Analysis of mixed finite element methods in mechanics

Daniela Capatina

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

Habilitation à Diriger des Recherches

Spécialité :
Mathématiques Appliquées et Applications des Mathématiques
présentée par

Daniela CAPATINA née PAPAGHIUC

Analyse de méthodes mixtes d'éléments finis en mécanique

Soutenue le 2 novembre 2011 devant le jury composé de :

M.	Mohamed	AMARA	professeur	Universite de Pau et des Pays de l'Adour, Pau
M.	Roland	BECKER	professeur	Universite de Pau et des Pays de l'Adour, Pau
Mme	Christine	BERNARDI	DR CNRS	Université Pierre et Marie Curie, Paris
M.	Ramón	CODINA	professeur	Universitat Politècnica de Catalunya, Barcelona
M.	Jean-Claude	NÉDÉLEC	professeur	Universite de Rennes 1, Rennes
M.	Jean-Marie	THOMAS	professeur	Universite de Pau et des Pays de l'Adour, Pau

Après avis des rapporteurs :

Mme	Christine	BERNARDI	DR CNRS	Université Pierre et Marie Curie, Paris
M.	Ramón	CODINA	professeur	Universitat Politècnica de Catalunya, Barcelona
M.	Endre	SÜLI	professeur	University of Oxford, Oxford

A la mémoire de mon père ...

Remerciements

Je suis très sensible à l'honneur que m'ont fait Christine Bernardi, Ramón Codina et Endre Süli en acceptant de rapporter sur mon travail. Je tiens à les remercier chaleureusement pour l'intérêt qu'ils ont porté à mon manuscrit et pour le temps qu'ils m'ont accordé.

Je suis très reconnaissante à Jean-Claude Nédélec d'avoir accepté de présider le jury et à mon ancien directeur de thèse, Jean-Marie Thomas, d'avoir accepté d'y participer. La présence de mes collaborateurs Mohamed Amara et Roland Becker dans le jury a été pour moi un réel plaisir.

A tous les membres du jury, j'exprime ici mes vifs remerciements.

Je tiens à adresser aussi mes remerciements à toutes les personnes avec qui j'ai été amenée à travailler ces années.

Je dois beaucoup à Mohamed Amara. Il m'a donné, le premier, l'occasion de co-encadrer des thèses et d'avoir des collaborations industrielles. Il m'a aussi aidée à élargir mes thèmes de recherche après la thèse, en me proposant de nouveaux sujets comme par exemple les milieux poreux. Je lui suis très reconnaissante pour ses conseils précieux et pour le travail qu'on a accompli ensemble, ainsi que pour son soutien pendant les années passées à la direction du Laboratoire.

Mes remerciements très chaleureux vont à Roland Becker, qui m'a fait profiter de ses connaissances scientifiques et m'a fait découvrir ainsi beaucoup de choses, concernant par exemple les méthodes stabilisées ou l'adaptation de maillage. Par l'ambiance de travail très stimulante qu'il a instaurée au sein de l'équipe Concha et par les objectifs scientifiques ambitieux qu'il s'est toujours fixé, il a gagné mon admiration. Roland, je te remercie pour toutes nos discussions (mathématiques et amicales) si enrichissantes et j'éspère qu'à force de "positiver", on va finir par montrer la positivité !

David Trujillo est un collaborateur proche mais aussi mon collègue de bureau et mon ami depuis de nombreuses années. David, merci du fond du coeur pour les innombrables fois où tu m'as dépannée grâce à tes vastes connaissances informatiques (mais pas seulement ...), pour la bonne humeur avec laquelle tu m'as supportée toutes ces années, pour ta disponibilité constante à mon égard et non en dernier lieu, pour la super ambiance qui a toujours régné dans notre bureau.

Je suis reconnaissante aussi à Peppino Terpollili de m'avoir permis de découvrir de plus près le milieu industriel, grâce au projet MoTher avec l'entreprise Total.

Je remercie chaleureusement Didier Graebling de m'avoir fait partager ses connaissances sur les fluides non-newtoniens, avec beaucoup de patience et d'humour. Vive les polymères !

Merci également à Robert Luce pour le travail très intéressant que nous avons commencé ensemble, nos séances de travail représentent des moments privilégiés dans ma semaine.

Je ne pourrais pas oublier mes anciens thésards Amna, Bertrand, Layal et Julie : ça a été une expérience très enrichissante et très agréable de travailler avec chacun d'entre vous. Je suis fière de vos réussites et je vous souhaite à tous beaucoup de succès pour la suite.

J'aimerais également remercier tous les membres de l'équipe-projet Concha au sein de laquelle j'ai effectué mes travaux de recherche ces dernières années. La solidarité et le vrai esprit d'èquipe dont nous faisons preuve comptent beaucoup à mes yeux. Merci encore à Roland, David, Robert, Didier, Eric et aux "jeunes" pour toutes les choses que j'ai apprises à leurs côtés, mais aussi pour leur amitié et pour les éclats de rire journaliers. Et Eric, cesse de m'appeler Conchita, s'il te plaît !

Le Laboratoire de Mathématiques Appliquées de l'Université de Pau m'a offert de très bonnes conditions de travail, je tiens donc à remercier ses directeurs successifs et tout le personnel administratif qui m'a toujours facilité les tâches. Je remercie chaleureusement mes collègues, dont beaucoup sont devenus de vrais amis au fil des années, d'avoir contribué à rendre non seulement productive mais très agréable cette période. Je joins à ces remerciements INRIA Bordeaux Sud-Ouest pour les diverses opportunités dont j'ai pu bénéficier.

Je souhaite aussi remercier mes amis, d'ici ou d'ailleurs (France, Roumanie, Etats-Unis, Canada, Italie ...) pour tous les bons moments passés ensemble, pour les encouragements et le réconfort qu'ils m'ont apportés. Je remercie plus particulièrement Noëlle, qui a été à mes côtés depuis mes premiers jours à Pau et avec qui j'ai partagé les fous rires comme les peines, les délicieux moments de papotage comme le stress du travail, la varicelle des enfants comme les recettes de cuisine !

Je voudrais rendre hommage à mes parents, je ne serais pas là où j'en suis aujourd'hui sans eux. Maman, merci d'avoir toujours été à mon écoute avec beaucoup de tact et de discrétion, de m'avoir aidée en toutes circonstances et d'avoir créé autour de toi une atmosphère optimiste. Papa, c'est grâce à toi que j'ai choisi ce métier : tu m'as transmis le goût pour les mathématiques, et de plus tu m'as montré l'exemple d'un chercheur reconnu et respecté par ses pairs. Tu n'es plus là aujourd'hui et je tiens à te dire combien tu me manques, même si tu seras toujours présent dans mon coeur.

Un grand merci à mon frère Nucu et à ma grand-mère adorée de m'avoir toujours encouragée et de savoir si bien comment m'extraire de mes préoccupations professionnelles ! Je remercie aussi toute ma belle-famille pour son soutien et mes trois petits neveux pour leurs rires joyeux lors des (trop rares ...) réunions de famille.

Enfin, je voudrais remercier mon époux et mon fils pour l'amour et la compréhension qu'ils m'ont toujours témoignés, et pour tous les weekends sacrifiés en faveur de mon habilitation.

Narcis, merci pour ta confiance et ton soutien inconditionnel (depuis 20 ans déjà ...), pour ta bonne humeur et ton esprit *zen*, pour ton sens de la dérision quant à l'importance des puissances de ε et de h, pour toutes les tâches ménagères que tu as prises en charge ces derniers temps (et j'espère que ça va continuer !) et pour tellement d'autres choses encore !

Quant à toi, Codrin, merci tout simplement d'exister et de m'apporter du bonheur chaque jour.

Contents

INTRODUCTION

Introduction (in French)

Ce mémoire résume mes principaux travaux de recherche en tant que Maître de Conférences au Laboratoire de Mathématiques Appliquées de l'Université de Pau et membre de l'équipe-projet INRIA Concha. Ils se situent dans le domaine de l'*Analyse Numérique des Equations aux Dérivées Partielles* et ils s'articulent autour de la modélisation, la discrétisation, l'analyse et la simulation numériques de différentes applications issues de la mécanique des milieux continus et des milieux poreux.

Un fil conducteur de mes activités de recherche est l'utilisation et l'étude des méthodes d'éléments finis et des formulations mixtes, appliquées aux problèmes linéaires et non-linéaires. Depuis les travaux de Brezzi [**45**] et Thomas [**162**] dans les années '70, l'analyse numérique des formulations mixtes a été au coeur de nombreuses applications, allant de la mécanique des fluides à l'éléctromagnétisme et en passant par l'élasticité linéaire.

Le traitement des cas réalistes constitue un autre aspect non négligeable de mes travaux ; à ce titre, j'ai été amenée à considérer des équations complexes (notamment en ingénierie pétrolière ou en mécanique des fluides non-newtoniens), des conditions de bord non-standard, des modèles multi-dimensionnels (en hydrodynamique fluviale) ainsi que des problèmes couplés (couplage thermo-mécanique, couplage de modèles, couplage fluide - milieu poreux).

Du point de vue de l'analyse numérique, quelques mots clés pour mes travaux de recherche sont : éléments finis (conformes, non-conformes, mixtes, Galerkin discontinus), méthodes stabilisées, analyse des formulations mixtes, estimations d'erreur, estimateurs d'erreur *a posteriori*, adaptation de modèle et de maillage, schémas robustes par rapport aux paramètres, verrouillage numérique, schémas positifs, algorithmique, implémentation et validation de code, développement dans la librairie C++ Concha[1].

Le développement d'une librairie en C++ dédiée à la mécanique des fluides est l'objectif majeur de l'équipe-projet Concha. Les points forts de cette librairie sont sa grande modularité associée à des méthodes numériques innovantes et performantes. Les domaines d'application visés vont de la turbulence aux écoulements de liquides polymères, en passant par les écoulements compressibles et le transfert de chaleur. La version actuelle de la librairie est composée de modules suivants : maillage, éléments finis, adaptation, calcul parallèle, modèles physiques, résolution et post-traitement. Le langage python est employé comme langage de script.

En ce qui concerne les domaines d'application, je me suis intéressée ces dernières années à la mécanique des fluides newtoniens (équations de Stokes et de Navier-Stokes, écoulements fluviaux à surface libre) et non-newtoniens (modèles de polymères de Giesekus et d'Oldroyd-B), mais aussi aux milieux poreux (équations de Darcy-Forchheimer avec transfert de chaleur, modèle multi-phasique de réservoir pétrolier) et au couplage fluide - milieu poreux. J'ai eu l'occasion d'étudier des écoulements incompressibles et compressibles. J'ai également pu considérer certains problèmes en mécanique des solides, aussi bien pendant ma thèse de doctorat (problème de transmission raide, élasticité presque incompressible) que par la suite (structures minces).

[1] *http://sites.google.com/site/conchapau*

Les publications auxquelles j'ai contribué sont placées à la fin de cette introduction et sont référencées dans le texte par des lettres (A, B et C) suivies de numéros, tandis que les références bibliographiques sont placées à la fin du document et sont indiquées par des numéros.

Je ne détaille pas ici les publications [A13], [A14], [A15], [C16] et [C17] issues de ma thèse de doctorat. Elle a été encadrée par Jean-Marie Thomas et portait sur la prévention du phénomène de verrouillage numérique, plus particulièrement pour le problème de transmission raide. Une brève description des résultats obtenus est donnée dans mon curriculum vitae.

J'ai choisi d'organiser ce mémoire en quatre chapitres, chacun avec ses propres notations, suivant le domaine des applications traitées. Le premier chapitre est dédié à l'élasticité linéaire, plus précisément aux modèles de plaques minces. J'ai regroupé dans le deuxième chapitre les applications en mécanique des fluides newtoniens, tandis que le troisième chapitre est dédié aux milieux poreux et au couplage avec un fluide newtonien. Le dernier chapitre traite des écoulements de fluides non-newtoniens. Pour finir, je présente quelques perspectives de recherche à court et moyen terme.

Dans la suite, je décris succinctement la contribution apportée dans chacun des chapitres, tout en présentant un bref état de l'art dans le domaine respectif, les motivations et le cadre dans lequel le travail a été effectué. Je donne ensuite la liste de mes publications et je précise quelques notations utilisées dans le document.

1. Applications en élasticité linéaire

Le choix des conditions de bord constitue une question de premier plan dans beaucoup de problèmes aux limites issus de la mécanique des milieux continus, car il a une influence non-négligeable sur le comportement de la solution. La prise en compte des conditions aux limites non-standard conduit souvent à de nouvelles formulations faibles, qui nécessitent le développement de schémas numériques adéquats. Cette thématique se retrouve dans plusieurs de mes travaux, voir par exemple le chapitre 2 pour ce qui concerne la mécanique des fluides. Quant au chapitre 1, il est consacré entièrement au traitement des conditions de bord physiques dans deux applications spécifiques en élasticité linéaire.

À la suite de ma thèse et dans la continuité de mes préoccupations sur les problèmes de structures minces, souvent concernés par le phénomène de verrouillage numérique, je me suis intéressée à deux modèles de plaques minces en flexion, les modèles de Kirchhoff-Love et de Reissner-Mindlin. Le travail réalisé est de nature théorique et a fait partie de la thèse de doctorat de Amna Chatti, que Mohamed Amara m'a proposé de co-encadrer sur la deuxième partie de sa thèse. Les résultats obtenus ont été publiés dans [A1], [A2] et repris dans [C1].

Modèles de plaques minces

Pour fixer les idées, soit $\Omega \subset \mathbb{R}^2$ la surface moyenne de la plaque et Γ sa frontière latérale, décomposée en trois parties disjointes Γ_0 , Γ_1 et Γ_2. Le cadre mécanique retenu est celui de l'élasticité linéaire, homogène et isotrope. Les inconnues du modèle de Kirchhoff-Love sont le déplacement transverse u et le moment de flexion $\underline{\sigma}$, qui est un tenseur symétrique d'ordre 2. Le modèle de Reissner-Mindlin a une inconnue supplémentaire, le vecteur rotation de la normale unitaire à la surface moyenne ; ses inconnues dépendent de la demi-épaisseur de la plaque ε et sont notées par u^ε, $\underline{\sigma}^\varepsilon$ et $\mathbf{r}^\varepsilon$.

La littérature sur l'approximation des modèles de plaques minces est tellement impressionnante qu'il est impossible de citer ici toutes les contributions. Néanmoins, la très grande majorité des travaux traitent le cas des plaques encastrées. Des conditions de Dirichlet et de Neumann sont imposées sur u dans le modèle de Kirchhoff-Love, tandis que dans le modèle de Reissner-Mindlin on impose des conditions de Dirichlet sur u^ε et $\mathbf{r}^\varepsilon$. En général, on élimine le moment de flexion des équations.

Nous considérons, en suivant Destuynder et Salaün [**72**], que la plaque est encastrée sur Γ_0, simplement fixée sur Γ_1 et que Γ_2 représente sa frontière libre, ce qui se traduit par les conditions

$$\begin{cases} u = 0, & \partial_n u = 0 & \text{sur } \Gamma_0 \\ u = 0, & \underline{\sigma}\mathbf{n}\cdot\mathbf{n} = 0 & \text{sur } \Gamma_1 \\ \underline{\sigma}\mathbf{n}\cdot\mathbf{n} = 0, & \partial_t(\underline{\sigma}\mathbf{n}\cdot\mathbf{t}) + \mathrm{div}\underline{\sigma}\cdot\mathbf{n} = 0 & \text{sur } \Gamma_2 \end{cases}$$

pour le modèle de Kirchhoff-Love, respectivement

$$\begin{cases} u^\varepsilon = 0, & \mathbf{r}^\varepsilon = 0 & & \text{sur } \Gamma_0 \\ u^\varepsilon = 0, & \mathbf{r}^\varepsilon\cdot\mathbf{t} = 0, & \underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{n} = 0 & \text{sur } \Gamma_1 \\ \mathbf{r}^\varepsilon\cdot\mathbf{t} = \partial_t u^\varepsilon, & \underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{n} = 0, & \partial_t(\underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{t}) + \mathrm{div}\underline{\sigma}^\varepsilon\cdot\mathbf{n} = 0 & \text{sur } \Gamma_2 \end{cases}$$

pour le modèle de Reissner-Mindlin. Ainsi, il n'est plus envisageable d'éliminer une inconnue du problème. De plus, le tenseur moment de flexion représente souvent la quantité d'intérêt pratique pour les ingénieurs, d'où l'importance de l'approcher soigneusement.

Pour chacun des modèles, nous avons proposé une formulation mixte dont l'inconnue principale est le moment de flexion et dont les multiplicateurs de Lagrange sont définis sur une partie du bord. On peut ensuite récupérer le déplacement et la rotation. Néanmoins, le tenseur et les fonctions-test associées doivent satisfaire la contrainte $\mathrm{D}(\underline{\tau}) = \mathrm{div}\,\mathrm{div}\underline{\tau} = 0$. Afin d'éviter sa discrétisation, l'idée est de décomposer ces tenseurs symétriques en appliquant deux fois le lemme de Tartar [**161**]. Pour le modèle de Kirchhoff-Love, la symmetrie est imposée dans l'espace tandis que pour le modèle de Reissner-Mindlin, elle est dualisée à l'aide d'un multiplicateur de Lagrange. Une idée similaire est utilisée par Destuynder et Salaün [**72**], mais ils appliquent une seule fois le lemme de Tartar et obtiennent un problème complètement différent.

J'ai obtenu ainsi une formulation mixte équivalente et bien-posée pour chacun des problèmes, dont l'inconnue principale n'est pas une variable physique. En revanche, les espaces employés sont des espaces de Sobolev classiques, tels que $H^1(\Omega)$ et $H^{1/2}(\Gamma)$. L'approximation numérique est faite par des éléments finis conformes classiques, pour lesquels j'ai pu montrer une condition *inf-sup*, uniformément par rapport au paramètre de discrétisation. Grâce au théorème de Babuška-Brezzi (voir par exemple [**47**]), on en déduit que les deux méthodes d'éléments finis proposées sont inconditionnellement convergentes et donnent un ordre de convergence optimal $O(h)$ lorsque la solution exacte est suffisamment régulière.

En ce qui concerne l'approximation des variables physiques, on récupère à l'aide de formules explicites locales le moment de flexion dans $\underline{L}^2(\Omega)$ et la rotation dans $H(curl;\Omega)$. On a également une approximation de $\mathrm{D}(\underline{\sigma})$ dans $H^{-1}(\Omega)$ pour le modèle de Kirchhoff-Love, respectivement de $\varepsilon\,\mathrm{div}\underline{\sigma}^\varepsilon$ dans $\mathbf{L}^2(\Omega)$ pour le modèle de Reissner-Mindlin. Le déplacement transverse s'obtient à l'aide d'un *post-process* simple (résolution d'un problème de Laplace par éléments finis P_1-continus).

Une suite intéressante à ces travaux serait de proposer une autre discrétisation, surtout pour le modèle de Reissner-Mindlin où l'on a approché une variable de $H^1(\Omega)$ par des éléments finis P_2-continus. Ce choix m'a permis d'établir une condition *inf-sup* uniforme à la fois par rapport à h et ε, tout en travaillant avec des approximations conformes. On pourrait envisager d'utiliser une méthode stabilisée, moins chère, et de faire son analyse *a priori* et *a posteriori*. En ce qui concerne les deux méthodes proposées dans [A1] et [A2], on peut facilement écrire des estimateurs d'erreur *a posteriori* de type résiduel, qui conduisent à une analyse *a posteriori* standard.

Phénomène de verrouillage numérique

Un autre point clef de la méthode proposée est que le problème discret de Reissner-Mindlin ne souffre pas de verrouillage numérique. Pour une présentation générale de ce phénomène, je renvoie par exemple à l'article [**16**], ou à ma thèse de doctorat. En effet, ce modèle de plaque mince dépend de manière singulière du petit paramètre ε qui caractérise l'épaisseur de la plaque. La difficulté majeure lors de sa discrétisation consiste à trouver une méthode d'éléments finis qui

soit uniformément convergente par rapport à ε (et à h, bien sûr), donc une approximation qui ne se détériore pas lorsque ε devient de plus en plus petit.

Il existe dans la littérature plusieurs discrétisations sans verrouillage du modèle de Reissner-Mindlin. Elles ont été proposées pour des plaques encastrées et en général, ne sont pas applicables pour des conditions aux limites complexes. Une présentation exhaustive des résultats existants et d'une bibliographie sur ce modèle (au moment de mes travaux) est donnée dans [**82**].

L'approche la plus utilisée pour de telles problèmes de perturbation singulière consiste à dualiser la contrainte imposée dans le cas limite $\varepsilon = 0$, ce qui conduit à une formulation mixte dont l'opérateur peut s'écrire (voir par exemple [**47**] et ses références) :

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & -\varepsilon^2\mathcal{C} \end{pmatrix}. \tag{0.0.1}$$

Dans notre cas, il s'agit de la contrainte $\mathbf{r} = \nabla u$ du modèle limite de Kirchhoff-Love. Il est en général difficile de trouver des espaces d'éléments finis *inf-sup* stables, qui ne soient pas trop coûteux du point de vue de l'implémentation. Pour pallier cet inconvenient, plusieurs solutions existent : on peut modifier certains opérateurs *via* des techniques d'intégration réduite ou ajouter un terme de stabilisation ou utiliser des éléments finis non-conformes (éventuellement enrichis avec des fonctions-bulles) comme dans [**10**]. Une autre solution consiste à écrire une formulation équivalente du problème, à l'aide du théorème de décomposition de Helmholtz et de deux inconnues supplémentaires (voir par exemple [**47**] pour une description détaillée). Enfin, on peut aussi employer des méthodes p ou $h\,p$ (cf. [**160**]), qui ont la réputation d'être bien adaptées pour des problèmes concernés par le phénomène de verrouillage, ou une méthode de moindres carrés comme dans [**43**], ou encore, plus récemment, une méthode de Galerkin discontinue [**9**].

Notre opérateur mixte du problème de Reissner-Mindlin est différent, puisqu'il s'écrit :

$$\begin{pmatrix} A+\varepsilon^2 A_0 & B & C \\ B^T & O & O \\ C^T & O & O \end{pmatrix}$$

avec A et B les mêmes que pour le modèle limite de Kirchhoff-Love, A_0 qui prend en compte l'inconnue supplémentaire du modèle de Reissner-Mindlin et C qui tient compte de la symétrie du tenseur moment de flexion. Cet opérateur n'est donc pas typique dans l'analyse des problèmes de perturbation singulière. Je l'ai déjà employé avec succès dans mes travaux de thèse, notamment pour le problème de transmission raide, pour montrer que l'approximation par éléments finis mixtes de Raviart-Thomas est sans verrouillage numérique.

2. Applications en mécanique des fluides newtoniens

Le chapitre 2 regroupe plusieurs sujets que j'ai abordés en mécanique des fluides numérique. Le point commun des applications présentées est qu'elles sont toutes décrites par les équations de Stokes ou de Navier-Stokes incompressibles, néanmoins différents aspects ont été traités (conditions aux limites non-standard, écoulements à surface libre, modélisation multi-dimensionnelle, stabilisation de Galerkin discontinue). Des résultats théoriques et numériques sont présentés.

