_y$ et posons β tel que $0 < \beta = \max(\beta_1, \beta_2)$. On assume que la borne supérieure sur X et Y est la même que dans la méthode De l'Isle-Owens c'est à dire que $0 < X \leq h/\sqrt{3}$ et $0 < Y \leq h/\sqrt{3}$. On a alors

$$W \leq \frac{2\epsilon h^2}{3} + \frac{2h^3\beta}{3\sqrt{3}}$$

$$\implies \epsilon - \frac{W}{h^2} \geq \frac{\epsilon}{3} - \frac{2h\beta}{3\sqrt{3}} > 0 \iff h < \frac{\epsilon\sqrt{3}}{2\beta}$$

Donc h doit être de l'ordre de grandeur de ϵ/β . Par exemple avec $h = \frac{\epsilon\sqrt{3}}{4\beta}$ on a

$$W \leq \frac{2\epsilon h^2}{3} + \frac{h^2\epsilon}{6} = \frac{5\epsilon h^2}{6}$$

et donc

$$\epsilon - W/h^2 \geq \frac{\epsilon}{6}$$

le résultat (3.7.4) devient

$$\|U\| \leq \frac{6}{\epsilon} \|f\|$$

Le schéma est donc stable lorsque $h \rightarrow 0$ mais nous n'avons pas trouvé de borne indépendante de ϵ pour assurer que la solution ne présentera pas d'oscillation lorsque $\epsilon \rightarrow 0$.

3.7.3. Stabilité de la méthode De l'Isle-Owens

Dans le cas particulier $h_x = h_y$, $\beta_1 = \beta_2$ on a $X = Y$ et on peut obtenir le même résultat de stabilité que pour la méthode Fatone et al.

Ensuite, lorsque $h_x, h_y \rightarrow 0$ il va exister un maillage suffisamment petit tel que

$$h_x < \frac{\sqrt{3}\epsilon}{|\beta_1|} \quad \text{et} \quad h_y < \frac{\sqrt{3}\epsilon}{|\beta_2|}$$

le développement limité $\sqrt{1+x} = 1 + \frac{x}{2} + O(x^2)$ pour $|x| < 1$ permet d'obtenir les approximations

$$X = \frac{\epsilon}{\beta_1} \left(\sqrt{1 + \frac{\beta_1^2 h_x^2}{3\epsilon^2}} - 1 \right) \approx \frac{\beta_1 h_x^2}{6\epsilon}$$

$$Y = \frac{\epsilon}{\beta_2} \left(\sqrt{1 + \frac{\beta_2^2 h_y^2}{3\epsilon^2}} - 1 \right) \approx \frac{\beta_2 h_y^2}{6\epsilon}$$

et pour $h_x = h_y = h$ on remarque que X et Y peuvent s'écrire comme (3.5.2) avec $\alpha = \frac{h^2}{6\epsilon}$.

Si le maillage est trop grossier pour utiliser l'approximation par le développement limité de la racine, on peut cependant assumer que $\frac{|\beta_1| h_x}{\sqrt{3}\epsilon} \gg 1$ et $\frac{|\beta_2| h_y}{\sqrt{3}\epsilon} \gg 1$ auquel cas on a les approximations

$$X \approx \frac{h_x}{\sqrt{3}}$$

$$Y \approx \frac{h_y}{\sqrt{3}}$$

On a mentionné précédemment que l'on pouvait choisir h_x différent de h_y afin que le point de collocation dans la méthode De l'Isle-Owens soit plus en accord avec l'orientation de $\vec{\beta}$.

Soit $\alpha > 0$, posons $h_x = \alpha\beta_1$ et $h_y = \alpha\beta_2$. Le point de collocation est maintenant situé sur la droite engendrée par $\vec{\beta}$ et passant par (x_i, y_j)

$$X \approx \frac{\alpha\beta_1}{\sqrt{3}}$$

$$Y \approx \frac{\alpha\beta_2}{\sqrt{3}}$$

Et on peut à partir d'ici conclure le même résultat de stabilité que Fatone et al. Nous n'avons pas non plus trouvé de borne indépendante de ϵ pour la méthode De l'Isle Owens.

3.8. ALGORITHME ET TEST NUMÉRIQUE

On teste l'efficacité des trois méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C) en résolvant numériquement sur un domaine carré en 2 dimensions le problème d'advection-diffusion présenté à la section précédente

$$\begin{cases} Lu = f & \text{dans } \Omega = (0, 1) \times (0, 1) \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (3.8.1)$$

Avec

$$L = -\epsilon\nabla^2 + \vec{\beta} \cdot \nabla = -\epsilon \frac{\partial^2}{\partial x^2} - \epsilon \frac{\partial^2}{\partial y^2} + \beta_1 \frac{\partial}{\partial x} + \beta_2 \frac{\partial}{\partial y}$$

3.8.1. Construction du problème de test

Supposons qu'on essaie de construire un problème d'advection-diffusion pour tester notre méthode de résolution numérique on aimerait que

1. le membre de droite f ainsi que les conditions frontières soient indépendantes de ϵ et $\vec{\beta}$
2. la solution u analytique soit connue
3. la solution u analytique soit calculable en tenant compte de la précision machine

Le critère 1 permet de s'assurer que la solution u dépend du facteur $\epsilon/||\vec{\beta}||$ et que pour $\epsilon \rightarrow 0$ une couche limite se crée. Les critères 2 et 3 assurent que l'erreur

$\|u^h - U\|$ ainsi que l'ordre de convergence de la solution pourront être calculés ce qui permettra de comparer adéquatement les méthodes.

Posons $f = 0$ de sorte que l'on puisse appliquer la technique de résolution par séparation de variable en posant :

$$u(x, y) = X(x)Y(y) \quad (3.8.2)$$

Pour avoir une solution non triviale on choisit des conditions frontières non-homogènes. Le problème à résoudre est :

$$\left\{ \begin{array}{l} Lu = 0 \quad \text{dans } \Omega = (0, 1) \times (0, 1) \\ u(x, 0) = 0 \\ u(x, 1) = 0 \\ u(1, y) = 0 \\ u(0, y) = g(y) \end{array} \right. \quad (3.8.3)$$

avec L défini par (3.2.1). En utilisant (3.8.2) on obtient

$$Lu = 0 \iff -\epsilon \frac{X''}{X} + \beta_1 \frac{X'}{X} = \lambda = \epsilon \frac{Y''}{Y} - \beta_2 \frac{Y'}{Y}$$

et les conditions frontières deviennent

$$Y(0) = Y(1) = X(1) = 0$$

$$Y(y)X(0) = g(y)$$

On doit résoudre les 2 équations différentielles homogènes d'ordre 2 linéaires à coefficients constants suivantes :

$$-\epsilon X'' + \beta_1 X' - \lambda X = 0 \quad (3.8.4)$$

$$-\epsilon Y'' + \beta_2 Y' + \lambda Y = 0 \quad (3.8.5)$$

Selon la nature (complexe ou réelles) des racines des équations caractéristiques de (3.8.4) et (3.8.5), les solutions de ce type d'équation sont sinusoidales ou exponentielles (voir par exemple [3]). Les conditions frontières homogènes pour la fonction Y imposent que le discriminant de l'équation caractéristique de (3.8.5) $-\epsilon t^2 + \beta_2 t + \lambda$, soit négatif pour obtenir une solution sinusoidale. On a donc

$$\beta_2^2 + 4\epsilon\lambda < 0 \implies \lambda < -\frac{\beta_2^2}{4\epsilon}$$

La solution Y qui satisfait $Y(0) = Y(1) = 0$ est

$$Y(y) = A_n \exp\left(\frac{\beta_2 y}{2\epsilon}\right) \sin(n\pi y)$$

avec

$$n \in \mathbb{N}, \quad \lambda = -\frac{4\epsilon^2 n^2 \pi^2 + \beta_2^2}{4\epsilon}$$

Cela entraîne que le discriminant de l'équation caractéristique de (3.8.4) satisfait

$$\beta_1^2 - 4\epsilon\lambda = \beta_1^2 + \beta_2^2 + 4\epsilon^2 n^2 \pi^2 > 0$$

Donc les racines de l'équation caractéristique pour X sont réelles et la solution exponentielle qui satisfait $X(1) = 0$ est

$$X(x) = B_n \exp\left(\frac{\beta_1 x}{2\epsilon} - \gamma\right) \sinh(\gamma(1-x)) \quad (3.8.6)$$

où

$$\gamma = \frac{\sqrt{\beta_1^2 - 4\epsilon\lambda}}{2\epsilon}$$

En combinant les constantes on obtient la famille de solutions

$$u_n(x, y) = X(x)Y(y) = C_n \sin(n\pi y) \exp\left(\frac{\beta_1 x + \beta_2 y}{2\epsilon} - \gamma\right) \sinh(\gamma(1-x)) \quad n \in \mathbb{N} \quad (3.8.7)$$

Comme L est linéaire, une somme de u_n est aussi une solution et pour une condition frontière donnée $X(0)Y(y) = g(y)$ la solution générale s'écrira potentiellement comme une série infinie

$$u(x, y) = \sum_{n=1}^{\infty} u_n(x, y)$$

Le calcul numérique de cette série génère cependant des erreurs d'arrondi et est souvent impraticable dû entre autres à la présence du terme $\exp\left(\frac{\beta_1 x + \beta_2 y}{2\epsilon}\right)$ qui varie considérablement dans le domaine lorsque $\epsilon \rightarrow 0$. Le critère 3 énoncé en début de section n'est donc pas satisfait pour un $\vec{\beta}$ arbitraire. Dans [15] on présente un exemple pour un cas particulier de problème de la forme (3.8.3) en posant $\beta_1 = 1$, $\beta_2 = 0$ et $g(y) = \sin \pi y$. La solution a alors pour seul terme $u_1(x, y)$ et est donnée par

$$u(x, y) = \exp\left(\frac{x}{2\epsilon}\right) \sin(\pi y) \frac{\sinh(\gamma(1-x))}{\sinh \gamma} \quad (3.8.8)$$

avec $\gamma = \frac{\sqrt{1 + 4\pi^2 \epsilon^2}}{2\epsilon}$

L'inconvénient dans le choix de ces paramètres est que la méthode de De l'Isle-Owens et celle de Fatone et al. coïncident sur ce problème puisque $\vec{\beta}$ est dans la direction d'un des axes.

On fera donc 2 séries de test. Dans la première on comparera l'erreur et l'ordre de convergence des solutions données par la méthode superconsistante et par la méthode centrée sur base du problème avec $\beta_2 = 0$. Dans la deuxième on comparera l'allure des surfaces des solutions numériques ainsi que leur variation totale pour les 3 méthodes De l'Isle Owens, Fatone et al. et centrée sur base d'un problème plus général mais dont la solution exacte est inconnue.

3.8.2. Résultats numériques d'un problème avec solution connue

On considère le problème

$$\left\{ \begin{array}{ll} -\epsilon u_{xx} - \epsilon u_{yy} + u_x = 0 & \text{dans } \Omega = (0, 1) \times (0, 1) \\ u(x, y) = 0 & \text{sur } \partial\Omega \setminus \{(x, y) | x = 0\} \\ u(0, y) = \sin(\pi y) & \end{array} \right. \quad (3.8.9)$$

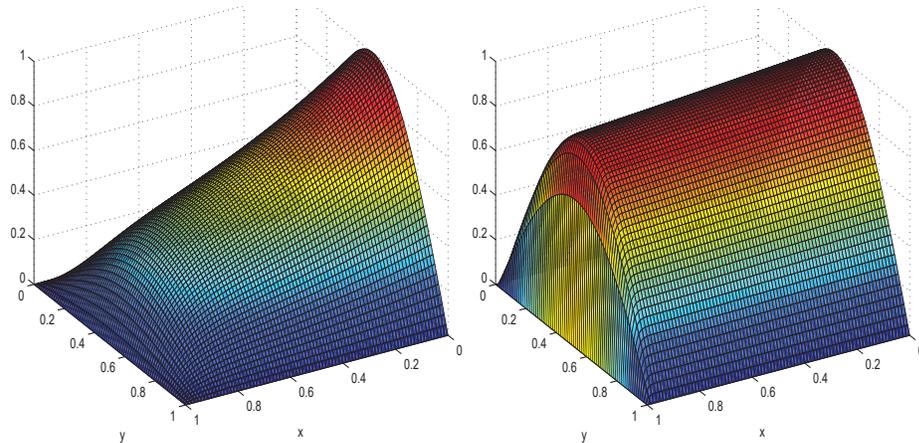


FIGURE 3.2. Graphique de la solution exacte (3.8.8) avec $\epsilon = 0.1$ (gauche) et $\epsilon = 0.01$ (droite)

La solution exacte u est donnée par (3.8.8). Pour calculer la solution discrète U on utilisera un maillage de taille $(n+1) \times (n+1)$ avec donc $h_x = h_y = h = 1/n$ et on pourra calculer l'erreur donnée par

$$e(h) = \frac{1}{(n-1)} \|U - u^h\|_2 = \sqrt{\frac{1}{(n-1)^2} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (U_{i,j} - u(x_i, y_j))^2} \quad (3.8.10)$$

Comme expliqué à la Section 2.4.3.1, la pente des graphiques de la Figure 3.3 correspond à l'ordre de convergence. Comme pour le problème en dimension 1, pour calculer adéquatement l'ordre de convergence il faut éliminer les portions de la courbe où n est trop petit ou trop grand. Pour des raisons de puissance de calcul limitée, nous ne pouvons pas utiliser un maillage aussi fin que lorsque nous avons résolu le problème en dimension 1 et pour $\epsilon \leq 1/1000$ l'ordre de convergence désiré pour la méthode superconsistante n'est jamais atteint (voir Tableau 3.8). Une possible explication serait que les maillages grossiers ne parviennent pas à "capturer" la couche limite lorsque celle-ci est trop fine. On avait déjà observé ce phénomène en dimension 1 dans la Figure 2.2 où l'ordre de convergence attendu apparaissait pour des maillages avec h comparable à $|\epsilon \log \epsilon|$ lorsque ϵ était petit. Si la convergence de l'erreur est limitée pour des maillages grossiers, on peut cependant, observer dans la Figure 3.3 ainsi que dans les Tableaux 3.5 à 3.7 que la magnitude de l'erreur pour la méthode superconsistante est nettement inférieure à celle de l'erreur pour la méthode centrée. La méthode superconsistante a donc encore comme en dimension 1 un net avantage quant à la stabilité des solutions. En effet, en observant les Figures 3.4 à 3.7 on voit que même pour ϵ aussi petit que 10^{-5} et avec des maillages grossiers ($n = 16, 32$) la surface de la solution pour la méthode superconsistante ne présente aucune instabilité. La méthode centrée au contraire échoue à calculer adéquatement la valeur de la solution proche de la frontière $x = 1$. Pour $\epsilon = 1/100$ on observe déjà des oscillations à la frontière lorsque n est trop petit et pour $\epsilon = 1/100000$ les oscillations sont présentes sur tout le domaine.

| n | e_{DO} | e_C |
|------|----------------|----------------|
| 8 | 0.004386 | 0.007341 |
| 16 | 0.001245 | 0.001757 |
| 32 | 0.0003195 | 0.0004317 |
| 64 | $8.037e - 005$ | 0.0001074 |
| 128 | $2.012e - 005$ | $2.682e - 005$ |
| 256 | $5.033e - 006$ | $6.703e - 006$ |
| 512 | $1.258e - 006$ | $1.676e - 006$ |
| 1024 | $3.146e - 007$ | $4.189e - 007$ |

TABLEAU 3.5. Comparaison des erreurs $\|U - u^h\|$ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.1$, $\beta_1 = 1$, $\beta_2 = 0$

| n | e_{DO} | e_C |
|------|-----------|-----------|
| 8 | 0.0178 | 3.384 |
| 16 | 0.01263 | 0.79 |
| 32 | 0.009006 | 0.2332 |
| 64 | 0.006569 | 0.1066 |
| 128 | 0.005175 | 0.04552 |
| 256 | 0.004247 | 0.01575 |
| 512 | 0.002368 | 0.00407 |
| 1024 | 0.0007994 | 0.0009389 |

TABLEAU 3.6. Comparaison des erreurs $\|U - u^h\|$ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = 1$, $\beta_2 = 0$

| n | e_{DO} | e_C |
|------|----------|---------|
| 8 | 0.01799 | 362.7 |
| 16 | 0.01272 | 95.56 |
| 32 | 0.008997 | 24.05 |
| 64 | 0.006362 | 5.861 |
| 128 | 0.004499 | 1.336 |
| 256 | 0.003181 | 0.3352 |
| 512 | 0.00225 | 0.1528 |
| 1024 | 0.001592 | 0.07562 |

TABLEAU 3.7. Comparaison des erreurs $\|U - u^h\|$ obtenues par les 2 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 1e - 005$, $\beta_1 = 1$, $\beta_2 = 0$

| ϵ | $Ordre_{DO}$ | $Ordre_C$ |
|------------|--------------|-----------|
| 1 | 2.00 | 2.00 |
| .1 | 1.98 | 2.01 |
| .01 | 1.97 | 2.02 |
| .001 | 1.20 | 2.03 |
| .0001 | 0.43 | 1.13 |
| .00001 | 0.50 | 1.07 |

TABLEAU 3.8. Ordre de convergence de l'erreur en fonction de $h = \frac{1}{n}$ (calculé comme la pente du graphique log-log) pour chacune des deux méthodes avec $\beta_1 = 1$, $\beta_2 = 0$

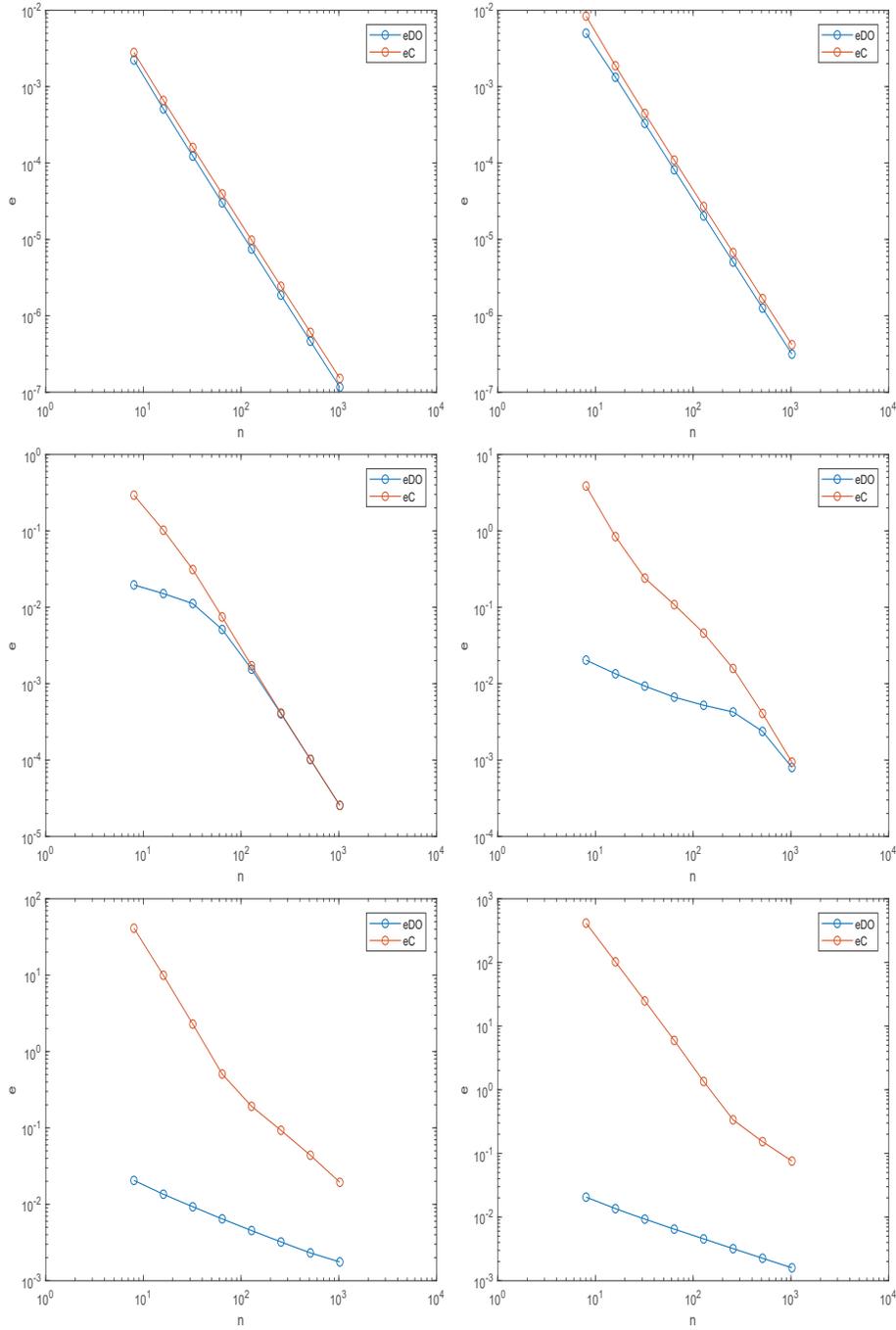


FIGURE 3.3. Graphique log-log de l'évolution de l'erreur $\|U - u^h\|$ en fonction de $h = \frac{1}{n}$ pour les 2 méthodes avec $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001, 0.00001$ (de gauche à droite puis de haut en bas), $\beta_1 = 1$ et $\beta_2 = 0$

