THÈSE

présentée par

Pierre SARAMITO

Ingénieur ENSIMAG 1986

Pour obtenir le titre de DOCTEUR

DE L'INSTITUT NATIONAL POLYTECHNIQUE
DE GRENOBLE

(Arrêté Ministériel du 23 juillet 1988)

Spécialité : Mécanique

Simulation numérique d'écoulements de fluides viscoélastiques par éléments finis incompressibles et une méthode de directions alternées–applications

Date de la soutenance : 5 mars 1990

Composition du Jury :

Président :	**Jean-Michel PIAU**
Rapporteurs :	**Jaques BARANGER**
	Jean-Claude SAUT
Examinateur :	**Jean-Robert CLERMONT**
	Yves FAUTRELLE
	Andre LATROBE

Thèse préparée au sein du laboratoire **IMG-INPG**

THÈSE

présentée par

Pierre SARAMITO

Ingénieur ENSIMAG 1986

Pour obtenir le titre de DOCTEUR

DE L'INSTITUT NATIONAL POLYTECHNIQUE
DE GRENOBLE

(Arrêté Ministériel du 23 juillet 1988)

Spécialité : Mécanique

Simulation numérique d'écoulements de fluides viscoélastiques par éléments finis incompressibles et une méthode de directions alternées–applications

Résumé :
Nous considérons la simulation numérique des écoulements de fluides viscoélastiques. Développant une approximation en temps par la méthode des directions alternées, nous proposons un algorithme entièrement nouveau permettant de découpler le calcul des contraintes de celui des vitesses. D'ordre deux en temps, cette méthode permet de plus le calcul rapide des solutions stationnaires. L'élément à divergence nulle de Thomas-Raviart est utilisé pour les vitesses, et celui de Lesaint-Raviart pour les contraintes. La méthode est appliquée au problème de l'écoulement de fluides du type Oldroyd dans une contraction brusque (problème de la marche).

Mots-clés :
SIMULATION NUMÉRIQUE, ÉQUATIONS AUX DÉRIVÉES PARTIELLES, MATÉRIAUX À MÉMOIRE, FLUIDES NON-NEWTONIEN, ÉLEMENTS FINIS, MÉTHODE DE DIRECTIONS ALTENÉES, PROBLÈME DE STOKES, PROBLÈME DE TRANSPORT, PRÉCONDITIONNEMENT, ASSEMBLAGE COMPACTÉ.

Cette thèse est le fruit de trois ans de recherches appliquées et théoriques au sein du groupe de Rhéologie des Fluides Industriels du laboratoire de l'Institut de Mécanique de Grenoble. Je tiens à remercier tous ses membres de l'accueil qu'ils m'ont réservé.

Je remercie également monsieur Jean-Michel Piau de m'avoir accueilli dans son groupe, de s'être intéressé à ce travail, et de m'avoir fait l'honneur de présider ce jury.

Je tiens à exprimer toute ma gratitude à monsieur Jacques Baranger pour les fructueuses discussions que nous avons eues, pour ses nombreux et précieux conseils, encouragements, et l'aide qu'il m'a apporté à la préparation de cette thèse.

Jean-Claude Saut m'a accueilli au Laboratoire de Mathématique de la faculté d'Orsay, où j'ai trouvé un environnement scientifique de très grande qualité. Je l'en remercie très sincèrement, ainsi que de l'intérêt et du soutient qu'il a apporté à ce travail.

Je remercie Jean-Robert Clermont qui a dirigé cette thèse, pour ses conseils et sa constante disponibilité, tout au long de l'élaboration de cette thèse.

Je tiens à remercier ici monsieur André Latrobe de m'avoir fait l'honneur de participer à ce jury et de s'être intéressé à ce travail.

Je remercie Yves Fautrelle d'avoir accepté de participer à ce jury.

Ma reconnaissance va également à messieurs P. Baras, J. F. Maitre, C. Paumier et P. Witomski, pour avoir assuré ma formation à l'Analyse Numérique et à la Mécanique des Fluides, et m'avoir donné le goût à la recherche dans ces domaines. J'ai beaucoup apprécié leurs précieux conseils et encouragements ainsi que l'accueil chaleureux du laboratoire TIM^3 *de l'IMAG.*

Que ma famille et mes amis saisissent aussi, à travers ce travail, quelque peu abstrait, une expression de ma profonde affection.

Contents

Notations

$\mathbb{N}$ ensemble des nombres entiers

$\mathbb{R}$ corps des réels

We nombre de Weissenberg

Re nombre de Reynolds

α paramétre de retard du modèle d'Oldroyd, dans $[0,1]$: $\alpha = 0$ correspond au modèle newtonien et $\alpha = 1$ à celui de Maxwell

Ω domaine ouvert de $\mathbb{R}^n$

$\Gamma = \partial\Omega$ fontière du domaine Ω

pp presque partout (égalité, définition), à un ensemble de mesure nulle près

$\vec{u} = (u_1, \ldots, u_n)$ champ de vecteur, défini pp dans Ω, et à valeur dans $\mathbb{R}^n$

$\vec{\nu}$ normale unitaire sortante de Ω, définie pp sur la frontière Γ

$\vec{\tau} = (-\nu_2, \nu_1)$ tangente unitaire à la frontière Γ, en dimension $n = 2$

Γ_- partie de Γ où $\vec{u}.\vec{\nu} < 0$, c'est-à-dire la frontière amont pour l'écoulement défini par $\vec{u}$ dans Ω

$B = \vec{u}.\vec{\nu}$, $M = |B|$ opérateurs scalaires, définis pp sur la frontière à $\vec{u}$ fixé

$\tau = (\tau_{i,j})_{i,j=1,\ldots,n}$ tenseur, défini pp dans Ω

$\tau^t = (\tau_{j,i})_{i,j=1,\ldots,n}$ transposé du tenseur τ

$\mathrm{tr}\,\tau = \sum_{i=1}^{n} \tau_{i,i}$ la trace du tenseur τ

$\vec{\nabla} = (\partial_1, \dots, \partial_n)$ opérateur de dérivation

$\vec{\nabla}\,\varphi = (\partial_1\varphi, \dots, \partial_n\varphi)$ gradient d'une fonction scalaire φ

$\Delta\varphi = \sum_{i=1}^{n} \frac{\partial_i^2\varphi}{\partial_i x_i^2}$ Laplacien d'une fonction scalaire φ

$\mathrm{div}\,\vec{u} = \vec{\nabla}.\vec{u} = \sum_{i=1}^{n} \partial_i u_i$ divergence d'un champ de vecteur $\vec{u}$

$\vec{\mathrm{rot}}\,\vec{u} = \vec{\nabla} \times \vec{u}$ rotationnel d'un champ de vecteur $\vec{u}$, où $\times$ désigne ici le produit vectoriel usuel

$\Delta\,\vec{u} = (\Delta u_1, \dots, \Delta u_n)$ Laplacien vectoriel d'un champ de vecteur $\vec{u}$

$\vec{\nabla}\vec{u} = (\partial_j u_i)_{i,j=1,\dots,n}$ tenseur gradient du champ de vecteur $\vec{u}$

$D(\vec{u}) = \frac{1}{2}(\vec{\nabla}\vec{u} + \vec{\nabla}\vec{u}^t)$ tenseur des taux de déformation du champ de vecteur $\vec{u}$

$\vec{\mathrm{div}}\,\tau = (\sum_{j=1}^{n} \partial_j \tau_{ij})_{i=1,\dots,n}$ divergence vectorielle d'un tenseur τ (divergences des vecteurs-ligne)

$\overset{\square}{\tau}$ dérivé matérielle (ou objective) d'un tenseur (relation (1.3))

$\vec{\mathrm{rot}}\,\varphi = (\partial_2\varphi, \dots, \partial_1\varphi)$ rotationnel vectoriel d'une fonction scalaire, dans le cas de la dimension $n = 2$

$\mathrm{rot}\,\vec{u} = \partial_1 u_2 - \partial_2 u_1$ rotationnel scalaire d'un champ de vecteur, dans le cas de la dimension $n = 2$

$W^{m,p}(\Omega)$ espace de Sobolev d'ordre m; $p \in]0, +\infty]$ (espace de Banach)

$H^m(\Omega) = W^{m,2}(\Omega)$ espace de Sobolev d'ordre m; $p = 2$ (espace de Hilbert)

$L^2(\Omega) = H^0(\Omega)$ espace des fonctions de carré sommable (pour la mesure de Lebesgue) dans Ω

$\|.\|_{m,p,\Omega}$ norme de l'espace $W^{m,p}(\Omega)$

$\|.\|_{m,\Omega} = \|.\|_{m,2,\Omega}$ norme de l'espace $H^m(\Omega)$

$(\varphi, \phi) = \int_\Omega \varphi\phi\,dx$ produit scalaire de $L^2(\Omega)$

$L^2_o(\Omega)$ sous espace espace de $L^2(\Omega)$ des fonctions de moyenne nulle dans Ω

$H(\mathrm{div},\Omega)$ espace des champ de vecteur de $L^2(\Omega)^n$ dont la divergence est de carré sommable

$L^2(\Omega)^{n^2}_s$ espace des tenseurs symériques dont les composantes sont de carré sommable

$D(\vec{\mathrm{div}},\Omega)$ espace des tenseurs de $L^2(\Omega)^{n^2}_s$ dont la divergence est de carré sommable

$\tilde{V}$ sous-espace de $L^2(\Omega)^n$ des champ de vecteur dont la divergence est nulle

$\tilde{D}$ sous-espace de $L^2(\Omega)^{n^2}$ composé de tenseurs déviateurs (de trace nulle) dont les composantes normales sont de moyenne nulle, et la divergence appartient à $\tilde{V}$

$\mathcal{T}_h$ triangulation par éléments finis du domaine Ω

Q_h sous-espace de dimension finie de $L^2(\Omega)$

Y_h sous-espace de dimension finie de $H(\mathrm{div},\Omega)$

T_h, S_h approximations de $D(\vec{\mathrm{div}},\Omega)$

π_h opérateur d'interpolation (projection) sur un espace de dimension finie, paramétré par $h > 0$

$\mathcal{T}(N)$, $\mathcal{W}(N)$ coût d'un algorithme en temps de calcul et place mémoire, respectivement

$\square$ fin de démonstration, d'exemple ou de remarque.

Introduction

Les motivations de l'étude

Certains phénomènes *visco-élastiques* spectaculaires, observés avec des polymères fluides, ne peuvent être décrits par les équations de Navier-Stokes. Ces effets sont caractérisés, entre autres, par une dépendance des propriétés élongationnelles et des effets de mémoire associés à l'élasticité du matériau.

Ces effets peuvent être très importants lors de la mise en forme des matériaux polymères : les *instabilités* d'écoulement, par exemple, limitent le taux de production lors de nombreuses opérations telles que l'extrusion de fibres. Des instabilités sont observées à très faible nombre de Reynolds, alors qu'un écoulement newtonien correspondant, est stable.

Le comportement visco-élastique est aussi à l'origine de motifs d'écoulement complexes, tels que des zones de recirculation de grande taille dans des régions confinées (*vortex*). Ces motifs d'écoulement prêtent à conséquence lors des processus d'extrusion des fibres de verre. Par exemple, l'accumulation d'impuretés dans les vortex, extrudées plus tard, conduisent à des défauts dans le matériau, sinon à la rupture des fibres.

Enfin, l'extrusion de matériaux viscoélastiques à travers un orifice est utilisée très souvent lors de la mise en forme des matériaux, et a été étudié expérimentalement depuis de nombreuses années.

Informatique et modélisation

Dans certains cas, partant d'hypothèses cinématiques, il est possible de décrire les effets visco-élastiques à l'aide de modèles monodimensionnels. Cependant, dans ce qui suit, nous nous intéresserons à l'étude de modèles de fluides visco-élastiques applicables à des géométries pour lesquelles ces hypothèses cinématiques ne peuvent être considérées.

Il existe de nombreux *modèles* exprimant en termes mathématiques les *phénomènes rhéologiques* observés. Ces modèles font généralement intervenir des *équations aux dérivées partielles*, souvent non-linéaires. L'obtention de propriétés quantitatives des résultats nécessite alors des calculs qui ont été considérés jusqu'à ces dernières années comme irréalisables.

Cependant, les développements récents de l'informatique et de l'*analyse numérique* permettent aujourd'hui d'aborder ce travail en abandonnant certaines hypothèses simplificatrices afin d'obtenir une simulation plus réaliste du phénomène. Ainsi, la *mécanique des fluides numérique* (Computational Fluid Dynamics, en anglais) est en passe de devenir un outil de mise au point aussi important que les *essais expérimentaux* : plus souple, plus facile à réaliser, et surtout plus économique, elle permet d'aborder des expériences délicates, voire impossible à reproduire en laboratoire.

Méthodes mathématiques

La plupart des modèles visant à décrire les comportements des fluides viscoélastiques mettent en avant des questions et des problèmes *complètement ouverts* (tels que l'existence et la régularité des solutions des équations). Il est cependant possible de se servir, comme point de départ pour l'étude, de théories générales, orientées vers les applications : ce sont les fondements de l'analyse numérique, préalable obligé au traitement sur ordinateur.

Pourquoi se limiter dans ce rapport à la *méthode des éléments finis* ? Il y a à cela une bonne raison : c'est la seule technique (en 1989) qui, en fait, puisse s'appliquer sans restriction sur la forme du domaine occupé par le fluide. Dans la pratique, nous nous sommes limités aux éléments de plus bas degré, car les écoulements sont généralement *singuliers* (au sens de Kondratiev et Grisvard [27]), et il est probable que les élements de degré (fixé) plus élevé ne soient pas plus précis; or ils sont plus difficiles à mettre en œuvre et entraînent des calculs plus longs.

Une bonne compréhension des *algorithmes de calcul* est nécessaire à leur utilisation efficace. Nous connaissons le fantastique essor des calculateurs, dû à l'évolution de la technologie : en 1957 les transistors remplaçaient les lampes, dans les années 1960 apparaissaient les premiers circuits imprimés, puis les premiers circuits intégrés de quelques dizaines de transistors par puce. En 1985, la technologie VLSI (Very Large Scale Integration) permet d'intégrer un million de portes logiques par puce. Ce que l'on connait moins, c'est que le gain de performance dû aux *progrès des méthodes mathématiques* est, dans certains domaines, au moins aussi important que celui dû à la révolution technologique. Par exemple, de 1973 à 1983, les capacités de calcul des ordinateurs les plus puissants ont été multipliées par 1000. Et dans un même temps, l'amélioration de certaines techniques numériques faisait gagner un autre facteur 1000 : ceci permet de calculer un avion complet en régime transonique en moins d'une nuit de Cray1 en 1985, et aujourd'hui, l'écoulement d'un matériau fluide visco-élastique dans une contraction brusque en moins de 10 minutes sur une station de travail.

Pour cette raison, nous avons tenu à présenter avec le plus grand soin les algorithmes issus des méthodes proposées cette thèse, avec une évaluation asymptotique du *temps de calcul* et de la *place-mémoire* nécessaire, en fonction de la taille des données.

Un problème complexe

La simulation numérique des écoulements de fluides visco-élastique donne actuellement lieu à une compétition internationale de première importance. Au cours de ces dernières années, des recherches intensives ont mis en évidence de nombreuses difficultés.

La plus importante est certainement la *perte de convergence* pour les grands temps de relaxation de la “mémoire” du fluide (*grands nombres de Weissenberg*). Il a été quelque peu admis que ces difficultés sont purement numériques : nature hyperbolique des équations, conditions de frontières, incapacité des méthodes d'éléments finis habituelles à aborder les singularités, ... Cependant, une autre source de difficulté provient certainement des modèles rhéologiques qui caractérisent les fluides visco-élastiques. Bien que la plupart de ces modèles puissent décrire bon nombre de comportement de ce type, les choses sont bien moins claires au *voisinage de singularités* issues de géométries complexes.

Enfin l'énorme *quantité de calculs* nécessaires à la plupart des méthodes utilisées pour ce type de simulation a constitué un obstacle majeur : on devait résoudre de très grands systèmes couplés, et ce type de calcul mobilisait des super-calculateurs pendant de très longues durées.

Sur quelques réponses récentes

Nous renvoyons à l'ouvrage de Keunigs [38] pour un historique très détaillé de l'énorme travail accompli en la simulation numérique des écoulements de fluides visco-élastiques.

En 1987, retenons le travail de Marchal et Crochet [15], qui mettent en évidence une technique de *décentrage vers l'amont* (upwind, en anglais), introduite par Hughes [32] sur des problèmes de diffusion-convection, puis étendue à d'autres problèmes (Hughes et Brooks [9]). Cette procédure tenant alors compte de la nature hyperbolique des équations, ils obtiennent ainsi des solutions numériques pour de grands nombres de Weissenberg. Cependant, cette méthode développe une approximation continue des contraintes : afin de s'assurer que l'approximation variationnelle mixte correspondante est bien-posée [7, 18], les auteurs sont contraints à utiliser une technique de maillages sous-éléments, ce afin d'obtenir des solutions acceptables pour des nombres de Weissenberg nuls. Enfin les temps de calcul sont non-négligeables sur des calculateurs CRAY.

Un pas important semble avoir été accompli dernièrement par M. Fortin et A. Fortin [16] : développant une approximation discontinue des contraintes, ils s'assurent ainsi que le problème approché est bien-posé pour les nombres de Weissenberg nuls et utilisent la technique de décentrage amont développée par Lesaint et Raviart [42, 41] dans le cadre de l'équation scalaire de transport de neutrons. Cependant, implémentée pour des éléments de degré élevé, le schéma n'est plus *monotone* (TVD en anglais, abréviation de Total Variation Decreasing) et les solutions approchées présentent des *oscillations* lorsque la solution exacte est peu régulière (près d'un coin rentrant). Enfin, le système obtenu reste très grand, et la convergence de la méthode découplée utilisée est lente : tout ceci conduit à des temps de calcul très élevés, limitant les expérimentations numériques.

Notre contribution aux recherches

Développant la méthode *l'éléments finis incompressibles* de Thomas et Raviart [21] pour les vitesses, (la solution approchée est à divergence nulle en vitesse), nous proposons (chapitre 2) plusieurs approximations originales des contraintes, en s'assurant que le problème est bien-posé pour des nombres de Weissenberg nuls. Dans le cas des éléments de plus bas degré, nous montrons de plus que le système obtenu est de *taille minimale* (asymptotiquement un degré de liberté par grandeur scalaire et par élément).

Abordant le problème des grands nombres de Weissenberg, nous étudions au chapitre 4 quelques techniques de décentrage vers l'amont pour le problème de *transport des contraintes*. Nous retenons alors les méthodes conduisant à des opérateurs monotones, garantissant ainsi des solutions non-oscillantes. Nous rappellons certaines estimations d'erreur associées.

Partant de la méthode originale des *directions alternées* dûe à Peaceman et Rachford [48], et s'inspirant des travaux de Glowinski, Mantel et Périaux [23] sur les équations de Navier-Stokes, nous présentons au chapitre 3 un nouvel algorithme permettant de découpler le calcul des vitesses de celui des contraintes. Inconditionnellement stable et d'ordre 2 en temps, cet algorithme permet en outre le calcul rapide des solutions stationnaires. Utilisant une méthode de *relaxations symétrisées par bloc* pour le calcul des contraintes, nous montrons que les coûts asymptotiques en temps de calculs et place mémoire sont liés directement à ceux du "solveur de Stokes" utilisé. Ce coût est alors optimal dans le contexte des *méthodes multigrilles* (voir [1], par exemple), et reste quasi-optimal pour la *méthode de gradient conjugé préconditionné* que nous avons implémentée.

Quel que soit le développement des ordinateurs, cet avantage d'optimalité est important. En effet, les besoins en simulation numérique sont quasi-illimités, et même lorsque la résolution des écoulements tridimensionnels de fluides visco-élastiques en géométrie à contraction brusque sera chose courante, on essaiera, par exemple, de modéliser l'influence des singularités de coin sur le gonflement de la surface libre située en aval.

Cependant, pour cette technique numérique, beaucoup de questions se posent : existe-il une solution pour les grands nombres de Weissenberg, est-elle unique, quelle est la nature des singularités de coin, quelles sont les conditions aux limites conduisant à un problème bien-posé, quelle peut être la justification de l'approximation du problème (en terme d'estimations d'erreur), peut-on obtenir de meilleures approximations des singularités ... Nous n'avons pas répondu à toutes ces questions, mais nous avons progressé, et les résultats obtenus, ainsi qu'une description détaillé des méthodes employées, sont rassemblés dans cette thèse.

Les expérimentations numériques décrites au chapitre 5 sont réalisées en géométrie à contraction 4 à 1, plane et axisymétrique. Les modèles du type Oldroyd nous permettent de comparer nos résultats avec les auteurs précités. Nous ne discuterons pas de sa justesse, comparé à d'autres modèles rhéologiques, car le problème associé comporte la plupart des difficultés de ce type de simulation numérique.

Enfin, l'annexe A rappelle les principaux résultats de la théorie de l'*approximation mixte abstraite*. Ces résultats sont utilisés au chapitre 2 lors de la présentation des méthodes d'élements finis. L'annexe B décrit l'implémentation de la méthode du gradient conjugé préconditionné, appliquée aux sous-problèmes de type Stokes évoqués au chapitre 4. L'annexe C présente la *structure de données* choisie lors de l'élaboration d'un logiciel convivial d'éléments finis. Ce logiciel a permis d'obtenir les résultats du chapitre 5.

Chapitre 1

Présentation du problème

La présentation de cette partie est purement formelle : les grandeurs tensorielles, vectorielles et scalaires intervenant dans les relations sont supposées suffisament régulières. Ce n'est que dans le chapitre suivant que nous introduirons les espaces fonctionnels appropriés.

1.1 Lois de comportement et de conservation

1.1.1 Les lois de comportement de type différentiel

Exprimons tout d'abord le tenseur des contraintes totales de Cauchy :

$$\sigma = -p.I + 2\eta_n.D(\vec{u}) + \tau \tag{1.1}$$

où p est le champ de pression hydrostatique, $\vec{u}$ le champ de vitesse, $D(\vec{u}) = \frac{1}{2}(\vec{\nabla}\vec{u} + \vec{\nabla}\vec{u}^t)$ le tenseur des taux de déformation, $\eta_n > 0$ la viscosité newtonienne (constante) et τ, le tenseur des contraintes supplémentaires.

Le tenseur τ satisfait à une équation de comportement. Nous considérons ici les lois différentielles du type Oldroyd [47], de la forme :

$$\lambda \overset{\square}{\tau} + \tau = 2\eta_v.D(\vec{u}) \tag{1.2}$$

où η_v est une seconde viscosité, $\lambda > 0$ le temps de relaxation et $(\overset{\square}{\cdot})$ désigne la dérivée objective d'un tenseur symétrique [47, 3] :

$$\overset{\square}{\tau} = \frac{\partial \tau}{\partial t} + (\vec{u}.\vec{\nabla})\tau + \tau.W(\vec{u}) - W(\vec{u}).\tau - a.(D(\vec{u}).\tau + \tau.D(\vec{u})) \tag{1.3}$$

avec $a \in [-1, 1]$ et $W(\vec{u}) = \frac{1}{2}(\vec{\nabla}\vec{u} - \vec{\nabla}\vec{u}^t)$ est le tenseur vorticité.

Remarque 1.1 Valeurs particulières des paramètres

Pour $a = 0$, nous obtenons en (1.3) la dérivée de Jaumann, pour $a = 1$, la dérivée convectée supérieure et pour $a = -1$ la dérivée convectée inférieure. Les comportements généralement adoptés correspondent à $a \in [0.8, 1]$

Pour $\lambda = 0$, nous retrouvons un fluide newtonien, de viscosité :

$$\eta = \eta_n + \eta_v \tag{1.4}$$

Enfin, pour $\eta_n = 0$, on retrouve le modèle de Maxwell. □

Remarque 1.2 Temps de retard des fluides visco-élastiques

En introduisant le tenseur

$$\tau_e = 2\eta_n.D(\vec{u}) + \tau$$

et à l'aide de la loi de comportement (1.2), on trouve que le tenseur τ_e satisfait à l'équation :

$$\tau_e + \lambda. \overset{\square}{\tau_e} = 2\eta(D(\vec{u}) + \lambda_r \overset{\square}{D}(\vec{u})) \tag{1.5}$$

où $\lambda_r = \frac{\lambda \eta_n}{\eta}$ est appelé *le temps de retard* du fluide, et où η a été introduit en (1.4). □

1.1.2 Lois de conservation

L'équation de la conservation de la quantité de mouvement s'écrit :

$$\rho(\frac{\partial \vec{u}}{\partial t} + (\vec{u}.\vec{\nabla})\vec{u}) - \vec{\mathrm{div}}\,\tau - \eta_n \Delta \vec{u} + \vec{\nabla} p = 0 \tag{1.6}$$

où $\rho > 0$ est la densité (supposée constante)

Dans la suite, le fluide sera supposé *lent* : nous ne considérerons pas le terme d'inertie $(\vec{u}.\vec{\nabla})\vec{u}$.

La conservation de la masse conduit à la relation d'incompressibilité :

$$\mathrm{div}\,\vec{u} = 0 \tag{1.7}$$

Enfin, la conservation des moments s'écrit $\sigma = \sigma^t$, d'où, de (1.1) : $\tau = \tau^t$.

1.1.3 Extension des modèles de type différentiels

Signalons qu'il est possible d'étendre ce modèle à d'autres modèles de type différentiel : Le modèle de Gisekus est obtenu en ajoutant le terme

$$\frac{\xi\lambda}{\eta_v}\tau$$

au terme de gauche de (1.2), où $\xi > 0$ est un paramètre constant. Pour $\xi = \frac{1}{2}$, on retrouve le modèle de Leonov.

Le modèle de Phan Thien et Tanner est obtenu en remplaçant l'équation (1.2) par :

$$\lambda \overset{\square}{\tau} + e^{\frac{\xi\lambda}{\eta_v}\mathrm{tr}(\tau)}\tau - 2\eta_v D(\vec{u}) = 0$$

où $\xi > 0$.

Le modèle de White-Metzner est obtenu en faisant varier les paramètres η_v, η_n et λ en fonction du second invariant du tenseur taux de déformation : $II_D = \mathrm{tr}(D(\vec{u})^2)$, si $\mathrm{tr}(\gamma)$ d'esigne la trace d'un tenseur γ. Les lois les plus connues sont de la forme :

la loi puissance :

$$\eta(z) = \eta_0 z^{\frac{r-2}{2}} \tag{1.8}$$

la loi de Carreau :

$$\eta(z) = \eta_\infty + (\eta_0 - \eta_\infty)(1 + k.z)^{\frac{r-2}{2}} \tag{1.9}$$

où $\eta_0 > \eta_\infty > 0$, $k > 0$ et$1 < r \leq 2$. Pour $\eta_v = \eta(II_D)$, $\eta_n = \lambda = 0$, nous retrouvons les fluides quasi-newtoniens (loi de puissance) et pseudo-plastiques (loi de Carreau), dont la résolution numérique par éléments finis est actuellement bien maîtrisé [5]. Dans le cas de certains mélanges de polymères, η_∞ peut être très petit. Avec $r = 2$, nous retouvons les fluides newtoniens.

Une autre classe de modèles est décrite par des lois de type intégral. Pour plus de détails, nous renvoyons à [38] et toutes les réferences qui y sont citées.

1.2 Conditions aux limites

Finalement, le système des équations doit être complété par un ensemble de conditions initiales et de frontières, dépendantes de la géométrie considérée.

Il faut mentionner que, pour la plupart des problèmes, les conditions de frontières ne sont pas évidentes, et des précautions doivent être prises pour assurer la fiabilité de la solution numérique.

Ainsi, de nombreux auteurs sont convaincus que la condition habituelle de *non-glissement* appliquée aux parois solides est une source de difficultés numériques. Cependant, la façon d'imposer une condition de glissement au voisinage des singularités (et conduisant à un problème bien-posé) reste inconnue.

Lorsque $a = 1$ et $\eta_n > 0$, le système (1.2)-(1.6)-(1.7) est de type mixte elliptique-hyperbolique [36, page 243]. Les caractéristiques réelles sont les lignes de courant, et les grandeurs transportées sur les caractéristiques sont les composantes du tenseur de contraintes τ.

Si nous considérons un domaine Ω borné simplement connexe de $\mathbb{R}^2$ ou $\mathbb{R}^3$, les conditions aux limites requises sont alors :

- une condition de type Dirichlet pour la vitesse :

$$\vec{u} = \vec{u}_\Gamma \text{ sur } \Gamma = \partial\Omega \tag{1.10}$$

- une condition sur la frontière amont pour les contraintes :

$$\tau = \tau_\Gamma \text{ sur } \Gamma_- \stackrel{\text{déf}}{=} \{x \in \Gamma, \vec{u}.\vec{\nu} < 0\} \tag{1.11}$$

  où $\vec{\nu}$ est la normale unitaire sortante à Ω sur Γ.

Les conditions à l'instant initial $t = 0$ s'écrivent :

$$\vec{u} = \vec{u}_o, \;\; \tau = \tau_o \text{ dans } \Omega \tag{1.12}$$

D'après la relation d'incompressibilité (1.7), la donnée $\vec{u}_\Gamma$ doit satisfaire une relation de compatibilité :

$$\int_\Gamma \vec{u}_\Gamma.\vec{\nu}\, \mathrm{d}\Gamma = 0 \tag{1.13}$$

où $\mathrm{d}\Gamma$ désigne une mesure superficielle sur la frontière Γ (supposée assez régulière).

1.3 Grandeurs sans dimension

Afin d'obtenir un système vérifié par des variables adimensionnelles, nous définissons le nombre de Weissenberg :

$$We = \lambda \frac{U}{L} \tag{1.14}$$

le nombre de Reynolds :

$$Re = \rho \frac{U.L}{\eta} \tag{1.15}$$

et un paramètre de retard [29] :

$$\alpha = \frac{\eta_v}{\eta} \tag{1.16}$$

où U et L sont respectivement une vitesse et une longueur caractéristique de l'écoulement. Le problème se ramène à chercher les grandeurs sans dimension, encore notées τ, $\vec{u}$ et p, définies dans Ω et satisfaisant à :

$$We\,(\frac{\partial \tau}{\partial t} + (\vec{u}.\vec{\nabla})\tau + \beta_a(\tau, \vec{\nabla}\vec{u})) \quad +\tau - 2\alpha D(\vec{u}) \quad = 0 \tag{1.17}$$

$$Re\,(\frac{\partial u}{\partial t} + (\vec{u}.\vec{\nabla})\vec{u}) \quad -\vec{\mathrm{div}}\tau - (1-\alpha)\Delta\vec{u} + \vec{\nabla}p \quad = 0 \tag{1.18}$$

$$\mathrm{div}\vec{u} \quad = 0 \tag{1.19}$$

avec les conditions aux limites et initiales du type (1.10)-(1.12) , et où la fonctionnelle bilinéaire β_a est définie par :

$$\beta_a(\tau, \vec{\nabla}\vec{u}) = \tau.W(\vec{u}) - W(\vec{u}).\tau - a.(D(\vec{u}).\tau + \tau.D(\vec{u})) \tag{1.20}$$

1.4 Résultats d'existence et d'unicité

Il existe différents théorèmes d'existence de solutions de (1.17)-(1.19), (1.10)-(1.12) suivant les hypothèses de régularité faites sur la frontière Γ, les données $\vec{u}_\Gamma$, τ_Γ, $\vec{u}_o$, τ_o, ou les paramètres We, α.

M. Renardy [52] obtient, par une méthode de point fixe, l'existence de solutions stationnaires pour toute valeur de α, avec une condition de Dirichlet homogène sur $\vec{u}$: $\vec{u} = 0$ sur Γ et des données petites.

C. Guilloppé et J. C. Saut [29] obtiennent un résultat global d'existence de solutions du problème non-stationnaire en supposant α petit (cas de solutions de polymère très diluées ou de solvant très visqueux, par exemple).

1.5 Problèmes ouverts

Dans le cas des équations de Navier-Stokes, l'existence d'écoulements stationnaires est connue sous des conditions moins restrictives que celles qui précèdent [21] : il n'est pas nécessaire de supposer les données petites ou la frontière régulière. Nous présenterons un certain nombre de raisons [53] qui rendent de semblables résultats plus difficiles, sinon impossibles à établir pour les fluides visco-élastiques.

1.5.1 Changement de type des équations

Le système des équations décrivant l'écoulement d'un fluide visco-élastique peut présenter un *changement de type* [36]. Un changement de type apparaît si la vitesse de propagation des ondes de cisaillement devient soit complexe, soit plus petite que la vitesse du fluide.

Le premier cas est associé aux instabilités de Hadamard. Dans le second cas, un écoulement stationnaire peut très bien exister. Cependant, la nature des conditions aux bords conduisant à un problème *bien-posé* est alors certainement différente du cas sous-critique.

1.5.2 Perte d'existence ou d'unicité de la solution

Certains modèles, tels que le fluide de Maxwell convecté supérieur (UCM) peuvent conduire, dans certains cas, à une perte d'existence ou d'unicité de la solution (voir [28], et le paragraphe 5.1.3).

1.5.3 Existence d'un point limite

L'existence d'un point limite pourrait très bien être la cause des difficultés rencontrées pour les grands nombres de Weissenberg ("the hight Weissenberg number problem" [38]).

Des expériences numériques postérieures [15] tendent à montrer qu'il s'agirait, dans la plupart des cas, d'effets numériques dûs aux méthodes employées, bien que ce problème apparaisse dans certains écoulements.

L'existence d'un point limite ne signifie pas nécessairement que les écoulements stationnaires cessent d'exister, mais seulement qu'ils ne peuvent plus être obtenus par des méthodes de cheminement [56].

1.5.4 Singularité de coin

De nombreux problèmes appliqués font intervenir des *domaines à coin*. Dans le cas newtonien, ceci ne présente pas de difficultés pour montrer l'existence d'une solution; en effet, l'existence d'écoulements stationnaires peut être établie sans analyser la nature de la *singularité* [27]. Il en est ainsi car il est possible de travailler à l'aide de normes faibles, et d'obtenir l'existence de champs de vitesse dans H^1.

Pour les fluides viscoélastiques, les contraintes sont obtenues à partir des vitesses par intégration le long des lignes de courant, et par conséquent l'existence et l'unicité des lignes de courant est essentielle pour ce type de résultats. Il semble ainsi qu'il ne soit pas possible d'éviter l'analyse des singularités de coin, même si on ne souhaite qu'un résultat d'existence. La nature de la singularité de coin est inconnue jusqu'à présent. Il est clair qu'elle dépend de l'équation constitutive.

Chapitre 2

Eléments finis incompressibles

L'objet de ce chapitre est de construire une classe de méthodes d'éléments finis. Le point de départ est l'élément fini incompressible de Thomas et Raviart [21] pour l'approximation des grandeurs vectorielles. Nous construisons alors plusieurs éléments pour l'approximation du tenseur des contraintes, et la *compatibilité* des approximations entre tenseur et vecteur est établie (autrement dit, le problème est bien posé pour des nombres de Weissenberg nuls).

2.1 Les motivations de ce choix

Le problème d'Oldroyd stationnaire ($\frac{\partial}{\partial t} = 0$) se présente comme une *perturbation singulière* [53] d'un problème de Stokes à trois champs. Ce dernier, obtenu pour $We = 0$, s'énonce comme suit :
chercher τ, $\vec{u}$, p, *définis dans un domaine* Ω *de* $\mathbb{R}^n$, *et vérifiant* :

$$\tau \quad -2\alpha D(\vec{u}) = 0 \text{ dans } \Omega \tag{2.1}$$
$$\vec{\mathrm{div}}\,\tau \quad +(1-\alpha)\Delta\vec{u} - \vec{\nabla}p = -\vec{f} \text{ dans } \Omega \tag{2.2}$$
$$\mathrm{div}\,\vec{u} = 0 \text{ dans } \Omega \tag{2.3}$$
$$\vec{u} = \vec{g} \text{ sur } \Gamma \tag{2.4}$$

où α est un paramètre susceptible de décrire l'intervalle $[0,1]$, $\vec{f}$ est donnée dans Ω, et $\vec{g}$ sur Γ.

Bien que linéaires, ces équations demandent une attention spéciale, essentiellement à cause de la contrainte d'incompressibilité (2.3).

Les méthodes classiques d'approximation du champ de vecteur $\vec{u}$ nécéssitent l'utilisation d'éléments sophistiqués, et la relation d'incompressibilité (2.3) n'est alors elle-même qu'approximativement satisfaite (champ de vecteur à divergence petite, mais non-nulle). D'autre part, il paraît essentiel de satisfaire *exactement* l'équation (2.3) pour aborder l'approximation des termes de type transport introduits par les modèles visco-élastiques.

Nous verrons qu'il est possible de construire *simplement* des éléments finis *incompressibles* [21] à condition de relaxer la continuité des vitesses tangentielles aux interfaces des élements, et d'introduire le vecteur *tourbillon* :

$$\vec{w} = (1-\alpha)\,\vec{\mathrm{rot}}\,\vec{u} \tag{2.5}$$

Remarquant que :

$$\vec{\mathrm{rot}}\,\vec{\mathrm{rot}}\,\vec{u} = -\Delta\vec{u} + \vec{\nabla}\,\mathrm{div}\,\vec{u} \tag{2.6}$$

le problème obtenu avec $\alpha = 0$ s'écrit encore :
chercher $\vec{u}$, $\vec{w}$, p, définis dans Ω, et vérifiant :

$$\begin{aligned}
\vec{w} - \vec{\mathrm{rot}}\,\vec{u} &= 0 \text{ dans } \Omega && (2.7)\\
-\vec{\mathrm{rot}}\,\vec{w} - \vec{\nabla}p &= -\vec{f} \text{ dans } \Omega && (2.8)\\
\mathrm{div}\,\vec{u} &= 0 \text{ dans } \Omega && (2.9)\\
\vec{u} &= \vec{g} \text{ sur } \Gamma && (2.10)
\end{aligned}$$

Nous saurons approcher (2.7)-(2.10) de façon standard en utilisant les éléments finis incompressibles de Thomas-Raviart [21].

Pour $\alpha = 1$, nous retrouvons un problème de Stokes à trois champs, qui s'énonce : *chercher $\vec{u}$, τ, p, définis dans Ω, et vérifiant* :

$$\begin{aligned}
\tau - 2D(\vec{u}) &= 0 \text{ dans } \Omega && (2.11)\\
\vec{\mathrm{div}}\,\tau - \vec{\nabla}p &= -\vec{f} \text{ dans } \Omega && (2.12)\\
\mathrm{div}\,\vec{u} &= 0 \text{ dans } \Omega && (2.13)\\
\vec{u} &= \vec{g} \text{ sur } \Gamma && (2.14)
\end{aligned}$$

et nous proposerons plusieurs éléments permettant de construire une approximation variationnelle de (2.11)-(2.14).

Il sera alors facile de déduire des deux méthodes précédentes une approximation du problème à quatre champs :

chercher τ,w,$\vec{u}$ et p, définis dans Ω, et vérifiant :

$$\begin{aligned}
\tau - 2\alpha D(\vec{u}) &= 0 \text{ dans } \Omega && (2.15)\\
\vec{w} - (1-\alpha)\,\vec{\mathrm{rot}}\,\vec{u} &= 0 \text{ dans } \Omega && (2.16)\\
-\vec{\mathrm{rot}}\vec{w} + \vec{\mathrm{div}}\,\tau - \vec{\nabla} p &= -\vec{f} \text{ dans } \Omega && (2.17)\\
\mathrm{div}\,\vec{u} &= 0 \text{ dans } \Omega && (2.18)\\
\vec{u} &= \vec{g} \text{ sur } \Gamma && (2.19)
\end{aligned}$$

Ce dernier problème n'est qu'une autre expression de (2.1)-(2.4).

Pour cela, nous commençerons par introduire certains espaces fonctionnels particulièrement adaptés à l'étude des écoulements incompressibles, ou de l'élasticité incompressible [55]. Les problèmes (2.11)-(2.14) et (2.7)-(2.10) présentant de nombreuses similitudes, nous en construirons des formulations variationnelles mixtes en suivant une méthode proposée dans [21]. présentant de nombreuses similitudes. Afin de s'assurer que les problèmes approchés restent bien-posés, nous énoncerons des conditions suffisantes de compatibilité entre les différents espaces de dimension finie, et proposerons concrètement plusieurs choix possibles pour ces espaces. Nous vérifierons ensuite qu'il est possible d'éliminer le tourbillon des équations en considérant l'aspect matriciel du problème. Enfin, il sera alors intéréssant de remarquer que la méthode proposée apparaît, dans une certaine mesure, comme une généralisation en élément finis d'un schéma aux différences finies “Marker and Cell” (M.A.C., ou schéma aux grilles décalées).

2.2 Le cadre fonctionnel

Les résultats classiques de ce paragraphe sont présentés sans les démonstrations. Le lecteur pourra consulter les références [6, 21], par exemple, pour plus de détails.

2.2.1 Les espaces fonctionnels usuels

On note $\mathcal{D}(\Omega)$ l'espace des fonctions de classe C^∞ à support compact sur Ω, et $\mathcal{D}'(\Omega)$ son dual, appellé espace des distributions. On note $< .,. >$ le produit de dualité entre $\mathcal{D}'(\Omega)$ et $\mathcal{D}(\Omega)$.

On note $\mathcal{D}(\overline{\Omega})$ l'espace défini par :

$$\mathcal{D}(\overline{\Omega}) = \{\varphi_{|\Omega};\ \varphi \in \mathcal{D}(\mathbb{R}^n)\} \tag{2.20}$$

Une fonction $f \in L^1_{loc}(\Omega)$ (localement intégrable au sens de Lebesgue) sera identifiée à une distribution par la relation :

$$< f, \varphi >= \int_\Omega f(x)\, \varphi(x)\, dx, \ \forall \varphi \in \mathcal{D}(\Omega) \tag{2.21}$$

On définit la dérivée au sens des distributions, et on note ∂^α :

$$< \partial^\alpha u, \varphi >= (-1)^{|\alpha|} < u, \partial^\alpha \varphi >, \ \forall \varphi \in \mathcal{D}(\Omega) \tag{2.22}$$

où $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ est un multi-indice, et $|\alpha| = \sum_{i=1}^n \alpha_i$.

Pour $u \in C^{|\alpha|}(\Omega)$, cette définition coïncide avec la définition classique de la dérivation :

$$\partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n} \tag{2.23}$$

Pour tout $m \in \mathbb{N}$, $p \in [1, +\infty]$, on définit par :

$$W^{m,p}(\Omega) = \{ v \in L^p(\Omega);\ \partial^\alpha v \in L^p(\Omega),\ \forall |\alpha| \le m \} \tag{2.24}$$

l'espace de Sobolev d'ordre m construit sur $L^p(\Omega)$. L'espace $W^{m,p}(\Omega)$ est un espace de Banach pour la norme :

$$\|u\|_{m,p,\Omega} = \begin{cases} \left(\sum_{|\alpha| \le m} \int_\Omega |\partial^\alpha u(x)|^p \, dx \right)^{\frac{1}{p}} & \text{si } p < +\infty \\ \sup_{|\alpha| \le m} \left(\text{supess}_{x \in \Omega} |\partial^\alpha u(x)| \right) & \text{si } p = +\infty \end{cases} \tag{2.25}$$

On utilise aussi la semi-norme :

$$|u|_{m,p,\Omega} = \begin{cases} \left(\sum_{|\alpha| = m} \int_\Omega |\partial^\alpha u(x)|^p \, dx \right)^{\frac{1}{p}} & \text{si } p < +\infty \\ \sup_{|\alpha| = m} \left(\text{supess}_{x \in \Omega} |\partial^\alpha u(x)| \right) & \text{si } p = +\infty \end{cases} \tag{2.26}$$

On note alors :

$$H^m(\Omega) = W^{m,2}(\Omega) \tag{2.27}$$

l'espace de Sobolev d'ordre m construit sur $L^2(\Omega)$. Muni du produit scalaire :

$$(u, v)_{m,\Omega} = \sum_{|\alpha| \le m} \int_\Omega \partial^\alpha u(x)\, \partial^\alpha v(x)\, dx \tag{2.28}$$

associé à la norme $\|.\|_{m,\Omega} = \|.\|_{m,2,\Omega}$, cet espace est un espace de Hilbert. Pour $m = 0$, on retrouve :

$$L^2(\Omega) = H^0(\Omega)$$

et on notera fréquemment :

$$(.,.) = (.,.)_{0,\Omega}$$

le produit scalaire de $L^2(\Omega)$.

Puisque $\mathcal{D}(\Omega) \subset H^m(\Omega)$, on peut définir la fermeture de $\mathcal{D}(\Omega)$ pour la norme $\|.\|_{m,\Omega}$:

$$H_0^m(\Omega) = \overline{\mathcal{D}(\Omega)}^{H^m(\Omega)} \tag{2.29}$$

On introduit alors $H^{-m}(\Omega)$, le dual de $H_0^m(\Omega)$, que l'on munit de la norme duale :

$$\|\Omega\|_{-m,\Omega} = \sup_{v \in H_0^m(\Omega), v \neq 0} \frac{|< f, v >|}{\|v\|_{m,\Omega}} \tag{2.30}$$

L'espace $H^{-m}(\Omega)$ est caractérisé comme étant l'espace des dérivées jusqu'à l'ordre m des éléments de $L^2(\Omega)$.

Enfin, on a le résultat très utile :

Théorème 2.1 *Inégalité de Friedrichs-Poincarré.*

Si Ω est simplement connexe et borné dans au moins une direction, alors la semi-norme $|.|_{m,\Omega}$ induit sur $H_0^m(\Omega)$ une norme, équivalente à $\|.\|_{m,\Omega}$.

En effet, ceci peut être décrit par l'inégalité :

$$\|v\|_{m,\Omega} \leq c_m |v|_{m,\Omega}, \ \forall v \in H_0^m(\Omega) \tag{2.31}$$

où c_m est une constante strictement positive.

2.2.2 Problèmes de trace et formules de Green

Problèmes de trace

Afin d'étudier plus précisément les traces de nos fonctions sur la frontière Γ, nous supposerons Γ bornée et de classe C^1 par morceaux. C'est-à-dire que Γ peut être paramétrée par une fonction σ de classe C^1.

Soit $d\sigma$ une mesure superficielle sur Γ, et $L^2(\Gamma)$ l'espace des fonctions de carré sommable pour cette mesure, muni de la norme :

$$\|v\|_{0,\Gamma} = \left(\int_\Gamma v(\sigma)^2 \, d\sigma \right)^{\frac{1}{2}} \tag{2.32}$$

On montre que l'application trace :

$$\gamma_o : \varphi \in \mathcal{D}(\overline{\Omega}) \longmapsto \gamma_o \varphi = \varphi_{|\Gamma}$$

peut être étendue par continuité en une application, encore notée γ_o, de $H^1(\Omega)$ dans $L^2(\Gamma)$.

On en déduit une caractérisation de $H^1_0(\Omega)$:

$$H^1_0(\Omega) = \operatorname{Ker} \gamma_o \tag{2.33}$$

et on définit l'image de γ_o dans $L^2(\Gamma)$:

$$H^{\frac{1}{2}}(\Gamma) \stackrel{\text{déf}}{=} \operatorname{Im} \gamma_o \tag{2.34}$$

qui est un sous-espace propre et dense de $L^2(\Gamma)$. De plus, l'espace $H^{\frac{1}{2}}(\Gamma)$ est un espace de Hilbert pour la norme définie par :

$$\|\mu\|_{\frac{1}{2},\Gamma} = \inf_{v \in H^1(\Omega), \gamma_o v = \mu} \|v\|_{1,\Omega} \tag{2.35}$$

et on note $H^{-\frac{1}{2}}(\Gamma)$ son dual, muni de la norme duale :

$$\|\lambda\|_{-\frac{1}{2},\Gamma} = \sup_{\mu \in H^{\frac{1}{2}}(\Gamma), \mu \neq 0} \frac{| < \lambda, \mu > |}{\|\mu\|_{\frac{1}{2},\Gamma}} \tag{2.36}$$

où $< .,. >$ désigne ici le produit de dualite entre $H^{-\frac{1}{2}}(\Gamma)$ et $H^{\frac{1}{2}}(\Gamma)$. On remarque que $< .,. >$ est une extension du produit scalaire de $L^2(\Gamma)$ dans le sens où, si $\lambda \in L^2(\Gamma)$, alors :

$$< \lambda, \mu > = \int_\Gamma \lambda(\sigma)\, \mu(\sigma)\, d\sigma \tag{2.37}$$

Formules de Green

Soit $\vec{\nu} = (\nu_1, \ldots, \nu_n)$ le vecteur de la normale unitaire sortante à Ω, qui existe presque partout sur Γ, puisque la frontière est supposée de classe C^1 par morceaux. Si $v \in H^2(\Omega)$, on peut définir la dérivée normale :

$$\frac{\partial v}{\partial \nu} = \sum_{i=1}^{n} \nu_i \gamma_o \left(\frac{\partial v}{\partial x_i} \right) \tag{2.38}$$

On montre, là encore, que l'application :

$$v \longmapsto \frac{\partial v}{\partial \nu}$$

est continue de $H^2(\Omega)$ dans $H^{\frac{1}{2}}(\Gamma)$.

Nous avons les formules de Green :

$$\int_\Omega u \frac{\partial v}{\partial x_i}\, dx = -\int_\Omega \frac{\partial u}{\partial x} v\, dx + \int_\Gamma u v \nu_i\, d\sigma \ , \forall i,\ 1 \leq i \leq n \tag{2.39}$$

ceci pour tout $u, v \in H^1(\Omega)$, et

$$(\vec{\nabla} u, \vec{\nabla} v) = (-\Delta u, v) + \int_\Gamma \frac{\partial u}{\partial \nu} v\, d\sigma \tag{2.40}$$

ceci pour tout $u \in H^2(\Omega),\ v \in H^1(\Omega)$.

2.2.3 L'espace $H(\mathrm{div}, \Omega)$

Cet espace est défini comme étant le domaine dans $L^2(\Omega)$ de l'opérateur divergence, soit :

$$H(\mathrm{div}, \Omega) = \{\vec{v} \in L^2(\Omega)^n;\ \mathrm{div}\, \vec{v} \in L^2(\Omega)\} \tag{2.41}$$

Muni de la norme :

$$\|\vec{v}\|_{H(\mathrm{div},\Omega)} = \left(\|\vec{v}\|^2_{0,\Omega} + \|\mathrm{div}\, \vec{v}\|^2_{0,\Omega}\right)^{\frac{1}{2}} \tag{2.42}$$

cet espace est un espace de Hilbert.

On montre que $\mathcal{D}(\overline{\Omega})$ est dense dans $H(\mathrm{div}, \Omega)$, et on peut ainsi étendre par continuité l'opérateur de trace normale à la frontière Γ :

$$\gamma_\nu : \vec{v} \in \mathcal{D}(\overline{\Omega})^n \longmapsto \gamma_\nu \vec{v} = (\vec{v}.\vec{\nu})_{|\Gamma}$$

en une application continue, encore notée γ_ν, de $H(\mathrm{div}, \Omega)$ dans $H^{-\frac{1}{2}}(\Gamma)$.

Par extension, $\gamma_\nu \vec{v}$ est appellée la composante normale de $\vec{v}$ sur Γ, et est souvent notée $(\vec{v}.\vec{\nu})_{|\Gamma}$, même lorsque $\vec{v}$ décrit $H(\mathrm{div}, \Omega)$.

La fermeture de $\mathcal{D}(\Omega)^n$ pour la norme $\|.\|_{H(\mathrm{div},\Omega)}$, notée :

$$H_0(\mathrm{div}, \Omega) = \overline{\mathcal{D}(\Omega)}^{H(\mathrm{div},\Omega)} \tag{2.43}$$

est caractérisée par :

$$H_0(\mathrm{div}, \Omega) = \mathrm{Ker}\, \gamma_\nu \tag{2.44}$$

Enfin, nous avons la formule de Green :

$$(\vec{v}, \vec{\nabla}\varphi) + (\mathrm{div}\, \vec{v}, \varphi) = < \gamma_\nu \vec{v}, \gamma_o \varphi > \tag{2.45}$$

ceci pour tout $\vec{v} \in H(\mathrm{div}, \Omega)$, $\varphi \in H^1(\Omega)$.

2.2.4 L'espace $H(\mathrm{rot}, \Omega)$

Rappelons que, dans le cas de la dimension $n = 2$, on a :

$$\vec{\mathrm{rot}}\,\varphi = (\partial_2\varphi, -\partial_1\varphi)\ ,\forall\varphi \in \mathcal{D}'(\Omega) \tag{2.46}$$
$$\mathrm{rot}\,\vec{v} = \partial_1 v_2 - \partial_2 v_1\ ,\forall\vec{v} \in \mathcal{D}'(\Omega)^2 \tag{2.47}$$

En dimension $n = 2$, l'identité (2.6) s'écrit :

$$\vec{\mathrm{rot}}\,\mathrm{rot}\,\vec{u} = -\Delta\vec{u} + \vec{\nabla}\,\mathrm{div}\,\vec{u} \tag{2.48}$$

L'espace $H(\mathrm{rot}, \Omega)$ est défini comme étant le domaine dans $L^2(\Omega)$ de l'opérateur rotationnel :

$$H(\mathrm{rot}, \Omega) = \{\vec{v} \in L^2(\Omega)^4;\ \mathrm{rot}\,\vec{v} \in L^2(\Omega)\} \tag{2.49}$$

Muni de la norme :

$$\|\vec{v}\|_{H(\mathrm{rot},\Omega)} = \left(\|\vec{v}\|^2_{0,\Omega} + \|\mathrm{rot}\,\vec{v}\|^2_{0,\Omega}\right)^{\frac{1}{2}} \tag{2.50}$$

cet espace est un espace de Hilbert.