Equations de Navier-Stokes avec conditions aux limites non-standard

Je me suis intéressée aux équations bi- et tri-dimensionnelles de Navier-Stokes munies des conditions aux limites non-standard. Ces travaux sont présentés dans la Section 2.1 et ont été publiés dans [A7], [A4] et [C5] ; les rapports [**4**] et [**5**] contiennent, quant à eux, des résultats techniques (estimation d'erreur dans $L^4(\Omega)$ pour la vitesse, conditions non-homogènes, cas 3D, résultats numériques supplémentaires) qui n'ont pas été inclus dans les publications, par souci de brièveté. Ce travail a été réalisé en collaboration avec Mohamed Amara et David Trujillo.

Les conditions aux limites considérées sont celles introduites dans [**64**], à savoir dans le cas 2D homogène :

$$\begin{cases} \mathbf{u}\cdot\mathbf{n}=0, & \mathbf{u}\cdot\mathbf{t}=0 & \text{sur } \Gamma_1 \\ \mathbf{u}\cdot\mathbf{t}=0, & p=0 & \text{sur } \Gamma_2 \\ \mathbf{u}\cdot\mathbf{n}=0, & \omega=0 & \text{sur } \Gamma_3, \end{cases}$$

où les bords $\Gamma_1, \Gamma_2, \Gamma_3$ forment une partition de $\partial\Omega$ et où $\mathbf{u}$ représente la vitesse du fluide, $\omega = curl\mathbf{u}$ la vorticité et $p = p_c + \frac{1}{2}\mathbf{u}\cdot\mathbf{u}$ la pression dynamique, avec p_c la pression cinématique. Dans le cas des équations de Stokes, la condition aux limites sur Γ_2 porte directement sur la pression cinématique. Les conditions aux limites en 3D s'obtiennent en remplaçant $\mathbf{u}\cdot\mathbf{t}$ par $\mathbf{u}\times\mathbf{n}$ et en posant $\boldsymbol{\omega} = \mathrm{curl}\mathbf{u}$.

Les équations 2D de Stokes munies de ces conditions aux limites ont été étudiées dans [**77**] et [**6**], et celles de Navier-Stokes dans [**64**].

Dans [**6**], les auteurs proposent une formulation vitesse-vorticité-pression du problème de Stokes dans les espaces $H(\mathrm{div}, curl; \Omega) \times L^2(\Omega) \times L^2(\Omega)$, discrétisée par des éléments finis conformes P_1-continus, respectivement P_0. La condition *inf-sup* discrète est alors évidente, tandis que la coercivité uniforme est obtenue en stabilisant les sauts à travers les arêtes de la vorticité et de la pression. Nous avons montré que l'on peut simplifier le terme de stabilisation et prendre en compte uniquement le saut de la pression et nous avons généralisé ensuite l'approche de [**6**] au cas non-linéaire.

Dans les deux autres références, les auteurs utilisent des formulations variationnelles ainsi que des espaces éléments finis complètement différents des nôtres. Ainsi, dans [**64**] ils calculent d'abord la vitesse dans l'espace $\mathbf{H}^1(\Omega)$ et récupèrent ensuite la pression à l'aide de sa dérivée normale sur Γ_3, tandis que dans [**77**] les auteurs introduisent une formulation à trois champs des équations de Stokes et cherchent la vitesse dans $H(\mathrm{div}, \Omega)$, la vorticité dans $H^1(\Omega)$ et la pression dans $L^2(\Omega)$. Le schéma numérique obtenu est plus coûteux et sa stabilité n'est garantie, dans le cas des conditions aux limites générales, que si l'on ajoute une stabilisation.

D'autres travaux sur l'approximation des équations de Stokes et Navier-Stokes munies des conditions aux limites non-standard utilisent la méthode spectrale [**30**], [**13**]. Les conditions aux limites traitées portent sur la composante normale de la vitesse et sur la composante tangentielle de la vorticité. Une autre discrétisation des équations de Navier-Stokes 3D, avec une frontière Dirichlet en plus des conditions précédentes, a été récemment étudiée dans [**34**].

En ce qui concerne ma contribution, j'ai d'abord montré, en vue de l'analyse du problème de Navier-Stokes, que l'opérateur de Stokes discret avec une donnée dans $\mathbf{L}^{4/3}(\Omega)$ satisfait des conditions de stabilité et de consistance, uniformément par rapport au paramètre de discrétisation. Une écriture équivalente du terme de convection $(\mathbf{u}\cdot\nabla)\mathbf{u}$ permet ensuite d réécrire l'opérateur de Navier-Stokes comme une composée entre l'opérateur de Stokes et un opérateur non-linéaire $\mathcal{G} : L^2(\Omega) \times L^2(\Omega) \times \mathbf{L}^4(\Omega) \to \mathbf{L}^{4/3}(\Omega)$, défini par $\mathcal{G}(\omega, p, \mathbf{u}) = \omega\mathbf{u}^\perp$ en 2D, respectivement $\mathcal{G}(\boldsymbol{\omega}, p, \mathbf{u}) = \boldsymbol{\omega}\times\mathbf{u}$ en 3D. Les aspects non-linéaires sont traités à l'aide d'une variante du théorème des fonctions implicites, cf. par exemple [**50**], [**150**]. Grâce aux propriétés précédentes de l'opérateur de Stokes, j'ai pu montrer que le problème discret de Navier-Stokes admet une unique solution dans un voisinage de la solution exacte, ainsi que des estimations d'erreur *a priori* et *a posteriori*. D'une part, on en déduit la convergence inconditionnelle de la méthode et, pour des solutions exactes régulières, l'ordre optimal de convergence $O(h)$. D'autre part, on obtient un estimateur d'erreur *a posteriori* de type résiduel. J'ai aussi établi dans [**4**], en utilisant un argument de dualité de type Aubin-Nitsche pour le problème non-linéaire stabilisé, un ordre de convergence amélioré pour la vitesse en norme $\mathbf{L}^4(\Omega)$, à savoir $O(h^{3/2})$ en 2D et $O(h^{5/4})$ en 3D.

En perspective, il pourrait être intéressant de traiter le problème de Navier-Stokes muni des conditions aux limites précédentes à l'aide d'autres formulations continues et discrètes. Je pense

en particulier aux méthodes d'éléments finis stabilisées, largement développées et utilisées ces dernières années pour la formulation classique en vitesse-pression. A ce titre, on peut citer la méthode bien connue SUPG de Brooks et Hughes [**51**], la stabilisation par arêtes de Burman et Hansbo [**53**] qui généralise à l'ordre élevé celle de Brezzi et Pitkäranta [**49**], ou encore les méthodes de stabilisation par projection de Codina [**62**], Codina et Blasco [**63**] ou de Becker et Braack [**23**].

Modélisation multi-dimensionnelle en hydrodynamique fluviale

La modélisation et simulation numérique de l'hydrodynamique fluviale est au coeur de nombreuses applications environnementales (inondations, transport de polluants, phénomènes de sédimentation etc.). En outre, de nombreux projets de recherche en biologie marine et ressources aquatiques nécessitent une connaissance précise des milieux fluviaux.

Les domaines de calcul sont de taille importante, pouvant remonter de l'embouchure du fleuve sur plusieurs dizaines de kilomètres, et le système à résoudre est complexe (équations de Navier-Stokes à surface libre). De ce fait, une simulation précise de tout le domaine est très coûteuse. Il est donc intéressant de disposer de modèles plus simples, 1D ou 2D, qui pourront être implémentés dans des régions adéquates du fleuve et couplés ensuite afin d'obtenir une solution numérique globale satisfaisante. L'idéal serait de disposer d'un outil de simulation de l'hydrodynamique fluviale avec détermination automatique des zones 1D, 2D et 3D lors du couplage des modèles.

En collaboration avec Mohamed Amara et David Trujillo, nous nous sommes intéressés d'une part, à la dérivation et l'approximation numériques des modèles hydrodynamiques multi-dimensionnels et d'autre part, au couplage de ces modèles 1D et 2D *via* des estimateurs *a posteriori*. Ma contribution a été d'ordre théorique ; les différents codes ont été développés par David Trujillo. L'application au fleuve Adour a fait l'objet d'un projet européen LITHEAU avec IFREMER[2], auquel nous avons participé.

Les premières versions de modèles hydrodynamiques, linéaires à chaque pas de temps car obtenus suite à la semi-discrétisation de la dérivée totale par la méthode des caractéristiques, sont considérées dans [A3], [C2], [C3], [C6].

Par la suite, j'ai obtenu et étudié des modèles non-linéaires, qui sont de plus hiérarchiques. Les résultats correspondants se trouvent dans les articles [B2], [B3], [B4] qui vont être soumis très prochainement et dans [C8], et ils sont détaillés dans la Section 2.2 de ce mémoire.

Enfin, une alternative intéressante à l'utilisation d'un modèle 3D, surtout au niveau de l'estuaire du fleuve où un modèle 2D est souvent insuffisant, consiste à coupler les modèles 2D horizontal et 2D vertical précédents. Un tel modèle 2.5D est décrit dans [C9] et étudié numériquement dans la thèse d'Agnès Petrau [**147**] financée par IFREMER.

Le problème physique de départ est décrit par les équations 3D instationnaires de Navier-Stokes dans un domaine à surface libre.

Lorsque le rapport entre les échelles verticale et horizontale est petit, plusieurs modèles simplifiés de type Saint Venant, appelés aussi *shallow water*, sont utilisés en hydraulique [**95**]. Ces modèles sont obtenus après intégration des équations 3D sur une colonne d'eau en 2D, respectivement sur une section transversale en 1D, avec hypothèse hydrostatique et approximation de Boussinesq. La fermeture du système intégré est réalisée *via* la modélisation d'un terme de friction par une formule empirique (Manning-Strickler, Chézy etc.). La validité expérimentale et la robustesse reconnues du système de Saint-Venant, ainsi que la grande quantité de méthodes numériques efficaces développées, en font le modèle le plus utilisé en mécanique des fluides à surface libre.

Plusieurs travaux sont dédiés à leur dérivation et justification, basées sur une analyse asymptotique. On peut citer les travaux de Gerbeau et Perthame [**89**], qui établissent rigoureusement

[2]Institut Français de Recherche pour l'Exploitation de la Mer

un modèle 1D de Saint-Venant en y incluant frottement et viscosité, dans le cas d'une géométrie et des conditions de bords simples. Cette étude a été étendue au cas 2D avec une topographie présentant de petites variations dans [**84**] et [**127**].

D'autres approches pour éviter la résolution 3D existent, comme par exemple les équations de Saint-Venant multi-couches [**12**]. Une autre étude sur les écoulements gravitaires en eaux peu profondes, sans restriction sur la topographie, a été réalisée dans [**35**]. Par ailleurs, une alternative consiste en l"utilisation des modèles de Navier-Stokes simplifiés. A titre d'exemple, un modèle quasi-3D a été obtenu dans [**134**] en intégrant suivant la verticale uniquement l'équation de continuité et en écrivant le système sous une forme 2D + 1D, grâce au choix de la discrétisation verticale. Le domaine est divisé en plusieurs couches (d'épaisseur fixe dans [**134**] ou définie en fonction de la topographie dans [**70**]) et la même approximation éléments finis de la vitesse horizontale est employée dans chaque couche. Ces modèles sont plus coûteux mais plus riches que les modèles 2D de type Saint-Venant, puisqu'ils permettent de calculer aussi une vitesse verticale ; ils sont surtout employés dans la modélisation des océans ou des lacs.

Notre approche a été de s'appuyer sur le cadre variationnel pour d'une part, obtenir des modèles hydrodynamiques sans se soucier de la fermeture du système approché et d'autre part, établir des estimations d'erreur entre le modèle initial et ses diverses approximations. En ce qui concerne le problème physique 3D, notons l'originalité des conditions aux limites considérées, en particulier sur la surface libre où l'on impose la pression et la force du vent.

La discrétisation en temps choisie découple l'équation de la surface libre des équations de mouvement et d'incompressibilité. Le schéma d'Euler implicite est utilisé pour chacun des systèmes conduisant ainsi, une fois le domaine du fluide déterminé, à un problème non-linéaire en vitesse et pression muni de conditions de bord non-standard.

J'ai proposé et étudié une formulation faible de ce dernier dans un espace de type $H(\mathrm{div}, \mathrm{curl})$, à partir de laquelle plusieurs modèles hydrodynamiques ont été dérivés en tant qu'approximations conformes sur des espaces adaptés. Plus précisément, nous avons obtenu un modèle 1D et un autre 2D vertical écrits en coordonnées curvilignes sur la courbe médiane, respectivement la surface longitudinale du fleuve, ainsi qu'un modèle 2D horizontal écrit sur la surface libre du fleuve. Ils fournissent tous des approximations tridimensionnelles de la vitesse et de la pression, tout en tenant compte de la géométrie du fleuve (courbure, largeur et bathymétrie variables).

J'ai montré que ces problèmes mixtes non-linéaires semi-discrétisés sont bien posés et j'ai établi des estimations d'erreur *a priori* et *a posteriori* entre le problème de départ et les modèles approchés, dans des normes pondérées par la hauteur d'eau ou par la largeur du fleuve. Pour chacun des modèles, j'ai proposé et analysé ensuite une approximation par éléments finis. Afin d'appliquer directement au cas discret les résultats obtenus dans le cas continu, nous avons choisi des espaces conformes et *inf-sup* stables, uniformément par rapport à la discrétisation en temps et en espace. D'autres approximations (non-conformes ou stabilisées) pourraient être envisagées, en adaptant alors l'analyse d'erreur.

J'ai pu justifier, à chaque pas de temps, des estimateurs d'erreur *a posteriori* entre le problème 3D et un modèle hydrodynamique générique discret, qui indiquent le domaine de validité du modèle d'un point de vue qualitatif et qui peuvent être utilisés pour le couplage adaptatif de ces modèles. Un des points forts de notre approche est que l'on a pu choisir les sous-espaces de projection de telle sorte que l'on obtienne une hiérarchie de modèles, le modèle 1D étant une approximation conforme des deux modèles 2D. En plus de fournir un cadre unifié pour l'analyse d'erreur, ceci facilite le couplage puisque les conditions de transmission entre les différents modèles sont implicitement contenues dans les formulations.

Le code développé est utilisé actuellement par IFREMER et emploie la vraie bathymétrie de l'Adour. Afin de valider numériquement les nouveaux modèles hydrodynamiques, des comparaisons avec les modèles 1D et 2D de Saint Venant ainsi que des simulations réalistes avec

des données de bathymétrie, débits et coefficients de marée fournies par IFREMER, ont été effectuées. Le code donne une bonne approximation de la vitesse du courant, et ce dans des configurations délicates comme l'écoulement autour d'une île ou près d'un méandre.

En perspective, plusieurs aspects de ce travail pourront être approfondis, comme par exemple la prise en compte de l'équation de la surface libre et de l'évolution en temps dans les estimateurs d'erreur *a posteriori*, ou étendus, comme le choix des espaces de discrétisation et l'analyse d'erreur dans un cadre plus large. De nouveaux aspects pourront être rajoutés, tels que l'adaptation de modèles par rapport à une fonctionnelle. Enfin, une modélisation plus précise, prenant en compte les variations de température et l'écoulement biphasique eau douce - eau salée au niveau de l'estuaire, pourrait également être envisagée.

Méthode de Galerkin discontinue pour les équations de Stokes

Mes activités au sein de l'EPI Concha m'ont amenée à m'intéresser aux méthodes de Galerkin discontinues (dG), pour lesquelles il y a eu un grand intérêt ces dernières années à cause de leur grande flexibilité et facilité d'implémentation, et ce malgré leur coût relativement élevé. La littérature sur les méthodes dG est devenue tellement impressionnante qu'il serait impossible de citer toutes les contributions; une présentation très récente des méthodes dG peut être trouvée dans le livre de Di Pietro et Ern [**74**]. Pour une approche unifiée dans le cadre elliptique je renvoie à[**8**] ; pour le cas hyperbolique, voir par exemple [**48**]. Enfin, pour ce qui concerne les méchanismes de stabilisation, voir[**46**].

J'ai plus particulièrement étudié les équations de Stokes. La Section 2.3 est consacrée à une méthode de Galerkin discontinue avec une stabilisation différente du terme visqueux. Les résultats, obtenus en collaboration avec Roland Becker et Julie Joie, sont publiés dans [A11] et [C14] et sont également inclus dans la thèse de doctorat de Julie Joie [**107**] que j'ai co-encadrée. Des détails supplémentaires sont donnés dans le rapport technique [**24**].

Notre approche suit la méthode de pénalisation intérieure (IP) symétrique, introduite par Wheeler [**169**] et Arnold [**7**] pour des problèmes elliptiques et étendue aux équations de Stokes et de Navier-Stokes par Girault, Rivière et Wheeler [**93**]. Sur chaque triangle, la vitesse appartient à l'espace $\boldsymbol{P}_k$ et la pression à P_{k-1} $(1 \leq k \leq 3)$; une stabilisation basée sur les sauts des vitesses à travers les arêtes est rajoutée afin de garantir existence, unicité et estimations d'erreur optimales pour la formulation discrète.

À la place, nous proposons de pénaliser la projection L^2-orthogonale sur P_{k-1} des sauts, en suivant une idée similaire de Hansbo et Larson [**100**] pour le problème d'élasticité linéaire. Ceci présente deux avantages considérables, décrits ci-dessous.

Tout d'abord, nous avons montré que la solution dG tend, lorsque le paramètre de stabilisation γ tend vers l'infini, vers la solution $\boldsymbol{P}_k^{\mathrm{noncf}} \times P_{k-1}^{\mathrm{disc}}$ non-conforme du problème de Stokes. Il est bien connu que ces espaces sont *inf-sup* stables, d'après [**66**] pour $k = 1$, [**88**] pour $k = 2$ et [**67**] pour $k = 3$; ce n'est pas le cas des espaces $\boldsymbol{P}_k^{\mathrm{conf}} \times P_{k-1}^{\mathrm{disc}}$ obtenus en passant à la limite dans la méthode IP classique. De plus, la constante de la condition *inf-sup* de notre schéma est indépendante de γ, tandis que celle de la méthode IP est en $O(1/\sqrt{\gamma})$. Ainsi, contrairement à [**93**], notre méthode est robuste pour des paramètres de stabilisation γ grands, phénomène que l'on a mis aussi en évidence numériquement.

Par une technique d'hybridisation, j'ai pu établir que l'ordre de convergence de la méthode dG vers la méthode non-conforme est en $O(1/\gamma)$. Pour ce faire, à l'aide d'un multiplicateur de Lagrange on a d'abord écrit le problème dG sous la forme matricielle (0.0.1) avec $\varepsilon^2 = 1/\gamma$ et on a appliqué ensuite la variante pour les formulations mixtes d'un résultat classique en optimisation. Ceci se rapproche d'ailleurs des techniques évoquées auparavant pour pallier le verrouillage numérique. Le point délicat mais indispensable ici est la condition *inf-sup* pour l'opérateur $\mathcal{B}$.

La preuve de cette dernière s'adapte également au cas où l'on utilise la stabilisation IP classique pour un problème elliptique de second ordre ; ainsi, on en déduit que la méthode dG avec pénalisation intérieure converge en $O(1/\gamma)$ vers la méthode conforme, tandis que la méthode dG avec notre stabilisation réduite converge vers la méthode non-conforme. Il est intéressant de noter que ce résultat reste valable lorsque l'on prend en compte dans la stabilisation uniquement les arêtes situées sur un bord Dirichlet, ce qui revient à utiliser la méthode de Nitsche (avec une approximation conforme ou non-conforme) pour traiter les conditions de bord de manière faible. Ces résultats semblent être nouveaux.

Le deuxième point positif de notre choix de stabilisation est lié à l'analyse d'erreur *a posteriori*. Dans le cadre elliptique, des estimateurs d'erreur *a posteriori* pour des méthodes de Galerkin discontinues ont été étudiés dans [**108**], [**123**] ; en ce qui concerne les méthodes *hp* adaptatives, on renvoie à [**101**].

Nous nous sommes intéressés à des estimateurs basés sur une reconstruction de flux dans $H(\mathrm{div},\Omega)$. Parmi les travaux récents sur ce sujet, on peut citer dans le cadre elliptique [**125**] pour une méthode d'éléments finis mixtes et [**78**] pour la méthode de Galerkin discontinue, ou encore [**110**] et [**166**] en lien avec des méthodes localement conservatives.

Nous avons ainsi pu construire un tenseur de $H(\mathrm{div},\Omega)$ localement conservatif et appartenant à l'espace d'éléments finis de Raviart-Thomas RT_{k-1}. Ceci nous a permis de définir un estimateur d'erreur *a posteriori* simple, qui semble nouveau pour l'approximation dG du problème de Stokes. Cet estimateur tend lui-aussi, lorsque $\gamma \to \infty$, vers un estimateur *a posteriori* pour l'approximation non-conforme, qui est équivalent dans le cas $k = 1$ à l'estimateur bien connu de [**69**]. Notons que dans [**78**] le flux appartient à un espace plus grand que le notre, à savoir RT_k.

D'autres approches pour la discrétisation dG des équations de Stokes et Navier-Stokes existent. On peut citer celles developpées par Bassi et Rebay [**22**] et par Cockburn et al. [**61**], qui introduisent le gradient de la vitesse comme troisième variable, discrétisée ensuite à l'aide des flux numériques différents. Nous avons étudié le lien entre la stabilisation usuelle de [**7**], celle de [**22**] et la notre.

Nous avons fait une analyse similaire pour la formulation en tenseur des vitesses de déformation $\underline{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T)$ du problème de Stokes, qui est équivalente à la formulation à trois champs et constitue, à ce titre, un point de départ intéressant pour traiter des fluides non-newtoniens. Dans ce cas, afin d'obtenir une inégalité de Korn discrète sur des espaces complètement discontinus, on rajoute un terme de stabilisation supplémentaire lorsque $k = 1$. Cette extension est présentée dans [C14] et n'est pas développée dans ce mémoire.

Les résultats théoriques ont été illustrés numériquement, des comparaisons avec la méthode de pénalisation intérieure ont aussi été effectuées. Le code est intégré dans la librairie Concha.

3. Applications en milieux poreux. Couplage avec milieu fluide

J'ai eu l'occasion de travailler dans le domaine de la modélisation et simulation numériques en milieux poreux, dans le cadre d'une collaboration avec l'entreprise Total. Mes travaux ont été effectués dans le cadre du projet MoTher (*Modélisation des Thermométries*), dont les responsables ont été Peppino Terpollili à Total et Mohamed Amara à l'Université de Pau.

Suite à l'émergence de nouvelles technologies d'acquisition avec l'apparition des fibres optiques, on dispose désormais des mesures de température dans les puits de pétrole. Le but final de MoTher est d'interpréter ces mesures afin de pouvoir déterminer la température initiale du réservoir et la répartition des débits par couche. Pour résoudre ces problèmes inverses, il est d'abord nécessaire de développer une approche directe, permettant de décrire d'un point de vue thermo-dynamique l'écoulement d'un fluide compressible dans un réservoir (milieu poreux) et un puits pétrolier (milieu fluide).

De plus, la modélisation themodynamique du modèle direct doit être suffisamment fine pour permettre d'interpréter par la suite de très petites variations de température (confirmées ultérieurement dans [**118**]). Il existe de nombreux codes dédiés à la simulation de puits et de réservoir pétroliers mais la plupart sont soit isothermes (RUBIS, ECLIPSE, REVEAL), soit négligent certains phénomènes physiques (STAR) qui jouent un rôle important lorsque l'on veut résoudre de tels problèmes.