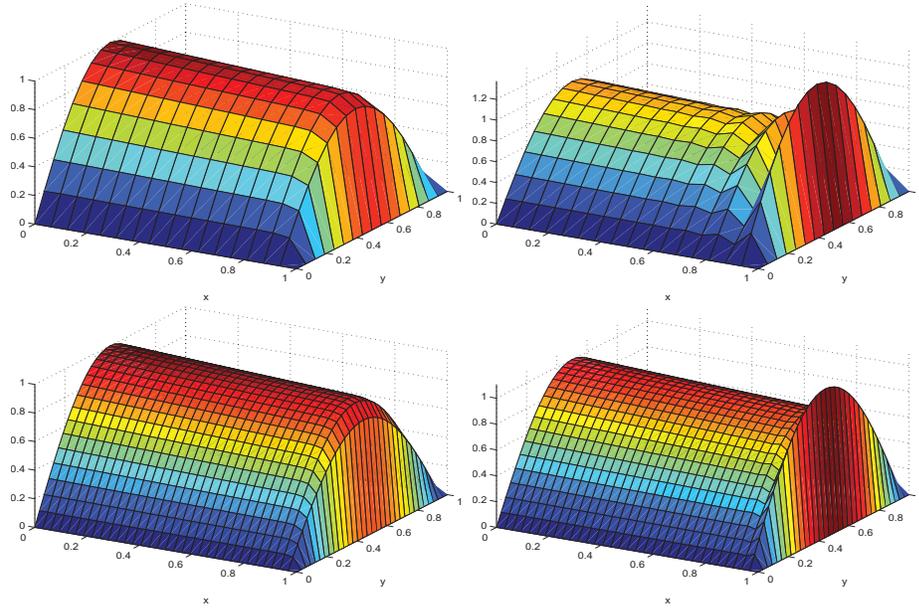


FIGURE 3.4. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 0.01$, $\beta_1 = 1$, $\beta_2 = 0$

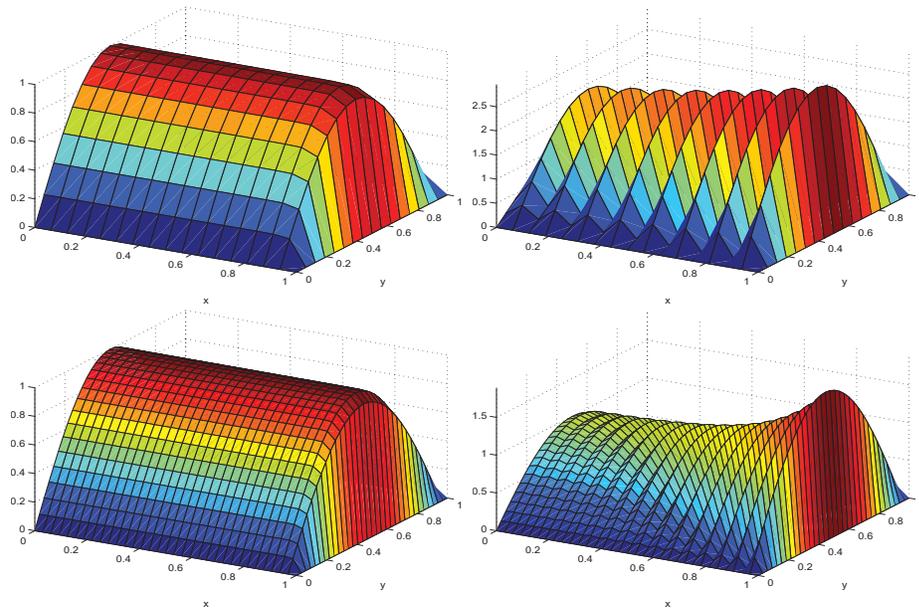


FIGURE 3.5. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 0.001$, $\beta_1 = 1$, $\beta_2 = 0$

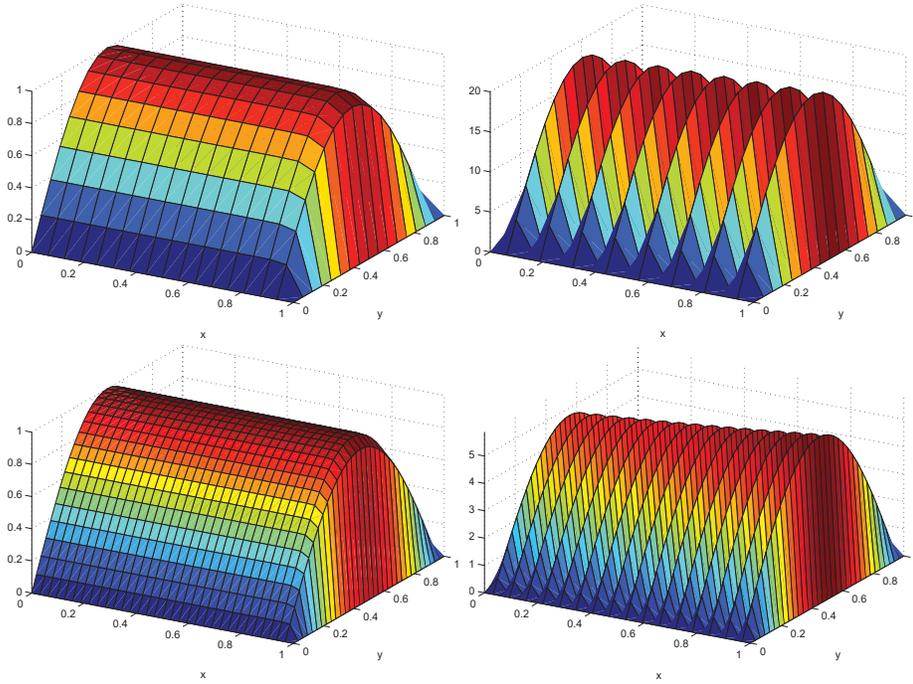


FIGURE 3.6. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 1e - 4$, $\beta_1 = 1$, $\beta_2 = 0$

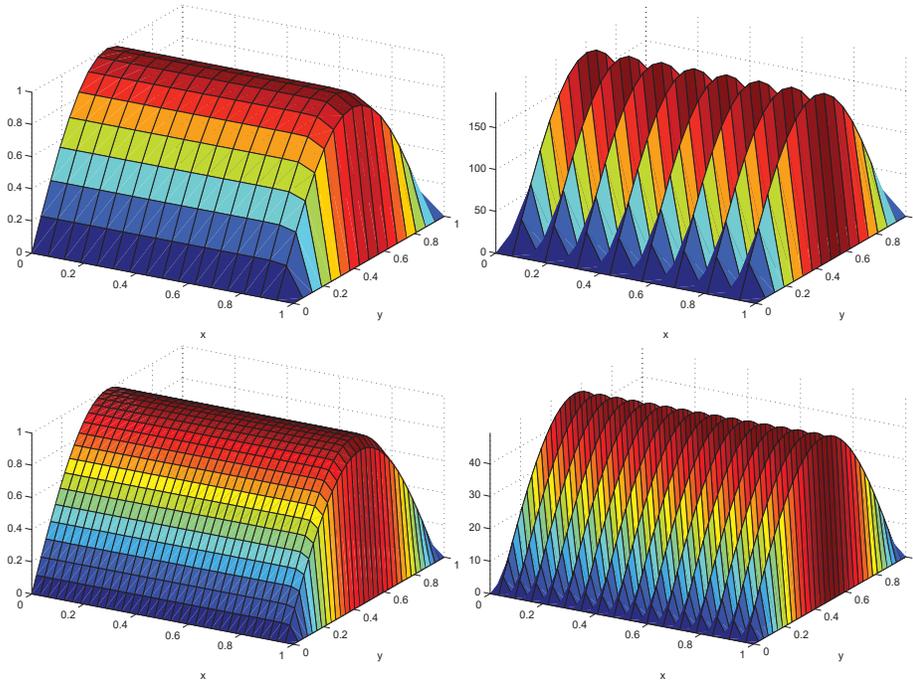


FIGURE 3.7. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), et centrée (droite) pour les maillages $n = 16$ (haut) et $n = 32$ (bas) et avec $\epsilon = 1e - 5$, $\beta_1 = 1$, $\beta_2 = 0$

3.8.3. Résultats numériques d'un problème avec solution exacte inconnue

Dans les prochaines simulations, on pose $f = 1$ avec les conditions Dirichlet homogènes à la frontière. On calculera la variation totale selon la formule suivante

$$var = \sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} |U_{i+1,j} - U_{i,j}| + \sum_{i=0}^{n_x-1} \sum_{j=0}^{n_y-1} |U_{i,j+1} - U_{i,j}|$$

La variation totale grandit en présence d'oscillations donc une valeur moins élevée témoigne en général d'une meilleure solution. Seulement, la solution exacte ne présentant pas une variation totale nulle on ne peut pas toujours conclure qu'une solution est plus mauvaise qu'une autre si sa variation totale est sensiblement plus élevée. Seuls des grands écarts de variations permettent de tirer des conclusions. Pour des petits écarts, on devra corroborer avec l'allure des surfaces des solutions sur les Figures présentées. On choisit le paramètre $\vec{\beta}$ comme un vecteur de longueur 1 de différentes orientations. On testera une orientation proche de $\pi/4$ puis une proche de $\pi/32$ puis une proche de 0. Dans chaque cas on testera 2 valeurs de ϵ soit $1/1000$ et $1/10000$. Dans tous les tests, la méthode centrée présente beaucoup d'oscillations que l'on peut observer autant dans les Figures 3.8 à 3.13 que dans les Tableaux de variation 3.9 à 3.14. Ces oscillations diminuent avec h mais pour ϵ petit le maillage doit être extrêmement fin pour les faire disparaître complètement. Ainsi on peut voir dans ces six Tableaux la nette différence entre les variations totales pour $\epsilon = 0.001$ et pour $\epsilon = 0.0001$ dans le cas de la méthode centrée. En comparaison, les 2 autres méthodes sont moins sensibles aux variations de ϵ .

Lorsque l'angle de $\vec{\beta}$ est proche de $\pi/4$, les 2 méthodes De l'Isle-Owens et Fatone et al. donnent des solutions semblables même pour des petites valeurs de ϵ comme on peut voir dans les Figures 3.8 et 3.9. De plus, dans les Tableaux 3.9 et 3.10, il n'y a pas de différence notable dans les variations totales entre les deux méthodes.

Des différences entre les deux méthodes De l'Isle-Owens et Fatone et al. apparaissent lorsque l'orientation du vecteur $\vec{\beta}$ s'approche d'un des axes. En l'occurrence, pour un angle proche de $\pi/32$ une oscillation apparaît à la frontière $y = 0$ pour la méthode De l'Isle Owens et $y = 1$ pour la méthode Fatone et al. comme le montre la Figure 3.10. Cette différence est à peine visible dans le calcul de la variation totale au Tableau 3.11. On constate aussi dans la Figure que l'oscillation pour la méthode Fatone et al. disparaît moins vite lorsque $h \rightarrow 0$. On remarque aussi que la méthode Fatone et al. est plus sensible au paramètre

ϵ . En effet l'oscillation sur la Figure 3.11 est plus grande et un raffinement du maillage n'améliore pas beaucoup la solution. Le Tableau 3.12 témoigne aussi de ce résultat et on voit que l'écart de variation entre la méthode De l'Isle Owens et celle de Fatone et al. est toujours important même lorsque $n = 256$. Finalement, pour $\vec{\beta}$ orienté très proche d'un axe c'est la méthode de Fatone et al. qui semble mieux fonctionner. On voit en effet qu'entre les Tableaux 3.13 et 3.14 et entre les Figures 3.12 et 3.13, la diminution du paramètre ϵ génère une oscillation proche de la frontière $y = 0$ pour la méthode de De l'Isle Owens. Aucune oscillation n'est visible pour la méthode de Fatone et al. même pour des maillages grossiers. On note cependant qu'un raffinement du maillage améliore la solution donnée par la méthode de De l'Isle Owens comme on le voit sur la Figure 3.13 pour $n = 32$. Des tests supplémentaires ont été fait avec $\vec{\beta}$ non constant en annexe. Dans les Figures E.1 à E.5 on a choisi $\vec{\beta}(x, y) = r \cdot ((y - 1/2)\vec{i} + (1/2 - x)\vec{j})$ pour simuler un champ circulaire. En faisant varier r on génère des instabilités. Pour $r = 1$ les deux méthodes fonctionnent bien (E.1). Pour $r = 5$ et $r = 10$, la méthode De l'Isle-Owens présente moins d'oscillations au centre du vortex comme on le voit sur les Figures E.2 et E.4. Un raffinement du maillage montre une nette amélioration pour les deux méthodes mais la méthode Fatone et al. reste toutefois moins performante basé sur l'allure de la solution sur les Figures E.3 et E.5. Il y a une moins bonne symétrie radiale sur la surface de droite.

En conclusion, les méthodes de De l'Isle Owens et de Fatone et al. sont nettement meilleures que le choix de collocation centré pour résoudre l'équation d'advection-diffusion. Il est cependant difficile de départager les deux méthodes car chacune apporte un avantage dépendamment de l'orientation de $\vec{\beta}$. Dans des situations où le champ de vitesse est orienté proche de la direction d'un des axes (mais sans coïncider) il semble que celle de Fatone et al. se comporte mieux. Lorsque ce n'est pas le cas, donc dans le reste de nos tests, les solutions données par la méthode de De l'Isle Owens sont sensiblement meilleures en particulier car les oscillations disparaissent plus rapidement lors d'un raffinement du maillage.

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|-----------|
| 8 | 18.49039 | 18.33277 | 289.46149 |
| 16 | 41.13094 | 41.03600 | 307.54297 |
| 32 | 86.15281 | 86.05528 | 393.00625 |
| 64 | 176.04861 | 175.91650 | 543.69434 |
| 128 | 355.85118 | 355.80947 | 666.11750 |
| 256 | 715.43979 | 715.38487 | 854.48928 |

TABLEAU 3.9. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$

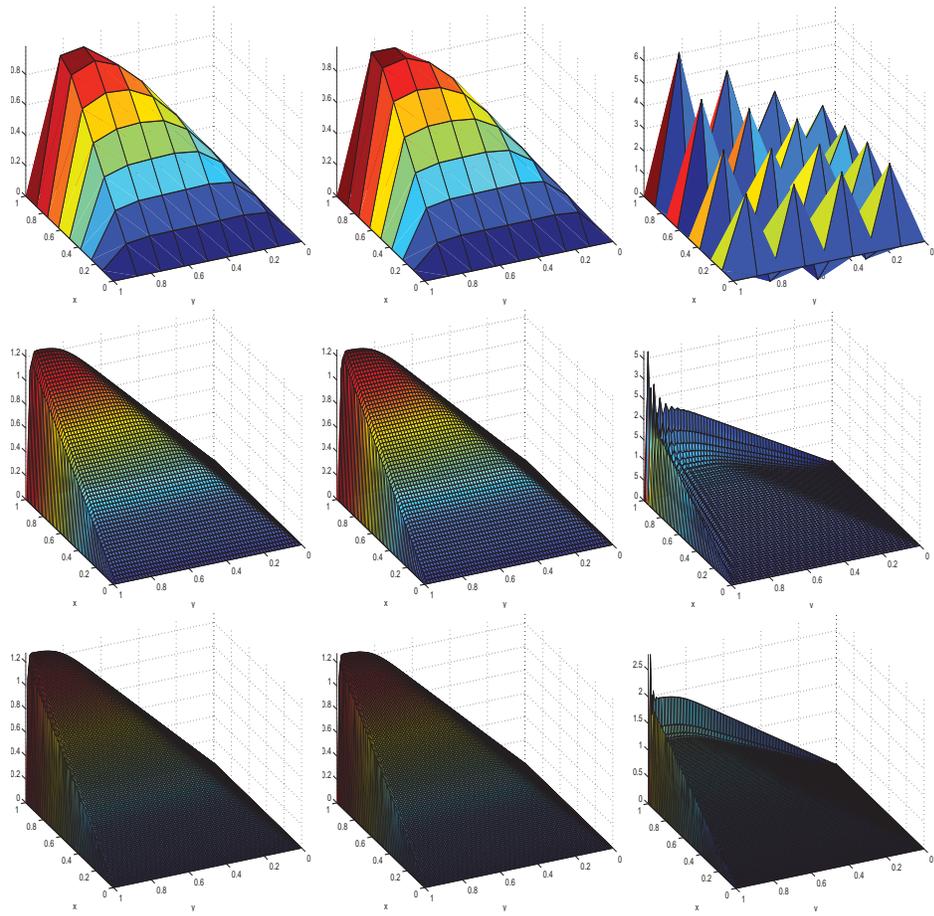


FIGURE 3.8. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 64, 128$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|------------|
| 8 | 18.59047 | 18.41837 | 2553.01935 |
| 16 | 41.35734 | 41.24074 | 2718.77471 |
| 32 | 86.66606 | 86.51552 | 3083.99174 |
| 64 | 176.93209 | 176.76763 | 3049.03086 |
| 128 | 357.24310 | 357.12447 | 4209.79106 |
| 256 | 718.01145 | 717.98454 | 5099.28595 |

TABLEAU 3.10. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$

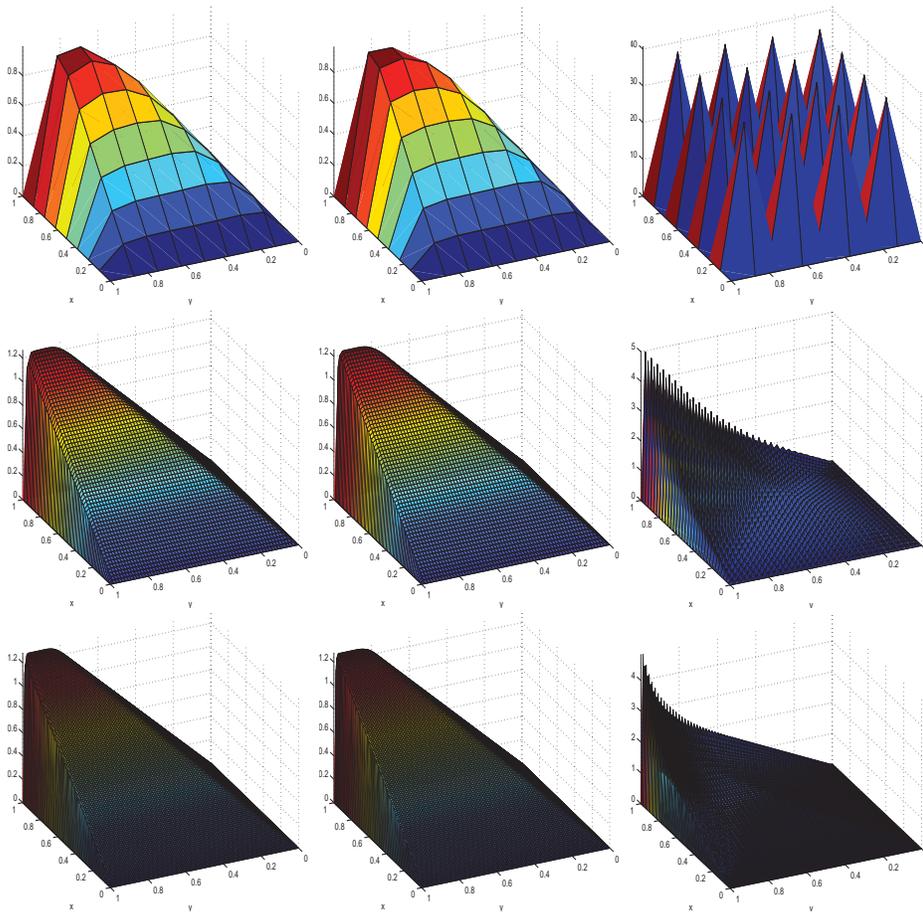


FIGURE 3.9. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 64, 128$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(230\pi/1024)$, $\beta_2 = \sin(230\pi/1024)$