Il est facile d'obtenir de nombreuses propriétés de l'espace $H(\mathrm{rot}, \Omega)$ à partir de celles de l'espace $H(\mathrm{div}, \Omega)$, en remarquant que

$$\vec{v} = (v_1, v_2) \in H(\mathrm{rot}, \Omega) \Longleftrightarrow \vec{w} = (v_2, -v_1) \in H(\mathrm{div}, \Omega)$$

et que $\vec{v}.\vec{\tau} = -\vec{w}.\vec{\nu}$, où $\vec{\tau} = (-\nu_2, \nu_1)$ est le vecteur unitaire tangent à la frontière. Ainsi, l'application trace tangente :

$$\gamma_\tau : \vec{v} \in H(\mathrm{rot}, \Omega) \longmapsto (\vec{v}.\vec{\tau})_{|\Gamma} \in H^{-\frac{1}{2}}(\Gamma)$$

est linéaire continue.

De plus, on a la formule de Green :

$$(\mathrm{rot}\,\vec{v}, \varphi) - (\vec{v}, \vec{\mathrm{rot}}\,\varphi) =< \gamma_\tau\vec{v}, \gamma_o\varphi > \tag{2.51}$$

ceci pour tout $\vec{v} \in H(\mathrm{rot}, \Omega)$, $\varphi \in H^1(\Omega)$.

2.2.5 L'espace $D(\vec{\operatorname{div}}, \Omega)$

On note souvent $L^2(\Omega)_s^{n^2}$ l'espace des tenseurs symétriques dont les composantes sont de carré sommable.

On muni cet espace du produit scalaire :

$$(\tau, \gamma) = \int_\Omega \tau : \gamma \, dx \tag{2.52}$$

où $\tau : \gamma$ est le produit contracté de deux tenseurs d'ordre n :

$$\tau : \gamma = \sum_{i,j=1}^{n} \tau_{ij} \gamma_{ij}$$

si $(\tau_{ij})_{1 \le i,j \le n}$ et $(\gamma_{ij})_{1 \le i,j \le n}$ sont les composantes dans une base $(\vec{x}_i)_{1 \le i \le n}$ des tenseurs τ et γ, respectivement.

Rappelons que $D(\vec{v})$ représente la partie symétrique du tenseur gradient $\vec{\nabla} \vec{v}$ d'un champ de vecteur $\vec{v}$, et que $\vec{\operatorname{div}}\, \gamma$ est un vecteur dont les composantes sont les divergences des vecteurs-ligne du tenseur τ dans la base considérée :

$$(\vec{\operatorname{div}}\, \tau)_i = \operatorname{div}((\tau_{ij})_{1 \le j \le n}) = \sum_{j=1}^{n} \partial_j \tau_{ij}$$

L'identité suivante est satisfaite :

$$2 \vec{\operatorname{div}}\, D(\vec{u}) = \Delta \vec{u} + \vec{\nabla} \operatorname{div} \vec{u} \tag{2.53}$$

Il est alors naturel d'introduire l'espace $D(\vec{\operatorname{div}}, \Omega)$ comme étant le domaine dans $L^2(\Omega)^n$ de l'opérateur $\vec{\operatorname{div}}$:

$$D(\vec{\operatorname{div}}, \Omega) = \{\gamma \in L^2(\Omega)_s^{n^2};\ \vec{\operatorname{div}}\, \gamma \in L^2(\Omega)^n\} \tag{2.54}$$

Cet espace, muni de la norme :

$$\|\gamma\|_{D(\vec{\operatorname{div}}, \Omega)} = \left(\|\gamma\|_{0,\Omega}^2 + \|\vec{\operatorname{div}}\, \gamma\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \tag{2.55}$$

est un espace de Hilbert. Des propriétés de $H(\operatorname{div}, \Omega)$ découlent celles de $D(\vec{\operatorname{div}}, \Omega)$. Remarquons en particulier la formule de Green :

$$(\gamma, D(\vec{v})) + (\vec{\operatorname{div}}\, \gamma, \vec{v}) = < \vec{v} \otimes \vec{\nu}, \gamma > \tag{2.56}$$

ceci pour tout $\gamma \in D(\vec{\operatorname{div}}, \Omega)$, $\vec{v} \in H^1(\Omega)^n$, et où $\vec{v} \otimes \vec{\nu}$ désigne le produit tensoriel des deux vecteurs $\vec{v}$ et $\vec{\nu}$, représenté par ses composantes $(v_i \nu_j)_{1 \le i,j \le n}$ dans la base $(\vec{x}_i)_{1 \le i \le n}$, si $(v_i)_{1 \le i \le n}$ et $(\nu_i)_{1 \le i \le n}$ sont les composantes de $\vec{v}$ et $\vec{\nu}$, respectivement.

Citons enfin le très utile

Théorème 2.2 *Inégalité de Korn*

Il existe une constante c_o strictement positive, telle que l'inégalité :

$$\|D(\vec{v})\|_{0,\Omega} \geq c_o|\vec{v}|_{1,\Omega} \tag{2.57}$$

ait lieu, pour tout $\vec{v} \in H_0^1(\Omega)^n$.

2.3 Résultats fondamentaux

Dans ce paragraphe sont rassemblées quelques propriétés des espaces introduits précédemment. Pour simplifier la présentation de ces résultats, nous supposerons que le domaine Ω est une partie ouverte simplement connexe de $\mathbb{R}^2$. Cependant, il est possible, sans grande difficulté, d'étendre ces résultats à un domaine multiplement connexe de $\mathbb{R}^2$. On pourra de plus consulter [21] pour le cas tridimensionnel.

2.3.1 Décomposition des champs de vecteur

Les résultats qui suivent découlent d'une propriété dûe à De Rham :

Théorème 2.3 *Si un champ de vecteur distribution $\vec{u} \in \mathcal{D}^{'}(\Omega)^n$ satisfait*

$$< \vec{u}, \vec{\phi} >= 0$$

pour tout $\phi \in \mathcal{D}(\Omega)^n$, $\operatorname{div} \vec{\phi} = 0$, alors, il existe une unique distribution $\varphi \in \mathcal{D}(\Omega)$ telle que

$$\vec{u} = \vec{\nabla}\varphi$$

Introduisons l'espace :

$$V = \{\vec{u} \in H_0^1(\Omega)^2,\ \operatorname{div} \vec{u} = 0\} \tag{2.58}$$

Lorsque nous établirons les théorèmes d'existence et d'unicité, nous utiliserons aussi le

Corollaire 2.4
L'opérateur divergence est un isomorphisme de $V^{\perp}$ dans

$$L_o^2(\Omega) \stackrel{déf}{=} \{q \in L^2(\Omega),\ \int_\Omega q\, dx = 0\} \tag{2.59}$$

Un autre résultat important est la décomposition orthogonale des champs de vecteurs de $L^2(\Omega)^n$:

Théorème 2.5 *Décomposition orthogonale de $L^2(\Omega)^n$*

Tout champ de vecteur $\vec{v} \in L^2(\Omega)^2$ admet la décomposition :

$$\vec{v} = \vec{\nabla} q + \vec{\operatorname{rot}}\, \phi \tag{2.60}$$

où $q \in H^1(\Omega)_{/\mathbb{R}}$ est l'unique solution de :

$$(\vec{\nabla} q, \vec{\nabla}\mu) = (\vec{v}, \vec{\nabla}\mu),\ \forall \mu \in H^1(\Omega) \tag{2.61}$$

et $\phi \in H_0^1(\Omega)$ est l'unique solution de

$$(\vec{\operatorname{rot}}\, \phi, \vec{\operatorname{rot}}\, \xi) = (\vec{v} - \vec{\nabla} q, \vec{\operatorname{rot}}\, \xi),\ \forall \xi \in H_0^1(\Omega) \tag{2.62}$$

Introduisant ensuite l'espace :

$$\tilde{V} = \{\vec{u} \in H_0(\mathrm{div}, \Omega);\ \mathrm{div}\, \vec{u} = 0\} \tag{2.63}$$

nous en déduisons que

$$L^2(\Omega)^n = \tilde{V} \oplus \tilde{V}^{\perp}$$

avec les caractérisations des espaces:

$$\tilde{V} = \{\vec{\mathrm{rot}}\, \phi,\ \phi \in H_0^1(\Omega)\} \tag{2.64}$$

$$\tilde{V}^{\perp} = \{\vec{\nabla} q,\ q \in H^1(\Omega)\} \tag{2.65}$$

Remarque 2.1 Fonction de courant

Lorsque $\vec{v} \in \tilde{V}$, ϕ, caractérisé par (2.62), représente la fonction de courant, et on a :

$$\vec{v} = \vec{\mathrm{rot}}\, \phi$$

□

2.3.2 Application aux tenseurs

La décomposition qui précéde trouve ici un corollaire immédiat qui nous sera très utile lors de la formulation mixte du problème de Stokes à l'aide de tenseurs.

Corollaire 2.6 *Représentation de vecteurs par des tenseurs déviateurs*

Tout champ de vecteur $\vec{u} = (u_1, u_2) \in L^2(\Omega)^2$ peut s'érire :

$$\vec{u} = \vec{\mathrm{div}}\, \chi \tag{2.66}$$

où χ est de la forme :

$$\chi = \begin{pmatrix} q & \phi \\ \phi & -q \end{pmatrix} \tag{2.67}$$

et, en posant $\vec{v} = (u_1, -u_2)$, q, ϕ sont caractérisés respectivement par (2.61) et (2.62).

Démonstration : Soit (u_1, u_2) les composantes de $\vec{u}$.
Posons $\vec{v} = (u_1, -u_2) \in L^2(\Omega)^2$. Alors, d'après le théorème de décomposition, il existe un unique $q \in H^1(\Omega)_{/\mathbb{R}}$ solution de (2.61) et un unique $\phi \in H_0^1(\Omega)$ solution de (2.62) tel que $\vec{v} = \vec{\nabla} q + \vec{\mathrm{rot}}\, \phi$.

Soit alors χ le tenseur déviateur défini par (2.67). On vérifie sans peine que :

$$\vec{\mathrm{div}}\, \chi = (v_1, -v_2) = \vec{u}$$

d'où le résultat. □

Remarque 2.2 Sur l'unicité de la représentation par un tenseur déviateur

On peut obtenir sans difficulté l'unicité de la représentation des vecteurs par des tenseur déviateurs en imposant de plus une condition sur la composante normale :

$$\int_\Omega q\,dx = 0$$

□

Introduisons alors le sous-espace fermé de $D(\vec{\text{div}}, \Omega)$:

$$\begin{aligned}\tilde{D} = \{\chi = &\begin{pmatrix} q & \phi \\ \phi & -q \end{pmatrix} \in D(\vec{\text{div}}, \Omega); \\ &\phi \in H_0^1(\Omega),\ q \in H^1(\Omega),\ \int_\Omega q\,dx = 0 \text{ et } \text{div}(\vec{\text{div}}\,\chi) = 0\}\end{aligned} \tag{2.68}$$

que l'on muni de la norme induite par celle de $D(\vec{\text{div}}, \Omega)$.

D'après ce qui précède, nous pouvons énoncer le

Théorème 2.7 *Les espaces $\tilde{V}$, $\tilde{D}$ et $H_0^1(\Omega)$ sont isomorphes.*

Plus précisément :

- $\vec{\text{rot}}$ *est un isomorphisme de $H_0^1(\Omega)$ dans $\tilde{V}$;*
- $\vec{\text{div}}$ *est un isomorphisme de $\tilde{D}$ dans $\tilde{V}$*

Ceci nous permet d'obtenir une inégalité du type (2.31) dans $\tilde{D}$:

Corollaire 2.8
Les normes $\|\vec{\text{div}}\,\chi\|_{0,\Omega}$ et $\|\chi\|_{D(\vec{\text{div}},\Omega)}$ sont équivalentes dans $\tilde{D}$.

Démonstration : Puisque l'opérateur $\vec{\text{div}}$ est continu bijectif de $\tilde{D}$ dans $\tilde{V}$, son inverse, noté $(\vec{\text{div}})^{-1}$, est aussi continu :

$$\|(\vec{\text{div}})^{-1}\vec{v}\|_{D(\vec{\text{div}},\Omega)} \leq c\,\|\vec{v}\|_{H(\text{div},\Omega)}$$

ceci pour tout $\vec{v} \in \tilde{V}$. En posant

$$\chi = (\vec{\text{div}})^{-1}\vec{v} \tag{2.69}$$

on obtient :

$$\|\chi\|_{D(\vec{\text{div}},\Omega)} \leq c\,\|\vec{\text{div}}\,\chi\|_{H(\text{div},\Omega)} \tag{2.70}$$

Or (2.69) établit une relation bi-univoque entre $\tilde{V}$ et $\tilde{D}$. Par conséquent, (2.70) est vérifié pour tout $\chi \in \tilde{D}$. □

2.4 Le problème de Stokes à trois champs

Dans ce paragraphe, on rappelle brièvement quelques résultats d'existence et d'unicité de la solution du problème de Stokes. Des résultats semblables sont ensuite présentés pour le problème de Stokes à trois champs (2.11)-(2.14).

On suppose ici que le domaine Ω est un ouvert borné simplement connexe de $\mathbb{R}^n$, et que sa frontière Γ est de classe C^1.

2.4.1 Rappels sur le problème de Stokes

Les grandeurs sans dimension $\vec{u}$, associé à la vitesse, et p à la pression, d'un fluide newtonien *lent* en écoulement stationnaire sont solution du problème de Stokes.

On a le résultat classique [21] suivant :

Théorème 2.9 *Résolution du problème de Stokes*
Soient $\vec{f} \in H^{-1}(\Omega)^n$, *et* $\vec{g} \in H^{\frac{1}{2}}(\Gamma)^n$ *tel que :*

$$\int_\Gamma \vec{g}.\vec{\nu}\, d\Gamma = 0 \tag{2.71}$$

Alors il existe un unique couple $(\vec{u}, p) \in H^1(\Omega)^n \times L^2_o(\Omega)$ *tel que*

$$-\nu\Delta\vec{u} + \vec{\nabla}p = \vec{f} \text{ dans } H^{-1}(\Omega)^n \tag{2.72}$$
$$\operatorname{div}\vec{u} = 0 \text{ dans } \Omega \tag{2.73}$$
$$\vec{u} = \vec{g} \text{ sur } \Gamma \tag{2.74}$$

où ν *est une constante strictement positive.*

Remarque 2.3 Problème de Stokes homogène
Pour $\vec{g} = 0$, le problème de Stokes équivaut encore à *trouver* $\vec{u} \in V$ *vérifiant :*

$$\nu(\vec{\nabla}\vec{u}, \vec{\nabla}\vec{v}) = < \vec{f}, \vec{v} >, \ \forall \vec{v} \in V \tag{2.75}$$

□

2.4.2 Résolution du problème de Stokes à trois champs

Rappelons l'identité (2.53), de sorte que les équations de Stokes s'écrivent aussi :

$$\begin{cases} \tau & -2\nu D(\vec{u}) & & = 0 \\ \vec{\operatorname{div}}\,\tau & & -\vec{\nabla}p & = -\vec{f} \\ & \operatorname{div}\vec{u} & & = 0 \\ & & \vec{u} & = \vec{g} \text{ sur } \Gamma \end{cases}$$

On a alors le

Théorème 2.10 *Résolution du problème de Stokes à trois champs*

Soit Ω *un ouvert borné connexe de* $\mathbb{R}^n$ *et* Γ *sa frontière, supposée de classe* C^1.

Soient $\vec{f} \in H^{-1}(\Omega)^n$, *et* $\vec{g} \in H^{\frac{1}{2}}(\Gamma)^n$ *satisfaisant (2.71).*

Alors, il existe un unique triplet $(\tau, \vec{u}, p) \in L^2(\Omega)_s^{n^2} \times H^1(\Omega)^n \times L_o^2(\Omega)$ *solution de*

$$
\begin{aligned}
(\tau - 2\nu D(\vec{u}), \gamma) &= 0 \quad \forall \gamma \in L^2(\Omega)_s^{n^2} && (2.76)\\
-(\tau, D(\vec{v})) + (p, \operatorname{div} \vec{v}) &= -(\vec{f}, \vec{v}) \quad \forall \vec{v} \in H_0^1(\Omega)^n && (2.77)\\
(\operatorname{div} \vec{u}, q) &= 0 \quad \forall q \in L^2(\Omega) && (2.78)\\
\vec{u} &= \vec{g} \text{ sur } \Gamma && (2.79)
\end{aligned}
$$

De plus $(\vec{u}, p)$ *est alors aussi l'unique solution du problème de Stokes (2.72)-(2.74), et on a*

$$\tau = 2\nu D(\vec{u}) \text{ dans } \Omega \tag{2.80}$$

Démonstration : D'après la relation (2.34), l'application trace γ_o est surjective de $H^1(\Omega)$ dans $H^{\frac{1}{2}}(\Gamma)$. Par conséquent, il existe un unique relèvement $\vec{u}_o \in H^1(\Omega)^n$ tel que

$$\vec{u}_o = \vec{g} \text{ sur } \Gamma \tag{2.81}$$

De la condition de compatibilité (2.71) sur $\vec{g}$, il vient alors, par la formule de Green :

$$\operatorname{div} \vec{u}_o = 0 \text{ dans } \Omega \tag{2.82}$$

On se place ensuite dans le cadre des problèmes variationnels abstraits (voir l'annexe A) avec :

$$
\begin{aligned}
\mathbf{X} &= L^2(\Omega)_s^{n^2}\\
\mathbf{M} &= V\\
\|.\|_{\mathbf{X}} &= \|.\|_{0,\Omega}\\
\|.\|_{\mathbf{M}} &= |.|_{1,\Omega}\\
a(\tau, \gamma) &= \frac{1}{2\nu}(\tau, \gamma), \ \forall \tau, \gamma \in L^2(\Omega)_s^{n^2}\\
b(\tau, \vec{v}) &= -(\tau, D(\vec{v})), \ \forall \tau \in L^2(\Omega)_s^{n^2}, \ \forall \vec{v} \in H_0^1(\Omega)^n\\
< l, \gamma > &= -b(\gamma, \vec{u}_o), \ \forall \gamma \in L^2(\Omega)_s^{n^2}\\
< \chi, \vec{v} > &= - < \vec{f}, \vec{v} >, \ \forall \vec{v} \in H_0^1(\Omega)^n
\end{aligned}
$$

On a alors :

$$\mathbf{V} = \{\gamma \in L^2(\Omega)_s^{n^2};\ (\gamma, D(\vec{v})) = 0,\ \forall \vec{v} \in V\}$$

Il reste à vérifier la $\mathbf{V}$-ellipticité de a, et la condition inf-sup sur b. L'ellipticité de a est immédiate puisque :

$$a(\gamma,\gamma) = \frac{1}{2\nu}\|\gamma\|_{0,\Omega}$$

En ce qui concerne la condition inf-sup sur b, il suffit d'établir que :

$$\sup_{\gamma \in L^2(\Omega)_s^{n^2}} \frac{-(D(\vec{v}),\gamma)}{\|\gamma\|_{0,\Omega}} \geq \beta\, |\vec{v}|_{1,\Omega},\ \forall \vec{v} \in V \tag{2.83}$$

Or, en choisissant $\gamma = -D(\vec{v})$, on obtient :

$$\sup_{\gamma \in L^2(\Omega)_s^{n^2}} \frac{-(D(\vec{v}),\gamma)}{\|\gamma\|_{0,\Omega}} \geq \|D(\vec{v})\|_{0,\Omega},\ \forall \vec{v} \in V$$

et, de l'inégalité de Korn (2.57), nous obtenons (2.83) avec $\beta = c_o$.

Par conséquent, d'après le théorème A.1 de Brezzi-Babuska, il existe un unique couple $(\tau, \vec{w}) \in L^2(\Omega)_s^{n^2} \times \mathbf{V}$ tel que :

$$\begin{cases} a(\tau,\gamma) & +b(\gamma,\vec{w}) & =< l,\gamma > & \forall \gamma \in L^2(\Omega)_s^{n^2} \\ b(\tau,\vec{v}) & & =< \gamma,\vec{v} > & \forall \vec{v} \in V \end{cases}$$

Interprétons à présent le problème résolu : autrement dit, $(\tau, \vec{u} = \vec{u}_o + \vec{w})$ est *la* solution de :

$$\frac{1}{2\nu}(\tau,\gamma) - (D(\vec{u}),\gamma) \qquad = 0, \qquad \forall \gamma \in L^2(\Omega)_s^{n^2} \tag{2.84}$$

$$-(\tau, D(\vec{v})) \quad = -(\vec{f},\vec{v}),\quad \forall \vec{v} \in \mathbf{V} \tag{2.85}$$

De (2.84), il vient $\tau = 2\nu D(\vec{u})$, et en reportant dans (2.85) :

$$-2\nu(D(\vec{u}), D(\vec{v})) = (-\vec{f},\vec{v}),\ \forall \vec{v} \in V \tag{2.86}$$

or, d'après le théorème 2.3 de De Rham, il existe un unique $p \in L_o^2(\Omega)$ tel que :

$$2\nu \operatorname{div} D(\vec{u}) = -\vec{f} + \vec{\nabla} p \text{ dans } H^{-1}(\Omega)^n \tag{2.87}$$

Enfin, du choix (2.81) de $\vec{u}_o$, il vient :

$$\vec{u} = \vec{g} \text{ sur } \Omega$$

et, puisque $\vec{w} \in V$:

$$\operatorname{div} \vec{u} = \operatorname{div} \vec{u}_o + \operatorname{div} \vec{w} = 0$$

De (2.86) et (2.87) et avec l'identité (2.53) , nous obtenons enfin :

$$-\nu\Delta\vec{u} + \vec{\nabla}p = \vec{f} \text{ dans } H^{-1}(\Omega)^n$$

Ainsi, $(\vec{u}, p)$ est bien la solution du problème de Stokes (2.72)-(2.74). □

2.4.3 Approximations classiques

Soient T_h, V_h et Q_h trois espaces de dimension finie, tels que

$$\begin{array}{rcl} T_h & \subset & L^2(\Omega)_s^{n^2} \\ V_h & \subset & H_0^1(\Omega)^n \\ Q_h & \subset & L_o^2(\Omega) \end{array}$$

et on note :

$$Q_{0,h} = Q_h \cap L_o^2(\Omega)$$

Le problème approché consiste à chercher $(\tau_h, \vec{u}_h, p_h) \in T_h \times V_h \times Q_{o,h}$ tels que :

$$\begin{array}{rcll} (\tau_h - 2\nu D(\vec{u}_h), \gamma_h) & = & 0,\ \forall \gamma_h \in T_h & (2.88) \\ -(\tau_h, D(\vec{v}_h)) + (p_h, \operatorname{div} \vec{v}_h) & = & -(\vec{f}, \vec{v}_h),\ \forall \vec{v}_h \in V_h & (2.89) \\ (\operatorname{div} \vec{u}_h, q_h) & = & 0,\ \forall q_h \in Q_h & (2.90) \\ \vec{u}_h = \vec{g} \text{ sur } \Gamma & & & (2.91) \end{array}$$

Pour ce problème, montrons le :

Théorème 2.11 *Existence et unicité de la solution approché*
Faisons les hypothèses suivantes :

i) $D(V_h) \subset T_h$

ii)

$$\inf_{q_h \in Q_h - \{0\}} \sup_{\vec{v}_h \in V_h - \{0\}} \frac{(q_h, \operatorname{div} \vec{v}_h)}{\|q_h\|_{0,\Omega} |\vec{v}_h|_{1,\Omega}} \geq c_1$$

où c_1 désigne une constante strictement positive.

Alors le problème (2.88)-(2.91) admet une solution unique.

DÉMONSTRATION : Posons $\vec{f} = 0$ et $\vec{g} = 0$, et soit $(\tau_h, \vec{u}_h, p_h)$ une solution du problème correspondant. En faisant $\gamma_h = \tau_h$ dans (2.88), il vient $\|\tau_h\|^2_{0,\Omega} = 2\nu(\tau_h, D(\vec{u}_h))$ De même, $q_h = p_h$ dans (2.90) conduit à $(\operatorname{div}\vec{u}_h, p_h) = 0$. Alors $\vec{v}_h = \vec{u}_h$ dans (2.89) conduit à $\tau_h = 0$.

Ainsi (2.88) s'écrit :

$$(D(\vec{u}_h), \gamma_h) = 0,\ \forall \gamma_h \in T_h$$

et l'hypothèse (i) permet de choisir $\gamma_h = D(\vec{u}_h)$, d'où $D(\vec{u}_h) = 0$. De l'inégalité de Korn (2.57), et puisque les normes induites par $\|.\|_{1,\Omega}$ et $|.|_{1,\Omega}$ sur $V_h \subset H^1_0(\Omega)$ sont équivalentes (théorm'eme friedrichs:poincarre), il vient $\vec{u}_h = 0$.

Enfin(2.89) s'écrit :

$$(\operatorname{div}\vec{v}_h, p_h) = 0,\ \forall \vec{v}_h \in V_h$$

L'hypothèse (ii) nous assure de l'existence d'un $\vec{v}_h \in V_h - \{0\}$ tel que $(\operatorname{div}\vec{v}_h, p_h) \geq c\,\|p_h\|_{0,\Omega}|\vec{v}_h|_{1,\Omega}$, où c est une valeur intermédiaire, $0 < c < c_1$; d'où $p_h = 0$. □

Afin de satisfaire la condition inf-sup (ii), nous sommes conduit à utiliser des éléments de degré élevé pour approcher le champ des vitesses. De plus, la solution approchée u_h ainsi obtenue ne satisfait généralement pas la relation d'incompressibilité : $\operatorname{div}\vec{u}_h$ reste "petit" mais non nul [31].

Il est cependant possible d'abaisser le degré des éléments en vitesse, à condition d'ajouter un nouveau degré de liberté (ou *bulle* : produit des coordonnées barycentriques) sur chaque élément. L'élément obtenu [2] satisfait alors la condition inf-sup. Un résultat similaire est obtenu en utilisant un procédé de régularisation du problème de Stokes pour l'élément $P_1 - P_1$ [8]. Si le nombre total de degrés de liberté est ainsi considérablement diminué, la relation d'incompressibilité n'est toujours pas exactement satisfaite.

Remarque 2.4 Choix de l'espace T_h

Dans tous les cas, on pourra choisir commodément $T_h = D(V_h)$ [4]. □

2.5 Deux formulations du problème de Stokes

Utilisant une approximation variationnelle mixte, nous présentons un procédé de construction de solutions approchées $\vec{u}_h$ à divergence nulle. L'espace d'approximation des contraintes devra alors satisfaire une condition de compatibilité que nous expliciterons. Proposant alors l'élément de Raviart-Thomas pour les vitesses, nous construirons quelques éléments associés pour les contraintes.

Nous allons appliquer ici la théorie abstraite rapportée en annexe A, afin d'obtenir des formulations variationelles des problèmes (2.11)-(2.14) et (2.7)-(2.10). Par commodité, et sans perte de généralité, nous supposerons $\vec{g} = 0$ (problème homogène). Pour simplifier l'exposé, on se restreindra au cas bidimensionnel.

2.5.1 La formulation vitesse-taux de déformation

Rappelons l'espace V introduit en (2.58). Puisque :

$$(\vec{\nabla}\vec{u}, \vec{\nabla}\vec{v}) = 2\,(D(\vec{u}), D(\vec{v})),\ \forall \vec{u}, \vec{v} \in V \tag{2.92}$$

le problème de Stokes (2.72)-(2.74) est équivalent à

$$(S_1) : \left\{ \begin{array}{l} \textit{trouver } \vec{u} \in V \textit{ tel que} \\ 2\nu(D(\vec{u}), D(\vec{v})) =< \vec{f}, \vec{v} >,\ \forall \vec{v} \in V \end{array} \right. \tag{2.93}$$

Pour commencer, remarquons qu'une autre description de l'espace V, à l'aide du tenseur taux de déformation, est :

$$V_1 = \{(\vec{v}, \gamma) \in H_0^1(\Omega) \times L^2(\Omega)^4_s;\ \operatorname{div} \vec{v} = 0 \text{ et } \gamma = D(\vec{v})\} \tag{2.94}$$

Ainsi, le problème (S_1) équivaut à

$$(P_1) : \left\{ \begin{array}{l} \textit{trouver } (\vec{u}, \tau) \in V_1 \textit{ tel que} \\ 2\nu(\tau, \gamma) =< \vec{f}, \vec{v} >,\ \forall (\vec{v}, \gamma) \in V_1 \end{array} \right. \tag{2.95}$$

Ce dernier problème se prête alors directement à une formulation mixte très utile.

2.5.2 La formulation vitesse-tourbillon

D'autre part, puisque :

$$(\vec{\nabla}\vec{u}, \vec{\nabla}\vec{v}) = (\operatorname{rot} \vec{u}, \operatorname{rot} \vec{v}),\ \forall \vec{u}, \vec{v} \in V \tag{2.96}$$

une formulation équivalente de (2.72)-(2.74) est aussi :

$$(S_2) : \left\{ \begin{array}{l} \textit{trouver } \vec{u} \in V \ \textit{tel que} \\ \nu(\mathrm{rot}\,\vec{u}, \mathrm{rot}\,\vec{v}) =< \vec{f}, \vec{v} >, \ \forall \vec{v} \in V \end{array} \right. \tag{2.97}$$

Remarquons, de façon analogue, qu'une autre description de l'espace V à l'aide du tourbillon est :

$$V_2 = \{(\vec{v}, \theta) \in H_0^1(\Omega) \times L^2(\Omega); \ \mathrm{div}\,\vec{v} = 0 \text{ et } \theta = \mathrm{rot}\,\vec{v}\} \tag{2.98}$$

Ainsi, le problème (S_2) équivaut à

$$(P_2) : \left\{ \begin{array}{l} \textit{trouver } (\vec{u}, \omega) \in V_2 \ \textit{tel que} \\ \nu(\omega, \theta) =< \vec{f}, \vec{v} >, \ \forall (\vec{v}, \theta) \in V_2 \end{array} \right. \tag{2.99}$$

qui se prête de même à une formulation tout aussi intéréssante.

2.6 Formulations mixtes associées

Les formulations (P_1) (vitesse-taux de déformation) et (P_2) (vitesse-tourbillon) ne sont cependant pas satisfaisantes, car l'approximation interne de V est délicate [30], et les approximations internes de V_1 et V_2 nécéssiteraient des éléments sophistiqués.

D'autre part, une approximation externe de V ne nous satisferait pas non plus, car la condition d'incompressibilité ne serait plus nécessairement satisfaite [50]; or cette condition reste cruciale lorsqu'on a en vue d'introduire les termes de transport des contraintes (traduisant l'effet mémoire du fluide visco-élastique).

Afin de contourner ces difficultés, nous allons introduire un multiplicateur de Lagrange λ correspondant aux contraintes :

$$\tau = D(\vec{u})$$

pour le problème (P_1), et

$$\omega = \mathrm{rot}\,\vec{u}$$

pour le problème (P_2). Nous relaxerons la régularité de nos fonctions-test tout en conservant les propriétés de divergence nulle dans les espaces V_1 et V_2. Pour cela, nous suivrons le modèle de la formulation mixte abstraite de l'annexe A.

2.6.1 Introduction d'un multiplicateur de Lagrange

Considérons tout d'abord le problème (P_1), et, puisqu'il n'y a pas de confusion, nous omettrons l'indice 1. Il s'agit précisément du type de problème étudié en annexe A, avec le choix suivant des espaces et des formes bilinéaires :

$$\mathbf{X} = \{\vec{v} \in H_0^1(\Omega)^2;\ \operatorname{div}\vec{v} = 0\} \times L^2(\Omega)_s^4$$

$$\mathbf{M} = L^2(\Omega)_s^4$$

$$\|v\|_{\mathbf{X}} = \left(|\vec{v}|_{1,\Omega}^2 + \|\gamma\|_{0,\Omega}^2\right)^{\frac{1}{2}}$$

pour tout $v = (\vec{v}, \gamma) \in \mathbf{X}$,

$$\|\mu\|_{\mathbf{M}} = \|\mu\|_{0,\Omega}$$

pout tout $\mu \in \mathbf{M}$,

$$a(u, v) = 2\nu(\tau, \gamma)$$

pour tout $u = (\vec{u}, \tau),\ v = (\vec{v}, \gamma) \in \mathbf{X}$,

$$b(v, \mu) = (D(\vec{v}) - \gamma, \mu)$$

pour tout $v = (\vec{v}, \gamma) \in \mathbf{X}$, $\mu \in \mathbf{M}$.

Alors le noyau défini par V est :

$$\mathbf{V} = \{v \in \mathbf{X};\ b(v, \mu) = 0,\ \forall \mu \in \mathbf{M}\}$$

et le problème mixte (Q) associé à (P) s'énonce :

$$(Q) : \left\{\begin{array}{llll} \textit{trouver } (u, \lambda) \in \mathbf{X} \times \mathbf{M} \textit{ tel que} & & & \\ a(u, v) & +b(v, \lambda) & =< \vec{f}, \vec{v} >, & \forall v \in \mathbf{X} \\ b(u, \mu) & & = 0 & \forall \mu \in \mathbf{M} \end{array}\right. \tag{2.100}$$

et nous avons le

Théorème 2.12
Le problème (Q) admet exactement une solution $(u, \lambda) \in \mathbf{X} \times \mathbf{M}$, et $u = (\vec{u}, D(\vec{u}))$, $\lambda = 2\nu D(\vec{u})$, où $\vec{u}$ est la solution du problème de Stokes (S).

Démonstration : Puisque (P) est équivalent à (S), il suffit de vérifier la condition inf-sup sur b. Pour cela, on choisit $v = (0, -\mu) \in \mathbf{X}$, et alors :

$$\sup_{v \in \mathbf{X},\, v \neq 0} \frac{b(v, \mu)}{\|v\|_{\mathbf{X}}} \geq \|\mu\|_{\mathbf{M}}$$

d'où $\beta = 1$.

Le problème (Q) admet donc une unique solution $(u, \lambda) \in \mathbf{X}$, où $u = (\vec{u}, D(\vec{u}))$ est *la* solution de (P).

Il reste donc à vérifier que $\lambda = 2\nu D(\vec{u})$. Pour cela, soit $v = (\vec{v}, \gamma) \in \mathbf{X}$ et $u = (\vec{u}, D(\vec{u}))$ la solution de (P). Alors :

$$\begin{aligned}
&a(u,v) + b(v, 2\nu D(\vec{u})) \\
&\quad = \quad 2\nu(D(\vec{u}), \gamma) + (D(\vec{v}) - \gamma, 2\nu D(\vec{u})) \\
&\quad = \quad 2\nu(D(\vec{u}), D(\vec{v})) \\
&\quad = \quad < \vec{f}, \vec{v} >
\end{aligned}$$

puisque $\vec{u}$ est la solution de (P). Or, comme λ est unique, on a montré $\lambda = 2\nu D(\vec{u})$. □

2.6.2 Relaxation de la régularité des fonctions-test

En suivant toujours le cadre variationnel abstrait présenté en annexe A, nous faisons le choix suivant des espaces :

$$\begin{aligned}
\tilde{\mathbf{X}} &= \tilde{V} \times L^2(\Omega)_s^4 \\
\tilde{\mathbf{M}} &= D(\vec{\mathrm{div}}, \Omega)
\end{aligned}$$

où $\tilde{V}$ a été défini en (2.63).

Les normes asociées sont :

$$\begin{aligned}
\|v\|_{\tilde{\mathbf{X}}} &= \left(\|\vec{v}\|_{0,\Omega}^2 + \|\gamma\|_{0,\Omega}^2\right)^{\frac{1}{2}}, \text{ pour tout } v = (\vec{v}, \gamma) \in \tilde{\mathbf{X}} \\
\|\mu\|_{\tilde{\mathbf{M}}} &= \left(\|\mu\|_{0,\Omega}^2 + \|\vec{\mathrm{div}}\,\mu\|_{0,\Omega}^2\right)^{\frac{1}{2}}, \text{ pour tout } \mu \in \tilde{\mathbf{M}}
\end{aligned}$$

On choisit alors les formes bilinéaires suivantes :

$$\begin{aligned}
\tilde{a}(u,v) &= a(u,v) = 2\nu(\tau, \gamma), & \forall u = (\vec{u}, \tau), v = (\vec{v}, \gamma) \in \tilde{\mathbf{X}} \\
\tilde{b}(v,\mu) &= -(\vec{v}, \vec{\mathrm{div}}\,\mu) - (\gamma, \mu), & \forall v = (\vec{v}, \gamma) \in \tilde{\mathbf{X}}, \mu \in \tilde{\mathbf{M}}
\end{aligned}$$

L'espace $\tilde{\mathbf{V}}$ obtenu est :

$$\tilde{\mathbf{V}} = \{(\vec{v}, \gamma) \in \tilde{\mathbf{X}};\ (\vec{v}, \vec{\mathrm{div}}\,\mu) = -(\gamma, \mu),\ \forall \mu \in \tilde{\mathbf{M}}\}$$

Rappelons à présent l'espace $\tilde{D}$ introduit en (2.68), et la correspondance bi-univoque établie entre $\tilde{V}$ et $\tilde{D}$ par la relation :

$$\vec{v} = \vec{\mathrm{div}}\,\chi$$

Ainsi $\tilde{\mathbf{V}}$ peut être décrit de façon équivalente par

$$\tilde{\mathbf{V}} = \{(\vec{\operatorname{div}}\chi, \gamma);\ \chi \in \tilde{D},\ \gamma \in L^2(\Omega)^4_s \text{ avec } (\vec{\operatorname{div}}\,\chi, \vec{\operatorname{div}}\,\gamma) = (\gamma, \mu),\ \forall \mu \in \tilde{\mathbf{M}}\} \tag{2.101}$$

Les problèmes $(\tilde{P})$ et $(\tilde{Q})$ associés à ces choix d'espace et de formes bilinéaires sont :

$$(\tilde{P}) : \begin{cases} \textit{trouver } u = (\vec{u}, \gamma) \in \tilde{\mathbf{V}} \textit{ tel que} \\ 2\nu(\tau, \gamma) = (\vec{f}, \vec{v}),\ \forall v = (\vec{v}, \gamma) \in \tilde{\mathbf{X}} \end{cases}$$

et :

$$(\tilde{Q}) : \begin{cases} \textit{trouver } u = (\vec{u}, \tau) \in \tilde{\mathbf{X}} \textit{ et } \lambda \in \tilde{\mathbf{M}} \textit{ tels que} \\ 2\nu(\tau, \gamma) - (\vec{v}, \vec{\operatorname{div}}\,\lambda) - (\gamma, \lambda) &= (\vec{f}, \vec{v}), & \forall v = (\vec{v}, \gamma) \in \tilde{\mathbf{X}} \\ -(\vec{u}, \vec{\operatorname{div}}\,\mu) - (\tau, \mu) &= 0 & \forall \mu \in \tilde{\mathbf{M}} \end{cases}$$

2.6.3 Existence et unicité de la solution

Vérifions à présent les conditions de compatibilité sur les espaces et les formes bilinéaires :

Injections continues et denses

Puisque l'injection canonique de $H^1(\Omega)$ dans $H(\operatorname{div}, \Omega)$ est continue et dense, il en va de même pour celle de $\mathbf{X}$ dans $\tilde{\mathbf{X}}$, et de celle de $\tilde{\mathbf{M}}$ dans $\mathbf{M}$.

Prolongement des formes bilinéaires

La forme bilinéaire $\tilde{a}$ coïncide avec a sur $\mathbf{X} \times \mathbf{X}$ de façon évidente. D'autre part, d'après la formule de Green (2.56), $\tilde{b}$ coïncide avec b sur $\mathbf{X} \times \tilde{\mathbf{M}}$.

Ellipticité sur le noyau

En ce qui concerne l'ellipticité de $\tilde{a}$ sur le noyau $\tilde{\mathbf{V}}$, on a :

$$\tilde{a}(v, v) = 2\nu\|\gamma\|^2_{0,\Omega},\ \text{pour tout } v = (\vec{v}, \gamma) \in \tilde{\mathbf{V}}$$

En rappelant l'expression (2.101) de $\tilde{\mathbf{V}}$, il vient, en choisissant $\vec{v} = \vec{\operatorname{div}}\,\chi$ et $\mu = \chi \in \tilde{\mathbf{M}}$:

$$\|\vec{v}\|^2_{0,\Omega} = \|\vec{\operatorname{div}}\,\chi\|^2_{0,\Omega} = (\gamma, \chi)$$

Or, en combinant l'inégalité de Cauchy-Schwartz :

$$(\gamma, \chi) \leq \|\gamma\|_{0,\Omega}\|\chi\|_{0,\Omega}$$

avec le résultat du corollaire 2.8 :

$$\|\gamma\|_{0,\Omega}^2 \leq \|\chi\|_{0,\Omega}^2 + \|\vec{\operatorname{div}}\,\chi\|_{0,\Omega}^2 \leq c\,\|\vec{\operatorname{div}}\,\chi\|_{0,\Omega}^2$$

nous obtenons :

$$\frac{1}{c}\|\vec{\operatorname{div}}\,\chi\|_{0,\Omega}\|\chi\|_{0,\Omega} \leq \|\vec{\operatorname{div}}\,\chi\|_{0,\Omega}^2 = (\gamma,\chi) \leq \|\gamma\|_{0,\Omega}\|\chi\|_{0,\Omega}$$

soit encore :

$$\|\vec{\operatorname{div}}\,\chi\|_{0,\Omega} \leq c\,\|\gamma\|_{0,\Omega}$$

Ainsi :

$$\begin{aligned}\tilde{a}(v,v) &= 2\nu\|\gamma\|_{0,\Omega}^2\\ &\geq \frac{2\nu}{c^2+1}\left(\|\gamma\|_{0,\Omega}^2 + \|\vec{\operatorname{div}}\,\chi\|_{0,\Omega}^2\right)\end{aligned}$$

D'où la $\tilde{V}$-ellipticité de $\tilde{a}$ avec

$$\tilde{\alpha} = \frac{2\nu}{c^2+1}$$

Existence et unicité de la solution faible

Etablissons tout d'abord le

Lemme 2.1 *Les espaces* $\mathbf{V}$ *et* $\tilde{\mathbf{V}}$ *sont identiques.*

Démonstration : On a bien sûr $\mathbf{V} \subset \tilde{\mathbf{V}}$. Montrons à présent que $\tilde{\mathbf{V}} \subset \mathbf{V}$. Soit $v = (\vec{v},\gamma) \in \tilde{\mathbf{V}}$, et étendons $\vec{v}$ et γ par zéro hors de Ω en :

$$\tilde{\vec{v}} = \begin{cases} \vec{v} & \text{dans } \Omega \\ 0 & \text{dans } \mathbb{R}^2 - \Omega \end{cases}$$

$$\tilde{\gamma} = \begin{cases} 2\gamma & \text{dans } \Omega \\ 0 & \text{dans } \mathbb{R}^2 - \Omega \end{cases}$$

Alors $\gamma \in L^2(\mathbb{R}^2)_s^4$ et $\tilde{\vec{v}} \in H(\operatorname{div},\mathbb{R}^2)$ avec $\operatorname{div}\vec{v} = 0$. De plus

$$\int_{\mathbb{R}^2} (\tilde{\vec{v}}.\vec{\operatorname{div}}\,\mu + \tilde{\gamma} : \mu)\,dx = 0,\ \forall \mu \in D(\vec{\operatorname{div}},\Omega)$$

et, en conséquence :

$$\tilde{\gamma} = D(\tilde{\vec{v}}) \in L^2(\mathbb{R}^2)_d^4$$

Il est facile de vérifier, par transformation de Fourier, que cette dernière relation entraîne $\vec{\tilde{v}} \in H^1(\mathbb{R}^2)^2$ et donc $\vec{v} \in H^1(\Omega)^2$. Maintenant, d'une part :

$$\vec{v}.\vec{\nu} = 0 \text{ sur } \Gamma, \text{ car } \vec{v} \in \tilde{\mathbf{V}}$$

et, d'autre part, par la formule de Green :

$$\begin{aligned} 0 &= -(\vec{v}, \vec{\mathrm{div}}\, \mu) - (\gamma, \mu) \\ &= (D(\vec{v}) - \gamma, \mu) + \int_\Gamma (\vec{v} \otimes \vec{\nu}) : \mu \, d\sigma \end{aligned}$$

ceci pour tout $\mu \in D(\vec{\mathrm{div}}, \Omega)$.

D'où $\vec{v} \otimes \vec{\nu} = 0$, et on vérifie facilement que :

$$\left.\begin{array}{ll} \vec{v}.\vec{\nu} & = 0 \text{ sur } \Gamma \\ \vec{v} \otimes \vec{\nu} & = 0 \text{ sur } \Gamma \end{array}\right\} \Rightarrow \vec{v} = 0 \text{ sur } \Gamma$$

d'où $\vec{v} \in H_0^1(\Omega)^2$.

□

De ce qui précède, et du théorème A.2 nous pouvons énoncer le

Théorème 2.13
Le problème $(\tilde{P})$ admet une solution unique $u = (\vec{u}, D(\vec{u})) \in \mathbf{V}$, où $\vec{u}$ est la solution du problème de Stokes.

De plus, si $D(\vec{u}) \in D(\vec{\mathrm{div}}, \Omega)$, alors les problèmes (Q) et $(\tilde{Q})$ sont équivalents.

2.6.4 Cas de la formulation vitesse-tourbillon

Bien sûr, tous les résultats précédent ont leur contrepartie en terme de tourbillon. Ainsi, en posant :

$$\begin{aligned} \mathbf{X} &= V \times L^2(\Omega) \\ \mathbf{M} &= L^2(\Omega) \end{aligned}$$

nous obtenons le problème (Q_2) associé à (P_2) :

$$(Q_2) : \left\{ \begin{array}{lll} \multicolumn{3}{l}{\textit{trouver } u = (\vec{u}, \omega) \in \mathbf{X} \textit{ et } \lambda \in L^2(\Omega) \textit{ tels que}} \\ \nu(\omega, \theta) + (\mathrm{rot}\, \vec{v} - \theta, \lambda) & = (\vec{f}, \vec{v}), & \forall v = (\vec{v}, \theta) \in \mathbf{X} \\ (\mathrm{rot}\, \vec{u} - \omega, \mu) & = 0 & \forall \mu \in \mathbf{M} \end{array} \right.$$

Le théorème A.1 nous assure alors que (Q) admet une unique solution

$$(u = (\vec{u}, \omega = \mathrm{rot}\, \vec{u}), \lambda = \nu \mathrm{rot}\, \vec{u}) \in \mathbf{X} \times \mathbf{M}$$

où $\vec{u}$ est la solution du problème de Stokes.

De façon similaire, on relaxe la régularité de (Q) en posant :

$$\begin{array}{rcl}\tilde{\mathbf{X}} & = & \tilde{\mathbf{V}} \times L^2(\Omega) \\ \tilde{\mathbf{M}} & = & H^1(\Omega)\end{array}$$

Conservons a inchangé, et exprimons $\tilde{b}$ par :

$$\tilde{b}(v,\mu) = (\vec{v}, \vec{\mathrm{rot}}\,\mu) - (\theta, \mu)$$

pour tout $v = (\vec{v}, \theta) \in \tilde{\mathbf{X}}$ et $\mu \in \tilde{\mathbf{M}}$.

La forme faible de (Q_2) est :

$$(\tilde{Q_2}) : \left\{ \begin{array}{lll} \multicolumn{3}{l}{\textit{trouver } u = (\vec{u}, \omega) \in \tilde{\mathbf{X}} \textit{ et } \lambda \in H^1(\Omega) \textit{ tels que}} \\ \nu(\omega, \theta) + (\vec{v}, \vec{\mathrm{rot}}\,\lambda) - (\theta, \lambda) & = (\vec{f}, \vec{v}), & \forall v = (\vec{v}, \theta) \in \tilde{\mathbf{X}} \\ (\vec{u}, \vec{\mathrm{rot}}\,\mu) - (\omega, \mu) & = 0 & \forall \mu \in H^1(\Omega) \end{array} \right.$$

Ici aussi, le théorème A.2 nous assure que le problème faible est bien-posé, et que, si $\mathrm{rot}\,\vec{u} \in H^1(\Omega)$, alors les problèmes (Q_2) et $(\tilde{Q_2})$ sont équivalents.

2.7 Méthode d'approximation mixte

Dans ce paragraphe, nous supposerons la solution du problème de Stokes $\vec{u}$ assez régulière, ce qui nous permettra d'utiliser indifféremment les formulations vitesse-taux de déformation et vitesse-tourbillon.

2.7.1 Résultats d'approximation abstraite

Commençons par présenter l'approximation du problème (Q_1), et puisqu'il n'y a pas de confusion, nous omettrons l'indice. Appliquant la méthode d'approximation abstraite à ce problème, nous introduisons trois espaces de dimension finie :

$$\begin{aligned} V_h &\subset \tilde{V} \\ T_h &\subset L^2(\Omega)_s^4 \\ \mathbf{M}_h &\subset D(\vec{\mathrm{div}}, \Omega) \end{aligned}$$

et nous noterons :

$$\mathbf{X}_h = V_h \times T_h$$

L'espace $\mathbf{V}_h$ est définit par :

$$\mathbf{V}_h = \{(\vec{v}_h, \gamma_h) \in V_h \times T_h;\ (\vec{v}_h, \vec{\mathrm{div}}\, \mu_h) = -(\gamma_h, \mu_h),\ \forall \mu_h \in M_h\} \quad (2.102)$$

Les lemmes qui suivent traitent de la condition inf-sup discrète sur $\tilde{b}$ et de la $\mathbf{V}_h$-ellipticité de $\tilde{a}$.

Lemme 2.2 $\mathbf{V}_h$*-ellipticité de* $\tilde{a}$

Supposons que l'application $\vec{\mathrm{div}}$ *est surjective de* $\mathbf{M}_h$ *dans* V_h.

Alors il existe deux constantes strictement positives, indépendantes de h, c *et* α^* *telles que les inégalités :*

$$\begin{aligned} \|\vec{v}_h\|_{0,\Omega} &\leq c\, \|\gamma_h\|_{0,\Omega} && (2.103) \\ \nu\, \|\gamma_h\|_{0,\Omega}^2 &\geq \alpha^*\, \|\vec{v_h}, \gamma_h\|_{\mathbf{X}}^2 && (2.104) \end{aligned}$$

aient lieu pour tout $(\vec{v}_h, \gamma_h) \in \mathbf{V}_h$.

DÉMONSTRATION :

- Par la définition (2.102) de $\mathbf{V}_h$, pour tout $(\vec{v}_h, \gamma_h) \in \mathbf{V}_h$, on a :

$$(\vec{v}_h, \vec{\mathrm{div}}\, \mu_h) = -(\gamma_h, \mu_h),\ \forall \mu_h \in \mathbf{M}_h \quad (2.105)$$

Puisque $V_h \subset \vec{\text{div}}\,\mathbf{M}_h$, il est possible de choisir dans la relation précédente $\mu_h \in \mathbf{M}_h$ tel que $\vec{v}_h = \vec{\text{div}}\,\mu_h$, et il vient alors, à l'aide de l'inégalité de Cauchy-Schwartz :

$$\|\vec{v}_h\|^2_{0,\Omega} \leq \|\gamma_h\|_{0,\Omega}\|\mu_h\|_{0,\Omega} \tag{2.106}$$

Or, de $\vec{v}_h = \vec{\text{div}}\,\mu_h \in V_h \subset \tilde{V}$ il vient de la définition (2.68) de $\tilde{D}$ que $\mu_h \in \tilde{D}$; puisque les normes $\|\vec{\text{div}}\,(.)\|_{0,\Omega}$ et $\|.\|_{D(\vec{\text{div}},\Omega)}$ sont équivalentes sur $\tilde{D}$ (corollaire 2.8), nous obtenons la première inégalité du lemme :

$$\|\mu_h\|_{0,\Omega} \leq c\,\|\vec{\text{div}}\,\mu_h\|_{0,\Omega} = c\,\|\vec{v}_h\|_{0,\Omega} \tag{2.107}$$

- De (2.106) et (2.107), il vient :

$$\|\vec{v}_h\|_{0,\Omega} \leq c\,\|\gamma_h\|_{0,\Omega},\ \forall(\vec{v}_h, \gamma_h) \in \mathbf{V}_h$$

En conséquence, l'application

$$(\vec{v}_h, \gamma_h) \in \mathbf{V}_h \longmapsto \|\gamma_h\|_{0,\Omega}$$

est une norme sur $\mathbf{V}_h$, équivalente à la norme induite par $\|.\|_{\tilde{\mathbf{X}}}$ sur $\mathbf{V}_h$. En particulier, la deuxième inégalité du lemme est vérifiée. □

Lemme 2.3 *Condition inf-sup discrète*
Si $M_h \subset T_h$, alors

$$\sup_{v_h \in \mathbf{X}_h} \frac{\tilde{b}(v_h, \mu_h)}{\|v_h\|_{\tilde{\mathbf{X}}}} \geq \|\mu_h\|_{0,\Omega},\ \forall \mu_h \in M_h$$

Démonstration : Soit $\mu_h \in M_h$. Comme $M_h \subset T_h$, le couple $v_h = (0, -\mu_h) \in \mathbf{X}_h = V_h \times T_h$, et, pour ce choix de v_h :

$$\frac{\tilde{b}(v_h, \mu_h)}{\|v_h\|_{\tilde{\mathbf{X}}}} = \|\mu_h\|_{0,\Omega}$$

d'où $\beta^* = 1$. □

Théorème 2.14 *Conditions de compatibilité des espaces V_h, T_h et $\mathbf{M}_h$*

- *Si les espaces T_h, M_h et V_h satisfont aux conditions de compatibilité :*

$$V_h \subset \vec{\text{div}}\,\mathbf{M}_h \ \text{ et }\ \mathbf{M}_h \subset T_h \tag{2.108}$$

alors le problème

$$(Q_h) : \begin{cases} \text{trouver } u_h = (\vec{u}_h, \tau_h) \in \mathbf{X}_h \text{ et } \lambda_h \in \mathbf{M}_h \text{ tels que} \\ 2\nu(\tau_h, \gamma_h) - (\vec{v}_h, \vec{\operatorname{div}}\, \lambda_h) - (\gamma_h, \lambda_h) \quad = (\vec{f}, \vec{v}_h), \quad \forall v_h = (\vec{v}_h, \gamma_h) \in \mathbf{X}_h \\ -(\vec{u}_h, \vec{\operatorname{div}}\, \mu_h) - (\tau_h, \mu_h) \quad = 0 \quad \forall \mu_h \in \mathbf{M}_h \end{cases}$$

admet une solution unique $(u_h = (\vec{u}_h, \tau_h), \lambda_h) \in \mathbf{X}_h \times \mathbf{M}_h$

- *Si, de plus* $\mathbf{M}_h = T_h$, *alors* $\lambda_h = 2\nu\tau_h$.