Le projet s'est construit principalement autour de deux thèses, que j'ai co-encadrées avec Mohamed Amara, sur les sujets suivants :

- modélisation réservoir et puits avec prise en compte de la thermométrie (thèse de Bertrand Denel, financement TOTAL, 2001-2004) ;
- couplage de modèles réservoir et puits. Modèle multi-phasique de réservoir (thèse de Layal Lizaik, bourse CIFRE [3], 2005-2008).

Ces travaux ont donné lieu aux publications [A5], [A6], [A8], [A9], [A10], [B1] et aux actes de conférences [C4] et [C7]. Le cas mono-phasique en géométrie 2D axisymmétrique est détaillé dans le chapitre 3 ; nous avons aussi considéré des écoulements 3D multi-phasiques dans le réservoir, décrits brièvement dans la dernière section du chapitre 3.

Modèle 2D axisymétrique de réservoir mono-phasique

Nous avons développé un modèle 2D axisymétrique de réservoir pétrolier qui couple une équation de Darcy-Forchheimer avec un bilan d'énergie exhaustif [**131**] ; ce dernier prend en compte, outre les effets diffusifs et convectifs, les effets de compressibilité et de dissipation visqueuse. Cette équation d'énergie quantifie ainsi le refroidissement ou l'échauffement du fluide produit avant qu'il entre dans le puits, et se distingue des modèles classiques qui supposent que le fluide produit entre dans le puits à température géothermique.

Du fait des vitesses importantes de filtration aux abords du puits, un terme quadratique en vitesse est rajouté dans l'équation classique de Darcy afin de tenir compte des pertes d'énergie cinétique. Dans notre travail, ce terme a été linéarisé suite à la discrétisation en temps ; quant à l'équation non-linéaire de Darcy-Forchheimer, elle a été récemment étudiée dans [**94**], où une formulation primale basée sur une approximation P_0 de la vitesse et P_1-nonconforme de la pression a été proposée.

La discrétisation en temps conduit à un système linéarisé à chaque pas de temps dont les inconnues sont la pression, le flux massique, la température et le flux de chaleur, la densité étant mise à jour à l'aide de l'équation d'état. J'ai montré l'existence et l'unicité de la solution du problème semi-discrétisé. Pour ce faire, j'ai d'abord négligé les termes de convection et j'ai obtenu une formulation mixte à laquelle j'ai appliqué une extension [**158**] du théorème de Babuška-Brezzi. J'ai ensuite étudié le problème complet à l'aide de l'alternative de Fredholm, sous une hypothèse de pas de temps suffisamment petit.

Nous avons ensuite proposé une discrétisation par éléments finis, dans laquelle les flux sont approchés par des éléments de Raviart-Thomas et les termes convectifs présents dans l'équation d'énergie sont traités par un schéma décentré à la Lesaint et Raviart [**119**]. J'ai montré que le problème discret admet une solution unique et une analyse *a posteriori* de l'erreur a été menée.

Le code a été validé par Bertrand Denel par de nombreux tests numériques, incluant des cas-test réels et des comparaisons avec des mesures et avec un autre logiciel. La valorisation industrielle du code a fait l'objet d'un contrat de confidentialité entre TOTAL, l'Université de Pau et la société KAPPA ENGINEERING[4]. Cette dernière a également participé au projet MOTHER, son rôle a été de transcrire les codes développés à l'UPPA aux standards de l'industrie.

[3] Convention Industrielle de Formation par la Recherche

[4] *www.kappaeng.com*

Modèle 1.5D axisymétrique de puits pétrolier

Nous avons ensuite considéré un modèle axisymétrique de puits vertical, basé sur des équations de type Navier-Stokes compressibles couplées avec la thermique. Afin de tenir compte de la direction privilégiée de l'écoulement, mais aussi pour diminuer le coût de calcul et pour éviter les instabilités numériques liées à un maillage 2D, un modèle 1.5D basé sur la dépendance explicite des inconnues par rapport à la variable radiale a été proposé.

Le problème non-linéaire obtenu après la discrétisation en temps est résolu par une méthode de point fixe autour de la densité. On calcule d'abord le flux massique $\mathbf{G} = \rho\mathbf{v}$ à partir de l'équation de conservation de la masse, *via* une formulation de Petrov-Galerkin et on récupère ensuite la vitesse et la pression, puis la température et le flux de chaleur, par le biais des formulations mixtes. La discrétisation en espace est effectuée sur un maillage spécial (une seule cellule dans la direction radiale) et utilise des éléments finis de Raviart-Thomas pour les flux, des constantes par morceaux pour la pression et la température, et des éléments Q_1- continus pour la vitesse. J'ai montré que ce choix d'espaces conduit à des formulations discrètes bien posées. Le code a ensuite été validé numériquement.

Un autre modèle de puits a aussi été considéré. Il diffère par le calcul de la vitesse radiale u_r, obtenue cette fois-ci à l'aide d'une projection du flux massique, tout en négligeant l'équation du moment correspondante. Ce modèle est un peu plus simple mais moins adapté pour le couplage avec le réservoir. En perspective, il serait intéressant de traiter le cas plus général de puits dévié.

Couplage de modèles de réservoir et de puits

Nous nous sommes intéressés enfin au couplage des modèles de puits et de réservoir précédemment introduits. Rappelons qu'il s'agit d'un couplage entre les équations de Darcy-Forchheimer et celles compressibles de Navier-Stokes, avec prise en compte des aspects thermiques et implicitement, d'une densité variable dans les deux milieux. De plus, les modèles à coupler sont de dimensions différentes (2D et 1.5D) et n'ont pas les mêmes inconnues. Ainsi, notre problème couplé fluide - milieu poreux est donc différent de ceux usuellement étudiés dans la littérature (cf. par exemple [**128**], [**114**], [**157**], [**92**], [**76**]).

Le couplage est réalisé en imposant des conditions de transmission adéquates à l'interface entre le puits et le réservoir. Ces conditions sont dualisées à l'aide des multiplicateurs de Lagrange et conduisent, à chaque pas de temps, à une formulation mixte dont l'opérateur s'écrit :

$$\begin{pmatrix} \mathcal{A} & \mathcal{I} \\ \mathcal{J} & 0 \end{pmatrix}, \text{ avec } \mathcal{A} = \begin{pmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{B}^T & -\mathtt{C} \end{pmatrix}.$$

Il s'agit d'un opérateur mixte non-standard. En effet, les formes bilineaires $\mathtt{A}(\cdot,\cdot)$ et $\mathtt{C}(\cdot,\cdot)$ sont ni symétriques, ni définies positives et de plus, les inconnues et les fonctions-test n'appartiennent pas au même espace. J'ai montré que l'opérateur $\mathcal{A}$ est injectif à l'aide du théorème de Babuška. À ce stade, je n'ai pas pu établir directement une deuxième condition *inf-sup* assurant l'existence de la solution, qui a été montrée plus loin à l'aide d'une méthode de Galerkin basée sur l'approximation éléments finis du problème.

Afin de tenir compte des données collectées sur le terrain et de pouvoir imposer un débit à la sortie du puits, nous avons opté pour une résolution globale du problème couplé. La discrétisation spatiale utilise les éléments finis précédemment employés dans chacun des modèles, les multiplicateurs de Lagrange à l'interface étant approchés par des constantes par morceaux. J'ai montré que le problème couplé discret admet une unique solution. Les estimations uniformes que j'ai pu établir pour l'opérateur discret ont permis ensuite d'en déduire l'existence d'une solution au problème continu. Le code obtenu a été validé par des tests numériques avec des données réelles, *via* des comparaisons avec les codes séparés puits et réservoir ainsi que par raffinement de maillage.

En ce qui concerne l'analyse mathématique du problème couplé, il y a plusieurs questions intéressantes qui pourront être traitées par la suite. On pourrait, dans un premier temps, considérer un modèle simplifié qui contient néanmoins les difficultés principales du modèle complet. Je pense par exemple au couplage Darcy-Stokes avec des densités et perméabilités variables, plus éventuellement un bilan d'énergie. Alors quelques sujets de réflexion sont : l'analyse de la discrétisation en temps, étude du problème non-linéaire à chaque pas de temps, estimations d'erreur, prise en compte de maillages non-conformes à l'interface etc.

Enfin, en perspective, une toute autre piste d'investigation pour le problème couplé serait l'utilisation des méthodes stabilisées pour la discrétistion, dans le but d'avoir un cadre unifié pour l'analyse et l'implémentation (voir par exemple [**17**]).

Modèle de réservoir multi-phasique multi-composantes

Enfin, nous avons considéré le cas complexe d'un modèle de réservoir multi-phasique (de type *black-oil* généralisé) anisotherme. Nous avons proposé une équation d'énergie originale pour le modèle multi-compositionnel considéré qui tient compte, comme dans le cas mono-phasique, de l'effet Joule-Thomson et de la dissipation, ainsi que la thermodynamique correspondante. Ceci constitue une différence majeure avec la modélisation d'injection de vapeur où une telle précision n'est pas nécessaire.

Nous sommes ensuite passés à l'approximation numérique et à l'implémentation. Compte tenu de la complexité du système à résoudre, et aussi par souci de cohérence avec les logiciels pétroliers existants, nous nous sommes orientés vers une discrétisation par volumes finis [**81**]. L'approche choisie en concertation avec TOTAL a été d'intégrer dans le code GPRS (*General Purpose Reservoir Simulation*) [**55**] développé à l'Université de Stanford, notre équation d'énergie et le module thermo-dynamique correspondant. Il est important de préciser que l'objectif de TOTAL est d'intégrer ces aspects thermodynamiques, une fois le code académique validé, dans un simulateur commercial développé en collaboration avec KAPPA ENGINEERING. Le choix du GPRS, qui utilise le même modèle de Coats et le même schéma volumes finis que le logiciel commercial, a donc été naturel.

Les premiers tests numériques effectués et la comparaison avec le code isotherme confirment le bon comportement du simulateur anisotherme développé. Sa validation complète devrait passer par d'autres cas-tests. En perspective, le modèle multi-phasique de réservoir devrait être couplé avec le modèle de puits.

4. Applications aux fluides non-newtoniens

Enfin, je me suis intéressée ces dernières années à la simulation numérique des écoulements complexes de fluides viscoélastiques.

Ces liquides ont en commun un comportement intermédiaire entre un liquide visqueux et un solide élastique, se traduisant entre autres par un effet mémoire et par l'apparition de contraintes normales aux plans de cisaillement. De plus, ils sont non-newtoniens, c'est à dire que leur viscosité est une fonction non-linéaire de la vitesse de déformation. Ces phénomènes ne sont pas prévus par les équations de Navier-Stokes.

Les liquides viscoélastiques sont omniprésents dans notre société : nous les rencontrons aussi bien dans l'agro-alimentaire ou la cosmétique que dans les biens de consommation courants tels que l'emballage, l'automobile, l'éléctroménager etc. Ils sont également présents en nous (sang, liquide synovial).

A l'heure actuelle mes travaux concernent l'approximation numérique des liquides viscoélastiques particuliers, les polymères. Ils ont été réalisés au sein de l'EPI Concha et sont présentés dans le chapitre 4, organisé en deux sections.

Simulation numérique de liquides polymères

Ce travail a été fait dans un cadre pluridisciplinaire, en collaboration avec Roland Becker et Didier Graebling, physicien au Laboratoire de Physico-Chimie des Polymères de l'Université de Pau. Nous avons co-encadré la thèse **[107]** de Julie Joie, thèse financée par le Conseil Régional d'Aquitaine (2007-2010). Les résultats obtenus sont publiés dans [A12], [C11], [C12] et [C13].

L'objectif est de développer un code de calcul pour la simulation des écoulements réalistes de liquides polymères.

Les principales difficultés sont dues aux propriétés intrinsèques de ces liquides et au couplage interne entre la viscoélasticité du liquide et l'écoulement, couplage quantifié par le nombre de Weissenberg. Ce paramètre physique est défini par $We = \lambda\dot{\gamma}$ avec λ le temps de relaxation et $\dot{\gamma}$ la vitesse de déformation. Pour la plupart des algorithmes, les nombres de Weissenberg élevés rencontrés en pratique soulèvent de sérieux problèmes de convergence, cf. **[109]** ou **[168]**.

De plus, le comportement rhéologique des liquides polymères est si complexe qu'il existe dans la littérature tout un ensemble d'équations constitutives pour le décrire, de manière plus ou moins réaliste.

Nous nous sommes concentrés sur les modèles de type différentiel et plus particulièrement sur le modèle non-linéaire de Giesekus, qui a l'avantage de reproduire convenablement les écoulements élongationnels, de cisaillement et mixtes. En outre, il ne requiert la connaissance que de deux paramètres caractéristiques du liquide facilement mesurables : sa viscosité et son temps de relaxation.

Neanmoins, d'autres modèles (Maxwell convecté, Oldroyd-B **[140]**, Phan-Thien Tanner **[148]**) ont également été considérés.

Notre but est de développer des schémas numériques stables et robustes pour des nombres de Weissenberg élevés. À ce jour, nous avons traité le cas bi-dimensionnel tout en négligeant les aspects thermiques; leur prise en compte ainsi que le passage en 3D font partie de mes perspectives.

Lors de la discrétisation par éléments finis de ces modèles non-linéaires à trois champs (vitesse, pression et tenseur des contraintes visqueuses), une attention particulière doit être portée aux choix des espaces d'approximation, mais aussi au traitement des termes convectifs.

Plusieurs méthodes mixtes bien posées existent dans la littérature, généralement basées sur une décomposition du tenseur des contraintes en une partie élastique et une partie visqueuse, conduisant à une formulation à quatre champs. Lorsque l'on y effectue un changement de variable, on obtient la méthode populaire EVSS (*Elastic Viscous Split Stress*) de **[153]** ou une variante cf. **[159]**. Sans changement de variable, on trouve la méthode DEVSS (*Discrete Elastic Viscous Split Stress*) de **[97]**.

Pour l'approximation des termes convectifs, deux approches sont principalement utilisées : une méthode de décentrage en amont SU ou SUPG (voir **[106]**, **[51]**) et une méthode de Galerkin discontinue, basée sur le schéma de Lesaint et Raviart **[119]**. Un avantage de la méthode dG, en plus de sa facilité d'implémentation, est que la condition de compatibilité entre l'espace discret de la vitesse et celui du tenseur des contraintes est facilement satisfaite. La première classe de méthodes a été initialement appliquée aux liquides viscoélastiques dans **[126]**, la deuxième dans **[87]**. Une présentation générale de ces diverses méthodes peut être trouvée dans **[15]** et **[142]**.

Nous avons choisi d'utiliser des éléments finis totalement discontinus pour approcher le tenseur des contraintes et des éléments finis non-conformes, de Crouzeix-Raviart **[66]** dans le cas triangulaire et de Rannacher-Turek **[154]** dans le cas quadrangulaire, pour la vitesse et la pression. Le terme convectif $\mathbf{u} \cdot \nabla \underline{\tau}$ est traité par un schéma décentré en amont à la Lesaint et Raviart. L'analyse de ces schémas numériques a été faite pour le problème de Stokes sous-jacent, et des vitesses de convergence optimales ont été obtenues dans les deux cas. Pour des maillages

en quadrilatères, on a montré qu'il est nécessaire d'ajouter un terme de régularisation afin d'avoir une méthode consistante ; le schéma se rapproche alors des méthodes EVSS et DEVSS citées auparavant.

Le code a été écrit dans la librairie C++ Concha et a été validé par le biais de différentes comparaisons : avec solution semi-analytique, avec des valeurs expérimentales et avec le code commercial Polyflow, reconnu dans le monde industriel pour sa fiabilité. Ces comparaisons nous ont permis d'illustrer le bon comportement du schéma mais aussi de montrer le réalisme du modèle de Giesekus par rapport à d'autres modèles rhéologiques. Nous avons pu montrer la supériorité de notre code par rapport à Polyflow, en termes de robustesse et de temps de calcul. Des nombres de Weissenberg supérieurs à 40, voire 70 ont été atteints pour des cas-tests de référence (canal, contractions 4:1 et 4:1:4, écoulement autour d'un cylindre) alors que Polyflow ne converge pas toujours.

Nous nous intéressons actuellement aux méthodes de résolution efficaces, afin de pouvoir utiliser des maillages fins (de l'ordre de 1 million d'éléments). Nous avons développé une méthode multi-grilles basée sur un lisseur de type Vanka. Ceci nous a permis de valider la méthode en comparant nos calculs de traînée aux résultats présents dans la littérature pour le modèle d'Oldroyd-B, dans le cas-test populaire d'écoulement autour d'un cylindre centré. Une validation similaire du modèle de Giesekus, qui passe par le développement d'une méthode multi-grilles adéquate, est en cours.

Schémas positifs pour une équation matricielle de type transport

Je travaille depuis peu sur ce sujet, en collaboration avec Roland Becker. Il est traité dans l'article [B5], qui va être soumis très prochainement, et il a été évoqué également dans [C13]. J'ai été amenée à m'y intéresser à la suite du travail sur la simulation numérique des polymères, dans le but d'expliquer le bon comportement du code, en particulier la meilleure robustesse pour des nombres de Weissenberg élevés du modèle numérique de Giesekus par rapport à d'autres modèles de polymères.

Les lois constitutives des fluides viscoélastiques peuvent être réécrites en termes du tenseur de conformation à la place du tenseur des contraintes. Il est connu que pour certains modèles, ce tenseur symétrique est défini positif, ce qui permet ensuite de montrer que l'énergie du système décroît (cf. par exemple **[124]** ou **[102]** pour le modèle de Oldroyd-B). Il est généralement accepté que les schémas numériques qui préservent la positivité en discret sont plus stables et permettent d'établir des estimations d'énergie en discret.

Le problème du nombre de Weissenberg élevé est souvent associé dans la littérature à la perte de positivité du tenseur de conformation au niveau discret. Il semble donc crucial pour la simulation numérique de pouvoir utiliser des schémas positifs. Pour obtenir de tels schémas, plusieurs approches ont été proposées ces dernières années.

Fattal et Kupferman ont proposé dans **[83]** de réécrire la loi constitutive en termes du logarithme du tenseur de conformation X, soit $\Psi = \ln X$. Pour ce faire, ils utilisent le fait que X est symétrique défini positif ainsi qu'une décomposition spécifique du tenseur gradient des vitesses, $\nabla \mathbf{u} = \Omega + S + NX^{-1}$ avec Ω et N anti-symétriques et S symétrique. Cette transformation à l'aide du logarithme introduit une non-linéarité supplémentaire au niveau du second membre ; aussi, elle rend non-linéaire même une équation de départ linéaire. On calcule ensuite une approximation Ψ_h de Ψ et on pose le tenseur de conformation discret $X_h = e^{\psi_h}$, qui est donc symétrique défini positif. Plusieurs travaux récents utilisent cette approche, comme par exemple **[104]** ou **[68]** mais à ce jour, il n'y a aucune analyse d'erreur pour cette méthode. Il serait intéressant d'étudier la méthode pour une équation plus simple comme l'équation de convection-diffusion-réaction, d'un point de vue théorique et numérique mais aussi par comparaison avec un schéma positif plus classique comme la méthode de Galerkin discontinue ; ce point fait partie de mes perspectives de recherche.

Afin de garantir la positivité du tenseur de conformation, une idée différente a été introduite par Lee et Xu dans [**116**], basée sur les similarités entre les lois constitutives de certains fluides viscoélastiques et les équations de Riccati. Ces dernières sont très utilisées en contrôle optimal et elles ont été largement étudiées, voir par exemple les livres [**113**], [**132**] pour une présentation générale des résultats existants ; un des premiers travaux sur leur approximation numérique a été fait par Nédélec [**136**].

Lee et Xu [**116**] ont interprété, à l'aide des dérivées de Lie, les lois constitutives comme des équations différentielles généralisées de Riccati, pour y appliquer ensuite des résultats connus pour ces équations. Leur approche repose de manière essentielle sur la méthode des caractéristiques employée pour la discrétisation. L'exemple typique considéré est le modèle instationnaire d'Oldroyd-B, pour lequel les auteurs pensent avoir établi la positivité de la solution. Un commentaire sur la validité de ce résultat est donné dans la sous-section 4.2.3.2 ; nous verrons qu'une condition sur le pas de temps ou sur le nombre de Weissenberg ne peut être évitée.

Les auteurs proposent également une autre méthode pour obtenir la positivité, en approchant la solution analytique du modèle d'Oldroyd-B sous forme intégrale. Néanmoins, cette idée paraît difficile à utiliser pour d'autres modèles tels que le modèle non-linéaire de Giesekus, pour lequel aucune solution analytique n'est connue.

Dans notre travail, nous avons adopté l'approche de [**116**] dans le but de l'appliquer au modèle de Giesekus et à une discrétisation de Galerkin discontinue. Pour ce faire, nous avons considéré le cadre plus général d'une équation matricielle stationnaire non-linéaire de type transport, qui inclut aussi bien la loi constitutive de Giesekus que celle d'Oldroyd-B.

En cohérence avec les travaux présentés dans la Section 4.1, nous l'avons discrétisée à l'aide du schéma de Lesaint et Raviart [**119**], adapté à un champs de vitesses discret donné. J'ai montré alors qu'une modification de la méthode de Newton conduit à une équation algébrique de Lyapunov, c'est à dire de Riccati sans terme quadratique (voir [**113**], [**132**]) sur chaque cellule du maillage, et j'ai pu conclure sous certaines hypothèses que les itérés de Newton convergent de manière monotone vers une solution définie positive.

Cette étude nous a permis de mieux comprendre les performances moindres du modèle d'Oldroyd-B par rapport au modèle de Giesekus. En effet, pour ce dernier les hypothèses se résument, grâce à la présence du terme non-linéaire, au choix de l'itéré initial, sans aucune condition sur le nombre de Weissenberg. Cependant, une condition de stabilité supplémentaire doit être satisfaite pour la loi constitutive d'Oldroyd-B, ce qui implique des limitations sur le nombre de Weissenberg. Cette condition, nécessaire pour une équation de Lyapunov mais pas pour une équation de Riccati, semble être négligée dans [**116**]. Enfin, notre étude est illustrée par des essais numériques.

Nos résultats théoriques sont certes incomplets et pourraient être améliorés, comme mentionné dans les perspectives. Néanmoins, il s'agit à notre connaissance du premier résultat de ce type sur le modèle de Giesekus, valable aussi bien dans le cas stationnaire qu'instationnaire, et ce indépendamment du pas de temps.

L'extension à d'autres schémas ainsi que l'application aux estimations d'énergie et à l'existence de solutions continue et discrète pour le modèle de Giesekus sont des travaux en cours, et à ce titre font partie de mes perspectives de recherche.

5. Liste de publications

Les articles de cette habilitation sont organisés en trois groupes et sont présentés, au sein de chaque groupe, par ordre chronologique. Les publications liées à ma thèse de doctorat apparaîssent à part.

Le premier groupe [A] contient les articles publiés dans des journaux internationaux avec comité de lecture. Le deuxième groupe [B] est formé d'articles qui vont être soumis très prochainement, et du rapport technique [B1] qui n'a pas été publié par ailleurs. D'autres rapports techniques qui complètent les publications des groupes [A] et [C] sont cités dans la bibliographie. Enfin, le troisième groupe [C] comprend des actes de congrès avec comité de lecture.