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|-----------|
| 8 | 19.01295 | 16.96903 | 315.23977 |
| 16 | 40.82372 | 40.90179 | 441.42831 |
| 32 | 86.44886 | 86.69990 | 490.52162 |
| 64 | 178.37046 | 176.71391 | 565.71711 |
| 128 | 361.82266 | 360.60617 | 701.75634 |
| 256 | 729.42168 | 728.04106 | 956.42445 |

TABLEAU 3.11. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$

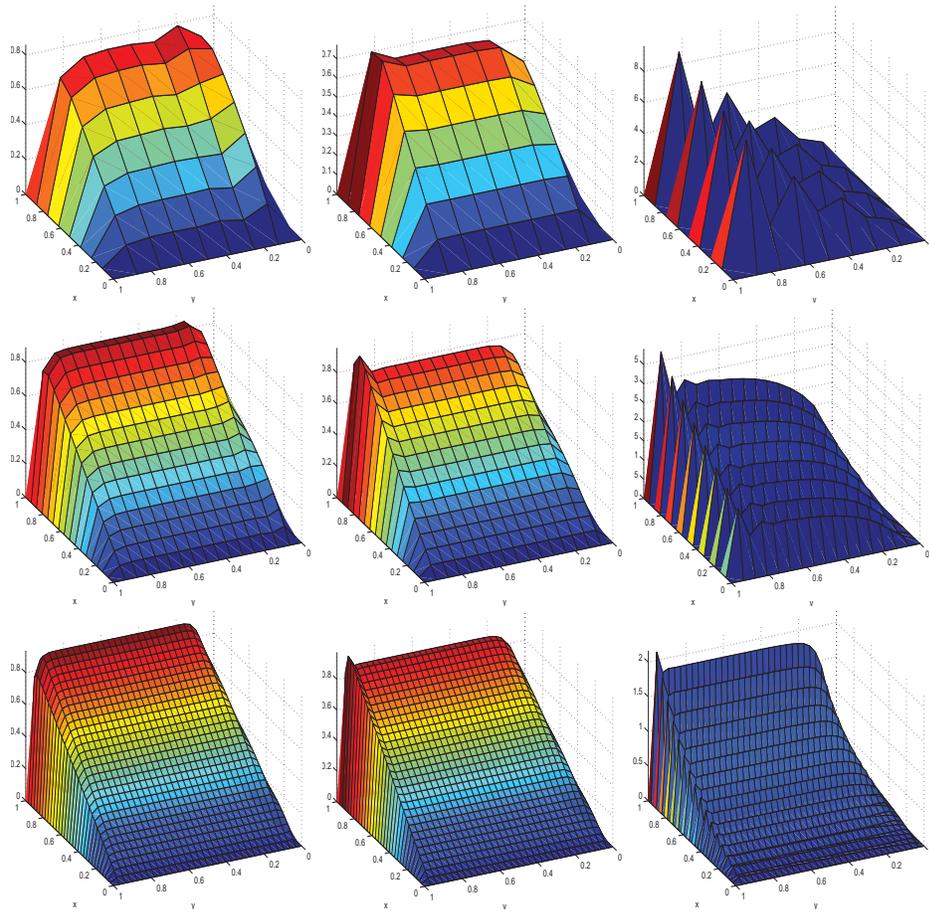


FIGURE 3.10. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|------------|
| 8 | 19.83705 | 17.14630 | 2510.51263 |
| 16 | 42.43226 | 42.48388 | 2654.43767 |
| 32 | 88.70430 | 98.05406 | 3638.71277 |
| 64 | 181.05003 | 218.01300 | 4628.46504 |
| 128 | 366.09888 | 435.78780 | 4935.26703 |
| 256 | 737.84584 | 795.96348 | 5265.77556 |

TABLEAU 3.12. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$

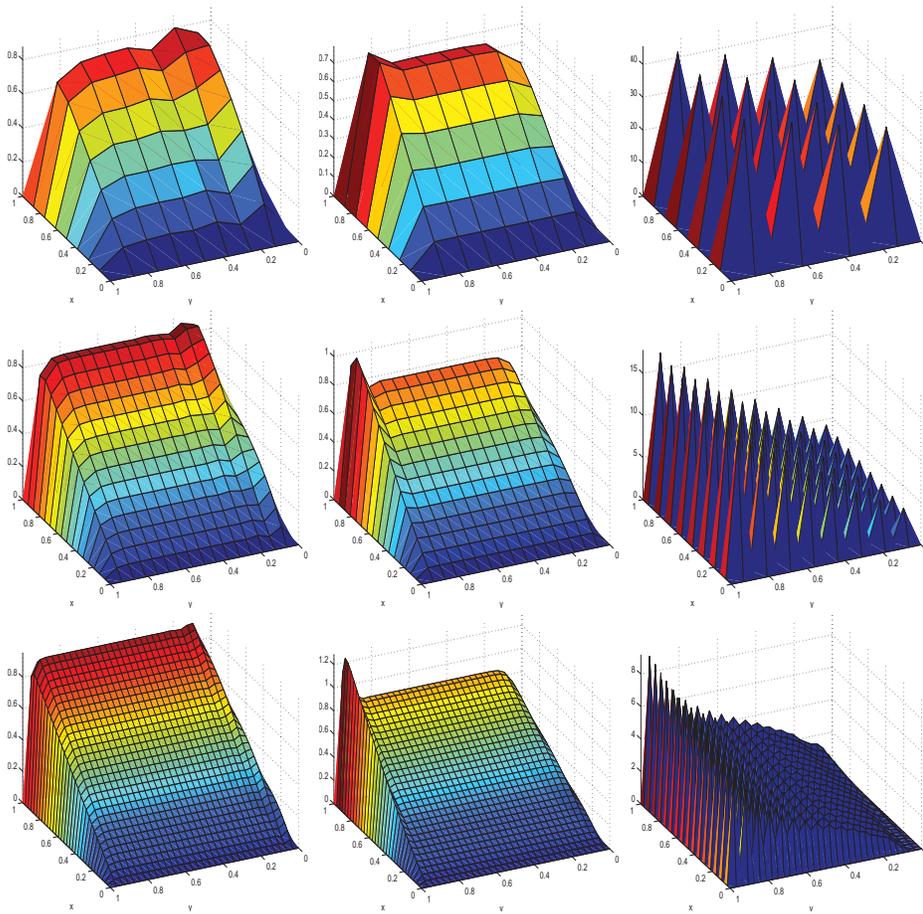


FIGURE 3.11. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(32\pi/1024)$, $\beta_2 = \sin(32\pi/1024)$

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|-----------|
| 8 | 18.06513 | 16.61171 | 453.07743 |
| 16 | 40.57547 | 39.61860 | 479.08769 |
| 32 | 87.23666 | 85.58855 | 500.27858 |
| 64 | 180.07063 | 178.10706 | 576.49729 |
| 128 | 365.07368 | 363.69618 | 712.78775 |
| 256 | 735.94302 | 734.44181 | 968.89768 |

TABLEAU 3.13. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$

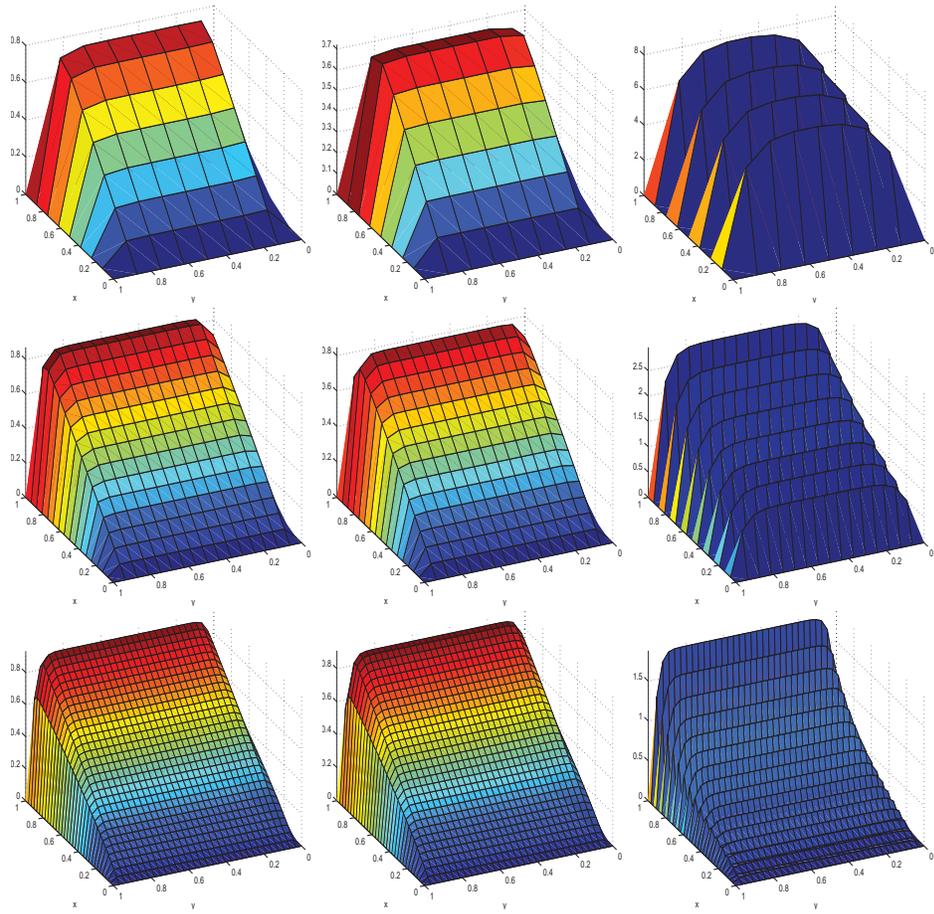


FIGURE 3.12. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$

| n | var_{DO} | var_F | var_C |
|-----|------------|-----------|------------|
| 8 | 20.75063 | 16.67218 | 4154.34995 |
| 16 | 43.45650 | 39.91693 | 4694.02532 |
| 32 | 89.39641 | 86.72268 | 4877.16497 |
| 64 | 183.59309 | 181.38724 | 4953.60733 |
| 128 | 373.33280 | 371.18596 | 5040.81982 |
| 256 | 752.83279 | 750.40721 | 5357.24163 |

TABLEAU 3.14. Comparaison des variations totales obtenues par les 3 méthodes pour différentes valeurs de $h = \frac{1}{n}$ avec $\epsilon = 0.0001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$

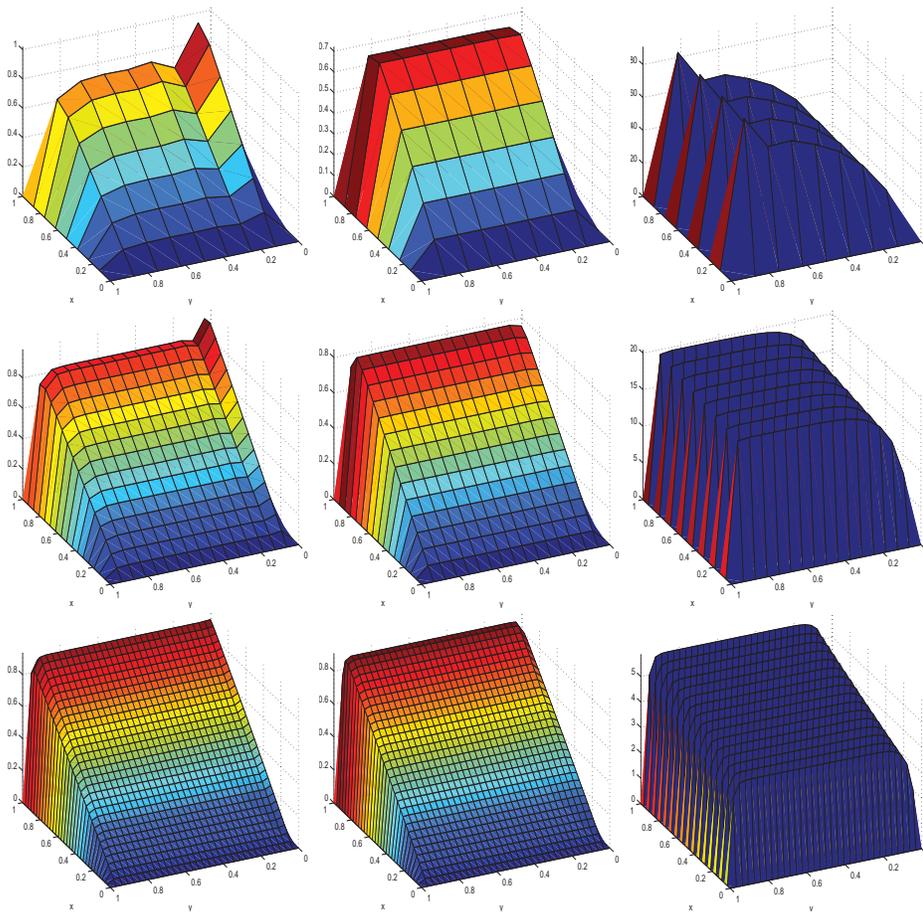


FIGURE 3.13. Solution discrète U obtenue par la méthode De l'Isle-Owens (gauche), Fatone et al. (milieu) et centrée (droite) pour les maillages $n = 8, 16, 32$ (de haut en bas) et avec $\epsilon = 0.0001$, $\beta_1 = \cos(1\pi/1024)$, $\beta_2 = \sin(1\pi/1024)$

Chapitre 4

APPLICATION À UN PROBLÈME DE NAVIER-STOKES

Au chapitre précédent, nous avons testé la discrétisation superconsistante sur un problème stationnaire en introduisant l'équation d'advection-diffusion. L'avantage de la méthode superconsistante se faisait sentir lorsque le terme diffusif était négligeable devant le terme advectif. On va maintenant tester la méthode sur un problème avec dépendance temporelle. On se servira des équations de Navier-Stokes qui décrivent la physique du comportement d'un fluide sous certaines hypothèses. Les équations résultantes auront une forme analogue à (3.1.13) sauf que le terme diffusif sera associé à la viscosité du fluide.

Dans la première section on explicite le développement menant à l'obtention des équations de Navier-Stokes pour un fluide newtonien incompressible. La théorie et les développements mathématiques sont tirés des ouvrages [4, 6, 19],

4.1. POSTULATS ET PRINCIPES GÉNÉRAUX DE MÉCANIQUE DES FLUIDES

4.1.1. Mécanique des milieux continus

Le fondement principal dans le développement des équations de Navier-Stokes repose sur le traitement du fluide comme un milieu continu, c'est à dire qu'on fait l'hypothèse que les fonctions qui décrivent le fluide (vitesse, pression, etc) sont des fonctions qui évoluent de façon continue dans l'espace.

C'est en réalité faux si on considère le fluide à l'échelle microscopique et qu'on étudie le mouvement des molécules. Dans des situations extrêmes de vitesse, de température ou de pression, un fluide aussi commun que l'air peut présenter des comportements qui ne sont pas décrit par la mécanique des milieux continus.

Cependant, dans la grande majorité des cas, l'hypothèse est valide car la taille des molécules et la distance entre elles sont suffisamment petites par rapport à la taille du système étudié.

4.1.2. Viscosité et fluide newtonien

Dans l'étude des milieux continus on distingue 2 catégories d'objet : les solides et les fluides. La distinction entre les deux réside dans la réaction de l'objet à une contrainte de cisaillement, c'est à dire une force appliquée de manière parallèle à la surface de l'objet. Si on applique une telle force sur un solide pour le déformer, sa déformation sera proportionnelle à la grandeur de la contrainte appliquée mais sera indépendante du temps d'application de la force. En comparaison, le fluide se déforme de manière continue tant que la contrainte est appliquée.

Le taux de déformation du fluide sous l'effet d'une contrainte de cisaillement dépend des forces de friction interne : c'est la propriété de viscosité. Lorsque le taux de déformation est linéairement proportionnel à la grandeur de la contrainte de cisaillement on dira que le fluide est newtonien.

Les équations de Navier-Stokes décrivent le comportement des écoulements pour les fluides visqueux newtoniens compressibles et incompressibles. Dans ce travail on s'intéresse seulement à la version incompressible des équations.

4.1.3. Fluide incompressible

Un fluide est incompressible si la densité ρ de chaque particule de fluide est constante. Pour être tout à fait rigoureux, tous les fluides sont compressibles, car ρ est toujours sensible à des variations de pression. Savoir si un fluide est fondamentalement incompressible est cependant sans intérêt. La question est plutôt de déterminer si les fluctuations de pression auxquelles est soumis le fluide dans le système n'entraînent qu'une variation négligeable de densité auquel cas on peut assumer que ρ est constant.

L'air par exemple peut être vu comme un fluide incompressible si la vitesse de l'écoulement est basse. Cependant, l'écoulement d'air dans une turbine ou sous l'aile d'un jet approche ou dépasse la vitesse du son et les changements de densité ne peuvent alors plus être négligés. La dynamique des fluides incompressible est donc l'étude des mouvements des fluides dans des situations où ils sont suffisamment bien décrits par la version incompressible des équations.

4.2. DÉRIVATION DES ÉQUATIONS DE NAVIER-STOKES

4.2.1. Représentation lagrangienne

Si on veut étudier un écoulement, deux approches s'offrent à nous. On peut observer l'évolution d'un point fixe dans l'écoulement (représentation eulérienne) ou on peut suivre une particule du fluide (représentation lagrangienne). Pour calculer l'évolution $\frac{d}{dt}$ d'une quantité Z d'un élément de fluide, il est souvent plus naturel d'utiliser la deuxième approche. Pour prendre en considération le déplacement de l'élément, on pose Z comme une fonction du temps t et de la position $\vec{r}(t)$. En coordonnées cartésiennes et en utilisant la règle de dérivation en chaîne on obtient

$$\begin{aligned} \frac{d}{dt} Z(\vec{r}(t), t) &= \frac{d}{dt} Z(x(t), y(t), z(t), t) \\ &= \frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial x} \frac{dx}{dt} + \frac{\partial Z}{\partial y} \frac{dy}{dt} + \frac{\partial Z}{\partial z} \frac{dz}{dt} \\ &= \frac{\partial Z}{\partial t} + \vec{u} \cdot \nabla Z \end{aligned} \quad (4.2.1)$$

Cette expression se généralise pour un système de coordonnées arbitraire en écrivant

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\vec{u} \cdot \nabla) \quad (4.2.2)$$

L'opérateur (4.2.2) est appelé la dérivée lagrangienne.

4.2.2. Conservation de la masse

On a vu au début du chapitre 3 comment obtenir l'équation de continuité (3.1.5). En utilisant la propriété suivante $\nabla \cdot (\rho \vec{u}) = \rho \nabla \cdot \vec{u} + \vec{u} \cdot \nabla \rho$ on peut réécrire l'équation sous la forme

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \vec{u} = 0 \quad (4.2.3)$$

Pour un fluide incompressible, $\frac{D\rho}{Dt}$ est constant et l'équation devient simplement

$$\nabla \cdot \vec{u} = 0 \quad (4.2.4)$$

4.2.3. Conservation de la quantité de mouvement

La façon la plus rigoureuse de dériver l'équation de la quantité de mouvement nécessite des éléments de calcul tensoriel que nous n'introduisons pas dans ce travail. Une autre approche est de simplement adapter la deuxième loi de Newton

$\vec{F} = m\vec{a}$ en posant $\vec{F} = \vec{s}$, $m = \rho$ et $\vec{a} = \frac{d\vec{u}}{dt}$. On reprend le développement (4.2.1) en posant $Z = \vec{u}$. et on trouve

$$\rho \frac{D\vec{u}}{Dt} = \vec{s} \quad (4.2.5)$$

4.2.4. Équations de Navier-Stokes pour un fluide newtonien incompressible

Les développements mathématiques et physiques de cette section utilisent beaucoup de raccourcis pour ne pas alourdir le travail. Les forces \vec{s} de l'équation (4.2.5) peuvent être de deux types : des forces agissant sur le volume (par exemple la gravité) et des forces agissant sur la surface du volume (contraintes). On néglige les forces du premier type et on décompose les forces de contraintes en 2 termes : le gradient de pression et la divergence du tenseur des contraintes visqueuses $\vec{\tau}$.

$$\frac{D\vec{u}}{Dt} = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}\nabla \cdot \vec{\tau} \quad (4.2.6)$$

Pour un fluide newtonien, on a émis l'hypothèse que les composantes de $\vec{\tau}$ sont des fonctions linéaires des composantes du gradient de vitesse et on peut montrer que

$$\vec{\tau} = \mu(\nabla\vec{u} + (\nabla\vec{u})^T) + \mu'(\nabla \cdot \vec{u})$$

où μ est le coefficient de viscosité dynamique et μ' , le coefficient de viscosité volumique. Pour un fluide incompressible, $\nabla \cdot \vec{u} = 0$ et on a

$$\nabla \cdot \vec{\tau} = \mu\nabla \cdot (\nabla\vec{u} + (\nabla\vec{u})^T) = \mu(\nabla \cdot \nabla\vec{u} + \nabla(\nabla \cdot \vec{u})) = \mu\nabla^2\vec{u}$$

L'équation finale est donc

$$\frac{D\vec{u}}{Dt} = -\frac{1}{\rho}\nabla p + \frac{\mu}{\rho}\nabla^2\vec{u} \quad (4.2.7)$$

C'est une équation vectorielle à trois composantes. En ajoutant l'équation d'incompressibilité (4.2.4) on obtient les équations de Navier-Stokes pour un fluide newtonien incompressible.