Démonstration :

• Le premier point s'obtient par application des lemmes 2.2 et 2.3, et des théorèmes généraux de l'annexe A.

• On vérifie ensuite que $(u_h = (\vec{u}_h, \tau_h), 2\nu\tau_h)$ est solution de (Q_h). Par unicité de la solution, on en déduit $\lambda_h = 2\nu\tau_h$. □

Remarque 2.5 Cas où $\mathbf{M}_h = T_h$

La condition de compatibilité s'écrit :

$$V_h \subset \vec{\operatorname{div}}\, T_h \tag{2.109}$$

Il est possible d'éliminer le multiplicateur de Lagrange $\lambda_h = 2\nu\tau_h$ du système (Q_h), et on obtient le problème $(\bar{P}_h)$: *trouver* $\vec{u}_h \in V_h$ *et* $\tau_h \in T_h$ *tels que*

$$-2\nu(\vec{\operatorname{div}}\, \tau_h, \vec{v}_h) \quad = (\vec{f}, \vec{v}_h), \quad \forall \vec{v}_h \in V_h \tag{2.110}$$

$$(\vec{u}_h, \vec{\operatorname{div}}\, \gamma_h) + (\tau_h, \gamma_h) \quad = 0 \quad \forall \gamma_h \in T_h \tag{2.111}$$

qui est une approximation à posteriori du problème :

$$(\bar{P}) : \begin{cases} \textit{trouver } \vec{u} \in \tilde{V} \textit{ et } \tau \in D(\vec{\operatorname{div}}, \Omega) \textit{ tels que} \\ -2\nu(\vec{\operatorname{div}}\, \tau, \vec{v}) \quad = (\vec{f}, \vec{v}), \quad \forall \vec{v} \in \tilde{V} \\ (\vec{u}, \vec{\operatorname{div}}\, \gamma) + (\tau, \gamma) \quad = 0 \quad \forall \gamma \in D(\vec{\operatorname{div}}, \Omega) \end{cases}$$

□

Théorème 2.15 *Majoration abstraite de l'erreur*

On suppose $\mathbf{M}_h = T_h$ *et* $V_h \subset \vec{\operatorname{div}}\, T_h$.

Alors, si $(u_h = (\vec{u}_h, \tau_h), \lambda_h = 2\nu\tau_h) \in \mathbf{X}_h \times \mathbf{M}_h$ *désigne la solution du problème* (Q_h) *et* $(u = (\vec{u}, \tau), \lambda = 2\nu\tau) \in \tilde{\mathbf{X}} \times \tilde{\mathbf{M}}$ *celle de* $(\tilde{Q})$, *on a la majoration abstraite de l'erreur :*

$$\begin{aligned} &\|\vec{u} - \vec{u}_h\|_{0,\Omega} + \|\tau - \tau_h\|_{0,\Omega} + \|2\nu\tau - \lambda_h\|_{0,\Omega} \\ &\leq c\left((1 + S(h))\left(\inf_{\vec{v}_h \in V_h} \|\vec{u} - \vec{v}_h\|_{0,\Omega} + \inf_{\gamma_h \in T_h} \|\tau - \gamma_h\|_{0,\Omega}\right)\right. \\ &\quad \left. + \inf_{\mu_h \in T_h} \left(\|2\nu\tau - \mu_h\|_{0,\Omega}^2 + \|\vec{\operatorname{div}}\,(2\nu\tau - \mu_h)\|_{0,\Omega}^2\right)^{\frac{1}{2}}\right) \end{aligned} \tag{2.112}$$

où :

$$S(h) = \sup_{\gamma_h \in M_h} \frac{\|\gamma_h\|_{\mathbf{M}}}{\|\gamma_h\|_{\tilde{\mathbf{M}}}}$$

Passons maintenent à l'approximation du problème $(\tilde{Q}_2)$. Pour cela, on introduit l'espace de dimension finie :

$$\Theta_h \subset L^2(\Omega)$$

On obtient facilement des résultats analogues (voir [21]) et la condition de compatibilité entre V_h et $\Theta_h = \mathbf{M}_h^{(2)} \subset H^1(\Omega)$ est :

$$V_h \subset \vec{\mathrm{rot}}\, \Theta_h \tag{2.113}$$

Il est là aussi possibe d'éliminer le multiplicateur de Lagrange $\lambda_h = \nu\omega_h$ du système obtenu. Le problème (P_h) devient alors :

$$(\bar{P}_{2;h}) : \begin{cases} \textit{trouver } \vec{u}_h \in V_h \ \textit{ et } \omega_h \in \Theta_h \ \textit{tels que} \\ \nu(\mathrm{rot}\,\omega_h, \vec{v}_h) \quad = (\vec{f}, \vec{v}_h), \quad \forall \vec{v}_h \in V_h \\ (\vec{u}_h, \mathrm{rot}, \theta_h) \quad = (\omega_h, \theta_h) \quad \forall \gamma_h \in T_h \end{cases}$$

qui est une approximation à posteriori du problème :

$$(\bar{P}_2) : \begin{cases} \textit{trouver } \vec{u} \in \tilde{V} \ \textit{ et } \omega \in H^1(\Omega) \ \textit{tels que} \\ \nu(\mathrm{rot}\,\omega, \vec{v}) \quad = (\vec{f}, \vec{v}), \quad \forall \vec{v} \in \tilde{V} \\ (\vec{u}, \mathrm{rot}\,\theta) \quad = (\omega, \theta) \quad \forall \theta_h \in H^1(\Omega) \end{cases}$$

Remarque 2.6 Les conditions de compatibilités (2.113) et (2.109) sont suffisantes, mais non nécessaires à l'obtention de problèmes approchés bien-posés. □

2.7.2 Approximation de la pression

Jusqu'à présent, nous avons éliminé la pression de nos problèmes, en ne travaillant qu'avec des champs de vecteurs à divergence nulle.

Reprenant la formulation vitesse-taux de déformation, et omettant l'indice 1, nous considérons maintenant les espaces :

$$\begin{aligned} \mathbf{Y} &= H_0^1(\Omega)^2 \times L^2(\Omega)_s^4 \\ Q &= L_o^2(\Omega) \end{aligned}$$

et le problème mixte

$$(P^{'}) : \begin{cases} \textit{trouver } (u = (\vec{u}, \tau), p, \lambda) \in \mathbf{V} \times L_o^2(\Omega) \times \mathbf{M} \textit{ tel que} \\ a(u, v) + b(v, \lambda) - (p, \operatorname{div} \vec{v}) = (\vec{f}, \vec{v}), \ \forall v = (\vec{v}, \gamma) \in \mathbf{Y} \end{cases}$$

ainsi que l'espace

$$\tilde{\mathbf{Y}} = H_0(\operatorname{div}, \Omega) \times L^2(\Omega)_d^4$$

et le problème mixte faible associé:

$$(\tilde{P}^{'}) : \begin{cases} \textit{trouver } (u = (\vec{u}, \tau), p, \lambda) \in \tilde{\mathbf{V}} \times L_o^2(\Omega) \times \tilde{\mathbf{M}} \textit{ tel que} \\ \tilde{a}(u, v) + \tilde{b}(v, \lambda) - (p, \operatorname{div} \vec{v}) = (\vec{f}, \vec{v}), \ \forall v = (\vec{v}, \gamma) \in \tilde{\mathbf{Y}} \end{cases}$$

Il est clair que, si $\lambda = 2\nu\tau \in D(\vec{\operatorname{div}}, \Omega)$, alors les problèmes $(P^{'})$ et $(\tilde{P}^{'})$ sont équivalents.

De plus, les noyaux de b et $\tilde{b}$, respectivement :

$$\begin{aligned} \mathbf{W} &= \{v = (\vec{v}, \gamma) \in \mathbf{Y};\ b(\vec{v}, \gamma; \mu) = 0,\ \forall \mu \in \mathbf{M}\} \\ \tilde{\mathbf{W}} &= \{v = (\vec{v}, \gamma) \in \tilde{\mathbf{Y}};\ \tilde{b}(\vec{v}, \gamma; \mu) = 0,\ \forall \mu \in \tilde{\mathbf{M}}\} \end{aligned}$$

coïncident.

Afin d'approcher le problème $(\tilde{P}^{'})$, on se donne deux espaces de dimension finie :

$$\begin{aligned} \mathbf{Y}_h &\subset \tilde{\mathbf{Y}} \\ Q_h &\subset L_o^2(\Omega) \end{aligned}$$

et on introduit l'espace :

$$\mathbf{W}_h = \{v_h = (\vec{v}_h, \gamma_h) \in \mathbf{Y}_h;\ \tilde{b}(v_h, \mu_h) = 0,\ \forall \mu_h \in \mathbf{M}_h\}$$

Le problème approché consiste alors à

$$(P_h^{'}) : \begin{cases} \textit{trouver } (u_h = (\vec{u}_h, \tau_h), p_h) \in \mathbf{Z}_h \times Q_h \textit{ tel que} \\ \tilde{a}(u_h, v_h) - (p_h, \operatorname{div} \vec{v}_h) = (\vec{f}, \vec{v}_h), \ \forall v_h = (\vec{v}_h, \gamma_h) \in \mathbf{W}_h \end{cases}$$

où on suppose que $\mathbf{Z}_h$ est défini par :

$$\mathbf{Z}_h = \{v_h = (\vec{v}_h, \gamma_h) \in \mathbf{W}_h;\ (q_h, \operatorname{div} \vec{v}_h) = 0,\ \forall q_h \in Q_h\} \tag{2.114}$$

Alors, si $u_h = (\vec{u}_h, \tau_h), p_h) \in \mathbf{Z}_h \times Q_h$ est une solution de $(P_h^{'})$, il est clair que $u_h = (\vec{u}_h, \tau_h)$ est solution de (P_h).

Réciproquement, on a le

Théorème 2.16
Si $\mathbf{Z}_h$ est défini par (2.114), et que :

$$\sup_{v_h=(\vec{v}_h,\gamma_h)\in\mathbf{W}_h} \frac{(q_h, \operatorname{div}\vec{v}_h)}{\|v_h\|_{\tilde{\mathbf{Y}}}} \geq c\,\|q_h\|_{0,\Omega} \tag{2.115}$$

alors, à toute solution $u_h = (\vec{u}_h, \tau_h) \in \mathbf{Z}_h$ de (P_h), on peut associer une unique fonction $p_h \in Q_h$ telle que le couple $(u_h = (\vec{u}_h, \tau_h), p_h)$ soit solution de $(P_h^{'})$.

Démonstration : Soit $u_h = (\vec{u}_h, \tau_h) \in \mathbf{Z}_h$ la solution de (P_h). Il vient de la définition (2.114) de $\mathbf{Z}_h$, de l'hypothèse (2.115) et par application du théorème A.1 de Brezzi-Babuska qu'il existe un unique élément $P_h \in Q_h$ tel que $(u_h = (\vec{u}_h, \tau_h), p_h)$ soit solution de $(P_h^{'})$. □

Remarque 2.7 Estimation abstraite de l'erreur sur la pression
Sous l'hypothèse $\lambda = 2\nu\tau \in H^1(\Omega)_d^4$, et avec une technique développée par Girault et Raviart, on peut montrer que :

$$\|p - p_h\|_{0,\Omega} \leq c\left(\|u - u_h\|_{\tilde{\mathbf{Y}}} + \inf_{\mu_h\in\mathbf{M}_h}\|\lambda - \mu_h\|_{1,\Omega} + \inf_{q_h\in Q_h}\|p - q_h\|_{0,\Omega}\right)$$

où $c > 0$ est une constante indépendante de h. □

2.8 Elements finis de degré l

Proposant l'élément de Thomas-Raviart pour approcher les vitesses, nous donnons ici plusieurs approximations concrètes des contraintes. La condition de compatibilité (2.109) prend alors une forme plus faible, que nous expliciterons.

Afin de simplifier la présentation de ce paragraphe, on suppose ici que Ω est un domaine polygonal de $\mathbb{R}^2$.

Soit $(\mathcal{T}_h)_{h>0}$ une famille de triangulation de $\overline{\Omega}$, composée de triangles ou de quadrilatères de diamètre majoré par h. On suppose la famille $(\mathcal{T}_h)_{h>0}$ *uniformément régulière* lorsque $h \to 0$:

$$\tau^* h \leq h_K \leq \sigma^* \rho_K, \ \forall K \in \mathcal{T}_h \tag{2.116}$$

où h_K est le diamètre de l'élément K, ρ_K son rayon, et σ^*, τ^* sont deux constantes strictement positives.

Soit $P_l(K)$ l'espace des polynômes de degré inférieur ou égal à l, par rapport aux deux variables x_1, x_2, si K est un triangle, et par rapport à chacune des variables, si K est un quadrilatère convexe.

Pour un entier $l \geq 1$, on définit les espace de dimension finie :

$$\Theta_h = \{\theta_h \in C^0(\bar{\Omega});\ \theta_{h|K} \in P_l(K),\ \forall K \in \mathcal{T}_h\} \tag{2.117}$$

$$\Theta_{0,h} = \Theta_h \cap L^2_o(\Omega) \tag{2.118}$$

$$\Phi_h = \Theta_h \cap H^1_0(\Omega) \tag{2.119}$$

Il reste à construire des espaces de dimension finie $\mathbf{Y}_h$, Q_h et T_h de sorte que les hypothèses sur $\mathbf{V}_h$ et des conditions de compatibilité soient vérifiées, avec, de plus, de bonnes propriétés d'approximation.

2.8.1 Elements incompressibles pour les vitesses

L'approximation concrète des champs de vecteurs est inspirée de [20, 51, 54].

Considérons un élément $K \in \mathcal{T}_h$ et l'élément associé $\hat{K}$ dans le plan de référence (ξ_1, ξ_2). Alors, il existe une application inversible $F \in P_1(\hat{K})$ qui transforme $\hat{K}$ en K. On note J le déterminant de la jacobienne de F (supposé strictement positif).

On établit une correspondance bi-univoque, d'une part, entre une fonction *scalaire* φ, définie sur K, et $\hat{\varphi}$, définie sur $\hat{K}$ par la relation :

$$\hat{\varphi} = \varphi \circ F \tag{2.120}$$

et, d'autre part, entre une fonction *vectorielle* $\vec{v} = (v_1, v_2)$ définie sur K et $\vec{\hat{v}} = (\hat{v}_1, \hat{v}_2)$ définie sur $\hat{K}$ par :

$$v_i \circ F = \frac{1}{J} \sum_{j=1}^{2} \frac{\partial F_i}{\partial \xi_j} \hat{v}_j \tag{2.121}$$

L'importance de la transformation (2.121) tient à la propriété d'invariance suivante [54] :

$$\left(\sum_{i=1}^{2} \frac{\partial v_i}{\partial x_i} \right) \circ F = \frac{1}{J} \sum_{i=1}^{2} \frac{\partial \hat{v}_i}{\partial \xi_i} \tag{2.122}$$

$$\int_{\partial K} \varphi \, \vec{v}.\vec{\nu}_K \, ds = \int_{\partial \hat{K}} \hat{\varphi} \, \vec{\hat{v}}.\vec{\nu}_{\hat{K}} \, d\hat{s} \tag{2.123}$$

ceci pour tout $\varphi \in L^2(\partial K)$, $\vec{v} \in H^1(K)^2$, où $\vec{\nu}_K$ et $\vec{\nu}_{\hat{K}}$ sont les vecteurs unitaires de la normale sortante le long de ∂K et $\partial \hat{K}$, respectivement.

On introduit un espace de dimension finie $\hat{\mathcal{V}}$ de vecteurs réguliers, définis sur $\hat{K}$, et on pose :

$$Y_h = \{\vec{v}_h \in H(\mathrm{div}, \Omega); \ \vec{\hat{v}}_{h|\hat{K}} \in \hat{\mathcal{V}}, \ \forall K \in \mathcal{T}_h\} \tag{2.124}$$

$$Y_{0,h} = Y_h \cap H_0(\mathrm{div}, \Omega) \tag{2.125}$$

$$\mathbf{Y}_h = Y_{0,h} \times T_h \tag{2.126}$$

L'espace $\hat{\mathcal{V}}$ est construit comme suit :

- Si $\hat{K}$ est le triangle de référence, on définit $\hat{\mathcal{V}}$ comme étant l'espace des fonctions $\vec{\hat{v}} = (\hat{v}_1, \hat{v}_2)$ de la forme :

  $$\hat{v}_1 = p_1 + \sum_{i=0}^{l-1} \alpha_i \xi_1^{l-i} \xi_2^i \tag{2.127}$$

  $$\hat{v}_2 = p_2 + \sum_{i=0}^{l-1} \alpha_i \xi_1^{l-i-1} \xi_2^{i+1} \tag{2.128}$$

  où p_1, $p_2 \in P_{l-1}$.

  Remarquons que, si $\vec{\hat{v}} \in \hat{\mathcal{V}}$ vérifie $\mathrm{div}\,\vec{\hat{v}} = 0$, alors on obtient $\alpha_i = 0, \ 0 \leq i \leq l-1$, ce qui entraîne $\vec{\hat{v}} \in (P_{l-1})^2$ et $\vec{\hat{v}} = \vec{\mathrm{rot}}\,\hat{\varphi}$, avec $\hat{\varphi} \in P_l$.

- si $\hat{K}$ est le carré de référence, $\hat{\mathcal{V}}$ est choisi comme étant l'espace des fonctions de la forme :

  $$\hat{v}_1 \in P_{l,l-1}; \ \hat{v}_2 \in P_{l-1,l} \tag{2.129}$$

où $P_{i,j}$ est l'espace des restrictions à $\hat{K}$ de tous les polynômes de la forme :

$$p(\xi_1, \xi_2) = \sum_{\substack{0 \leq \alpha_1 \leq i \\ 0 \leq \alpha_2 \leq j}} c_{\alpha_1,\alpha_2}\, \xi_1^{\alpha_1} \xi_2^{\alpha_2} \tag{2.130}$$

De même, si $\hat{v} \in \hat{\mathcal{V}}$ vérifie $\operatorname{div} \vec{\hat{v}} = 0$, alors il existe un unique $\hat{\varphi} \in Q(\hat{K})$ tel que $\vec{\hat{v}} = \vec{\operatorname{rot}}\, \hat{\varphi}$.

Exemple 2.1 Cas des éléments de plus bas degré ($l = 1$)

Dans le cas triangulaire, les éléments de $\hat{\mathcal{V}}$ sont de la forme :

$$\begin{aligned} \hat{v}_1 &= c_1 + \alpha\xi_1 \\ \hat{v}_2 &= c_2 + \alpha\xi_2 \end{aligned}$$

et $\operatorname{div} \vec{\hat{v}} = 0$ entraîne $\alpha = 0$: $\hat{v}_1$ et $\hat{v}_2$ sont constants sur $\hat{K}$.

Dans le cas du carré de référence, les éléments de $\hat{\mathcal{V}}$ sont de la forme :

$$\begin{aligned} \hat{v}_1 &= c_1 + \alpha\xi_1 \\ \hat{v}_2 &= c_2 + \beta\xi_2 \end{aligned}$$

et $\operatorname{div} \vec{\hat{v}} = 0$ entraîne $\alpha = -\beta$. □

On vérifie que $\vec{\hat{v}} \in \hat{\mathcal{V}}$ est déterminé de façon unique par :

- Les valeurs de $\vec{\hat{v}}.\vec{\nu}_{\hat{K}}$ en l points distincts de chaque arête $\hat{S}$ de $\hat{K}$;
- les valeurs des moments :
$$\int_{\hat{K}} \hat{v}_i p_i \, d\xi$$
où $p_i \in P_{l-2}$, $1 \leq i \leq l$, si $\hat{K}$ et le triangle de référence, et $p_1 \in P_{l-2,l-1}$, $p_2 \in P_{l-1,l-2}$, si $\hat{K}$ est le carré de référence.

Dans [54] on trouve le

Lemme 2.4 *condition suffisante d'appartenance à* $H(\operatorname{div}, \Omega)$

Soit $\vec{v} \in L^2(\Omega)^2$, *régulière dans chaque élément* $K \in \mathcal{T}_h$.

Si, de plus, $\vec{v}$ *satisfait à*

$$\vec{v}_{|K_1}.\vec{\nu}_{K_1} + \vec{v}_{|K_2}.\vec{\nu}_{K_2} = 0, \text{ sur } \partial K_1 \cap \partial K_2 \tag{2.131}$$

pour tout couple d'éléments adjacents K_1, K_2, *alors* $\vec{v} \in H(\operatorname{div}, \Omega)$.

Il suffit donc de raccorder $\vec{v}_{h|K}.\vec{\nu}_K$ en $l+1$ points distincts de chaque arête pour que la relation (2.131) soit satisfaite et que $\vec{v}_h \in H(\mathrm{div}, \Omega)$. Ainsi, les degrés de liberté de $\vec{v}_h \in Y_h$ peuvent être choisis comme étant :

- Les valeurs de $\vec{v}_h.\vec{\nu}_S$ en l points distincts de chaque arête S de $\mathcal{T}_h$;
- les moments pour chaque $K \in \mathcal{T}_h$.

Remarque 2.8 Lorsque K est un triangle ou un parallélogramme, la transformation F est affine, et l'espace $\hat{\mathcal{V}}$ est invariant par F. □

Exemple 2.2 cas des éléments de plus bas degré (suite)
Les degrés de libertés de $v_h \in V_h$ pourront être les flux de v_h à travers les arêtes de $\mathcal{T}_h$. Ainsi, la dimension de Y_h est égale au nombre d'arêtes de $\mathcal{T}_h$. □

On pose ensuite :

$$Q_h = \{q_h \in L_o^2(\Omega);\ q_{h|K} \in P_{l-1}(K),\ \forall K \in \mathcal{T}_h\} \tag{2.132}$$
$$V_h = \{\vec{v}_h \in Y_{0,h};\ (\mathrm{div}\,\vec{v}_h, q_h) = 0,\ \forall q_h \in Q_h\} \tag{2.133}$$

Remarquons que $q_h \in Q_h$ n'est pas nécessairement continu aux interfaces des éléments.

Maintenant, pour un $\vec{v}_h \in V_h$ donné, on définit $\mathrm{div}_h\,\vec{v}_h$ comme étant la projection de $\mathrm{div}\,\vec{v}_h$ sur Q_h :

$$(\mathrm{div}_h\,\vec{v}_h, q_h) = (\mathrm{div}\,\vec{v}_h, q_h),\ \forall q_h \in Q_h$$

Lorsque l'application $F = F_K$ est affine (c'est-à-dire lorsque l'élément K est un triangle ou un parallélogramme), il vient de la propriété d'invariance (2.122) que :

$$(\mathrm{div}_h\,\vec{v}_h)_{|K} = \mathrm{div}\,\vec{v}_{h|K},\ \forall K \in \mathcal{T}_h$$

Ceci n'est plus vrai lorsque K est un quadrilatère général, mais la propriété suivante reste vraie [20]

$$\mathrm{div}\,\vec{v}_h = 0 \Longleftrightarrow \mathrm{div}_h\,\vec{v}_h = 0,\ \forall \vec{v}_h \in Y_h$$

En conséquence, nous avons la caractérisation suivante de V_h :

$$V_h = \{\vec{v} \in Y_h;\ \mathrm{div}\,\vec{v}_h = 0\} \subset \tilde{V} \tag{2.134}$$

Cette propriété remarquable nous assure la $\mathbf{V}_h$-ellipticité de $\tilde{a}$. Ainsi le problème (P_h) constitue une *approximation interne* du problème $(\tilde{P})$.

La vérification de la condition in-sup discrète nécéssite des outils plus sophistiqués que ceux présentés ici (consulter [54], par exemple). Cependant, lorsque la triangulation $\mathcal{T}_h$ ne comporte que des triangles et des parallélogrammes, et à l'aide de la propriété d'invariance (2.122), on obtient une nouvelle caractérisation de V_h :

$$\begin{aligned} V_h &= \{\vec{v}_h \in Y_h;\ \operatorname{div}\vec{v}_h = 0\} \\ &= \{\vec{\operatorname{rot}}\,\varphi_h;\ \varphi_h \in \Phi_h\} \end{aligned}$$

Ceci nous assure que, dans ce cas, la condition de compatibilité (2.113) entre les espaces V_h et Θ_h est satisfaite.

2.8.2 Approximations des contraintes

Cas des éléments rectangulaires

Considérons l'espace :

$$\begin{aligned} T_h^{(1)} = \{\gamma_h &= \begin{pmatrix} n_{1;h} & d_h \\ d_h & n_{2;h} \end{pmatrix} n_{1;h|K}, n_{2;h|K} \in P_{l-1}(K); \\ &\text{et } d_h \in \Theta_h\} \end{aligned} \tag{2.135}$$

Cet espace ne constitue pas une approximation conforme de $D(\vec{\operatorname{div}}, \Omega)$. En effet, la divergence d'un tenseur de $T_h^{(1)}$ n'est pas nécessairement dans $L^2(\Omega)^2$. Pour cet espace, la condition de compatibilité (2.109) n'a plus de sens, et nous la remplaçons par une forme plus faible :

$$\vec{v}_h \in V_h;\ (\vec{v}_h, \vec{\operatorname{div}}\gamma_h) = 0, \forall \gamma_h \in T_h \Longrightarrow \vec{v}_h = 0 \tag{2.136}$$

D'une part, la condition (2.109) entraîne (2.136), et d'autre part, nous avons le

Lemme 2.5
Si la condition (2.136) est satisfaite, alors le problème ($\bar{P}_h$) admet une solution unique.

Démonstration : Posons $\vec{f} = 0$, et soit $(\vec{u}_h, \tau_h)$ une solution du problème correspondant. En faisant $\vec{v}_h = \vec{u}_h$ dans (2.110), on obtient :

$$(\vec{\operatorname{div}}\,\tau_h, \vec{u}_h) = 0$$

et en remplaçant dans (2.111) avec $\gamma_h = \tau_h$, il vient $\tau_h = 0$.

Reprennons (2.110), qui s'écrit maintenent :

$$(\vec{u}_h, \vec{\operatorname{div}}\, \gamma_h) = 0, \ \forall \gamma_h \in T_h$$

D'après l'hypothèse (2.136), ceci donne $\vec{u}_h = 0$ □

Commençons par montrer le

Théorème 2.17 *d'approximation des contraintes*

Supposons la triangulation composée de rectangles à côtés parrallèles aux axes.

Les espaces V_h et $T_h = M_h$ étant définis comme en (2.102) et (2.135), respectivement, le problème (2.110)-(2.111) admet une solution unique.

Démonstration : Il suffit de vérifier que $T_h^{(2)}$ satisfait à la condition (2.136).

Pour fixer les idées, choisissons $l = 1$, le principe de la démonstration étant le même pour les éléments de degré plus élevé.

Soit $\vec{v}_h = (v_{1;h}, v_{2:h}) \in V_h$. La composante $v_{1;h}$ (resp. $v_{2;h}$) est continue par rapport à x_1 (resp. x_2) en tant que composante normale sur les arêtes verticales (resp. horizontales). Par conséquent, les dérivées $\frac{\partial v_{1;h}}{\partial x_1}$ et $\frac{\partial v_{2;h}}{\partial x_2}$ sont bien définies, en tant qu'éléments de Q_h.

Considérons un tenseur $\gamma_h \in T_h^{(2)}$ de la forme :

$$\gamma_h = \begin{pmatrix} \gamma_{11;h} & 0 \\ 0 & \gamma_{22;h} \end{pmatrix} \tag{2.137}$$

Nous avons alors :

$$(\vec{v}_h, \vec{\operatorname{div}}\, \gamma_h) = -(\frac{\partial v_{1;h}}{\partial x_1}, \gamma_{11;h}) - (\frac{\partial v_{2;h}}{\partial x_2}, \gamma_{22;h})$$

Sur un élément $K \in \mathcal{T}_h$, les composantes de $\vec{v}_h$ sont de la forme :

$$\begin{aligned} v_{1;h}(x_1, x_2) &= c_{1,K} + \alpha_K x_1 \\ v_{2;h}(x_1, x_2) &= c_{2,K} + \alpha_K x_2 \end{aligned}$$

où $c_{1,K}$, $c_{2,K}$ et $\alpha_K \in \mathbb{R}$.

Posant alors :

$$\gamma_h = \begin{pmatrix} -\alpha_K & 0 \\ 0 & \alpha_K \end{pmatrix} \in T_h^{(2)}$$

Nous obtenons immédiatement $(\vec{v}_h, \vec{\mathrm{div}}\,\gamma_h) = 0 \Longrightarrow \alpha_K = 0,\ \forall K \in \mathcal{T}_h$. D'autre part, par continuité des composantes normales le long des arêtes, et de la condition $\vec{v}_h.\vec{\nu} = 0$ sur Γ, il vient $c_{i,K} = 0,\ \forall K \in \mathcal{T}_h,\ i = 1, 2$. D'où $\vec{v}_h = 0$.

Bien entendu, ceci reste encore valable lorsque les éléments restent assez proches de parallélogramme. □

Remarque 2.9 Autres approximations des contraintes

Nous n'avons pas eu besoin d'hypothèses sur l'approximation du champ des contraintes de cisaillement d_h pour montrer que le problème $(\bar{P}_h)$ est bien-posé.

Nous avons choisi $d_h \in \Theta_h$ pour obtenir de bonnes propriétés d'interpolation avec le minimum de degrés de liberté. □

Cas général

Un choix concret pour l'espace des variables Lagrangiennes est :

$$M_h = \{\mu_h \in L^2(\Omega)^4_s;\ \mu_{h;1,.}, \mu_{h;2,.} \in Y_h\} \tag{2.138}$$

où $\mu_{h;i,.} = (\mu_{h;i,1}, \mu_{h;i,2})$ désigne le i-ième vecteur ligne de μ_h, $1 \le i \le 2$. En effet, il est clair que, si $\mu_h \in M_h$, alors $\vec{\mathrm{div}}\,\mu_h \in Q_h^2$. D'autre part, $V_h \subset \vec{\mathrm{div}}\,M_h$.

Lorsque nous souhaiterons approcher le terme de transport des contraintes (chapitre 4), il nous sera agréable d'utiliser une approximaton discontinue des composantes du tenseur des contraintes, avec l'espace :

$$T_h^{(2)} = \{\gamma_h = \begin{pmatrix} n_{1;h} & d_h \\ d_h & n_{2;h} \end{pmatrix};\ n_{1;h|K}, n_{2;h|K} \text{ et } d_{h|K} \in P_l(K).\} \tag{2.139}$$

Il est clair que $M_h \subset T_h^{(2)}$. Ainsi, le théorème 2.14 s'applique, et le problème (Q_h) admet une unique solution.

2.9 Le problème à quatre champs approché

2.9.1 Aproximation des conditions de frontière

Actuellement, si Y_h est défini par (2.124), alors l'espace des traces normales (voir page 2.2.3) des fonctions de Y_h sur Γ, noté $\gamma_\nu(Y_h)$, est l'espace des fonctions scalaires sur Γ, de degré l sur chaque arête frontière de $\mathcal{T}_h$ (ces fonctions n'étant pas nécessairement continue).

Notre problème est de construire une approximation $\vec{g}_h$ de la condition de frontière $\vec{g}$ telle que :

$$\int_\Gamma \vec{g}_h.\vec{\nu}\, d\Gamma = 0 \tag{2.140}$$

En effet, si la projection orthogonale de $\vec{g}.\vec{\nu} \in \gamma_\nu(Y)$ sur $\gamma_\nu(Y_h)$ est notée $\pi_h \vec{g}.\vec{\nu}$, on a généralement :

$$\int_\Gamma \pi_h \vec{g}.\vec{\nu}\, d\Gamma \neq 0 \tag{2.141}$$

Ceci représente un inconvénient majeur à la construction d'un relèvement à divergence nulle $\vec{u}_h$ de $\vec{g}$ dans V_h.

Afin de contourner cette difficulté, on définit $\vec{g}_h$ de la façon suivante : sur chaque arête frontière $S \subset \Gamma$, on pose :

$$(\vec{g}_h.\vec{\nu})_{|S} = (\pi_h \vec{g}.\vec{\nu})_{|S} - \frac{1}{|\Gamma|}\int_\Gamma \pi_h \vec{g}.\vec{\nu}\, d\Gamma \tag{2.142}$$

et on vérifie que (2.142) entraîne (2.140).

2.9.2 Le problème approché

Considérons le problème approché (S_h) :
trouver $(\tau_h, \lambda_h, \omega_h, \vec{u}_h, p_h) \in T_h \times M_h \times \Theta_h \times Y_h \times Q_{0,h}$ *tel que,*
pour tout $(\gamma_h, \mu_h, \theta_h, \vec{v}_h, q_h) \in T_h \times M_h \times \Theta_h \times Y_{0,h} \times Q_h$*, on ait :*

$$(\tau_h, \gamma_h) - (\gamma_h, \lambda_h) = 0 \tag{2.143}$$

$$-(\tau_h, \mu_h) - 2\alpha(\vec{u}_h, \vec{\operatorname{div}}\, \mu_h) = 2\alpha \int_\Gamma (\vec{g}_h \otimes \vec{\nu}) : \mu_h\, d\Gamma \tag{2.144}$$

$$(\omega_h, \theta_h) - (1-\alpha)(\vec{u}_h, \vec{\operatorname{rot}}\, \theta_h) = (1-\alpha)\int_\Gamma (\vec{g}_h.\vec{\tau})\theta_h\, d\Gamma \tag{2.145}$$

$$(\vec{v}_h, \vec{\operatorname{div}}\, \lambda_h) - (\vec{v}_h, \vec{\operatorname{rot}}\, \omega_h) + (p_h, \operatorname{div} \vec{v}_h) = -(\vec{f}, \vec{v}_h) \tag{2.146}$$

$$(q_h, \operatorname{div} \vec{u}_h) = 0 \tag{2.147}$$

$$\vec{u}_h.\vec{\nu} = \vec{g}_h.\vec{\nu} \text{ sur } \Gamma \tag{2.148}$$

$$\tag{2.149}$$

Montrons le

Théorème 2.18 *Existence et unicité de la solution de* (S_h)

Supposons que les espaces T_h*,* M_h*,* Θ_h*,* Y_h *et* Q_h *satisfont aux conditions de compatibilité : (2.108) (2.113), et (2.115).*

Alors le problème (S_h) *admet une solution unique.*

Démonstration : Si $\alpha = 0$ ou 1, ce travail a déjà été accompli, et nous supposerons donc $0 < \alpha < 1$. Posons donc $\vec{f} = 0$ et $\vec{g}_h = 0$, et soit $(\tau_h, \lambda_h, \omega_h, \vec{u}_h, p_h)$ une solution du problème correspondant. En faisant $\gamma_h = \tau_h$ dans (2.143) et $\mu_h = \lambda_h$ dans (2.144), il vient $\|\tau_h\|_{0,\Omega}^2 = -2\alpha(\vec{u}_h, \vec{\operatorname{div}}\, \lambda_h)$. De façon similaire, le choix $\theta_h = \omega_h$ dans (2.144) permet d'obtenir $\|\omega_h\|_{0,\Omega}^2 = (1-\alpha)(\vec{\operatorname{rot}}\, \omega_h, \vec{u}_h)$. Enfin, (2.146) conduit à $(p_h, \operatorname{div} \vec{u}_h) = 0$. Choisissons alors $\vec{v}_h = \vec{u}_h$ dans (2.145). De ce qui précéde, on obtient :

$$\frac{1}{2\alpha}\|\tau_h\|_{0,\Omega}^2 + \frac{1}{1-\alpha}\|\omega_h\|_{0,\Omega}^2 = 0$$

d'où $\tau_h = 0$ et $\omega_h = 0$.

Le reste de la démonstration est identique à celle du lemme 2.5. □

Remarque 2.10 Cas où $M_h = T_h$

Dans le cas où $M_h = T_h$, il est possible d'éliminer le multiplicateur de Lagrange λ_h du problème (S_h). (voir remarque 2.5).

Le problème obtenu apparaît comme une approximation à posteriori de la formulation variationelle suivante de (2.15)-(2.19) :
trouver $(\tau, \omega, \vec{u}, p) \in D(\vec{\operatorname{div}}, \Omega) \times H^1(\Omega) \times H(\operatorname{div}, \Omega) \times L_o^2(\Omega)$ *tel que,*
pour tout $(\gamma, \theta, \vec{v}, p) \in D(\vec{\operatorname{div}}, \Omega) \times H^1(\Omega) \times H_0(\operatorname{div}, \Omega) \times L^2(\Omega)$*, on ait :*

$$(FVS) : \left\{ \begin{array}{rcl} (\tau, \gamma) + 2\alpha(\vec{u}, \vec{\operatorname{div}}\, \gamma) & = & 0 \\ (\omega, \theta) - (1-\alpha)(\vec{u}, \vec{\operatorname{rot}}\, \theta) & = & (1-\alpha) \int_\Gamma \vec{g}.\vec{\tau}\, d\Gamma \\ (\vec{\operatorname{div}}\, \tau, \vec{v}) - (\vec{\operatorname{rot}}\, \omega, \vec{v}) + (p, \operatorname{div} \vec{v}) & = & -(\vec{f}, \vec{v}) \\ (\operatorname{div} \vec{u}, q) & = & 0 \\ \vec{u}.\vec{\nu} & = & \vec{g}.\vec{\nu} \text{ sur } \Gamma \end{array} \right.$$

□

Remarque 2.11 Choix concret des espaces d'approximation

De ce qui précède, nous en déduisons que lorsque les espaces T_h et M_h sont donnés par (2.139), (2.138) ou bien que $M_h = T_h$ est donné par (2.135) et que les espaces Θ_h, Y_h et Q_h, sont donnés respectivement par (2.117), (2.124) et (2.132), alors le problème (S_h) admet une solution unique. □

2.9.3 Formulation matricielle; élimination du tourbillon

Afin de simplifier la prśentation de ce paragraphe, nous supposerons que les espaces M_h et T_h coïncident.

Soient

$$\begin{aligned} N_\tau &= \dim T_h \\ N_\omega &= \dim \Theta_h \\ N_u &= \dim Y_{0,h} \\ N_p &= \dim Qh \end{aligned} \tag{2.150}$$

les dimensions des espaces approchés, et $(\gamma_h^{(i)})_{1\le i\le N_\tau}$, $(\theta_h^{(i)})_{1\le i\le N_\omega}$, $(\vec{v}_h^{(i)})_{1\le i\le N_u}$, $(q_h^{(i)})_{1\le i\le N_p}$, les bases associées (construites à partir des degrés de libertés des élements).

On introduit les matrices B, R et D, définies par :

$$B_{ij} = (\vec{v}^{(i)}, \vec{\operatorname{div}}\, \gamma_h^{(j)}),\ 1 \le i \le N_u,\ 1 \le j \le N_\tau \tag{2.151}$$

$$R_{ij} = -(\vec{v}^{(i)}, \vec{\operatorname{rot}}\, \theta_h^{(j)}),\ 1 \le i \le N_u,\ 1 \le j \le N_\omega \tag{2.152}$$

$$D_{ij} = (q^{(i)}, \operatorname{div} \vec{v}_h^{(j)}),\ 1 \le i \le N_p,\ 1 \le j \le N_u \tag{2.153}$$

associées respectivement aux opérateurs $\vec{\operatorname{div}}$, $-\vec{\operatorname{rot}}$ et div, ainsi que les matrices de masse M_τ et M_ω :

$$M_{\tau;ij} = (\gamma_h^{(i)}, \gamma_h^{(j)}),\ 1 \le i,j \le N_\tau \tag{2.154}$$

$$M_{\omega;ij} = (\theta_h^{(i)}, \theta_h^{(j)}),\ 1 \le i,j \le N_\omega \tag{2.155}$$

Le problème (S_h) peut se mettre sous la forme matricielle :

$$\begin{pmatrix} \frac{1}{2\alpha}M_\tau & 0 & B^t & 0 \\ 0 & \frac{1}{1-\alpha}M_\omega & R^t & 0 \\ B & R & 0 & D^t \\ 0 & 0 & D & 0 \end{pmatrix} \begin{pmatrix} \tau \\ \omega \\ u \\ p \end{pmatrix} = \begin{pmatrix} \tau_b \\ \omega_b \\ 0 \\ p_b \end{pmatrix} \tag{2.156}$$

où τ, ω, u et p sont les vecteurs des composantes de τ_h, ω_h, $\vec{u}_h$ et p_h dans les bases précédentes, et τ_b, ω_b et p_b sont les seconds membres issus de la condition à la frontière (2.147).

Il est alors possible d'éliminer le tourbillon ω du système (2.156) pour obtenir

$$\begin{pmatrix} \frac{1}{2\alpha}M_\tau & B^t & 0 \\ B & C & D^t \\ 0 & D & 0 \end{pmatrix} \begin{pmatrix} \tau \\ u \\ p \end{pmatrix} = \begin{pmatrix} \tau_b \\ \omega_b \\ u_b \\ p_b \end{pmatrix} \tag{2.157}$$

où on a posé :

$$C = -(1-\alpha)R.(M^\omega)^{-1}R^t \quad (2.158)$$
$$u_b = -(1-\alpha)R.(M^\omega)^{-1}\omega_b \quad (2.159)$$

Remarque 2.12 matrices de masse diagonales

En utilisant des formules de quadratures appropriées (Mass-Dumping, en anglais), il est possible de rendre les matrices M_τ et M_ω diagonales, ce qui rend leur inversion effective plus aisée. □

2.9.4 Perturbation du système linéaire

Lorsque nous introduirons les termes de transport des contraintes, traduisant l'effet-mémoire du fluide, le système prendra la forme :

$$\begin{pmatrix} A(u) & B^t & 0 \\ B & C & D^t \\ 0 & D & 0 \end{pmatrix} \begin{pmatrix} \tau \\ u \\ p \end{pmatrix} = \begin{pmatrix} \tau_b \\ \omega_b \\ u_b \\ p_b \end{pmatrix} \quad (2.160)$$

où

$$A(u) = \frac{1}{2\alpha}M_\tau + \frac{We}{2\alpha}T(u) \quad (2.161)$$

et la matrice de perturbation $T(u)$ est définie par :

$$T(u)_{ij} = (\gamma^{(i)}, (\vec{u}_h.\vec{\nabla})_h\, \gamma_h^{(j)} + \beta(\gamma_h^{(j)}, \vec{\nabla}\,\vec{u}_h)),\ 1 \le i,j \le N_\tau \quad (2.162)$$

où $\vec{u}_h \in Y_h$ est le champ de vecteur de composantes $u = (u_1, \dots, u_{N_u})$ dans la base $(\vec{v}_h^{(i)})_{1\le i\le N_u}$. L'opérateur $(\vec{u}_h.\vec{\nabla})_h$ représente une approximation de l'opérateur de transport $(\vec{u}_h.\vec{\nabla})$. Le chapitre 4 est consacré à l'étude de cette approximation.

2.9.5 Lien avec la méthode des différences finies

Avant de terminer ce chapitre, il sera intéressant de noter que la méthode d'éléments finis étudiée n'est autre qu'une généralisation d'un schéma aux différences finies. Ce schéma apparaît lui même comme une extension au problème de Stokes à trois champs d'un schéma très classique pour le problème de Stokes [46] ou le problème de Navier-Stokes parabolisé [44].

Pour cela, considérons un maillage rectangulaire uniforme, de pas (h_1, h_2) dont la maille est représenté sur la figure 2.1, avec les notations :

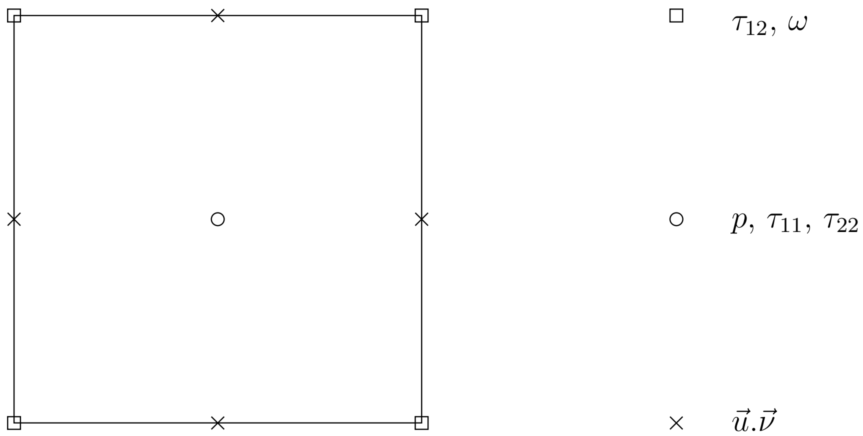


Figure 2.1: Interprétation en différences finies de la méthode

$$\tau_h = \begin{pmatrix} \tau_{11;h} & \tau_{12;h} \\ \tau_{12;h} & \tau_{22;h} \end{pmatrix}$$

$$\begin{aligned} \vec{u}_h &= (u_h, v_h) \\ f &= (f^{(1)}, f^{(2)}) \end{aligned}$$

$$\begin{aligned}
f^{(k)}_{i,j} &= f^{(k)}(ih_1, jh_2),\ \ k = 1, 2 \\
\tau_{11;i+\frac{1}{2},j+\frac{1}{2}} &= \tau_{11;h}\big((i+\tfrac{1}{2})h_1, (j+\tfrac{1}{2})h_2\big) \\
\tau_{12;i,j} &= \tau_{12;h}(ih_1, jh_2) \\
\tau_{22;i+\frac{1}{2},j+\frac{1}{2}} &= \tau_{22;h}\big((i+\tfrac{1}{2})h_1, (j+\tfrac{1}{2})h_2\big) \\
\omega_{i,j} &= \omega_h(ih_1, jh_2) \\
u_{i,j+\frac{1}{2}} &= u_h\big(ih_1, (j+\tfrac{1}{2})h_2\big) \\
v_{i+\frac{1}{2},j} &= v_h\big((i+\tfrac{1}{2})h_1, jh_2\big) \\
p_{i+\frac{1}{2},j+\frac{1}{2}} &= p_h\big((i+\tfrac{1}{2})h_1, (j+\tfrac{1}{2})h_2\big)
\end{aligned}$$

En utilisant la formule des trapèzes pour évaluer les intégrales intervenant dans les produits scalaires :

$$\int_K f\,dK \approx \frac{h_1 h_2}{4}\left(f_{i,j} + f_{i+1,j} + f_{i,j+1} + f_{i+1,j+1}\right)$$

on obtient un schéma aux différences finies sur une maille intérieure au domaine, pour le problème de Stokes à trois champs :

$$\tau_{11;i+\frac{1}{2},j+\frac{1}{2}} = -\tau_{22;i+\frac{1}{2},j+\frac{1}{2}} = 2\alpha\frac{u_{i+1,j+\frac{1}{2}} - u_{i,j+\frac{1}{2}}}{h_1} \tag{2.163}$$

$$\tau_{12;i,j} = \alpha\left(\frac{v_{i+\frac{1}{2},j} - v_{i-\frac{1}{2},j}}{h_1} + \frac{u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}}}{h_2}\right) \tag{2.164}$$

$$\omega_{i,j} = \alpha\left(\frac{v_{i+\frac{1}{2},j} - v_{i-\frac{1}{2},j}}{h_1} + \frac{u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}}}{h_2}\right) \tag{2.165}$$

$$\begin{aligned} &-\frac{\tau_{11;i+\frac{1}{2},j+\frac{1}{2}} - \tau_{11;i-\frac{1}{2},j+\frac{1}{2}}}{h_1} - \frac{\tau_{12;i,j+1} - \tau_{i,j}}{h_2} \\ &+ \frac{\omega_{i,j+1} - \omega_{i,j}}{h_2} + \frac{p_{i+\frac{1}{2},j+\frac{1}{2}} - p_{i-\frac{1}{2},j+\frac{1}{2}}}{h_1} = f^{(1)}_{i,j+\frac{1}{2}} \end{aligned} \tag{2.166}$$

$$\begin{aligned} &-\frac{\tau_{12;i+1,j} - \tau_{i,j}}{h_1} - \frac{\tau_{22;i+\frac{1}{2},j+\frac{1}{2}} - \tau_{22;i+\frac{1}{2},j-\frac{1}{2}}}{h_2} \\ &- \frac{\omega_{i+1,j} - \omega_{i,j}}{h_1} + \frac{p_{i+\frac{1}{2},j+\frac{1}{2}} - p_{i+\frac{1}{2},j-\frac{1}{2}}}{h_2} = f^{(2)}_{i+\frac{1}{2},j} \end{aligned} \tag{2.167}$$

$$\frac{u_{i+1,j+\frac{1}{2}} - u_{i,j+\frac{1}{2}}}{h_1} + \frac{v_{i+\frac{1}{2},j+1} - v_{i+\frac{1}{2},j}}{h_2} = 0 \tag{2.168}$$

En combinant les équations (2.163) à (2.168), nous obtenons en particulier :

$$\begin{aligned} &\frac{-u_{i+1,j+\frac{1}{2}} + 2u_{i,j+\frac{1}{2}} - u_{i-1,j+\frac{1}{2}}}{h_1^2} + \frac{-u_{i,j+\frac{3}{2}} + 2u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}}}{h_2^2} \\ &+ \frac{p_{i+\frac{1}{2},j+\frac{1}{2}} - p_{i-\frac{1}{2},j+\frac{1}{2}}}{h_1} = f^{(1)}_{i,j+\frac{1}{2}} \end{aligned} \tag{2.169}$$

$$\begin{aligned} &\frac{-v_{i+\frac{3}{2},j} + 2v_{i+\frac{1}{2},j} - v_{i-\frac{1}{2},j}}{h_1^2} + \frac{-v_{i+\frac{1}{2},j+1} + 2v_{i+\frac{1}{2},j} - v_{i+\frac{1}{2},j-1}}{h_2^2} \\ &+ \frac{p_{i+\frac{1}{2},j+\frac{1}{2}} - p_{i+\frac{1}{2},j-\frac{1}{2}}}{h_2} = f^{(2)}_{i+\frac{1}{2},j} \end{aligned} \tag{2.170}$$

$$\frac{u_{i+1,j+\frac{1}{2}} - u_{i,j+\frac{1}{2}}}{h_1} + \frac{v_{i+\frac{1}{2},j+1} - v_{i+\frac{1}{2},j}}{h_2} = 0 \tag{2.171}$$

Ce dernier système est n'est autre que le très classique schéma “Marcker and Cell” [57] (schéma aux grilles entrelacées) appliqué ici au problème de Stokes. Nous savons que ce schéma est en $\mathcal{O}(h^2)$ pour la norme L^2 de l'erreur, et se prête très bien à une résolution par méthodes multigrilles (voir [46], par exemple).

Remarquons enfin que la solution approchée (u_h, p_h) de ce dernier problème est indépendante du paramètre α. Si cette propriété peut paraître immédiate pour une solution du problème continu (2.1)-(2.4), il est tout à fait remarquable que ceci puisse être encore vérifié par la solution du problème approché.

Chapitre 3

Approximation en temps par le θ-schéma

L'idée-clé de la méthode présentée ici est d'utiliser la décomposition associée à la méthode des directions alternées pour *découpler* les deux principales difficultées du problème, à savoir la non-linéarité de l'équation (1.17) et la relation d'incompressibilité (1.19).

Il s'agit d'une application de la méthode originale des *directions alternées* dûe à de Peaceman et Rachford [48]. Une démarche similaire a été utilisée par Glowinski, Mantel et Périaux [23, 22] pour approcher la solution stationnaire des équations de Navier-Stokes.

Le découplage de ces difficultées permet de résoudre le problème de façon bien plus efficace en diminuant le temps de calcul et en augmentant la précision des résultats numériques. L'algorithme obtenu est entièrement nouveau, et conduit à certaines reflexions sur les valeurs pouvant être prises par les paramètres de la méthode, et par les coûts asympotiques des calculs.

Comme d'habitude (voir page 15), l'écoulement sera supposé lent, et nous ne considèrerons pas le terme d'inertie dans (1.18).

3.1 Présentation du θ-schéma

Soit H un espace de Hilbert sur $\mathbb{R}$. On considère l'opérateur A, continu, défini de H dans H', et le problème :
trouver $u \in L^\infty(H, \mathbb{R})$ *solution de* :

$$\begin{aligned} m\frac{\mathrm{d}u}{\mathrm{d}t} + A(u) &= 0 & (3.1)\\ u(0) &= u_o & (3.2) \end{aligned}$$

où $u_o \in H$ et m est réel (constant).

A une décomposition de l'opérateur A de la forme :

$$A = A_1 + A_2 \tag{3.3}$$

nous associons la suite $(u^n)_{n\in\mathbb{N}},\ u^n \in H$ définie par récurrence :

$$u^0 = u_o \tag{3.4}$$

et pour tout $n \geq 0$, u^{n+1} est defini (implicitement) à partir de u^n par les trois relations :

$$m\frac{u^{n+\theta}-u^n}{\theta\,\Delta t} \quad +A_1u^{n+\theta} \quad = -A_2u^n \tag{3.5}$$

$$m\frac{u^{n+1-\theta}-u^{n+\theta}}{(1-2\theta)\Delta t} \quad +A_2u^{n+1-\theta} \quad = -A_1u^{n+\theta} \tag{3.6}$$

$$m\frac{u^{n+1}-u^{n+1-\theta}}{\theta\,\Delta t} \quad +A_1u^{n+1} \quad = -A_2u^{n+1-\theta} \tag{3.7}$$

en ayant décomposé le pas de temps $(t, t+\Delta t)$, $\Delta t > 0$, en trois fractions $(t, t+\theta\,\Delta t)$; $(t+\theta\,\Delta t, t+(1-\theta)\Delta t)$; $(t+(1-\theta)\Delta t, t+\Delta t)$, où $\theta \in]0, \frac{1}{2}[$.

Lorsque A_1 et A_2 sont linéaires, $\theta \in\ ]0,1[$ et $\Delta t > 0$, cette méthode de directions alternées, dite *θ-schéma*, est inconditionnellement stable, et l'erreur de troncature est en $O(\Delta t^2)$.

Cette méthode paraît en particulier bien adaptée à la recherche des solutions stationnaires d'un problème aux limites (nous pouvons choisir Δt arbitrairement grand).