Articles dans des journaux avec comité de lecture

[A1] M. Amara, D. Capatina-Papaghiuc, A. Chatti : *Bending Moment Mixed Method for the Kirchhoff-Love Plate Model*, SIAM J. Num. Anal., vol. 40, n. 5, p. 1632-1649, 2002

[A2] M. Amara, D. Capatina-Papaghiuc, A. Chatti : *New Locking-Free Method for the Reissner-Mindlin Plate Model*, SIAM J. Num. Anal., vol. 40, n. 4, p. 1561-1582, 2002

[A3] M. Amara, D. Capatina-Papaghiuc, D. Trujillo : *Hydrodynamical modelling and multidimensional approximation of estuarian river flows*, Comput. Vis. Sci., vol. 6, n. 2-3, p. 39-46, 2004

[A4] M. Amara, D. Capatina-Papaghiuc, E. Chacon Vera, D. Trujillo : *Vorticity-velocity-pressure formulation for Navier-Stokes equations*, Comput. Vis. Sci., vol. 6, n. 2-3, p. 47-52, 2004

[A5] M. Amara, D. Capatina-Papaghiuc, B. Denel, P. Terpolilli : *Mixed Finite Element Approximation for a Coupled Petroleum Reservoir Model*, M2AN, vol. 39, n. 2, p. 349-376, 2005

[A6] M. Amara, D. Capatina-Papaghiuc, B. Denel, P. Terpolilli : *Numerical modelling of flow with heat transfer in petroleum reservoir*, Int. J. Numer. Method. Fluids, vol. 47, n. 8, p. 955-962, 2005

[A7] M. Amara, D. Capatina-Papaghiuc, D. Trujillo: *Stabilized finite element method for Navier-Stokes equations with physical boundary conditions*, Math. of Comp., vol. 76, n. 259, p. 1195-1217, 2007

[A8] M. Amara, D. Capatina, L. Lizaik : *Numerical coupling of 2.5D reservoir and 1.5D wellbore models in order to interpret thermometrics*, Int. J. Numer. Method. Fluids, vol. 56, n. 8, p. 1115-1122, 2008

[A9] M. Amara, D. Capatina, L. Lizaik : *Coupling of Darcy-Forchheimer and compressible Navier-Stokes equations with heat transfer*, SIAM J. Sci. Comp., vol. 31, n. 2, p. 1470-1499, 2009

[A10] D. Capatina, L. Lizaik, P. Terpolilli : *Numerical modeling of multi-component multiphase flows in petroleum reservoirs with heat transfer*, Appl. Analysis, vol. 88, n. 10-11, p. 1509-1525, 2009

[A11] R. Becker, D. Capatina, J. Joie : *Connections between discontinuous Galerkin and nonconforming finite element methods for the Stokes equations*, Numerical Methods for Partial Differential Equations, 2011 *(published online DOI 10.1002/num.20671)*

[A12] R. Becker, D. Capatina, D. Graebling, J. Joie : *Nonconforming finite element approximation of the Giesekus model for polymer flows*, Computers and Fluids, vol. 46, p. 142-147, 2011

Articles issus de la thèse de doctorat

[A13] D. Capatina-Papaghiuc, J.-M. Thomas : *Non conforming finite element methods without numerical locking*, Numerische Mathematik, vol. 81, p. 163-186, 1998

[A14] D. Capatina-Papaghiuc, N. Raynaud : *Numerical Approximation of Stiff Transmission Problems by Mixed Finite Element Methods*, M2AN, vol. 32, n°5, p. 611-629, 1998

[A15] D. Capatina-Papaghiuc, J.-M. Thomas : *Approximations non conformes de problèmes dépendant d'un petit paramètre : résultat de convergence uniforme*, C. R. Acad. Sci., t. 325, Série I, p. 97-100, 1997

Rapports techniques et articles en préparation

[B1] M. Amara, D. Capatina, B. Denel : *Petroleum wellbore models,* Preprint LMA UPPA n°08, 34 p, *http://hal.archives-ouvertes.fr/docs/00/19/23/91/PDF/0708.pdf,* 2007

[B2] M. Amara, D. Capatina, D. Trujillo : *Variational multi-dimensional modeling of an estuarian river flow*, à soumettre, 2011

[B3] D. Capatina, A. Petrau, D. Trujillo : *Numerical approximation of new 2D vertical and 1D hydrodynamic models*, à soumettre, 2011

[B4] M. Amara, D. Capatina, D. Trujillo : *Derivation and numerical approximation of a 2D horizontal model*, Technical Report INRIA, 26p*, http://hal.inria.fr/inria-00342858,* à soumettre

[B5] R. Becker, D. Capatina : *Numerical analysis of a matrix-valued transport equation of Riccati type*, en préparation

Actes de conférences avec comité de lecture

[C1] M. Amara, D. Capatina-Papaghiuc, A. Chatti : *New Mixed Formulations for Thin Plate Models with Physical Boundary Conditions*, ECCOMAS, Barcelona, septembre 2000

[C2] M. Amara, D. Capatina-Papaghiuc, D. Trujillo : *Hydrodynamical modeling and numerical approximation of river flows*, Monografias del Seminario Matematico Garcia de Galdeano (7th International Conference Zaragoza-Pau on Applied Mathematics and Statistics), vol. 27, p. 41-48, 2003

[C3] M. Amara, D. Capatina-Papaghiuc, D. Trujillo: *Derivation, analysis and numerical approximation of new 2D and 1D hydrodynamical models*, ECCOMAS, Jyväskylä, 2004

[C4] M. Amara, D. Capatina-Papaghiuc, B. Denel, P. Terpollili : *Numerical modeling of temperature for flows in porous media and petroleum wellbore*, ECCOMAS, Jyväskylä, 2004

[C5] M. Amara, D. Capatina, D. Trujillo : *Numerical simulation of the 3D Navier-Stokes equations with non standard boundary conditions*, TAMTAM, Tunis, 2005

[C6] M. Amara, D. Capatina-Papaghiuc, D. Trujillo : *Variational approach for the hydrodynamical multidimensionnal modelling*, International Conference on Adaptive Modelling and Simulation ADMOS, Barcelona, 2005

[C7] M. Amara, D. Capatina-Papaghiuc, L. Lizaik : *Numerical coupling of petroleum wellbore and reservoir models with heat transfer*, ECCOMAS CFD, Egmond und Zee, 2006

[C8] M. Amara, D. Capatina, D. Trujillo : *Hierarchical Multiscale Model Adaption in Fluvial Hydrodynamics*, FLUCOME, Tallahassee (USA), 2007

[C9] M. Amara, D. Capatina, A. Petrau, D. Trujillo : *A Quasi 3D Estuarian River Flow Modeling*, ECCOMAS, Venise, 2008

[C10] R. Becker, D. Capatina : *Adaptive finite elements for viscoelastic flows*, ECCOMAS CFD, Lisbonne, 2010

[C11] R. Becker, D. Capatina, J. Joie : *Nonconforming finite element approximation for polymer flows at large Weissenberg numbers*, ECCOMAS CFD, Lisbonne, 2010

[C12] R. Becker, D. Capatina, D. Graebling, J. Joie *: Finite element discretization for the numerical simulation of polymer flows*, Monografias del Seminario Matematico Garcia de Galdeano (10th International Conference Zaragoza-Pau on Applied Mathematics and Statistics), vol. 35, p. 57 - 64, 2010

[C13] R. Becker, D. Capatina : *Finite element approximation of Giesekus model for polymer flows*, in Numerical Mathematics and Advanced Applications 2009 (Proceedings of 8th ENUMATH), Springer, p. 133-142, 2010

[C14] R. Becker, D. Capatina, J. Joie : *A new DG method for the Stokes problem with a priori and a posteriori error analysis*, in Numerical Mathematics and Advanced Applications 2009 (Proceedings of 8th ENUMATH), Springer, p. 143-152, 2010

[C15] N. Barrau, D. Capatina : *Numerical simulation of anisothermal flows of Newtonian fluids*, Monografias del Seminario Matematico Garcia de Galdeano (11th International Conference Zaragoza-Pau on Applied Mathematics and Statistics), vol. 35, p. 37-46, 2011 (to appear)

Actes issus de la thèse de doctorat

[C16] D. Capatina-Papaghiuc : *Robust Nonconforming Approximations of Parameter Dependent Problems*, 4th World Congress of Computational Mechanics WCCM (MS), Buenos Aires, Argentine, 29 juin–2 juillet 1998

[C17] D. Capatina-Papaghiuc : *Nonconforming methods and numerical locking. Applications*, 5èmes Journées de Mathématiques Appliquées Pau - Saragosse, Jaca, septembre 1997

6. Notation

We agree to write the vectors in bold letters, the second-order tensors in underlined letters and we employ the summation convention of Einstein; the product of two tensors will be denoted by $\underline{\tau} : \underline{\sigma} = \tau_{ij}\sigma_{ij}$. For any scalar function v, any 2D vector field $\mathbf{v} = (v_1, v_2)$ and any second order tensor $\underline{v} = (v_{ij})_{1\leq i,j\leq 2}$, we denote:

$$\mathrm{div}\mathbf{v} = \partial_1 v_1 + \partial_2 v_2, \qquad \mathrm{div}\underline{v} = \begin{pmatrix} \partial_1 v_{11} + \partial_2 v_{12} \\ \partial_1 v_{21} + \partial_2 v_{22} \end{pmatrix},$$

$$\mathrm{curl}v = (\partial_2 v, -\partial_1 v), \quad curl\mathbf{v} = \partial_1 v_2 - \partial_2 v_1, \quad \mathrm{curl}\mathbf{v} = \begin{pmatrix} \partial_2 v_1 & -\partial_1 v_1 \\ \partial_2 v_2 & -\partial_1 v_2 \end{pmatrix}$$

and we equally put:

$$\mathbf{v}^{\perp} = (-v_2, v_1), \quad \nabla\mathbf{v} = (\partial_i v_j)_{1\leq i,j\leq 2}, \quad \underline{\varepsilon}(\mathbf{v}) = \frac{1}{2}(\nabla\mathbf{v} + \nabla\mathbf{v}^T).$$

Similar notations are introduced in 3D, in particular for $\mathbf{v} = (v_1, v_2, v_3)$ we denote $\mathrm{curl}\mathbf{v} = \nabla \times \mathbf{v}$.

As usually, for a given domain ω of $\mathbb{R}^n$ we shall denote by $L^2(\omega)$ the space of square integrable functions for the Lebesgue measure on ω and we put:

$$\begin{aligned} H^1(\omega) &= \left\{u \in L^2(\omega);\ \nabla u \in \left(L^2(\omega)\right)^n\right\}, \\ H(\mathrm{div},\omega) &= \left\{\mathbf{u} \in \left(L^2(\omega)\right)^n;\ \mathrm{div}\mathbf{u} \in L^2(\omega)\right\} \end{aligned}$$

and, for $n = 2$ and $n = 3$ respectively

$$\begin{aligned} H(\mathrm{div}, curl;\omega) &= \left\{\mathbf{v} \in \mathbf{L}^2(\omega);\ \mathrm{div}\mathbf{v} \in L^2(\omega),\ curl\mathbf{v} \in L^2(\omega)\right\}, \\ H(\mathrm{div},\mathrm{curl};\omega) &= \left\{\mathbf{v} \in \mathbf{L}^2(\omega);\ \mathrm{div}\mathbf{v} \in L^2(\omega),\ \mathrm{curl}\mathbf{v} \in \mathbf{L}^2(\omega)\right\}. \end{aligned}$$

We agree to denote the vector spaces by bold letters, in particular for real numbers $p > 1$, $s > 0$ we denote $L^p(\Omega)^n$, $H^s(\Omega)^n$, $H(\mathrm{div},\omega)^n$ by $\mathbf{L}^p(\Omega)$, $\mathbf{H}^s(\Omega)$, $\mathbf{H}(\mathrm{div},\omega)$ respectively. For a given Hilbert space V, we also put $\underline{V} = \{\underline{\tau} = (\tau_{ij});\ \tau_{ij} \in V,\ 1 \leq i,j \leq 2\}$. The notation $\|\cdot\|_{\mathcal{L}(Y)}$ stands for the norm of a linear continuous operator of $\mathcal{L}(Y)$, where Y is a Banach space.

For a given boundary $\Gamma \subset \partial\omega$, we denote by $\langle\cdot,\cdot\rangle_\Gamma$ the duality product between $H_{00}^{1/2}(\Gamma)$ and its topological dual space $H^{-1/2}(\Gamma)$; we recall that $H_{00}^{1/2}(\Gamma)$ is the space of traces on Γ of

functions in $H^1(\omega)$ which vanish on $\partial\omega \setminus \Gamma$. We denote $\mathbf{n}$ and $\mathbf{t}$ the unit outward normal vector, respectively a unit tangent vector to the boundary Γ. We use the notation $\partial_t v = \nabla v \cdot \mathbf{t}$ and $\partial_n v = \nabla v \cdot \mathbf{n}$ for the tangential, respectively the normal derivative on the boundary of a scalar function v in 2D.

For a polygonal or polyhedral domain Ω, we agree to denote by $(\mathcal{T}_h)_{h>0}$ a regular family of triangulations, such that $\overline{\Omega} = \cup_{K\in\mathcal{T}_h} K$. For every cell K of $\mathcal{T}_h$, we denote by h_K its diameter and we define the discretization parameter $h = \max_{K\in\mathcal{T}_h} h_K$. For $n = 2$, we denote by $\mathcal{E}_h$, $\mathcal{E}_h^{int}$ and $\mathcal{E}_h^{\partial}$ the set of edges, of internal edges and of edges situated on $\Gamma = \partial\Omega$ respectively, and by h_e the length of the edge e. Similar notations are used for the faces of the triangulation in 3D. On every internal edge e such that $\{e\} = \partial T_1 \cap \partial T_2$, we define once for all the unit normal $\mathbf{n}_e$; for a given function φ with $\varphi_{/T_i} \in \mathcal{C}(T_i)$ $(1 \le i \le 2)$, we define on e: $\varphi^{\rm in}(\mathbf{x}) = \lim_{\varepsilon\to 0} \varphi(\mathbf{x} - \varepsilon\mathbf{n}_e)$, $\varphi^{\rm ex}(\mathbf{x}) = \lim_{\varepsilon\to 0} \varphi(\mathbf{x} + \varepsilon\mathbf{n}_e)$ as well as the jump $[\varphi] = \varphi^{\rm in} - \varphi^{\rm ex}$. If the edge belongs to $\partial\Omega$, then $\mathbf{n}_e$ is the outward normal $\mathbf{n}$ and the jump coincides with the trace. In addition, let x^- the negative part of a real number x defined by $x^- = \min\{0, x\}$ and let $x^+ := x - x^-$. For a given vector $\boldsymbol{\beta}$, we use the classical notations

$$\partial\Omega^- = \{x \in \partial\Omega;\ \boldsymbol{\beta}(x) \cdot \mathbf{n}(x) \le 0\}, \quad \partial\Omega^+ := \partial\Omega \setminus \partial\Omega^-.$$

For any $k \in \mathbb{N}$, we denote by P_k the space of polynomials of total degree $\le k$ and by π_k^ω the $L^2(\omega)$-orthogonal projection operator on P_k. We shall use the notation RT_k for the Raviart-Thomas finite element space.

We say that a matrix A is *positive* if it is symmetric and positive definite, and *nonnegative* if it is symmetric positive semidefinite, and we write $A > 0$ and $A \ge 0$ respectively. The notation $A \ge B$ for two matrices A, B means that $A - B \ge 0$. Let us denote by $\mathcal{M}$ the vector space of 2×2-matrices and let $L^p(\Omega, \mathcal{M})$ be the space of matrix-valued functions whose coefficients belong to $L^p(\Omega)$, $1 \le p \le \infty$. We consider similar notations for the subspace $\mathcal{M}_{\rm sym}$ of symmetric 2×2-matrices.

CHAPTER 1

APPLICATIONS IN LINEAR ELASTICITY

CHAPTER 1

Applications in linear elasticity

This chapter is devoted to two aplications in solid mechanics, more precisely in linear elasticity, developed after my PhD thesis. They concern the Kirchhoff-Love and the Reissner-Mindlin thin plate models endowed with non-standard boundary conditions, involving the stress tensor.

The results of this chapter were published in [A1] and [A2]. To summarize, we propose well-posed formulations of the two plate models which take well into account the prescribed boundary conditions. Their discretizations by conforming finite elements are shown to be unconditionally convergent and optimal whenever the exact solution is sufficiently smooth. Moreover, the approximation of the Reissner-Mindlin model is shown to be locking-free.

1.1. Two thin plate models

For technical reasons, we assume that $\Omega \subset \mathbb{R}^2$ is a connected polygonal domain without cuts and that its boundary Γ is decomposed in three disjoint parts, $\Gamma = \bar{\Gamma}_0 \cup \bar{\Gamma}_1 \cup \bar{\Gamma}_2$ with $\Gamma_0 \neq \emptyset$ and $\Gamma_0 \cup \Gamma_1$ connected. The material of the plate is supposed to be homogeneous and isotropic.

The equations describing the Kirchhoff-Love model are, cf. for instance **[72]**:

$$
\left\{
\begin{array}{ll}
\Delta^2 u = \mathrm{div}(\mathrm{div}\underline{\sigma}) = f & \text{in } \Omega \\
\sigma_{ij} = (1-\nu)\partial_{ij} u + \nu \Delta u \delta_{ij} & \text{in } \Omega \\
u = 0, \quad \partial_n u = 0 & \text{on } \Gamma_0 \\
u = 0, \quad \underline{\sigma}\mathbf{n}\cdot\mathbf{n} = 0 & \text{on } \Gamma_1 \\
\underline{\sigma}\mathbf{n}\cdot\mathbf{n} = 0, \quad \partial_t(\underline{\sigma}\mathbf{n}\cdot\mathbf{t}) + \mathrm{div}\underline{\sigma}\cdot\mathbf{n} = 0 & \text{on } \Gamma_2,
\end{array}
\right. \tag{1.1.1}
$$

where $\nu > 0$ is the Poisson coefficient and $f \in L^2(\Omega)$. The problem includes the case of a clamped plate (obtained for $\Gamma_1 = \Gamma_2 = \emptyset$ and modeled by a classical biharmonic problem), as well as the case of a simply supported plate (obtained when $\Gamma_0 = \Gamma_2 = \emptyset$).

The corresponding Reissner-Mindlin model is governed by the following equations:

$$
\left\{
\begin{array}{ll}
-\dfrac{1}{1-\nu}\mathrm{div}\underline{\sigma}^\varepsilon + \dfrac{1}{\varepsilon^2}(\mathbf{r}^\varepsilon - \nabla u^\varepsilon) = 0 & \text{in } \Omega \\
\dfrac{1-\nu}{\varepsilon^2}\mathrm{div}(\mathbf{r}^\varepsilon - \nabla u^\varepsilon) = f & \text{in } \Omega \\
\sigma_{ij}^\varepsilon = (1-\nu)\,\varepsilon_{ij}(\mathbf{r}^\varepsilon) + \nu(\mathrm{div}\mathbf{r}^\varepsilon)\delta_{ij} & \text{in } \Omega \\
u^\varepsilon = 0, \quad \mathbf{r}^\varepsilon = 0 & \text{on } \Gamma_0 \\
u^\varepsilon = 0, \quad \mathbf{r}^\varepsilon\cdot\mathbf{t} = 0, \quad \underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{n} = 0 & \text{on } \Gamma_1 \\
\mathbf{r}^\varepsilon\cdot\mathbf{t} = \partial_t u^\varepsilon, \quad \underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{n} = \partial_t(\underline{\sigma}^\varepsilon\mathbf{n}\cdot\mathbf{t}) + \mathrm{div}\underline{\sigma}^\varepsilon\cdot\mathbf{n} = 0 & \text{on } \Gamma_2,
\end{array}
\right. \tag{1.1.2}
$$

with the unknowns u^ε, $\mathbf{r}^\varepsilon$ and $\underline{\sigma}^\varepsilon$ and with ε a small parameter characterizing the plate's thickness.

The previous choice of boundary conditions ensures cf. Destuynder and Salaün **[72]** that as ε tends towards 0, one has:

$$
\begin{array}{ll}
u^\varepsilon \overset{H^1(\Omega)}{\longrightarrow} u, & \mathbf{r}^\varepsilon \overset{H^1(\Omega)}{\longrightarrow} \mathbf{r} = \nabla u, \\
\underline{\sigma}^\varepsilon \overset{L^2(\Omega)}{\longrightarrow} \underline{\sigma}, & \mathrm{div}\underline{\sigma}^\varepsilon \overset{H^{-1}(\Omega)}{\longrightarrow} \mathrm{div}\underline{\sigma}.
\end{array}
$$

In order to give a mathematical framework for the analysis of (1.1.1) and (1.1.2), we introduce the following Hilbert spaces:

$$V = \left\{ v \in H^1(\Omega);\ v = 0 \text{ on } \Gamma_0 \cup \Gamma_1 \right\},$$
$$\underline{X} = \left\{ \underline{\tau} = (\tau_{ij})_{1 \le i,j \le 2};\ \tau_{ij} \in L^2(\Omega),\ \mathrm{D}(\underline{\tau}) \in L^2(\Omega) \right\},$$
$$\underline{X}^\varepsilon = \left\{ \underline{\tau} \in \underline{X};\ \varepsilon\, \mathrm{div}\underline{\tau} \in \mathbf{L}^2(\Omega) \right\},$$

endowed with the norms

$$\|v\|_V = |v|_{1,\Omega},$$
$$\| \underline{\tau} \|_{\underline{X}} = (\| \underline{\tau} \|_{0,\Omega}^2 + \| \mathrm{D}(\underline{\tau}) \|_{0,\Omega}^2)^{1/2},$$
$$\| \underline{\tau} \|_{\underline{X}^\varepsilon} = (\| \underline{\tau} \|_{0,\Omega}^2 + \varepsilon^2 \| \mathrm{div}\underline{\tau} \|_{0,\Omega}^2 + \| \mathrm{D}(\underline{\tau}) \|_{0,\Omega}^2)^{1/2}$$

where $\mathrm{D}(\underline{\tau}) = \mathrm{div}(\mathrm{div}\underline{\tau}) = \partial_{ij}\tau_{ij}$.

Remark 1.1.1. The choice of the weighted norm on $\underline{X}^\varepsilon$ is justified by the fact that even in the case of a clamped plate with Ω convex, one only has the following regularity result cf. for instance [**47**]:

$$\mathbf{r}^\varepsilon \in \mathbf{H}^2(\Omega), \quad u^\varepsilon \in H^2(\Omega),$$
$$\|\mathbf{r}^\varepsilon\|_{2,\Omega} + \|u^\varepsilon\|_{2,\Omega} + \varepsilon \|\mathrm{div}\underline{\sigma}^\varepsilon\|_{1,\Omega} \le c \|f\|_{0,\Omega}.$$

The above estimate for $\mathrm{div}\underline{\sigma}^\varepsilon$ cannot be improved even for a smoother domain and a smoother loading f; $\mathrm{div}\underline{\sigma}^\varepsilon$ is not uniformly bounded in $\mathbf{H}^1(\Omega)$ because of boundary layers in the Reissner-Mindlin model (cf. for instance [**11**]).