Finalement, pour un fluide visqueux, on doit spécifier la condition limite de non-glissement qui suit l'hypothèse que la vitesse tangentielle du fluide sur une paroi solide est nulle. Pour une frontière impénétrable, la vitesse perpendiculaire est aussi nécessairement nulle. On pose donc :

$$\vec{u} = \vec{0} \text{ sur une frontière impénétrable et immobile}$$

4.2.5. Nombre de Reynolds

Le nombre de Reynolds est un nombre sans dimension qui représente le rapport entre les forces d'inertie et les forces visqueuses. Si u_0 et L sont respectivement les échelles caractéristiques de vitesse et de taille du domaine, on a

$$Re = \frac{u_0 L \rho}{\mu}$$

Dans le problème de cavité entraînée que nous résoudrons, on s'intéresse aux équations adimensionnées c'est à dire qu'on posera $u_0 = L = \rho = 1$ de sorte que le nombre de Reynolds dépend uniquement de la viscosité μ . Nous réécrivons (4.2.7) de la façon suivante :

$$\frac{D\vec{u}}{Dt} = -\nabla p + \epsilon \nabla^2 \vec{u} \quad (4.2.8)$$

où $\epsilon := 1/Re$ sera un paramètre très petit puisqu'on s'intéresse aux écoulements à grand nombre de Reynolds.

En hydrodynamique, lorsqu'on observe l'écoulement d'un fluide près d'une surface immobile, la vitesse du fluide est nulle sur la frontière et tend vers la vitesse globale de l'écoulement plus on s'éloigne de la paroi. La relation entre la distance de la paroi et la vitesse du fluide dépend de la viscosité de celui-ci. Cette viscosité induit un frottement entre les couches voisines de fluide et plus le fluide est visqueux, plus le changement de vitesse est progressif. À l'inverse, si la viscosité est faible, la région où elle a une influence sur la vitesse sera très mince proche de la paroi. On appelle cette région la couche limite.

Mathématiquement, plus le nombre de Reynold est grand, plus la couche limite est mince et plus la résolution numérique est difficile.

4.2.6. Problème de cavité entraînée

Pour tester notre méthode de discrétisation superconsistante on résout le problème de cavité entraînée "*lid driven cavity*". C'est un problème de référence pour tester des méthodes numériques car une panoplie de résultats et d'analyses existent dans la littérature. Le problème en 2 dimensions s'énonce ainsi : On considère un fluide dans un domaine rectangulaire $\Omega = (0, l_x) \times (0, l_y)$ (voir Figure 4.1) et on pose sa vitesse $\vec{u}(x, y, t) := (u(x, y, t), v(x, y, t))$ comme une fonction vectorielle à dépendance spatiale et temporelle. On définit les conditions frontières telles que $v = 0$ sur $\partial\Omega$, $u(x, 0, t) = u(0, y, t) = u(l_x, y, t) = 0$ et $u(x, l_y, t) = u_0$. Le fluide est initialement au repos $\vec{u}(x, y, 0) = 0, \forall (x, y) \in \Omega$ et on étudie l'évolution de \vec{u} dans le temps.

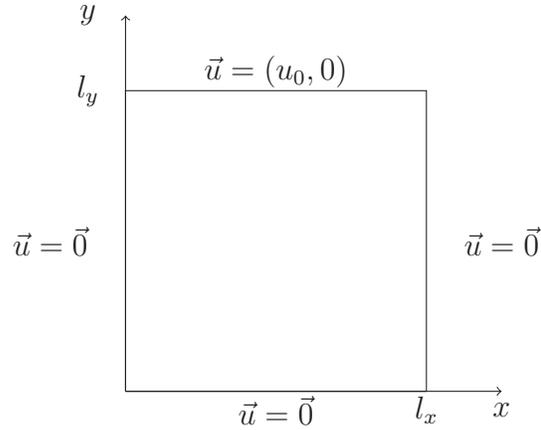


FIGURE 4.1. Domaine et conditions frontières pour le problème de cavité entraînée.

Pour différentes valeurs de viscosité, la condition en $y = l_y$ entraîne la création d'un ou plusieurs tourbillons (vortex) dans le domaine. Soit $T > 0$ suffisamment grand, il est possible que ces tourbillons se stabilisent lorsque $t > T$ et on a alors un régime permanent $\frac{\partial}{\partial t} \vec{u}(x, y, t) = 0$, $t > T$. Il est possible aussi de trouver une solution périodique de période L telle que $\vec{u}(x, y, t) = \vec{u}(x, y, t + L)$, $\forall (x, y) \in \Omega$, $t > T$. Une solution totalement chaotique est aussi une possibilité. Dans la littérature, on s'intéresse au lien entre la grandeur de la viscosité du fluide et l'existence de solutions en régime permanent ou périodique.

Notre objectif n'est pas de résoudre avec précision le problème de cavité entraînée mais plutôt de comparer notre méthode superconsistante à la méthode centrée et celle de Fatone et al. sur une base uniforme. Dans cette optique, les paramètres que nous choisirons sont ceux qui sont les plus couramment utilisés soit

$$l_x = l_y = 1 \text{ et } u_0 = 1$$

Aussi, nous limiterons la grandeur du nombre de Reynolds à des valeurs telles qu'on peut assumer qu'une solution en régime permanent existe. Basé sur les résultats des différentes publications, nous choisirons $Re \leq 5000$.

4.3. ÉQUATIONS DU PROBLÈME

On considère un fluide incompressible dans un rectangle $[0, 1] \times [0, 1]$ (Figure 4.1). Soit \hat{e}_x et \hat{e}_y des vecteurs formant une base orthonormée de \mathbb{R}^2 , on note la vitesse $\vec{u} = u\hat{e}_x + v\hat{e}_y$ et la pression p . Soit Re le nombre de Reynolds, en posant $\epsilon = 1/Re$, les équations à résoudre sont les équations de Navier-Stokes adimensionnées (4.2.4) et (4.2.8). En 2 dimensions elles s'écrivent :

$$u_t - \epsilon(u_{xx} + u_{yy}) + uu_x + vv_y + p_x = 0 \quad (4.3.1)$$

$$v_t - \epsilon(v_{xx} + v_{yy}) + uv_x + vv_y + p_y = 0 \quad (4.3.2)$$

$$u_x + v_y = 0 \quad (4.3.3)$$

Avec les conditions frontières

$$u(0, y) = 0 \quad u(1, y) = 0 \quad u(x, 0) = 0 \quad u(x, 1) = 1$$

$$v(0, y) = 0 \quad v(1, y) = 0 \quad v(x, 0) = 0 \quad v(x, 1) = 0$$

En notation plus compacte les équations (4.3.1) (4.3.2) et (4.3.3) se réécrivent :

$$\vec{u}_t - \epsilon \nabla^2 \vec{u} + \vec{u} \cdot \nabla \vec{u} + \nabla p = \vec{0} \quad (4.3.4)$$

$$\nabla \cdot \vec{u} = 0 \quad (4.3.5)$$

et l'équation en régime permanent associée à (4.3.4) est

$$-\epsilon \nabla^2 \vec{u} + \vec{u} \cdot \nabla \vec{u} + \nabla p = \vec{0} \quad (4.3.6)$$

On s'intéresse aux solutions de (4.3.6) mais plutôt que de résoudre directement l'équation, on utilise (4.3.4) et on calcule la solution à différents pas de temps jusqu'à ce qu'elle n'évolue plus selon un certain critère de tolérance. À ce moment on considère que le régime permanent est atteint et que la solution satisfait (4.3.6).

4.3.1. Méthodes de projection

Pour construire le système d'équations discrétisé correspondant à (4.3.1)-(4.3.3) on considère une solution discrète composée des approximations de la vitesse et de la pression au pas de temps k soit respectivement : $\vec{U}^k = (U^k, V^k)$ et P^k . On aura aussi une condition initiale (U^0, V^0) . Il ne sera pas nécessaire de choisir une condition initiale pour la pression.

Pour contourner la difficulté que présente le couplage des inconnues u et p , une approche typique est d'utiliser une méthode à pas fractionnaires qui consiste à décomposer les équations à résoudre en une ou plusieurs étapes intermédiaires. On utilisera une approche à 2 pas, donc une seule étape intermédiaire. Dans le premier pas on résout l'équation de la quantité de mouvement (4.3.1) et (4.3.2) en négligeant le terme de pression. On obtient une solution provisoire que l'on corrige au deuxième pas en introduisant l'équation d'incompressibilité.

Ce type de méthode est appelé méthode de projection et notre décomposition suit l'idée présenté dans [17]. Les deux étapes sont respectivement

$$\frac{U^* - U^k}{\Delta t} - \theta\epsilon(U_{xx}^* + U_{yy}^*) - (1 - \theta)\epsilon(U_{xx}^k + U_{yy}^k) + U^k U_x^* + V^k U_y^* = 0 \quad (4.3.7)$$

$$\frac{V^* - V^k}{\Delta t} - \theta\epsilon(V_{xx}^* + V_{yy}^*) - (1 - \theta)\epsilon(V_{xx}^k + V_{yy}^k) + U^k V_x^* + V^k V_y^* = 0 \quad (4.3.8)$$

puis

$$\frac{U^{k+1} - U^*}{\Delta t} - \theta\epsilon(U_{xx}^{k+1} + U_{yy}^{k+1}) + \theta\epsilon(U_{xx}^* + U_{yy}^*) = -(P^{k+1})_x \quad (4.3.9)$$

$$\frac{V^{k+1} - V^*}{\Delta t} - \theta\epsilon(V_{xx}^{k+1} + V_{yy}^{k+1}) + \theta\epsilon(V_{xx}^* + V_{yy}^*) = -(P^{k+1})_y \quad (4.3.10)$$

En additionnant (4.3.7) et (4.3.9) puis (4.3.8) et (4.3.10) on retrouve une approximation de (4.3.1) et (4.3.2) où le terme diffusif est une combinaison de $\Delta \vec{U}^{k+1}$ et $\Delta \vec{U}^k$.

$$\frac{U^{k+1} - U^k}{\Delta t} - (1 - \theta)\epsilon(U_{xx}^k + U_{yy}^k) - \theta\epsilon(U_{xx}^{k+1} + U_{yy}^{k+1}) + U^k U_x^* + V^k U_y^* + (P^{k+1})_x = 0 \quad (4.3.11)$$

$$\frac{V^{k+1} - V^k}{\Delta t} - (1 - \theta)\epsilon(V_{xx}^k + V_{yy}^k) - \theta\epsilon(V_{xx}^{k+1} + V_{yy}^{k+1}) + U^k V_x^* + V^k V_y^* + (P^{k+1})_y = 0 \quad (4.3.12)$$

Puisqu'on connaît la valeur de (U^k, V^k) , le premier système (4.3.7)-(4.3.8) est linéaire et on pourra résoudre l'équation matricielle associée pour trouver (U^*, V^*) . Le deuxième système (4.3.9)-(4.3.10) a 3 inconnues U^{k+1}, V^{k+1} et P^{k+1} . En notation vectorielle il s'écrit

$$\frac{\vec{U}^{k+1} - \vec{U}^*}{\Delta t} - \theta\epsilon(\Delta \vec{U}^{k+1} - \Delta \vec{U}^*) = -\nabla P^{k+1} \quad (4.3.13)$$

En appliquant la divergence de chaque côté et en exigeant que $\nabla \cdot \vec{U}^{k+1} = 0$ on obtient une équation de Poisson que l'on peut résoudre pour trouver P^{k+1}

$$\left(\frac{1}{\Delta t} - \theta\epsilon\Delta\right) (\nabla \cdot \vec{U}^*) = \Delta P^{k+1} \quad (4.3.14)$$

Ensuite on calcule le champ ∇P^{k+1} . L'équation (4.3.13) se réécrit

$$\left(\frac{1}{\Delta t} - \theta\epsilon\nabla^2\right) (\vec{U}^{k+1} - U^*) = -\nabla P^{k+1} \quad (4.3.15)$$

ce qui implique

$$\vec{U}^* = \vec{U}^{k+1} + \left(\frac{1}{\Delta t} - \theta\epsilon\nabla^2\right)^{-1} \nabla P^{k+1} \quad (4.3.16)$$

en supposant que l'opérateur $\left(\frac{1}{\Delta t} - \theta\epsilon\nabla^2\right)$ est inversible, on peut utiliser cette dernière expression pour exprimer \vec{U}^* en fonction de \vec{U}^{k+1} dans (4.3.11) et (4.3.12). Après substitution on obtient :

$$\begin{aligned} \frac{\vec{U}^{k+1} - \vec{U}^k}{\Delta t} - (1 - \theta)\epsilon\nabla^2\vec{U}^k - \theta\epsilon\nabla^2\vec{U}^{k+1} + \vec{U}^k \cdot \nabla\vec{U}^{k+1} + \nabla P^{k+1} \\ = -\vec{U}^{k+1} \cdot \nabla \left[\left(\frac{1}{\Delta t} - \theta\epsilon\nabla^2\right)^{-1} \nabla P^{k+1} \right] \end{aligned} \quad (4.3.17)$$

Lorsqu'on atteint un régime permanent, $\vec{U}^k = \vec{U}^{k+1}$, et la solution trouvée satisfait l'équation suivante

$$-\epsilon\nabla^2\vec{U}^{k+1} + \vec{U}^{k+1} \cdot \nabla\vec{U}^{k+1} + \nabla P^{k+1} = -\vec{U}^{k+1} \cdot \nabla \left[\left(\frac{1}{\Delta t} - \theta\epsilon\nabla^2\right)^{-1} \nabla P^{k+1} \right] \quad (4.3.18)$$

C'est une discrétisation de l'équation (4.3.6) mais avec un terme d'erreur au membre de droite. En pratique on choisira Δt de l'ordre de ϵ donc le terme $\theta\epsilon\nabla^2$ est négligeable devant $\frac{1}{\Delta t}$. Ainsi le terme d'erreur est de l'ordre de Δt .

4.3.1.1. Choix de θ

En prenant $\theta = 0$, le deuxième système (4.3.9)-(4.3.10) est considérablement simplifié mais on perd le terme diffusif dans le premier système (4.3.7)-(4.3.8). Comme notre objectif est de tester l'efficacité de la méthode superconsistante lorsque le terme diffusif est négligeable devant le terme advectif ce choix n'est pas adéquat.

Puisque la discrétisation de la dérivée temporelle est consistante d'ordre 2 au temps $k + \frac{1}{2}$, le choix raisonnable est de prendre $\theta = 1/2$ de sorte que le terme diffusif soit une interpolation centrée entre $\nabla^2\vec{U}^k$ et $\nabla^2\vec{U}^{k+1}$. Les deux systèmes présentent cependant des termes supplémentaires au membre de droite ce qui complique la résolution.

Pour les besoins du test, nous allons choisir $\theta = 1$ dans le premier système et $\theta = 0$ dans le deuxième. Avec ce choix, l'équation contenant le terme d'erreur une fois le régime permanent atteint (4.3.18) n'est plus la même et on peut vérifier qu'on obtient plutôt

$$-\epsilon\nabla^2\vec{U}^{k+1} + \vec{U}^{k+1} \cdot \nabla\vec{U}^{k+1} + \nabla P^{k+1} = \Delta t \left(\epsilon\nabla^2 - \vec{U}^{k+1} \cdot \nabla \right) \nabla P^{k+1} \quad (4.3.19)$$

Par rapport à (4.3.18), on a un nouveau terme de l'ordre de $\epsilon\Delta t$ donc négligeable. Par rapport au choix $\theta = 1/2$, le choix $\theta = 1$ dans le premier système et $\theta = 0$ dans le deuxième fait donc bénéficier d'une simplicité d'implémentation sans grande perte en précision. Une idée présentée dans [5] pour améliorer la précision de la solution finale est de rajouter les composantes \vec{i} et \vec{j} du terme $-\Delta t (\epsilon\nabla^2 - \vec{U}^k \cdot \nabla) \nabla P^k$ aux membres de droite de (4.3.7) et (4.3.8) respectivement. Ainsi lorsque le régime permanent est atteint et que l'on peut considérer que $P^k = P^{k+1}$ et $\vec{U}^k = \vec{U}^{k+1}$, le terme que l'on a ajouté va éliminer le membre de droite de (4.3.19). Cette idée est cependant apparue tard dans notre recherche et nous ne l'avons pas mise en œuvre pour ne pas avoir à recommencer les simulations qui sont couteuses en temps. De plus, on rappelle que l'objectif est de comparer la méthode de De l'Isle Owens avec celle de Fatone et al. et la discrétisation centrée. Or le terme d'erreur est le même pour les trois méthodes donc cette nouvelle approche n'aurait probablement pas affecté nos conclusions.

En résumé, à chaque pas de temps, il faut résoudre trois systèmes : calcul de (U^*, V^*) , calcul de P^{k+1} et calcul de (U^{k+1}, V^{k+1}) et les équations associées sont dans l'ordre

$$\left(1 - \epsilon\Delta t \frac{\partial^2}{\partial x^2} - \epsilon\Delta t \frac{\partial^2}{\partial y^2} + \Delta t U^k \frac{\partial}{\partial x} + \Delta t V^k \frac{\partial}{\partial y}\right) U^* = U^k \quad (4.3.20)$$

$$\left(1 - \epsilon\Delta t \frac{\partial^2}{\partial x^2} - \epsilon\Delta t \frac{\partial^2}{\partial y^2} + \Delta t U^k \frac{\partial}{\partial x} + \Delta t V^k \frac{\partial}{\partial y}\right) V^* = V^k \quad (4.3.21)$$

$$\nabla^2 P^{k+1} = \frac{1}{\Delta t} (U_x^* + V_y^*) \quad (4.3.22)$$

$$U^{k+1} = U^* - \Delta t P_x^{k+1} \quad (4.3.23)$$

$$V^{k+1} = V^* - \Delta t P_y^{k+1} \quad (4.3.24)$$

Les conditions frontières pour U et V sont celles décrites sur la figure 4.1 avec $u_0 = 1$. Pour les fluides visqueux, la condition frontière pour la pression P est plus délicate à spécifier et dépend du problème étudié. Pour le problème de cavité entraînée, et en général pour les problèmes d'écoulement présentant la condition de non-glissement aux frontières, il est d'usage pour des raisons de stabilité [1, 8] de prescrire la condition Neumann homogène pour la pression dans la résolution de (4.3.22). C'est aussi le choix fait dans l'article [18] qui a inspiré une grande partie de notre code `MATLAB`.