3.2 Application au problème d'Oldroyd

3.2.1 Choix d'une décomposition

Considérons la décomposition :

$$A_1(\tau,\vec{u},p) = \begin{bmatrix} \frac{\omega}{2\alpha}\tau & -D(\vec{u}) & \\ \vec{\mathrm{div}}\tau & +(1-\alpha)\Delta\vec{u} & -\vec{\nabla}p \\ & \mathrm{div}\vec{u} & \end{bmatrix} \tag{3.8}$$

où l'opérateur Δ est associé à la condition au bord de type Dirichlet (1.10),

$$A_2(\tau,\vec{u},p) = \begin{bmatrix} \frac{We}{2\alpha}((\vec{u}.\vec{\nabla})\tau + \beta(\tau,\vec{\nabla}\vec{u})) + \frac{1-w}{2\alpha}\tau \\ 0 \\ 0 \end{bmatrix} \tag{3.9}$$

où l'opérateur $(\vec{u}.\vec{\nabla})$ est associé à la condition à la frontière amont (1.11) et ω est un paramètre de relaxation associé à cette décomposition.

La matrice diagonale réelle m est définie par :

$$m = \text{diag}(\frac{We}{2\alpha}, -Re, 0) \tag{3.10}$$

3.2.2 Les sous-problèmes obtenus

Il est alors possible de définir (formellement) une suite $(\tau^n, \vec{u}^n, p^n)_{n\in\mathbb{N}}$ par :

$$(\tau^0, \vec{u}^0, p^0) = (\tau_o, \vec{u}_o, p_o) \tag{3.11}$$

et où, $\forall n \geq 0$, le terme $(\tau^{n+1}, \vec{u}^{n+1}, p^{n+1})$ est défini implicitement en trois étapes :

première étape : τ^n et u^n étant connus, on cherche $(\tau^{n+\theta}, \vec{u}^{n+\theta}, p^{n+\theta})$ solution de :

$$\begin{aligned}
\frac{We}{2\alpha}\frac{\tau^{n+\theta} - \tau^n}{\theta\,\Delta t} &+ \frac{\omega}{2\alpha}\tau^{n+\theta} - D(\vec{u}^{n+\theta}) && (3.12)\\
&= -\frac{We}{2\alpha}((\vec{u}^n.\vec{\nabla})\tau^n + \beta(\tau^n, \vec{\nabla}\vec{u}^n)) - \frac{1-\omega}{2\alpha}\tau^n &&\\
-Re\frac{\vec{u}^{n+\theta} - \vec{u}^n}{\theta\,\Delta t} &+ \vec{\text{div}}\,\tau^{n+\theta} + (1-\alpha)\Delta\vec{u}^{n+\theta} - \vec{\nabla}p^{n+\theta} = 0 && (3.13)\\
\text{div}\,\vec{u}^{n+\theta} &= 0 && (3.14)\\
\vec{u}^{n+\theta} &= \vec{u}_\Gamma((n+\theta)\Delta t) \text{ sur } \Gamma && (3.15)
\end{aligned}$$

deuxième étape : $(\tau^{n+\theta}, \vec{u}^{n+\theta}, p^{n+\theta})$ étant connu, chercher $(\tau^{n+1-\theta}, \vec{u}^{n+1-\theta})$ solution de :

$$\begin{aligned}
\frac{We}{2\alpha}\frac{\tau^{n+1-\theta}-\tau^{n+\theta}}{(1-2\theta)\Delta t} &+ \frac{We}{2\alpha}((\vec{u}^{n+1-\theta}.\vec{\nabla})\tau^{n+1-\theta} + \beta(\tau^{n+1-\theta}, \vec{\nabla}\vec{u}^{n+1-\theta})) \\
+\frac{1-\omega}{2\alpha}\tau^{n+1-\theta} &= -\frac{\omega}{2\alpha}\tau^{n+\theta} + D(\vec{u}^{n+\theta}) \qquad (3.16) \\
-Re\frac{\vec{u}^{n+1-\theta}-\vec{u}^{n+\theta}}{(1-2\theta)\Delta t} &= -\vec{\mathrm{div}}\tau^{n+\theta} - (1-\alpha)\Delta\vec{u}^{n+\theta} + \vec{\nabla}p^{n+\theta} \qquad (3.17) \\
\tau^{n+1-\theta} &= \tau_\Gamma((n+1-\theta)\Delta t) \text{ sur } \Gamma_- \qquad (3.18)
\end{aligned}$$

troisième étape (symétrisation) : $\tau^{n+1-\theta}$ et $u^{n+1-\theta}$ étant connus, on cherche $(\tau^{n+1}, \vec{u}^{n+1}, p^{n+1})$ solution de :

$$\begin{aligned}
\frac{We}{2\alpha}\frac{\tau^{n+1}-\tau^{n+1-\theta}}{\theta\,\Delta t} &+ \frac{\omega}{2\alpha}\tau^{n+1} - D(\vec{u}^{n+1}) \qquad (3.19) \\
&= -\frac{We}{2\alpha}((\vec{u}^{n+1-\theta}.\vec{\nabla})\tau^{n+1-\theta} \\
&+ \beta(\tau^{n+1-\theta}, \vec{\nabla}\vec{u}^{n+1-\theta})) - \frac{1-\omega}{2\alpha}\tau^{n+1-\theta} \\
-Re\frac{\vec{u}^{n+1}-\vec{u}^{n+1-\theta}}{\theta\,\Delta t} &+ \vec{\mathrm{div}}\tau^{n+1} + (1-\alpha)\Delta\vec{u}^{n+1} - \vec{\nabla}p^{n+1} = 0 \qquad (3.20) \\
\mathrm{div}\,\vec{u}^{n+1} &= 0 \qquad (3.21) \\
\vec{u}^{n+1} &= \vec{u}_\Gamma((n+1)\Delta t) \text{ sur } \Gamma \qquad (3.22)
\end{aligned}$$

3.3 Découplage du calcul des composantes

En fait, les principales difficultées de l'approximation numérique proviennent en particulier :

- du traitement de la contrainte d'incompressibilité (1.19) et du terme non-linéaire de transport $(\vec{u}.\vec{\nabla})\tau$;
- du fait que les composantes τ et $\vec{u}$ soient couplées par le terme de transport et par les équations (1.17) et (1.19).

L'intérêt de la méthode vient de la possibilité de découpler ces difficultés.

3.3.1 Première étape : sous-problème de type Stokes

Remarquons tout d'abord que $\tau^{n+\theta}$ s'exprime en fonction de τ^n, $\vec{u}^n$ et $\vec{u}^{n+\theta}$ d'après la relation (3.12) selon :

$$\begin{aligned} \tau^{n+\theta} = \frac{1}{We + \omega\theta\,\Delta t} & \quad (\quad (We - (1-\omega)\theta\,\Delta t)\tau^n \\ & \quad - \quad (We\,\theta\,\Delta t)\gamma^n + 2\alpha\theta\,\Delta t\,D(\vec{u}^{n+\theta})\ ) \end{aligned} \tag{3.23}$$

où :

$$\gamma^n = (\vec{u}^n.\vec{\nabla})\tau^n + \beta(\tau^n, \vec{\nabla}\vec{u}^n) \tag{3.24}$$

Remarquons que la relation d'incompressibilité (3.14) sur $\vec{u}^{n+\theta}$ permet d'établir l'identité :

$$2\,\vec{\mathrm{div}}\,D(\vec{u}^{n+\theta}) = \Delta\vec{u}^{n+\theta} \tag{3.25}$$

Substituons alors $\tau^{n+\theta}$ dans (3.13) par son expression donnée en (3.23), et utilisons l'identité précédente.

Nous obtenons ainsi une relation de la forme :

$$\lambda\vec{u}^{n+\theta} - \eta\Delta\vec{u}^{n+\theta} = \vec{f}^{n+\theta} \tag{3.26}$$

où nous avons posé successivement :

$$\lambda = \frac{Re}{\theta\,\Delta t} \tag{3.27}$$

$$\eta = 1 - \alpha\frac{We - (1-\omega)\theta\,\Delta t}{We + \omega\theta\,\Delta t} \tag{3.28}$$

et :

$$\begin{aligned}\vec{f}^{n+\theta} &= \lambda\vec{u}^n \\ &+ \frac{1}{We+\omega\theta\,\Delta t}\Big((We-(1-\omega)\theta\,\Delta t)\,\vec{\operatorname{div}}\tau^n \\ &- (We\theta\,\Delta t)\,\vec{\operatorname{div}}\gamma^n\Big)\end{aligned} \tag{3.29}$$

Or $\vec{f}^{n+\theta}$ se calcule *explicitement* (via (3.24)) en fonction de τ^n et $\vec{u}^n$, connus à l'étape 1. Ainsi, la recherche de $\vec{u}^{n+\theta}$ et $p^{n+\theta}$ satisfaisant les équations (3.26), (3.14) et (3.15) constitue un sous-problème de type Stokes. Remarquons que le paramètre η est strictement positif pour toute valeur de $\alpha \in\ [0,1]$, $\theta \in\ ]0,\frac{1}{2}[$, $\omega \in\ ]0,1[$, $We \geq 0$ et $\Delta t > 0$.

Par conséquent, ce sous-problème est bien-posé, et $\vec{u}^{n+\theta}$, $p^{n+\theta}$ sont ainsi définis de façon unique. De plus nous savons approcher ces grandeurs par des méthodes performantes (voir l'annexe B).

Enfin, $\tau^{n+\theta}$ s'obtient de façon *explicite* à partir de la relation (3.23). Nous avons ainsi construit $(\tau^{n+\theta}, \vec{u}^{n+\theta}, p^{n+\theta})$ à partir de τ^n et u^n, et ainsi résolu l'étape 1.

3.3.2 Deuxième étape : sous-problème de type transport

Passons à présent à la seconde étape. De (3.13) et (3.17), et puisque $Re > 0$, nous obtenons :

$$\frac{\vec{u}^{n+\theta}-\vec{u}^n}{\theta\,\Delta t} = \frac{\vec{u}^{n+1-\theta}-\vec{u}^{n+\theta}}{(1-2\theta)\Delta t}$$

d'où nous en déduisons une expression *explicite* de $\vec{u}^{n+1-\theta}$:

$$\vec{u}^{n+1-\theta} = \frac{1-\theta}{\theta}\,\vec{u}^{n+\theta} - \frac{1-2\theta}{\theta}\,\vec{u}^n \tag{3.30}$$

où les grandeurs $\vec{u}^n$ et $\vec{u}^{n+\theta}$ sont déjà connues, respectivement à l'étape 1 et 2. Par conséquent, $\vec{u}^{n+1-\theta}$ est connu.

La relation (3.16) peut se mettre sous la forme :

$$(\vec{u}^{n+1-\theta}.\vec{\nabla})\tau^{n+1-\theta} + \beta(\tau^{n+1-\theta}, \vec{\nabla}\vec{u}^{n+1-\theta}) + \nu\tau^{n+1-\theta} = g^{n+1-\theta} \tag{3.31}$$

avec

$$\nu = \frac{We+(1-\omega)(1-2\theta)\Delta t}{We(1-2\theta)\Delta t} \tag{3.32}$$

$$g^{n+1-\theta} = \left(\frac{We-\omega(1-2\theta)\Delta t}{We(1-2\theta)\Delta t}\right)\tau^{n+\theta} + \frac{2\alpha}{We}D(\vec{u}^{n+\theta}) \tag{3.33}$$

Or $g^{n+1-\theta}$ se calcule directement à partir de $\tau^{n+\theta}$ et $\vec{u}^{n+\theta}$ qui sont maintenant connus. Par conséquent, la recherche de $\tau^{n+1-\theta}$ à partir de (3.31) et (3.18) constitue un système linéaire du premier ordre ($\nu > 0$).

Sous certaines hypothèses sur les valeurs des paramètres Δt et ω (voir au chapitre 4, page 89), ce sous-problème admet une solution. et $\tau^{n+1-\theta}$ est bien défini.

Il est ainsi possible (modulo un bon choix de ω et Δt) de résoudre l'étape 2, et de construire $(\tau^{n+1-\theta}, \vec{u}^{n+1-\theta})$.

3.3.3 Troisième étape : symétrisation

Nous construirons $(\tau^{n+1}, \vec{u}^{n+1}, p^{n+1})$ à partir de $(\tau^{n+1-\theta}, \vec{u}^{n+1-\theta})$ en résolvant l'étape 3 de la même façon que l'étape 1 (en remplaçant les indices $(n+\theta)$ et n par $(n+1)$ et $(n+1-\theta)$, respectivement, dans les relations (3.23) à (3.29)).

3.3.4 Conclusion

Ainsi, il nous a été possible de construire $(\tau^{n+1}, \vec{u}^{n+1}, p^{n+1})$ à partir de $(\tau^{n}, \vec{u}^{n}, p^{n})$ en utilisant une succession d'opérateurs linéaires. D'autre part, puisque $(\tau^{0}, \vec{u}^{0}, p^{0})$ est connu à l'aide de (3.11), nous pouvons construire, par récurrence sur n, $(\tau^{n}, \vec{u}^{n}, p^{n})$ pour tout $n \in \mathbb{N}$.

3.4 Obtention d'un algorithme

3.4.1 Ecriture de l'algorithme

Des remarques qui précèdent, nous en déduisons un algorithme de construction de $(\tau^{n+1}, \vec{u}^{n+1}, p^{n+1})$ à partir de $(\tau^{n}, \vec{u}^{n}, p^{n})$. Dans la suite, le symbole := désignera un calcul explicite de l'expression de droite, affecté à la grandeur située à gauche.

étape 1 : τ^n et $\vec{u}^n$ sont connus

- calculer successivement :

$$\begin{aligned}
\gamma^n &:= (\vec{u}^n.\vec{\nabla})\tau^n + \beta(\tau^n, \vec{\nabla}\vec{u}^n) && (3.34)\\
\vec{f}^{n+\theta} &:= \lambda\vec{u}^n + \left(\frac{We - (1-\omega)\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right) \vec{\mathrm{div}}\,\tau^n \\
&\quad - \left(\frac{We\,\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right) \vec{\mathrm{div}}\,\gamma^n && (3.35)
\end{aligned}$$

- le couple $(\vec{u}^{n+\theta}, p^{n+\theta})$ est caractérisé comme étant l'unique solution du problème de type Stokes :

$$\lambda\vec{u}^{n+\theta} - \eta\Delta\vec{u}^{n+\theta} + \vec{\nabla}p^{n+\theta} = \vec{f}^{n+\theta} \text{ dans } \Omega \tag{3.36}$$

$$\operatorname{div} u^{n+\theta} = 0 \text{ dans } \Omega \tag{3.37}$$

$$\vec{u}^{n+\theta} = \vec{u}_\Gamma((n+\theta)\Delta t) \text{ sur } \Gamma \tag{3.38}$$

  où λ et η sont donnés respectivement par (3.27) et (3.28).

- calculer :

$$\tau^{n+\theta} := \left(\frac{We - (1-\omega)\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)\tau^n - \left(\frac{We\,\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)\gamma^n + \left(\frac{2\,\alpha\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right) D(\vec{u}^{n+\theta}) \tag{3.39}$$

étape 2 : $\tau^{n+\theta}$, $\vec{u}^{n+\theta}$, $\vec{u}^n$ et $D(\vec{u}^{n+\theta})$ étant connus,

- calculer successivement :

$$g^{n+1\theta} := \left(\frac{We - \omega(1-2\theta)\Delta t}{We(1-2\theta)\Delta t}\right)\tau^{n+\theta} + \frac{2\alpha}{We}D(\vec{u}^{n+\theta}) \tag{3.40}$$

$$\vec{u}^{n+1-\theta} := \frac{1-\theta}{\theta}\,\vec{u}^{n+\theta} - \frac{1-2\theta}{\theta}\,\vec{u}^n \tag{3.41}$$

- $\tau^{n+1-\theta}$ est solution de :

$$(\vec{u}^{n+1-\theta}.\vec{\nabla})\,\tau^{n+1-\theta} + \beta(\tau^{n+1-\theta}, \vec{\nabla}\vec{u}^{n+1-\theta}) + \nu\tau^{n+1-\theta} = g^{n+1-\theta} \tag{3.42}$$

$$\tau^{n+1-\theta} = \tau_\Gamma((n+1-\theta)\Delta t) \tag{3.43}$$

étape 3 : remplaçer $n+\theta$ et n par $n+1$ et $n+1-\theta$, respectivement, dans l'étape 1.

3.4.2 Etude des paramètres de la méthode

Un bon choix [23] de θ est : $\theta = 1 - \frac{\sqrt{2}}{2}$.

Lorsque nous choisirons Δt arbitrairement grand (pour obtenir rapidement les solutions stationnaires), les sous-problèmes resterons asymptotiquement bien-posés. En particulier :

$$\lim_{\Delta t \to +\infty} \eta = 1 + \alpha\tfrac{1-\omega}{\omega} > 0 \quad \forall\alpha \in [0,1],\ \forall\omega \in]0,1[\tag{3.44}$$

$$\lim_{\Delta t \to +\infty} \nu = \tfrac{1-\omega}{We} > 0 \quad \forall\omega \in]0,\frac{1}{2}[\tag{3.45}$$

Le fluide étant supposé lent, nous avons négligé le terme d'inertie, et les solutions stationnaires sont ici indépendantes du nombre de Reynolds Re.

Lorsque nous souhaiterons étudier le modèle de Maxwell ($\alpha = 1$), les sous-problèmes seront encore bien-posés :

$$\eta_{(\alpha=1)} = \frac{\theta\,\Delta t}{We + \omega\theta\,\Delta t} \tag{3.46}$$

En comparant (3.32) et (3.44) remarquons que l'introduction du paramètre de relaxation ω permet de contrôler la valeur de ν pour les grands pas de temps.

D'autre part, pour de grands nombres de Weissenberg, les sous-problèmes peuvent aussi être envisagés :

$$\lim_{We \to +\infty} \eta \quad = 1 - \alpha > 0 \qquad \forall \alpha \in \ [0, 1[\tag{3.47}$$

$$\lim_{We \to +\infty} \nu \quad = \tfrac{1}{(1-2\theta)\Delta t} > 0 \quad \forall \theta \in \ ]0, \frac{1}{2}[\tag{3.48}$$

avec la restriction (4.18) sur le pas de temps présentée au chapitre 4, pour nous assurer que le sous-problème de type transport admet une solution.

Enfin, nous présentons au chapitre suivant un choix du pas de temps Δt, indépendant des paramètres physiques We et α, et nous assurant que le sous-problème de transport des contraintes (3.42)-(3.43) est bien-posé.

3.5 Aspect matriciel du θ-schéma

Nous laissons au lecteur le soin d'adapter l'algorithme précédent au problème approché, et nous passerons directement à sa formulation en termes matriciels. Il faut cependant noter que l'algorithme obtenu utilise explicitement l'incompressibilité de la solution approchée, par la relation (3.25). Ainsi, l'adaptation de cet algorithme, associé à une méthode d'éléments finis ne satisfaisant pas exactement la relation d'incompressibilité nécéssite certaines précautions.

3.5.1 Forme matricielle de l'algorithme

Reprenons les notations matricielles introduites au paragraphe 2.9.3, et définissons de plus les matrices de masse M_u et M_p associées aux champs de vecteurs et de pression, respectivement, et défines par :

$$
\begin{aligned}
M_{u;ij} &= (\vec{v}_h^{(i)}, \vec{v}_h^{(j)}),\ 1 \leq i,j \leq N_u && (3.49)\\
M_{p;ij} &= (q_h^{(i)}, q_h^{(j)}),\ 1 \leq i,j \leq N_q && (3.50)
\end{aligned}
$$

Un pas du θ-schéma sous sa forme matricielle est alors effectué en trois étapes :

étape 1 : les vecteurs τ^n, u^n étant connus, calculer $\tau^{n+\theta}$, $u^{n+\theta}$, $p^{n+\theta}$ et $d^{n+\theta}$ comme suit :

- assembler la matrice $T(u^n)$, définie en (2.162) et le second membre τ_b^n, associé à l'opérateur de transport $(\vec{u}_h^n.\vec{\nabla})_h + \beta(.,\vec{\nabla}\vec{u}_h^n)$ avec des conditions sur la frontière amont du type $\tau_h^n = \tau_\Gamma^n$ sur Γ_- (voir l'annexe C).

- calculer explicitement :

$$
\begin{aligned}
\gamma^n &:= M_\tau^{-1}.(T(u^n).\tau^n - \tau_b^n) && (3.51)\\
f^{n+\theta} &:= \lambda M_u.u^n + \left(\frac{We - (1-\omega)\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)(B^t.\tau^n - div\tau_b^n)\\
&\quad - \left(\frac{We\,\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)(B^t.\gamma^n - div\gamma_b^n) && (3.52)
\end{aligned}
$$

  où $div\tau_b^n$ et $div\gamma_b^n$ sont les seconds membres associés à l'opérateur $\vec{\mathrm{div}}$, et calculés à partir des degrés de liberté imposés sur la frontière amont pour les tenseur τ^n et γ^n, respectivement.

- résoudre le système :

$$\begin{pmatrix} C_{(\lambda,\eta)} & D \\ D^t & 0 \end{pmatrix} \begin{pmatrix} u^{n+\theta} \\ p^{n+\theta} \end{pmatrix} = \begin{pmatrix} f^{n+\theta} + u^{n+\theta}_{(\lambda,\eta);b} \\ p_b^{n+\theta} \end{pmatrix} \tag{3.53}$$

  où la matrice symétrique définie positive $C_{(\lambda,\eta)}$:

$$C_{(\lambda,\eta)} \overset{\text{déf}}{=} \lambda M^u + \eta C \tag{3.54}$$

  est associée à l'opérateur de Helmholtz $(\lambda I - \eta\Delta)$ avec conditions aux limites du type Dirichlet $\vec{u} = \vec{u}_\Gamma$ sur Γ. On a noté $u_{(\lambda,\eta);b}$ le second membre correspondant (voir l'annexe C), calculé à partir des degrés de liberté imposés

- calculer explicitement :

$$\begin{aligned} d^{n+\theta} &:= Bu^{n+\theta} - d_b^{n+\theta} && (3.55) \\ \tau^{n+\theta} &:= \left(\frac{We - (1-\omega)\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)\tau^n - \left(\frac{We\,\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right)\gamma^n \\ &\quad + \left(\frac{2\,\alpha\theta\,\Delta t}{We + \omega\theta\,\Delta t}\right) M_\tau^{-1}.d^{n+\theta} && (3.56) \end{aligned}$$

  où $d_b^{n+\theta}$ est le second membre construit à partir des degrés de liberté imposés de $\vec{u}_h$ sur Γ, et associé à l'opérateur $D(.)$ (voir l'annexe C).

étape 2 : les vecteurs $\tau^{n+\theta}$, $u^{n+\theta}$, $d^{n+\theta}$ et u^n étant connus, calculer $\tau^{n+1-\theta}$ et $u^{n+1-\theta}$ comme suit :

- calculer explicitement :

$$\begin{aligned} g^{n+1-\theta} &:= \left(\frac{We - \omega(1-2\theta)\Delta t}{We(1-2\theta)\Delta t}\right) M_\tau \tau^{n+\theta} + \frac{2\alpha}{We} d^{n+\theta} && (3.57) \\ u^{n+1-\theta} &:= \frac{1-\theta}{\theta}\, u^{n+\theta} - \frac{1-2\theta}{\theta}\, u^n && (3.58) \end{aligned}$$

- assembler la matrice $A_\nu = \nu.M_\tau + T(u^{n+1-\theta})$ et le second membre $\tau_b^{n+1-\theta}$, associés à l'opérateur $\nu.I + (\vec{u}_h^{n+1-\theta}.\vec{\nabla})_h + \beta(.,\vec{\nabla}\vec{u}_h^{n+1-\theta})$ avec des conditions sur la frontière amont du type $\tau_h^n = \tau_\Gamma^n$ sur Γ_- (voir l'annexe C).

- résoudre :

$$A_\nu.\tau^{n+1-\theta} = g^{n+1-\theta} + \tau_b^{n+1-\theta} \tag{3.59}$$

étape 3 : remplacer $n+\theta$ et n par $n+1$ et $n+1-\theta$, respectivement, dans l'étape 1.

3.6 Etude du coût asymptotique de la méthode

Soit l'espace $H_h = T_h \times V_h \times P_h$, de dimension finie $N = N_\tau + N_u + N_p$. Nous allons évaluer le temps de calcul et la place mémoire nécessaire à une itération de la méthode lorsque N devient grand.

3.6.1 Temps de calcul

Le temps de calcul d'une itération de la méthode des directions alternées est :

$$\mathcal{T}_{DA}(N) = 2\,\mathcal{T}_S(N) + \mathcal{T}_T(N) + \mathcal{T}_L(N) \tag{3.60}$$

où

> $\mathcal{T}_S(N)$ est le temps d'une résolution d'un sous-problème du type Stokes,
>
> $\mathcal{T}_T(N)$ celui de la résolution d'un problème non symétrique du type transport,
>
> $\mathcal{T}_L(N)$ est le temps pris par les assemblages de matrices, produits matrices-vecteurs et combinaisons linéaires entre vecteurs.

L'assemblage des matrices M_u, M_τ, B et D peut être fait une fois pour toutes. Les paramètres λ et η étant fixés, il en va de même pour la matrice $C_{(\lambda,\eta)}$. Nous montrons à l'annexe C, qu'en utilisant un stockage compacté des matrices, le coût de ces opérations est en $\mathcal{O}(N)$, ainsi que le coût d'un produit matrice-vecteur. D'où :

$$\mathcal{T}_L(N) = \mathcal{O}(N) \tag{3.61}$$

D'autre part, lorsqu'il n'existe pas de lignes de courant fermées, il est possible, à l'aide d'une numérotation appropriée, de rendre la matrice A_ν triangulaire inférieure (voir le chapitre 4). De plus, cette matrice étant très creuse, le coût de la résolution est alors en $\mathcal{O}(N)$. Dans le cas général, nous avons résolu (3.59) par une méthode du type *sur-relaxations successives symétrisées par bloc* : balayage suivant le sens de l'écoulement, puis dans le sens inverse, ceci bloc par bloc, puis dans l'ordre de blocs inverse (voir au chapitre 4). La matrice A_ν étant représentée sur ordinateur à l'aide d'une structure compactée, le coût d'une itération est en $\mathcal{O}(N_\tau)$ (voir l'annexe C). Afin de diminuer le nombre d'itérations nécessaire à l'obtention d'une solution $\tau^{n+1-\theta}$ satisfaisante de (3.59), on démarrera les itérations avec la valeur précédente : $\tau^{n+\theta}$, généralement assez proche de $\tau^{n+1\theta}$. Lors d'expérimentations de ce procédé, nous avons constaté une convergence en très peu d'itérations, et un coût asymptotique optimal :

$$\mathcal{T}_T(N) = \mathcal{O}(N_\tau) \tag{3.62}$$

Enfin, classiquement, nous avons :

$$\mathcal{T}_S(N) = \mathcal{O}(N_u) \tag{3.63}$$

par la méthode "full-multigrille" [1, 46] (optimale) et :

$$\mathcal{T}_S(N) = N_u^{1+\varepsilon} \log N_u \tag{3.64}$$

par la méthode du gradient conjugé préconditionné ($\varepsilon \approx \frac{1}{5}$, quasi-optimale, voir aussi en annexe B). Les paramètres λ et η étant fixés, nous assemblons, puis factorisons cette matrice une fois pour toutes sous forme de Choleski (voir l'annexe B) lors de l'initialisation des itérations du θ-schéma :

$$C_{(\lambda,\eta)} = L^t L \tag{3.65}$$

Cette factorisation permettra de résoudre par une méthode directe les sous-problèmes du type :

$$C_{(\lambda,\eta)}.u = f \tag{3.66}$$

Ces problèmes apparaissent dans l'algorithme du gradient conjugé, nous pourrons alors les résoudre en un temps de l'ordre de $\mathcal{O}(d.N_u)$, où d désignant la largeur de bande de la matrice $C_{(\lambda,\eta)}$. Nous avons implémenté cette dernière méthode pour réaliser les expérimentations numériques du chapitre 5.

Enfin, dans les deux cas (multigrille et gradient conjugé) , il est encore possible de réduire le nombre d'itérations nécessaires à l'obtention d'une solution $(\vec{u}^{n+\theta}, p^{n+\theta})$ (resp. $(\vec{u}^{n+1}, p^{n+1})$) convenable en partant de la valeur précédente $(\vec{u}^n, p^n)$ (resp. $\vec{u}^{n+1-\theta}, p^{n+1-\theta})$), qui est généralement voisine.

Ainsi, l'algorithme d'une itération du θ-schéma est optimal lorsqu'il est associé à un solveur full-multigrille, et reste quasi-optimal lorsqu'on utilise un solveur gradient-conjugé-préconditionné.

Notons enfin que l'utilisation d'une méthode d'éléments finis utilisant une base à divergence nulle [30] permettrait de ramener le problème de Stokes à un système linéaire dont la matrice, à structure bande, peut être factorisée une fois pour toutes sous la forme de Choleski. Le coût d'une résolution d'un sous-problème de Stokes est alors en $\mathcal{O}(d.N_u)$.

3.6.2 Place mémoire

L'occupation asymptotique de la mémoire lorsque $N \to +\infty$ est :

$$\mathcal{W}_{DA}(N) \simeq \mathcal{W}_S(N) + \mathcal{W}_T(N) + \mathcal{W}_L(N) \tag{3.67}$$

où

> $\mathcal{W}_S(N)$ est la place mémoire nécessaire à la résolution d'un sous-problème du type Stokes,
>
> $\mathcal{W}_T(N)$ celle nécessaire à la résolution d'un problème non symétrique du type transport,
>
> $\mathcal{W}_L(N)$ est la place occupée par les matrices assemblées et les vecteurs.

L'algorithme nécessite le stockage permanent des matrices B, D et des champs τ, u, p, tandis que les champs γ, f, d et u à l'itération précédente peuvent être gérés par un allocateur dynamique de mémoire (en langage de programation C [37] ou Pascal, par exemple).

La structure creuse de la matrice A_ν du système (3.59) peut être déterminée une fois pour toutes (voir l'annexe C), les coefficients étant recalculés à chaque itération.

De l'étude du procédé de stockage des matrices présentée en annexe C, nous avons immédiatement :

$$\mathcal{W}_L(N) = \mathcal{W}_T(N) = \mathcal{O}(N) \tag{3.68}$$

La méthode de gradient conjugé préconditionné pour le sous-problème conduit à une occupation mémoire asymptotique :

$$\mathcal{W}_S(N) = \mathcal{O}(d.N) \tag{3.69}$$

correspondant à la forme deCholeski de la matrice $C_{(\lambda,\eta)}$ si d est la largeur de bande de la matrice $C(\lambda, \eta)$ (voir l'annexe B). Notons que la largeur de bande peut être considérablement réduite en utilisant une numérotation appropriée des éléments. Dans le cas d'un solveur multigrille, l'encombrement-mémoire reste optimal.

Par conséquent, la place en mémoire nécessaire à la méthode est asymptotiquement liée à celle due au solveur de Stokes.

3.6.3 Conclusion

Le coût d'une itération de cette méthode est donc directement lié à celui du solveur utilisé pour résoudre le sous-problème de type Stokes.

- les méthodes multigrilles conduisent à un coût optimal en temps de calcul et en occupation mémoire;
- la méthode du gradient-conjugé préconditionné reste quasi-optimale;
- l'utilisation de méthode d'élément finis utilisant une base à divergence nulle permettrait d'obtenir un temps de calcul optimal, tandis que l'encombrement en mémoire resterait quasi-optimal.

Chapitre 4

Transport des contraintes

Ce chapitre est consacré à l'approximation des sous-problèmes du type transport apparaîssant dans l'algorithme des directions alternées étudié au chapitre précédent. Nous sommes donc conduit à étudier le problème (T) de transport d'un champ de contrainte :
trouver τ, défini sur Ω, et tel que :

$$\begin{aligned}(\vec{u}.\vec{\nabla})\tau + \beta_a(\tau, \vec{\nabla}\vec{u}) + \nu\tau &= f \text{ dans } \Omega \qquad (4.1)\\ \tau &= g \text{ sur } \Gamma \qquad (4.2)\end{aligned}$$

où $\nu > 0$, le champ de vecteur $\vec{u} \in W^{1,\infty}(\Omega)$, à divergence nulle, est supposé connu, $f \in L^2(\Omega)^4_s$, et la forme bilinéaire β_a est définie en (1.20), chapitre 1.

Il s'agit de retenir, parmi les méthodes existantes, celles qui approchent l'opérateur de transport $(\vec{u}.\vec{\nabla})$ par un opérateur monotone (TVD, en anglais en abrégé, pour Total Variation Decreasing). En effet, les méthodes ne possédant pas cette propriété peuvent conduire à des solutions approchées présentant des oscillations, lorsque la solution exacte n'est pas assez régulière. Un second critère de choix sera la facilité avec laquelle nous saurons résoudre le système linéaire obtenu.

Dans un premier temps, formulant le problème (T) dans le cadre des systèmes symétriques positifs de Friedrichs [19], nous énoncerons une condition suffisante à l'existence d'une solution.

Puis, approchant ce problème par une méthode d'éléments finis classiques, nous montrerons, à l'aide d'un exemple simple, que la méthode n'est pas monotone.

Une classe de méthodes d'approximation est ensuite appliquée à (T) : méthodes de *diffusion artificielle* [32], de *diffusion suivant les lignes de courant* [9, 45, 34], et la méthode de *Galerkin discontinue* [41, 42, 33]. Afin de mettre en évidence les propriétés de ces méthodes, certains calculs seront développés sur des exemples simples.

Enfin, terminant ce chapitre par une stratégie de choix de méthode selon l'espace de dimension finie considéré pour approcher les contraintes, nous proposerons un algorithme de résolution du système linéaire obtenu.

4.1 Le problème continu

La formulation du problème (T) en un système symétrique de Friedrichs [19] nous conduit à chercher un tenseur symétrique τ, défini sur Ω, et satisfaisant à :

$$A\tau \stackrel{\text{déf}}{=} \nu\tau + (\vec{u}.\nabla)\tau + \beta_a(\tau, \vec{\nabla}\vec{u}) = f \quad \text{dans } \Omega \tag{4.3}$$

$$(M - B)(\tau - g) = 0 \quad \text{sur } \Gamma \tag{4.4}$$

où l'on a posé :

$$B = \vec{u}.\vec{\nu} \tag{4.5}$$

$$M = |B| \tag{4.6}$$

4.1.1 Notations, définitions et propriétés

Notations

Dans la suite, $H^k(\Omega)_s^4$ désignera l'espace de Sobolev d'ordre k des tenseurs symétriques, construit sur l'espace $L^2(\Omega)_s^4$, et dont la norme sera encore notée $\|.\|_{k,\Omega}$.

On notera aussi, pour tout $\tau, \gamma \in L^2(\Gamma)_s^4$:

$$\langle\tau, \gamma\rangle_\Gamma = \int_\Gamma \tau : \gamma \, ds$$

où ds désigne une mesure superficielle sur Γ et

$$\|\tau\|_{0,\Gamma} = \langle\tau, \tau\rangle_\Gamma^{\frac{1}{2}}$$

Définitions

Nous appellerons *adjoint formel* de A, et nous noterons A^*, l'opérateur défini, pour tout $\gamma \in L^2(\Omega)_s^4$ par :

$$A^*\gamma \stackrel{\text{déf}}{=} \nu\gamma - (\vec{u}.\vec{\nabla})\gamma - \beta_{-a}(\gamma, \vec{\nabla}\vec{u}) \tag{4.7}$$

Remarquons que l'ajoint formel de A est associé à une forme bilinéaire de paramétre $-a$.

On introduit la partie symétrique de A :

$$C_o\tau \stackrel{\text{déf}}{=} \nu\tau + \frac{1}{2}(\beta_a - \beta_{-a})(\tau, \vec{\nabla}\vec{u}) \tag{4.8}$$

L'opérateur A, défini en (4.3), est dit *positif* au sens de Friedrichs [19] si il existe une constante c_o strictement positive telle que l'on ait, pour toute matrice symétrique τ d'ordre 2, l'inégalité :

$$(C_o\tau):\tau \;\geq\; c_o\tau:\tau \text{ dans } \Omega \tag{4.9}$$

Enfin, nous dirons que $\tau \in L^2(\Omega)_s^4$ est *solution faible* du problème (T) si, pour tout $\gamma \in L^2(\Omega)_s^4$ satisfaisant les *conditions aux limites adjointes* :

$$(M+B)\gamma = 0 \text{ sur } \Gamma \tag{4.10}$$

on ait :

$$(\tau, A^*\gamma) = (f,\gamma) + \frac{1}{2}\langle\!\langle (M-B)g, \gamma\rangle\!\rangle_\Gamma \tag{4.11}$$

Premières propriétés

Pour tout $\tau, \gamma \in \mathbb{R}_s^4$, les identités suivantes sont vérifiées :

$$(W(\vec{u}).\tau):\gamma \;=\; -\tau:(W(\vec{u}).\gamma) \text{ dans } \Omega \tag{4.12}$$
$$(D(\vec{u}).\tau):\gamma \;=\; \tau:(D(\vec{u}).\gamma) \text{ dans } \Omega \tag{4.13}$$

Nous en déduisons une formule de Green : $\forall \tau, \gamma \in H^1(\Omega)_s^4$:

$$(A\tau, \gamma) = (\tau, A^*\gamma) + \langle\!\langle B\tau, \gamma\rangle\!\rangle_\Gamma \tag{4.14}$$

En particulier, pour $\gamma = \tau$, (4.12) s'écrit :

$$(W(\vec{u}).\tau):\tau = 0 \text{ dans } \Omega \tag{4.15}$$

De la définition (1.20) de la forme bilinéaire β_a, de l'identité précédente, il vient :

$$(C_o.\tau):\tau = \nu\,\tau:\tau - 2a\,(D(\vec{u}).\tau):\tau \text{ dans } \Omega,\ \forall \tau \in \mathbb{R}_s^4 \tag{4.16}$$

En utilisant la norme définie en (2.25), nous obtenons alors une expression plus explicite de la propriété de positivité (4.9) :

$$\nu - 2|a|.\|D(\vec{u})\|_{0,\infty,\Omega} > 0 \tag{4.17}$$

4.1.2 Résultat d'existence et unicité

Dans le cadre général des systèmes symétriques, Friedrichs [19] a démontré un résultat d'existence de la solution faible. Ce résultat s'énonce ici :

Théorème 4.1 *Existence d'une solution faible*
Soit $\vec{u} \in W^{1,\infty}(\Omega)^2$.
Si la condition de positivité (4.17) est satisfaite, alors le problème (T) *admet une solution faible* $\tau \in L^2(\Omega)^4_s$.

Remarque 4.1 sur la condition de positivité
• La condition (4.17) est suffisante, mais non nécessaire à l'existence d'une solution faible dans $L^2(\Omega)^4_s$.
• Il existe différents résultats d'existence de solutions fortes de (T), suivant les hypothèses de régularité faites sur la frontière Γ ou les données $\vec{u}$, f et g. □

4.1.3 Étude de la condition de positivité

Lorsque le problème (T) provient de l'algorithme de la θ-méthode étudié au chapitre 3, le paramètre ν est défini par la relation (3.32), page 74. Les paramètres physiques We, α et numérique θ étant supposés fixés, il est toujours possible de choisir ω et Δt de sorte que la condition (4.17) soit satisfaite :

si $|a|\, We \|D(\vec{u})\|_{0,\infty,\Omega} \leq \frac{1}{2}$, il suffit d'imposer $\omega < \omega_c$, ceci $\forall \Delta t > 0$, où

$$\omega_c \stackrel{\text{déf}}{=} \min(\frac{1}{2}, 1 - 2|a|.\|D(\vec{u})\|_{0,\infty,\Omega})$$

si $|a|\, We \|D(\vec{u})\|_{0,\infty,\Omega} > \frac{1}{2}$, on imposera $\Delta t < \Delta t_c$, ceci $\forall \omega \in]0, \frac{1}{2}[$, où

$$\Delta t_c \stackrel{\text{déf}}{=} \frac{We}{(1-2\theta)\,(2|a|\, We\|D(\vec{u})\|_{0,\infty,\Omega} - 1 + \omega)}$$

Remarquons que pour $a = 0$, le pas de temps pourra être choisis arbitrairement grand.

Dans la pratique, et pour $a \neq 0$, nous serons dans le deuxième cas ($a = 1$, We et $\|D(\vec{u})\|_{0,\infty,\Omega}$ assez grand). Une façon plus commode de satisfaire cette condition est alors de choisir $\Delta t = \Delta t^* < \Delta t_c$, ceci $\forall \omega \in]0, \frac{1}{2}[$, où :

$$\Delta t^* \stackrel{\text{déf}}{=} \lim_{We \to \infty} \Delta t_c = \frac{1}{2|a|(1-2\theta)\|D(\vec{u})\|_{0,\infty,\Omega}} \tag{4.18}$$

Ce choix de Δt a alors l'avantage d'être indépendant du paramètre physique We du problème. Dans la pratique, il sera facile d'obtenir une estimation numérique de Δt^* à partir de $D(\vec{u})$.

Ainsi, par un choix approprié des paramètres de la θ-méthode, il sera toujours possible de considérer que le sous-problème (T) admet une solution faible. Dans la suite, nous supposerons qu'un tel choix est réalisé, et que la condition de positivité est satisfaite.

4.1.4 Formulation variationnelle

Le problème de la recherche des solutions faibles de (T) admet une solution unique τ dans $H^1(\Omega)^4_s$, qui est aussi l'unique solution du problème variationnel :

$$(FVT) : (A\tau - f, \gamma) + \frac{1}{2}\langle\!\langle (M - B)(\tau - g), \gamma \rangle\!\rangle_\Gamma = 0 \tag{4.19}$$

En effet, en faisant $\gamma = \tau$ dans (4.19) et en utilisant la condition de positivité (4.9) et la formule de Green (4.14), nous obtenons l'estimation a priori :

$$\|\tau\|_{0,\Omega} + \|\sqrt{M}\tau\|_{0,\Gamma} \leq c\,(\|f\|_{0,\Omega} + \|g\|_{0,\Gamma} \tag{4.20}$$

où $c > 0$ désigne une constante indépendante des données $\vec{u}$, f et g ; cette inégalité entraîne l'unicité.

4.2 Position du problème

La solution du problème (T) (4.1)-(4.2) n'est généralement pas *globalement régulière*, sauf lorsque la frontière Γ et les données f et g le sont suffisamment. En particulier, si g est discontinue en $x_o \in \Gamma_-$, la solution τ sera *discontinue* le long de la caractéristique $x(s)$, ou ligne de courant, donnée par :

$$\frac{dx}{ds}(s) = u(x(s)) \tag{4.21}$$

$$x(0) = x_o \tag{4.22}$$

Nous appliquons ici la méthode de Galerkin classique au problème (T), et rappelons une estimation d'erreur. La perte de stabilité de la solution approchée sera mise en évidence en considérant un problème monodimentionnel.

4.2.1 Un résultat d'interpolation

Considérons $(\mathcal{T}_h)_{h>0}$ une famille de triangulation du domaine Ω, vérifiant la *condition d'angle minimum* habituelle [12], et indexée par le paramètre h, représentant le diamètre maximum des éléments $K \in \mathcal{T}_h$. Pour simplifier, nous pourrons supposer la triangulation *uniformément régulière* (voir au paragraphe 2.8, page 54).

Soit k un entier positif. On notera T_h le sous–espace de $H^1(\Omega)^4_s$ des tenseurs symétriques d'ordre 2, dont les composantes sont des fonctions continues, polynômiales par morceaux sur $\mathcal{T}_h$, et de degré inférieur ou égal à k. La théorie classique de l'approximation [12] nous fournit les résultats d'interpolation suivants :

Lemme 4.1 *interpolation dans T_h*

Pour tout $\tau \in H^{s+1}(\Omega)^4_s$, $1 \le s \le k$, il existe une interpolée *dans T_h, notée $r_h\tau$, et qui satisfait aux inégalités :*

$$\|\tau - r_h\tau\|_{0,\Omega} + h\|\tau - r_h\tau\|_{1,\Omega} \le ch^{s+1}\|\tau\|_{s+1,\Omega} \tag{4.23}$$

$$\|\tau - r_h\tau\|_{0,\Gamma} \le ch^{s+\frac{1}{2}}\|\tau\|_{s+1,\Omega} \tag{4.24}$$

où c désigne une constante indépendante de h.

4.2.2 Approximation de Galerkin classique

Cette méthode consiste à chercher $\tau_h \in T_h$ tel que l'égalité :

$$(FVT1)_h : \ (A\tau_h - f, \gamma_h) + \frac{1}{2}\langle\!\langle (M-B)(\tau - g), \gamma_h \rangle\!\rangle_\Gamma = 0 \tag{4.25}$$

soit vérifiée pour tout $\gamma_h \in T_h$.

Lorsque la condition de positivité est vérifiée, l'existence et l'unicité de la solution de $(FV1)_h$ découlent de l'estimation a priori (4.20) . De plus nous avons le

Théorème 4.2 *d'estimation de l'erreur*

Soit $\vec{u} \in W^{1,\infty}(\Omega)^2$.

Si $\tau \in H^{s+1}(\Omega)^4_s$ est la solution de (FVT), et $\tau_h \in T_h$, la solution de $(FVT1)_h$, on a, pour tout s, $1 \le s \le k$:

$$\|\tau - \tau_h\|_{0,\Omega} + \|\sqrt{M}(\tau - \tau_h)\|_{0,\Gamma} \le ch^s \|\tau\|_{s+1,\Omega} \tag{4.26}$$

Démonstration : La formule de Green (4.14) appliquée à l'erreur $\epsilon_h = \tau - \tau_h$ nous permet d'écrire :

$$(A\epsilon_h, \epsilon_h) = (\nu\epsilon_h - 2aD(\vec{u})\epsilon_h, \epsilon_h) + \frac{1}{2}\langle\!\langle B\epsilon_h, \epsilon_h \rangle\!\rangle_\Gamma$$

Ensuite, de la propriété de positivité (4.16), appliquée à ϵ_h, et en sommant sur le domaine Ω, il vient :

$$(\nu\epsilon_h - 2aD(\vec{u})\epsilon_h, \epsilon_h) \ge c_o \|\epsilon_h\|^2_{0,\Omega}$$

Des deux relations précédentes, nous obtenons :

$$(A\epsilon_h, \epsilon_h) + \frac{1}{2}\langle\!\langle (M-B)\epsilon_h, \epsilon_h \rangle\!\rangle_\Gamma \ge c_o\|\epsilon_h\|^2_{0,\Omega} + \frac{1}{2}\|\sqrt{M}\epsilon_h\|^2_{0,\Gamma}$$

D'autre part, puisque τ satisfait (FVT), et τ_h satisfait $(FVT1)_h$, leur différence ϵ_h satisfait, pour tout $\gamma_h \in \mathcal{T}_h$:

$$(A\epsilon_h, \gamma_h) + \frac{1}{2}\langle\!\langle (M-B)\epsilon_h, \gamma_h \rangle\!\rangle_\Gamma = 0$$

En choisissant $\gamma_h = \tau_h - r_h\tau$, et en sommant avec l'inégalité précédente, on obtient :

$$(A\epsilon_h, \eta_h) + \frac{1}{2}\langle\!\langle (M-B)\epsilon_h, \eta_h \rangle\!\rangle_\Gamma \ge c_o\|\epsilon_h\|^2_{0,\Omega} + \frac{1}{2}\|\sqrt{M}\epsilon_h\|^2_{0,\Gamma}$$

où $\eta_h = \tau - r_h\tau$ désigne l'erreur d'interpolation.

En utilisant l'inégalité de Cauchy-Schwartz, puis la continuité de l'opérateur A dans $H^1(\Omega)_s^4$, on majore le premier terme :

$$(A\epsilon_h, \eta_h) \leq c_1 \|\epsilon_h\|_{1,\Omega} \|\eta_h\|_{0,\Omega}$$

Pour le deuxième terme, on a :

$$\langle\!\langle (M - B)\epsilon_h, \eta_h \rangle\!\rangle_\Gamma \leq \sup_{\sigma_h \in \mathcal{T}_h} \frac{\|\eta_h.\sigma_h\|_{0,\Gamma}}{\|\sigma_h\|_{0,\Omega}} \|\epsilon_h\|_{0,\Omega}$$

et les propriétés de l'erreur d'interpolation (4.23) et (4.23) nous conduisent à (4.26). □

Remarque 4.2 Amélioration de ce résultat.

Lorsque la triangulation est composée de rectangles identiques, à côtés parallèles aux axes de coordonnées, et pour $k = 1$, on montre dans [42], avec certaines hypothèses de régularité sur τ, que :

$$\|\tau - \tau_h\|_{0,\Omega} + \|\sqrt{M}(\tau - \tau_h)\|_{0,\Gamma} = \mathcal{O}(h^2) \qquad (4.27)$$

et que cette estimation est optimale dans ce cas. □

Remarque 4.3 conditions de frontière imposées fortement

Lorsque la fonction g est assez régulière, nous préfèrerons imposer fortement les conditions aux bords, en cherchant $\tau_h \in T_h$, $\tau_h = g_h$ sur Γ_-, satisfaisant

$$(A\tau_h - f, \gamma_h) = 0 \qquad (4.28)$$

pour tout $\gamma_h \in \mathcal{T}_h$, $\gamma_h = 0$ sur Γ_-. où g_h est une interpolation (au sens de T_h) de g sur Γ_-. □

Remarque 4.4 Régularité de la solution et convergence

La méthode de Galerkin n'est définie que pour $k \geq 1$. En particulier, avec $k = 1$, le théorème 4.2 suppose que la solution τ appartient à $H^2(\Omega)_s^4$ pour établir la convergence de la méthode lorsque $h \to 0$. En particulier, lorsque la solution exacte est discontinue, le résultat précédent ne nous assure plus la convergence.

De plus, nous allons voir dans ce qui suit que la solution approchée peut être *très oscillante*, alors que la solution exacte, *très régulière*, ne présente pas ces propriétés. □

4.2.3 Perte de stabilité de la méthode

Supposons ici le champ de vecteur $\vec{u}$ constant, de composantes $(u, 0)$, $u > 0$ dans un système de coordonnées (x_1, x_2), $f = 0$ et $\Omega =]0, 1[^2$. On pose

$$\epsilon = \frac{\nu}{u}$$

En projetant (T) sur les axes de coordonnée, nous obtenons trois problèmes de Cauchy de la forme :

$$\begin{aligned} \frac{\partial \tau}{\partial x_1}(x_1, x_2) + \epsilon\tau(x_1, x_2) &= 0 \ , \forall x_1 \in]0, 1[\qquad (4.29) \\ \tau(0, x_2) &= g(0, x_2) \qquad (4.30) \end{aligned}$$

où x_2 décrit l'intervalle $]0, 1[$.

La solution de (4.29)-(4.30) est connue :

$$\tau(x_1, x_2) = g(0, x_2).e^{-\epsilon x_1}, 0 \leq x_1 \leq 1 \qquad (4.31)$$

Commençons par introduire un découpage de $[0, 1]$ en I intervalles égaux de longueur $h = 1/I$. Nous choisirons pour T_h, l'espace, de dimension $I + 1$, des éléments de $C^0([0, 1])$ dont la restriction à chaque intervalle

$$[ih, (i + 1)h], \ 0 \leq i \leq I - 1$$

est un polynôme de degré inférieur ou égal à 1. Nous avons :

$$T_h \subset H^1(0, 1)$$

Il est clair que la famille $\varphi_i, \ 0 \leq i \leq I$ des éléments de T_h, définie par :

$$\varphi_i(jh) = \delta_{ij} \ , \ 0 \leq j \leq I$$

constitue une base de T_h. Un élément $\gamma_h \in T_h$ sera ainsi parfaitement déterminé par ses composantes $\gamma_h(ih)$, $0 \leq i \leq I$ dans cette base.

Notons encore que l'opérateur d'interpolation r_h est ici défini pour tout $\gamma \in H^1(0, 1)$ par :

$$r_h\gamma = \sum_{i=0}^{I} \gamma_h(ih)\, \varphi_i \ \in T_h$$

Avec ce choix de T_h, la méthode de Galerkin consiste à chercher la fonction $\tau_h = \sum_{j=0}^{I} \tau_j\varphi_j \ \in T_h$, où les nombres $\tau_j = \tau_h(jh)$ sont solution du système linéaire :

$$\begin{aligned} &\sum_{j=0}^{I} \left(\int_0^1 (\frac{\mathrm{d}\varphi_j}{\mathrm{d}x_1} + \epsilon\varphi_j)\varphi_i \, dx_1 \right) \tau_j \\ &+ \ (\tau_0 - g(0, x_2))\varphi_i(0) = 0, \ 0 \leq i \leq I \qquad (4.32) \end{aligned}$$

L'estimation (4.27) s'applique, et l'erreur est en $\mathcal{O}(h^2)$, ce qui est optimal.

La solution exacte $\tau(.,x_2)$ étant *monotone décroissante* par rapport à la première variable, il paraît très souhaitable que la solution approchée τ_h ait des propriétés analogues.

En évaluant les intégrales apparaissant dans (4.32), le système s'écrit encore :

$$\left\{\begin{array}{lll} \tau_1 - \tau_0 & +\epsilon h(2\tau_0 + \tau_1)/3 + 2\tau_0 & = 2g(0,x_2) \\ \tau_{i+1} - \tau_{i-1} & +\epsilon h(\tau_{i+1} + 4\tau_i + \tau_{i-1})/3 & = 0 \ , 1 \leq i \leq I-1 \\ \tau_I - \tau_{I-1} & +\epsilon h(2\tau_I + \tau_{I-1})/3 & = 0 \end{array}\right. \tag{4.33}$$

La matrice associée à ce système est tridiagonale. Cette matrice est à *diagonale strictement dominante* dès que $\epsilon h > 3/2$, et alors les méthodes de résolution itératives classiques (Jacobi, Gauss-Seidel, SOR, ...) convergent.