For any $f \in L^2(\Omega)$, let

$$\underline{X}^f = \{\underline{\tau} \in \underline{X};\ \mathrm{D}(\underline{\tau}) = f\} \neq \emptyset, \qquad \underline{X}^{\varepsilon,f} = \{\underline{\tau} \in \underline{X}^\varepsilon;\ \mathrm{D}(\underline{\tau}) = f\} \neq \emptyset.$$

It is useful to introduce the boundary value problem

$$\left\{ \begin{array}{ll} \Delta\phi^f = f & \text{in } \Omega \\ \phi^f = 0 & \text{on } \Gamma_0 \cup \Gamma_1 \\ \partial_n \phi^f = 0 & \text{on } \Gamma_2, \end{array} \right. \tag{1.1.3}$$

whose unique solution ϕ^f belongs to V. One can then decompose the solutions $\underline{\sigma} \in \underline{X}^f$ and $\underline{\sigma}^\varepsilon \in \underline{X}^{\varepsilon,f}$ of (1.1.1) and (1.1.2) as follows:

$$\underline{\sigma} = \underline{\sigma}^0 + \phi^f \underline{I}, \qquad \underline{\sigma}^\varepsilon = \underline{\sigma}^{\varepsilon,0} + \phi^f \underline{I}$$

with $\underline{\sigma}^0 \in \underline{X}^0$ and $\underline{\sigma}^{\varepsilon,0} \in \underline{X}^{\varepsilon,0}$.

It has been proved in [A1] that $\underline{\mathcal{D}}(\overline{\Omega})$ is a dense subspace of $\underline{X}$ and that the trace operators

$$\gamma_0 \ : \ (\underline{\mathcal{D}}(\overline{\Omega}), \|\cdot\|_{\underline{X}}) \longrightarrow H^{-1/2}(\Gamma), \qquad \gamma_0(\underline{\tau}) = \underline{\tau}\mathbf{n} \cdot \mathbf{n},$$
$$\gamma_1 \ : \ (\underline{\mathcal{D}}(\overline{\Omega}), \|\cdot\|_{\underline{X}}) \longrightarrow H^{-3/2}(\Gamma), \qquad \gamma_1(\underline{\tau}) = \partial_t(\underline{\tau}\mathbf{n} \cdot \mathbf{t}) + \mathrm{div}\underline{\tau} \cdot \mathbf{n}$$

are linear and continuous, so they can be extended by continuity on the whole space $\underline{X}$. Moreover, the following Green formula holds for any $v \in H^2(\Omega)$ and any $\underline{\tau} \in \underline{X}$:

$$\int_\Omega \mathrm{D}(\underline{\tau}) v \mathrm{dx} = \int_\Omega \tau_{ij} \partial_{ij} v \mathrm{dx} - \langle \gamma_0(\underline{\tau}), \partial_n v \rangle_{-\frac{1}{2},\frac{1}{2},\Gamma} + \langle \gamma_1(\underline{\tau}), v \rangle_{-\frac{3}{2},\frac{3}{2},\Gamma}.$$

1.2. Bending moment formulations

For each of the previous plate models, a first variational formulation with the bending moment as main unknown was proposed, leading to a three-fields formulation for the Kirchhoff-Love model and a four-fields formulation for the Reissner-Mindlin one.

As regards the Kirchhoff-Love model, we consider the following formulation:

$$\left\{\begin{array}{ll} \underline{\sigma}^0 \in \underline{X}^0,\ (u_0,u_1)\in M\times N & \\ \forall \underline{\tau}\in \underline{X}^0, & a(\underline{\sigma}^0,\underline{\tau})+b(\underline{\tau},(u_0,u_1)) = -a(\phi^f \underline{I},\underline{\tau}) \\ \forall (v_0,v_1)\in M\times N, & b(\underline{\sigma}^0,(v_0,v_1)) = \langle \phi^f, v_1\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}, \end{array}\right. \tag{1.2.1}$$

where the Hilbert spaces are defined by

$$\begin{aligned} M &= \left\{ v\in H^{3/2}(\Gamma);\ v=0 \text{ on } \Gamma_0\cup\Gamma_1\right\},\\ N &= \left\{ v\in H^{1/2}(\Gamma);\ v=0 \text{ on } \Gamma_0\right\} \end{aligned}$$

and the continuous bilinear forms on $\underline{X}\times\underline{X}$ and $\underline{X}\times(M\times N)$ are:

$$\begin{aligned} a(\underline{\sigma},\underline{\tau}) &= \frac{1}{1-\nu}\int_\Omega \underline{\sigma}:\underline{\tau}\,\mathrm{dx} - \frac{\nu}{1-\nu^2}\int_\Omega (\mathrm{tr}\underline{\sigma})(\mathrm{tr}\underline{\tau})\,\mathrm{dx},\\ b(\underline{\tau},(\mu,\lambda)) &= \langle \gamma_1(\underline{\tau}),\mu\rangle_{-\frac{3}{2},\frac{3}{2},\Gamma} - \langle \gamma_0(\underline{\tau}),\lambda\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}. \end{aligned}$$

THEOREM 1.2.1. *Problem (1.2.1) has a unique solution $(\underline{\sigma}^0,u_0,u_1)$. Moreover,*

$$\left\{\begin{array}{ll} \underline{\sigma}^0+\phi^f\underline{I}=\underline{\sigma} & \text{in } \Omega\\ u_0=u & \text{on } \Gamma\\ u_1=\partial_n u & \text{on } \Gamma, \end{array}\right.$$

where $(\underline{\sigma},u)$ satisfies the Kirchhoff-Love equations (1.1.1).

Proof. The well-posedness results from the Babuška-Brezzi theorem [**47**]; in particular, we have checked the *inf-sup* condition for $b(\cdot,\cdot)$:

$$\sup_{\underline{\tau}\in\underline{X}^0}\frac{b(\underline{\tau},(v_0,v_1))}{\|\underline{\tau}\|_{\underline{X}}} \geq c(\|v_0\|_{\frac{3}{2},\Gamma}+\|v_1\|_{\frac{1}{2},\Gamma}).$$

The proof's details and the interpretation in the sense of distributions are given in [A1]. ■

Concerning now the Reissner-Mindlin problem, we introduce additional bilinear forms $a^\varepsilon(\cdot,\cdot)$ and $c(\cdot,\cdot)$ defined on $\underline{X}^\varepsilon\times\underline{X}^\varepsilon$ and on $\underline{X}^\varepsilon\times L^2(\Omega)$ by:

$$\begin{aligned} a^\varepsilon(\underline{\sigma},\underline{\tau}) &= a(\underline{\sigma},\underline{\tau})+\varepsilon^2\,a_0(\underline{\sigma},\underline{\tau}),\\ a_0(\underline{\sigma},\underline{\tau}) &= \frac{1}{1-\nu}\int_\Omega \mathrm{div}\underline{\sigma}\cdot\mathrm{div}\underline{\tau}\,\mathrm{dx},\\ c(\underline{\tau},\mu) &= \int_\Omega(\tau_{12}-\tau_{21})\,\mu\,\mathrm{dx}. \end{aligned}$$

The role of $c(\cdot,\cdot)$ is to dualize the symmetry of the bending tensor, while $a_0(\cdot,\cdot)$ takes into account the new variable which is the rotation vector.

Then we propose the following variational formulation:

$$\left\{\begin{array}{ll} \underline{\sigma}^{\varepsilon,0}\in\underline{X}^{\varepsilon,0},\ (u_0^\varepsilon,r_0^\varepsilon)\in M\times N,\ \lambda^\varepsilon\in L^2(\Omega) & \\ \forall\underline{\tau}\in\underline{X}^{\varepsilon,0}, & a^\varepsilon(\underline{\sigma}^{\varepsilon,0},\underline{\tau})+b(\underline{\tau},(u_0^\varepsilon,r_0^\varepsilon))+c(\underline{\tau},\lambda^\varepsilon) = -a^\varepsilon(\phi^f\underline{I},\underline{\tau})\\ \forall(\zeta,\eta)\in M\times N, & b(\underline{\sigma}^{\varepsilon,0},(\zeta,\eta)) = \langle\phi^f,\eta\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}\\ \forall\mu\in L^2(\Omega), & c(\underline{\sigma}^{\varepsilon,0},\mu) = 0. \end{array}\right. \tag{1.2.2}$$

The main tool for its well-posedness is again the Babuška-Brezzi theorem. The most technical point is the proof of the *inf-sup* condition for $d(\cdot,\cdot)$, defined on $\underline{X}^{\varepsilon,0}\times(M\times N\times L^2(\Omega))$ by:

$$d(\underline{\tau},(\zeta,\eta,\mu))=b(\underline{\tau},(\zeta,\eta))+c(\underline{\tau},\mu).$$

LEMMA 1.2.2. *To any $(\zeta,\eta,\mu)\in M\times N\times L^2(\Omega)$, one can associate a tensor $\underline{\tau}\in\underline{X}^{\varepsilon,0}$ such that:*

$$\left\{\begin{array}{l} d(\underline{\tau},(\zeta,\eta,\mu))\geq c(\|\mu\|_{0,\Omega}+\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma})^2\\ \|\ \underline{\tau}\ \|_{\underline{X}^{\varepsilon}}\leq c(\|\mu\|_{0,\Omega}+\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma}).\end{array}\right. \tag{1.2.3}$$

Proof. First, one constructs $\underline{\tau}_1$ such that the *inf-sup* condition for $b(\cdot,\cdot)$ holds. Let $(\zeta,\eta)\in M\times N$ and let $\mathbf{q}=(\partial_t\zeta)\,\mathbf{t}+\eta\,\mathbf{n}\in\mathbf{H}^{1/2}(\Gamma)$ which satisfies, since $|\Gamma_0|\neq 0$,

$$\|\mathbf{q}\|_{1/2,\Gamma}\leq c(\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma}).$$

One next considers the auxiliary problem:

$$\left\{\begin{array}{ll}\boldsymbol{\Delta}\mathbf{w}=\mathbf{0} & \text{in }\Omega\\ \mathbf{w}=\mathbf{q} & \text{on }\Gamma\end{array}\right. \tag{1.2.4}$$

and chooses $\underline{\tau}_1=-\boldsymbol{\nabla}\mathbf{w}\in\underline{X}^{\varepsilon,0}$. Obviously, $\mathrm{div}\underline{\tau}_1=0$ and $b(\underline{\tau}_1,(\zeta,\eta))=|\mathbf{w}|^2_{1,\Omega}$, so it follows that:

$$b(\underline{\tau}_1,(\zeta,\eta))\geq c(\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma})^2,\qquad \|\ \underline{\tau}_1\ \|_{\underline{X}^{\varepsilon}}\leq c(\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma}).$$

Next, let $\underline{\tau}_2\in\underline{X}^{\varepsilon,0}$ such that $\underline{\tau}=\underline{\tau}_1+\underline{\tau}_2$ satisfies the relations (1.2.3) and $\mathrm{div}\underline{\tau}=0$. The latter relation ensures that the *inf-sup* condition holds uniformly with respect to ε. The trick is to construct a tensor $\underline{\tau}_2$ with vanishing traces, such that $b(\underline{\tau}_2,(\zeta,\eta))=0$ so one will only have to check the *inf-sup* condition for $c(\cdot,\cdot)$ now. For any $\mu\in L^2(\Omega)$, let $P(\mu)=\dfrac{1}{|\Omega|}\int_\Omega\mu\,\mathrm{dx}$ and consider $\lambda=\mu-P(\mu)-curl\mathbf{w}$, which belongs to $L^2_0(\Omega)$ and satisfies:

$$\|\lambda\|_{0,\Omega}\leq c(\|\mu\|_{0,\Omega}+\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma}).$$

It is well-known (cf. **[91]** for instance) that there exists $\mathbf{v}\in\mathbf{H}^1_0(\Omega)$ such that:

$$\mathrm{div}\,\mathbf{v}=\lambda\text{ in }\Omega,\qquad |\mathbf{v}|_{1,\Omega}\leq c\,\|\lambda\|_{0,\Omega}.$$

Therefore, the function $\boldsymbol{\varphi}=\mathbf{v}+\dfrac{P(\mu)}{2}\mathbf{x}$ satisfies the next two conditions:

$$\mathrm{div}\boldsymbol{\varphi}=\mu-curl\mathbf{w}\text{ in }\Omega,\qquad |\boldsymbol{\varphi}|_{1,\Omega}\leq c(\|\mu\|_{0,\Omega}+\|\eta\|_{1/2,\Gamma}+\|\zeta\|_{3/2,\Gamma}).$$

The boundary Γ being polygonal, one finally deduces that $\partial_t\boldsymbol{\varphi}\cdot\mathbf{n}=\partial_t(\partial_t\boldsymbol{\varphi}\cdot\mathbf{t})=0$ on Γ. The conclusion follows by choosing next $\underline{\tau}_2=-\mathrm{curl}\boldsymbol{\varphi}\in\underline{X}^{\varepsilon,0}$. ■

It was also proved in [A2] that:

THEOREM 1.2.3. *Let $(\underline{\sigma}^\varepsilon, (u_0^\varepsilon, r_0^\varepsilon), \lambda^\varepsilon)$ be the solution of (1.2.2). Then $\underline{\sigma}^\varepsilon$ is the bending moment calculated by (1.1.2) and*

$$\begin{cases} r_0^\varepsilon = \mathbf{r}^\varepsilon \cdot \mathbf{n} & \text{on } \Gamma \\ u_0^\varepsilon = u^\varepsilon & \text{on } \Gamma \\ \lambda^\varepsilon = -\dfrac{1}{2} curl \mathbf{r}^\varepsilon & \text{in } \Omega \end{cases}$$

where $(\underline{\sigma}^\varepsilon, \mathbf{r}^\varepsilon, u^\varepsilon)$ satisfies the equations (1.1.2).

In order to recover the other initial unknowns of (1.1.2), one can now solve for the transverse displacement u^ε the second-order elliptic problem:

$$\begin{cases} \Delta u^\varepsilon = \dfrac{1}{1+\nu} \mathrm{tr}\underline{\sigma}^\varepsilon - \dfrac{\varepsilon^2}{1-\nu} f & \text{in } \Omega \\ u^\varepsilon = 0 & \text{on } \Gamma_0 \cup \Gamma_1 \\ u^\varepsilon = u_0^\varepsilon & \text{on } \Gamma_2, \end{cases} \tag{1.2.5}$$

while the rotation vector $\mathbf{r}^\varepsilon$ is given by the relation:

$$\mathbf{r}^\varepsilon = \frac{\varepsilon^2}{1-\nu} \mathrm{div}\underline{\sigma}^\varepsilon + \nabla u^\varepsilon. \tag{1.2.6}$$

1.3. Equivalent mixed formulations

In order to avoid the discretization of the constraint $\mathrm{div}(\mathrm{div}\underline{\tau}) = 0$ imposed on the test-functions of (1.2.1) and (1.2.2), an equivalent formulation based on a special decomposition of the spaces $\underline{X}^0$ and $\underline{X}^{\varepsilon,0}$ was introduced. The unknowns of the new variational problems now belong to classical Sobolev spaces such as $H^1(\Omega)$, $H^{1/2}(\Gamma)$, $L^2(\Omega)$, which are easy to approximate by conforming finite elements.

1.3.1. Characterization of constrained sub-spaces. Any $\underline{\tau} \in \underline{X}^0$ satisfies the constraint $\mathrm{D}(\underline{\tau}) = 0$. Applying Tartar's lemma (cf. **[91]** or **[161]** for instance), one gets the existence of a unique $\rho \in L_0^2(\Omega)$ such that $\mathrm{div}\underline{\tau} = \mathrm{curl}\rho$. One more application of the same lemma gives the existence of a unique function $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega) \cap \mathbf{L}_0^2(\Omega)$ such that

$$\underline{\tau} = \mathrm{curl}\boldsymbol{\varphi} + \rho \underline{J}, \qquad \underline{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

REMARK 1.3.1. It follows that for $\underline{\tau} \in \underline{X}^f$ there exist unique functions $\widetilde{\boldsymbol{\varphi}} \in \mathbf{H}^1(\Omega) \cap \mathbf{L}_0^2(\Omega)$ and $\widetilde{\rho} \in L_0^2(\Omega)$ such that $\underline{\tau} = \mathrm{curl}\widetilde{\boldsymbol{\varphi}} + \widetilde{\rho}\underline{J} + \phi^f \underline{I}$.

The trace operators can then be expressed on $\underline{X}^0$ in the following way:

$$\gamma_0(\underline{\tau}) = -\partial_t \boldsymbol{\varphi} \cdot \mathbf{n}, \qquad \gamma_1(\underline{\tau}) = -\partial_t(\partial_t \boldsymbol{\varphi} \cdot \mathbf{t}). \tag{1.3.1}$$

If moreover the tensor $\underline{\tau} \in \underline{X}^0$ is symmetric, then $2\rho = \mathrm{div}\boldsymbol{\varphi}$ and $\underline{\tau}$ can be written as below:

$$\underline{\tau} = \begin{pmatrix} \partial_2 \varphi_1 & (\partial_2 \varphi_2 - \partial_1 \varphi_1)/2 \\ (\partial_2 \varphi_2 - \partial_1 \varphi_1)/2 & -\partial_1 \varphi_2 \end{pmatrix}, \tag{1.3.2}$$

with a unique function $\boldsymbol{\varphi}$ now belonging to

$$\mathbf{H} = \left\{ \boldsymbol{\varphi} \in \mathbf{H}^1(\Omega); \ \int_\Omega \boldsymbol{\varphi} \, \mathrm{dx} = \mathbf{0}, \ \int_\Omega \mathrm{div}\boldsymbol{\varphi} \, \mathrm{dx} = 0 \right\}.$$

Thanks to Korn's inequality, $\mathbf{H}$ is a Hilbert space endowed with the norm $|\cdot|_{1,\Omega}$.

In a similar way, one obtains that for any $\underline{\tau} \in \underline{X}^{\varepsilon,0}$ there exist unique functions $\rho \in H^1(\Omega)/\mathbb{R}$ and $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega)/\mathbb{R}^2$ such that :

$$\underline{\tau} = \mathrm{curl}\boldsymbol{\varphi} + \rho\underline{J}. \tag{1.3.3}$$

We also consider the Hilbert spaces:

$$\begin{aligned}\mathbf{K} &= \left\{\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega);\ \int_\Omega \boldsymbol{\varphi}\,\mathrm{dx} = \mathbf{0}\right\}\\ W^\varepsilon &= \left\{\rho \in L^2_0(\Omega);\ \varepsilon\mathrm{curl}\rho \in \mathbf{L}^2(\Omega)\right\},\end{aligned}$$

endowed with the norms $|\cdot|_{1,\Omega}$ and $\|\cdot\|_{0,\Omega} + \varepsilon|\cdot|_{1,\Omega}$ respectively, and we put $\mathbf{Y}^\varepsilon = \mathbf{K} \times W^\varepsilon$.

Note that contrarily to the Kirchhoff-Love model, the symmetry of the bending moment is no longer imposed because this would lead to a function $\boldsymbol{\varphi}$ too regular, and hence more difficult to approximate by continuous low-order finite elements.

1.3.2. New formulation of the Kirchhoff-Love model. By means of the decomposition (1.3.2), we obtain a new equivalent formulation of (1.2.1), whose main unknown belongs to $\mathbf{H}$. For this purpose, let us define the bilinear continuous form $A(\cdot,\cdot)$ on $\mathbf{H} \times \mathbf{H}$:

$$\begin{aligned}A(\boldsymbol{\psi},\boldsymbol{\varphi}) &= a\left(\mathrm{curl}\boldsymbol{\psi} + \frac{1}{2}(\mathrm{div}\boldsymbol{\psi})\underline{J},\ \mathrm{curl}\boldsymbol{\varphi} + \frac{1}{2}(\mathrm{div}\boldsymbol{\varphi})\underline{J}\right)\\ &= \frac{1}{1-\nu}\int_\Omega \left[\partial_2\psi_1\partial_2\varphi_1 + \partial_1\psi_2\partial_1\varphi_2 + \frac{1}{2}(\partial_2\psi_2 - \partial_1\psi_1)(\partial_2\varphi_2 - \partial_1\varphi_1)\right]\mathrm{dx}\\ &\quad -\frac{\nu}{1-\nu^2}\int_\Omega (\partial_2\psi_1 - \partial_1\psi_2)(\partial_2\varphi_1 - \partial_1\varphi_2)\,\mathrm{dx}.\end{aligned}$$

Let us also compute from (1.3.1), for any $(v_0, v_1) \in M \times N$,

$$b\left(\mathrm{curl}\boldsymbol{\varphi} + \frac{1}{2}(\mathrm{div}\boldsymbol{\varphi})\underline{J},\ (v_0,v_1)\right) = -\langle\partial_t(\partial_t\boldsymbol{\varphi}\cdot\mathbf{t}), v_0\rangle_{-\frac{3}{2},\frac{3}{2},\Gamma} + \langle\partial_t\boldsymbol{\varphi}\cdot\mathbf{n}, v_1\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma} = -\langle\partial_t\nabla w, \boldsymbol{\varphi}\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}$$

where w is any function of $H^2(\Omega)$ satisfying $w = v_0$, $\partial_n w = v_1$ on Γ. This leads us to introduce a new bilinear form $B(\cdot,\cdot)$ on $\mathbf{H} \times \mathbf{Z}$ by setting

$$B(\boldsymbol{\varphi}, \mathbf{q}) = -\langle\partial_t\mathbf{q}, \boldsymbol{\varphi}\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}$$

where

$$\mathbf{Z} = \left\{\mathbf{q} \in \mathbf{H}^{1/2}(\Gamma);\ \mathbf{q} = \mathbf{0} \text{ on } \Gamma_0,\ \mathbf{q}\cdot\mathbf{t} = 0 \text{ on } \Gamma_1,\ \int_\Gamma \mathbf{q}\cdot\mathbf{t}\,\mathrm{ds} = 0\right\}$$

is endowed with the usual norm $\|\cdot\|_{1/2,\Gamma}$. To any $\mathbf{q} \in \mathbf{Z}$, one can associate a unique couple $(v_0, v_1) \in M \times N$ by putting $\mathbf{q} = (\partial_t v_0)\mathbf{t} + v_1\mathbf{n}$.

We also introduce the linear continuous forms $F(\cdot)$ and $G(\cdot)$ on $\mathbf{H}$, respectively $\mathbf{Z}$ by

$$\begin{aligned}F(\boldsymbol{\varphi}) &= -\frac{1}{1+\nu}\int_\Omega \phi^f(\partial_2\varphi_1 - \partial_1\varphi_2)\,\mathrm{dx},\\ G(\mathbf{q}) &= \int_\Gamma \phi^f\mathbf{q}\cdot\mathbf{n}\mathrm{ds}\end{aligned}$$

and consider the following mixed variational problem:

$$\left\{\begin{array}{lll}\boldsymbol{\psi} \in \mathbf{H},\ \mathbf{p} \in \mathbf{Z} & & \\ \forall\boldsymbol{\varphi} \in \mathbf{H}, & A(\boldsymbol{\psi},\boldsymbol{\varphi}) + B(\boldsymbol{\varphi},\mathbf{p}) & = F(\boldsymbol{\varphi})\\ \forall\mathbf{q} \in \mathbf{Z}, & B(\boldsymbol{\psi},\mathbf{q}) & = G(\mathbf{q})\end{array}\right. \tag{1.3.4}$$

which is shown to admit a unique solution. I present below the proof of the *inf-sup* condition for $B(\cdot,\cdot)$.