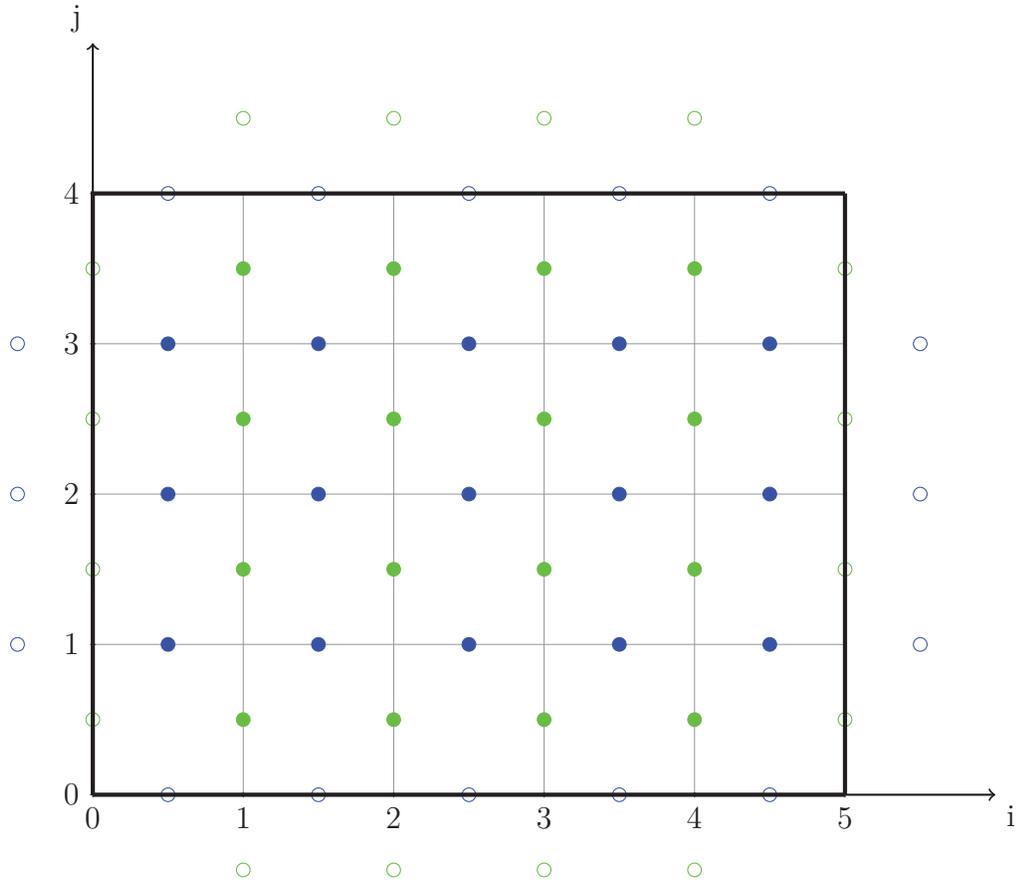


FIGURE 4.2. Exemple de grille de représentation uniforme décalée. Les points verts et bleus constituent la grille de U^k et de V^k respectivement.

4.4. DISCRÉTISATION DES ÉQUATIONS

On utilise une grille uniforme décalée "staggered grid" (voir Figure 4.2). La vitesse horizontale U^k est calculée aux points $(i, j + \frac{1}{2})$ et la vitesse verticale V^k est calculée aux points $(i + \frac{1}{2}, j)$. La pression est calculée au centre de la cellule aux points $(i + \frac{1}{2}, j + \frac{1}{2})$. Ce choix permet d'utiliser des discrétisations centrées, simples à implémenter pour les équations (4.3.22), (4.3.23) et (4.3.24) tout en gardant une consistance d'ordre 2.

On réfère à [18] pour les détails de la discrétisation de (4.3.22) à (4.3.24). En particulier, la gestion des conditions frontières doit être faite avec soin car certains points frontières de la grille décalée sont à l'extérieur du domaine. Aussi, la matrice associée à la résolution de la pression dans l'équation (4.3.22) est singulière car la condition Neumann homogène ne décrit le champ de pression qu'à une constante près. Une infinité de solutions satisfont donc l'équation. Comme on ne s'intéresse qu'au gradient de pression, cette constante n'influe pas sur le champ de vitesse.

Cependant pour la résolution numérique il est important de travailler avec des matrices régulières. On suit donc l'idée de [18] pour rendre la matrice régulière en s'assurant que la solution engendrée reste correcte.

L'influence de notre travail se situe dans la résolution des équations (4.3.20) et (4.3.21) où la méthode superconsistante aura un impact dû à la présence d'un terme advectif et d'un terme diffusif. Pour garder la nomenclature des chapitres précédents, on condense la notation en posant

$$L = 1 - \epsilon \Delta t \frac{\partial^2}{\partial x^2} - \epsilon \Delta t \frac{\partial^2}{\partial y^2} + \Delta t \beta_1 \frac{\partial}{\partial x} + \Delta t \beta_2 \frac{\partial}{\partial y} \quad (4.4.1)$$

avec β_1 et β_2 les valeurs interpolées de U^k et V^k respectivement aux points de collocation de l'équation. Le système se réécrit

$$LU^* = U^k$$

$$LV^* = V^k$$

Chaque équation utilise ses propres grilles de représentation et de collocation $(\vec{\tau}_u, \vec{\mu}_u)$ et $(\vec{\tau}_v, \vec{\mu}_v)$. Le développement de l'opérateur discrétisé suit la même démarche qu'au chapitre 3. L'opérateur est cependant sensiblement différent et il en va de même pour le calcul des points de collocation. Après avoir posé $\chi(x, y)$ comme dans (3.3.2) et en posant $X = x_i - x$ et $Y = y_j - y$ on trouve les équations suivantes à résoudre pour la méthode De l'Isle-Owens

$$X^3 - 3\beta_1 \Delta t X^2 - X (6\Delta t \epsilon + h_x^2) + \beta_1 \Delta t h_x^2 = 0 \quad (4.4.2)$$

$$Y^3 - 3\beta_2 \Delta t Y^2 - Y (6\Delta t \epsilon + h_y^2) + \beta_2 \Delta t h_y^2 = 0 \quad (4.4.3)$$

Avec χ défini par (3.5.1) et en posant $X = \alpha \beta_1$ et $Y = \alpha \beta_2$, les points de collocation proposés par Fatone et al. se calculent en trouvant α la plus petite racine strictement positive de

$$\begin{aligned} & \alpha^5 \beta_1^2 \beta_2^2 - 6\alpha^4 \beta_1^2 \beta_2^2 \Delta t + \alpha^3 \left(-6\Delta t \epsilon (\beta_1^2 + \beta_2^2) - \beta_1^2 h_y^2 - \beta_2^2 h_x^2 \right) \\ & + 4\alpha^2 \Delta t (\beta_1^2 h_y^2 + \beta_2^2 h_x^2) + \alpha (6\Delta t \epsilon (h_x^2 + h_y^2) + h_x^2 h_y^2) \\ & - 2\Delta t h_x^2 h_y^2 = 0 \end{aligned} \quad (4.4.4)$$

On note que le calcul des couples (X, Y) nécessite la connaissance de X et Y puisque β_1 et β_2 sont les valeurs de U^k et V^k interpolées sur leur grille de collocation. On devra donc choisir une valeur initiale de β_1 et β_2 , s'en servir pour calculer X et Y , puis recalculer β_1 et β_2 en interpolant les vitesses au nouveau point de collocation et ainsi de suite jusqu'à ce que les valeurs de X et Y aient convergé selon un certain critère de tolérance. Pour accélérer les calculs

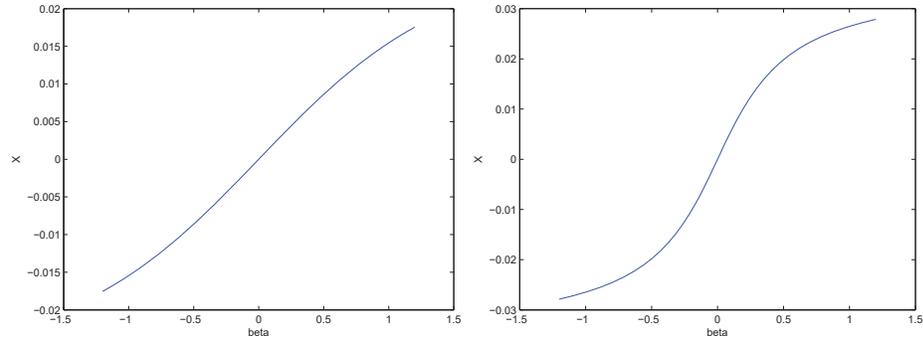


FIGURE 4.3. Solution minimale X de l'équation (4.4.2) en fonction de β_1 pour $Re = 100$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite)

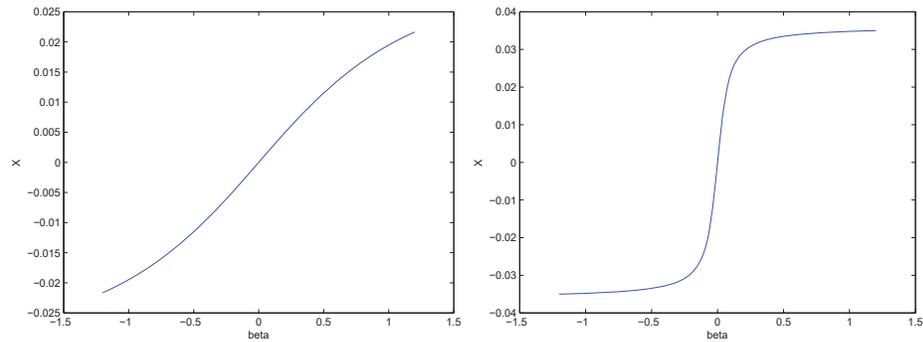


FIGURE 4.4. Solution minimale X de l'équation (4.4.2) en fonction de β_1 pour $Re = 5000$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite)

des racines de (4.4.2)-(4.4.3), comme les solutions dépendent de façon continue à β_1 et β_2 respectivement et que les autres paramètres (ϵ , h_x, h_y , Δt) sont constants au cours de la simulation, on peut initialiser un tableau de valeurs de X en fonction de β_1 et de Y en fonction de β_2 et utiliser de simples interpolations pendant la simulation. Puisque que nous fixons $u = 1$ sur la frontière nord, on peut assumer que les valeurs de β_1 et β_2 seront toujours dans l'intervalle $[-1, 1]$.

Le principe est le même pour le calcul des solutions minimales de (4.4.4) sauf qu'on aura une relation à 2 variables indépendantes et une interpolation en 2 dimensions à faire. Les Figures 4.3 à 4.5 montrent des exemples de ces relations pour certains choix de paramètres. Pour calculer les interpolations, nous utiliserons une méthode par spline implémentée dans MATLAB.

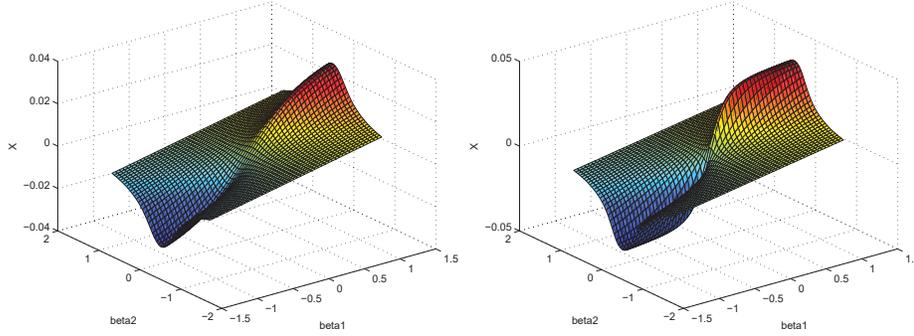


FIGURE 4.5. Valeur de $X = \alpha\beta_1$ tel que α est la solution minimale de l'équation (4.4.4) et en fonction de β_1 et β_2 . Les autres paramètres sont fixés : $Re = 5000$, $n_x = 16$ et $\Delta t = 0.025$ (gauche), $\Delta t = 0.4$ (droite)

4.5. ANALYSE DES RÉSULTATS NUMÉRIQUES

4.5.1. Présentation des résultats

Des simulations ont été effectuées avec différents nombre de Reynolds, différents choix de Δt et différentes résolutions $n_x = n_y = n$. Plus précisément : $Re \in \{100, 400, 1000, 3200, 5000\}$, $\Delta t \in \{0.001, 0.00625, 0.0125, 0.025, 0.05, 0.1, 0.2, 0.4\}$, et $n \in \{16, 32, 64, 128, 256\}$.

Nous avons testé la méthode de De l'Isle Owens, l'alternative de Fatone et al. présentée au chapitre 3 et aussi une discrétisation centrée ($X = Y = 0$). Le critère d'arrêt pour décider que le régime permanent est atteint est

$$\frac{1}{dt} \left(\max(|U^{k+1} - U^k|) + \max(|V^{k+1} - V^k|) \right) < 10^{-6}$$

Les données numériques obtenues des simulations sont les vitesses U^k et V^k en différents points de la cavité. Pour analyser la justesse de ces résultats nous pourrions dans un premier temps nous intéresser à la valeur de la vitesse horizontale U le long de la ligne verticale passant par le centre de la cavité $x = 1/2$ (profil de vitesse U) et la valeur de la vitesse verticale V le long de la ligne horizontale passant par le centre de la cavité $y = 1/2$ (profil de vitesse V).

Pour chacune des 3 méthodes, on compare ces valeurs aux résultats tirés de Ghia et al. [14]. Pour ce faire, on interpole par splines cubiques nos profils de vitesse aux mêmes points que ceux de l'article et on calcule la norme $\|\cdot\|_2$ de la différence.

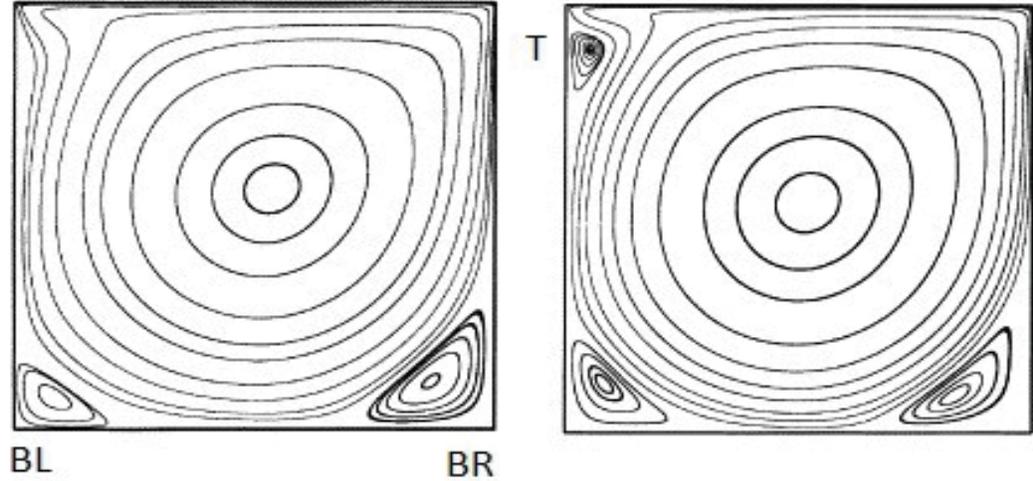


FIGURE 4.6. Lignes de courant de la solution en régime permanent pour $Re=1000$ (gauche) et $Re=3200$ (droite). Ces images sont tirées de Yang et Fu [11]

Soit \vec{U}_G et \vec{V}_G les profils de référence, et soit \vec{U} et \vec{V} nos profils interpolés, on s'intéresse donc aux erreurs :

$$e_U = \|\vec{U}_G - \vec{U}\|_2 \quad (4.5.1)$$

$$e_V = \|\vec{V}_G - \vec{V}\|_2 \quad (4.5.2)$$

Plutôt que de répertorier les erreurs calculées pour les profils U et V séparément, les tableaux qui suivent contiennent la somme des deux erreurs $e_U + e_V$ pour les différents choix des paramètres n, Re, dt .

Dans un deuxième temps, on s'intéressera aux lignes de courant du champ de vitesse \vec{U} . Ces lignes de courant sont les courbes de niveau de la fonction de courant Φ qui satisfait

$$\Phi_x = V$$

$$\Phi_y = -U$$

Cette fonction existe car le champ de vitesse est à divergence nulle et ainsi Φ satisfait $\Phi_{xy} - \Phi_{yx} = V_y + U_x = 0$.

Pour trouver Φ il suffira de résoudre l'équation de Poisson suivante

$$\nabla^2 \Phi = V_x - U_y$$

La Figure 4.6 montre l'allure des lignes de courant attendues pour $Re = 1000$ et $Re = 3200$. On remarque qu'en plus du vortex principal, des vortex secondaires se créent dans les coins en bas à gauche de la cavité ainsi qu'en bas à droite. Pour

$Re = 3200$, un vortex supplémentaire en haut à gauche apparaît. Il est présent aussi pour $Re = 5000$.

On comparera la localisation du centre de ces vortex avec les résultats de Ghia et al. [14]. Pour garder la nomenclature employée dans cet article, on nommera les vortex secondaires BL, "bottom left", BR, "bottom right" et T, "top" (voir Figure 4.6).

On pourra aussi comparer avec les résultats de Ghia et al. la valeur de Φ au centre des vortex. Les simulations pour générer les lignes de courant et calculer les valeurs associées aux vortex ont été effectuées avec $\Delta t = 0.00625$. Dans la plupart des tests, pour des maillages trop fins $n \geq 128$, les 3 méthodes donnent des résultats assez semblables. Les sections suivantes afficheront les comparatifs pour des petites valeurs de n . Les simulations sur les maillages plus fins serviront au calcul de l'ordre de convergence de l'erreur.

4.5.1.1. $Re=100$

| $n \backslash \Delta t$ | 0.4 | 0.2 | 0.1 | 0.05 | 0.025 | 0.0125 | 0.00625 |
|-------------------------|-------|-------|-------|-------|-------|--------|---------|
| 16 | 0.419 | 0.376 | 0.350 | 0.331 | 0.317 | 0.309 | 0.304 |
| 32 | 0.320 | 0.247 | 0.201 | 0.165 | 0.138 | 0.121 | 0.112 |
| 64 | 0.292 | 0.206 | 0.150 | 0.107 | 0.075 | 0.055 | 0.043 |
| 128 | 0.286 | 0.191 | 0.129 | 0.083 | 0.049 | 0.027 | 0.014 |

TABLEAU 4.1. Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode de De l'Isle-Owens pour $Re=100$

| $n \backslash \Delta t$ | 0.4 | 0.2 | 0.1 | 0.05 | 0.025 | 0.0125 | 0.00625 |
|-------------------------|-------|-------|-------|-------|-------|--------|---------|
| 16 | 0.417 | 0.372 | 0.342 | 0.316 | 0.293 | 0.275 | 0.262 |
| 32 | 0.320 | 0.247 | 0.200 | 0.163 | 0.135 | 0.116 | 0.104 |
| 64 | 0.292 | 0.206 | 0.150 | 0.107 | 0.075 | 0.054 | 0.042 |
| 128 | 0.286 | 0.191 | 0.129 | 0.083 | 0.049 | 0.027 | 0.014 |

TABLEAU 4.2. Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode de Fatone et al. pour $Re=100$

| $n \backslash \Delta t$ | 0.4 | 0.2 | 0.1 | 0.05 | 0.025 | 0.0125 | 0.00625 |
|-------------------------|-------|-------|-------|-------|-------|--------|---------|
| 16 | 0.441 | 0.403 | 0.383 | 0.371 | 0.363 | 0.358 | 0.356 |
| 32 | 0.327 | 0.256 | 0.212 | 0.179 | 0.155 | 0.140 | 0.132 |
| 64 | 0.294 | 0.208 | 0.153 | 0.111 | 0.080 | 0.060 | 0.048 |
| 128 | 0.286 | 0.191 | 0.130 | 0.084 | 0.050 | 0.028 | 0.015 |

TABLEAU 4.3. Différence entre les profils de vitesse de Ghia et al. et ceux calculés par la méthode centrée pour $Re=100$

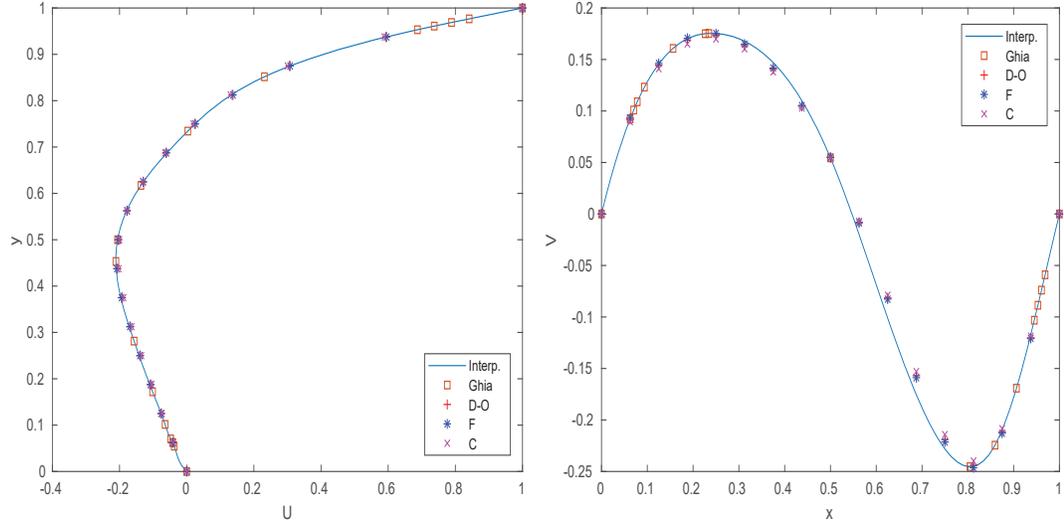


FIGURE 4.7. Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 100$, $\Delta t = 0.00625$ et $n = 32$

| | <i>Ghia</i> | De l'Isle-Owens | Fatone | Centré |
|--------|-------------|-----------------|------------|------------|
| Φ | -0.103 | -0.101 | -0.102 | -0.1 |
| x | 0.617 | 0.616 | 0.615 | 0.619 |
| y | 0.734 | 0.74 | 0.739 | 0.743 |
| Φ | $1.75e-06$ | $7.06e-06$ | $7.14e-06$ | $6.92e-06$ |
| x | 0.0313 | 0.039 | 0.039 | 0.039 |
| y | 0.0391 | 0.041 | 0.041 | 0.041 |
| Φ | $1.25e-05$ | $2.47e-05$ | $2.53e-05$ | $2.36e-05$ |
| x | 0.945 | 0.939 | 0.939 | 0.94 |
| y | 0.0625 | 0.0671 | 0.0681 | 0.0671 |