Cependant pour $\epsilon h = 3$, la solution est :

$$\begin{aligned} \tau_0 &= \frac{2}{3} g(0,x_2) \\ \tau_i &= 0 \ , 1 \leq i \leq I \end{aligned}$$

ce qui est loin d'être satisfaisant.

En recherchant la solution sous la forme :

$$\tau_i = \alpha r_1^i + \beta r_2^i \ , \ 0 \leq i \leq I \tag{4.34}$$

nous obtenons r_1 et r_2 comme étant les deux racines de :

$$(1 + \frac{\epsilon h}{3}) r^2 + \frac{\epsilon h}{3} r + \frac{\epsilon h}{3} - 1 = 0 \tag{4.35}$$

et α et β sont donnés par la première et la dernière des équations du système (4.33)

Soit r_2 la plus petite des racines de (4.35); r_2 est toujours négative, et diminue avec ϵh lorsque :

$$\epsilon h < 3 \tag{4.36}$$

Or l'autre racine r_1 est toujours dans l'intervalle $]0,1[$, et les coefficients α et β sont non–nuls. Par conséquent la solution τ_h *oscille*, et lorsque (4.36) est vérifiée, ces oscillations ont tendance à s'amplifier lorsque ϵh diminue. Dans ce cas, nous dirons que *le schéma n'est pas monotone.*

Nous pouvons aussi recourir à des formules de quadrature pour calculer les intégrales intervenant dans (4.32). Par exemple, en utilisant la formule des trapèzes :

$$\int_{ih}^{(i+1)h} \psi(x)\, dx \approx \frac{h^2}{2} (\psi((i+1)h) + \psi(ih)) \tag{4.37}$$

pour toute fonction ψ continue sur $[ih, (i+1)h]$. Cette formule est exacte lorsque ψ est un polynôme de degré inférieur ou égal à 1. Le système (4.32) devient :

$$\left\{\begin{array}{lll} \tau_1 - \tau_0 & +\epsilon h \tau_0 + 2\tau_0 & = 2g(0, x_2) \\ \tau_{i+1} - \tau_{i-1} & +2\epsilon h \tau_i & = 0 \ , 1 \leq i \leq I-1 \\ \tau_I - \tau_{I-1} & +\epsilon h \tau_I & = 0 \end{array}\right. \tag{4.38}$$

et par un raisonnement analogue, nous constatons que le schéma n'est plus monotone dès que :

$$\epsilon h < 1 \tag{4.39}$$

Enfin, il est possible d'imposer fortement les conditions à l'amont $x_1 = 0$, ce qui revient à remplacer la première équation de (4.38) par $\tau_0 = g(0, x_2)$, mais ceci nous conduit encore à une perte de monotonie dès que (4.39) est vérifiée. Les résultats numériques sont représentés sur la figure 4.1

Conclusion

La méthode de Galerkin est précise, mais peu stable, et des oscillations peuvent alors dénaturer complètement la solution approchée lorsque le pas de discrétisation diminue.

Il y a plusieurs solutions à ce problème :

- Nous pouvons résoudre les systèmes linéaires par une méthode qui converge, même si le système devient singulier. Ainsi, Wornom et Hafez [58] ont montré qu'en utilisant un *balayage correspondant au sens de l'écoulement* dans une méthode de relaxation, il est possible de retrouver une solution physique.

- Une autre méthode consiste à modifier le schéma de façon à renforcer la diagonale du système. C'est l'objet des méthodes de diffusion qui conduisent à des *schémas décentrés*.

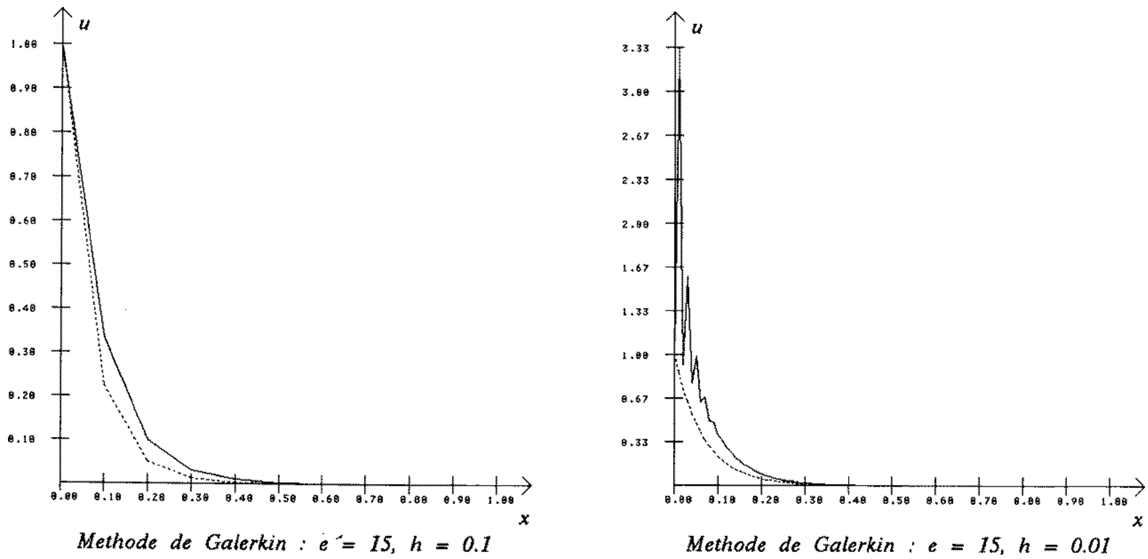


Figure 4.1: Méthode de Galerkin classique : perte de monotonie

4.3 Monotonie, stabilité et dominance diagonale

Dans la suite, nous allons étudier les méthodes d'approximation utilisant des techniques de décentrage. Nous rechercherons parmi ces méthodes celles qui nous assurent la stabilité de la solution approché. Cette dernière propriété est directement liée à la monotonie de la matrice du système linéaire obtenu.

L'objet de ce court paragraphe est de présenter quelques caractérisations d'une matrice monotone non-singulière.

Une matrice $A = (a_{i,j})_{1 \le i,j \le n}$ est à *diagonale dominante* si

$$|a_{i,i}| \geq \sum_{j=1, j \neq i}^{n} |a_{i,j}| \ , \ \forall i, \ 1 \leq i \leq n \tag{4.40}$$

et à *diagonale strictement dominante* si

$$|a_{i,i}| > \sum_{j=1, j \neq i}^{n} |a_{i,j}| \ , \ \forall i, \ 1 \leq i \leq n \tag{4.41}$$

Une matrice à diagonale strictement dominante sera non-singulière.

La matrice A est dite *monotone* si A est inversible, et tous les coefficients de A^{-1} sont positifs.

Les conditions suivantes sont équivalentes à : A est une matrice monotone non-singulière :

i) A est non-singulière, et $A + D$ est non-singulière, pour toute matrice diagonale D positive.

ii) toute valeur propre de A est de partie réelle strictement positive.

iii) $a_{i,i} > 0, \ \forall i, \ 1 \leq i \leq n$, et A est à diagonale strictement dominante.

Dans l'étude pratique des méthodes décentrées, nous chercherons à vérifier que la propriété (iii) est satisfaite. La propriété (ii) nous assure alors la stabilité de la solution approchée, quel que soit le second membre et la condition à l'amont considérés. La propriété (i) nous permet de considérer des variantes du problème initial (sous-problèmes apparaissant dans l'algorithme des directions alternées, au chapitre 3).

Il existe bien d'autres caractérisations de ces matrices : par exemple, si le graphe associé à la matrice A est fortement connexe, et que A est à diagonale dominante, alors A est monotone non-singulière Pour plus de détails, nous renvoyons à [24, page 136].

4.4 Les méthodes de diffusion

Cette classe de méthodes utilise une approximation continue des grandeurs transportées. L'idée est d'introduire un petit terme de diffusion dans le problème de transport, ce terme étant destiné à tendre vers zéro avec le paramètre h de discrétisation. La diagonale de la matrice obtenue est alors renforcée, et la solution approchée plus régulière.

4.4.1 La méthode de diffusion artificielle

Nous introduisons un terme de diffusion supplémentaire :

$$k_h(\vec{\nabla}\tau_h, \vec{\nabla}\gamma_h) \tag{4.42}$$

dans la formulation $(FVT1)_h$. Le paramètre k_h désigne le *coefficient de diffusion* scalaire.

En imposant fortement les conditions de frontière (remarque 4.3), il s'agit donc de chercher $\tau_h \in T_h$, $\tau_h = g_h$ sur Γ_-, satisfaisant à :

$$(FVT2)_h : k_h(\vec{\nabla}\tau_h, \vec{\nabla}\gamma_h) + (A\tau_h - f, \gamma_h) = 0 \tag{4.43}$$

ceci pour tout $\gamma_h \in \mathcal{T}_h$, $\gamma_h = 0$ sur Γ_-, et où g_h est une approximation (au sens de T_h)de g sur Γ.

Contrairement à la méthode précédente, la solution approchée possède ici une bonne stabilité, due à la présence du terme de diffusion (4.42). Mais cette solution est peu précise, l'erreur étant au mieux en $\mathcal{O}(h)$

Remarquons que cette méthode revient à approcher, par la méthode de Galerkin, le problème modifié :

$$\begin{aligned} -\vec{\nabla}.(k_h\vec{\nabla}\tau) + A\tau &= f \text{ dans } \Omega \\ (M-B)(\tau - g) &= 0 \text{ sur } \Gamma \end{aligned}$$

4.4.2 Choix du paramètre de diffusion

Le paramètre de diffusion doit être choisi le plus petit possible, afin de ne pas trop perturber le problème initial. D'autre part, il doit être assez grand pour assurer la stabilité de la méthode.

Reprenons l'exemple monodimensionnel (4.29)-(4.30) présenté au paragraphe 4.2.3. La méthode de diffusion artificielle nous conduit à résoudre le système linéaire :

$$\left\{\begin{array}{l}\tau_0 = g(0,x_2)\\ -\frac{2k_h}{h}(\tau_{i+1} - 2\tau_i + \tau_{i-1}) + \tau_{i+1} - \tau_{i-1}\\ \qquad\qquad + \epsilon h(\tau_{i+1} + 4\tau_i + \tau_{i-1})/3 = 0\ ,\ 1 \le i \le I-1\\ \frac{2k_h}{h}(\tau_I - \tau_{I-1}) + \tau_I - \tau_{I-1} + \epsilon h(2\tau_I + \tau_{I-1})/3 = 0\end{array}\right. \tag{4.44}$$

en imposant fortement la condition à l'amont.

En suivant le même raisonnement qu'au paragraphe 4.2.3, nous obtenons la plus petite valeur de k_h conservant la monotonie du schéma (4.44) pour tout $\epsilon > 0$:

$$\bar{k}_h = \frac{h}{2} \tag{4.45}$$

Pour ce choix de k_h, le système (4.44) devient :

$$\left\{\begin{array}{lll}\tau_0 - g(0,x_2) & & = 0\\ \tau_i - \tau_{i-1} & +\epsilon h(\tau_{i+1} + 4\tau_i + \tau_{i-1})/3 & = 0\ ,\ 1 \le i \le I-1\\ 2(\tau_I - \tau_{I-1}) & +\epsilon h(2\tau_I + \tau_{I-1})/3 & = 0\end{array}\right. \tag{4.46}$$

En utilisant la formule des trapèzes (4.37) pour évaluer les intégrales, on trouve encore :

$$\bar{k}_h = \frac{h}{2}\ ,\ (u > 0) \tag{4.47}$$

et le système linéaire s'écrit alors :

$$\left\{\begin{array}{lll}\tau_0 - g(0,x_2) & & = 0\\ \tau_i - \tau_{i-1} & +\epsilon h \tau_i & = 0\ ,\ 1 \le i \le I-1\\ & \tau_I & = 0\end{array}\right. \tag{4.48}$$

La matrice associée à ce système est triangulaire inférieure, et la résolution, quasi-explicite, donne :

$$\begin{array}{rcl}\tau_i & = & (\dfrac{1}{1+\epsilon h})^i g(0,x_2)\ ,\ 0 \le i \le I-1\\ \tau_I & = & (\dfrac{1}{1+\epsilon h})^{I-1}\dfrac{g(0,x_2)}{1+\frac{\epsilon h}{2}}\end{array}$$

Cette solution est représentée sur la figure 4.2 On pourra comparer ces résultats avec ceux de la figure 4.1.

Ces schémas sont à rapprocher de la méthode de Crank-Nicholson [50] lorsque x_1 est interprétée comme une variable de temps.

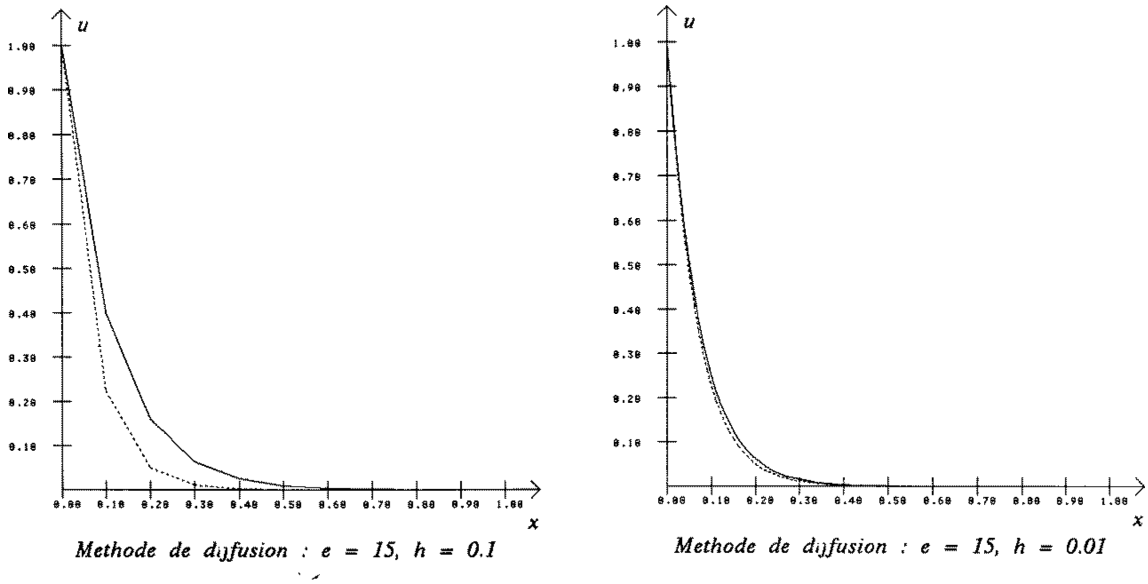


Figure 4.2: Méthode de diffusion : monotonie

Les relations du type (4.45) ou (4.47) se généralisent de façon empirique au cas bidimensionnel [32]. Considérons par exemple un élément quadrilatéral quelconque K, et soient $(b_i)_{1\leq i\leq 4}$ les milieux de ses arêtes, comme indiqué sur la figure 4.3

Les B_i sont ordonnés sur ∂K suivant le sens positif. Nous pouvons alors choisir :

$$\bar{k}_h = \frac{1}{2}(|\vec{u}.\ \overrightarrow{b_2 b_4}\ | + |\vec{u}.\ \overrightarrow{b_3 b_1}\ |) \tag{4.49}$$

Si on souhaite pouvoir différencier la formulation variationnelle $(FV2)_h$ par rapport à $\vec{u}$ en $\vec{u} = 0$, on choisira plutôt [9] :

$$\bar{k}_h = \frac{\sqrt{u^2 + v^2}}{2}(|b_2 b_4| + |b_3 b_1|) \tag{4.50}$$

ce qui nous permettrait d'utiliser éventuellement une méthode de type Newton ou quasi-Newton pour résoudre le système complet des équations.

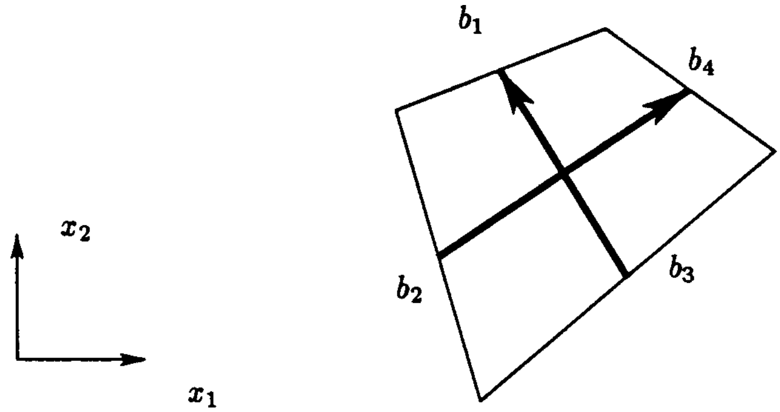


Figure 4.3: Elément quadrilatéral K

4.4.3 Exemple bidimensionnel

Les équations projetées

Rapportons le plan $\mathbb{R}^2$ au système de coordonnées (x_1, x_2), et considérons, pour fixer les idées, le domaine d'écoulement :

$$]-\frac{5}{2}, 0[\times]-1, 1[\cup[0, \frac{5}{2}, [\times]-\frac{1}{c}, \frac{1}{c}[$$

où c représente le taux de contraction. Soient (u_1, u_2) les composantes du champ de vecteur $\vec{u}$. Nous supposerons l'écoulement symétrique par rapport à l'axe $x_2 = 0$, c'est-à-dire :

$$\begin{array}{rcl} u_1(x,-y) & = & u_1(x,y) \\ u_2(x,-y) & = & -u_2(x,y) \end{array} \tag{4.51}$$

en tout point (x, y) du domaine. Si le tenseur τ, solution du problème (T), est représenté par la matrice :

$$\begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{bmatrix}$$

dans la base associée au système de coordonnées (x_1, x_2), l'hypothèse de symétrie (4.51) conduit à :

$$\begin{array}{rcl} \tau_{11}(x,-y) & = & \tau_{11}(x,y) \\ \tau_{12}(x,-y) & = & -\tau_{12}(x,y) \\ \tau_{22}(x,-y) & = & \tau_{22}(x,y) \end{array} \qquad (4.52)$$

en tout point (x, y) du domaine.

Par conséquent, il nous suffira de chercher τ sur le demi-domaine :

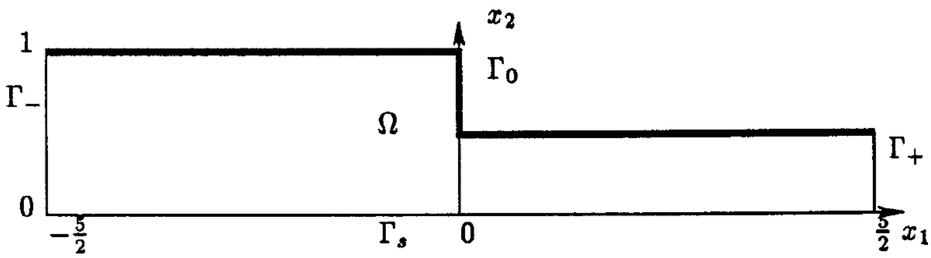


Figure 4.4: Domaine Ω pour $c = 2$

$$\Omega =]-\frac{5}{2}, 0[\times]0, 1[\cup[0, \frac{5}{2}, [\times]0, \frac{1}{c}[\qquad (4.53)$$

représenté sur la figure 4.4. Les équations, projetées sur les axes, s'écrivent :

$$\left\{ \begin{array}{llll} (\nu - 2a\frac{\partial u}{\partial x_1} + \vec{u}.\vec{\nabla})\tau_{11} & +(w-ad)\tau_{12} & & = f_{11} \\ -(w+ad)\tau_{11} & +2(\nu + \vec{u}.\vec{\nabla})\tau_{12} & +(w-ad)\tau_{22} & = 2f_{12} \\ & -(w+ad)\tau_{12} & +(\nu - 2a\frac{\partial v}{\partial x_2} + \vec{u}.\vec{\nabla})\tau_{22} & = f_{22} \end{array} \right. \qquad (4.54)$$

avec :

$$\begin{array}{rcl} w & = & \dfrac{\partial v}{\partial x_1} - \dfrac{\partial u}{\partial x_2} \\ d & = & \dfrac{\partial v}{\partial x_1} + \dfrac{\partial u}{\partial x_2} \end{array}$$

Nous supposerons l'écoulement "amont-aval" dans la marche :

$$\left\{ \begin{array}{rll} u \geq 0,\ v = 0 & \text{sur} & \Gamma_+ \cup \Gamma_- \\ u = v = 0 & \text{sur} & \Gamma_0 \\ \frac{\partial u}{\partial x_1} = v = 0 & \text{sur} & \Gamma_s \end{array} \right. \qquad (4.55)$$

et nous décomposerons la frontière $\Gamma = \partial\Omega$ suivant :

$$\begin{array}{rll} \text{l'amont :} & \Gamma_- & = \{-\frac{5}{2}\}\times]0,1[\\ \text{l'aval :} & \Gamma_+ & = \{\frac{5}{2}\}\times]0,\frac{1}{c}[\\ \text{l'axe de symétrie :} & \Gamma_s & =]-\frac{5}{2},\frac{5}{2}[\times\{0\} \\ \text{et la paroi :} & \Gamma_0 & = \Gamma \\ (\Gamma_+ \cup \Gamma_- \cup \Gamma_s) & & \end{array}$$

Formulation variationnelle

Considérons ici le problème où τ_{11} et τ_{22} sont donnés, et cherchons une solution τ_{12} de (4.54) avec une condition à l'amont du type

$$\tau_{12|\Gamma} = \tau_\Gamma \text{ sur } \Gamma_-$$

Lorsqu'il n'y aura pas d'ambiguïté, on omettra dans la suite de ce paragraphe, l'indice 12.

La formulation variationnelle de ce problème revient donc à :
chercher $\tau \in H^1(\Omega)$ satisfaisant pour tout $\varphi \in H^1(\Omega)$:

$$\int_\Omega ((\vec{u}.\vec{\nabla})\tau + \nu\tau - f)\,\varphi\,dx - \int_{\Gamma_-} (\vec{u}.\vec{n})(\tau - \tau_\Gamma)\varphi\,ds = 0 \tag{4.56}$$

où on a posé :

$$f = f_{12} - \frac{ad}{2}(\tau_{22} + \tau_{11}) + \frac{w}{2}(\tau_{22} - \tau_{11})$$

L'espace d'approximation

Pour obtenir une approximation variationnelle de la solution τ, nous allons construire un sous-espace T_h, de dimension finie de l'espace $H^1(\Omega)$.

Soient deux entiers positifs L et M vérifiant :

$$2L = 5cM$$

et la triangulation uniforme de Ω, représentée sur la figure 4.5, dont les noeuds sont les points :

$$\begin{array}{ll} x_{l,m} = (lh, mh), & \\ \quad\text{pour} & 0 \le l \le L \text{ et } 0 \le m \le cM \\ \quad\text{ou} & L+1 \le l \le 2L \text{ et } 0 \le m \le M \end{array}$$

où h est le paramètre d'approximation :

$$h = \frac{1}{cM}$$

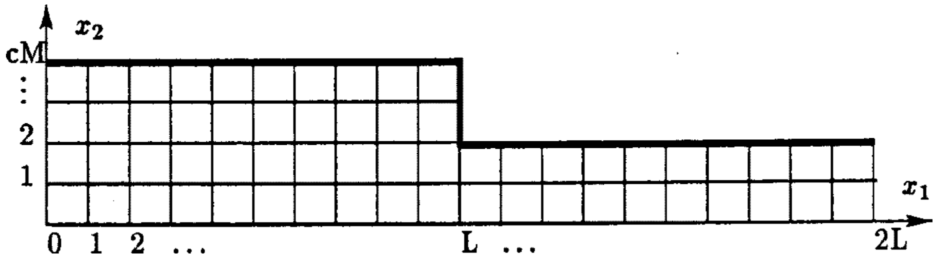


Figure 4.5: Maillage de $\bar{\Omega}$ avec $c = 2$, $L = 10$ et $M = 2$

En notant :

$$K_{l,m} = [lh, (l+1)h] \times [mh, (m+1)h]$$

la triangulation de $\bar{\Omega}$ est :

$$\begin{aligned} \mathcal{T}_h = \{K_{l,m}, \quad & 0 \le l \le L-1 \text{ et } 0 \le m \le cM-1 \\ \text{ou} \quad & L \le l \le 2L-1 \text{ et } 0 \le m \le M-1\} \end{aligned}$$

Soit $Q_1(K)$ l'espace des polynômes à deux variables sur le quadrilatère K de degré inférieur ou égal à 1 par rapport à *chacune* des variables. $Q_1(K)$ est l'espace de degré 4 engendré par les polynômes : 1, x_1,x_2, et x_1x_2.

Nous choisirons alors pour sous-espace de dimension finie de $H^1(\Omega)$:

$$T_h = \{\varphi_h \in C^0(\bar{\Omega}),\ \varphi_{h/K_{l,m}} \in Q_1(K_{l,m}),\ \forall K \in \mathcal{T}_h\}$$

Propriétés de l'espace approché

Une fonction $\varphi \in Q_1(K_{l,m})$ est déterminée de manière unique par ses valeurs aux sommets du carré $K_{l,m}$:

$$\begin{aligned} \varphi(x,y) = \varphi(x_{l,m}) &+ \frac{\varphi(x_{l+1,m}) - \varphi(x_{l,m})}{h}(x - lh) \\ &+ \frac{\varphi(x_{l,m+1}) - \varphi(x_{l,m})}{h}(y - mh) \\ &+ \frac{\varphi(x_{l+1,m+1}) + \varphi(x_{l,m}) - \varphi(x_{l,m+1}) - \varphi(x_{l+1,m})}{h}(x - lh)(y - mh) \end{aligned}$$

Soit, d'autre part, une fonction φ, définie sur $\bar{\Omega}$, et dont la restriction à chaque carré $K_{l,m}$ coïncide avec un polynôme de $Q_1(K_{l,m})$. Alors la restriction de φ à une arête du carré $K_{l,m}$ est une fonction affine, et il suffit donc d'assurer la continuité de φ aux extrémités des arêtes pour que la fonction φ soit continue sur toute cette arête.

Ainsi, φ est continue sur $\bar{\Omega}$ si et seulement si elle est continue aux $c(M+1)(L+1)+L(M+1)$ noeuds du maillage. En conséquence, une fonction de T_h est déterminée de façon unique par les valeurs qu'elle prend aux noeuds du maillage, et :

$$\dim(T_h) = c(M+1)(L+1) + L(M+1)$$

• Soit $T_{0,h}$ le sous-espace de T_h des fonctions nulles aux noeuds de Γ_-. Toute fonction φ_h appartenant à T_h peut s'écrire

$$\varphi_h = \varphi_h^0 + \varphi_h^b$$

où $\varphi_h^0 \in T_{0,h}$ et $\varphi_h = \varphi_h^0$ sur tous les noeuds de sur $\bar{\Omega} - \Gamma_-$. Une telle décomposition est unique, et φ_h^b ne dépend que des valeurs de φ_h aux noeuds appartenant à Γ_-. Nous noterons :

$$N_\tau = \dim(T_{0,h}) = (c+1)(M+1)L$$

la dimension de T_h^0.

• Numérotons à présent les sommets de $\bar{\Omega}-\Gamma_-$ de 1 à N_τ. Ceci équivaut à choisir une bijection N de

$$\begin{aligned} \mathcal{N} = \{(l,m), \quad & 1 \le l \le L \text{ et } 0 \le m \le cM+1 \\ \text{ou} \quad & L+1 \le l \le 2L \text{ et } 0 \le m \le M+1\} \\ \text{dans} \quad \mathcal{I} = \{i \in \mathbb{N}, \quad & 0 \le i \le N_\tau\} \end{aligned}$$

Nous noterons alors x_i le noeud :

$$x_i = x_{N^{-1}(i)}$$

Par exemple, une numérotation de haut en bas, puis de gauche à droite (qui suit le sens de l'écoulement principal) correspond au choix de bijection :

$$N(l,m) = \begin{cases} 1 + m + (cM+1)l & 1 \le l \le L-1 \\ & 0 \le m \le cM \\ (cM+1)L + 1 + m + (M+1)l & L \le l \le 2L \\ & 0 \le m \le M \end{cases} \tag{4.57}$$

Soit $\varphi_i,\ 1 \le i \le N_\tau$ la fonction de $T_{0,h}$ définie par :

$$\varphi_i(a_j) = \delta_{i,j}\ ,\ 0 \le i,j \le N_\tau$$

Il est clair que la famille $(\varphi_i)_{1 \le i \le N_\tau}$ constitue une base de $T_{0,h}$. Les composantes de $\varphi_h^0 \in T_{0,h}$ dans cette base sont les nombres $\varphi_h^0(x_i),\ 1 \le i \le N_\tau$.

Approximation variationnelle

Introduisons la forme bilinéaire associée à la méthode de diffusion artificielle, définie par :

$$a(\tau,\varphi) = \int_\Omega k\vec{\nabla}\tau.\vec{\nabla}\varphi\,dx + \int_\Omega (\vec{u}.\vec{\nabla}\tau)\varphi\,dx + \nu\int_\Omega \tau\varphi\,ds \tag{4.58}$$

pour tout $\tau, \varphi \in T_0 = \{\varphi \in H^1(\Omega),\ \varphi_{/\gamma_-} = 0\}$

Nous cherchons à présent :

$$\tau_h = \tau_h^b + \sum_{j=1}^{n_\tau} \tau_j\varphi_j \in T_h$$

telle que $\tau_h^b = \tau_\Gamma$ aux noeuds de Γ_-, et où les nombres $\tau_j = \tau(x_j),\ 1 \leq j \leq N_\tau$ sont solutions du système linéaire :

$$\sum_{i=1}^{N_\tau} a(\varphi_j,\varphi_i)\tau_j = (f,\varphi_i) \quad 1 \leq i \leq N_\tau \tag{4.59}$$

La matrice $a(\varphi_j,\varphi_i),\ 1 \leq i,j \leq N_\tau$ est *creuse*, car $a(\varphi_j,\varphi_i) = 0$, sauf si les points x_i et x_j sont sommet d'un même élément $K \in \mathcal{T}_h$. Avec un choix de numérotation convenable, nous obtenons une matrice à *structure-bande*, c'est-à-dire :

$$a(\varphi_i,\varphi_j) = 0, \quad \text{pour } |i-j| > d_{max}$$

où d_{max}, la *largeur de bande*, devient petite devant N_τ lorsque h tend vers zéro.

Ainsi, pour le choix de numérotation (4.57), nous établissons facilement que :

$$d_{max} = cM + 2$$

De plus, la matrice possède, au plus 9 coefficients non-nuls par ligne et par colonne.

L'utilisation de formules de quadrature appropriées pour évaluer les quantités $a(\varphi_j,\varphi_i)$ et (f,φ_i) va nous permettre d'obtenir un système linéaire dont la matrice n'a plus que 5 coefficients non-nuls par ligne et par colonne.

Utilisation de formules de quadrature

Dans la suite, il nous sera plus agréable de noter les inconnues

$$\tau_{l,m} = \tau_{N(l,m)}\ ,\ \forall (l,m) \in \mathcal{N}$$

Ainsi :

$$\tau_{l,m} = \tau_h(a_{l,m}) \ , \ \forall (l,m) \in \mathcal{N}$$

et $\varphi_{l,m}$ désignera la fonction de T_h^0 valant 1 au noeud $x_{l,m}$ et zéro aux autres noeuds du maillage.

Nous utilisons la généralisation bidimensionnelle de la formule des trapèzes (4.37) : si ψ est une fonction continue sur le carré $K_{l,m}$, on a la formule de quadrature :

$$\int_{K_{l,m}} \psi(x)\, dx \approx \frac{h^2}{4} \left(\psi(x_{l,m}) + \psi(x_{l+1,m}) + \psi(x_{l,m+1}) + \psi(x_{l+1,m+1}) \right) \tag{4.60}$$

qui est exacte lorsque $\psi \in Q_1(K_{l,m})$. Nous pouvons alors définir le produit scalaire $(.,.)_h$ entre deux fonctions τ_h, φ_h de T_h^0 par :

$$\begin{aligned}(\tau_h, \varphi_h)_h = \sum_{K_{l,m} \in \mathcal{T}_h} \frac{h^2}{4} &\left(\tau_h(x_{l,m})\varphi_h(x_{l,m}) + \tau_h(x_{l+1,m})\varphi_h(x_{l+1,m}) \right. \\ &\left. + \tau_h(x_{l,m+1})\varphi_h(x_{l,m+1}) + \tau_h(x_{l+1,m+1})\varphi_h(x_{l+1,m+1}) \right)\end{aligned} \tag{4.61}$$

et la forme bilinéaire $a_h(.,.)$ par :

$$a_h(\tau_h, \varphi_h) = (k\frac{\partial \tau_h}{\partial x_1}, \frac{\partial \varphi_h}{\partial x_1})_h + (k\frac{\partial \tau_h}{\partial x_2}, \frac{\partial \varphi_h}{\partial x_2})_h + (\vec{u}.\vec{\nabla}\tau_h, \varphi_h)_h + \nu(\tau_h, \varphi_h)_h \tag{4.62}$$

Avec les définitions précédentes, la méthode d'approximation variationnelle revient à chercher

$$\tau_h^* = \tau_h^b + \sum_{(p,q)\in\mathcal{N}} \tau_{p,q}^* \varphi_{p,q} \ \in T_h^0$$

telle que $\tau_h^b = \tau_\Gamma$ aux noeuds de Γ_-, et où les nombres :

$$\tau_{p,q}^* = \tau_{p,q}^*(a_{p,q}) \ , \ (p,q) \in \mathcal{N}$$

sont solution du système linéaire :

$$\sum_{(p,q)\in\mathcal{N}} a_h(\varphi_{p,q}, \varphi_{l,m})\tau_{p,q}^* = (f, \varphi_{l,m})_h \ , \ \forall (l,m) \in \mathcal{N} \tag{4.63}$$

Remarquons que nous avons ainsi modifié la matrice et le second membre, du système à résoudre. On pourra vérifier que ce problème est encore bien posé. De plus, si τ_h est la solution construite à partir de (4.59), et si τ_h^* est la solution de (4.63), on peut montrer que :

$$\|\tau_h - \tau_h^*\|_{0,\Omega} \leq ch\|\tau\|_{0,\Omega}$$

où c est une constante indépendante de h, et τ, la solution de (4.56).

Ainsi, le choix de la formule de quadrature (4.60) dans l'intégration numérique n'affecte pas l'ordre de la méthode d'éléments finis. Dans la suite, nous omettrons l'astérisque pour désigner la solution de (4.63).

Le schéma obtenu

Nous supposons ici que le champ $\vec{u}$ est obtenu à l'aide de l'élément de Thomas-Raviart de degré $l = 1$ presenté au chapitre 2, soit $\vec{u} \in V_h$. Sur tout élement $K_{l,m}$ nous avons donc :

$$\begin{aligned} u_1(x_1, x_2) &= u_{l+1,m+\frac{1}{2}}.(x_1 - lh) + u_{l,m+\frac{1}{2}}.((l+1)h - x_1) \\ u_2(x_1, x_2) &= v_{l+\frac{1}{2},m+1}.(x_2 - mh) + v_{l+\frac{1}{2},m}.((m+1)h - x_2) \end{aligned}$$

avec les notations :

$$\begin{aligned} u_{l,m+\frac{1}{2}} &= u_1(lh, (m+\tfrac{1}{2})h) \\ v_{l+\frac{1}{2},m} &= u_2((l+\tfrac{1}{2})h, mh) \end{aligned}$$

Utilisons l'expression (4.49) pour le coefficient de diffusion de la méthode. Sur une maille $K_{l,m}$ intérieure au domaine Ω, le schéma obtenu s'écrit,

$$\begin{aligned} &-(k_{l-\frac{1}{2},m} + hu_{l,m})\tau_{l-1,m} - (k_{l,m+\frac{1}{2}} - hv_{l,m})\tau_{l,m+1} + (4k_{l,m} + \nu h^2)\tau_{l,m} \\ &-(k_{l,m-\frac{1}{2}} + hv_{l,m})\tau_{l,m-1} - (k_{l+\frac{1}{2},m} - hu_{l,m})\tau_{l+1,m} = h^2 f_{l,m} \\ &\text{ceci } \forall (l,m) \in \mathcal{N}. \end{aligned} \tag{4.64}$$

avec la condition sur la parois amont imposée fortement :
$\tau_{0,m} = \tau_\Gamma(-\frac{5}{2}, mh),\ 0 \leq m \leq cM$.

Nous avons noté dans (4.64) :

$$\begin{aligned} f_{l,m} &= f(lh, mh) \\ u_{l,m} &= \tfrac{1}{2}(u_{l,m-\frac{1}{2}} + u_{l,m+\frac{1}{2}}) \\ v_{l,m} &= \tfrac{1}{2}(v_{l-\frac{1}{2},m} + u_{l+\frac{1}{2},m})\ ,\ (l < 2L) \end{aligned}$$

et :

$$\begin{aligned} k_{l-\frac{1}{2},m} &= \tfrac{h}{8}(4|v_{l-\frac{1}{2},m}| + |u_{l,m+\frac{1}{2}}| + |u_{l,m-\frac{1}{2}}| + |u_{l-1,m+\frac{1}{2}}| + |u_{l-1,m-\frac{1}{2}}|) \\ k_{l,m-\frac{1}{2}} &= \tfrac{h}{8}(4|u_{l,m-\frac{1}{2}}| + |v_{l+\frac{1}{2},m}| + |v_{l-\frac{1}{2},m}| + |v_{l+\frac{1}{2},m-1}| + |v_{l-\frac{1}{2},m-1}|) \\ k_{l,m} &= \tfrac{1}{4}(k_{l-\frac{1}{2},m} + k_{l+\frac{1}{2},m} + k_{l,m-\frac{1}{2}} + k_{l,m+\frac{1}{2}}) \end{aligned} \tag{4.65}$$

Avec les expressions (4.65), le schéma (4.64) est *monotone*. La matrice associée au système linéaire est alors à *diagonale strictement dominante* : la solution τ_h n'oscille pas, et les méthodes itératives classiques convergent.

Cependant, l'expression (4.49) du paramètre de diffusion artificielle ne conduit pas aux plus petites valeurs possibles de $k_{l,m-\frac{1}{2}}$ et $k_{l-\frac{1}{2},m}$ conservant la monotonie du schéma.

Amélioration du schéma

Il est possible d'améliorer le schéma (4.64) :

$$\begin{array}{l} -(|u_{l,m}| + u_{l,m})\tau_{l-1,m} - (|v_{l,m}| - v_{l,m})\tau_{l,m+1} + 2(|u_{l,m}| + |v_{l,m}| + \nu h)\tau_{l,m} \\ -(|v_{l,m}| + v_{l,m})\tau_{l,m-1} - (|u_{l,m}| - u_{l,m})\tau_{l+1,m} = 2hf_{l,m} \\ \text{ceci } \forall (l,m) \in \mathcal{N} \end{array} \tag{4.66}$$

auquel on ajoute les conditions pour les noeuds appartenant à la parois amont Γ_-.

La matrice associée au système ainsi obtenu n'a que 3 coefficients non nuls par ligne. De plus, remarquons que, si l'écoulement est tel que $u_1 > 0$ et $u_2 < 0$ dans Ω (sans recirculations), et avec le choix de numérotation (4.57), la résolution est quasi-explicite, la matrice du système linéaire étant triangulaire inférieure.

Dans le cas d'un écoulement présentant des lignes de courant fermées, nous utiliserons une méthode itérative de résolution. La matrice du système n'étant pas symétrique, on obtiendra un opérateur d'iteration symétrique par une méthode de Gauss-Seidel symétrisée ou de de sur-relaxations successives symétrisées (SSOR, voir l'annexe C).

Remarquons enfin que ce dernier schéma correspond à l'introduction de deux termes de diffusion artificielle, dirigés suivant les axes de coordonnées (diffusion paraxiale).

4.4.4 Diffusion suivant les lignes de courant

Ce schéma, appelé en anglais *streamline diffusion*, a été proposé par Hughes et al. [9], puis étudié par Johnson et al. [34]. On introduit le terme de diffusion

$$k_h.(\vec{u}.\vec{\nabla}\tau_h, \vec{u}.\vec{\nabla}\gamma_h) \tag{4.67}$$

dans la formulation $(FVT1)_h$. Le paramètre k_h désigne toujours le coefficient de diffusion scalaire. La méthode originale repose donc sur le problème modifié :

$$\begin{aligned} -\vec{\nabla}.(K_h\vec{\nabla}\tau) + A\tau &= f \text{ dans } \Omega && (4.68)\\ (B-M)(\tau-g) &= 0 \text{ sur } \Gamma && (4.69)\end{aligned}$$

où le tenseur de diffusion K_h est défini par :

$$K_h = k_h \vec{u} \otimes \vec{u} \qquad (4.70)$$

dont les axes principaux sont portés par $\vec{u}$.

Cependant, il est possible d'introduire le terme (4.67) sans modifier l'équation constitutive (voir Chistie & al [11]). L'idée est la suivante : on remplace les fonctions-test γ_h dans $(FVT1)_h$ par des fonctions ayant plus de "poids" à l'amont qu'à l'aval : $\gamma_h + k_h(\vec{u}.\vec{\nabla})\gamma_h$. Cette formulation de "Petrov-Galerkin" s'énonce :
trouver $\tau_h \in T_h$, $\tau_h = g_h$ sur Γ_- tel que :

$$(FVT3)_h : (A\tau_h - f, \gamma_h + k_h(\vec{u}.\vec{\nabla})\gamma_h) = 0 \qquad (4.71)$$

ceci pour tout $\gamma_h \in T_h$, $\gamma_h = 0$ sur Γ_-, et où g_h est une approximation (au sens de T_h) de g sur Γ_-. Ceci revient en fait à appliquer la méthode de Galerkin au problème :

$$\begin{aligned} -\vec{\nabla}.(K_h\vec{\nabla}\tau) + A\tau &= f + k_h(\vec{u}.\vec{\nabla})f \text{ dans } \Omega && (4.72)\\ (B-M)(\tau-g) &= 0 \text{ sur } \Gamma && (4.73)\end{aligned}$$

mais cette fois, la solution exacte τ est encore solution de ce problème.

Pour cette méthode, nous avons le résultat [34] suivant :

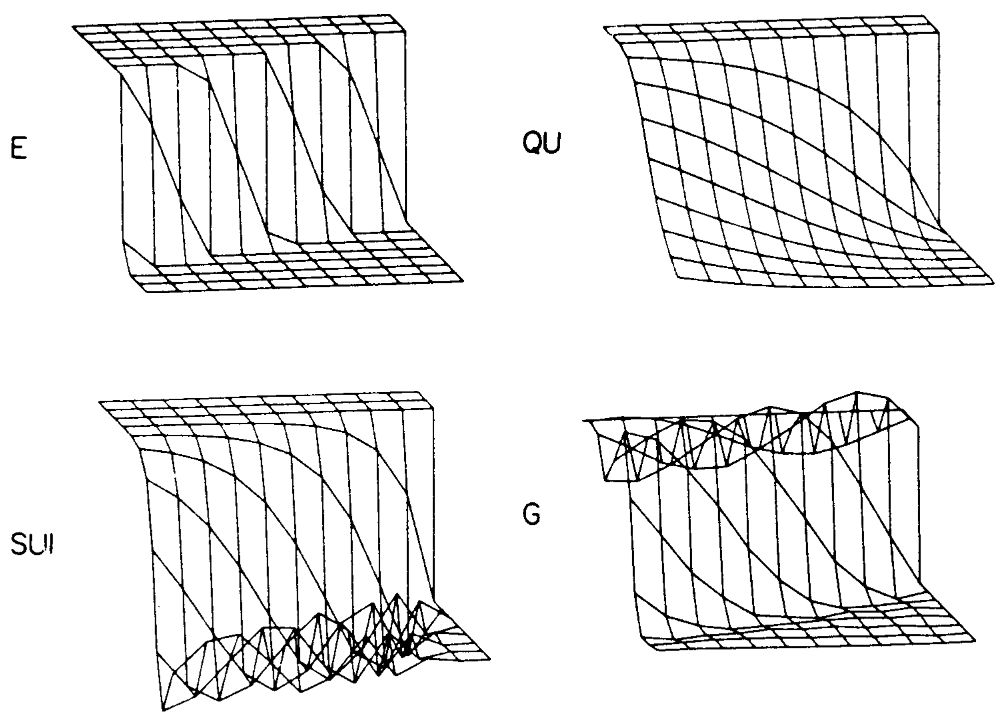
Théorème 4.3
Soit $\vec{u} \in W^{1,\infty}(\Omega)^2$.

- *le problème $(FVT3)_h$ admet une unique solution;*
- *Si $\tau \in H^{s+1}(\Omega)^4_s$ est la solution de (FVT), et $\tau_h \in T_h$, la solution de $(FVT3)_h$, on a, pour tout s, $1 \le s \le k$:*

$$\|\tau - \tau_h\|_{0,\Omega} \le c h^{s+\frac{1}{2}} \|\tau\|_{s+1,\Omega} \qquad (4.74)$$

4.4.5 Conclusion sur les méthodes de diffusion

L'introduction d'un terme de diffusion artificielle (4.42) permet d'obtenir des solutions non-oscillantes, mais peu précises. En fait, on résoud alors un problème plus visqueux, donc de solution plus régulière. Si la solution exacte est peu régulière, la solution approchée reste très regulière, le problème est alors dénaturé (voir figure 4.6.c ou [9, page 219]).



(E) solution exacte,
(QU) méthode de diffusion artificielle,
(G) méthode de Galerkin,
(SU1) méthode de diffusion suivant les lignes de courant.

Figure 4.6: Méthodes de Galerkin et de diffusion, d'après Hughes et Brook

De plus, nous avons vu que le terme de diffusion ne correspond pas au plus petit terme conservant la monotonie du schéma, et qu'il est possible d'optimiser la méthode dans le cas d'un maillage rectangulaire à côtés parallèles aux axes. D'autre part, le coefficient de diffusion $\bar{k}_h$, défini en (4.49), est une généralisation empirique du cas monodimensionnel aux éléments quadrilatéraux. Pour un élément triangulaire, il n'existe pas de telle relation à notre connaissance.

La méthode de diffusion suivant les lignes de courant paraît plus précise, mais moins stable : il n'est pas clair que le schéma soit monotone, et lorsque la solution est peu régulière, la solution approchée présente encore des oscillations (voir figure 4.6.d). D'autre part, cette méthode semble là aussi limitée à des élément quadrilatéraux. Enfin, la résolution n'est plus explicite, même dans le cas d'un maillage rectangulaire et pour des lignes de courant non-fermées.

4.5 Méthodes de Galerkin discontinues

Si la solution est approchée par des fonctions polynômiales discontinues aux interfaces des éléments, il est possible d'introduire un décentrage dans un schéma qui contient les intégrales sur les arêtes, en utilisant la valeur à droite où à gauche de l'arête, suivant la direction de l'écoulement. Lesaint et Raviart [42, 41] ont introduit une première méthode de ce type et donné des estimations d'erreur pour les systèmes symétriques positifs de Friedrichs.

4.5.1 principe de la méthode

Dans le cadre de notre problème de transport des contraintes (T), on désignera par S_h le sous-espace de $L^2(\Omega)_s^4$ des tenseurs symétriques d'ordre 2, dont les composantes sont des fonctions polynômiales par morceaux sur $\mathcal{T}_h$, et de degré inférieur ou égal à k.

On recherche alors $\tau_h \in S_h$, $\tau_h = g_h$ sur Γ_- telle que :

$$(FVT4)_h : \sum_{K\in\mathcal{T}_h} (A\tau_h - f, \sigma_h)_{0,K} + \langle \frac{M_K - B_K}{2}[\tau_h], \sigma_h\rangle_{\partial K} = 0 \qquad (4.75)$$

ceci pour tout $\gamma_h \in S_h$, $\gamma_h = 0$ sur Γ_-, et où g_h est une approximation (au sens de T_h) de g sur Γ_-.

On a posé :

> $[\sigma] = \sigma^{int} - \sigma^{ext}$ le saut d'une fonction σ à la traversé d'une arête d'un élément K , où σ^{int} et σ^{ext} sont les traces intérieures et extérieures de σ sur ∂K, respectivement.
>
> $B_K = \vec{u}.\vec{\nu}_K$, où $\vec{\nu}_K$ est le vecteur de la normale unitaire sortante à K sur ∂K.

D'autre part, M_K vérifie :

i) $M_{K_1} = M_{K_2}$ sur $\partial K_1 \cap \partial K_2$, ceci $\forall K_1, K_2 \in \mathcal{T}_h$

ii) $M_K \geq 0$ sur ∂K

iii) $\forall \tau, \gamma \in \mathrm{I\!R}_s^4$:

$$(M_K - B_K)\tau:\gamma \leq M_K\tau:\tau + c\gamma:\gamma \qquad (4.76)$$

où c est une constante indépendante de h et des données f et g.

Remarque 4.5 choix de M_K

Il est possible de prendre

$$M_K = |B_K| \text{ sur } \partial K \tag{4.77}$$

mais nous verrons que ce n'est pas le seul choix possible. □

Remarque 4.6 sur les discontinuités

Si la solution exacte τ est régulière, nous espérons que la solution approchée de $(FVT4)_h$ le sera aussi. Dans ce sens, nous avons ajouté à la formulation classique un terme a priori petit :

$$\frac{1}{2}\langle M_K - B_K[\tau_h], \sigma_h\rangle_{\partial K} \tag{4.78}$$

exprimant les discontinuités de la solution approchée aux interfaces des éléments. □

4.5.2 Résolution du problème approché; monotonie

Avec le choix (4.77) de M_K, et du fait de la discontinuité des fonctions de S_h, le problème se décompose en une famille de problèmes élémentaires :
trouver $\tau_h \in S_h$ *tel que*

$$(A\tau_h - f, \sigma_K)_{0,K} - \langle\!\langle(\vec{u}.\nu_K)[\tau_h], \sigma_K\rangle_{\partial K_-} \tag{4.79}$$

ceci $\forall \sigma_K \in P_k(K)$, $\forall K \in Th$, et où ∂K_- désigne la partie frontière de K où le flux de $\vec{u}$ est entrant :

$$\partial K_- = \{x \in \partial K;\ (\vec{u}.\vec{\nu}_K)(x) < 0\} \tag{4.80}$$

Lorsque l'écoulement ne contient pas de lignes de courant fermée, il sera possible de résoudre le problème en un balayage du maillage, partant des élément situés à l'amont, et suivant les lignes de courant (en numérotant convenablement les éléments, par exemple). Dans le cas général, on montre que ce problème admet une solution unique (voir [41, 42]).

Johnson et Pitkäranta ont montré [35] que l'erreur associée au schéma $(FVT4)_h$ est en $\mathcal{O}(h^{k+\frac{1}{2}})$ pour la norme L^2 de l'erreur, et en $\mathcal{O}(h^{k+1})$ si la triangulation est uniforme.

Remarque 4.7 cas des éléments de plus bas degré

Ce dernier résultat de convergence nous permet de considérer la méthode pour $k = 0$, correspondant à des solution approchées constantes par élément. □

Enfin, pour $k \leq 1$, et le choix (4.77) de M_K, ce schéma est monotone, alors que pour $k \geq 2$, cette propriété n'est plus conservée [25].

4.5.3 Exemple monodimensionnel

Reprenons l'exemple monodimensionnel (4.29)-(4.30) du paragraphe 4.2.3. L'espace de dimension finie $S_h$est composé des fonctions dont la restriction à chaque intervalle $[jh,(j+1)h],\ 0 \le j \le I-1$ est un polynôme de degré inférieur ou égal à k.

La méthode de Galerkin discontinue nous conduit à chercher $\tau_h \in S_h$, $\tau_h(0) = g$ tel que :

$$\sum_{j=0}^{I-1} \left(\int_{jh}^{(j+1)h} (\frac{\partial \tau_h}{\partial x} + \epsilon \tau_h)\sigma_h \, dx \right. \\ + \left. \frac{1+\theta_j}{2}(\tau_j^+ - \tau_j^-)\sigma_j^+ + \frac{1-\theta_{j+1}}{2}(\tau_{j+1}^+ - \tau_{j+1}^-)\sigma_j^- \right) = 0 \quad (4.81)$$

pour tout $\sigma_h \in S_h$, $\sigma_h = 0$ sur Γ_-.

On a noté, $\forall j$, $0 \le j \le I$:

$$\sigma_j^+ = \lim_{\substack{x \to jh \\ x > jh}} \sigma_h(x) \quad \text{et} \quad \sigma_j^- = \lim_{\substack{x \to jh \\ x < jh}} \sigma_h(x)$$

et les quantités θ_j sont positives et représentent les nombres $M_K(jh)$.

Le système (4.81) contient $(k+1)I$ équations et autant d'inconnues. Remarquons que le choix

$$\theta_j = 1,\ 0 \le j \le I$$

conduit à I problèmes indépendants, comportant chacun $k+1$ inconnues :

$$\int_{jh}^{(j+1)h} (\frac{\partial \tau_h}{\partial x} + \epsilon \tau_h)\sigma_j \, dx + (\tau_j^+ - \tau_j^-)\sigma_j = 0 \quad (4.82)$$

pour tout $\sigma_j \in P_k(jh,(j+1)h)$, $1 \le j \le I$.

La solution $\tau_h \in S_h$ de (4.82) vérifie la majoration d'erreur [42] :

$$\|\tau - \tau_h\|_{L^2(0,1)} \le ch^{k+1}\|\tau\|_{H^{k+1}(0,1)} \quad (4.83)$$

ceci $\forall k \ge 0$. Ceci nous permet d'utiliser des fonctions constantes par intervalle ($k = 0$).