LEMMA 1.3.2. *There exists $c > 0$ such that*

$$\sup_{\boldsymbol{\varphi}\in\mathbf{H}} \frac{B(\boldsymbol{\varphi},\mathbf{q})}{|\boldsymbol{\varphi}|_{1,\Omega}} \geq c\,\|\mathbf{q}\|_{\frac{1}{2},\Gamma}\,.$$

Proof. Let any $\mathbf{q}\in\mathbf{Z}$ and let the auxiliary boundary value problem (1.2.4). On the one hand, there exists a unique $\mathbf{z}\in\mathbf{H}^1(\Omega)\cap\mathbf{L}^2_0(\Omega)$ such that $\mathrm{curl}\mathbf{w}=\nabla\mathbf{z}$. Moreover, one has that

$$\int_\Omega \mathrm{div}\mathbf{z}\,\mathrm{dx} = \int_\Gamma \mathbf{w}\cdot\mathbf{t}\,\mathrm{ds} = 0$$

so $\mathbf{z}$ belongs to $\mathbf{H}$. On the other hand, since $\partial_t\mathbf{q} = -(\mathrm{curl}\mathbf{w})\mathbf{n}\in\mathbf{H}^{-1/2}(\Gamma)$ one has that

$$\sup_{\boldsymbol{\varphi}\in\mathbf{H}} \frac{B(\boldsymbol{\varphi},\mathbf{q})}{|\boldsymbol{\varphi}|_{1,\Omega}} \geq \frac{\int_\Omega \mathrm{curl}\mathbf{w}:\nabla\mathbf{z}\mathrm{dx}}{|\mathbf{z}|_{1,\Omega}} = |\mathbf{w}|_{1,\Omega}\,.$$

But $\mathbf{q}=\mathbf{0}$ on Γ_0 and Poincaré's inequality together with the trace theorem yield that $|\mathbf{w}|_{1,\Omega}\geq c\,\|\mathbf{q}\|_{\frac{1}{2},\Gamma}$, which ends the proof. ∎

As regards now the link with the solution $(\underline{\sigma},u)$ of problem (1.1.1), one directly has

$$\left\{\begin{array}{ll} \underline{\sigma} = \mathrm{curl}\boldsymbol{\psi} + \dfrac{1}{2}(\mathrm{div}\boldsymbol{\psi})\underline{J} + \phi^f\underline{I} & \text{in } \Omega \\ \nabla u = \mathbf{p} \quad \text{on } \Gamma & \end{array}\right. \tag{1.3.5}$$

whereas the displacement u is given by the second order elliptic problem:

$$\left\{\begin{array}{ll} \Delta u = \dfrac{1}{1+\nu}(\mathrm{tr}\underline{\sigma}) = \dfrac{1}{1+\nu}(-curl\boldsymbol{\psi} + 2\phi^f) & \text{in } \Omega \\ u = 0 & \text{on } \Gamma_0\cup\Gamma_1 \\ \partial_n u = \mathbf{p}\cdot\mathbf{n} & \text{on } \Gamma_2. \end{array}\right. \tag{1.3.6}$$

1.3.3. New formulation of the Reissner-Mindlin model. To any $\underline{\sigma},\ \underline{\tau}\in\underline{X}^{\varepsilon,0}$ we now associate by means of (1.3.3) the corresponding couples $(\boldsymbol{\psi},\xi)$, $(\boldsymbol{\varphi},\rho)\in\mathbf{Y}^\varepsilon$ and we introduce the following bilinear form on $\mathbf{Y}^\varepsilon\times\mathbf{Y}^\varepsilon$:

$$A^\varepsilon(\cdot,\cdot) = A(\cdot,\cdot) + \varepsilon^2 A_0(\cdot,\cdot)$$

where:

$$\begin{aligned} A((\boldsymbol{\psi},\xi),(\boldsymbol{\varphi},\rho)) = a(\underline{\sigma},\underline{\tau}) &= \frac{1}{1-\nu}\int_\Omega \left[(\xi-\partial_1\psi_1)(\rho-\partial_1\varphi_1)+(\xi-\partial_2\psi_2)(\rho-\partial_2\varphi_2)\right]\mathrm{dx} \\ &\quad + \frac{1}{1-\nu}\int_\Omega (\partial_2\psi_1\partial_2\varphi_1+\partial_1\psi_2\partial_1\varphi_2)\,\mathrm{dx} \\ &\quad - \frac{\nu}{1-\nu^2}\int_\Omega (\partial_2\psi_1-\partial_1\psi_2)(\partial_2\varphi_1-\partial_1\varphi_2)\,\mathrm{dx}, \\ A_0((\boldsymbol{\psi},\xi),(\boldsymbol{\varphi},\rho)) = a_0(\underline{\sigma},\underline{\tau}) &= \frac{1}{1-\nu}\int_\Omega \mathrm{curl}\xi\cdot\mathrm{curl}\rho\,\mathrm{dx}. \end{aligned}$$

We also define the continuous forms $B(\cdot,\cdot)$, $C(\cdot,\cdot)$, $F^\varepsilon(\cdot)$ on $\mathbf{Y}^\varepsilon\times\mathbf{Z}$, $\mathbf{Y}^\varepsilon\times L^2(\Omega)$ and $\mathbf{Y}^\varepsilon$ respectively, by:

$$\begin{aligned} B((\boldsymbol{\varphi},\rho),\mathbf{q}) &= -\langle\partial_t\mathbf{q},\boldsymbol{\varphi}\rangle_{-\frac{1}{2},\frac{1}{2},\Gamma}, \\ C((\boldsymbol{\varphi},\rho),\lambda) = c(\underline{\tau},\lambda) &= \int_\Omega \lambda\,(2\rho-\mathrm{div}\boldsymbol{\varphi})\,\mathrm{dx}, \\ F^\varepsilon((\boldsymbol{\varphi},\rho)) &= -a^\varepsilon(\phi^f\underline{I},\underline{\tau}). \end{aligned}$$

I have then established in [A2] the following result.

THEOREM 1.3.3. *The mixed problem:*

(1.3.7)
$$\left\{\begin{array}{ll} (\boldsymbol{\psi}^\varepsilon,\xi^\varepsilon)\in\mathbf{Y}^\varepsilon,\ \mathbf{p}^\varepsilon\in\mathbf{Z},\ \lambda^\varepsilon\in L^2(\Omega) & \\ \forall((\boldsymbol{\varphi},\rho))\in\mathbf{Y}^\varepsilon, & A^\varepsilon((\boldsymbol{\psi}^\varepsilon,\xi^\varepsilon),(\boldsymbol{\varphi},\rho))+B((\boldsymbol{\varphi},\rho),\mathbf{p}^\varepsilon)+C((\boldsymbol{\varphi},\rho),\lambda^\varepsilon) = F^\varepsilon((\boldsymbol{\varphi},\rho)) \\ \forall\mathbf{q}\in\mathbf{Z}, & B((\boldsymbol{\psi}^\varepsilon,\xi^\varepsilon),\mathbf{q}) = G(\mathbf{q}) \\ \forall\mu\in L^2(\Omega), & C((\boldsymbol{\psi}^\varepsilon,\xi^\varepsilon),\mu) = 0 \end{array}\right.$$

satisfies the hypotheses of the Babuška-Brezzi theorem, uniformly with respect to ε.

Proof. For any $(\boldsymbol{\varphi},\rho)\in\mathbf{Y}^\varepsilon$, let $\underline{\tau}=\mathrm{curl}\boldsymbol{\varphi}+\rho\underline{J}$. One obviously has that

$$A((\boldsymbol{\varphi},\rho),(\boldsymbol{\varphi},\rho))=a(\underline{\tau},\underline{\tau})\geq c\parallel\underline{\tau}\parallel_{0,\Omega}^2.$$

In order to prove the uniform $\mathbf{Y}^\varepsilon$-ellipticity of $A^\varepsilon(\cdot,\cdot)$, it is sufficient to show

$$|\boldsymbol{\varphi}|_{1,\Omega}^2+\|\rho\|_{0,\Omega}^2\leq c\parallel\underline{\tau}\parallel_{0,\Omega}^2. \tag{1.3.8}$$

Note that

$$\parallel\underline{\tau}\parallel_{0,\Omega}^2=|\varphi_1|_{1,\Omega}^2+|\varphi_2|_{1,\Omega}^2+2\|\rho\|_{0,\Omega}^2-2\int_\Omega\rho\,(\partial_1\varphi_1+\partial_2\varphi_2)\,\mathrm{dx}.$$

According to **[91]**, there exists a positive constant k such that, for $\rho\in L_0^2(\Omega)$,

$$k\,\|\rho\|_{0,\Omega}\leq\|\nabla\rho\|_{-1,\Omega}=\parallel\mathrm{div}\underline{\tau}\parallel_{-1,\Omega}\leq k_1\parallel\underline{\tau}\parallel_{0,\Omega}.$$

Thanks to Young's inequality, one obtains $|\boldsymbol{\varphi}|_{1,\Omega}\leq c\parallel\underline{\tau}\parallel_{0,\Omega}$ which implies the estimate (1.3.8).

In order to establish the uniform *inf-sup* condition for $D(\cdot,\cdot)=B(\cdot,\cdot)+C(\cdot,\cdot)$ on $\mathbf{Y}^\varepsilon\times(\mathbf{Z}\times L^2(\Omega))$, we fix an arbitrary couple $(\mathbf{q},\mu)\in\mathbf{Z}\times L^2(\Omega)$ and we construct $(\boldsymbol{\varphi},\rho)\in\mathbf{Y}^\varepsilon$ such that

$$\begin{array}{rcl} D((\boldsymbol{\varphi},\rho),(\mathbf{q},\mu)) & \geq & c(\|\mathbf{q}\|_{1/2,\Gamma}^2+\|\mu\|_{0,\Omega}^2), \\ |\boldsymbol{\varphi}|_{1,\Omega}+\|\rho\|_{0,\Omega}+\varepsilon\,|\rho|_{1,\Omega} & \leq & c(\|\mathbf{q}\|_{1/2,\Gamma}+\|\mu\|_{0,\Omega}). \end{array}$$

We actually take $\rho=0$, so we only have to construct $\boldsymbol{\varphi}\in\mathbf{K}$.

To any $\mathbf{q}\in\mathbf{Z}$, we first associate $\boldsymbol{\varphi}_1\in\mathbf{H}^1(\Omega)/\mathbb{R}^2$ exactly as in Lemma 1.3.2, such that:

$$\begin{array}{rcll} |\boldsymbol{\varphi}_1|_{1,\Omega} & \leq & c\,\|\mathbf{q}\|_{1/2,\Gamma}, & \mathrm{div}\boldsymbol{\varphi}_1\in L_0^2(\Omega), \\ B((\boldsymbol{\varphi}_1,\rho),\mathbf{q}) & \geq & c\,\|\mathbf{q}\|_{1/2,\Gamma}^2, & \forall\rho\in W^\varepsilon. \end{array}$$

Next, for any $\mu\in L^2(\Omega)$ we introduce $\lambda\in L_0^2(\Omega)$ defined by:

$$\lambda=-\mu+P(\mu)-\mathrm{div}\boldsymbol{\varphi}_1,$$

to which we associate $\boldsymbol{\varphi}_2\in\mathbf{H}_0^1(\Omega)$ satisfying:

$$\mathrm{div}\boldsymbol{\varphi}_2=\lambda,\qquad\qquad|\,\boldsymbol{\varphi}_2\,|_{1,\Omega}\leq c\,\|\lambda\|_{0,\Omega}.$$

Finally, we set :

$$\boldsymbol{\varphi}=\boldsymbol{\varphi}_1+\boldsymbol{\varphi}_2-\frac{P(\mu)}{2}\mathbf{x}$$

and choose $\boldsymbol{\varphi}_1$ (which is unique up to a constant) such that $\int_\Omega \boldsymbol{\varphi}_1 \mathrm{dx} = 0$. Then $\boldsymbol{\varphi} \in \mathbf{H}$ and moreover:

$$\begin{aligned} C((\boldsymbol{\varphi},0),\mu) &= \|\mu\|^2_{0,\Omega}\,, \\ |\boldsymbol{\varphi}|_{1,\Omega} &\leq c(\|\mathbf{q}\|_{1/2,\Gamma} + \|\lambda\|_{0,\Omega} + \|P(\mu)\|_{0,\Omega}) \leq c(\|\mathbf{q}\|_{1/2,\Gamma} + \|\mu\|_{0,\Omega}). \end{aligned}$$

The boundary Γ being polygonal, one has that $\partial_t \boldsymbol{\varphi} = \partial_t \boldsymbol{\varphi}_1 - \dfrac{P(\mu)}{2}\mathbf{t}$, which implies:

$$B((\boldsymbol{\varphi},0),\mathbf{q}) = B((\boldsymbol{\varphi}_1,0),\mathbf{q}).$$

One can now deduce the *inf-sup* condition, uniformly with respect to ε. ■

The proof of the next statement is quite technical and can be found in [A2].

THEOREM 1.3.4. *Let $((\boldsymbol{\psi}^\varepsilon,\xi^\varepsilon),\mathbf{p}^\varepsilon,\lambda^\varepsilon)$ be the unique solution of (1.3.7). Then:*

$$\left\{\begin{array}{ll} \underline{\sigma}^\varepsilon = \mathrm{curl}\boldsymbol{\psi}^\varepsilon + \xi^\varepsilon \underline{J} + \phi^f \underline{I} & \textit{in } \Omega \\ \mathbf{r}^\varepsilon = \mathbf{p}^\varepsilon & \textit{on } \Gamma \\ -\dfrac{1}{2}\mathit{curl}\mathbf{r}^\varepsilon = \lambda^\varepsilon & \textit{in } \Omega \end{array}\right. \tag{1.3.9}$$

where $(\underline{\sigma}^\varepsilon, u^\varepsilon, \mathbf{r}^\varepsilon)$ is the solution of (1.1.2).

Finally, it is important to note that the physical variables are recovered from the solution of problem (1.3.7) by means of relations (1.3.9) and (1.2.6) for the bending moment and the rotation, and of the next elliptic problem for the transverse displacement:

$$\left\{\begin{array}{ll} \Delta u^\varepsilon = \dfrac{1}{1+\nu}\mathrm{tr}\underline{\sigma}^\varepsilon - \dfrac{\varepsilon^2}{1-\nu} f & \text{in } \Omega \\ u^\varepsilon = 0 & \text{on } \Gamma_0 \cup \Gamma_1 \\ \partial_t u^\varepsilon = \mathbf{p}^\varepsilon \cdot \mathbf{t} & \text{on } \Gamma_2 \end{array}\right. . \tag{1.3.10}$$

1.4. Finite element approximation

Let $(\mathcal{T}_h)_h$ a regular family of triangulations of Ω consisting of triangles, $\mathcal{E}_h^1$ the set of edges situated on $\Gamma_1 \cup \Gamma_2$ and$\mathcal{T}_h^*$ denote the set of triangles $K \in \mathcal{T}_h$ which have at least an edge in $\mathcal{E}_h^1$.

We first consider a P_1 - continuous finite element approximation ϕ_h^f of ϕ^f, solution of

$$\left\{\begin{array}{l} \phi_h^f \in V_h \\ \forall v_h \in V_h, \quad \int_\Omega \nabla\phi_h^f \cdot \nabla v_h \,\mathrm{dx} = \int_\Omega f v_h \,\mathrm{dx} \end{array}\right. \tag{1.4.1}$$

where $V_h = W_h \cap V$ and

$$W_h = \left\{ v_h \in \mathcal{C}^0(\bar{\Omega});\ v_{h|K} \in P_1,\ \forall K \in \mathcal{T}_h \right\}.$$

The regularity results for the Laplace operator ensure cf. **[96]**, **[112]** that there exists $b \in]\dfrac{1}{2}, 1]$ ($b = 1$ if Ω is convex) such that

$$| \phi^f - \phi_h^f |_{1,\Omega} = \inf_{v_h \in V_h} | \phi^f - v_h |_{1,\Omega} \leq c h^b \| f \|_{0,\Omega}\,.$$

1.4.1. Discrete Kirchhoff-Love formulation. We consider the following finite dimensional spaces $\mathbf{H}_h \subset \mathbf{H}$ and $\mathbf{Z}_h \subset \mathbf{Z}$:

$$\begin{aligned}
\mathbf{H}_h^1 &= \Big\{\boldsymbol{\varphi}_h \in \mathbf{H}^1(\Omega);\ \forall K \in \mathcal{T}_h,\ \boldsymbol{\varphi}_{h|K} \in \mathbf{P}_1 \text{ if } K \notin \mathcal{T}_h^* \\
&\qquad\qquad \text{and } \boldsymbol{\varphi}_{h|K} \in \mathbf{P}_2 \text{ if } K \in \mathcal{T}_h^*\Big\}, \\
\mathbf{H}_h &= \mathbf{H} \cap \mathbf{H}_h^1, \\
\mathbf{Z}_h &= \left\{\mathbf{q}_h \in \mathbf{Z};\ \mathbf{q}_h \in \boldsymbol{C}^0(\Gamma) \text{ and } \forall e \in \mathcal{E}_h^1,\ \mathbf{q}_{h|e} \in \mathbf{P}_1\right\}.
\end{aligned}$$

The degrees of freedom of $\boldsymbol{\varphi}_h \in \mathbf{H}_h$ are its values at the nodes of the triangulation, to which we add the values at the midpoints of the edges belonging to $\mathcal{E}_h^1$.

We write down the discrete version of the continuous problem (1.3.4) as follows:

$$\left\{\begin{array}{ll}
\boldsymbol{\psi}_h \in \mathbf{H}_h,\ \mathbf{p}_h \in \mathbf{Z}_h & \\
\forall \boldsymbol{\varphi}_h \in \mathbf{H}_h, \quad A(\boldsymbol{\psi}_h, \varphi_h) + B(\boldsymbol{\varphi}_h, \mathbf{p}_h) & = F_h(\boldsymbol{\varphi}_h) \\
\forall \mathbf{q}_h \in \mathbf{Z}_h, \quad B(\boldsymbol{\psi}_h, \mathbf{q}_h) & = G_h(\mathbf{q}_h).
\end{array}\right. \tag{1.4.2}$$

The linear forms $F_h(\cdot)$ and $G_h(\cdot)$ are obtained from $F(\cdot)$ and $G(\cdot)$ by replacing ϕ^f by ϕ_h^f.

Then one can show the uniform *inf-sup* condition below:

LEMMA 1.4.1. *There exists a positive constant c independent of h such that*

$$\forall \mathbf{q}_h \in \mathbf{Z}_h, \quad \sup_{\boldsymbol{\varphi}_h \in \mathbf{H}_h} \frac{B(\boldsymbol{\varphi}_h, \mathbf{q}_h)}{|\boldsymbol{\varphi}_h|_{1,\Omega}} \geq c \parallel \mathbf{q}_h \parallel_{1/2,\Gamma}.$$

Proof. We apply once more Fortin's trick (see **[47]**, **[158]**), by using the continuous *inf-sup* condition and the interpolation operator $\widetilde{\boldsymbol{\mathcal{I}}_h} : \mathbf{H} \cap \boldsymbol{C}^0(\overline{\Omega}) \to \mathbf{H}_h$ defined hereafter. Let $\boldsymbol{\mathcal{I}}_{1h}$ be the classical Lagrange interpolation operator which satisfies, for any $\boldsymbol{\varphi} \in \mathbf{H}^1(\Omega) \cap \boldsymbol{C}^0(\overline{\Omega})$,

$$(\boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi})_{|K} \in \mathbf{P}_1 \quad \text{and} \quad \boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi}(N) = \varphi(N),$$

for every triangle K and every vertex N of $\mathcal{T}_h$. We also introduce the operator $\boldsymbol{\mathcal{I}}_{2h}$ defined by $(\boldsymbol{\mathcal{I}}_{2h}\boldsymbol{\varphi})_{|K} \in \mathbf{P}_2$ and

$$\boldsymbol{\mathcal{I}}_{2h}\boldsymbol{\varphi}(N) = \mathbf{0}, \quad \int_e (\boldsymbol{\varphi} - \boldsymbol{\mathcal{I}}_{2h}\boldsymbol{\varphi})\,\mathrm{ds} = \mathbf{0},$$

for every vertex N of $\mathcal{T}_h$ and every edge $e \in \mathcal{E}_h$.

Then we put (see also **[47]**) on every $K \in \mathcal{T}_h$:

$$\boldsymbol{\mathcal{I}}_h\boldsymbol{\varphi} = \left\{\begin{array}{ll}
\boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi} & \text{if} \quad K \notin \mathcal{T}_h^* \\
\boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi} + \boldsymbol{\mathcal{I}}_{2h}(\boldsymbol{\varphi} - \boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi}) & \text{if} \quad K \in \mathcal{T}_h^*
\end{array}\right.,$$

which clearly has the property:

$$\forall e \in \mathcal{E}_h^1, \quad \int_e \boldsymbol{\mathcal{I}}_h\boldsymbol{\varphi}\,\mathrm{ds} = \int_e \boldsymbol{\varphi}\,\mathrm{ds}. \tag{1.4.3}$$

If $\boldsymbol{\varphi} \in \mathbf{H} \cap \boldsymbol{C}^0(\overline{\Omega})$ then we only have $\boldsymbol{\mathcal{I}}_h\boldsymbol{\varphi} \in \mathbf{H}_h^1$, hence we construct

$$\widetilde{\boldsymbol{\mathcal{I}}_h}\boldsymbol{\varphi} = \boldsymbol{\mathcal{I}}_h\boldsymbol{\varphi} - a\mathbf{x} + \mathbf{b} \in \mathbf{H}_h, \qquad a \in \mathbb{R},\ \mathbf{b} \in \mathbb{R}^2.$$

Let us now come back to the proof of the uniform *inf-sup* condition for problem (1.4.2). To any $\mathbf{q}_h \in \mathbf{Z}_h$, we associate exactly as in Lemma 1.3.2 a function $\mathbf{z} \in \mathbf{H}$ such that

$$\frac{B(\mathbf{z}, \mathbf{q}_h)}{|\mathbf{z}|_{1,\Omega}} \geq c\,\|\mathbf{q}_h\|_{\frac{1}{2},\Gamma} \geq |\mathbf{z}|_{1,\Omega}.$$

We note that $\nabla\mathbf{z} = \mathrm{curl}\mathbf{w}$ with $\mathbf{w} \in \mathbf{H}^1(\Omega)$, $\Delta\mathbf{w} = \mathbf{0}$ in Ω and $\mathbf{w} = \mathbf{q}_h$ on Γ. Classical regularity results yield cf. **[96]**, **[112]** that $\mathbf{z} \in \mathbf{H}^{1+a}(\Omega)$ with $a > 0$, so $\boldsymbol{\mathcal{I}}_h\mathbf{z}$ is well-defined.