TABLEAU 4.4. Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 100$, $n = 32$, $\Delta t = 0.00625$

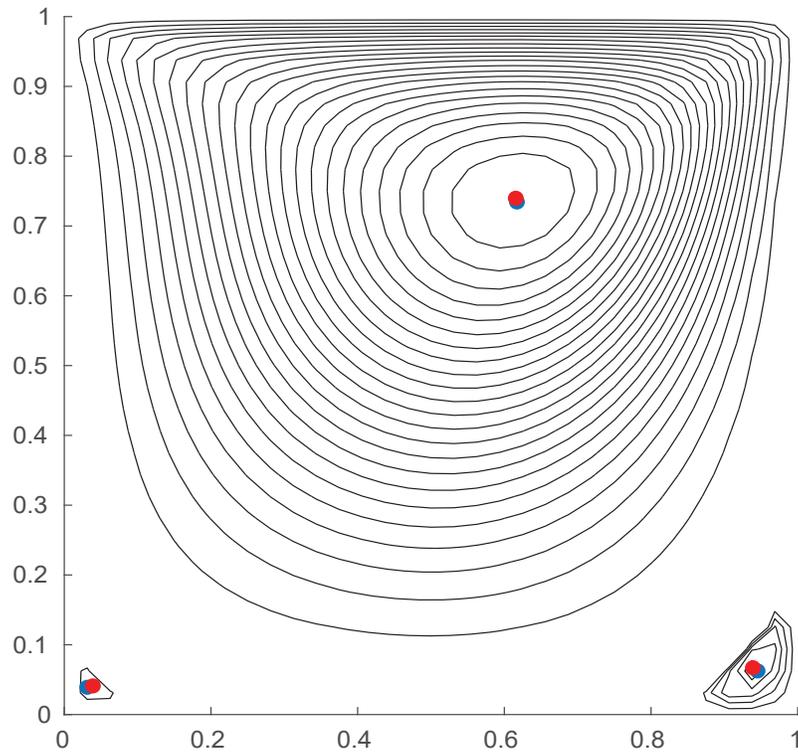


FIGURE 4.8. Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 100$, $n = 32$ et pour la méthode De l'Isle-Owens. La localisation des centres de référence (Ghia) est en bleu

4.5.1.2. $Re=400$

| $n \backslash \Delta t$ | D-O | F | C |
|-------------------------|-------|-------|-------|
| 16 | 0.867 | 0.582 | 1.118 |
| 32 | 0.354 | 0.274 | 0.481 |
| 64 | 0.199 | 0.188 | 0.226 |
| 128 | 0.172 | 0.171 | 0.178 |

TABLEAU 4.5. Valeur de $\|\vec{U}_G - \vec{U}\|_2 + \|\vec{V}_G - \vec{V}\|_2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 400$, $\Delta t = 0.00625$

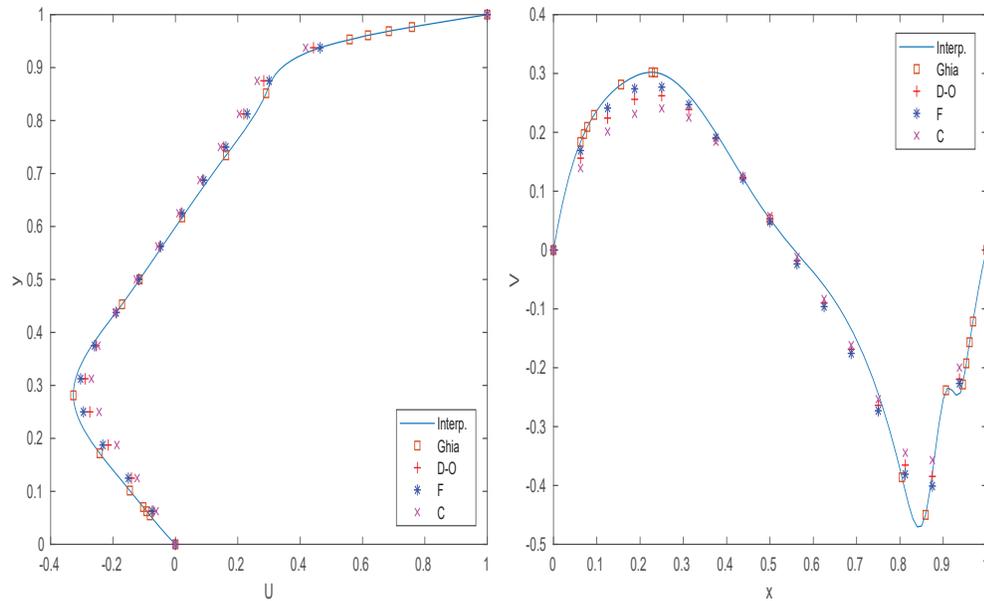


FIGURE 4.9. Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 400$, $\Delta t = 0.00625$ et $n = 32$

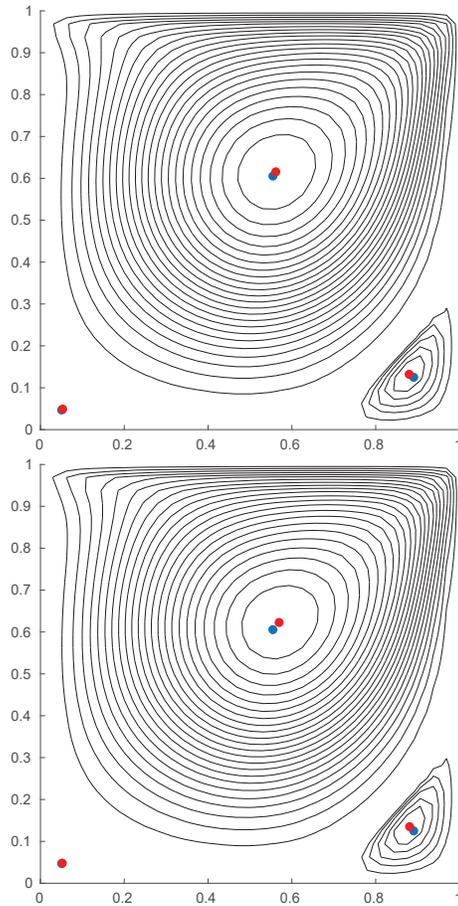


FIGURE 4.10. Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 400$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), et centrée (bas). La localisation des centres de référence (Ghia) est en bleu

| | <i>Ghia</i> | De l'Isle-Owens | Fatone | Centré |
|--------|--------------|-----------------|--------------|--------------|
| Φ | -0.114 | -0.103 | -0.107 | -0.0975 |
| x | 0.555 | 0.562 | 0.557 | 0.57 |
| y | 0.606 | 0.616 | 0.613 | 0.623 |
| Φ | $1.42e - 05$ | $2.19e - 05$ | $2.51e - 05$ | $1.79e - 05$ |
| x | 0.0508 | 0.0531 | 0.0551 | 0.0521 |
| y | 0.0469 | 0.049 | 0.049 | 0.048 |
| Φ | 0.000642 | 0.000734 | 0.000809 | 0.000687 |
| x | 0.891 | 0.88 | 0.877 | 0.881 |
| y | 0.125 | 0.132 | 0.131 | 0.135 |

TABLEAU 4.6. Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 400$, $n = 32$, $\Delta t = 0.00625$

4.5.1.3. $Re=1000$

| n \ Δt | D-O | F | C |
|----------------|-------|-------|-------|
| 16 | 1.159 | 0.631 | 1.468 |
| 32 | 0.711 | 0.577 | 0.962 |
| 64 | 0.296 | 0.250 | 0.394 |
| 128 | 0.180 | 0.174 | 0.205 |

TABLEAU 4.7. Valeur de $\|\vec{U}_G - \vec{U}\|_2 + \|\vec{V}_G - \vec{V}\|_2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 1000$, $\Delta t = 0.00625$

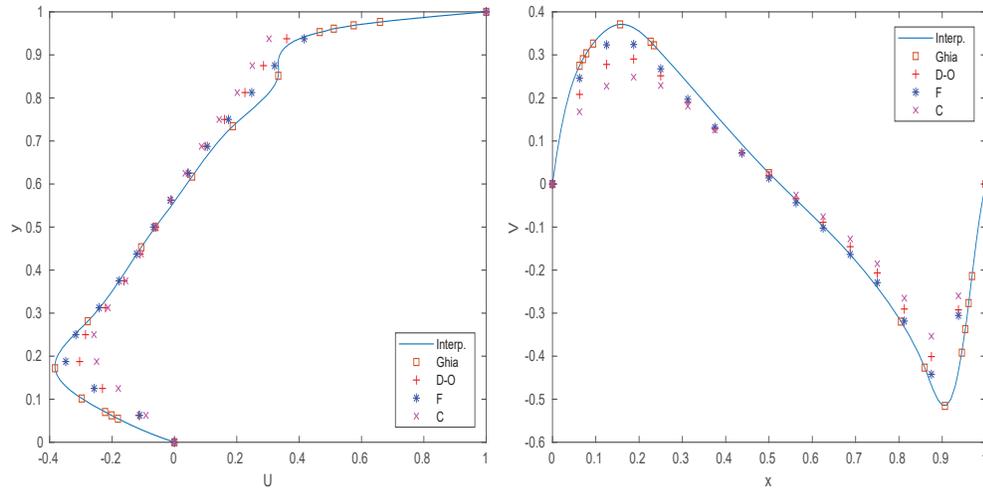


FIGURE 4.11. Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 1000$, $\Delta t = 0.00625$ et $n = 32$

| | <i>Ghia</i> | De l'Isle-Owens | Fatone | Centré |
|--------|-------------|-----------------|----------|--------------|
| Φ | -0.118 | -0.0951 | -0.105 | -0.0846 |
| x | 0.531 | 0.535 | 0.527 | 0.544 |
| y | 0.562 | 0.58 | 0.58 | 0.588 |
| Φ | 0.000231 | 0.00015 | 0.000205 | $9.66e - 05$ |
| x | 0.0859 | 0.0811 | 0.0841 | 0.0791 |
| y | 0.0781 | 0.0711 | 0.0751 | 0.0681 |
| Φ | 0.00175 | 0.00188 | 0.00213 | 0.00173 |
| x | 0.859 | 0.85 | 0.839 | 0.854 |
| y | 0.109 | 0.126 | 0.126 | 0.131 |

TABLEAU 4.8. Localisation du centre des vortex Primaire-BL-BR (de haut en bas) et valeur de Φ en ces points pour $Re = 1000$, $n = 32$, $\Delta t = 0.00625$

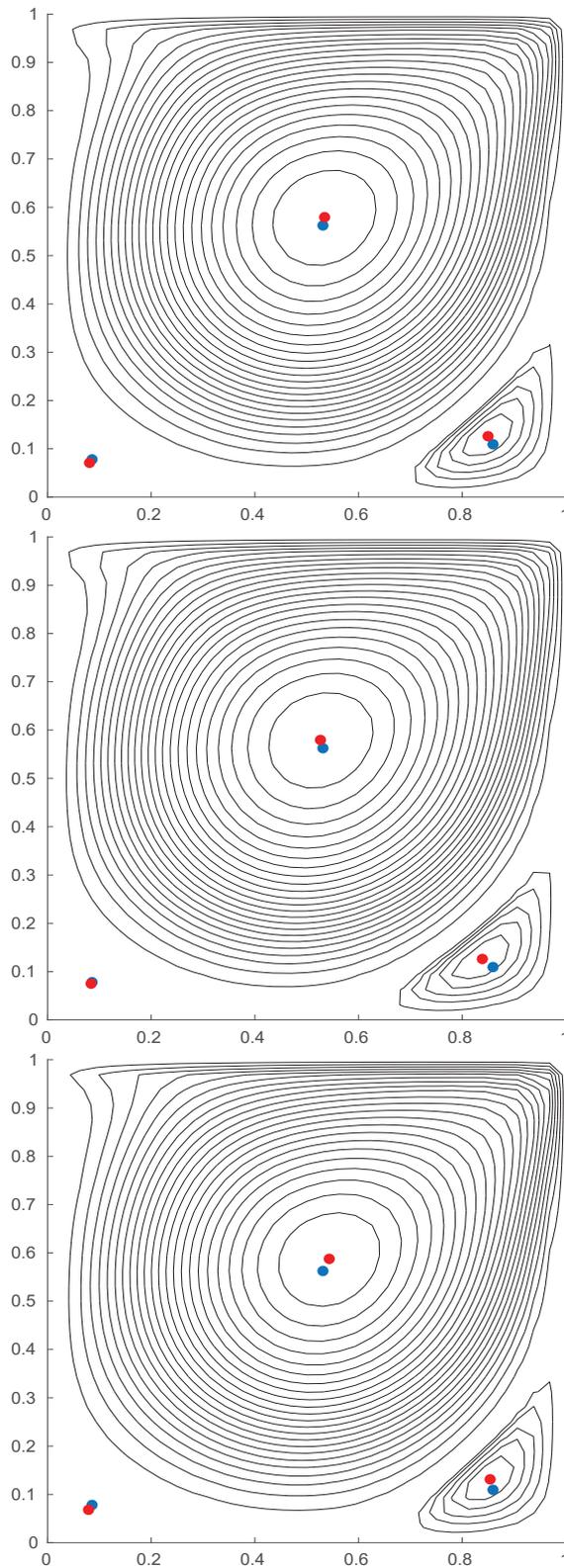


FIGURE 4.12. Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 1000$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu

4.5.1.4. $Re=3200$

| n \ Δt | D-O | F | C |
|----------------|-------|-------|-------|
| 16 | 1.650 | 0.802 | 1.949 |
| 32 | 1.226 | 0.691 | 1.514 |
| 64 | 0.642 | 0.497 | 0.897 |
| 128 | 0.401 | 0.375 | 0.480 |

TABLEAU 4.9. Valeur de $\|\vec{U}_G - \vec{U}\|_2 + \|\vec{V}_G - \vec{V}\|_2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 3200$, $\Delta t = 0.00625$

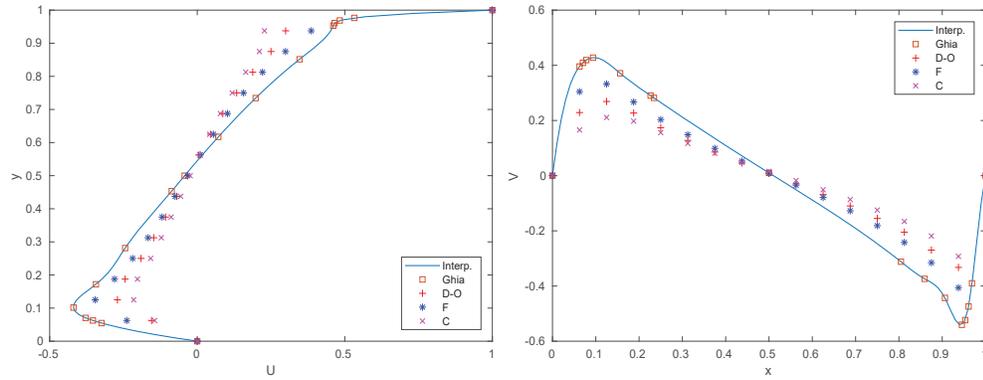


FIGURE 4.13. Comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 3200$, $\Delta t = 0.00625$ et $n = 32$

| | <i>Ghia</i> | De l'Isle-Owens | Fatone | Centré |
|--------|-------------|-----------------|----------|----------|
| Φ | -0.12 | -0.0775 | -0.0944 | -0.0652 |
| x | 0.516 | 0.522 | 0.52 | 0.537 |
| y | 0.547 | 0.557 | 0.552 | 0.552 |
| Φ | 0.000978 | 0.000552 | 0.000793 | 0.000479 |
| x | 0.0859 | 0.0851 | 0.0831 | 0.0881 |
| y | 0.109 | 0.103 | 0.109 | 0.107 |
| Φ | 0.00314 | 0.00249 | 0.00206 | 0.00225 |
| x | 0.812 | 0.81 | 0.825 | 0.84 |
| y | 0.0859 | 0.103 | 0.0961 | 0.101 |
| Φ | 0.000728 | 0.000335 | 0.000402 | 0.000326 |
| x | 0.0547 | 0.038 | 0.033 | 0.044 |
| y | 0.898 | 0.887 | 0.888 | 0.917 |

TABLEAU 4.10. Localisation du centre des vortex Primaire-BL-BR-T (de haut en bas) et valeur de Φ en ces points pour $Re = 3200$, $n = 32$, $\Delta t = 0.00625$

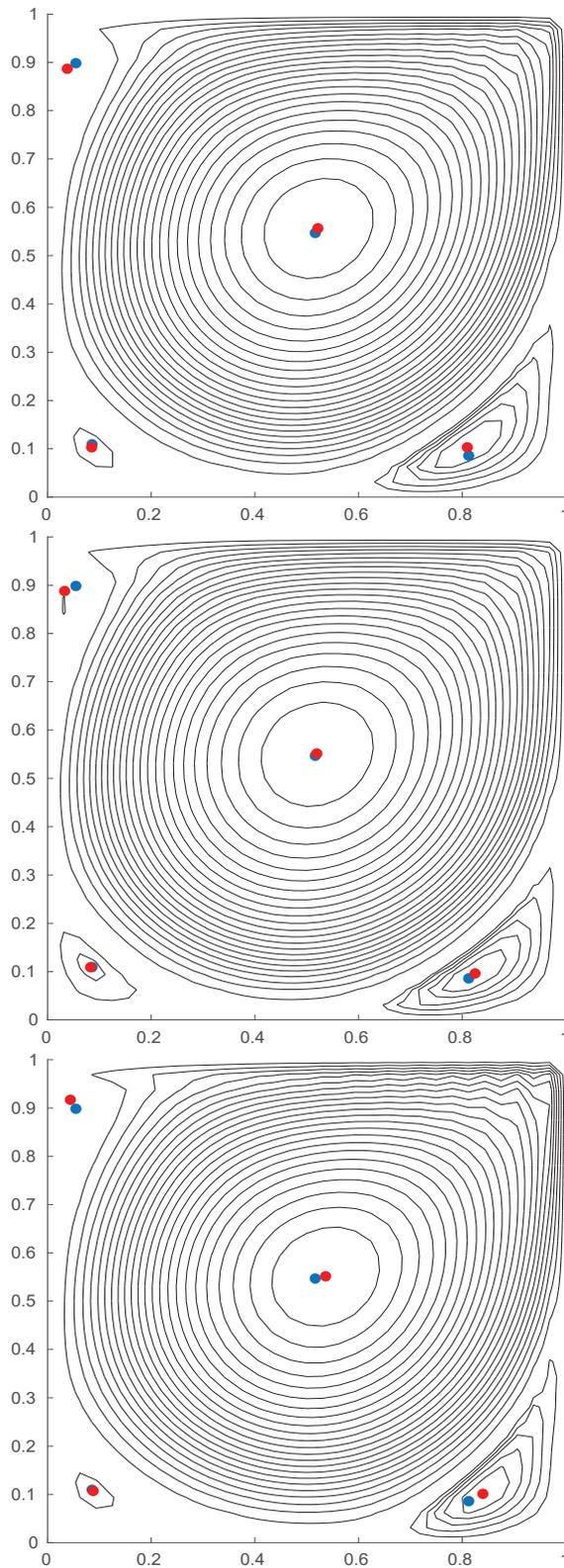


FIGURE 4.14. Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 3200$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu

4.5.1.5. $Re=5000$

| n \ Δt | D-O | F | C |
|----------------|-------|-------|-------|
| 32 | 1.337 | 0.755 | 1.695 |
| 64 | 0.845 | 0.696 | 1.141 |
| 128 | 0.505 | 0.468 | 0.623 |