Considérons à présent, pour $k = 0$, le choix $\theta_j = \theta \ge 0$, $\theta_0 \ge 0$ et $\theta_I = 1$. Le problème (4.81) conduit au système :

$$\tau_0 = g \quad (4.84)$$

$$\begin{aligned}
\epsilon h\tau_{\frac{1}{2}} + \frac{1+\theta_0}{2}(\tau_{\frac{1}{2}} - \tau_0) + \frac{1-\theta}{2}(\tau_{\frac{3}{2}} - \tau_{\frac{1}{2}}) &= 0 \qquad (4.85)\\
\epsilon h\tau_{j+\frac{1}{2}} + \frac{1+\theta}{2}(\tau_{j+\frac{1}{2}} - \tau_{j-\frac{1}{2}}) + \frac{1-\theta}{2}(\tau_{j+\frac{3}{2}} - \tau_{j+\frac{1}{2}}) &= 0, \qquad (4.86)\\
1 \leq j \leq I-1 &\\
\epsilon h\tau_{I-\frac{1}{2}} + \frac{1+\theta}{2}(\tau_{I+\frac{1}{2}} - \tau_{I-\frac{3}{2}}) &= 0 \qquad (4.87)
\end{aligned}$$

où $\tau_{j+\frac{1}{2}}$ désigne la valeur de τ_h sur l'intervalle $]jh,(j+1)h[$. Ce schéma n'est monotone pour toute valeur de ϵ que si $\theta \geq 1$.

Cette méthode coïncide avec un schéma aux différences finies où les dérivées sont évaluées en :

$$x_{j+\frac{1}{2},\theta} = (j + \frac{1-\theta}{2})h$$

Ce schéma est d'autant plus décentré vers l'amont que θ est grand. Lorsque $\theta \neq 0$ ou 1, on peut montrer [41] que l'erreur est au mieux en $\mathcal{O}(h^{\frac{1}{2}})$.

Le schéma centré correspond à $\theta_0 = \theta_I = 1$ et $\theta_j = 0,\ 0 \leq j \leq I-1$, et la solution approchée τ_h vérifie alors l'estimation d'erreur [41] :

$$\|\tau - \tau_h\|_{0,\Omega} + |\tau((I-\tfrac{1}{2})h) - tau_{I-\frac{1}{2}}| \leq ch\|\tau\|_{2,\Omega} \qquad (4.88)$$

Le schéma décentré classique est obtenu pour $\theta_j = 1,\ 0 \leq j \leq I$, l'estimation (4.83) s'applique l'erreur est en $\mathcal{O}(h)$, et le système s'écrit :

$$\tau_0 = g \qquad (4.89)$$

$$\begin{aligned}
\epsilon h\tau_{\frac{1}{2}} + \tau_{\frac{1}{2}} - \tau_0 &= 0 \qquad (4.90)\\
\epsilon h\tau_{j+\frac{1}{2}} + \tau_{j+\frac{1}{2}} - \tau_{j-\frac{1}{2}} &= 0,\ 1 \leq j \leq I-1 \qquad (4.91)
\end{aligned}$$

dont la solution est :

$$\tau_{j+\frac{1}{2}} = (1+\epsilon h)^{-j-1},\ 0 \leq j \leq I-1 \qquad (4.92)$$

et est représentée sur la figure 4.7. On pourra comparer ces résultats avec ceux des figures 4.1 et 4.2.

Ainsi, les méthodes les plus précises correspondent à $\theta = 0$ (schéma centré) et $\theta = 1$ (schéma décentré), pour lequelles la méthode de plus bas degré conduit à une erreur en $\mathcal{O}(h)$. Seul le schéma décentré est monotone pour toute valeur de ϵ, , et correspond au choix (4.77) de M_K.

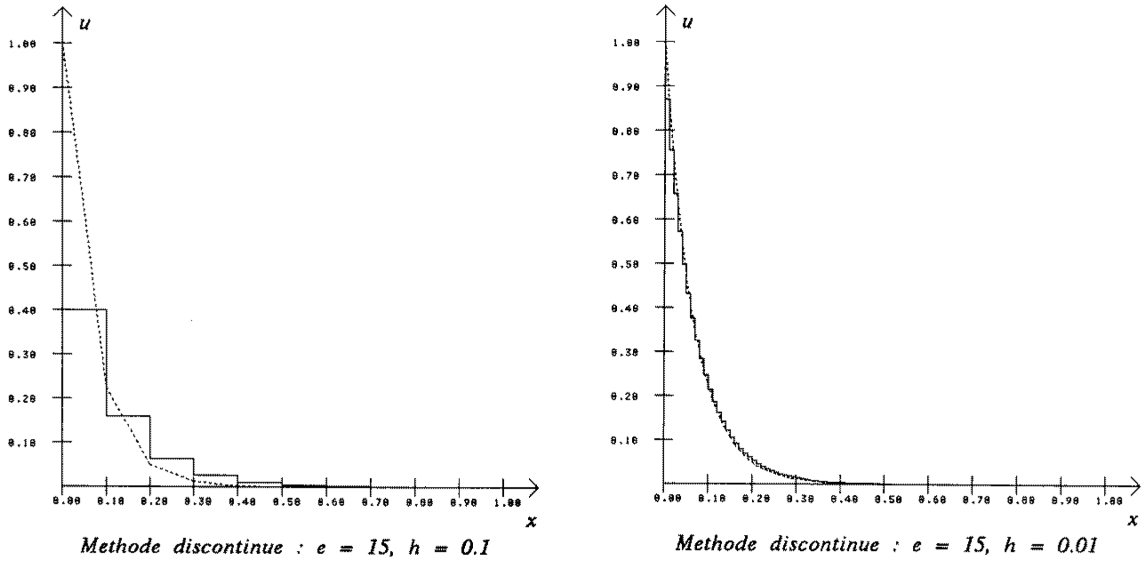


Figure 4.7: Méthode discontinue

4.5.4 Relation avec la méthode des différences finies

Reprenons l'exemple bidimensionnel paragraphe 4.4.3. L'espace de dimension finie S_h est composé des fonctions dont la restriction à chaque carré $K_{l,m}$ est un polynôme de degré inférieur ou égal à k. Nous allons considérer le cas des éléments de plus bas dgré ($k = 0$) et le choix (4.77) de M_K. Le schéma obtenu sur une maille intérieure au domaine est :

$$\begin{aligned}
&(u_{l+1,m+\frac{1}{2}} - |u_{l+1,m+\frac{1}{2}}|)(\tau_{l+\frac{3}{2},m+\frac{1}{2}} - \tau_{l+\frac{1}{2},m+\frac{1}{2}}) \\
&\quad + \quad (v_{l+\frac{1}{2},m+1} - |v_{l+\frac{1}{2},m+1}|)(\tau_{l+\frac{1}{2},m+\frac{3}{2}} - \tau_{l+\frac{1}{2},m+\frac{1}{2}}) \\
&\quad + \quad (u_{l,m+\frac{1}{2}} + |u_{l,m+\frac{1}{2}}|)(\tau_{l+\frac{1}{2},m+\frac{1}{2}} - \tau_{l-\frac{1}{2},m+\frac{1}{2}}) \\
&\quad + \quad (v_{l+\frac{1}{2},m} + |v_{l+\frac{1}{2},m}|)(\tau_{l+\frac{1}{2},m+\frac{1}{2}} - \tau_{l+\frac{1}{2},m-\frac{1}{2}}) \\
&\quad + \quad 2\nu h \tau_{l+\frac{1}{2},m+\frac{1}{2}} = 2h f_{l+\frac{1}{2},m+\frac{1}{2}} \qquad (4.93)
\end{aligned}$$

Il est facile de vérifier que ce schéma est monotone, du fait de la relation d'incompressibilité discrète (2.168) associé à l'interpolation du champ de vecteur $\vec{u} \in V_h$ par la méthode Marcker and Cell. Il s'agit du schéma aux différences finies"Donor Cell" (schéma décentré, aux grilles entrecroisées).

Remarquons encore que si l'une des composantes du champs de vecteur $\vec{u}$ est de signe constant, la résolution est quasi-explicite.

4.5.5 conclusion

La méthode de Galerkin discontinue conduit à des schémas monotones pour les éléments de degré 0 ou 1. De plus, cette méthode est précise, et s'applique à des triangulations composées de triangles ou de quadriltères.

D'autre part, lorsqu'il n'existe pas de lignes de courant fermées, il est possible, avec une numérotation appropriée, de résoudre le problème de façon quasi-explicite.

4.6 Aspect matriciel et résolution

4.6.1 Choix de la méthode d'approximation

Utilisant l'élément de Raviart-Thomas de degré l pour approcher le champs de vecteur vitesse, nous proposons une stratégie de choix de la méthode de décentrage selon l'espace d'approximation choisi pour les contraintes (paragraphe 2.8.2, page 58).

Lorsque nous choisissons d'approcher les contraintes à l'aide de l'espace $T_h^{(1)}$ (voir paragraphe 2.8.2), nous utiliserons :

- la méthode discontinue de Lesaint et Raviart ($k = l - 1$) pour les composantes normales τ_{11} et τ_{22};
- la méthode de diffusion suivant les lignes de courant ($k = l$) pour la composante de cisaillement τ_{12}.

Ceci nous assure un faible nombre de degrés de liberté et une méthode suffisamment précis.

En considérant les espaces de contraintres $T_h^{(2)}$ et de variables Lagrangiennes M_h, nous utiliserons exclusivement la méthode discontinue de Lesaint et Raviart . Le nombre de degrés de liberté est alors sensiblement augmenté, mais les schémas obtenus sont monotones dans le cas des éléments de plus bas degré $l = 1$.

L'espace $T_h^{(2)}$ conduit au plus grand nombre de degrés de liberté. Cependant, les composantes du tenseur sont approchées de façon identique, et ceci constitue un gain de place mémoire et de temps d'assemblage, car l'opérateur de transport scalaire approché $(\vec{u}.\vec{\nabla})_h$ n'est calculé et stocké qu'une seule fois.

4.6.2 Résolution du système linéaire

Le système linéaire à résoudre est de la forme :

$$\begin{pmatrix} A_1 & 0 & A_{13} \\ 0 & A_2 & A_{23} \\ A_{31} & A_{32} & A_3 \end{pmatrix} \begin{pmatrix} \tau_{11} \\ \tau_{22} \\ \tau_{12} \end{pmatrix} = \begin{pmatrix} f_{11} \\ f_{22} \\ f_{12} \end{pmatrix} \tag{4.94}$$

où les coefficients des matrices sont définis par :

$$a_{1;i,j} = (\gamma_{11;i}, (\nu + (\vec{u}.\vec{\nabla})_h)\gamma_{11;j}) \tag{4.95}$$

$$a_{2;i,j} = (\gamma_{22;i}, (\nu + (\vec{u}.\vec{\nabla})_h)\gamma_{22;j}) \tag{4.96}$$

$$
\begin{aligned}
a_{3;i,j} &= 2(\gamma_{12;i}, (\nu + (\vec{u}.\vec{\nabla})_h)\gamma_{12;j}) && (4.97)\\
a_{13;i,j} &= a_{31;j,i} = (\gamma_{11;i}, (w - ad)\gamma_{12;j}) && (4.98)\\
a_{23;i,j} &= a_{32;j,i} = -(\gamma_{22;i}, (w + ad)\gamma_{12;j}) && (4.99)
\end{aligned}
$$

avec $w = \partial_1 u_2 - \partial_2 u_1$, $d = \partial_1 u_2 + \partial_2 u_1$, et $(\gamma_{lm;i})_{1\leq i\leq N_{lm}}$ est la base associée à la composante lm du tenseur, $lm \in \{11, 22, 12\}$. La composante f_{lm} du second membre est construite à partir du tenseur f et des degrés de liberté imposés sur la frontière amont, $lm \in \{11, 22, 12\}$. Remarquons que la composante de cisaillement est associée à un "poid" double, issu du produit scalaire de $L^2(\Omega)^4_s$.

Pour résoudre ce problème non symétrique, nous avons utilisé une méthode de relaxations successives symétrisées (SSOR), bloc par bloc, puis dans l'ordre inverse des blocs. Une itération de cette méthode s'énonce :

τ^n étant connu, on calcule τ^{n+1} en cinq étapes :

i) $\tau_{11}^{n+\frac{1}{2}}$ est obtenu en une itération SSOR sur le système :

$$A_1\tau_{11}^{n+\frac{1}{2}} = f_{11} - A_{13}\tau_{12}^n$$

ii) $\tau_{22}^{n+\frac{1}{2}}$ est obtenu en une itération SSOR sur le système :

$$A_2\tau_{22}^{n+\frac{1}{2}} = f_{22} - A_{23}\tau_{12}^n$$

i) τ_{12}^{n+1} est obtenu en deux itérations SSOR sur le système :

$$A_3\tau_{12}^{n+1} = f_{12} - A_{31}\tau_{11}^{n+\frac{1}{2}} - A_{32}\tau_{22}^{n+\frac{1}{2}}$$

ii) τ_{22}^{n+1} est obtenu en une itération SSOR sur le système :

$$A_2\tau_{22}^{n+1} = f_{22} - A_{23}\tau_{12}^{n+1}$$

ii) enfin, τ_{11}^{n+1} est obtenu en une itération SSOR sur le système :

$$A_1\tau_{11}^{n+1} = f_{11} - A_{13}\tau_{12}^{n+1}$$

Cette méthode a l'avantage de présenter un opérateur d'itération symétrique.

Chapitre 5

Expérimentations numériques

Nous avons rassemblé dans ce chapitre les résultats numériques obtenus par les méthodes décrites aux chapitres précédent. La présentation est faite de telle sorte qu'il est facile de comparer ces résultats avec ceux obtenus par d'autres méthodes [39, 15, 16].

Ce chapitre commence par une présentation des modèles viscoélastiques utilisés en considérant des problèmes monodimensionnels : écoulements de cisaillement simple, d'élongation, et de Poiseuille. Ces problèmes mettent en évidence certaines propriétés fondamentales des fluides viscoélastiques, ainsi que les dificultés caractérisques associées à ce type de problème.

Le problème de l'écoulement d'un fluide viscoélastique dans un domaine à contraction brusque est ensuite décrit : certaines grandeurs caractéristiques sont introduites : saut de pression au passage de la région d'entrée (en terme de bilan d'énergie), longueur et intensité des zones de recirculation (en terme de fonction de courant).

Présentant les performances du "solveur de Stokes", nous étudions le préconditionnement de l'algorithme du gradient conjugé par la méthode du Lagrangien augmenté (voir en annexe B) pour une géométrie de contraction, plane ou axisymétrique. L'étude du coût asymptotique et de la convergence de la méthode d'éléments finis mixtes de Thomas-Raviart sera ensuite abordée en considérant deux familles de maillages de ce type de domaine.

Expérimentant l'algorithme des directions alternées (θ-méthode) présenté au chapitre 3, nous étudions l'apect qualitatif des solutions : lorsque le nombre de Weissenberg varie, nous étudierons les les contraintes élongationelles et les profils de vitesse sur l'axe, ainsi que la longueur et l'intensité des zones de recirculation. La convergence de la méthode d'éléments finis pour les grands nombres de Weissenberg est aussi abordée.

5.1 Etudes des modèles en écoulements simples

Afin de mettre en évidence certaines propriétés caractéristiques des modèles visco-élastiques, et les limitations de la famille de modèles considérée, nous étudierons tout d'abord quelques écoulements simples.

5.1.1 Ecoulements viscosimétriques

Régime permanent

Considérons l'écoulement en cisaillement simple décrit par la figure 5.1

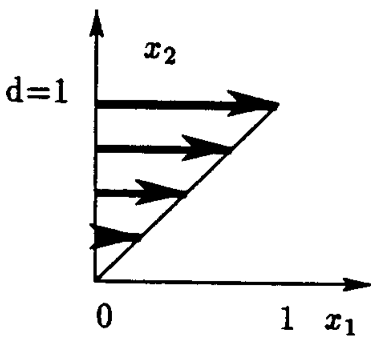


Figure 5.1: Ecoulement en cisaillement simple

Le champ des vitesses est donné par $u(x_1, x_2) = \dot{\gamma} x_2$. De l'équation constitutive du modèle d'Oldroyd (1.17), nous obtenons la contrainte de cisaillement : $\tau_{12}(x_1, x_2) = \alpha\dot{\gamma}$. La *viscosité de cisaillement*, notée η_s est définie par : $\eta_s = \tau_{12}/\dot{\gamma}$ et nous obtenons :

$$\eta_s = \alpha \tag{5.1}$$

La *première différence des contraintes normales*, notée N_1, est définie par $N_1 = \tau_{11} - \tau_{22}$, et, pour le modèle d'Oldroyd :

$$N_1 = 2\alpha\dot{\gamma}^2 \tag{5.2}$$

Ce modèle nous permet donc d'introduire des efforts normaux, inexistants pour les fluides newtoniens. Cependant, la relation (5.2) correspond à une loi experimentale obtenue lorsque $\dot{\gamma}$ reste petit. Pour les grandes valeurs du taux de cisaillement $\dot{\gamma}$, cette relation ne coïncide plus avec les résultats expérimentaux sur certains fluides viscoélastiques.

Régime transitoire

Considérons un démarrage en cisaillement simple :
pour $t \leq 0$, $u(t, x_1, x_2) = 0$ et pour $t > 0$, $u(t, x_1, d) = \dot{\gamma} d$. La contrainte de cisaillement est de la forme :

$$\tau_{12}(t, x_1, x_2) = \alpha \frac{\partial u}{\partial y}(t, x_1, x_2)(1 - e^{-\frac{t}{We}})$$

et la *viscosité de démarrage en cisaillement*, notée η_s^+, est définie par :

$$\eta_s^+ = \frac{\tau_{12}}{\frac{\partial u}{\partial t}}$$

Nous obtenons ici :

$$\eta_s^+(t) = \alpha(1 - e^{-\frac{t}{We}}), \ t \geq 0 \tag{5.3}$$

De même, lors d'un arrêt en cisaillement simple : $u(t, x_1, x_2) = \dot{\gamma} x_2, \ t \leq 0$ et $u(t, x_1, d) = 0, \ t > 0$, nous obtenons une *viscosité d'arrêt en cisaillement* :

$$\eta_s^- = \alpha e^{-\frac{t}{We}}, \ t \geq 0 \tag{5.4}$$

Ainsi, le modèle utilisé permet d'introduire un temps de relaxation (ssocié à We). Cependant, pour certains fluides réels, les expériences de démarrage en cisaillement simple font apparaître un dépassement (overshoot, en anglais) de la viscosité de cisaillement asymptotique η_s, alors qu'ici η_s^+ est strictement croissante dans le temps.

5.1.2 Ecoulement en élongation

Régime permanent

Un tel écoulement est obtenu pour un tenseur taux-de-déformation diagonal :

$$D(\vec{u}) = \dot{\varepsilon} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} \tag{5.5}$$

Dans le cas de l'équation constitutive du modèle d'Oldroyd (1.17), les contraintes normales sont :

$$\tau_{11} = \frac{2\alpha\dot{\varepsilon}}{1 - 2a\,We\,\dot{\varepsilon}} \tag{5.6}$$

$$\tau_{22} = \tau_{33} = \frac{-\alpha\dot{\varepsilon}}{1 + a\,We\,\dot{\varepsilon}} \tag{5.7}$$

La *viscosité élongationnelle*, notée η_e, définie par :

$$\eta_e = \frac{\tau_{11} - \tau_{22}}{\dot{\varepsilon}}$$

vaut ici :

$$\eta_e(\dot{\varepsilon}) = \frac{3\alpha}{(1 + a\,We\,\dot{\varepsilon})(1 - 2a\,We\,\dot{\varepsilon})} \tag{5.8}$$

et les variations de η_e en fonction de $\dot{\varepsilon}$ sont représentées sur la figure 5.2 pour $a\,We = 1$ et $\alpha = 1$. Ainsi η_e augmente avec le taux délongation $\dot{\varepsilon}$, jusqu'à

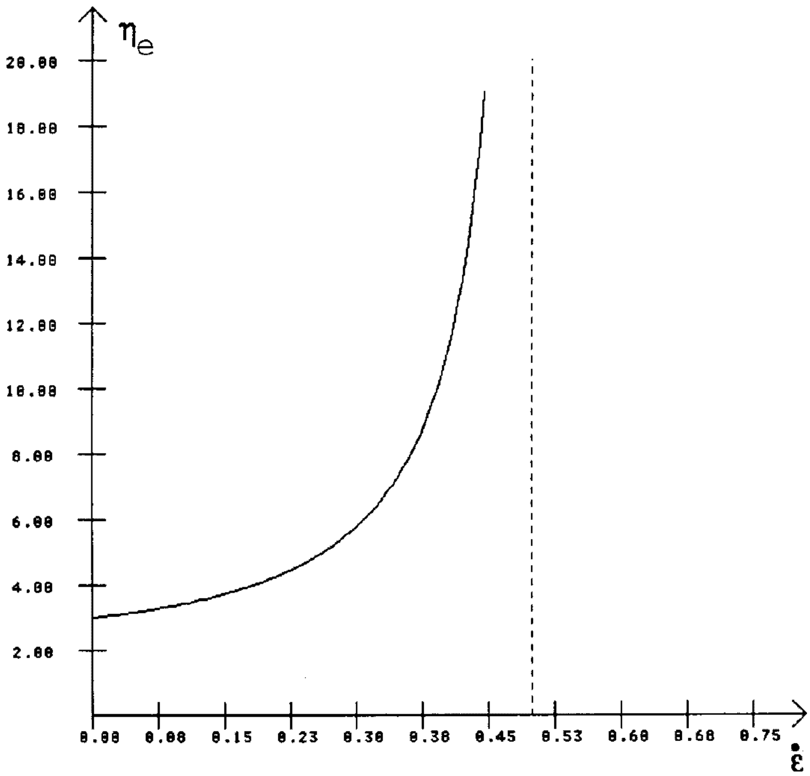


Figure 5.2: Viscosité élongationnelle du modèle d'Oldroyd

une valeur critique $\dot{\varepsilon}_c = \frac{1}{2}a\,We$, où η_e est infini. Au delà de cette valeur critique, la viscosité élongationnelle donnée par le modèle est négative, ce qui n'a plus de sens physique.

Remarquons que pour le modèle newtonien ($We = 0$), nous retrouvons la *viscosité élongationnelle de Trouton* : $\eta_e = 3\alpha$.

Démarrage en élongation

Considérons à présent le cas où $D(\vec{u}) = 0$, $t \leq 0$, et $D(\vec{u})$ est donné par (5.5) pour $t > 0$. Les contraintes normales sont alors :

$$\tau_{11} = \frac{2\alpha\dot{\varepsilon}}{1 - 2a\,We\,\dot{\varepsilon}}\left(1 - e^{-(1-2a\,We\,\dot{\varepsilon})\frac{t}{We}}\right) \tag{5.9}$$

$$\tau_{22} = \tau_{33} = \frac{-alpha\dot{\varepsilon}}{1 + a\,We\,\dot{\varepsilon}}\left(1 - e^{-(1+a\,We\,\dot{\varepsilon})\frac{t}{We}}\right) \tag{5.10}$$

et la *viscosité de démarrage en élongation*, notée η_e^+, est ici :

$$\eta_e^+(t,\dot{\varepsilon}) = \eta_e(\dot{\varepsilon}) - \alpha\left(\frac{2}{1-2a\,We\,\dot{\varepsilon}}e^{-(1-2a\,We\,\dot{\varepsilon})\frac{t}{We}} + \frac{1}{1+a\,We\,\dot{\varepsilon}}e^{-(1+a\,We\,\dot{\varepsilon})\frac{t}{We}}\right) \tag{5.11}$$

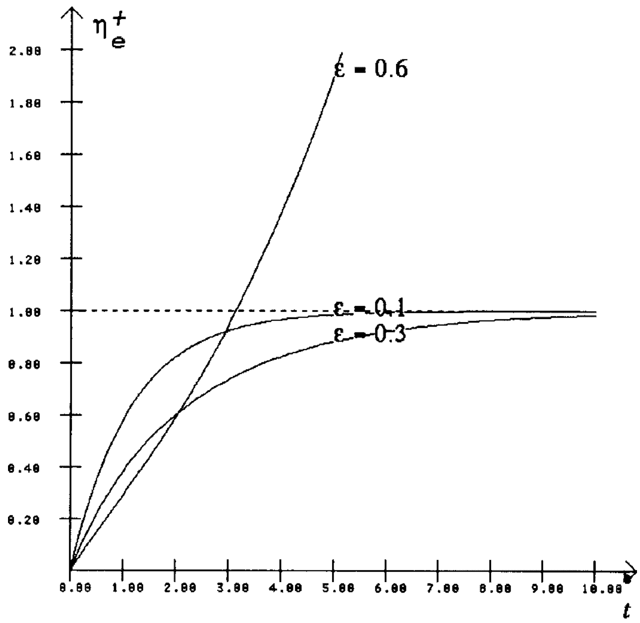


Figure 5.3: Démarrage en élongation pour le modèle d'Oldroyd

Ainsi, lorsque $\dot{\varepsilon} < \dot{\varepsilon}_c$, $\eta_e^+(t,\dot{\varepsilon})$ augmente jusqu'à attendre un plateau de valeur $\eta_e(\dot{\varepsilon})$. Par contre, lorsque $\dot{\varepsilon} > \dot{\varepsilon}_c$, la viscosité de démarrage en élongation augmente indéfiniment, et ce comportement ne correspond plus à des fluides réels. Ces variations sont représentées sur la figure 5.3 pour $a = We = \alpha = 1$. Le modèle d'Oldroyd, s'il permet de traduire les effets élongationnels de fluides réels, reste limité à de petites valeurs du taux d'élongation.

5.1.3 Ecoulements de Poiseuille

Guilloppé et Saut [28] ont étudié les écoulements de Poiseuille pour le modèle d'Oldroyd, dont nous rappelons quelques résultats , et présentons une méthode de calcul des solutions en régime permanent. Considérons donc un écoulement permanent de Poiseuille dans un conduit de rayon unité. La vitesse est de la forme $\vec{u}(x_1, x_2) = (u(x_2), 0)$. L'équation consitutive du modèle d'Oldroyd (1.17) conduit à :

$$\tau_{12} = \frac{\alpha\dot{\gamma}}{1 + (1 - a^2)We^2\,\dot{\gamma}^2} \tag{5.12}$$

$$\tau_{11} = \frac{\alpha(1 + a)We\,\dot{\gamma}^2}{1 + (1 - a^2)We^2\,\dot{\gamma}^2} \tag{5.13}$$

$$\tau_{22} = -\frac{\alpha(1 - a)We\,\dot{\gamma}^2}{1 + (1 - a^2)We^2\,\dot{\gamma}^2} \tag{5.14}$$

où on a posé $\dot{\gamma} = \frac{\mathrm{d}u}{\mathrm{d}x_2}$.

Soit $f > 0$ la force de pression exercée dans la conduite. La conservation de la quantité de mouvement (1.18) se ramène à :

$$-(1 - \alpha)\frac{\mathrm{d}^2 u}{\mathrm{d}{x_2}^2} - \frac{\mathrm{d}\tau_{12}}{\mathrm{d}x_2} - f = 0 \text{ dans }]0, 1[\tag{5.15}$$

à laquelle nous associons les conditions aux limites de type Dirichlet-Neumann :

$$\frac{\mathrm{d}u}{\mathrm{d}x_2}(0) = u(1) = 0 \tag{5.16}$$

En intégrant (5.15) avec les conditions aux limites (5.16), et en utilisant (5.12), nous obtenons l'équation caractéristique :

$$(1 + (1 - \alpha)(1 - a^2)We^2\,\dot{\gamma}^2)\dot{\gamma} = (1 + (1 - a^2)We\,\dot{\gamma}^2)fx_2 = 0 \tag{5.17}$$

Cas des dérivées convectées supérieures et inférieures

Lorsque $|a| = 1$, la solution est :

$$\tau_{12}(x_2) = -\alpha f x_2 \tag{5.18}$$

$$\tau_{11}(x_2) = \alpha(1 + a)We\,f^2 x_2^2 \tag{5.19}$$

$$\tau_{22}(x_2) = -\alpha(1 - a)We\,f^2 x_2^2 \tag{5.20}$$

$$u(x_2) = \frac{f}{2}(1 - x_2^2) \tag{5.21}$$

$$p(x_2) = -f x_1 \tag{5.22}$$

Cas des dérivées intermédiaires

Suivant [28], nous posons $k = We\sqrt{1-a^2}$, $x = kfx_2$ et

$$\delta(x) = -k\dot{\gamma}(\frac{x}{kf}) \tag{5.23}$$

de telle sorte que l'équation caractéristique (5.17) s'écrive :

$$(1 + (1-\alpha)\delta^2)\delta = (1+\delta^2)x \tag{5.24}$$

Les courbes cractéristiques $\delta = \delta(x)$ sont représentées sur la figure 5.4 pour quelque valeurs remarquables de α. Lorsque $\alpha > 8/9$, les courbes admettent

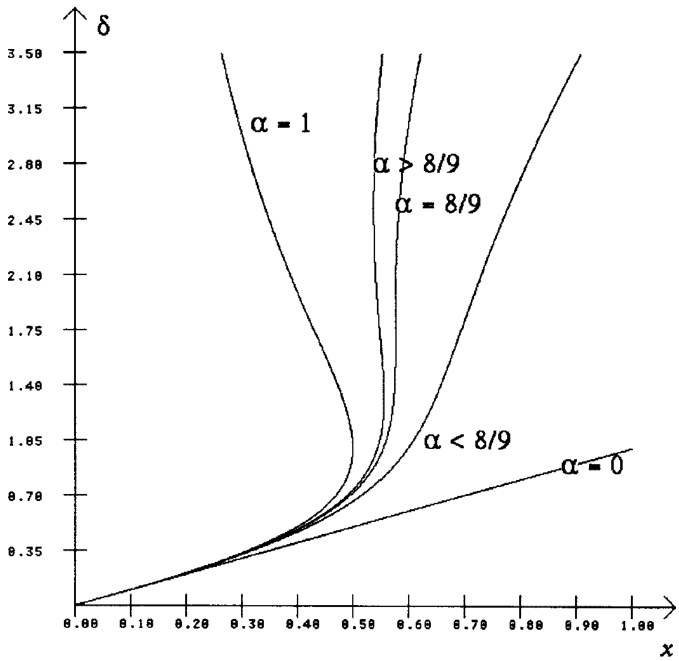


Figure 5.4: Courbes caractéristiques de l'écoulement de Poiseuille

des points de rebroussement, donnés par les racines de l'équation bicarrée :

$$\alpha(3+\delta^2)\delta^2 = (1+\delta^2)^2 \tag{5.25}$$

dont le discriminant est $\alpha(9\alpha - 8)$.

Pour $\alpha = 0$, nous retrouvons la solution newtonienne (figure 5.5.a). Pour $0 < \alpha < 8/9$, le problème admet une unique solution $u \in C^\infty(0,1)$ (figure 5.5.b). Pour $\alpha = 8/9$, δ n'est pas dérivable en $x = 1/\sqrt{3}$, et donc, si

$3k^2f^2 > 1$, la solution u sera seulement de classe C^1 (figure 5.5.c). Dans le cas contraire, la solution est de classe C^∞. Lorsque $1 < \alpha < 8/9$, la courbe caractéristique admet deux points de retournement, aux abscisses notées x_m et x_M, $0 < x_m < x_M$, et racines de l'équation bicarrée (5.25). Alors, si $kf \leq x_m$, le problème admet une solution unique, de classe C^∞, si $x_m < kf \leq x_M$, le problème admet un continuum de solutions u de classe C^0, dont une est de classe C^∞ (figure 5.5.d), et si $kf > x_M$, le problème admet un continuum de solutions u de classe C^0, mais aucune n'est plus régulière (figure 5.5.e). Enfin, pour $\alpha = 1$, si $kf \leq \frac{1}{2}$, il existe une unique solution finie $u \in C^\infty(0,1)$ (figure 5.5.f), tandis que si $kf > \frac{1}{2}$, il n'existe aucune solution. Nous avons représenté en pointillé la solution newtonienne associée à $We = 0$, afin de pouvoir comparer les débits obtenus.

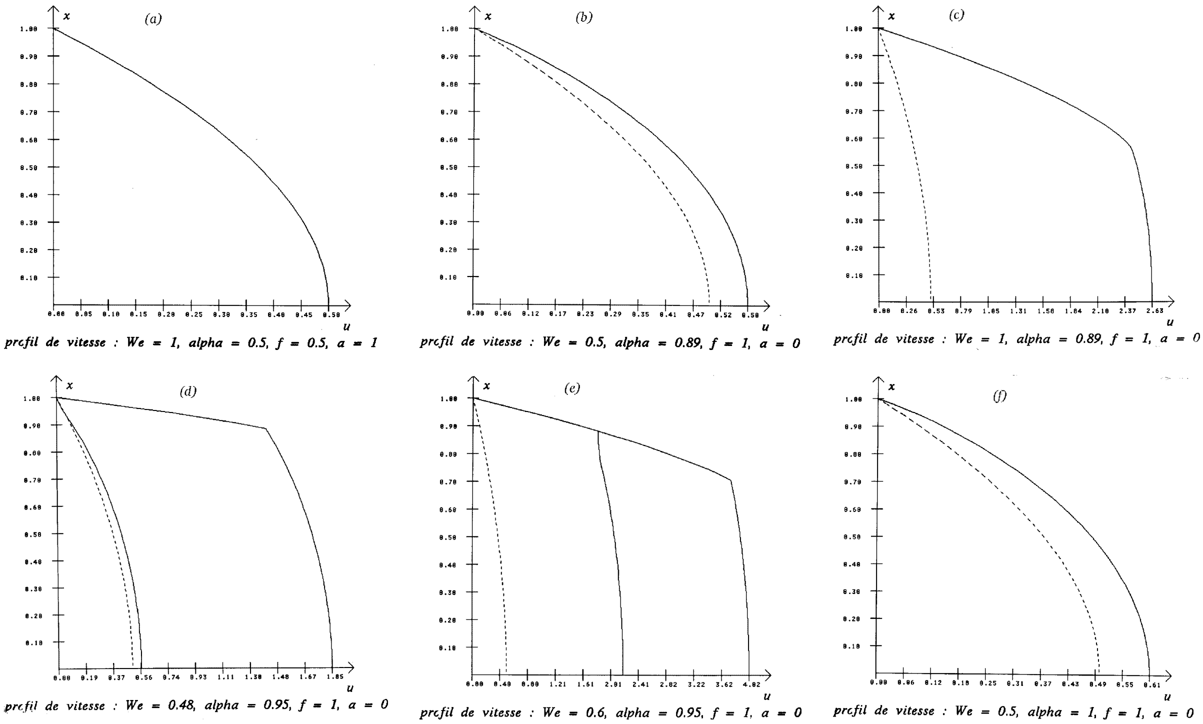


Figure 5.5: Profils de vitesse pour l'écoulement de Poiseuille

Remarquons que l'équation caractéristique (5.17) correspond, à x fixé, à une équation de degré inférieur ou égal à trois en δ : si $\alpha \in [8/9, 1[$ et $x \in [x_m, x_M]$, cette équation admet trois racines réeles; si $\alpha = 1$ et $x \leq \frac{1}{2}$, deux racines réeles; si $\alpha = 1$ et $x > \frac{1}{2}$, aucune racine; enfin, dans tous les autres cas, il exite une unique racine.

En choisissant la plus petite des racines, si elle existe, il est donc possible de calculer directement une solution, qui sera la solution régulière, dans le cas où une telle solution existe. Lorsqu'il y a multiplicité des racines, nous avons choisis la plus grande, pour les calculs de la figure 5.5.

Pour cela, considérons une subdivision $(x_{2;i})_{0\leq i\leq N+1}$ de $[0,1]$, $x_{2;i} = ih$, de pas $h = \frac{1}{N+1}$. Posant $x_i = kfx_{2;i}$, nous pouvons calculer $\delta_i = \delta(x_i)$ d'après (5.17). Les valeurs exactes du tenseur des contraintes aux noeuds de la subdivision se déduisent de (5.12)-(5.14) et (5.23). Pour calculer une solution approchée u_h, affine sur chaque intervalle $[x_{2;j-1}, x_{2;j}]$, nous utiliserons par exemple la formule des trapèzes :

$$\int_{(j-1)h}^{jh} \psi(x)\,dx \approx \frac{h}{2}(\psi(jh) + \psi((j-1)h))$$

et définirons u_h par ses valeurs u_i au noeuds de la subdivision, soit :

$$u_i = \frac{h}{2k}\left(\sum_{j=i}^{N} \delta_j + \delta_{j+1}\right) \tag{5.26}$$

Conclusion

Les solutions non-régulières conduisent à un débit asymptotiquement infini lorsque $\alpha \to 1$, ce qui constitue une limite du modèle physique. D'autre part, pour $\alpha > 8/9$, $|a| < 1$ et f fixé, il existe un point limite We_{max} au delà duquel il n'existe plus de solution ne sont régulière. Dans le cas $|a| = 1$, les solutions restent régulières pour toute valeur de α et We.

L'étude de la stabilité de ces solutions est présentée dans [28].

5.2 Présentation du problème de la marche

L'extrusion de matériaux viscoélastiques à travers un orifice est utilisée très souvent lors de la mise en forme des matériaux, et a été étudié depuis de nombreuses années.

Dans la *région d'entrée*, située à l'amont d'un orifice axisymétrique, le fluide subit une *accélération* lorsqu'il passe d'un conduit large à un conduit de section plus petite.

L'écoulement est approximativement *viscosimétrique* près des parois, et *élongationnel* le long de l'axe : il est donc de type *complexe*. Les fonctions viscosimétriques et d'élongation sont amenées à être des mécanismes importants de l'écoulement dans la région d'entrée.

S'inspirant des études expérimentales présentés dans [49], nous présentons ici quelques paramètres sans dimension succeptibles de décrire les écoulements de matériaux viscoélastiques dans un domaine à contraction brusque. Ces paramètres seront utilisés pour présenter les résultats numériques obtenus.

5.2.1 Description de l'écoulement

Des résultats numériques et expérimentaux obtenus pour les fluides newtoniens lents mettent en évidence une recirculation dont la taille et la forme ne dépendent que du taux de contraction :

$$C = \frac{r_1}{r_2} \tag{5.27}$$

où r_1 (resp. r_2) est le rayon de la section amont (resp. aval) (voir figure 5.6). Dans ce cas, la recirculation ressemble à un tore concave, limité par la paroi du tube amont, la parois de la contraction, et le coin r'́entrant. Cette recirculation s'étend à l'amont jusqu'à une distance L_1, et on note X le rapport de L_1 et du diamètre D :

$$X = \frac{L_1}{2r_1} \tag{5.28}$$

Pour les taux de contraction plus grand que 4, ce rapport reste sensiblement constant :

$$X_0 \approx 0.17 \tag{5.29}$$

Ce résultat s'applique encore à certains écoulements viscoélastiques, à condition que le débit reste suffisament petit, afin de maintenir la validité de l'approximation newtonienne.

L'angle β, sur la figure 5.6, représente l'*angle d'entrée du fluide*.

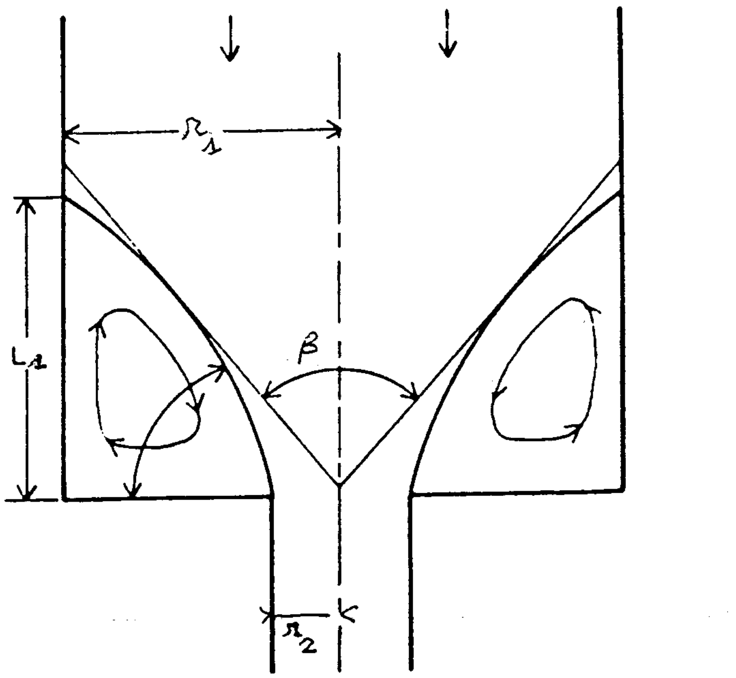


Figure 5.6: Description de l'écoulement dans une contraction brusque

Le calcul du champ des vitesses dans une contraction circulaire a été réalisé [39] pour des fluides purement visqueux dont la viscosité dépend du taux de cisaillement (loi de Carreau (1.9)). L'effet de la dépendance en cisaillement est de réduire l'extension du tore concave : pour des débits importants, la recirculation s'éloigne du coin ré-entrant pour se replier dans le coin sortant entre le tube large et la paroi, et ce, jusqu'à ce qu'il n'y ait quasiment plus de recirculation. La dimension de la zone de recirculation diminue aussi lorsque le nombre de Reynolds (représentant l'inertie) augmente, ou lorsque l'index de puissance r de la loi (1.9) diminue.

À titre d'exemple, de grandes zones de recirculation ont été observées expérimentalement pour de nombreux fluides viscoélastiques, [49]. Ces effets résultent probablement de propriètes élastiques du fluide, et peuvent être interprétés en terme de propriètés élongationnelles : les fluides dont la viscosité élongationnelle dépend du taux d'élongation mettent en évidence de grandes zones de recirculation, alors que des fluides analogues, à viscosité élongationnelle constante, développent de petites recirculations. La taille de la recirculation dépend de l'amplitude avec laquelle la viscosité élongationnelle dépasse la valeur Troutonnienne. Lorsque cette amplitude est très grande, les contraintes dominantes de l'écoulement dans la région de l'orifice sont les contraintes d'élongation, et le fluide entrant dans l'orifice vient surtout de la région axiale. Le fluide aux alentours est alors soumis à des contraintes plus faibles, et peut circuler seulement localement, comme un processus invoquant deux fluides différents.

5.2.2 Conditions aux limites

Nous restreindrons ici cette étude aux écoulements dans un domaine à contraction brusque comportant un plan ou un axe de symétrie. Nous pouvons par conséquent nous ramener au demi-domaine Ω correspondant à $x_2 > 0$ (marche, voir figure 5.7).

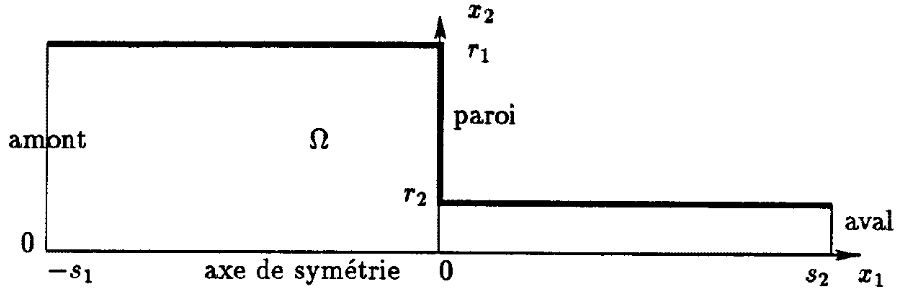


Figure 5.7: Domaine à contraction $C = 4$

Du point de vue de la loi de comportement, seul le cas de la dérivée convectée supérieure ($a = 1$) est abordé. En supposant l'écoulements de Poiseuille établi aux sections amont : $x_1 = s_1$ et aval : $x_1 = s_2$ (autrement dit, le domaine est "suffisament" long), nous obtenons de (5.21) les conditions de type Dirichlet aux sections amont et aval de la forme :

$$u_1(s_i, x_2) = \frac{f_i}{2}\left(1 - (\frac{x_2}{r_i})^2\right), \ x \in]0, r_i[, \ i = 1, 2 \tag{5.30}$$

où $f_1, f_2 > 0$. Le long de l'axe de symétrie, la condition est de type mixte Dirichlet-Neumann :

$$\frac{\partial u_1}{\partial x_2}(x_1, 0) = u_2(x_1, 0) = 0, \ s_1 \leq x_1 \leq s_2 \tag{5.31}$$

et $\vec{u} = 0$ le long de la parois.

De (5.30) et de la condition de compatibilité (1.13) traduisant la conservation du débit :

$$q_1 = q_2$$

où :

$$q_i = \begin{cases} f_i r_i/3 & \text{en géométrie plane} \\ f_i r_i^2/8 & \text{en géométrie axisymétrique} \end{cases} \tag{5.32}$$

nous obtenons :

$$f_2 = \begin{cases} C f_1, & \text{en géométrie plane} \\ C^2 f_1, & \text{en géométrie axisymétrique} \end{cases} \tag{5.33}$$

Les contraintes à la section amont sont alors données par :

$$\tau_{11}(s1, x_2) = 2\alpha We\, f_1^2 x_2^2 / r_1^4 \tag{5.34}$$

$$\tau_{22}(s1, x_2) = 0 \tag{5.35}$$

$$\tau_{12}(s1, x_2) = -\alpha f_1 x_2 / r_1^2 \tag{5.36}$$

ceci $\forall x_2,\ 0 \le x_2 \le r_1$.

Notant τ_w la contrainte de cisaillement à la section aval le long de la paroi, un adimensionnement intéréssant correspond à choisir f_2 tel que $|\tau_w| = 1$. Nous obtenons ainsi de (5.12) et (5.33) :

$$f_1 = \begin{cases} r_2/C, & \text{en géométrie plane} \\ r_2/C^2, & \text{en géométrie axisymétrique} \end{cases} \tag{5.37}$$

5.2.3 Chute de pression

La pression physique, notée $\bar{p}$, est définie par :

$$\bar{p} = -\frac{1}{n} \mathrm{tr}\, \sigma \tag{5.38}$$

où n est le nombre de dimension ($n = 2$ en géométrie plane), et $n = 3$ en géométrie axisymétrique), et σ le tenseur des contraintes totales de Cauchy.

La chute de pression dans le domaine Ω est alors définie par :

$$\delta p = \frac{1}{Q} \int_\Gamma \bar{p} \vec{u}.\vec{\nu}\, d\Gamma \tag{5.39}$$

où Q est le débit physique :

$$Q = \begin{cases} q_1 = q_2 & \text{en géométrie plane} \\ \pi q_1 = \pi q_2 & \text{en géométrie axisymétrique} \end{cases}$$

En intégrant la relation (5.38) par partie, du fait de la relation d'incompressibilité (1.19), et utilisant l'expression (1.1) du tenseur de Cauchy, il vient :

$$\delta p = \frac{1}{Q} \left(\int_\Gamma \vec{\nabla} p.\vec{u}\, dx - \frac{1}{n} \int_\Gamma \mathrm{tr}\, \tau (\vec{u}.\vec{\nu})\, d\Gamma \right)$$

Par la relation de conservation de la quantité de mouvement (1.18), nous pouvons exprimer δp uniquement en fonction de $\vec{u}$ et τ. Faisant apparaître le tourbillon $\mathrm{rot}\,\vec{u}$ par l'identité (2.48), nous appliquons la formule de Green (2.51) pour obtenir :

$$\begin{aligned}\delta p = &-\frac{1-\alpha}{Q}\left(\|\mathrm{rot}\,\vec{u}\|^2_{0,\Omega} - \int_\gamma \mathrm{rot}\,\vec{u}(\vec{u}.\vec{\tau})\,d\Gamma\right)\\ &+\frac{1}{Q}\left((\vec{\mathrm{div}}\,\tau,\vec{u}) - \frac{1}{n}\int_\Gamma \mathrm{tr}\,\tau(\vec{u}.\vec{\nu})\,d\Gamma\right)\end{aligned} \tag{5.40}$$

Pour un écoulement de Poiseuille dans le domaine amont : $\Omega_1 =]-s_1, 0[\times]0, r_1[$ (resp. aval : $\Omega_2 =]0, s_2[\times]0, r_2[$) nous obtenons :

$$\delta p_i = \begin{cases} -s_i f_i/r_i^2 & \text{en géométrie plane} \\ -2s_i f_i/\pi r_i^2 & \text{en géométrie axisymétrique} \end{cases} \tag{5.41}$$

La chute de pression due au passage du fluide dans la région d'entrée est alors définie par :

$$\delta p_c = \delta p - (\delta p_1 + \delta p_2)$$

Lorsque les longueurs des sections amont et aval tendent vers l'infini, cette quantité tend vers une valeur ne dépendant que de la nature du fluide.

Remarquant que

$$\delta p_2/\tau_w = \begin{cases} s_2 & \text{en géométrie plane} \\ 2s_2 & \text{en géométrie axisymétrique} \end{cases} \tag{5.42}$$

nous pouvons exprimer cette perte en terme de longeur équivalente :

$$L = \begin{cases} \delta p_c/\tau_w & \text{en géométrie plane} \\ \delta p_c/2\tau_w & \text{en géométrie axisymétrique} \end{cases} \tag{5.43}$$

Ainsi, un écoulement de Poiseuille de débit q_2, en régime permenent dans un conduit de longeur L et de rayon r_2 provoque une variation de pression équivalente à celle obtenu par le passage du fluide dans la région d'entrée.

Pour de faibles nombres de Weissenberg, ce paramètre tend vers une valeur constante, correspondant à un écoulement de fluide newtonien :

$$L_0 \approx 0.59 \tag{5.44}$$

5.2.4 Intensité des recirculations

La fonction de courant, notée ϕ, est caractérisée par le problème variationnel : *trouver* $\phi \in H^1(\Omega)$, $\phi = \phi_\Gamma$ *sur* Γ *telle que :*

$$(\vec{\mathrm{rot}}\,\phi, \vec{\mathrm{rot}}\,\xi) = (\vec{u}, \vec{\mathrm{rot}}\,\xi),\ \forall \xi \in H_0^1(\Omega) \tag{5.45}$$

où $\vec{u}$ est le champ des vitesses. et ϕ_Γ est la condition aux frontières.

Les conditions aux sections amont et aval sont :

$$\phi_\Gamma(s_i, x_2) = q\varphi(\frac{x_2}{r_i}),\ 0 \leq x_2 \leq r_i \tag{5.46}$$

où :

$$\varphi(x) = \begin{cases} \frac{1}{2}(x-1)^2(x+2) & \text{en géométrie plane} \\ (x-1)^2(x+1)^2 & \text{en géométrie axisymétrique} \end{cases} \tag{5.47}$$

La condition le long de l'axe de symétrie est

$$\phi_\Gamma(x_1, 0) = \phi_0,\ s_1 \leq x_1 \leq s_2 \tag{5.48}$$

où $\phi_0 = q$.

Dans la suite, nous représenterons l'écoulement par les lignes isovaleurs de ϕ/ϕ_0. L'écoulement principal correspond domaine $\phi > 0$, tandisque la zone de recirculation est caractérisée par $\phi < 0$, et la ligne de séparation interieure par $\phi = 0$.

L'*intensité des recirculations* est alors définie par

$$\psi = -\frac{1}{Q} \min_{x \in \Omega} \phi(x) \tag{5.49}$$

et nous représenterons les variations de ψ en fonction de We. Lorsque le champ de vitesse approché $\vec{u}_h \in V_h$, la fonction de courant associée, notée ϕ_h, appartient à l'espace Θ_h, et nous noterons ψ_h l'intensité des recirculations obtenues.

5.2.5 Triangulations et méthode d'éléments finis

Le domaine d'écoulement Ω considéré ici correspond à $r_1 = C = 4$, $r_2 = 1$, $s_1 = 64$ et $s_2 = 150$ (le choix de s_1 et s_2 sera justifié dans ce qui suit). Nous considérons ici deux familles de triangulation du domaine $\bar{\Omega}$, et la méthode d'élements finis de plus bas degré ($l = 1$) associée au choix des espaces de dimension finie T_h, Y_h et Q_h, donnés respectivement par (2.139), (2.124) et (2.132).

Partant d'une triangulation grossière $\mathcal{T}_0^{(1)}$, la première famille, notée $(\mathcal{T}_i^{(1)})_{i\geq 0}$, est obtenue par raffinements uniformes successifs d'un facteur (approximatif) de $1/\sqrt{2}$. Les maillages $\mathcal{T}_i^{(1)}$, $0 \leq i \leq I = 3$ sont représentés sur la figure 5.8. La longeur des arêtes des éléments voisins du coin ré-entrant sur une triangulation $\mathcal{T}_i^{(1)}$, $i \geq 0$, sont donnés par la relation :

$$h_i^{(1)} = \frac{1}{3 \times 2^{i/2}}$$

La deuxième famille de triangulation, notée $(\mathcal{T}_i^{(2)})_{i\geq 0}$, est obtenue par raffinements successifs de la triangulation grossière $\mathcal{T}_0^{(2)} = \mathcal{T}_0^{(1)}$ au voisinage du coin ré-entrant, et de telle sorte que $\mathcal{T}_i^{(1)}$ et $\mathcal{T}_i^{(2)}$ aient le même nombre d'éléments, et que la longeur des arêtes des éléments voisins du coin ré-entrant sur la triangulation $\mathcal{T}_i^{(2)}$, $i \geq 0$, soit donnée par la relation :

$$h_i^{(2)} = \frac{1}{3 \times 2^{i}}$$

Les maillages $\mathcal{T}_i^{(2)}$, $i = 0, \ldots, 4$ sont représentés sur la figure 5.9.