Then by considering the discrete function $\mathbf{z}_h = \widetilde{\boldsymbol{\mathcal{I}}_h}\mathbf{z}$ and using (1.4.3) we obtain:

$$B(\mathbf{z}, \mathbf{q}_h) = B(\boldsymbol{\mathcal{I}}_h\mathbf{z}, \mathbf{q}_h) = B(\mathbf{z}_h, \mathbf{q}_h).$$

The last equality follows from the fact that

$$B(a\mathbf{x} + \mathbf{b}, \mathbf{q}_h) = a \int_\Gamma \mathbf{q}_h \cdot \mathbf{t}\mathrm{ds} = 0.$$

On the other hand, we obtain by standard arguments such as the passage to the reference and the Bramble-Hilbert lemma that $|\boldsymbol{\mathcal{I}}_h\mathbf{z}|_{1,\Omega} \leq c\,|\mathbf{z}|_{1,\Omega}$. Finally, we obtain that

$$\sup_{\boldsymbol{\varphi}_h \in \mathbf{H}_h} \frac{B(\boldsymbol{\varphi}_h, \mathbf{q}_h)}{|\boldsymbol{\varphi}_h|_{1,\Omega}} \geq \frac{B(\mathbf{z}_h, \mathbf{q}_h)}{|\ \mathbf{z}_h\ |_{1,\Omega}} \geq c\,\|\mathbf{q}_h\|_{\frac{1}{2},\Gamma}$$

which completes the lemma's proof. ■

One can now deduce, thanks to the Babuška-Brezzi theory, the next *a priori* error bound.

THEOREM 1.4.2. *The variational problem (1.4.2) has a unique solution, which satisfies:*

$$|\boldsymbol{\psi} - \boldsymbol{\psi}_h|_{1,\Omega} + \|\mathbf{p} - \mathbf{p}_h\|_{\frac{1}{2},\Gamma} \leq c\{\inf_{\boldsymbol{\varphi}_h \in \mathbf{H}_h} |\boldsymbol{\psi} - \boldsymbol{\varphi}_h|_{1,\Omega} + \inf_{\mathbf{q}_h \in \mathbf{Z}_h} \|\mathbf{p} - \mathbf{q}_h\|_{\frac{1}{2},\Gamma} + \inf_{v_h \in V_h} |\ \phi^f - v_h\ |_{1,\Omega}\}$$

with a constant c independent of the discretization.

1.4.2. Discrete Reissner-Mindlin formulation. The approximation of the additional unknown $\xi^\varepsilon \in W^\varepsilon$ is achieved in W_h, while for $\lambda^\varepsilon \in L^2(\Omega)$ and $\psi^\varepsilon \in \mathbf{K}$ we employ the finite dimensional spaces

$$\begin{aligned} L_h &= \left\{\lambda_h \in L^2(\Omega);\ \ \forall K \in \mathcal{T}_h,\ \ \lambda_{h|K} \in P_0\right\}, \\ \mathbf{K}_h^1 &= \left\{\boldsymbol{\varphi}_h \in \mathbf{H}^1(\Omega);\ \ \forall K \in \mathcal{T}_h,\ \ \boldsymbol{\varphi}_{h|K} \in \mathbf{P}_2\right\}, \\ \mathbf{K}_h &= \mathbf{K} \cap \mathbf{K}_h^1. \end{aligned}$$

For $\varepsilon \neq 0$ the space W^ε coincides algebraically with $H^1(\Omega)$, so $W_h \subset W^\varepsilon$. The norm considered on W_h is the weighted norm previously defined on W^ε.

For the sake of simplicity, we denote:

$$\mathbf{Y}_h = \mathbf{K}_h \times W_h \subset \mathbf{Y}^\varepsilon$$

and we consider the following discrete version of (1.3.7):

$$\text{(1.4.4)} \quad \begin{cases} (\boldsymbol{\psi}_h^\varepsilon, \xi_h^\varepsilon) \in \mathbf{Y}_h,\ \mathbf{p}_h^\varepsilon \in \mathbf{Z}_h,\ \lambda_h^\varepsilon \in L_h \\ \forall (\boldsymbol{\varphi}_h, \rho_h) \in \mathbf{Y}_h, \quad A^\varepsilon((\boldsymbol{\psi}_h^\varepsilon, \xi_h^\varepsilon), (\boldsymbol{\varphi}_h, \rho_h)) + B((\boldsymbol{\varphi}_h, \rho_h), \mathbf{p}_h^\varepsilon) + C((\boldsymbol{\varphi}_h, \rho_h), \lambda_h^\varepsilon) \\ \qquad\qquad = F_h^\varepsilon((\boldsymbol{\varphi}_h, \rho_h)) \\ \forall \mathbf{q}_h \in \mathbf{Z}_h, \qquad B((\boldsymbol{\psi}_h^\varepsilon, \xi_h^\varepsilon), \mathbf{q}_h) = G_h(\mathbf{q}_h) \\ \forall \mu_h \in L_h, \qquad C((\boldsymbol{\psi}_h^\varepsilon, \xi_h^\varepsilon), \mu_h) = 0. \end{cases}$$

The most technical point is again the discrete *inf-sup* condition, shown in [A2].

LEMMA 1.4.3. *There exists c, independent of h and ε, such that for any* $(\mathbf{q}_h, \mu_h) \in \mathbf{Z}_h \times L_h$,

$$\sup_{(\boldsymbol{\varphi}_h, \rho_h) \in \mathbf{Y}_h} \frac{B((\boldsymbol{\varphi}_h, \rho_h), \mathbf{q}_h) + C((\boldsymbol{\varphi}_h, \rho_h), \mu_h)}{|\boldsymbol{\varphi}_h|_{1,\Omega} + \|\rho_h\|_{0,\Omega} + \varepsilon\,|\rho_h|_{1,\Omega}} \geq c(\|\mathbf{q}_h\|_{1/2,\Gamma} + \|\mu_h\|_{0,\Omega}).$$

Proof. We make use of the continuous *inf-sup* condition established in Theorem 1.3.3 and of the continuous interpolation operator $\boldsymbol{\mathcal{L}}_h : \mathbf{H}^1(\Omega) \cap \boldsymbol{C}^0(\overline{\Omega}) \to \mathbf{K}_h^1$, defined on every triangle $K \in \mathcal{T}_h$ by:

$$\boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi} = \boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi} + \boldsymbol{\mathcal{I}}_{2h}(\boldsymbol{\varphi} - \boldsymbol{\mathcal{I}}_{1h}\boldsymbol{\varphi}).$$

It clearly satisfies the properties:

$$\begin{aligned} \int_e \boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi}\,\mathrm{ds} &= \int_e \boldsymbol{\varphi}\,\mathrm{ds}, \quad \forall e \in \mathcal{E}_h, \\ \int_K \mathrm{div}(\boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi})\,\mathrm{dx} &= \int_K \mathrm{div}\boldsymbol{\varphi}\,\mathrm{dx}, \quad \forall K \in \mathcal{T}_h. \end{aligned}$$

Note that for $\boldsymbol{\varphi} \in \mathbf{K} \cap \boldsymbol{C}^0(\overline{\Omega})$, one only has that $\boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi}$ belongs to $\mathbf{K}_h^1$, and not to $\mathbf{K}$.

To any $\mathbf{q}_h \in \mathbf{Z}_h$, we associate $\boldsymbol{\varphi}_1 \in \mathbf{H}^1(\Omega)/\mathbb{R}^2$ as in the proof of Theorem 1.3.3. Since $\mathbf{q}_h \in \mathbf{H}^1(\Gamma)$, we get as previously that $\boldsymbol{\varphi}_1 \in \boldsymbol{C}^0(\overline{\Omega})$. Then by considering the discrete function $\boldsymbol{\varphi}_{1h} = \boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi}_1 \in \mathbf{K}_h^1$, we obtain:

$$\frac{B((\boldsymbol{\varphi}_{1h},0),\mathbf{q}_h)}{|\,\boldsymbol{\varphi}_{1h}\,|_{1,\Omega}} = \frac{B((\boldsymbol{\varphi}_1,0),\mathbf{q}_h)}{|\,\boldsymbol{\varphi}_{1h}\,|_{1,\Omega}} \geq c\frac{B((\boldsymbol{\varphi}_1,0),\mathbf{q}_h)}{|\boldsymbol{\varphi}_1|_{1,\Omega}} \geq c\,\|\mathbf{q}_h\|_{1/2,\Gamma}\,.$$

Next, following the proof of Theorem 1.3.3, to any $\mu_h \in L_h$ we associate $\boldsymbol{\varphi}_2 \in \mathbf{H}_0^1(\Omega)$ such that

$$\mathrm{div}\boldsymbol{\varphi}_2 = -\mu_h + P(\mu_h) - \mathrm{div}\boldsymbol{\varphi}_{1h}, \qquad |\,\boldsymbol{\varphi}_2\,|_{1,\Omega} \leq c(\|\mu_h\|_{0,\Omega} + \|\mathrm{div}\boldsymbol{\varphi}_{1h}\|_{0,\Omega}).$$

We put $\boldsymbol{\varphi}_h' = \boldsymbol{\varphi}_{1h} + \boldsymbol{\mathcal{L}}_h\boldsymbol{\varphi}_2 - \dfrac{P(\mu_h)}{2}\mathbf{x}$ which belongs to $\mathbf{K}_h^1$ and consider $\boldsymbol{\varphi}_h = \boldsymbol{\varphi}_h' - P(\boldsymbol{\varphi}_h') \in \mathbf{K}_h$. Then we have:

$$\begin{aligned} C((\boldsymbol{\varphi}_h,0),\mu_h) &= \|\mu_h\|_{0,\Omega}^2\,, \\ B((\boldsymbol{\varphi}_h,0),\mathbf{q}_h) &= B((\boldsymbol{\varphi}_{1h},0),\mathbf{q}_h), \\ |\boldsymbol{\varphi}_h|_{1,\Omega} &\leq c(\|\mathbf{q}_h\|_{1/2,\Gamma} + \|\mu_h\|_{0,\Omega}) \end{aligned}$$

which ends the proof. ∎

The previous result immediately implies the well-posedness of the mixed problem (1.4.4), as well as the *a priori* error bound:

$$\begin{aligned} &|\boldsymbol{\psi}^\varepsilon - \boldsymbol{\psi}_h^\varepsilon|_{1,\Omega} + \|\xi^\varepsilon - \xi_h^\varepsilon\|_{0,\Omega} + \varepsilon\,|\xi^\varepsilon - \xi_h^\varepsilon|_{1,\Omega} + \|\mathbf{p}^\varepsilon - \mathbf{p}_h^\varepsilon\|_{1/2,\Gamma} + \|\lambda^\varepsilon - \lambda_h^\varepsilon\|_{0,\Omega} \\ &\quad \leq c\Big\{\inf_{\boldsymbol{\varphi}_h\in\mathbf{K}_h}|\boldsymbol{\psi}^\varepsilon - \boldsymbol{\varphi}_h|_{1,\Omega} + \inf_{\rho_h\in W_h}(\|\xi^\varepsilon - \rho_h\|_{0,\Omega} + \varepsilon\,|\xi^\varepsilon - \rho_h|_{1,\Omega}) \\ &\quad + \inf_{\mathbf{q}_h\in\mathbf{Z}_h}\|\mathbf{p}^\varepsilon - \mathbf{q}_h\|_{1/2,\Gamma} + \inf_{\mu_h\in L_h}\|\lambda^\varepsilon - \mu_h\|_{0,\Omega} + \inf_{v_h\in V_h}|\,\phi^f - v_h\,|_{1,\Omega}\Big\} \end{aligned}$$

with a constant c independent of both h and ε.

1.4.3. Approximation of the physical variables. It is now easy to recover the quantities of interest, that is the bending moment, the displacement and the rotation.

Let us first consider the Kirchhoff-Love model. In view of (1.3.5) and (1.3.6), we set:

$$\underline{\sigma}_h = \mathrm{curl}\boldsymbol{\psi}_h + \frac{1}{2}(\mathrm{div}\boldsymbol{\psi}_h)\underline{J} + \phi_h^f\underline{I}$$

and then we solve

$$\begin{cases} u_h \in V_h \\ \forall v_h \in V_h, \quad \int_\Omega \nabla u_h \cdot \nabla v_h \mathrm{dx} = \dfrac{-1}{1+\nu}\int_\Omega (\mathrm{tr}\underline{\sigma}_h) v_h \mathrm{dx} + \int_{\Gamma_2} \mathbf{p}_h \cdot \mathbf{n} v_h \mathrm{ds}. \end{cases}$$

If the solution $\underline{(\sigma, u)}$ of the initial Kirchhoff-Love model satisfies

$$\underline{\sigma} \in (H^a(\Omega))^4, \qquad u \in H^{2+a}(\Omega), \qquad 0 < a \leq 1,$$
$$\| \underline{\sigma} \|_{a,\Omega} + \| u \|_{2+a,\Omega} \leq c \| f \|_{0,\Omega},$$

then standard interpolation results imply that

$$\| \underline{\sigma} - \underline{\sigma}_h \|_{0,\Omega} + \| u - u_h \|_{1,\Omega} + \| D(\underline{\sigma}) - D(\underline{\sigma}_h) \|_{-1,\Omega} \leq c h^{\min\{a,b\}} \| f \|_{0,\Omega} .$$

As regards now the Reissner-Mindlin model, the approximated bending moment is given by

$$\underline{\sigma}_h^\varepsilon = \mathrm{curl}\boldsymbol{\psi}_h^\varepsilon + \xi_h^\varepsilon \underline{J} + \phi_h^f \underline{I}$$

while u_h^ε is obtained by discretizing the variational formulation of (1.3.10). Note that the pre-processing of ϕ_h^f and the post-processing of the displacement (for both models) are very simple: one has to solve twice a Laplace problem, whose matrix is computed only once. The discrete rotation vector $\mathbf{r}_h^\varepsilon$ is given by (1.2.6), while the multiplier λ_h^ε represents a piecewise constant approximation of $curl\mathbf{r}^\varepsilon$.

In order to obtain the convergence rate of the discretization method, we assume that the exact solution of (1.1.2) satisfies:

$$\mathbf{r}^\varepsilon \in \mathbf{H}^{1+a}(\Omega), \quad u^\varepsilon \in H^{1+a}(\Omega),$$
$$\|\mathbf{r}^\varepsilon\|_{1+a,\Omega} + \|u^\varepsilon\|_{1+a,\Omega} + \varepsilon \| \mathrm{div}\underline{\sigma}^\varepsilon \|_{a,\Omega} \leq c \|f\|_{0,\Omega} .$$

This hypothesis is verified in convex domains with $a = 1$, at least for clamped plates (cf. for instance [**47**]).

Then we deduce, with c independent of the plate's thickness ε and of h, that:

$$\| \underline{\sigma}^\varepsilon - \underline{\sigma}_h^\varepsilon \|_{0,\Omega} + \varepsilon \| \mathrm{div}(\underline{\sigma}^\varepsilon - \underline{\sigma}_h^\varepsilon) \|_{0,\Omega} + |u^\varepsilon - u_h^\varepsilon|_{1,\Omega} + \|\mathbf{r}^\varepsilon - \mathbf{r}_h^\varepsilon\|_{0,\Omega} \leq c h^{\min\{a,b\}} \|f\|_{0,\Omega} .$$

CHAPTER 2

APPLICATIONS IN NEWTONIAN FLUID MECHANICS

CHAPTER 2

Applications in Newtonian fluid mechanics

This chapter gathers together several applications of numerical modeling by finite elements in fluid mechanics. All problems treated here are governed by the incompressible Stokes or Navier-Stokes equations. Nevertheless, each section addresses a different topic.

Thus, in Section 2.1 a low-order conforming approximation of the steady Navier-Stokes equations endowed with non-standard boundary conditions is considered. Section 2.2 is devoted to the derivation, study and coupling of new 2D and 1D hierarchical models in fluvial hydrodynamics. Finally, Section 2.3 deals with the *a priori* and *a posteriori* analysis of a new discontinuous Galerkin method for the Stokes equations.

2.1. Navier-Stokes equations with non-standard boundary conditions

The next results are taken from the papers [A7], [A4] and from the technical report [**4**]. For sake of brievety, only the 2D case is studied here; the extension to 3D can be found in [**5**] or [C5].

To summarize, we study a velocity-vorticity-pressure fomulation of the incompressible Navier-Stokes equations with the boundary conditions of [**64**]. A low-order conforming finite element approximation, based on piecewise constant elements for the vorticity and the pressure and on continuous, piecewise linear elements for the velocity is proposed. To ensure the well-posedness of the corresponding discrete Stokes problem, a stabilization term taking into account the jumps of the pressure across the edges is added. Next, based on the stability and the consistency properties of the discrete Stokes operator, we establish that the discrete Navier-Stokes problem is well-posed and we obtain *a priori* and *a posteriori* error estimates. The theoretical results are illustrated by numerical experiments.

2.1.1. Functional framework. Let Ω be a simply connected bounded domain of $\mathbb{R}^2$, with a polygonal boundary $\Gamma = \partial\Omega$. We consider the stationary incompressible Navier-Stokes equations

$$\begin{cases} (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p = \mathbf{f} & \text{in } \Omega \\ \qquad\qquad \mathrm{div}\mathbf{u} = 0 & \text{in } \Omega \end{cases}$$

and impose the following boundary conditions:

$$\begin{cases} \mathbf{u}\cdot\mathbf{n} = 0, \quad \mathbf{u}\cdot\mathbf{t} = 0 & \text{on } \Gamma_1 \\ \mathbf{u}\cdot\mathbf{t} = 0, \quad p + \frac{1}{2}\mathbf{u}\cdot\mathbf{u} = 0 & \text{on } \Gamma_2 \\ \mathbf{u}\cdot\mathbf{n} = 0, \quad curl\mathbf{u} = 0 & \text{on } \Gamma_3, \end{cases}$$

where $\Gamma_1, \Gamma_2, \Gamma_3$ are disjoint and form a partition of $\Gamma = \partial\Omega$.

For the sake of simplicity, we present here only the case of homogeneous boundary conditions; the non-homogeneous case is treated in [**4**].

The above boundary conditions are introduced in [**64**] and they are weaker than the classical ones, where the velocity is given all over the boundary. They apply, for instance, in a pipe flow (cf. Fig. 2.1.1): one can impose either the injection velocity $\mathbf{u} = \mathbf{u}_0$ or the pressure on the inlet boundary Γ_{in}, a no-slip condition $\mathbf{u} = \mathbf{0}$ on the lower boundary Γ_{low} and the pressure at the tube exit Γ_{out}, with an unknown velocity field. Finally, on the upper boundary Γ_{up}, the domain's geometry and the axisymmetry hypothesis lead to imposing a null vorticity.

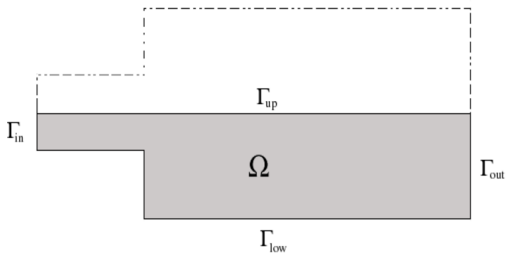


FIGURE 2.1.1. Domain with non-standard boundary conditions

By means of the relation:

$$(\mathbf{u}\cdot\nabla)\mathbf{u} = (curl\mathbf{u})\mathbf{u}^{\perp} + \frac{1}{2}\nabla(\mathbf{u}\cdot\mathbf{u})$$

where $\mathbf{u}^{\perp} = (u_2, -u_1)$, the problem can be equivalently written as follows:

$$\left\{\begin{array}{ll} \nu\mathrm{curl}\omega + \nabla\widetilde{p} + \omega\mathbf{u}^{\perp} = \mathbf{f} & \text{in } \Omega \\ \omega = curl\mathbf{u} & \text{in } \Omega \\ \mathrm{div}\mathbf{u} = 0 & \text{in } \Omega, \end{array}\right. \tag{2.1.1}$$

together with the boundary conditions

$$\left\{\begin{array}{lll} \mathbf{u}\cdot\mathbf{n} = 0, & \mathbf{u}\cdot\mathbf{t} = 0 & \text{on } \Gamma_1 \\ \mathbf{u}\cdot\mathbf{t} = 0, & \widetilde{p} = 0 & \text{on } \Gamma_2 \\ \mathbf{u}\cdot\mathbf{n} = 0, & \omega = 0 & \text{on } \Gamma_3. \end{array}\right. \tag{2.1.2}$$

The unknowns are now the velocity field $\mathbf{u}$, the dynamic pressure $\widetilde{p} = p + \frac{1}{2}\mathbf{u}\cdot\mathbf{u}$, denoted by p in the sequel, and the scalar vorticity ω. The kinematic viscosity $\nu > 0$ is given and for the sake of simplicity, we take $\mathbf{f} \in \mathbf{L}^{4/3}(\Omega)$ and we suppose that $|\Gamma_2| > 0$.

The analysis of the nonlinear Navier-Stokes problem is based on the properties of the associated linear Stokes operator. Therefore, we first consider the Stokes equations associated to (2.1.1), with data $\mathbf{g} \in \mathbf{L}^{4/3}(\Omega)$ and with the boundary conditions (2.1.2). We next recall some results from **[6]**.

Let us introduce the Hilbert space:

$$\mathbf{M} = \{\mathbf{v} \in H(\mathrm{div}, curl; \Omega); \quad \mathbf{v}\cdot\mathbf{n}|_{\Gamma_1\cup\Gamma_3} = \mathbf{v}\cdot\mathbf{t}|_{\Gamma_1\cup\Gamma_2} = 0\}.$$

Both $H(\mathrm{div}, curl; \Omega)$ and $\mathbf{M}$ are endowed with the norm $\|\mathbf{v}\|^2_{\mathbf{M}} = \|\mathbf{v}\|^2_{0,\Omega} + \|\mathrm{div}\mathbf{v}\|^2_{0,\Omega} + \|curl\mathbf{v}\|^2_{0,\Omega}$.

Under the hypothesis

$$(\mathbf{H1}) \qquad \{\mathbf{v} \in \mathbf{M};\ \mathrm{div}\mathbf{v} = curl\mathbf{v} = 0 \quad \text{a.e. in } \Omega\} = \{\mathbf{0}\},$$

the seminorm $|\mathbf{v}|_{\mathbf{M}} = (\|\mathrm{div}\mathbf{v}\|^2_{0,\Omega} + \|curl\mathbf{v}\|^2_{0,\Omega})^{1/2}$ is a norm on $\mathbf{M}$, equivalent to $\|\cdot\|_{\mathbf{M}}$.

REMARK 2.1.1. The hypothesis **(H1)** is true in particular if one of the following situations hold : $|\Gamma_1| > 0$, or $|\Gamma_1| = |\Gamma_3| = 0$, or $|\Gamma_1| = 0$ and $|\Gamma_3| > 0$ with Γ_3 simply connected.

Another key point is that the space $\mathbf{M}$ is next assumed to be continuously embedded in $\mathbf{H}^s(\Omega)$, for some $s \in]1/2, 1]$. This is not a restrictive hypothesis; it is satisfied (see **[65]**) if there are no nonconvex corners at the intersection of $\overline{\Gamma}_1 \cup \overline{\Gamma}_2$ and $\overline{\Gamma}_3$. The embedding holds with $s = 1$ if Ω is a convex polygon, or if Ω is a Lipschitz-continuous domain and $|\Gamma_2| = |\Gamma_3| = 0$ (cf. for instance **[91]**). Then the Sobolev embedding theorem implies that $\mathbf{M} \subset \mathbf{L}^4(\Omega)$ and that the traces of elements of $\mathbf{M}$ belong to $\mathbf{L}^2(\Gamma)$.