TABLEAU 4.11. Valeur de $\|\vec{U}_G - \vec{U}\|_2 + \|\vec{V}_G - \vec{V}\|_2$ pour les méthodes de De l'Isle-Owens, Fatone et al. et centrée avec $Re = 5000$, $\Delta t = 0.00625$

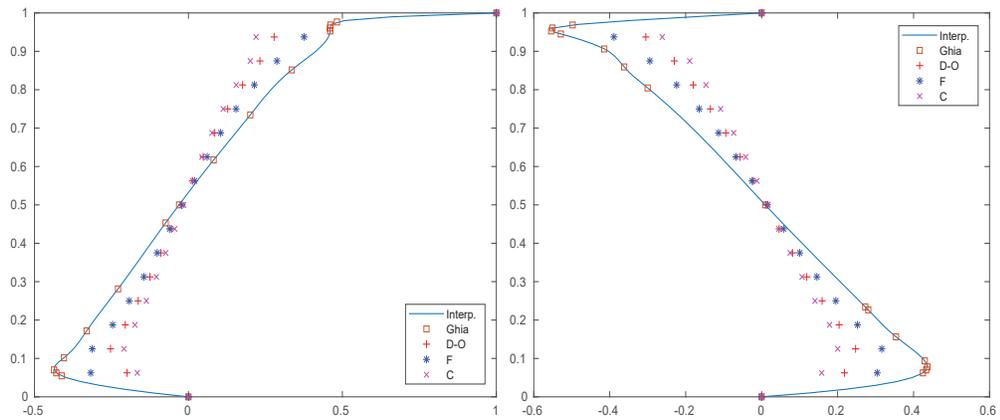


FIGURE 4.15. comparaison des profils de vitesse U (gauche) et V (droite) pour les méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C). $Re = 5000$, $\Delta t = 0.00625$ et $n = 32$

| | <i>Ghia</i> | De l'Isle-Owens | Fatone | Centré |
|--------|-------------|-----------------|----------|----------|
| Φ | -0.119 | -0.072 | -0.0921 | -0.0604 |
| x | 0.512 | 0.527 | 0.528 | 0.54 |
| y | 0.535 | 0.546 | 0.543 | 0.538 |
| Φ | 0.00136 | 0.000776 | 0.00111 | 0.00076 |
| x | 0.0703 | 0.0801 | 0.0791 | 0.0851 |
| y | 0.137 | 0.126 | 0.136 | 0.133 |
| Φ | 0.00308 | 0.0018 | 0.00138 | 0.0019 |
| x | 0.809 | 0.826 | 0.857 | 0.848 |
| y | 0.0742 | 0.0851 | 0.0731 | 0.0831 |
| Φ | 0.00146 | 0.000639 | 0.000734 | 0.000792 |
| x | 0.0625 | 0.0551 | 0.036 | 0.0591 |
| y | 0.91 | 0.889 | 0.884 | 0.92 |

TABLEAU 4.12. Localisation du centre des vortex Primaire-BL-BR-T (de haut en bas) et valeur de Φ en ces points pour $Re = 5000$, $n = 32$, $\Delta t = 0.00625$

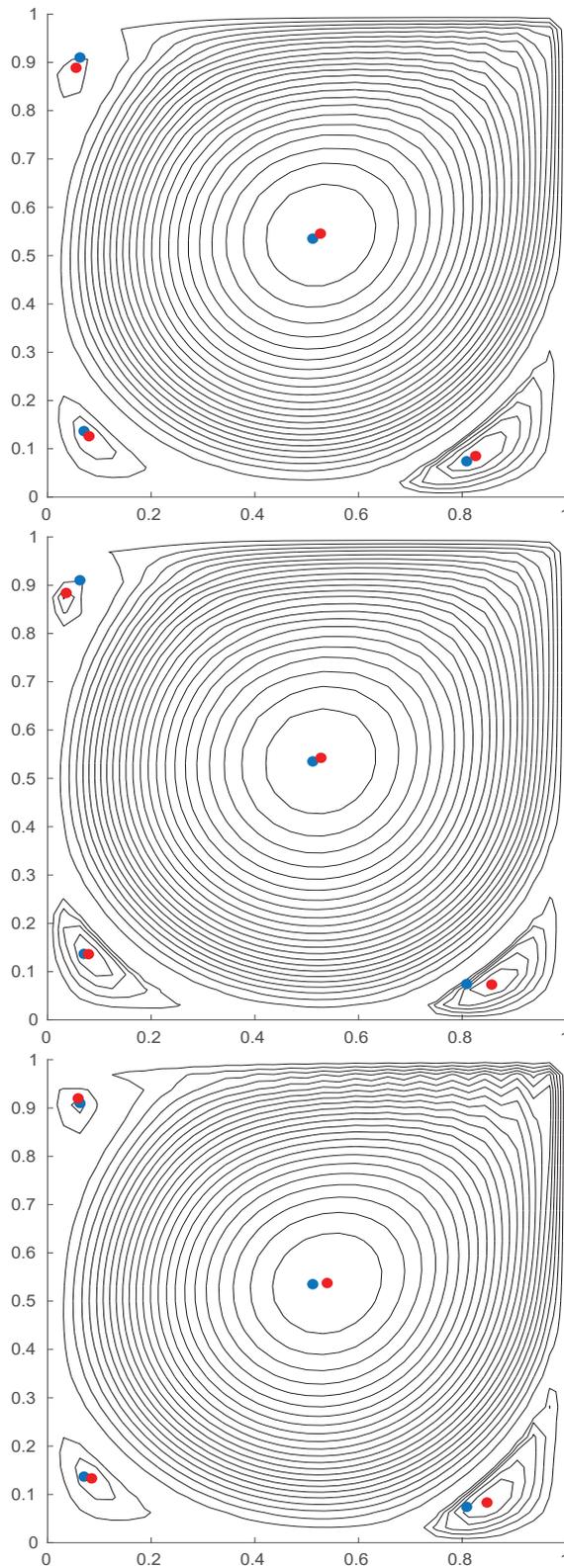


FIGURE 4.16. Lignes de courant et localisation du centre des vortex (en rouge) pour $Re = 5000$, $n = 32$ et pour les méthodes De l'Isle Owens (haut), Fatone et al. (milieu) et centrée (bas). La localisation des centres de référence (Ghia) est en bleu

4.5.1.6. *Ordre de convergence*

Nous avons généré une grande quantité de simulations avec $\Delta t = 0.001$ et avec plusieurs maillages $16 \leq n \leq 160$ pour que l'erreur liée au pas de temps soit négligeable et ainsi espérer extraire avec précision l'ordre de convergence par rapport à h . Pour $Re = 3200, 5000$, ces données étaient trop longues à récupérer car le régime permanent est atteint beaucoup plus tard dans les simulations donc nous montrons seulement les résultats pour $Re = 100, 400, 1000$.

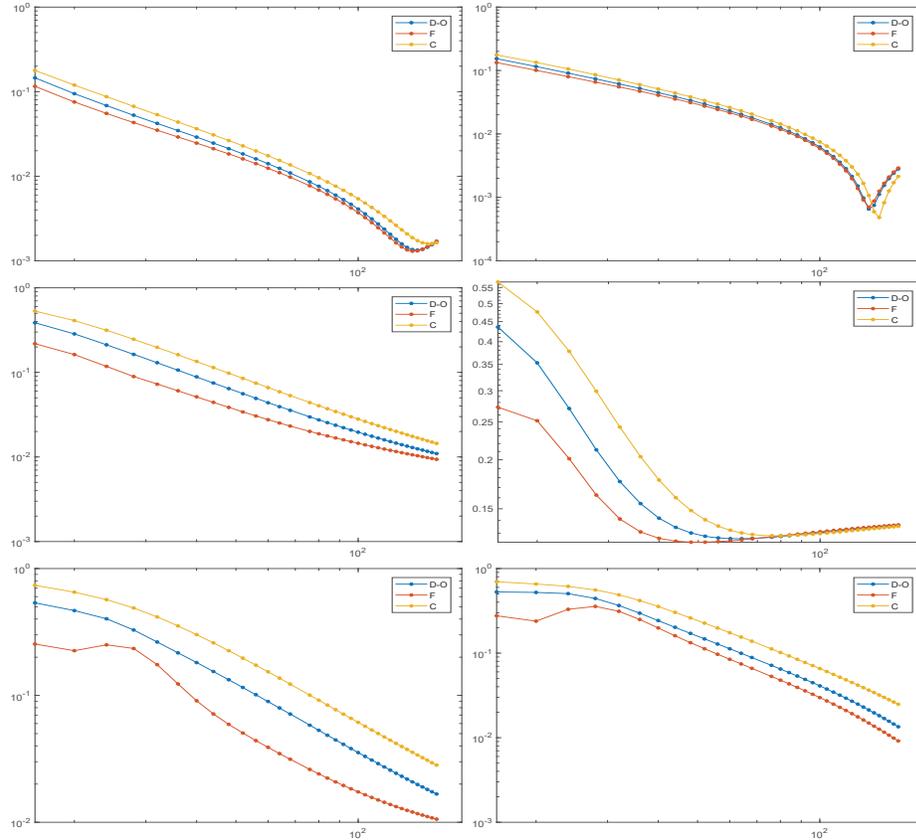


FIGURE 4.17. Graphique log-log de $e_U = \|\vec{U}_G - \vec{U}\|_2$ (gauche) et $e_V = \|\vec{V}_G - \vec{V}\|_2$ (droite) pour $Re = 100, 400, 1000$ (de haut en bas) et pour les 3 méthodes De l'Isle-Owens (D-O), Fatone et al. (F) et centrée (C)

| Re | D-O | | F | | C | |
|------|------|------|------|------|------|------|
| | eU | eV | eU | eV | eU | eV |
| 100 | 1.79 | 1.54 | 1.70 | 1.48 | 1.79 | 1.55 |
| 400 | 1.60 | 1.38 | 1.37 | 1.07 | 1.65 | 1.53 |
| 1000 | 1.77 | 2.08 | 1.58 | 2.20 | 1.76 | 1.94 |

TABLEAU 4.13. Ordre de convergence de l'erreur en fonction de h

4.5.2. Discussion des résultats

Il est normal que nos résultats ne soient pas particulièrement précis lorsqu'on les compare aux données de Ghia et al. [14] car à chaque pas de temps, ces derniers utilisent une méthode itérative sur plusieurs grilles de plus en plus fines (méthode multigrilles) et qui permet de réduire considérablement l'erreur. Plutôt que de résoudre les équations de Navier-Stokes avec précision nous cherchons plutôt à démontrer qu'en faisant un simple décalage du point de collocation, le schéma superconsistant résultant est plus performant que les schémas centrés de dérivée première et seconde.

Par le membre de droite de l'équation (4.3.19), on s'attend à avoir une erreur de l'ordre de Δt en temps. C'est ce que l'on peut vérifier pour les profils de vitesse dans les Tableaux 4.1 à 4.3 car lorsqu'on divise le pas de temps par 2, l'erreur est au mieux divisée par 2 donc la convergence n'est pas plus rapide que linéaire.

On s'attend aussi à une convergence quadratique en h . Si on divise par deux le pas de discrétisation entre deux simulations on s'attend alors à diviser l'erreur par 4. Cette convergence commence à être plus notable dans la dernière colonne des Tableaux 4.1 à 4.3 car l'erreur liée à Δt devient négligeable par rapport à celle liée à h . Pour les valeurs de $Re > 100$, seuls les résultats avec $\Delta t = 0.00625$ sont montrés car c'est là que l'erreur liée à h domine.

Dans la Figure 4.17 on voit que les 3 méthodes convergent à peu près de la même façon. Dans le Tableau 4.13 on a calculé l'ordre de convergence et on voit qu'il est au moins plus que linéaire bien qu'il ne soit pas tout à fait d'ordre 2. Étonnamment, la convergence est meilleure lorsque Re augmente mais ceci est probablement dû au fait que pour des maillages grossiers, l'erreur liée à h est particulièrement grande par rapport à celle liée à Δt et donc on observe plus facilement une convergence quadratique lorsque h diminue.

Pour Re plus petit, l'erreur liée à Δt est moins négligeable devant celle liée à h et impose plus facilement une limite sur l'ordre de convergence.

Comme les schémas sont tous d'ordre 2 dans l'espace et que la seule différence est dans la localisation du point de collocation lors de la résolution des équations (4.3.20) et (4.3.21), pour des maillages fins il y a peu de différence entre les 3 méthodes. Nous avons donc fait la plupart des comparatifs pour $n = 32$ afin de voir quelle méthode performe le mieux sur des maillages grossiers.

Globalement, lorsqu'on regarde la Figure 4.17 et les Tableaux des profils de vitesse 4.5, 4.7, 4.9 et 4.11, il est clair que la méthode superconsistante est préférable à la méthode centrée. En effet, l'erreur liée aux champs de vitesse est toujours plus grande pour (C) que pour (D-O) et l'écart grandit lorsque le nombre de

Reynolds augmente. En observant les Figures 4.10, 4.12, 4.14 et les Tableaux associées 4.6, 4.8, 4.10, on voit que le centre des vortex est de moins en moins juste pour les deux méthodes à mesure que Re augmente. Cependant le décalage est moins grave pour (D-O) que pour (C). Pour $Re = 3200$ et $Re = 5000$ on remarque dans les Figures 4.14 et 4.16 que les lignes de courant pour la méthode centrée présentent plus d'oscillations que pour la méthode (D-O).

Lorsqu'on compare (D-O) et (F), il semble que la deuxième approche performe mieux. Tout d'abord, l'erreur est systématiquement plus grande pour (D-O) dans les profils de vitesse et ce pour toutes les valeurs de n testées. Aussi, les lignes de courant pour $Re = 5000$ (Figure 4.16) présentent encore moins d'oscillations avec (F) qu'avec (D-O).

Nous avons vu au chapitre précédent que lorsque le champ de vitesse est orienté proche de la direction d'un des axes les solutions données par la méthode de Fatone et al. sont meilleures que celles de De l'Isle-Owens. Dans le problème de cavité entraînée, le champ de vitesse est parallèle aux frontières dans la couche limite donc on soupçonne que cela explique les moins bons résultats donnés par (D-O).

En regardant les Figures 4.7, 4.9 , 4.11 , 4.13 et 4.15 on voit en effet que c'est surtout proche des frontières que les différences entre les trois méthodes apparaissent le plus et que les solutions de Fatone et al. sont systématiquement plus proches de celles de Ghia.

Par rapport aux résultats du chapitre 3 nous n'avons pas réussi à trouver des choix de paramètre où la méthode (D-O) performe particulièrement mieux que (F).

Remarque 4.5.1. *Pour les plus grandes valeurs de Re soit 3200 et 5000, il est possible que le critère d'arrêt soit trop élevé et que le régime permanent ne soit pas encore atteint. Nous ne pouvons donc pas comparer adéquatement les vortex secondaires puisque ceux-ci n'ont peut être pas tout à fait fini de se stabiliser.*

CONCLUSION

Dans ce travail, nous avons considéré les discrétisations construites à partir d'un polynôme d'interpolation local et nous avons montré qu'en choisissant le point de collocation superconsistant basé sur la taille de l'espace Q des polynômes p tel que le résidu est nul sur la grille de collocation, le schéma gagne un ordre de consistance. Les simulations numériques ont montré que ces schémas sont particulièrement stables au sens où ils diminuent grandement la présence d'oscillations non désirées dans les solutions numériques de problèmes présentant une couche limite. En dimension 2, le choix de notre point de collocation diffère du choix fait par Fatone et al. [10] bien que les deux approches augmentent la taille de Q et la consistance du schéma. Notre choix est basé sur l'annulation de certains termes dans le développement de Taylor et est donc applicable à la plupart des opérateurs linéaires. L'approche de Fatone et al. est basée sur la forme de l'opérateur d'advection-diffusion et il n'est pas clair si on peut généraliser la méthode à d'autres types d'opérateur tout en assurant une augmentation de l'ordre de consistance. Dans le cas précis de l'opérateur d'advection-diffusion, il semble que l'approche de Fatone et al. soit plus robuste car elle dépend directement du champ de vitesse et s'adapte mieux lorsque celui-ci est très variable dans l'espace. Si, pour certains choix de paramètres dans les problèmes du Chapitre 3, notre méthode se comportait mieux, les résultats du Chapitre 4 sont plus catégoriques quant à l'avantage de la méthode de Fatone et al. Dans ce travail nous n'avons pas transcrit de résultats numériques associés à des discrétisations sur des grilles non uniformes c'est à dire où h n'est pas constant. Il n'est pas rare de vouloir utiliser ce genre de grille lorsqu'on veut étudier avec plus de précision ce qu'il se passe dans certaines régions du domaine. Ainsi, au chapitre 3 et 4 on aurait pu exiger que la distance entre les points soit plus petite proche des frontières, là où il y a une couche limite. Des premiers tests ont été effectués pour comparer la méthode de De l'Isle Owens à celle de Fatone et al. sur de telles grilles et les résultats étaient prometteurs pour la première.

Annexe A

DÉVELOPPEMENT EN SÉRIE DE TAYLOR

Pour une fonction lisse à une seule variable $u(x) \in C^\infty(\mathbb{R})$ autour de $x = a$

$$u(x) = \sum_{i=0}^{\infty} u^{(i)}(a) \frac{(x-a)^i}{i!}$$

Pour une fonction lisse à 2 variables $u(x, y) \in C^\infty(\mathbb{R}^2)$ autour de $(x, y) = (a, b)$

$$u(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} C_{i,j} (x-a)^i (y-b)^j$$

$$\begin{aligned} \frac{\partial^{k+l}}{\partial x^k \partial y^l} u(x, y) &= \sum_{i=k}^{\infty} \sum_{j=l}^{\infty} C_{i,j} \frac{i!}{(i-k)!} (x-a)^{i-k} \frac{j!}{(j-l)!} (y-b)^{j-l} \\ \frac{\partial^{k+l}}{\partial x^k \partial y^l} u(a, b) &= C_{k,l} k! l! \end{aligned}$$

et donc

$$u(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{i! j!} \frac{\partial^{i+j}}{\partial x^i \partial y^j} u(a, b) (x-a)^i (y-b)^j$$

Annexe B

INÉGALITÉS DISCRÈTES

Théorème B.0.1. (*Cauchy-Schwartz*)

Soit $a = (a_{i,j})$ et $b = (b_{i,j})$ avec $i = 0, \dots, n$ et $j = 0, \dots, m$ alors

$$\sum_{i=0}^n \sum_{j=0}^m a_{i,j} b_{i,j} \leq \sqrt{\sum_{i=0}^n \sum_{j=0}^m a_{i,j}^2} \sqrt{\sum_{i=0}^n \sum_{j=0}^m b_{i,j}^2} = \|a\| \|b\|$$

DÉMONSTRATION. Soit $t \in \mathbb{R}$, on pose $P(t) = \|a + tb\|^2$. Ainsi,

$$\begin{aligned} P(t) &= \sum_{i=0}^n \sum_{j=0}^m (a_{i,j} + tb_{i,j})^2 \\ &= \sum_{i=0}^n \sum_{j=0}^m (a_{i,j}^2 + 2ta_{i,j}b_{i,j} + t^2b_{i,j}^2) \\ &= \|a\|^2 + 2t \sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} + t^2\|b\|^2 \end{aligned} \tag{B.0.1}$$

Par construction, $P(t)$ est positif ou nul $\forall t$ donc le polynôme de degré 2 $\|a\|^2 + 2t \sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} + t^2\|b\|^2$ est aussi positif ou nul pour tout t . Son discriminant est donc négatif ou nul c.-à-d.

$$4 \left(\sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} \right)^2 - 4\|a\|^2\|b\|^2 \leq 0$$

De là découle

$$\left(\sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} \right)^2 \leq \|a\|^2\|b\|^2$$

Et finalement

$$\sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} \leq \left| \sum_{i=0}^n \sum_{j=0}^m a_{i,j}b_{i,j} \right| \leq \|a\| \|b\|$$

□

B-ii

Théorème B.0.2. (Poincaré) Soit $u = [u_0, u_1, \dots, u_n]$ avec les conditions frontiere $u_0 = u_n = 0$ alors

$$\frac{1}{n^2} \|u\|^2 \leq \sum_{i=0}^{n-1} (u_{i+1} - u_i)^2$$

DÉMONSTRATION. Soit $u_k = \max_i |u_i|$

$$\|u\|^2 = \sum_{i=0}^n u_i^2 \leq n u_k^2 = n \left(\sum_{i=0}^{k-1} (u_{i+1} - u_i) \right)^2 \leq n^2 \sum_{i=0}^{n-1} (u_{i+1} - u_i)^2$$

□

Théorème B.0.3. (Poincaré 2D) Soit $u = (u_{i,j})$, $i = 0, \dots, n$ et $j = 0, \dots, m$ avec les conditions frontiere $u_{0,j} = u_{n,j} = u_{i,0} = u_{i,m} = 0$ alors

$$\|u\|^2 = \sum_{i=0}^n \sum_{j=0}^m u_{i,j}^2 \leq n^2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (u_{i+1,j} - u_{i,j})^2 + m^2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (u_{i,j+1} - u_{i,j})^2$$

DÉMONSTRATION. Soit $0 \leq i \leq n$, on pose

$$\|u_i\|^2 = \sum_{j=0}^m u_{i,j}^2$$

Similairement avec $0 \leq j \leq m$:

$$\|u_j\|^2 = \sum_{i=0}^n u_{i,j}^2$$

Ainsi

$$\begin{aligned} \|u\|^2 &= \sum_{i=0}^n \sum_{j=0}^m u_{i,j}^2 \\ &\leq \sum_{i=0}^n \|u_i\|^2 + \sum_{j=0}^m \|u_j\|^2 \\ &= \sum_{i=0}^{n-1} \|u_i\|^2 + \sum_{j=0}^{m-1} \|u_j\|^2 \\ &\leq m^2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (u_{i,j+1} - u_{i,j})^2 + n^2 \sum_{j=0}^{m-1} \sum_{i=0}^{n-1} (u_{i+1,j} - u_{i,j})^2 \end{aligned} \tag{B.0.2}$$

La troisième ligne utilise les conditions frontieres et la dernière utilise (B.0.2)

□

Annexe C

IDENTITÉS SOMMATION

$$\begin{aligned}
 \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i+1,j+1} &= \sum_{i=2}^n \sum_{j=2}^m U_{i,j} U_{i-1,j-1} = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i-1,j-1} \\
 \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i,j+1} &= \sum_{i=1}^{n-1} \sum_{j=2}^m U_{i,j} U_{i,j-1} = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i,j-1} \\
 \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i+1,j} &= \sum_{i=2}^n \sum_{j=1}^{m-1} U_{i,j} U_{i-1,j} = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i-1,j} \\
 \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i+1,j-1} &= \sum_{i=2}^n \sum_{j=0}^{m-2} U_{i,j} U_{i-1,j+1} = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i-1,j+1}
 \end{aligned}$$

Dans chaque cas, la première égalité vient simplement d'un changement d'indices et la deuxième est vraie grâce aux conditions frontière.