D'autre part, la longeur des arêtes des éléments augmente selon une progression géométrique de facteur constant lorsque on s'éloigne d'un voisinage du coin ré-entrant. La table 5.1 indique les pas $h_i^{(1)}$, $h_i^{(2)}$, les nombres de faces NF_i, de sommets NS_i, de degrés de liberté N_i, ainsi que le rapport N_i/NF_i. Remarquons que ce rapport tend vers 6 lorsque i devient grand; or la méthode d'éléments finis approche 6 champs scalaires : τ_{11}, τ_{22}, τ_{12}, u_1, u_2 et p. Par conséquent la méthode utilise asymptotiquement *un degré de liberté par élément et par champ scalaire*, ce qui est optimal. Ceci peut

i	$h_i^{(1)}$	$h_i^{(2)}$	NF_i	NS_i	N_i	N_i/NF_i
0	0.3333	0.3333	144	184	945	6.56
1	0.2357	0.1667	299	356	1909	6.38
2	0.1667	0.0833	576	655	3615	6.27
3	0.1179	0.0416	1196	1309	7403	6.18

Table 5.1: Caractéristiques des deux familles de triangulations

être déduit de la relation donnant le nombre total de degré de liberté pour la méthode d'éléments finis utilisée :

$$N_i = \dim(T_h \times Y_h \times Q_h) = 3NF_i + NS_i + NA_i$$

où NA est le nombre d'arêtes, donné par la relation d'Euler :

$$NA_i = NF_i + NS_i - 1$$

Remarquant que lorsque i devient grand, $NS_i \approx NF_i$, nous obtenons :

$$N_i \approx 6NF_i$$

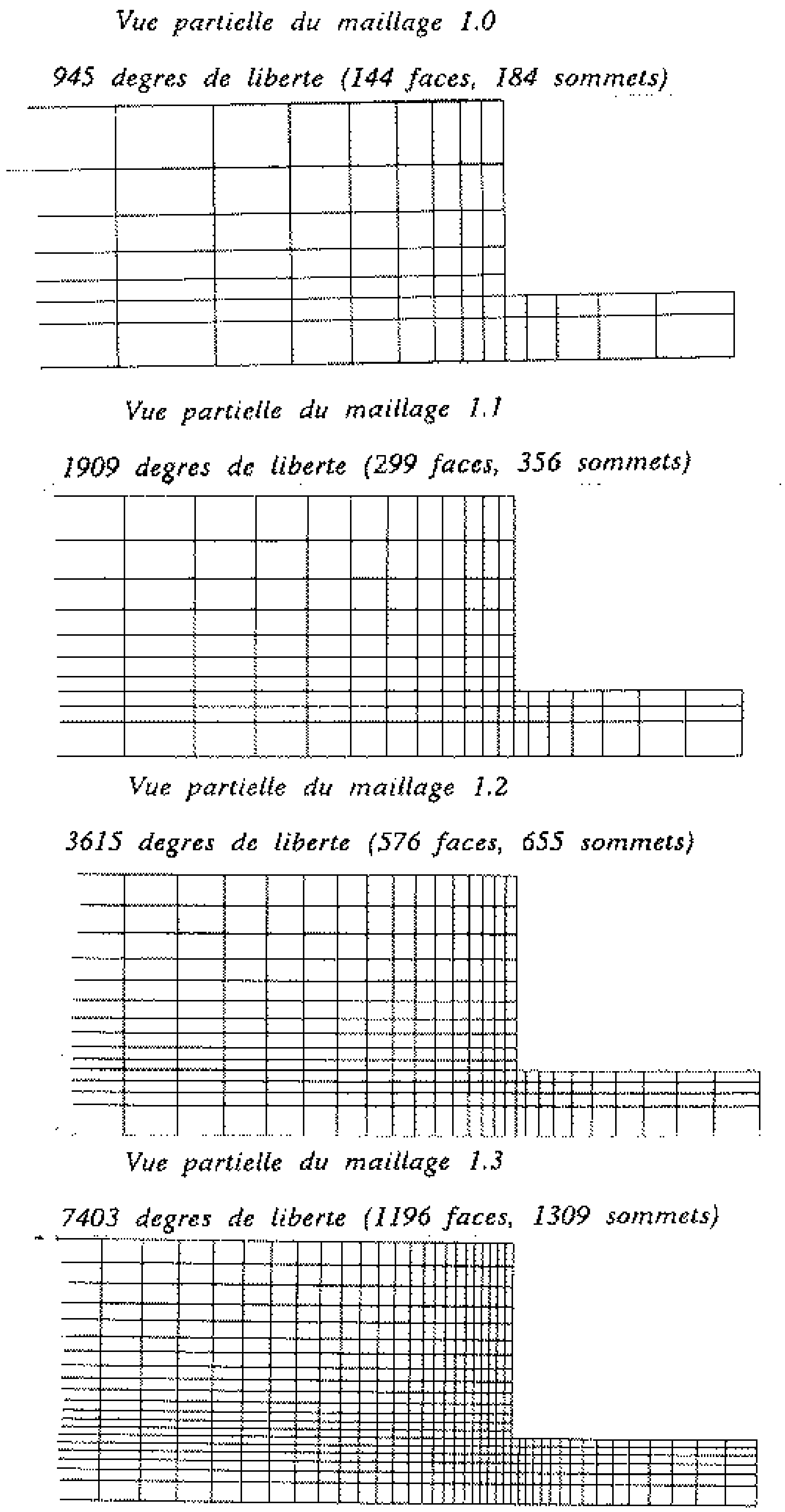


Figure 5.8: Une première famille de maillages du domaine (vues partielles)

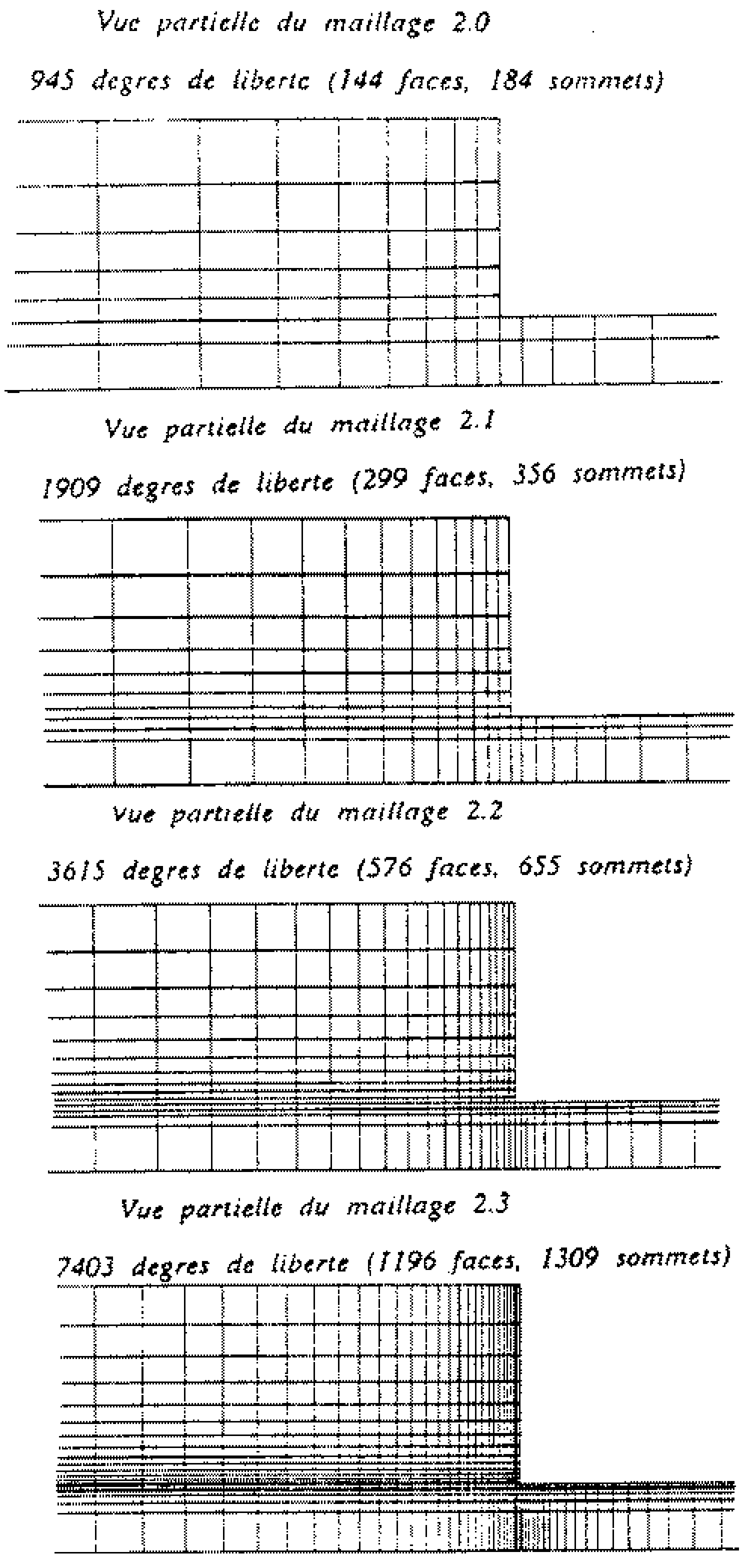


Figure 5.9: Une deuxième famille de maillages du domaine (vues partielles)

5.3 Etude de la résolution du problème de Stokes

Nous présentons ici quelques résultats concernant la résolution rapide du problème de Stokes, et les écoulements d'un fluide newtonien dans une marche, en géométrie plane et axisymétrique.

5.3.1 Préconditionnement

Pour une description de l'algorithme du gradient conjugé préconditionné, le lecteur pourra se rapporter à [17] ou à l'annexe B. La figure 5.10 met en évidence l'accélération de la convergence de l'algorithme par l'utilisation

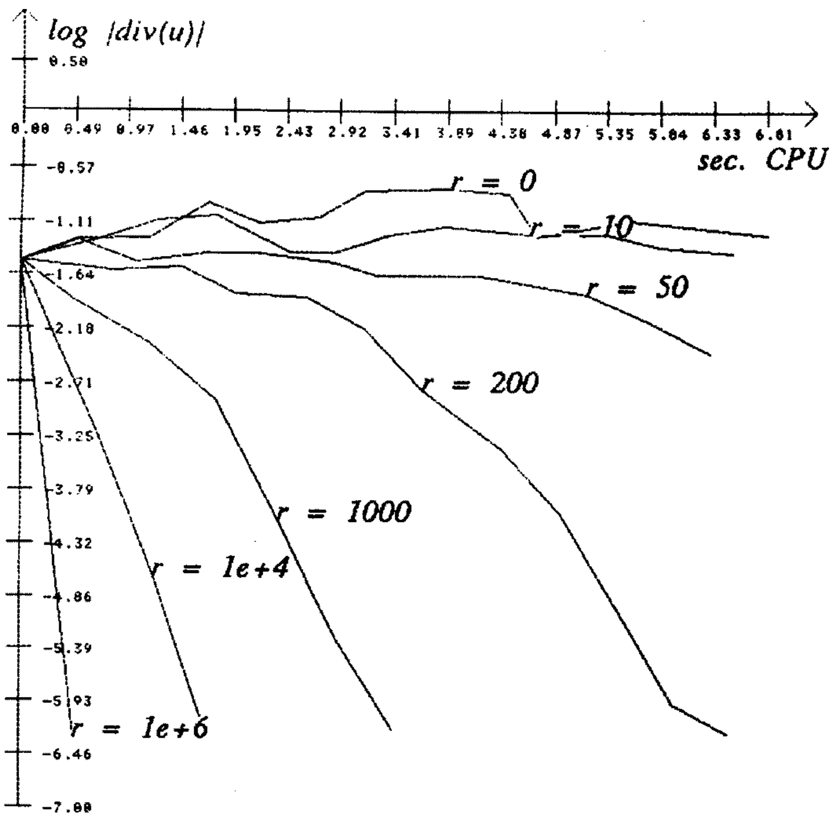


Figure 5.10: Influence du préconditionnement sur la convergence de l'algorithme du G.C.

du préconditionnement. Ce préconditionnement correspond à la méthode du Lagrangien Augmenté, paramétré par r. Pour $r = 0$, nous retrouvons l'algorithme du gradient conjugé classique.

Cette accélération de la convergence est d'autant plus évidente que la résolution du problème est ici initialement mal conditionnée. Ainsi, sans utiliser de préconditionnement ($r = 0$), le résidu sur la contrainte d'incompressibilité est de l'ordre de 10^{-1} au bout de 50 itérations (25 secondes CPU). Un préconditionnement avec $r = 10^6$ conduit à un résidu de l'ordre de 10^{-11} au bout de 2 itérations (0.5 secondes CPU).

5.3.2 Coût de l'algorithme

Considérons à présent l'algorithme du gradient conjugé préconditionné avec $r = 10^6$. Nous recherchons le coût en temps de calcul $\mathcal{T}(N)$ de l'algorithme en fonction de la taille des données N sous la forme :

$$\mathcal{T}(N) = c.N^k \log N \tag{5.50}$$

Une estimation de k est donée par :

$$k_i = \frac{\log\left(\frac{\mathcal{T}_i \log N_{i-1}}{\mathcal{T}_{i-1} \log N_i}\right)}{\log\left(\frac{N_i}{N_{i-1}}\right)} \quad 1 \leq i \leq I = 3$$

De même, en recherchant $\mathcal{W}$ sous la forme :

$$\mathcal{W}(N) = c.N^{k'} \tag{5.51}$$

une estimation de $k_i^{'}$ est :

$$k_i^{'} = \frac{\log\left(\mathcal{W}_i/\mathcal{W}_{i-1}\right)}{\log\left(N_i/N_{i-1}\right)} \quad 1 \leq i \leq I = 3$$

Les résultats sont rassemblés dans le tableau 5.2. Nous constatons l'efficacité

i	N_i	$\mathcal{T}_i$	k_i	$\mathcal{W}_i$	$k_i^{'}$
0	945	0.50		28	
1	1909	1.45	1.37	63	1.05
2	3615	3.40	1.20	137	1.08
3	7403	9.74	1.46	333	1.12

Table 5.2: Coût de l'algorithme du G.C.P.

de la méthode : les coûts en temps de calcul doublent approximativement lorque double le nombre d'inconnues. D'autre part, l'occupation en place mémoire apparaît ici linéaire.

5.3.3 Convergence de la méthode d'éléments finis

Commençons par rechercher la longeur s_1 du domaine amont correspondant à un écoulement de Poiseuille établi à l'amont. La table 5.3 met en évidence la nécessité d'une section aval assez allongée. Ces grandeurs ont été calculées à l'aide du maillage 1.2, avec une section aval $s_2 = 8$ et en géométrie axisymétrique. Pour une géométrie plane, les résultats sont similaires, et la

s_1	L	ψ
25	0.49	1.87×10^{-3}
50	0.55	1.87×10^{-3}
75	0.58	1.88×10^{-3}
100	0.59	1.875×10^{-3}

Table 5.3: Influence de la longeur du domaine amont sur la perte de pression et l'activité des recirculations

longeur à l'amont $s_1 = 75$ est apparue suffisante.

Nous considérons à présent les deux suites de solutions approchées associées aux deux familles de triangulation et à la méthode d'éléments finis définis au paragraphe 5.2.5. Nous allons comparer les taux de convergence de ces deux suites.

La table 5.4 donne l'activité de la recirculation, notée ψ_i, et associée à une triangulation $\mathcal{T}_i$, d'une famille fixée.

Les résultats donnés pour les deux familles de triangulations sont présentés dans la table 5.4, pour une géométrie axisymétrique. Nous constatons que la deuxième famille de triangulation possède des propriétés de convergence bien plus importantes que la première. Ceci est certainement dû au fait que la solution est singulière au voisinage de la singularité : la première famille approche mal cette singularité, et les solutions approchées associées sont peu précises dans la région d'entrée. Par contre, la concentration d'éléments de petites tailles développés par la deuxième famille permet d'augmenter les propriétés de convergence dans la zone d'entrée : l'erreur est divisée approximativement par deux lorsqu'on double le nombre d'éléments du maillage.

La figure 5.11 représente les lignes de courant de la solution approchée obtenue sur la triangulation $\mathcal{T}_3$ (famille 2), pour une géométrie plane (a) et axisymétrique (b).

première famille de triangulation :

i	ψ_i	$\vert\psi_i - \psi_I\vert \times 10^3$
0	1.95×10^{-3}	0.88
1	1.77×10^{-3}	0.70
2	1.24×10^{-3}	0.17
3	1.07×10^{-3}	

deuxième famille de triangulation :

i	ψ_i	$\vert\psi_i - \psi_I\vert \times 10^3$
0	1.951×10^{-3}	0.076
1	1.835×10^{-3}	0.040
2	1.896×10^{-3}	0.021
3	1.875×10^{-3}	

Table 5.4: Convergence pour les deux familles de triangulations

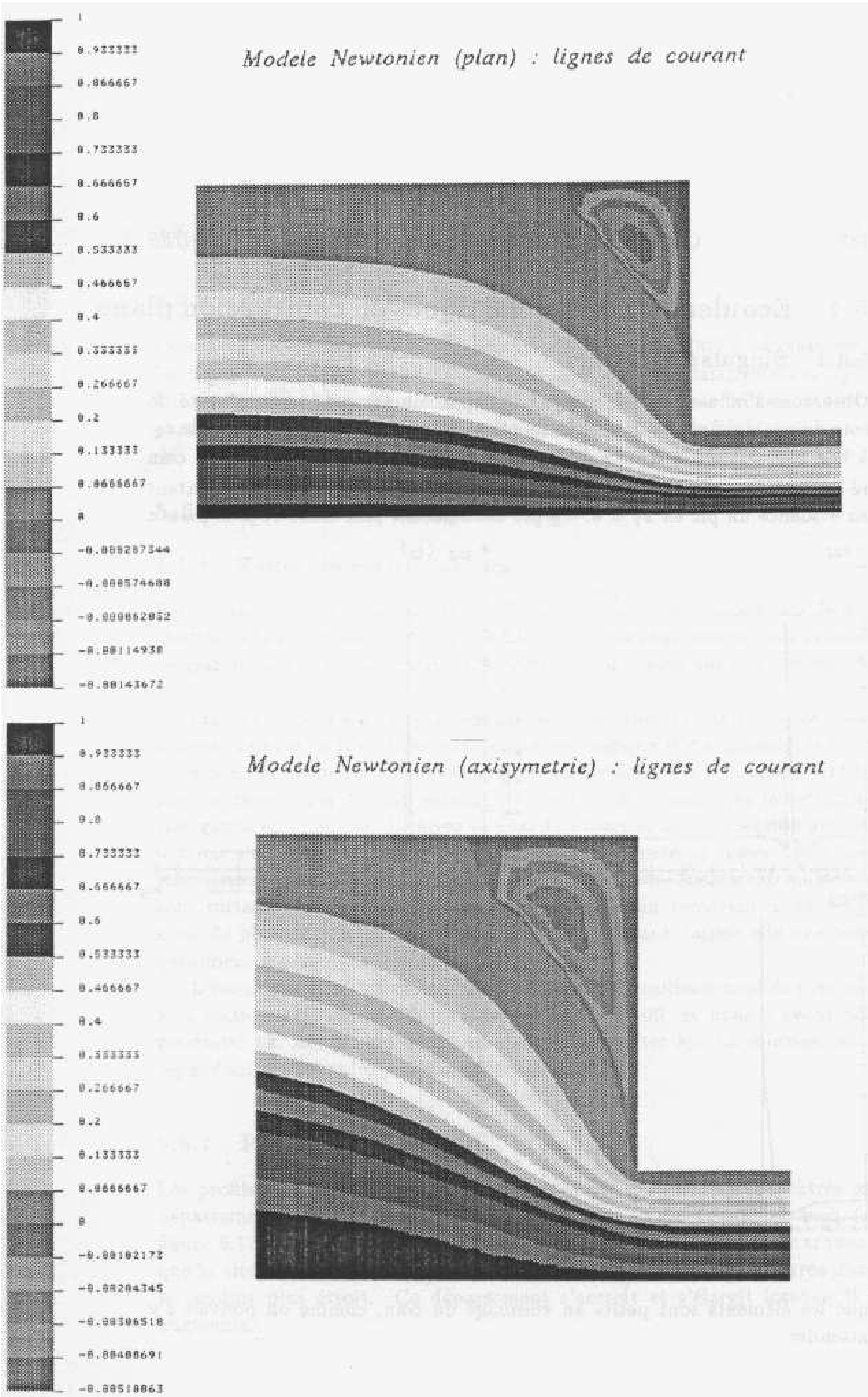


Figure 5.11: Lignes de courant pour un fluide newtonnien en contraction plane (a) et axisymétrique (b)

5.4 Ecoulements viscoélastiques en contraction plane

5.4.1 Singularité de coin

Observons l'influence du maillage sur l'approximation de la singularité de coin de τ_{22}. La figure 5.12 reprśente une section de la contrainte normale τ_{22} à $We = 1$ suivant la ligne horizontale $x_2 = r_2 - h_i^{(2)}$, passant près du coin ré-entrant, et pour les triangulations $\mathcal{T}_i^{(2)}$, $i = 0, \ldots, 3$. Ces sections en évidence un pic en $x_1 = 0.$ à $We = 4$, Ce est d'autant plus élevé et plus

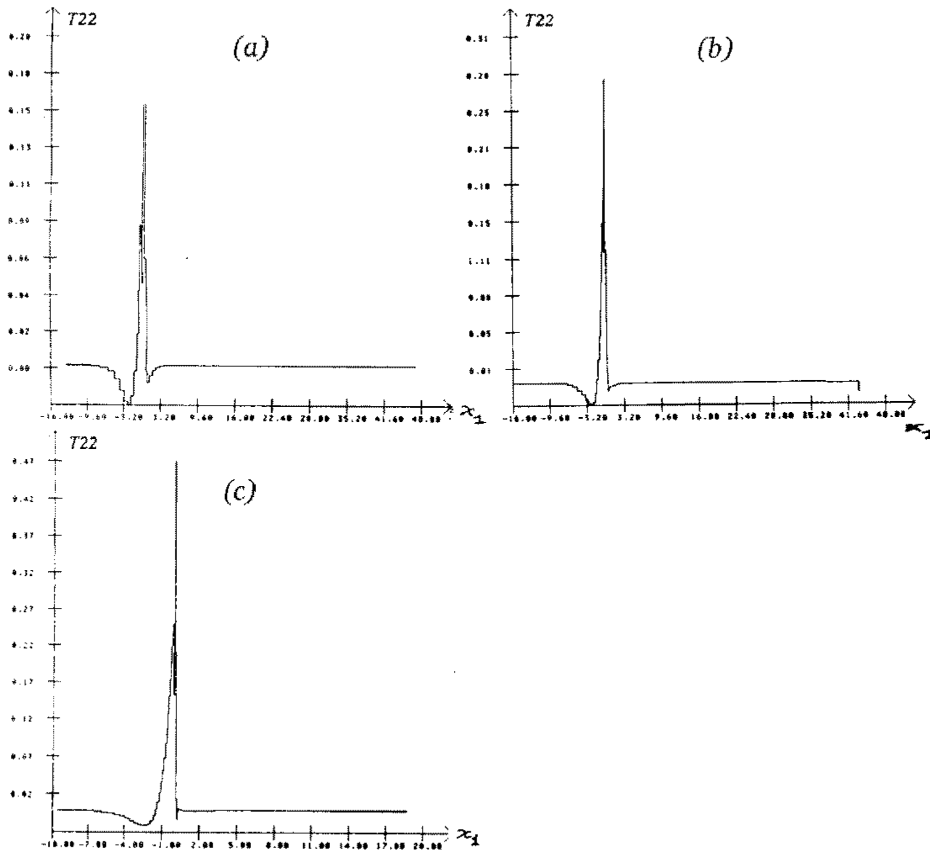


Figure 5.12: Approximation de la singularité de coin

pointu que les éléments sont petits au voisinage du coin, comme on pouvait s'y attendre.

5.4.2 Zones de recirculations

Commentons à présent les recirculations obtenues. La figure 5.13 représente les lignes de courant obtenues à $We = 1$ et4 sur la triangulation $\mathcal{T}_2^{(2)}$. L'intensité et la longeur des zones de recirculation reste faible. D'autre part, le foyer d'activité a tendance à s'étendre vers le coin réentrant.

5.5 Géométrie axisymetrique

5.5.1 Zones de recirculations

Par opposition aux écoulements en géométrie plane, les écoulements de fluides viscoélastiques dans contraction axisymétrique sont connus pour présenter de grandes zones de recirculation, et différent fortement des écoulements de fluides newtoniens.

Les figure 5.14 à 5.16 représentent de telles zones de recirculation pour diverses valeurs de We. Nous constatons que, lorsque We augmente, le foyer de la zone de recirculation s'éloigne du coin sortant ($We = 0$, figure 5.11.b) pour se rapprocher du coin entrant ($We = 4$ à 20), tandisque la recirculation gagne en intensité. Le foyer va ensuite s'avancer dans la section amont, tout en en s'élargissant ($We = 20$ à 50), mais l'intensité va rester désormais sensiblement constante. Enfin, la zone d'activité va finalement quitter le coin sortant et s'étendra ($We = 60$ à 90) du coin réentrant à un point situé de plus en plus loin sur la parois amont, laissant derière elle une zone quasiment morte dans le coin sortant.

L'écoulement de Poiseuille n'est plus établi avec suffisament de précision à la section aval ($s_2 = 200$) au delà de $We = 100$, et nous n'avons pas poursuivi les calculs, qui nécéssiteraient d'augmenter s_2. La solution reste cependant stationnaire jusqu'à cette valeur.

5.5.2 Profils de vitesse sur l'axe

Les profils de vitesse le long de l'axe présentent dans la région d'entrée un dépassement de la valeur obtenue lorsque l'écoulement est établi en aval. La figure 5.17 représente ces variations pour $We = 1$ et 4. Nous remarquons que la vitesse présente un dépassement au niveau de la section d'entrée dans le conduit plus étroit. Ce dépassement s'accroît et s'élargit lorsque We augmente.

5.5.3 Perte de pression

Les calculs des pertes de pression nous ont conduit à utiliser un domaine long à l'aval ($s_2 = 150$), afin de s'assurer que l'écoulement de Poiseuille est établi à l'aval. Cependant, la perte de pression L_h due au passage dans la région d'entrée correspond à la différence de deux grandes quantités, et, pour les triangulations utilisées ici, cette quantité ne semble pas calculée avec assez de précision pour pouvoir conclure à une convergence lorsque s_2 devient grand (on trouvera une remarque semblable dans [14]).

D'autre part, nous avons constaté que L_h varie de façon importante avec la taille des éléments situés au voisinage du coin réentrant. Or les grandeurs τ et $\vec{u}$ intervenant dans l'expression (5.40) sont singulières au coin réentrant, ainsi que leurs dérivées. Une autre raison serait l'incapacité des méthodes d'éléments finis classiques à approcher correctement les singularité. La connaissance de la nature de de la singularité de coin nous fait actuellement défaut, et les méthodes d'éléments finis adaptés aux singularités [40] sont par conséquent inapplicables pour l'instant.

L'utilistion de méthodes d'élements finis auto-adaptatives [43, 1] permettrait de contourner cette difficulté. Cependant, ces méthodes nécéssitent une estimation *a posteriori* des erreurs sur chaque élément, et dont la somme majore l'erreur globale. Nous ne connaissons pas actuellement de telles estimations pour le problème des écoulements visco-élastiques.

5.6 Temps de calcul

La table 5.5 représente le temps CPU nécéssaire ̀l'obtention d'une solution à $We = 4$ suivant la triangulation utilisée (machine `apollo DN 4000`).

triangulation	N	temps CPU
0	945	55s
1	1909	3mn 28s
2	3615	10mn 16s
3	7403	28mn 32s

Table 5.5: Temps de calcul sur `apollo DN 4000` pour $We = 4$.

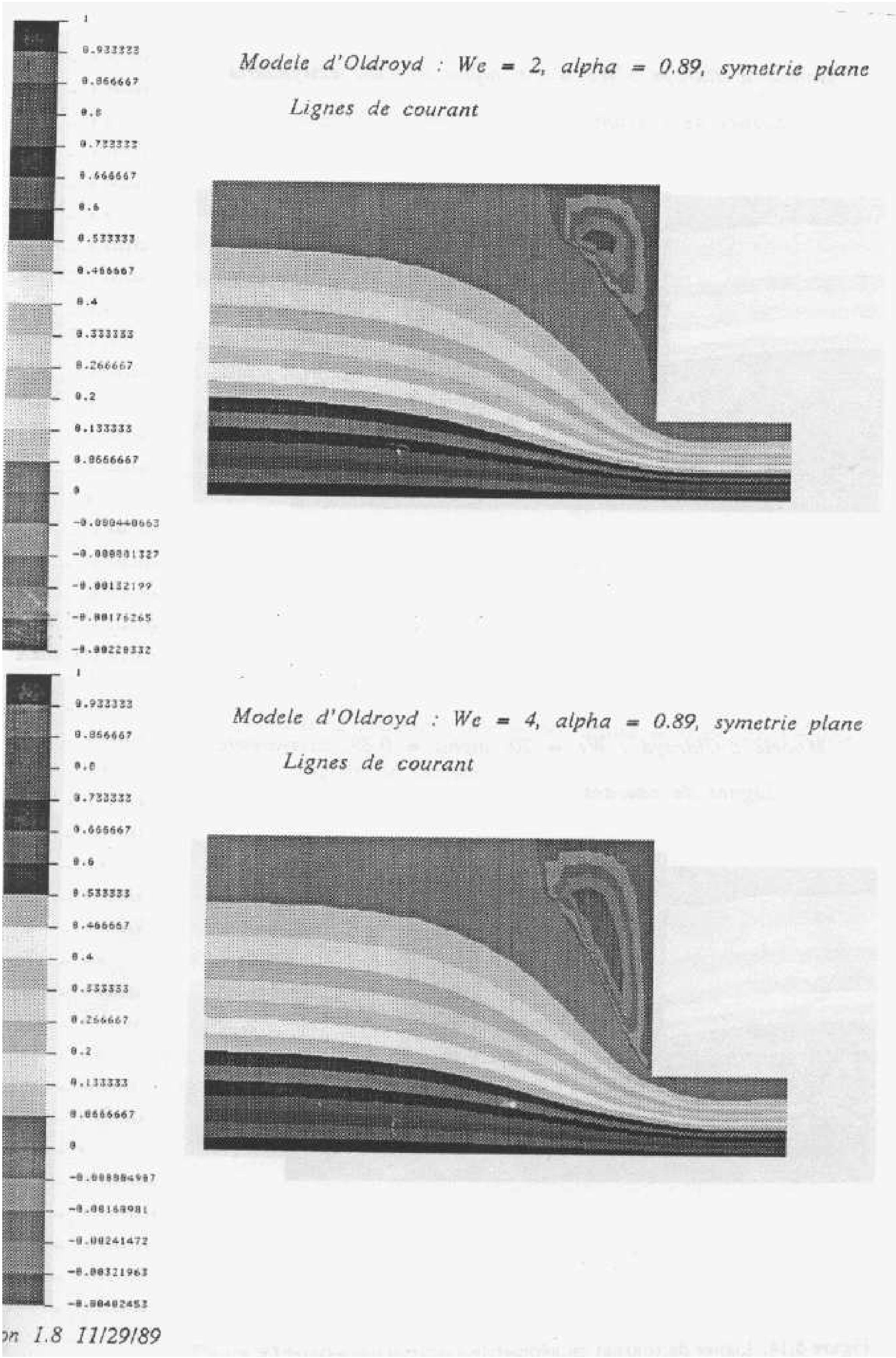


Figure 5.13: Lignes de courant en géométrie plane

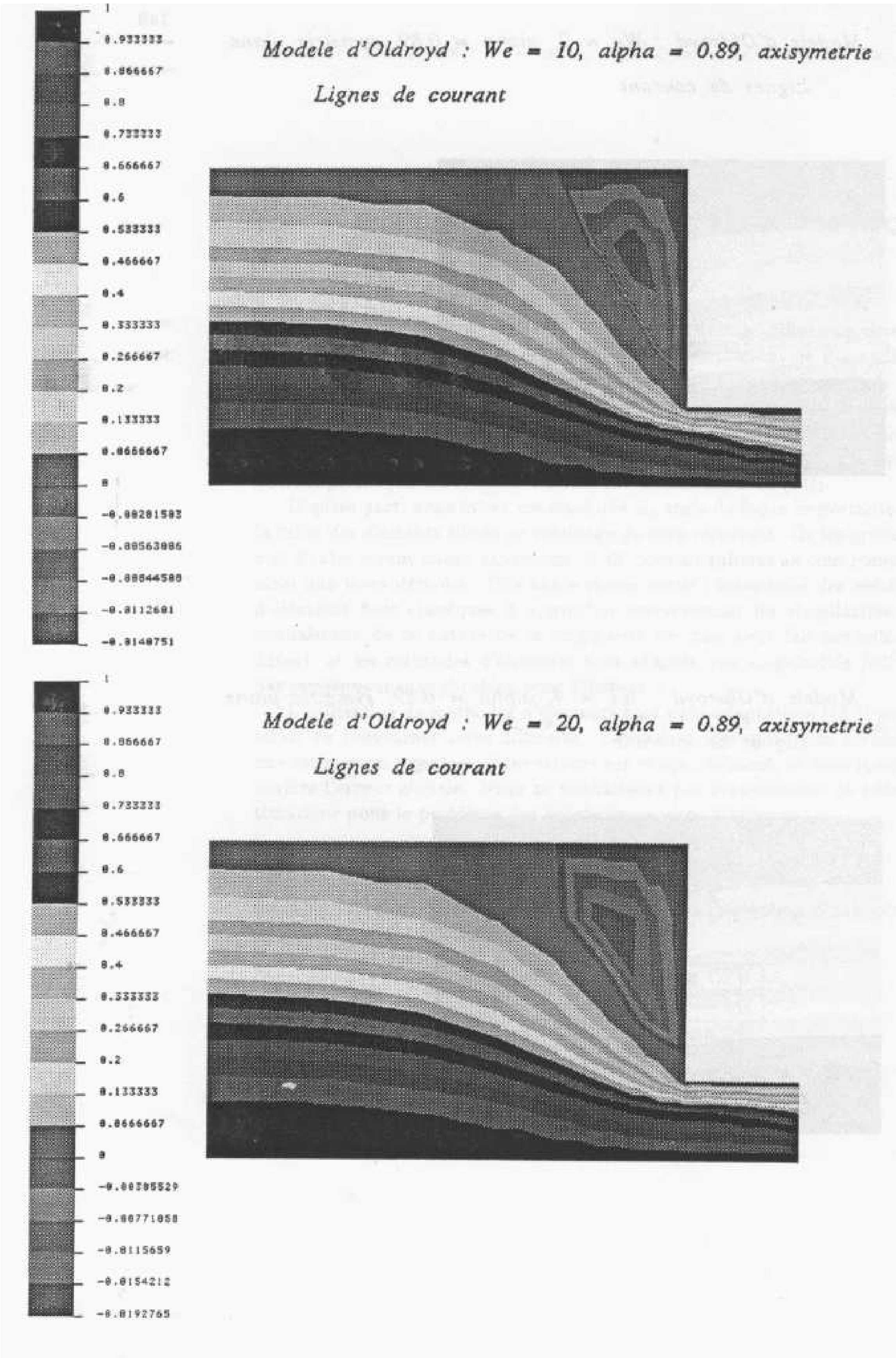


Figure 5.14: Lignes de courant en géométrie axisymétrique (part I)

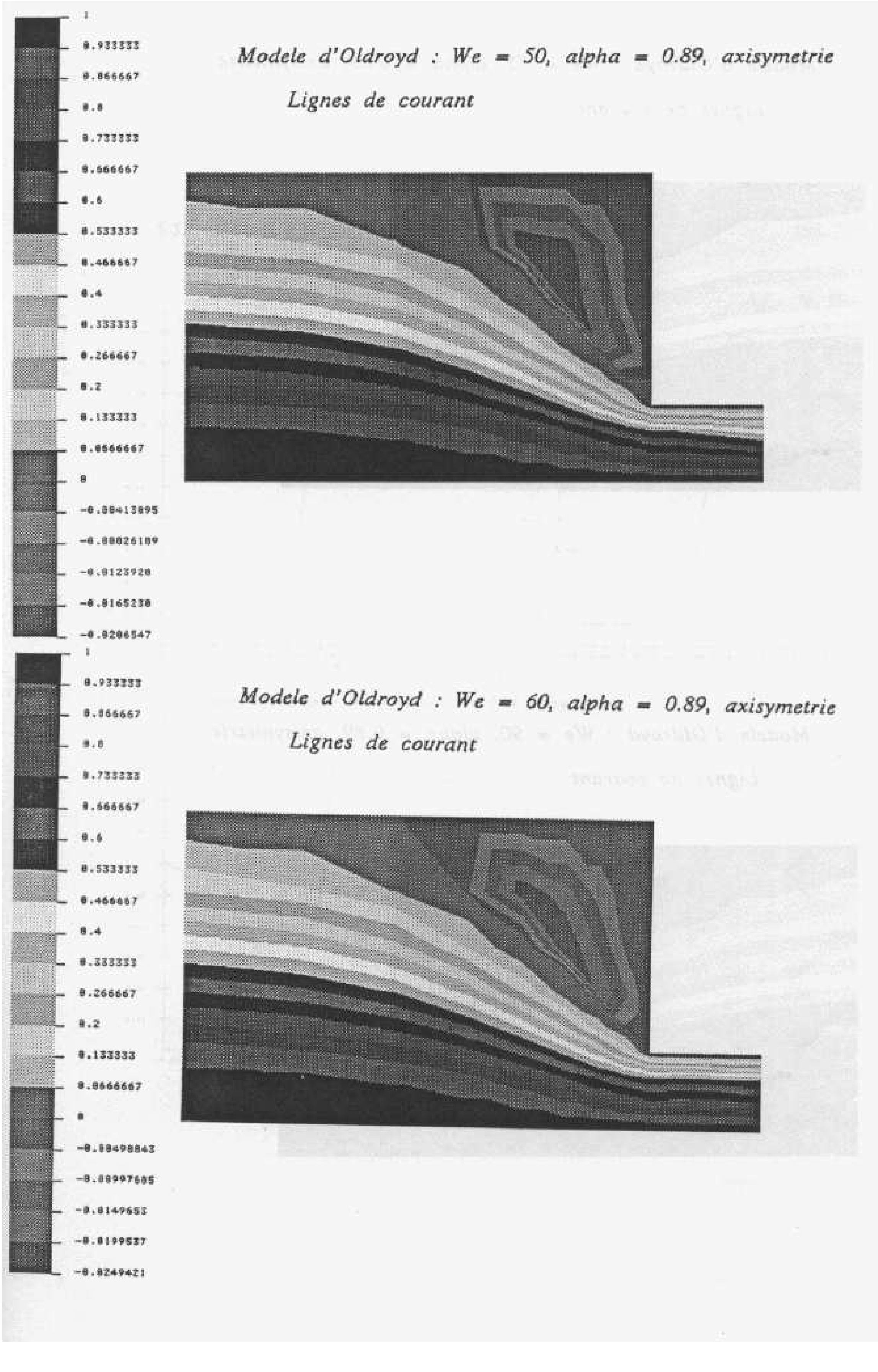


Figure 5.15: Lignes de courant en géométrie axisymétrique (Part II)

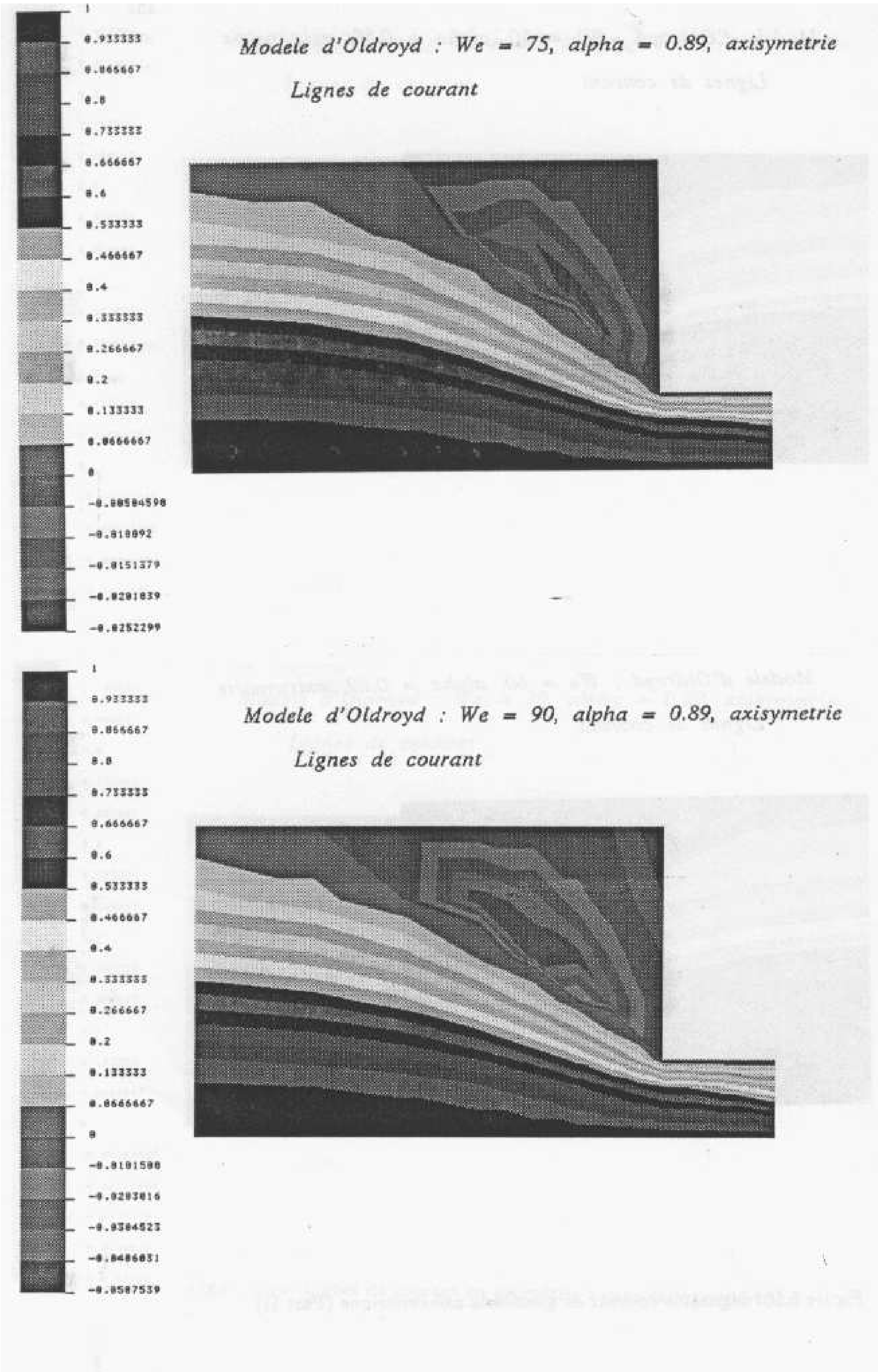


Figure 5.16: Lignes de courant en géométrie axisymétrique (part III)

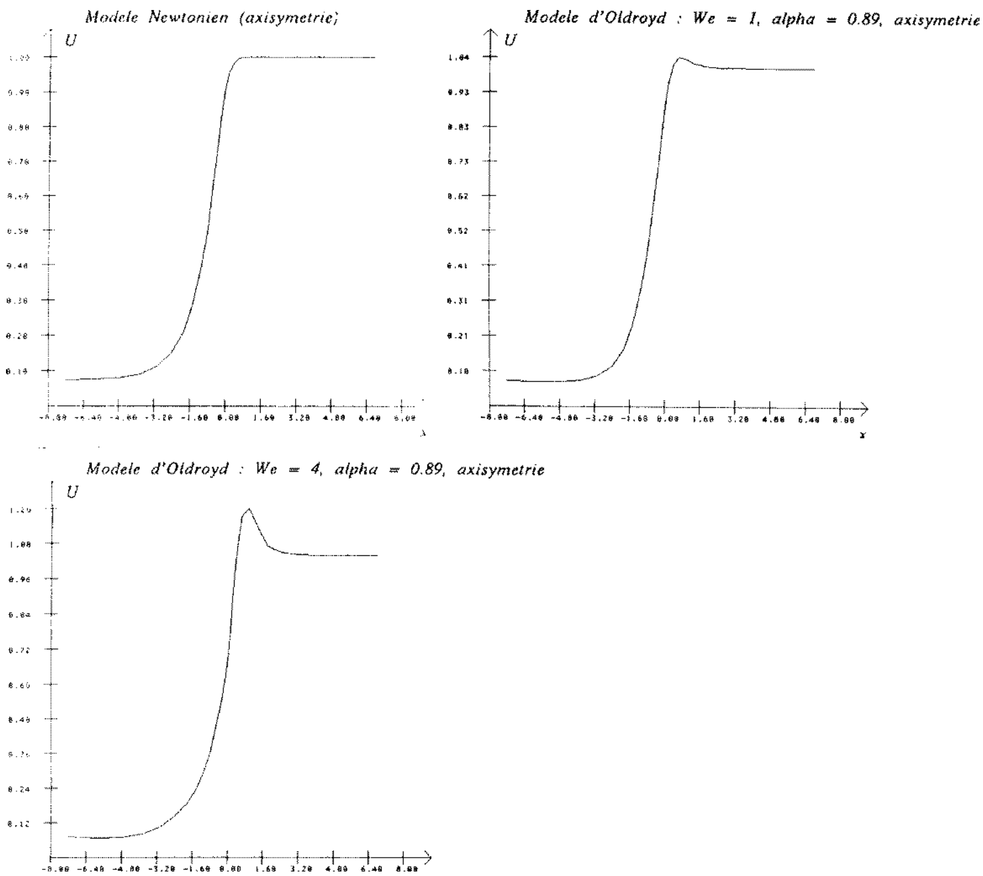


Figure 5.17: Profils de vitesse sur l'axe

Appendix A

Problèmes variationnels abstraits

L'objet de cette annexe est de faciliter la lecture du chapitre 2. Nous présentons ici les problèmes variationnels dans un cadre abstrait ainsi que les énoncés des principaux résultats associés, fréquemment cités lors de la construction de la méthode d'éléments finis.

Pour les démonstrations de ces résultats, le lecteur pourra consulter [7, 54, 21].

A.1 Enoncé et résolution

A.1.1 Enoncé du problème abstrait

Soient $\mathbf{X}$ et $\mathbf{M}$ deux espaces de Hilberts, de normes respectives $\|.\|_{\mathbf{X}}$ et $\|.\|_{\mathbf{M}}$, et $\mathbf{X}^{'}$, $\mathbf{M}^{'}$ les espaces duaux. Lorsqu'il n'y aura pas d'ambiguïté, on notera indifférement $< .,. >$ la forme bilinéaire de dualité entre les espaces $\mathbf{X}^{'}$ et $\mathbf{X}$, ou $\mathbf{M}^{'}$ et $\mathbf{M}$.

On introduit deux formes bilinéaires continues :

$$a : \mathbf{X} \times \mathbf{X} \longmapsto \mathrm{I\!R} \tag{A.1}$$

$$b : \mathbf{X} \times \mathbf{M} \longmapsto \mathrm{I\!R} \tag{A.2}$$

de normes :

$$\|a\| = \sup_{\substack{u,v \in \mathbf{X} \\ u \neq 0, v \neq 0}} \frac{a(u,v)}{\|u\|_{\mathbf{X}} \|v\|_{\mathbf{X}}} \tag{A.3}$$

$$\|b\| = \sup_{\substack{u\in\mathbf{X},\mu\in\mathbf{M}\\ u\neq 0,\mu\neq 0}} \frac{b(u,\mu)}{\|u\|_{\mathbf{X}}\|\mu\|_{\mathbf{M}}} \tag{A.4}$$

et le problème variationnel :

$$(Q) : \left\{ \begin{array}{lll} \multicolumn{3}{l}{\textit{pour tout } l \in \mathbf{X}', \chi \in \mathbf{M}', \textit{ trouver } (u,\lambda) \in \mathbf{X}\times\mathbf{M} \textit{ tel que}} \\ a(u,v) & +b(v,\lambda) & =< l, v >, \ \forall v \in \mathbf{X} \\ b(u,\mu) & & =< \chi, \mu >, \ \forall \mu \in \mathbf{M} \end{array} \right.$$

A.1.2 Définition des opérateurs et des espaces

Afin d'étudier le problème (Q), nous aurons besoin de notations supplémentaires. Associons aux formes a et b deux opérateurs linéaires continus $A \in \mathcal{L}(\mathbf{X};\mathbf{X}')$ et $B \in \mathcal{L}(\mathbf{X};\mathbf{M}')$, définis par :

$$< Au, v > \ = a(u,v), \quad \forall u, v \in \mathbf{X} \tag{A.5}$$

$$< Bv, \mu > \ = b(v,\mu), \quad \forall v \in \mathbf{X}, \ \forall \mu \in \mathbf{M} \tag{A.6}$$

Soit $B^* \in \mathcal{L}(\mathbf{M};\mathbf{X}')$ l'opérateur adjoint de B, c'est-à-dire :

$$< B^*\mu, v > = < Bv, \mu > = b(v,\mu), \ \forall \mu \in \mathbf{M}, \ \forall v \in \mathbf{X} \tag{A.7}$$

et nous pouvons vérifier que

$$\|A\|_{\mathcal{L}(\mathbf{X};\mathbf{X}')} = \|a\| \tag{A.8}$$

$$\|B\|_{\mathcal{L}(\mathbf{X};\mathbf{M}')} = \|b\| \tag{A.9}$$

Avec ces opérateurs, nous obtenons une formulation équivalente de (Q) :

$$(Q') : \left\{ \begin{array}{lll} \multicolumn{3}{l}{\textit{trouver } (u,\lambda) \in \mathbf{X}\times\mathbf{M} \textit{ tel que}} \\ Au & +B^*\lambda & = l, \ \text{dans } \mathbf{X}' \\ Bu & & = \chi, \ \text{dans } \mathbf{M}' \end{array} \right.$$

Le noyau de B, et les espaces translatés associés sont définis par :

$$\mathbf{V} = \mathrm{Ker} B \tag{A.10}$$

$$\mathbf{V}(\chi) = \{v \in \mathbf{X}; \ Bv = \chi\}, \ \forall \chi \in \mathbf{M}' \tag{A.11}$$

Remarquons qu'il est possiblet d'écrire, de façon équivalente :

$$\mathbf{V}(\chi) = \{v \in \mathbf{X}; \ b(v,\mu) = < \chi, \mu >, \ \forall \mu \in \mathbf{M}\} \tag{A.12}$$

$$\mathbf{V} = \mathbf{V}(0) \tag{A.13}$$

De plus, la continuité de B entraîne que $\mathbf{V}$ est un sous-espace fermé de $\mathbf{X}$.

A.1.3 Formulation sur le noyau

Maintenant, nous pouvons associer à (Q) le problème :

$$(P) : \begin{cases} \textit{trouver } u \in \mathbf{V}(\chi) \textit{ tel que} \\ a(u,v) =< l, v >, \ \forall v \in \mathbf{V} \end{cases}$$

Il est clair que, si $(u, \lambda) \in \mathbf{X} \times \mathbf{M}$ est une solution de (Q), alors $u \in \mathbf{V}(\chi)$ et u est une solution de (P). Le reste de ce paragraphe est consacré à établir la réciproque, ainsi que l'existence et l'unicité de la solution, sous des hypothèses qui vont suivre.

A.1.4 Résolution du problème

Pour cela, définissons l'ensemble polaire de $\mathbf{V}$:

$$\mathbf{V}^o = \{g \in \mathbf{X}^{'}; \ < g, v >= 0, \ \forall v \in \mathbf{V}\} \subset \mathbf{X}^{'} \tag{A.14}$$

Commençons par le

Lemme A.1 *de la condition inf-sup*
Sont équivalentes :

i) il existe une constante $\beta > 0$ telle que :

$$\inf_{\mu \in \mathbf{M}} \sup_{v \in \mathbf{X}} \frac{b(v,\mu)}{\|v\|_{\mathbf{X}} \|\mu\|_{\mathbf{M}}} \geq \beta \tag{A.15}$$

ii) l'opérateur B^* est un isomorphisme de $\mathbf{M}$ dans $\mathbf{V}^o$, et

$$\|B^* \mu\|_{\mathbf{X}'} \geq \beta \, \|\mu\|_{\mathbf{M}}, \ \forall \mu \in \mathbf{M} \tag{A.16}$$

iii) l'opérateur B est un isomorphisme de $\mathbf{V}^\perp$ dans $\mathbf{M}^{'}$, et

$$\|Bv\|_{\mathbf{M}'} \geq \beta \, \|v\|_{\mathbf{X}}, \ \forall v \in \mathbf{X} \tag{A.17}$$

Ceci nous permet d'énoncer un important résultat dû à Brezzi [7] :

Théorème A.1 *Existence et unicité de la solution de (Q)*
On fait les hypothèses suivantes :

i) la forme bilinéaire a est elliptique sur le noyau $\mathbf{V}$:

$$\exists \alpha > 0 \textit{ tel que } a(v,v) \geq \alpha \, \|v\|_{\mathbf{X}}, \ \forall v \in \mathbf{V} \tag{A.18}$$

ii) la forme bilinéaire b satisfait la condition inf-sup (A.15).

Alors (P) admet une unique solution $u \in \mathbf{V}(\chi)$, et il existe un unique $\lambda \in \mathbf{M}$ tel que (u, λ) soit l'unique solution de (Q).

De plus, l'application

$$(l, \chi) \longmapsto (u, \lambda)$$

est un isomorphisme de $\mathbf{X}' \times \mathbf{M}'$ dans $\mathbf{X} \times \mathbf{M}$.

Remarque A.1 La condition inf-sup est nécessaire et suffisante

Si a est $\mathbf{V}$-elliptique, alors l'application

$$(l, \chi) \longmapsto (u, \lambda)$$

est un isomorphisme de $\mathbf{X}' \times \mathbf{M}'$ dans $\mathbf{X} \times \mathbf{M}$ si et seulement si la condition inf-sup est vérifiée.

En effet, on sait que cette condition est nécessaire, et il reste à montrer qu'elle est suffisante.

Soit $\chi \in \mathbf{X}'$ et (u, λ) une solution de (Q) avec, pour second membre $(0, \chi)$. Alors $Bu = \chi$, et ainsi $B(\mathbf{X}) = \mathbf{M}'$. Cela entraîne que B est un isomorphisme de $\mathbf{V}^\perp$ dans $\mathbf{M}'$.

En vertu du lemme A.1, la condition inf-sup (A.15) est satisfaite. □

A.2 Formulation mixte d'un problème abstrait

Dans ce paragraphe, on reprend à nouveau le problème abstrait, mais en se plaçant dans un cadre plus *faible*, conduisant à une formulation mixte, dont on en déduira une approximation.

A.2.1 Extension des formes et des espaces

Nous reprenons ici les problèmes (P) et (Q) introduits précédemment, en supposant que les hypothèses (i) et (ii) du théorème A.1 sont satisfaites.