On a aussi

$$\begin{aligned}
 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})^2 &= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} U_{i,j}^2 - 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} U_{i,j} U_{i+1,j} + \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} U_{i+1,j}^2 \\
 &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j}^2 - 2 \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} U_{i+1,j} + \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j}^2 \\
 &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (2U_{i,j}^2 - 2U_{i,j} U_{i+1,j})
 \end{aligned}$$

Et similairement

$$\begin{aligned}
 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})^2 &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (2U_{i,j}^2 - 2U_{i,j} U_{i,j+1}) \\
 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i+1,j} - U_{i,j+1})^2 &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (2U_{i,j}^2 - 2U_{i,j} U_{i-1,j+1})
 \end{aligned}$$

C-ii

$$\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j+1})^2 = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (2U_{i,j}^2 - 2U_{i,j}U_{i+1,j+1})$$

On a donc

$$\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (U_{i,j}U_{i,j-1} - 2U_{i,j}^2 + U_{i,j}U_{i,j+1}) = - \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})^2 \quad (\text{C.0.1})$$

$$\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (U_{i,j}U_{i-1,j} - 2U_{i,j}^2 + U_{i,j}U_{i+1,j}) = - \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})^2 \quad (\text{C.0.2})$$

$$\begin{aligned} \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (U_{i,j}U_{i,j+1} - U_{i,j}U_{i,j-1}) &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} (U_{i,j}U_{i+1,j} - U_{i,j}U_{i-1,j}) \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j}(U_{i-1,j-1} + U_{i-1,j+1} - U_{i+1,j-1} - U_{i+1,j+1}) \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j}(U_{i-1,j-1} - U_{i-1,j+1} + U_{i+1,j-1} - U_{i+1,j+1}) = 0 \end{aligned} \quad (\text{C.0.3})$$

$$\begin{aligned} &\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j}(U_{i-1,j-1} - U_{i-1,j} + U_{i-1,j+1} - U_{i,j-1} - U_{i,j+1} + U_{i+1,j-1} - U_{i+1,j} + U_{i+1,j+1}) \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} 2U_{i,j}U_{i+1,j+1} - 2U_{i,j}U_{i+1,j} + 2U_{i,j}U_{i-1,j+1} - 2U_{i,j}U_{i,j+1} \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} 2U_{i,j}U_{i+1,j+1} - 2U_{i,j}^2 + 2U_{i,j}^2 - 2U_{i,j}U_{i+1,j} + 2U_{i,j}U_{i-1,j+1} - 2U_{i,j}^2 + 2U_{i,j}^2 - 2U_{i,j}U_{i,j+1} \\ &= - \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j+1})^2 + \sum_{i=0}^{n-1} \sum_{j=1}^{m-1} (U_{i,j} - U_{i+1,j})^2 - \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i+1,j} - U_{i,j+1})^2 + \sum_{i=1}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})^2 \\ &= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} -(U_{i,j} - U_{i+1,j+1})^2 + (U_{i,j} - U_{i+1,j})^2 - (U_{i+1,j} - U_{i,j+1})^2 + (U_{i,j} - U_{i,j+1})^2 \end{aligned} \quad (\text{C.0.4})$$

$$\begin{aligned}
& \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} U_{i,j} (U_{i-1,j-1} - U_{i-1,j+1} - U_{i+1,j-1} + U_{i+1,j+1}) \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} 2U_{i,j} U_{i+1,j+1} - 2U_{i,j} U_{i-1,j+1} \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} 2U_{i,j} U_{i+1,j+1} - 2U_{i,j}^2 + 2U_{i,j}^2 - 2U_{i,j} U_{i-1,j+1} \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} -(U_{i,j} - U_{i+1,j+1})^2 + (U_{i+1,j} - U_{i,j+1})^2
\end{aligned} \tag{C.0.5}$$

$$\begin{aligned}
& \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j+1})^2 - (U_{i+1,j} - U_{i,j+1})^2 \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j} + U_{i+1,j} - U_{i+1,j+1})^2 - (U_{i+1,j} - U_{i,j} + U_{i,j} - U_{i,j+1})^2 \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})^2 + (U_{i+1,j} - U_{i+1,j+1})^2 - (U_{i+1,j} - U_{i,j})^2 - (U_{i,j} - U_{i,j+1})^2 \\
&+ 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})(U_{i+1,j} - U_{i+1,j+1}) - 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})(U_{i+1,j} - U_{i,j}) \\
&= 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})(U_{i+1,j} - U_{i+1,j+1}) + (U_{i,j} - U_{i,j+1})(U_{i,j} - U_{i+1,j})
\end{aligned} \tag{C.0.6}$$

$$\begin{aligned}
& \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [(U_{i,j} - U_{i+1,j}) + (U_{i+1,j} - U_{i+1,j+1})]^2 \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})^2 + \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})^2 \\
&+ 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})(U_{i+1,j} - U_{i+1,j+1})
\end{aligned} \tag{C.0.7}$$

$$\begin{aligned}
& \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [(U_{i,j} - U_{i,j+1}) + (U_{i,j} - U_{i+1,j})]^2 \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i+1,j})^2 + \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})^2 \\
&+ 2 \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (U_{i,j} - U_{i,j+1})(U_{i,j} - U_{i+1,j})
\end{aligned} \tag{C.0.8}$$

Annexe D

RÉSULTATS NUMÉRIQUES SUPPLÉMENTAIRES DU CHAPITRE 2

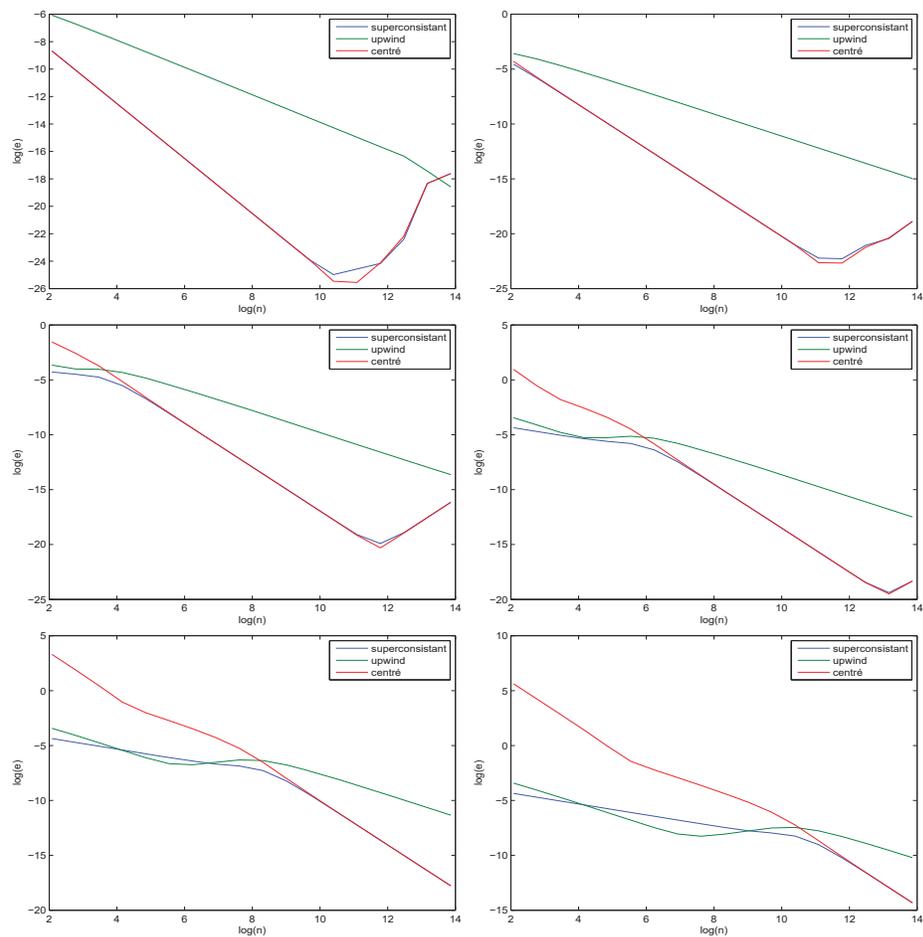


FIGURE D.1. Graphique log-log de l'erreur en fonction de n pour $\epsilon = 1, 0.1, 0.01, 0.001, 0.0001, 0.00001$ (lecture de gauche à droite puis de haut en bas). Le problème résolu est l'équation d'advection-diffusion (2.4.1) avec membre de droite $f = x$ et $u = 0$ sur la frontière. La solution exacte est donnée par (2.4.14)

Annexe E

RÉSULTATS NUMÉRIQUES SUPPLÉMENTAIRES DU CHAPITRE 3

Résolution du problème d'advection-diffusion (3.1.13) avec $f = 1$ et conditions Dirichlet homogènes à la frontière. On pose $\vec{\beta} = r((y - 1/2)\vec{i} + (1/2 - x)\vec{j})$ La méthode de De l'Isle-Owens est montrée à gauche et celle de Fatone et al. à droite

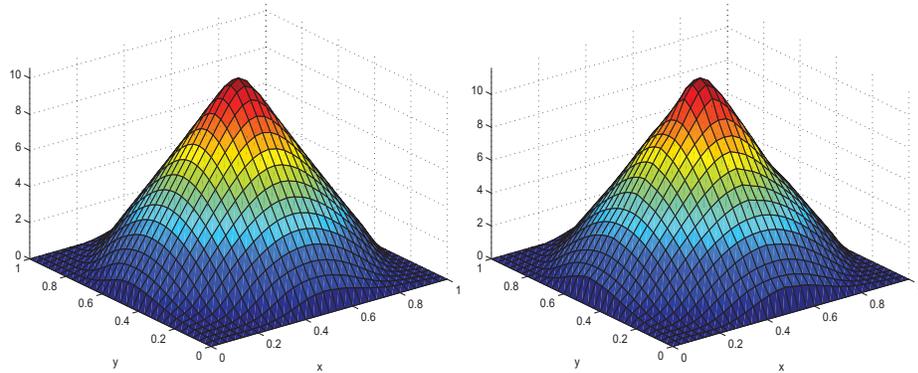


FIGURE E.1. $\epsilon = 0.001$, $r = 1$, $n_x = 30$, $n_y = 30$

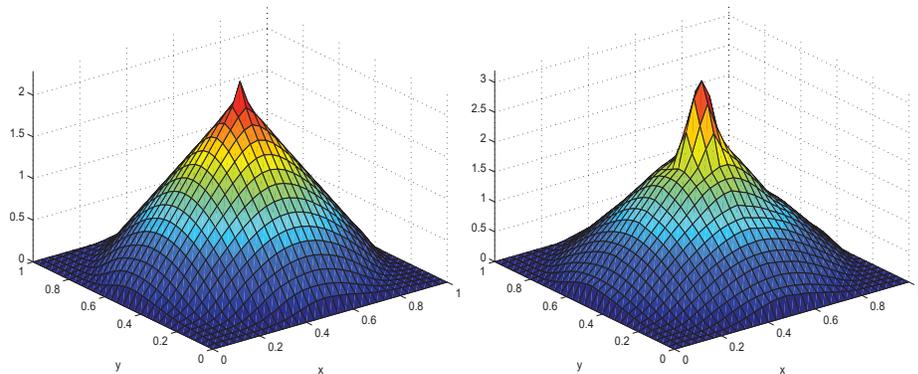
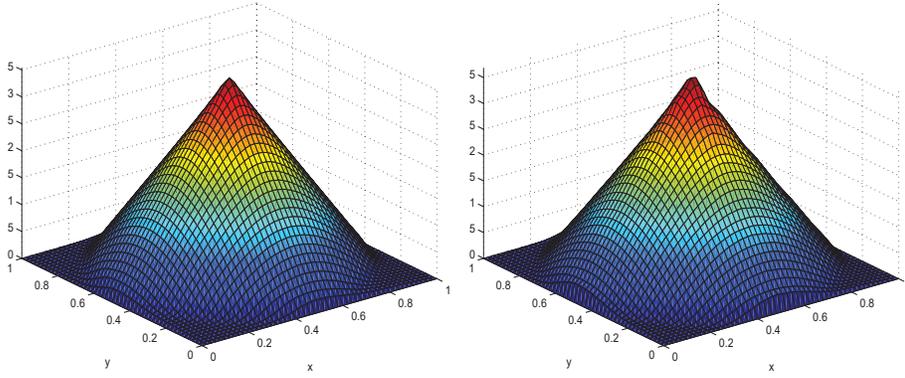
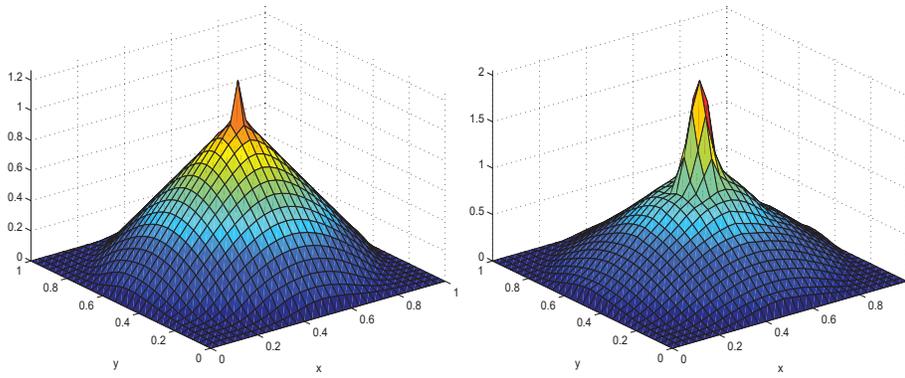
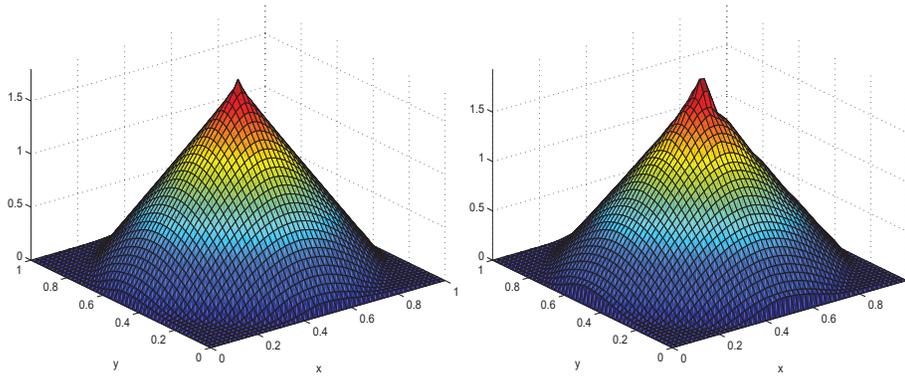


FIGURE E.2. $\epsilon = 0.001$, $r = 5$, $n_x = 30$, $n_y = 30$

FIGURE E.3. $\epsilon = 0.001$, $r = 5$, $n_x = 50$, $n_y = 50$ FIGURE E.4. $\epsilon = 0.001$, $r = 10$, $n_x = 30$, $n_y = 30$ FIGURE E.5. $\epsilon = 0.001$, $r = 10$, $n_x = 50$, $n_y = 50$

Bibliographie

- [1] R. A. Bernatz, K. D. Carlson, C. J. Chen, and W. Lin. *Finite Analytic Method in Flows and Heat Transfer*. Taylor & Francis, New York, 2000.
- [2] R.B. Bird, E.N. Lightfoot, and W.E. Stewart. *Transport Phenomena* (revised second ed.) John Wiley & Sons. *New York*, 2007.
- [3] W. E. Boyce, R.C. DiPrima, and C.W. Haines. *Elementary differential equations and boundary value problems*, volume 9. Wiley, New York, 1969.
- [4] P. Charbonneau. *PHY-3140 Hydrodynamique*, 2014.
- [5] T. C. Corke, E. Erturk, and C. Gökçöl. Numerical solutions of 2-d steady incompressible driven cavity flow at high Reynolds numbers. *Int. J. Numer. Meth. Fl.*, 48(7) :747–774, 2005.
- [6] O. Darrigol. *Worlds of Flow : A history of hydrodynamics from the Bernoullis to Prandtl*. Oxford University Press, 2005.
- [7] J.P. Demailly. *Analyse numérique et équations différentielles*. Collection Grenoble sciences. EDP Sciences, 2006.
- [8] M. Deville, M. Israeli, and S.A. Orszag. Boundary conditions for incompressible flows. *J. Sci. Comput*, pages 75–111, 03 1986.
- [9] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [10] L. Fatone, D. Funaro, and R. Giova. Finite-difference schemes for transport-dominated equations using special collocation nodes. *Numer. Meth. Part. D. E.*, 21(4) :649–671, 2005.
- [11] L.-M. Fu and R.-J. Yang. Thermal and flow analysis of a heated electronic component. *Int. J. Heat. Mass. Tran.*, 44(12) :2261 – 2275, 2001.
- [12] D. Funaro. A note on second-order finite-difference schemes on uniform meshes for advection–diffusion equations. *Numer. Meth. Part. D. E.*, 15(5) :581–588, 1999.
- [13] D. Funaro. Superconsistent discretizations. *J. Sci. Comput.*, 17(1-4) :67–79, 2002.
- [14] K.N. Ghia, U. Ghia, and C.T. Shin. High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method. *J. Comput. Phys.*, 48 :387–411, December 1982.

- [15] D.F. Griffiths, J.W. Dold, and D.J. Silvester. *Essential Partial Differential Equations : Analytical and Computational Aspects*. Springer Undergraduate Mathematics Series. Springer International Publishing, 2015.
- [16] V. John and P. Knobloch. A computational comparison of methods diminishing spurious oscillations in finite element solutions of convection-diffusion equations. *Programs and Algorithms of Numerical Mathematics*, pages 122–136, 2006.
- [17] J. B. Lorente, R. Codina, and A. Huerta. A fractional-step method for the incompressible Navier-Stokes equations related to a predictor-multicorrector algorithm. *Int. J. Numer. Meth. Fl.*, 28(10) :1391–1419, 1998.
- [18] B. Seibold. A compact and fast MATLAB code solving the incompressible Navier-Stokes equations on rectangular domains. *Massachusetts Institute of Technology*, 2008.
- [19] S. Whitaker. *Introduction to Fluid Mechanics*. Prentice-Hall, 1968.