Afin de donner une formulation plus faible de (Q), nous introduisons tout d'abord deux espaces de Hilbert $\tilde{\mathbf{X}}$ et $\tilde{\mathbf{M}}$, normés respectivement par $\|.\|_{\tilde{\mathbf{X}}}$ et $\|.\|_{\tilde{\mathbf{M}}}$, et tels que les injections de $\mathbf{X}$ dans $\tilde{\mathbf{X}}$ et de $\tilde{\mathbf{M}}$ dans $\mathbf{M}$ soient continues et denses, ce qu'on note :

$$\mathbf{X} \overset{d}{\subset} \tilde{\mathbf{X}} \tag{A.19}$$

$$\tilde{\mathbf{M}} \overset{d}{\subset} \mathbf{M} \tag{A.20}$$

Ensuite, considérons les deux formes bilinéaires continues :

$$\tilde{a} : \tilde{\mathbf{X}} \times \tilde{\mathbf{X}} \longmapsto \mathbb{R} \tag{A.21}$$

$$\tilde{b} : \tilde{\mathbf{X}} \times \tilde{\mathbf{M}} \longmapsto \mathbb{R} \tag{A.22}$$

et posons :

$$\|\tilde{a}\| = \sup_{\substack{u,v\in\mathbf{X} \\ u\neq 0, v\neq 0}} \frac{\tilde{a}(u,v)}{\|u\|_{\tilde{\mathbf{X}}}\|v\|_{\tilde{\mathbf{X}}}} \tag{A.23}$$

$$\|\tilde{b}\| = \sup_{\substack{u\in\mathbf{X},\mu\in\mathbf{M} \\ u\neq 0, \mu\neq 0}} \frac{\tilde{b}(u,\mu)}{\|v\|_{\tilde{\mathbf{X}}}\|\mu\|_{\tilde{\mathbf{X}}}} \tag{A.24}$$

Ces deux formes sont des extensions de a et b au sens où :

$$\tilde{a}(u,v) = a(u,v), \quad \forall u,v \in \mathbf{X} \tag{A.25}$$

$$\tilde{b}(v,\tilde{\mu}) = b(v,\tilde{\mu}), \quad \forall v \in \mathbf{X},\ \forall \tilde{\mu} \in \tilde{\mathbf{M}} \tag{A.26}$$

Nous supposerons de plus que $l \in \tilde{\mathbf{X}}'$ et $\chi \in \tilde{\mathbf{M}}'$.

A.2.2 Formulation variationnelle mixte

Considèrons le problème :

$$(\tilde{Q}) : \begin{cases} \textit{trouver } (\tilde{u},\tilde{\lambda}) \in \tilde{\mathbf{X}} \times \tilde{\mathbf{M}} \textit{ tel que} \\ \tilde{a}(\tilde{u},v) \quad +b(v,\tilde{\lambda}) \quad =< l,v >,\ \forall v \in \tilde{\mathbf{X}} \\ \tilde{b}(\tilde{u},\mu) \qquad\qquad\quad =< \chi,\mu >,\ \forall \mu \in \tilde{\mathbf{M}} \end{cases}$$

Pour tout $\chi \in \mathbf{M}'$, définissons la variété :

$$\tilde{\mathbf{V}}(\chi) = \{v \in \tilde{\mathbf{X}};\ \tilde{b}(v,\mu) =< \chi,\mu >,\ \forall \mu \in \tilde{\mathbf{M}}\} \tag{A.27}$$

et le sous-espace fermé de $\tilde{\mathbf{X}}$:

$$\tilde{\mathbf{V}} = \tilde{\mathbf{V}}(0) \tag{A.28}$$

Puisque $\tilde{b}$ est une extension (A.26) de b sur $\tilde{\mathbf{X}} \times \tilde{\mathbf{M}}$, on a :

$$\mathbf{V} \subset \tilde{\mathbf{V}} \tag{A.29}$$

$$\mathbf{V}(\chi) \subset \tilde{\mathbf{V}}(\chi) \tag{A.30}$$

Au problème $(\tilde{Q})$, nous associons alors le problème :

$$(\tilde{P}) : \begin{cases} \textit{trouver } \tilde{u} \in \tilde{\mathbf{V}}(\chi) \textit{ tel que} \\ \tilde{a}(\tilde{u},v) =< l,v >,\ \forall v \in \tilde{\mathbf{V}} \end{cases}$$

A.2.3 Existence et unicité de la solution faible

Afin de pouvoir analyser les problèmes $(\tilde{P})$ et $(\tilde{Q})$, nous supposertons $\tilde{a}$ $\tilde{\mathbf{V}}$-elliptique :

$$\exists \tilde{\alpha} > 0 \text{ tel que } \tilde{a}(v,v) \geq \tilde{\alpha}\, \|v\|^2_{\tilde{\mathbf{X}}},\ \forall v \in \tilde{\mathbf{V}} \tag{A.31}$$

D'autre part, il n'y a plus besoin de condition inf-sup sur $\tilde{b}$, mais de (A.26) et (A.31) nous déduisons que $\tilde{b}$ vérifie :

$$\sup_{v\in\tilde{\mathbf{X}}} \frac{\tilde{b}(v,\mu)}{\|v\|_{\tilde{\mathbf{X}}}} \geq \frac{1}{C} \sup_{v\in\mathbf{X}} \frac{\tilde{b}(v,\mu)}{\|v\|_{\mathbf{X}}} \geq \frac{\beta}{C}\|\mu\|_{\mathbf{M}},\ \forall \mu \in \tilde{\mathbf{M}} \tag{A.32}$$

où C est la constante de l'injection continue de $\mathbf{X}$ dans $\tilde{\mathbf{X}}$.

Cette propriété inf-sup est plus faible que la condition inf-sup ordinaire, car elle fait intervenir $\|.\|_{\mathbf{M}}$, au lieu de $\|.\|_{\tilde{\mathbf{M}}}$. Le théorème A.1 de Brezzi-Babuska ne suffit donc pas à montrer que $(\tilde{Q})$ est bien-posé. Le résultat suivant contourne cette difficulté :

Théorème A.2 *Existence et unicité de la solution de* $(\tilde{Q})$
Soit (u, λ) *la solution de* (Q).

- *Si* $\tilde{a}$ *est* $\tilde{\mathbf{V}}$*-elliptique, alors le problème* $(\tilde{P})$ *admet exactement une solution :* $\tilde{u} \in \tilde{\mathbf{V}}(\chi)$.

  De plus, si $\tilde{u} \in \mathbf{V}(\chi)$ *ou si* $\mathbf{V}$ *est dense dans* $\tilde{\mathbf{V}}$*, alors* $\tilde{u} = u$.

- *Si, de plus* $\lambda \in \tilde{\mathbf{M}}$*, alors* (u, λ) *est l'unique solution de* $(\tilde{Q})$.

Remarque A.2
Au chapitre 2, nous étudions deux applications où l'approximation de $(\tilde{P})$ et $(\tilde{Q})$ est plus simple que celle de (P) et (Q). □

A.3 Approximation mixte abstraite

Dans ce paragraphe, nous supposerons toutes les hypothèses du théorème A.2 satisfaites.

A.3.1 Approximation variationnelle mixte

Pour tout $h > 0$, soient $\mathbf{X}_h$ et $\mathbf{M}_h$ deux espaces de dimension finie vérifiant :

$$\mathbf{X}_h \subset \tilde{\mathbf{X}} \tag{A.33}$$

$$\mathbf{M}_h \subset \tilde{\mathbf{M}} \tag{A.34}$$

On approche le problème $(\tilde{Q})$ par

$$(Q_h) : \begin{cases} \textit{trouver } (u_h, \lambda_h) \in \mathbf{X}_h \times \mathbf{M}_h \textit{ tel que} \\ \tilde{a}(u_h, v_h) \quad +b(v_h, \lambda_h) \quad =< l, v_h >, \ \forall v_h \in \mathbf{X}_h \\ \tilde{b}(u_h, \mu_h) \qquad\qquad\qquad =< \chi, \mu_h >, \ \forall \mu_h \in \mathbf{M}_h \end{cases}$$

A nouveau, on définit :

$$\mathbf{V}_h(\chi) = \{v_h \in \mathbf{X}_h;\ \tilde{b}(v_h, \mu_h) =< \chi, \mu_h >,\ \forall \mu_h \in \mathbf{M}_h\} \quad (A.35)$$

$$\mathbf{V}_h = \mathbf{V}_h(0) \quad (A.36)$$

Puis, à (Q_h), on associe le problème :

$$(P_h) : \begin{cases} \textit{trouver } u_h \in \mathbf{V}_h(\chi) \textit{ tel que} \\ \tilde{a}(u_h, v_h) =< l, v_h >, \ \forall v_h \in \mathbf{V}_h \end{cases}$$

En général, $\mathbf{V}_h$ n'est pas inclu dans $\tilde{\mathbf{V}}$, et (P_h) constitue une *approximation externe* de $(\tilde{P})$.

A.3.2 Existence, unicité et estimation d'erreur

Afin d'obtenir des estimations d'erreur, on fait les hypothèses :

i) $\tilde{a}$ est uniformément $\mathbf{V}_h$-elliptique : il existe une constante $\alpha^* > 0$, indépendante de h, telle que

$$\tilde{a}(v_h, v_h) \geq \alpha^* \|v_h\|^2_{\tilde{\mathbf{X}}}, \ \forall v_h \in \mathbf{V}_h \quad (A.37)$$

ii) $\tilde{b}$ satisfait une condition inf-sup faible discrète : il existe une constante $\beta^* > 0$, indépendante de h, telle que

$$\sup_{v_h \in \mathbf{X}_h, v_h \neq 0} \frac{\tilde{b}(v_h, \mu_h)}{\|v_h\|_{\tilde{\mathbf{X}}}} \geq \beta^* \|\mu_h\|_{\mathbf{M}}, \ \forall \mu_h \in \mathbf{M}_h \quad (A.38)$$

qui sont les analogues discrets des hypothèses (A.31) et (A.32).

On a le

Théorème A.3 *Existence, unicité et estimations abstraites d'erreur*

- *On suppose* $\mathbf{V}_h$ *non-vide, et* $\tilde{a}$ $\mathbf{V}_h$*-elliptique. Alors* (P_h) *admet une unique solution* $u_h \in \mathbf{V}_h(\chi)$ *et on a l'estimation d'erreur :*

$$\|u - u_h\|_{\tilde{\mathbf{X}}} \leq (1 + \frac{\|\tilde{a}\|}{\alpha^*} \inf_{v_h \in \mathbf{V}_h(\chi)} \|u - v_h\|_{\tilde{\mathbf{X}}} + \frac{1}{\alpha^*} \inf_{\mu_h \in \mathbf{M}_h} \sup_{v_h \in \mathbf{V}_h} \frac{|\tilde{b}(v_h, \lambda - \mu_h)|}{\|v_h\|_{\tilde{\mathbf{X}}}} \quad (A.39)$$

- *On suppose que* $\tilde{b}$ *vérifie la condition inf-sup faible discrète. Alors* $\mathbf{V}_h 9\chi) \neq \emptyset$ *et* (Q_h) *admet exactement une solution* (u_h, λ_h)*, où* u_h *est la solution de* (P_h)*. De plus,* λ_h *vérifie l'estimation d'erreur :*

$$\begin{aligned}\|\lambda - \lambda_h\|_{\mathbf{M}} &= \frac{\|\tilde{a}\|}{\beta^*} \|u - u_h\|_{\tilde{\mathbf{X}}} \\ &+ \inf_{\mu_h \in \mathbf{M}_h} \left(\frac{\|\tilde{b}\|}{\beta^*} \|\lambda - \mu_h\|_{\tilde{\mathbf{M}}} + \|\lambda - \mu_h\|_{\mathbf{M}} \right)\end{aligned} \tag{A.40}$$

A.3.3 Estimation plus fine de l'erreur

L'estimation (A.40) n'est *pas optimale*, car on donne une majoration pour $\|\lambda - \lambda_h\|_{\mathbf{M}}$ en termes de $\|\lambda - \mu_h\|_{\tilde{\mathbf{M}}}$, alors qu'en général, $\tilde{\mathbf{M}}$ est strictement inclu dans $\mathbf{M}$.

Il est habituellement difficile d'évaluer directement une expression comme

$$\inf_{v_h \in \mathbf{V}_h(\chi)} \|u - v_h\|_{\tilde{\mathbf{X}}}$$

En fait, il est possible de réduire cette expression à une erreur d'approximation dans $\mathbf{X}_h \supset \mathbf{V}_h(\chi)$, bien que ce procédé ne soit pas toujours optimal.

Puisque M_h est de dimension finie, il existe une constante $S(h)$ telle que

$$\|\mu_h\|_{\tilde{\mathbf{M}}} \leq S(h) \, \|\mu_h\|_{\mathbf{M}}, \ \forall \mu_h \in \mathbf{M}_h \tag{A.41}$$

Avec ces remarques, on a le

Corollaire A.4 *Estimation plus fine de l'erreur*

Avec les hypothèses du théorème précédent, le problème admet une solution unique (u_h, λ_h)*, et il existe une constante* $c > 0$*, ne dépendant que de* α^**,* $\|\tilde{a}\|$ *et* $\|\tilde{b}\|$ *telle que*

$$\begin{aligned}&\|u - u_h\|_{\tilde{\mathbf{X}}} + \|\lambda - \lambda_h\|_{\tilde{\mathbf{M}}} \\ &\quad \leq \quad c\,(1 + S(h)) \left[\inf_{v_h \in \mathbf{X}_h} \|u - v_h\|_{\tilde{\mathbf{X}}} + \inf_{\mu_h \in \mathbf{M}_h} \|\lambda - \mu_h\|_{\tilde{\mathbf{M}}} \right]\end{aligned} \tag{A.42}$$

Remarque A.3

Dans les applications du chapitre 2, S(h) dépend de la dimension de $\mathbf{M}_h$, et, plus précisément :

$$\lim_{\dim \mathbf{M}_h \to +\infty} S(h) \longrightarrow +\infty$$

□

Appendix B

Algorithme du gradient conjugé

L'objet de cette annexe est de présenter l'algorithme du *gradient conjugé* [17, 50, 21] utilisé pour résoudre les sous-problèmes de type Stokes.

On considère donc les systèmes linéaires de la forme :

$$\begin{pmatrix} A & B \\ B^t & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \tag{B.1}$$

où A est une matrice carrée définie symétrique positive $N \times N$, B une matrice rectangulaire $N \times M$, $f \in \mathbb{R}^N$ et $g \in \mathbb{R}^M$.

B.1 L'algorithme du gradient conjugé

On remarque que le système précédent est équivalent à

$$(B^t A^{-1} B)p = B^t A^{-1} f - g \tag{B.2}$$

$$u = -A^{-1}Bp + A^{-1}f \tag{B.3}$$

En appliquant l'algorithme du gradient conjugé [13] à (B.2), et faisant intervenir u à l'aide de (B.3), on obtient le très classique :

algorithme du gradient conjugé appliqué au problème (B.1)

Soit p_o donné (quelconque);

Choisir une matrice de préconditionnement C, définie symétrique positive, $M \times M$;

- préliminaires :
  poser :

$$\begin{aligned} p^0 &:= p_o && \text{(B.4)} \\ s^0 &:= f - B.p^0 && \text{(B.5)} \end{aligned}$$

  résoudre :

$$A.u^0 = s^0 \tag{B.6}$$

- pour $n \geq 0$, on suppose u^n et p^n connus, et on calcule u^{n+1} et p^{n+1} suivant :

  – calculer l'opposé du gradient g^n et la direction conjugée w^n suivant :

    * si $n = 0$, résoudre :

$$C.g^0 = B^t.u^0 - g \tag{B.7}$$

    et poser

$$w^0 = g^0 \tag{B.8}$$

    * si $n \geq 1$ poser :

$$\begin{aligned} g^n &:= g^{n-1} - \rho_{n-1}.z^{n-1} && \text{(B.9)} \\ \lambda_n &:= \frac{\|g^n\|_C}{\|g^{n-1}\|_C} && \text{(B.10)} \\ w^n &:= g^n + \lambda_n.w^{n-1} && \text{(B.11)} \end{aligned}$$

  – calculer ensuite la direction de descente v^n pour le champ u^n en posant :

$$s^n := B.w^n \tag{B.12}$$

  et en résolvant :

$$A.v^n = s^n \tag{B.13}$$

  puis le pas de descente ρ_n en résolvant :

$$C.z^n = Bt.v^n \tag{B.14}$$

  et en posant :

$$\rho_n := \frac{\|g^n\|_C}{(z^n, w^n)_C} \tag{B.15}$$

– calculer enfin p^{n+1} et u^{n+1} en une étape de descente :

$$p^{n+1} := p^n + \rho_n.w^n \tag{B.16}$$

$$u^{n+1} := u^n - \rho_n.v^n \tag{B.17}$$

La norme $\|.\|_C$ est définie pour tout $w \in \mathbb{R}^M$ par

$$\|w\|_C = w^t.C.w$$

B.2 Choix d'un préconditionnement

Soit $M^* = \text{rang}(B) \leq M$ le rang de la matrice B. L'algorithme converge théoriquement en au plus M^* itérations. Du fait des erreurs d'arrondis dûes à la représention informatique des réels, ce résultat n'intervient pas en pratique. D'autre part, dans les applications $\text{rang}(B) = M - 1$ est grand et on souhaite obtenir un résultat convenable en beaucoup moins d'itérations. Un test d'arrêt satisfaisant consiste à évaluer la norme de la contrainte

$$\|B^t.u^n - g\| \tag{B.18}$$

et à stopper l'algorithme si cette quantité est suffisamment petite.

Il s'agit aussi de choisir la matrice de préconditionnement C pour que l'algorithme converge le plus vite possible. Lorsque le système (B.1) provient de la discrétisation des problèmes de type Stokes de la forme :

$$\begin{aligned} (\lambda I - \eta\Delta)\vec{u} \quad + \vec{\nabla} p &= \vec{f} \text{ dans } \Omega \\ -\text{div}\, \vec{u} &= g \text{ dans } \Omega \\ \vec{u} &= \vec{u}_\Gamma \text{ sur } \Gamma \end{aligned}$$

Cahouët et et Chabard ont montré [10] qu'un bon choix pour C est :

$$C = (\eta I_h^{-1} - \lambda(\Delta_h)^{-1})^{-1}$$

avec des conditions aux bords de type Neumann pour le Laplacien scalaire Δ_h, et où I_h est la matrice de masse pour les pressions, encore notée M_p au paragraphe 3.5.

Lorsque λ est petit devant η (cas des grands pas de temps dans l'algorithme du θ-schéma), ce choix peut donc se ramener à un préconditionnement par la matrice de masse :

$$C = M_p$$

rendue diagonale à l'aide d'une formule de quadrature appropriée, et donc facile à inverser dans (B.7) et (B.14).

B.3 La méthode du Lagrangien Augmenté

Une autre approche possible pour accélérer la convergence de cet algorithme est la méthode du *Lagrangien Augmenté*, introduite par Fortin et Glowinski [17]. Pour cela, on remplace (B.1) par le problème équivalent :

$$\begin{pmatrix} (A + rB.B^t) & B \\ B^t & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f + rB.g \\ g \end{pmatrix} \tag{B.19}$$

et on applique la méthode précédente.

L'algorithme obtenu converge théoriquement d'autant plus vite que r est grand. Cependant, le conditionnement de la matrice $A_r = A + rB.B^t$ tend vers l'infini avec r. La résolution des sous-problèmes (B.6) ou (B.13) s'en trouve affectée, particulièrement si on utilise une méthode itérative de résolution des systèmes linéaires (Gauss-Seidel [24], par exemple).

Par conséquent, on préfère factoriser la matrice A_r, une fois pour toutes (lors de l'initialisation de l'algorithme du θ-schéma) sous sa forme de Choleski. Ceci nous permettra de choisir de grandes valeurs de r et d'obtenir un résultat satisfaisant en très peu d'itérations (on pourra choisir le paramètre r aussi grand que possible et ne conduisant pas à un débordement du format "réel" en machine lors de la factorisation).

B.4 Factorisation de Choleski

Afin de compléter cette annexe, nous rappelons le procédé de factorisation de Choleski [13] d'une matrice $A = (a_{ij})_{1 \leq i,j \leq N}$, symétrique définie positive :

$$A = L.L^t$$

B.4.1 Algorithme abstrait

La matrice triangulaire inférieure $L = (l_{ij})_{1 \leq i,j \leq N}$. est alors définie par :

- $j = 1$:

$$l_{1,1} = \sqrt{a_{1,1}}, \ i = 1 \tag{B.20}$$

$$l_{i,1} = \frac{a_{i,1}}{l_{1,1}}, \ 2 \leq i \leq N \tag{B.21}$$

- $2 \leq j \leq N$:

$$l_{j,j} = \left(a_{j,j} - \sum_{k=1}^{j-1} l_{j,k}^2 \right)^{\frac{1}{2}}, \ i = j \tag{B.22}$$

$$l_{i,j} = \frac{1}{l_{j,j}} \left(a_{i,j} - \sum_{k=1}^{j-1} l_{i,k} l_{j,k} \right), \ j+1 \leq i \leq N \tag{B.23}$$

On vérifie facilement que le temps de factorisation est en $\mathcal{O}(N^3)$, et l'encombrement mémoire en $\mathcal{O}(N^2)$, ce qui est considérable.

B.4.2 Cas d'une matrice à structure-bande

La factorisation de Choleski respecte la structure bande de la matrice [13] : si d est la largeur de bande de A, définie comme étant le plus petit entier $\tilde{d}$ tel que :

$$a_{i,j} = 0, \ |i - j| \geq \tilde{d}$$

alors la matrice L est aussi à structure bande, de largeur d.

Cette propriété nous permet d'optimiser le temps de calcul de L et la place mémoire occupée. Nous utilisons alors une nouvelle définition de L :

- $j = 1$

$$l_{1,1} = \sqrt{a_{1,1}} \quad i = 1 \tag{B.24}$$

$$l_{i,1} = \frac{a_{i,1}}{l_{1,1}} \quad 2 \leq i \leq d \tag{B.25}$$

- $2 \leq j \leq N$

$$l_{j,j} = \left(a_{j,j} - \sum_{k=\max(1,j-d+1)}^{j-1} l_{j,k}^2 \right)^{\frac{1}{2}}, \ i = j \tag{B.26}$$

$$l_{i,j} = \frac{1}{l_{j,j}} \left(a_{i,j} - \sum_{k=\max(1,i-d+1)}^{j-1} l_{i,k} l_{j,k} \right), \ j+1 \leq i \leq \min(N, j+d-1) \tag{B.27}$$

On vérifie que le temps de factorisation est en $\mathcal{O}(N.d^2)$, et l'encombrement mémoire en $\mathcal{O}(N.d)$.

En numérotant de façon appropriée les éléments, on peut réduire de façon très importante la largeur de la bande, et ainsi le coût de la factorisation, et l'encombrement en mémoire. Par exemple, si le domaine est alongé (marche, convergent, ...) on prendra soin de numéroter les éléments suivant la plus petite longeur puis suivant la plus grande. Dans ce cas $d \ll N$. Rappelons que cette méthode se généralise sur un domaine quelconque (algorithme de numérotation frontale, ...).

B.5 Coût de la résolution directe

Il s'agit maintenant de résoudre des systèmes linéaires de la forme :

$$(L.L^t).u = b \tag{B.28}$$

ce qui est réalisé en deux étapes :

i) on résoud le système triangulaire inférieur $L.v = b$ en un balayage descendant :

pour $i = 1$ à N :

$$v_i := \frac{1}{l_{i,i}} \left(b_i - \sum_{j=\max(1,i-d+1)}^{i-1} l_{i,j} v_j \right) \tag{B.29}$$

ii) puis le système triangulaire supérieur $L^t.u = v$, en un balayage remontant :

pour $i = N$ à 1, par pas de -1 :

$$u_i := \frac{1}{l_{i,i}} \left(v_i - \sum_{j=i+1}^{mini(N,i+d-1)} l_{j,i} u_j \right) \tag{B.30}$$

On vérifie que le coût de cette résolution est en $\mathcal{O}(d.N)$. Remarquons d'autre part qu'il est possible de faire coïncider l'implémentation des tables u et v lors de cette résolution : il n'y a donc pas d'encombrement mémoire supplémentaire.

B.6 Coût de la méthode

La factorisation étant réalisée une fois pour toutes, les sous-problèmes (B.6) et (B.13) peuvent donc être résolus en un temps de calcul $\mathcal{O}(d.N)$.

Ainsi le coût d'une itération de la méthode du gradient conjugé reste en $\mathcal{O}(d.N + M)$.

Lorsque la précision souhaitée sur la contrainte est fixée par le test d'arrêt (B.18) , le temps de calcul global de cet algorithme (sans compter le temps de factorisation) est estimé à :

$$N^{1+\varepsilon} \log N$$

avec $\varepsilon \approx \frac{1}{5}$.

L'algorithme nécéssite le stockage permanent des champs u, p, g, w, z, tandis que le champ s peut être alloué temporairement (par un allocateur dynamique de mémoire [37]).

L'encombrement asymptotique en mémoire est donc directement lié à celui de la matrice de Choleski L.

Appendix C

Assemblage des matrices d'éléments finis

Les matrices issues de l'approximation de problèmes aux limites par la méthode des éléments finis sont généralement très *ceuses*. Pour des raisons d'encombrement en mémoire des ordinateurs, et afin d'optimiser les algorithmes opérant sur ces matrices, il est essentiel de réduire le volume de stockage au minimum.

Pour cela, nous considérons dans cette annexe une *structure de données* ne représentant que les éléments non-nuls d'une matrice : ce procédé n'étant pas standard des logiciels habituellement commercialisés, il nécéssitera ici quelques developpements.

Les matrices considérées sont obtenues par addition de *matrices élémentaires*. Cette opération, appelée encore *assemblage*, ne pose aucun problème si la structure de la matrice globale est connue d'avance. Malheureusement, cette structure est généralement inconnue.

Après avoir introduit quelques notations, et rappelé certaines propriétés d'implémentation de la méthode des éléments finis, nous donnons un algorithme abstrait d'assemblage.

Ensuite, nous présentons une structure de données [26] pour les matrices creuses, ainsi qu'un algorithme d'assemblage concret associé.

Enfin, nous examinerons certains algorithmes opérant concrètement sur une telle structure de donnée : produits matrice-vecteur et méthode de surrelaxation successives symétrisée, et montrerons l'optimalité en temps de calcul et place mémoire.

C.1 Présentation du problème

C.1.1 Notations

Soient X et Y deux espaces de Hilbert, et a une forme bilinéaire de $X \times Y$ dans $\mathbb{R}$. On note $X^{'}$ (resp. $Y^{'}$) le dual de X (resp. Y), et $< .,. >$ la forme de dualité entre $X^{'}$ et X ou entre $Y^{'}$ et Y, lorsqu'il n'y aura pas de confusion.

Soient $X_h \subset X$ et $Y_h \subset Y$ des sous espaces de dimension finie $N = \mathrm{dim} X_h$ et $M = \mathrm{dim} Y_h$ On désigne par $(u_h^{(i)})_{1\leq i\leq N}$ et $(v_h^{(j)})_{1\leq j\leq M}$ deux bases de X_h et Y_h, respectivement.

On définit alors la matrice $A = (a_{i,j})_{\substack{1\leq i\leq N\\ 1\leq j\leq M}}$ par :

$$a_{i,j} = a(u_h^{(i)}, v_h^j)$$

l'opérateur $\mathrm{A} \in \mathcal{L}(Y; X^{'})$ par

$$< \mathrm{A}v, u >= a(u,v),\ \forall u \in X,\ \forall v \in Y \tag{C.1}$$

et son adjoint $\mathrm{A}^* \in \mathcal{L}(X; Y^{'})$ par :

$$< \mathrm{A}^* u, v >= a(u,v),\ \forall u \in X,\ \forall v \in Y \tag{C.2}$$

Nous considérerons que la matrice A (resp. A^t) représente l'opérateur A (resp. A^*) sur $\mathcal{L}(Y_h; X_h)$ (resp. $\mathcal{L}(X_h; Y_h)$).

C.1.2 Matrices élémentaires

Soit $\mathcal{T}_h$ une triangulation d'un domaine Ω. Les bases $(u_h^{(i)})_{1\leq i\leq N}$ et $(v_h^{(j)})_{1\leq j\leq M}$ sont associées à des noeuds de la triangulation : $(x_i)_{1\leq i\leq N}$ et $(y_i)_{1\leq i\leq M}$ respectivement. On supposera que le supports d'une fonction de base $u_h^{(i)}$ (resp. $v_h^{(j)}$) est réduit aux éléments voisins du noeud x_i (resp. y_j) associé.

Ainsi le coefficient $a_{i,j}$ de la matrice A s'écrit encore :

$$a_{i,j} = \int_{\mathrm{supp}(u_h^{(i)})\cap \mathrm{supp}(v_h^{(j)})} \mathrm{A}v_h^{(j)}\, u_h^{(i)}\, d\Omega = \sum_{\substack{K\in\mathcal{T}_h\\ x_i,y_j\in K}} \int_K \mathrm{A}v_h^{(j)}\, u_h^{(i)}\, dK \tag{C.3}$$

et en notant

$$e_{i,j}^K = \int_K \mathrm{A}v_h^{(j)}\, u_h^{(i)}\, dK \tag{C.4}$$

nous faisons apparaître la matrice élémentaire $E_K = (e^K_{i,j})_{\substack{1\leq i\leq N\\ 1\leq j\leq M}}$

Ainsi la matrice globale A se décompose en une somme finie de matrices élémentaires :

$$A = \sum_{K\in\mathcal{T}_h} E_K \tag{C.5}$$

C.1.3 Extraction des degrés de liberté bloqués

Une fonction $u_h \in X_h$ (resp. $v_h \in Y_h$) se décompose sur la base $(u_h^{(i)})_{1\leq i\leq N}$ (resp. $(v_h^{(j)})_{1\leq j\leq M}$) suivant :

$$u_h = \sum_{j=1}^{N} u_i\, u_h^{(i)} \ , \ v_h = \sum_{j=1}^{M} v_i\, v_h^{(i)} \tag{C.6}$$

Cependant, du fait de conditions aux limites associées à l'opérateur A, certaines de ces composantes peuvent être imposées (*degrés de liberté bloqués*), ce que on note :

$$u_i = \bar{u}_i,\ v_j = \bar{v}_j \tag{C.7}$$

Quitte à renuméroter les fonctions de base, nous supposerons que les n (resp. m) premiers degrés de libertés de u_h (resp. v_h) sont libres, et les $N-n$ (resp. $M-m$) suivants bloqués.

Ainsi, nous pouvons réduire la matrice A en $\tilde{A}$:

$$\tilde{A} = (a_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq m}} \tag{C.8}$$

On fait apparaître deux vecteurs second-membres $b \in \mathbb{R}^n$ et $c \in \mathbb{R}^m$, construits à partir de $(\bar{v}_j)_{M-m\leq j\leq M}$ et $(\bar{u}_i)_{N-n\leq i\leq N}$, respectivement, et tels que on ait :

$$a(u_h, v_h) = U^t.(\tilde{A}.V - b) = V^t.(\tilde{A}^t.U - c) \tag{C.9}$$

pour tout $U = (u_i)_{1\leq i\leq n} \in \mathbb{R}^n$, $V = (v_j)_{1\leq j\leq m} \in \mathbb{R}^m$. Dans la suite, et lorsqu'il n'y aura pas d'ambiguïté, nous noterons encore A la matrice ainsi réduite.

C.1.4 Algorithme abstrait d'assemblage

Soit K un élément de $\mathcal{T}_h$, E_K la matrice élémentaire et b_K et c_K les seconds membres associés, construits à partir des degrés de liberté bloqués sur K.

Ces matrices et vecteurs sont associés à une numérotation locale à l'élément K :

$$
\begin{aligned}
E_K &= (e^K_{i,j})_{\substack{1\le i\le n_K\\ 1\le j\le m_K}} && \text{(C.10)}\\
b_K &= (b^K_i)_{1\le i\le n_K} && \text{(C.11)}\\
c_K &= (c^K_j)_{1\le j\le m_K} && \text{(C.12)}
\end{aligned}
$$

Les tables $iglob_K(i),\ 1 \le i \le n_K$ et $jglob_K(j),\ 1 \le j \le m_K$ associées à l'élément K vont nous permettre de passer de cette numérotation locale à la numérotation globale. Par exemple, pour des espaces de fonctions continues affines par morceaux sur $\mathcal{T}_h$, et dans le cas d'éléments triangulaires, on a $n_K = 3$, et la numérotation des noeuds de la triangulation nous permet d'accèder directement à la numérotation globale des degrés de liberté.

Nous pouvons énoncer un
algorithme abstrait d'assemblage

- mise à zéro de la matrice globale A et des seconds membres b et c;
- pour tout élément $K \in \mathcal{T}_h$:
  - calculer la matrice E_K et les seconds membres b_K et c_K;
  - effectuer les sommations :

$$
\begin{aligned}
A_{iglob(i),jglob(j)} &:= A_{iglob(i),jglob(j)} + e^K_{i,j},\ 1 \le i \le n_K,\ 1 \le j \le m_K && \text{(C.13)}\\
b_{iglob(i)} &:= b_{iglob(i)} + b^K_i,\ 1 \le i \le n_K && \text{(C.14)}\\
c_{jglob(j)} &:= c_{jglob(j)} + c^K_j,\ 1 \le j \le m_K && \text{(C.15)}
\end{aligned}
$$

C.2 Représentation compactée des matrices

C.2.1 La strucure de données

La matrice A est génériaement très creuse : en notant

$$NZ(A) = \text{card}\{(i,j);\ 1 \le i \le n,\ 1 \le j \le m,\ a_{i,j} \ne 0\} \qquad \text{(C.16)}$$

le nombre de coefficients non-nuls de A, on a généralement :

$$\lim_{\substack{n\to+\infty\\ m\to+\infty}} \frac{NZ(A)}{nm} = 0 \qquad \text{(C.17)}$$

Il est alors essentiel de réduire le volume des données pour stocker la matrice A en mémoire des ordinateurs.

S'inspirant de [26], nous présentons une méthode de stockage des matrices d'éléments finis ne conservant que les valeurs non-nulles des coefficients. Un coefficient non-nul sera représenté par un triplet :

(*valeur*, indice de *ligne*, indice de *colonne*)

L'ensemble des coefficients non-nuls est alors stocké dans une table de NZ triplets, par ordre d'indice ligne puis colonne croissant.

Connaissant seulement un couple d'indices (i, j), il n'est alors pas possible d'accéder directement à la valeur du coefficient $a_{i,j}$. Pour permettre de rechercher rapidement un coefficient donné, on dispose de plus d'une table *début-ligne(i)*, $1 \leq i \leq n + 1$, contenant l'index dans les tables précédentes, du premier élément non-nul de la ligne i. Par convention, *début-ligne*$(n + 1) = NZ + 1$. Ainsi, la
fonction *accés (i,j)*
retourne en résultat un *index* dans les tables *valeur*, *ligne* et *colonne* lorsque le coefficient $a_{i,j} \neq 0$, et un code *nul* lorsque $a_{i,j} = 0$ L'algorithme est le suivant :

- *jcour* := j
- *index* := *début-ligne (i)*
- tant que *index* < *début-ligne (i+1)* et que *colonne(index)* < j

  incrémenter l'*index*.

- à la sortie de cette boucle :
  si *index* ≥ *début-ligne(i+1)* ou *colonne(index)* > j, c'est que le coefficient est nul, et alors :

  index := nul

  sinon, *colonne(index)* = j et la valeur de l'*index* est correcte.

Lorsque la matrice A est carrée $(n = m)$ et tous les coefficients diagonaux non-nuls , il est intéréssant d'accéder directement à un coefficient diagonal, ceci afin d'optimiser certains algorithmes opérant sur la matrice. Pour cela, on disposera d'une table *diag* de n index dans la table *valeur*, et telle que *valeur(diag(i))* contienne la valeur du coefficient $a_{i,i}$.

C.2.2 première étape : détermination de la structure creuse

Listes d'éléments non-nuls

La détermination de cette structure de donnée nécéssite une phase préliminaire. Lors de cette première phase, on construit une table *chaîne(i)*, $1 \leq i \leq n$, de n listes chaînées. Chaque *chaîne(i)* contient la liste des éléments non-nuls de la ligne i, par ordre d'indice de colonne croissant.

Lors de cette première phase, un nouveau coefficient $a_{i,j} \neq 0$ sera représenté par un maillon j dans la liste *chaîne(i)*.

Un maillon sera en fait constitué de deux champs : le premier est l'indice de colonne j, et le second, une information *suivant* permettant d'accéder au maillon suivant dans la liste. Le champ *suivant* pourra prendre la valeur *fin* lorsque la liste sera terminée.

On supposera dans la suite que la fonction *alloue-maillon(j,suiv)* alloue un espace en mémoire, y charge la valeur du couple *(j,suiv)*, et renvoie une information *pointeur* permettant d'accéder aux champs. S'inspirant du langage C [37] dans lequel nous avons implémenté cet algorithme, on notera dans la suite *pointeur*→*j* et *pointeur*→*suivant* l'accès aux champs j et *suivant*, respectivement, d'un maillon désigné par la variable *pointeur*.

Enfin, l'entier *NZ* compte le nombre de maillons effectivement insérés.

L'algorithme suivant insère à sa place, s'il n'y existe pas, un maillon d'indice colonne j dans une *chaîne(i)* :

procédure *insérer (i,j)*

- si *chaîne(i)* = *fin* ou sinon, si *chaîne(i)*→*j* > *j*, on insère le maillon en tête de liste :

  chaîne(i) := *alloue-maillon (j, chaîne (i))*

  incrémenter *NZ*

- sinon, on parcourt la liste *chaîne(i)* à la recherche du maillon j :

  courant := *chaîne(i)*

  tant que *courant* ≠ *fin* et que *courant*→*j* < *j*, passer au maillon suivant :

  précédent := *courant*

  courant := *courant*→*suivant*

  Lorsque cette boucle s'arête, il y a deux possibilités :

  soit *courant* = *fin* et on va insérer le maillon à la fin de la liste, soit *courant*$\rightarrow$*j* $\geq j$, et alors, si le maillon n'existe pas déjà, on l'insère entre le maillon précédent et le maillon courant :

si *courant* = *fin* ou sinon, si *courant*$\rightarrow$*j* $> j$:

précédent$\rightarrow$*suivant* := *alloue-maillon (j, précedent*$\rightarrow$*suivant)*
incrémenter *NZ*

La construction de la table *chaîne(i)*, $1 \leq i \leq n$ est alors réalisée en un balayage de la triangulation $\mathcal{T}_h$:
algorithme de construction des structures chaînées

- initialisations :
  NZ := 0
  chaîne(i) := *fin*, $\forall i,\ 1 \leq i \leq n$

- pour tout $K \in \mathcal{T}_h$,

  accéder à la structure de la matrice élémentaire E_K :

$$n_K, m_K, iglob_K, jglob_K$$

  pour tout i, j, $1 \leq i \leq n_K$, $1 \leq j \leq m_K$,

  insérer $(iglob_K(i), jglob_K(j))$

Lorsque cet algorithme s'arête, la variable *NZ* contient le nombre de coeficients (à priori) non-nuls de la matrice.

Construction de la structure creuse

Ceci nous permet d'allouer dynamiquement les tables *ligne* et *colonne* de *NZ* indices exactement. La table *début-ligne* contiendra $n + 1$ index (par commodité), et, si la matrice est carrée, la table *diag* en contiendra n.

En déroulant les listes chaînées, nous allons pouvoir remplir ces tables. Nous supposerons que la procédure *libérer-maillon(pointeur)* dé-alloue l'emplacement d'un maillon désigné par *pointeur*.
algorithme de construction de la structure creuse

- initialisations :
  index := 0
  début-ligne(n+1) := $NZ + 1$, par commodité

- pour tout i, $1 \leq i \leq n$:

début-ligne(i) := *index*

courant := *chaîne(i)*

puis, on déloule la liste *chaîne(i)* :
tant que *courant* $\neq$ *fin*,

$*$ on charge les informations du maillon dans la structure de données :
ligne(index) := *i*
colonne(index) := *courant*$\rightarrow$*j*
si la matrice est carrée et que j = courant$\rightarrow$j, alors, de plus :
diag(i) := *index*

$*$ il reste à passer au maillon suivant en libérant le précédent :
précédent := *courant*
courant := *courant*$\rightarrow$*suivant*
libérer-maillon(précédent)
incrémenter *index*

lorsque l'algorithme s'arrête, on a $index = NZ$.

C.2.3 Deuxième étape : assemblage effectif

La structure creuse de la matrice étant déterminée, il suffit de transposer ici l'algorithme abstrait d'assemblage : on parcourt une deuxième fois la triangulation, et somme les matrices élémentaires dans la matrice creuse globale. Pour cela, on dispose des tables *valeur*, b et c de NZ, n et m réels, respectivement.
algorithme concret d'assemblage

- mise à zéro des tables *valeur*, b et c;
- pour tout $K \in \mathcal{T}_h$,
  - accèder aux matrices et second-membres élémentaires :

$$n_K, m_K, E_K, b_K, cK, iglob_K, jglob_K$$

  - sommer dans le second membre c :

$$c(jglob_K(j)) := c(jglob_K(j)) + c_K(j),\ \forall j,\ 1 \leq j \leq m_K$$

  - pour tout i, $1 \leq i \leq n_K$, faire :

sommer dans le second membre b :

$$b(iglob_K(i)) := b(iglob_K(i)) + b_K(i)$$

pour tout j, $\leq j \leq m_K$, on recherche l'index du coefficient correspondant dans la structure de donnée, puis on somme :

index := *accès*($iglob_K(i)$,$jglob_K(j)$)
valeur(index) := *valeur(index)* +$E_K(i,j)$

C.3 Algorithmes opérant sur la structure creuse

C.3.1 Produits matrice-vecteur

Disposant de cette structure creuse, nous pouvons réaliser des produits matrice-vecteur avec un coût en $\mathcal{O}(NZ)$.

algorithme du produit $U := A.V$

pour $i := 1$ à n

$U(i) := 0$
pour *index* := *début-ligne(i)* à *début-ligne(i* $+ 1$*)* -1
$U(i) := U(i)+$ *valeur(index).* V*(colonne(index))*

Dans la pratique, nous n'utiliserons pas l'accès indexé dans les tables V, *début-ligne*, *ligne* et *valeur*, dont le parcourt est séquentiel. En effet, il est alors possible de parcourir ces tables à l'aide de pointeurs [37, page 85], ce qui augmente considérablement l'efficacité de cet algorithme.

algorithme du produit $V := A^t.U$

$V(j) := 0$, $1 \leq j \leq m$
pour *index* := 1 à *NZ*

V(colonne(index)) := V(colonne(index)) + valeur(index).U(ligne(index))

Remarquons qu'ici aussi, il est possible d'optimiser le parcours séquentiel des tables *valeur*, *ligne* et *colonne*.

C.3.2 Sur-relaxations successives symétrisées

Rappellons brièvement le principe de la méthode. On suppose la matrice carrée ($n = m$), à diagonale strictement dominante (voir page 98). Une itération de la méthode de sur-relaxations successives (SOR, en anglais en abrégé) de paramètre $\omega \in]0,2[$ sur le système :

$$A.U = b$$

sécrit :

$$U(i) := (1-\omega)U(i) + \frac{\omega}{a_{i,i}}\left(b(i) - \sum_{j=1}^{n} a_{i,j}U(i)\right), \text{ ceci pour } i = 1 \text{ à } n$$

En utilisant la structure creuse des matrices, l'algorithme s'énonce :
algorithme d'une itération SOR

- une itération SOR :
  pour $i = 1$ à n,
  - une boucle pour calculer la somme :
    somme := 0
    pour *index* := *début-ligne(i)* à *début-ligne(i + 1)*−1 faire
        somme := *somme* + *valeur(index).U(colonne(index))*
  - puis corriger $U(i)$:
    $U(i) := (1-\omega).U(i) - \omega.somme/valeur(diag(i))$

Une itération de la méthode SOR-inverse consiste à executer les affectations précédentes en parcourrant pour $i = n$ à 1, par pas de -1. Enfin, une itération de la méthode de sur-relaxations successives symétrisées (SSOR, en anglais en abrégé) enchaîne les deux parcourts précédents.

C.3.3 Coûts optimaux

Les matrices issues de la méthode des éléments finis sont très creuses, et on a en général :

$$NZ = \mathcal{O}(N)$$

Par exemple, si la matrice provient de l'approximation de l'opérateur harmonique Δ, que on utilise des fonctions continues affines par morceaux sur un maillage rectangulaire à côtés parrallèles aux axes, et à condition d'évaluer les intégrales à l'aide de la formule des trapèzes, on trouve :

$$NZ \approx 5N$$

Ainsi, les algorithmes d'assemblage, de produits matrice-vecteur et d'une itération SSOR conduisent à des coûts optimaux en place-mémoire et en temps de calcul.

Bibliography

[1] E. M. Abdalass. *Résolution performante du probème de Stokes par mini-éléments, maillages auto-adaptatifs et méthodes multigrilles—applications.* thèse, Ecole Centrale de Lyon, 1987.

[2] D. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the stokes equations. Preprint, University of Pavia, 1983.

[3] G. Astarita and G. Marrucci. *Principle of Non-Newtonian Fluid Mechanics.* Mac Graw Hill, 1974.

[4] J. Baranger. Communication personelle. 1989.

[5] J. Baranger, P. Georget, and K. Najib. Error estimates for a mixed finite element method for a non-newtonian flow. *Journal of Non-Newtonian Fluid Mechanics*, 23:415–421, 1987.

[6] H. Brezis. *Analyse fonctionelle. Théorie et applications.* Masson, 1983.

[7] F. Brezzi. On the existence, uniqueness and approximation of saddle point problems arising from lagrangian multiplier. *RAIRO, série Analyse Numérique*, R2:129–151, 1974.

[8] F. Brezzi and J. Pitkaranta. On the stabilisation of finite element approximation of the stokes equations. *Notes on Numerical Fluid Mechanics*, vol 10, Efficient Solution of Elliptic Systems, Proceeding:11–19, 1984.

[9] A. N. Brooks and T. J. R Hughes. Streamline-upwind/petrov-galerkin formulation for convection dominated flow with particular emphasis on the incompressible navier-stokes equations. *Comp. Meth. in Appl. Mech and Eng.*, 32:199–259, 1982.

[10] J. Cahouët and J. P. Chabard. Some fast 3-d finite element solver for generlized stokes problems. Technical Report HE/41/87.03, EDF-LNH, 1987.

[11] I. Christie, D. F. Griffiths, A. R. Mitchell, and O. C. Zienckiewicz. Finite element methods for second order differential equations with significant first order derivatives. *Int. J. Numer. Meth. Eng.*, 10:1389–1396, 1976.

[12] P. G. Ciarlet. *The finite element method for elliptic problems.* North–Holland, Amsterdam, 1978.

[13] P. G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation.* Masson, 1985.

[14] M. J. Crochet, Debaut, and J. M. Marchal. Numerical simulation of highly viscoelstic flows though an abrupt contraction. *Journal of Non-Newtonian Fluid Mechanics*, 29:119–146, 1988.

[15] M. J. Crochet and J. M. Marchal. A new mixed finite element for calculating viscoelastic flow. *J. of Non-Newtonian Fluids Mechanics*, 26:77–114, 1987.

[16] M. Fortin and A. Fortin. A new approach for the fem simulation of viscoelastic flow. *Journal of Non-Newtonian Fluid Mechanics*, 32:295–310, 1989.

[17] M. Fortin and R. Gowinski. *Lagrangian Augmented Methods.* North Holland, 1981.

[18] M. Fortin and R. Pierre. On the convergence of the mixed method of crochet and marchal for viscoelastic flow. *to appear in Comput. Meth in Appl. Mech. Eng.*, 1989.

[19] K. O. Friedrichs. Symetric positive differential equations. *Comm. on Pure an Appl. Math.*, II:333–418, 1958.

[20] V. Girault and P. A. Raviart. An analysis of a mixed finite element method for the navier-stokes equations. *Numerische Mathematik*, 33:235–271, 1979.

[21] V. Girault and P. A. Raviart. *Finite Element Methods for Navier-Stokes Equations—Theory and Algorithms.* Springer Verlag, 1986.

[22] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems.* Springer series in Computational Physics, 1984.

[23] R. Glowinski and J. Périaux. *Numerical Methods for Nonlinear Problems in Fluid Dynamics, Proceeding of the International Seminar on Scientific Super-Computer, Feb. 2–6.* 1987.

[24] G. H. Gobub and G. A. Meurant. *Résolution numérique des grands systèmes linéaires.* CEA-EDF-INRIA. Eyrolles, 1983.

[25] J. B. Goodman and R. J. Leveque. On the accuracy of a stable scheme for two dimensional conservation laws. *Soc. Ind. Appl. Math. Numer. Anal.*, 25:268–284, 1988.

[26] J. P. Grégoire. Assemblage dans la méthode des éléments finis. Bulletin de la Direction des Études et Recherches, série C 1, EDF, 1984.

[27] P. Grisvard. Problèmes aux limites dans les polygônes. mode d'emploi. Buletin de la direction des études et recherches, série c, EDF, 1986.

[28] C. Guilloppé and J. C. Saut. Global existence and one-dimensional nonlinear stability of shearing motions of viscoelastic fluids of oldroyd type. *à paraître dans : Modélisation Mathémetique et Analyse Numérique*, 1989.

[29] C. Guiloppé and J. C. Saut. Résultat d'existence pour les fluides viscoélastiques à loi de comportement de type différentiel. *Compte Rendu de l'Académie des Sciences de Paris*, t.305, série I:489–492, 1987.

[30] F. Hecht. A non-conforming P^1 basis with free divergence in $\mathbb{R}^3$. *RAIRO, série Analyse Numérique*, 1983.

[31] P. Hood and C. Taylor. A numerical solution of the navier-stokes equations using the finite element technique. *Comp. and Fluids*, 1:73–100, 1973.

[32] T. J. R. Hughes. A simple scheme for developing 'upwind' finite elements. *Int. J. for Num. Meth. in Eng.*, 12:1359–1365, 1978.

[33] J. Jaffre. Approximation par une méthode d'éléments finis mixtes d'une équation du type diffusion–convection linéaire stationnaire. Rapport de Recherche 367, LABORIA-INRIA, 1979.

[34] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Comp. Meth. in Appl. Mech. and Eng.*, 45:285–312, 1984.

[35] C. Johnson and J. Pitkäranta. An analysis of the discontinuous galerkin method for a scalar hyperbolic equation. *Math. of Comp.*, 46:1–26, 1986.

[36] D. D. Joseph, M. Renardy, and J. C. Saut. Hyperbolicity and change of type in the flow of viscoelastic fluids. *Arch. Ration. Mech. Anal.*, pages 213–251, 1985.

[37] B. W. Kernighan and D. M. Richtie. *Le langage C.* Masson, 1984.

[38] R. Keunings. *Fundamentals of Computer Modeling for Polymer Processing.* C.L Tucker III, Carl Hanser Verlag, to appear, 1987.

[39] M. E. Kim, R. A. Brown, and R. C. Armstrong. The roles of inertia and shear-thinning in the flow of an inelastic liquid though an axisymetric sudden contraction. *Journal of Non-Newtonian Fluid Mechanics*, 13:341–363, 1983.

[40] D. Leguillon and E. Sanchez-Palencia. *Computation of singular solution in elliptic problems and ellasticity.* Masson, Paris, 1987.

[41] P. Lesaint. *Sur la résolution des systèmes hyperboliques du premier ordre par des méthodes d'éléments finis.* thèse d'état, Université Pierre et Marie Curie, 1974.

[42] P. Lesaint and P. A. Raviart. *On finite element methods for solving the neutron transport equation.* Carl de Borr, Academic Press, 1974.

[43] J. F. Maitre. Eléments finis auto-adaptatifs (notes de cours). 1988.

[44] F. Nataf. *Approximation paraxiale pour les fluides incompressibles. Etude mathématique et numérique.* thèse de l'Ecole Polytechnique, Paris, 1989.

[45] U. Nävert. *A finite element method for convection–diffusion problems.* thesis, Chalmers Univ. of Technology, 1982.

[46] P. Nigon. *Une nouvelle classe de méthodes multigrilles pour les problèmes mixtes—Application à la résolution rapide du problème de Stokes.* thèse, Ecole Centrale de Lyon, 1984.

[47] J. G. Oldroyd. On the formultion of rheological equations of states. *Proc. Roy. Soc. London*, A 200:523–541, 1950.

[48] D. W. Peaceman and H. H. Rachford. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Ind. Appl. Math.*, 3:28–41, 1955.

[49] J. M. Piau, N. El Kissi, and B. Tremblay. Low reynolds number flow visualisation of linear and branched silicones upstream of orifices dies. *Journal of Non-Newtonian Fluid Mechanics*, 1988.

[50] O. Pironneau. *Méthode des éléments finis pour les fluides.* Masson, Paris, 1988.

[51] P. A. Raviart. *Méthodes d'éléments finis en mécanique des fluides.* CEA-EDF-INRIA. Eyrolles, 1981.

[52] M. Renardy. Existence of slow steady flows of viscoelastic fluids with differential constitutive equations. *Z. Angew. Math. u Mech*, 65:449–451, 1985.

[53] M. Renardy. Recent advances in the mathematical theory of steady flow of viscoelastic fluids. *J. of Non-Newtonian Fluids Mechanics*, 29:11–24, 1988.

[54] J. E. Roberts and J.M. Thomas. Mixed and hybrid methods. Rapport de Recherche 737, INRIA, 1987, à paraître comme un chapitre dans Handbock of Numerical Analysis, vol.3 : Finite Element Methods, P. G. Ciarlet and J. L. Lions editors, North Holland, Amsterdam.

[55] R. Stenberg. Some new families of finite elements for the stokes equations. Rapport de Recherche 995, INRIA, 1988.

[56] H. Wacker. *Continuation methods.* H. Wacker, Academic Press, 1978.

[57] J. E. Welch, F. H. Harlow, J. P. Shannon, and B. J. Daly. The mac method. Repport, Los Alamos, New Mexico LA 34 25, Los Alamos Scientific Laboratory, 1965.

[58] S. F. Wormon and M. W. Hafez. Implicit conservative sheme for the euler equations. *AIAA journal 24*, 2:215–223, 1986.

List of Figures