Then we put $\mathbf{X} = \mathbf{L}^2(\Omega)$, we define for all $\boldsymbol{\sigma} = (\omega, p) \in \mathbf{X}$, $\boldsymbol{\tau} = (\theta, q) \in \mathbf{X}$ and $\mathbf{v} \in \mathbf{M}$:

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\tau}) &= \nu \int_\Omega \omega\theta \, \mathrm{dx}, \\ b(\boldsymbol{\tau}, \mathbf{v}) &= -\nu \int_\Omega \theta \, curl\mathbf{v} \, \mathrm{dx} + \int_\Omega q \mathrm{div}\mathbf{v} \, \mathrm{dx}, \\ l(\mathbf{v}) &= -\int_\Omega \mathbf{g} \cdot \mathbf{v} \, \mathrm{dx} \end{aligned}$$

and we consider the following three-fields mixed variational formulation of the Stokes problem:

$$\left\{ \begin{array}{ll} (\boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{X} \times \mathbf{M} & \\ a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \mathbf{u}) = 0 & \forall \boldsymbol{\tau} \in \mathbf{X}, \\ b(\boldsymbol{\sigma}, \mathbf{v}) = l(\mathbf{v}) & \forall \mathbf{v} \in \mathbf{M}. \end{array} \right. \tag{2.1.3}$$

Problem (2.1.3) can be easily shown to satisfy the Babuška-Brezzi conditions, so one can define a linear continuous operator

$$\mathcal{S} : \mathbf{L}^{4/3}(\Omega) \to \mathbf{X} \times \mathbf{L}^4(\Omega), \quad \mathcal{S}(\mathbf{g}) = (\boldsymbol{\sigma}, \mathbf{u})$$

where $(\boldsymbol{\sigma}, \mathbf{u})$ is the unique solution of the Stokes problem (2.1.3).

For the simplicity of notation, we denote from now on the Banach space $\mathbf{X} \times \mathbf{L}^4(\Omega)$ by $\mathbf{Y}$. By introducing the nonlinear operator

$$\mathcal{G} : \mathbf{Y} \to \mathbf{L}^{4/3}(\Omega), \quad \mathcal{G}(\boldsymbol{\sigma}, \mathbf{u}) = \omega \mathbf{u}^\perp,$$

the Navier-Stokes equations (2.1.1) can be put in a general nonlinear setting as follows:

$$\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u}) = (\mathbf{0}, \mathbf{0}) \tag{2.1.4}$$

where the mapping $\mathcal{F}$ is defined by

$$\mathcal{F} : \mathbf{Y} \to \mathbf{Y}, \quad \mathcal{F}(\boldsymbol{\tau}, \mathbf{v}) = (\boldsymbol{\tau}, \mathbf{v}) - \mathcal{S}(\mathbf{f} - \mathcal{G}(\boldsymbol{\tau}, \mathbf{v})).$$

We assume next that there exists a solution $(\boldsymbol{\sigma}, \mathbf{u})$ such that $\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u}) = 0$ and $D\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u})$ is an isomorphism on $\mathbf{Y}$.

2.1.2. Finite element discretization. We are interested in the numerical approximation of the nonlinear problem (2.1.4). For this purpose, we first consider a discretization of the associated linear problem (2.1.3). We assume that each triangulation is compatible with the boundary conditions and moreover, that Γ_2 contains at least one vertex. We employ conforming finite elements of lowest-order:

$$\begin{aligned} L_h &= \{q \in L^2(\Omega);\ q_{|K} \in P_0 \ \forall K \in \mathcal{T}_h\}, \\ \mathbf{X}_h &= L_h \times L_h, \\ \mathbf{M}_h &= \{\mathbf{v} \in \boldsymbol{\mathcal{C}}^0(\overline{\Omega});\ \mathbf{v}_{|K} \in \mathbf{P}_1 \ \forall K \in \mathcal{T}_h\} \cap \mathbf{M}. \end{aligned}$$

The previous choice of spaces ensures that the uniform *inf-sup* condition, which represents the main difficulty in the velocity-pressure formulation of the Stokes problem, is obviously satisfied. However, the bilinear form $a(\cdot, \cdot)$ is not coercive on the discrete kernel $\mathbf{V}_h$ of $b(\cdot, \cdot)$,

$$\mathbf{V}_h = \{\boldsymbol{\tau} \in \mathbf{X}_h;\ b(\boldsymbol{\tau}, \mathbf{v}) = 0 \ \forall \mathbf{v} \in \mathbf{M}_h\}.$$

In order to retrieve its coercivity, we use stabilization, that is we replace $a(\cdot, \cdot)$ by $a(\cdot, \cdot) + \beta \, A_h(\cdot, \cdot)$ where $\beta > 0$ is a stabilization parameter, which can be chosen independently of h.

In **[6]**, the stabilization term $A_h(\cdot, \cdot)$ was defined by means of the jumps of both the pressure and the vorticity across the edges. I have shown that it is actually sufficient to stabilize only

the pressure, leading to similar theoretical and numerical results. So let $A_h : \mathbf{X}_h \times \mathbf{X}_h \to \mathbb{R}$ be defined by

$$A_h(\boldsymbol{\delta}, \boldsymbol{\tau}) = \sum_{e \in \mathcal{E}_h^{int} \cup \Gamma_2} h_e \int_e [r][q] \mathrm{ds}, \quad \forall \boldsymbol{\delta} = (\rho, r),\ \boldsymbol{\tau} = (\theta, q) \in \mathbf{X}_h$$

and let the following approximation of (2.1.3):

$$\left\{ \begin{array}{ll} (\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h \times \mathbf{M}_h & \\ a(\boldsymbol{\sigma}_h, \tau) + \beta\, A_h(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \mathbf{u}_h) = 0 & \forall \boldsymbol{\tau} \in \mathbf{X}_h \\ b(\boldsymbol{\sigma}_h, \mathbf{v}) = l(\mathbf{v}) & \forall \mathbf{v} \in \mathbf{M}_h. \end{array} \right. \tag{2.1.5}$$

It is useful to introduce the seminorm $|\cdot|_h$ defined for all $\boldsymbol{\tau} \in \mathbf{X}_h$ by $|\boldsymbol{\tau}|_h = \sqrt{A_h(\boldsymbol{\tau}, \boldsymbol{\tau})}$.

Then I have shown that the new bilinear form $a(\cdot,\cdot) + \beta\, A_h(\cdot,\cdot)$ is uniformly $\mathbf{V}_h$-elliptic and $\mathbf{X}_h$-continuous with respect to the $\mathbf{L}^2(\Omega)$-norm of $\mathbf{X}$.

LEMMA 2.1.2. *There exist two positive constants independent of h such that:*

$$\begin{aligned} \|q\|_{0,\Omega} \le c_1 \left(\|\theta\|_{0,\Omega} + |\boldsymbol{\tau}|_h \right), \qquad & \forall \boldsymbol{\tau} = (\theta, q) \in \mathbf{V}_h, \\ |\boldsymbol{\tau}|_h \le c_2 \|q\|_{0,\Omega}, \qquad & \forall \boldsymbol{\tau} = (\theta, q) \in \mathbf{X}_h. \end{aligned}$$

It is also useful to establish the next result, which can be found in [A7]:

LEMMA 2.1.3. *To any $\boldsymbol{\tau} = (\theta, q) \in \mathbf{X}_h$, one can associate a function $\boldsymbol{\phi}_h \in \mathbf{H}^1(\Omega) \cap \mathbf{M}$ such that:*

$$|\boldsymbol{\phi}_h|_{1,\Omega} \le c|\boldsymbol{\tau}|_h \quad \textit{and} \quad A_h(\boldsymbol{\delta}, \boldsymbol{\tau}) = \int_\Omega r \mathrm{div} \boldsymbol{\phi}_h \mathrm{dx}, \quad \forall \boldsymbol{\delta} = (\rho, r) \in \mathbf{X}_h$$

where c is independent of both the discretization and the stabilization parameters.

Gathering together the previous lemmas, it follows that the mixed formulation (2.1.5) fulfills the hypotheses of the Babuška-Brezzi theorem, uniformly with respect to h. Hence the discrete problem (2.1.5) is well-posed and one can now introduce the discrete Stokes operator as follows:

$$\mathcal{S}_h : \mathbf{L}^{4/3}(\Omega) \to \mathbf{Y}, \quad \mathcal{S}_h(\mathbf{g}) = (\boldsymbol{\sigma}_h, \mathbf{u}_h)$$

where $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h \times \mathbf{M}_h$ is the unique solution of (2.1.5). Obviously, $\mathcal{S}_h$ is a linear and continuous operator, which satisfies the condition:

$$(\mathbf{A1}) \qquad \forall \mathbf{g} \in \mathbf{L}^{4/3}(\Omega), \quad \|\mathcal{S}_h(\mathbf{g})\|_{\mathbf{Y}} \le c \|\mathbf{g}\|_{\mathbf{L}^{4/3}(\Omega)}$$

with c a positive constant independent of h but depending on β.

Moreover, for smooth data $\mathbf{g} \in \mathbf{L}^2(\Omega)$ one gets that $\mathcal{S}_h$ satisfies the following error bound.

THEOREM 2.1.4. *Let $\mathbf{g} \in \mathbf{L}^2(\Omega)$ and let $\bar{\boldsymbol{\sigma}}_h$ be the $\mathbf{L}^2(\Omega)$-projection of $\boldsymbol{\sigma}$ on $\mathbf{X}_h$, where $(\boldsymbol{\sigma}, \mathbf{u}) = \mathcal{S}(\mathbf{g})$. Then the following estimate holds:*

$$\|(\mathcal{S} - \mathcal{S}_h)(\mathbf{g})\|_{\mathbf{X} \times \mathbf{M}} \le C \{ h \|\mathbf{g}\|_{0,\Omega} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_h\|_{\mathbf{X}} + \inf_{\mathbf{v}_h \in \mathbf{M}_h} |\mathbf{u} - \mathbf{v}_h|_{\mathbf{M}} \},$$

where C is a constant independent of h (but depending on β). If moreover $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbf{H}^1(\Omega) \times \mathbf{H}^2(\Omega)$, the method has an optimal convergence rate $O(h)$.

In the general situation of less regular data, one can establish:

THEOREM 2.1.5. *For any $\mathbf{g} \in \mathbf{L}^{4/3}(\Omega)$, one has:*

$$\|(\mathcal{S} - \mathcal{S}_h)(\mathbf{g})\|_{\mathbf{Y}} \quad \le \quad C \left\{ h^{1/2} \|\mathbf{g}\|_{\mathbf{L}^{4/3}(\Omega)} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_h\|_{\mathbf{X}} + \inf_{\mathbf{v}_h \in \mathbf{M}_h} |\mathbf{u} - \mathbf{v}_h|_{\mathbf{M}} \right\}.$$

Theorem 2.1.5 yields the unconditional convergence of the approximation, that is $\mathcal{S}_h$ satisfies

$$(\mathbf{A2}) \qquad \forall \mathbf{g} \in \mathbf{L}^{4/3}(\Omega), \quad \lim_{h\to 0} \|(\mathcal{S}-\mathcal{S}_h)(\mathbf{g})\|_{\mathbf{Y}} = 0.$$

Let us now look at the numerical approximation of the Navier-Stokes problem (2.1.4):

$$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) = 0, \tag{2.1.6}$$

where the mapping $\mathcal{F}_h$ is defined by:

$$\mathcal{F}_h : \mathbf{Y} \to \mathbf{Y}, \quad \mathcal{F}_h(\boldsymbol{\tau}, \mathbf{v}) = (\boldsymbol{\tau}, \mathbf{v}) - \mathcal{S}_h(\mathbf{f} - \mathcal{G}(\boldsymbol{\tau}, \mathbf{v})).$$

We remark that if $(\boldsymbol{\sigma}_h, \mathbf{u}_h)$ is solution of equation (2.1.6), then $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbf{X}_h \times \mathbf{M}_h$. The functional $\mathcal{F}_h$ is differentiable and for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbf{Y}$, one has:

$$D\mathcal{F}_h(\boldsymbol{\tau}, \mathbf{v}) = \mathcal{I} + \mathcal{S}_h \circ D\mathcal{G}(\boldsymbol{\tau}, \mathbf{v}).$$

2.1.3. Analysis of the discrete nonlinear problem. The analysis of the discrete problem (2.1.6) uses a well-known result based on the implicit function theorem, which was first established in **[50]**. Some variants can be found in **[150]** or in **[54]**. In order to apply here the general result of **[150]**, we suppose that the Stokes operator $\mathcal{S}$ satisfies the following regularity assumption:

(H2) there exists $a > 0$ such that $\mathcal{S} : \mathbf{L}^{4/3}(\Omega) \to \mathbf{H}^a(\Omega) \times \mathbf{H}^{1+a}(\Omega)$ is well-defined and continuous.

REMARK 2.1.6. This condition holds, for instance, with $a = 1$ whenever Ω is a convex polygon and $|\Gamma_2| = |\Gamma_3| = 0$, cf. **[91]**.

I have then established:

THEOREM 2.1.7. *Assume **(H2)**. Then the nonlinear mapping $\mathcal{F}_h$ fulfils the conditions:*
***(C1)** there exists a positive constant c independent of h such that, for any $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbf{Y}$:*

$$\|D\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u}) - D\mathcal{F}_h(\boldsymbol{\tau}, \mathbf{v})\|_{\mathcal{L}(\mathbf{Y})} \le c\,\|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\tau}, \mathbf{v})\|_{\mathbf{Y}}$$

(C2) $\lim_{h\to 0} \|\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u})\|_{\mathbf{Y}} = 0$
***(C3)** there exists $h_0 > 0$ such that for any $h < h_0$, $D\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u})$ is an isomorphism of $\mathbf{Y}$ and*

$$\left\|D\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u})^{-1}\right\|_{\mathcal{L}(\mathbf{Y})} \le 2\left\|D\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u})^{-1}\right\|_{\mathcal{L}(\mathbf{Y})} \ .$$

Proof. Note that $D\mathcal{G}(\boldsymbol{\sigma}, \mathbf{u}) : \mathbf{Y} \to \mathbf{L}^{4/3}(\Omega)$ is defined for any $\boldsymbol{\delta} = (\rho, r) \in \mathbf{X}$, $\mathbf{w} \in \mathbf{L}^4(\Omega)$ by $D\mathcal{G}(\boldsymbol{\sigma}, \mathbf{u})(\boldsymbol{\delta}, \mathbf{w}) = \omega\mathbf{w}^\perp + \rho\mathbf{u}^\perp$.

The stability property **(A1)** together with Hölder's and Cauchy-Schwarz inequalities imply the condition **(C1)**. The consistency property **(C2)** follows from **(A2)** together with the relation

$$\|\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u})\|_{\mathbf{Y}} = \|(\mathcal{S}-\mathcal{S}_h)(\mathbf{f} - \mathcal{G}(\boldsymbol{\sigma}, \mathbf{u}))\|_{\mathbf{Y}}\,.$$

In order to prove **(C3)**, we write that

$$D\mathcal{F}_h(\boldsymbol{\sigma}, \mathbf{u}) = D\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u}) \circ (\mathcal{I} + \mathcal{B}_h)$$

with $\mathcal{B}_h = D\mathcal{F}(\boldsymbol{\sigma},\mathbf{u})^{-1} \circ (D\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u}) - D\mathcal{F}(\boldsymbol{\sigma},\mathbf{u}))$. It is known that if $\|\mathcal{B}_h\|_{\mathcal{L}(\mathbf{Y})} < 1$, then $D\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u})$ is an isomorphism and the next bound holds:

$$\left\|D\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u})^{-1}\right\|_{\mathcal{L}(\mathbf{Y})} \leq \frac{\left\|D\mathcal{F}(\boldsymbol{\sigma},\mathbf{u})^{-1}\right\|_{\mathcal{L}(\mathbf{Y})}}{1 - \|\mathcal{B}_h\|_{\mathcal{L}(\mathbf{Y})}}. \tag{2.1.7}$$

Note that

$$\|\mathcal{B}_h\|_{\mathcal{L}(\mathbf{Y})} \leq \left\|D\mathcal{F}(\boldsymbol{\sigma},\mathbf{u})^{-1}\right\|_{\mathcal{L}(\mathbf{Y})} \|(\mathcal{S} - \mathcal{S}_h) \circ D\mathcal{G}(\boldsymbol{\sigma},\mathbf{u})\|_{\mathcal{L}(\mathbf{Y})}.$$

As a consequence of **(H2)**, one deduces from Theorem 2.1.5 that for any $\mathbf{g} \in \mathbf{L}^{4/3}(\Omega)$:

$$\|(\mathcal{S} - \mathcal{S}_h)(\mathbf{g})\|_{\mathbf{Y}} \quad \leq \quad ch^{\alpha} \|\mathbf{g}\|_{\mathbf{L}^{4/3}(\Omega)}$$

where $\alpha = \min(\frac{1}{2}, a)$. Since $D\mathcal{G}(\boldsymbol{\sigma},\mathbf{u})$ is a bounded operator from $\mathbf{Y}$ to $\mathbf{L}^{4/3}(\Omega)$, it follows that

$$\lim_{h\to 0} \|(\mathcal{S} - \mathcal{S}_h) \circ D\mathcal{G}(\boldsymbol{\sigma},\mathbf{u})\|_{\mathcal{L}(\mathbf{Y})} = 0$$

which yields $\|\mathcal{B}_h\|_{\mathcal{L}(\mathbf{Y})} < \frac{1}{2}$ for $h < h_0$. Together with (2.1.7), this ends the proof. ■

Then the next statement is true, according to **[150]**:

THEOREM 2.1.8. *Assume **(H2)**. Then there exist $h_1 > 0$ and $\delta > 0$ such that, for all $h < h_1$, problem (2.1.6) has a unique solution satisfying $\|(\boldsymbol{\sigma},\mathbf{u}) - (\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{Y}} \leq \delta$. Moreover, the following* a priori*, respectively* a posteriori *estimates hold:*

$$\|(\boldsymbol{\sigma},\mathbf{u}) - (\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{Y}} \quad \leq \quad c\,\|\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u})\|_{\mathbf{Y}} \tag{2.1.8}$$

$$\|(\boldsymbol{\sigma},\mathbf{u}) - (\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{Y}} \quad \leq \quad c'\,\|\mathcal{F}(\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{Y}} \tag{2.1.9}$$

with c, c' independent of the discretization.

The condition **(C2)** yields that the approximation method for the Navier-Stokes problem is unconditionally convergent and its convergence rate is given by an upper bound for $\|\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u})\|_{\mathbf{Y}}$. If $\mathbf{f} \in \mathbf{L}^2(\Omega)$ and $(\boldsymbol{\sigma},\mathbf{u}) \in \mathbf{H}^1(\Omega) \times \mathbf{H}^2(\Omega)$, then one deduces from Theorem 2.1.4 the same convergence rate $O(h)$ as for the Stokes problem:

$$\|(\boldsymbol{\sigma},\mathbf{u}) - (\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{Y}} \leq ch(\|\mathbf{f} - \mathcal{G}(\boldsymbol{\sigma},\mathbf{u})\|_{0,\Omega} + |\boldsymbol{\sigma}|_{1,\Omega} + |\mathbf{u}|_{2,\Omega}).$$

REMARK 2.1.9. By means of a technical Aubin-Nitsche argument, I have shown that the convergence rate for the velocity in $\mathbf{L}^4(\Omega)$-norm is improved to $O(h^{3/2})$. A detailed proof is given in **[4]**.

Theorem 2.1.8 also says that an upper bound of $\|\mathcal{F}(\boldsymbol{\sigma}_h,\mathbf{u}_h)\|_{\mathbf{X}\times\mathbf{M}}$ is an *a posteriori* error estimator. Let us compute for any $\boldsymbol{\tau} = (\theta, q) \in \mathbf{X}$ and $\mathbf{v} \in \mathbf{M}$, the quantity $\langle \mathcal{F}(\boldsymbol{\sigma}_h,\mathbf{u}_h), (\boldsymbol{\tau},\mathbf{v})\rangle$ where $\langle\cdot,\cdot\rangle$ is the scalar product of $\mathbf{X}\times\mathbf{M}$. By taking $\boldsymbol{\tau}_h = \mathbf{0}$ and $\mathbf{v}_h = \boldsymbol{\mathcal{R}}_h\mathbf{v}$ with $\boldsymbol{\mathcal{R}}_h$ a local regularization operator of Clément type, one gets after integration by parts that:

$$\langle \mathcal{F}(\boldsymbol{\sigma}_h,\mathbf{u}_h), (\boldsymbol{\tau},\mathbf{v})\rangle =$$
$$\textstyle\sum_{K\in\mathcal{T}_h} \left(\int_K \eta_{K,1}\theta \mathrm{dx} + \int_K \eta_{K,2} q \mathrm{dx} - \int_K \boldsymbol{\eta}_{K,3}\cdot(\mathbf{v} - \boldsymbol{\mathcal{R}}_h\mathbf{v})\mathrm{dx}\right) + \sum_{e\in\mathcal{E}_h} \int_e \boldsymbol{\eta}_e \cdot (\mathbf{v} - \boldsymbol{\mathcal{R}}_h\mathbf{v})\mathrm{ds}.$$

Here above, we have employed the following residuals, on every triangle K:

$$\eta_{K,1} = \nu(\omega_h - curl\,\mathbf{u}_h), \quad \eta_{K,2} = \mathrm{div}\mathbf{u}_h, \quad \boldsymbol{\eta}_{K,3} = \mathbf{f} - \omega_h\mathbf{u}_h^{\perp},$$

respectively on every edge e:

$$\boldsymbol{\eta}_e = \nu[\omega_h]\mathbf{t} - [p_h]\mathbf{n},$$

where the jump $[\cdot]$ is equal to the trace on a Dirichlet boundary and vanishes on the other boundary edges.

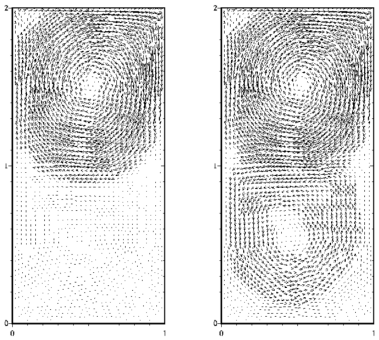


FIGURE 2.1.2. Velocity for Re = 5000 (left) and Re = 20000 (right)

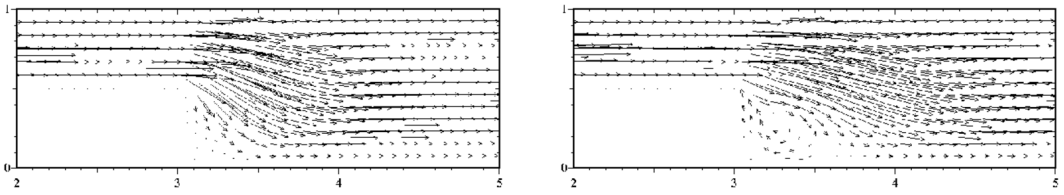


FIGURE 2.1.3. Velocity near the step for Re = 10 (left) and Re = 1000 (right)

Thanks to interpolation error estimates for the operator $\boldsymbol{\mathcal{R}}_h$ on $\mathbf{M} \subset \mathbf{H}^s(\Omega)$ with $s \in]1/2, 1]$, cf. for instance **[31]** or **[91]**, we finally obtain the *a posteriori* error bound

$$\|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\|_{\mathbf{Y}} \leq C(\sum_{K \in \mathcal{T}_h} \eta(K)^2)^{1/2}$$

where C is independent of h and β and where the local error estimator $\eta(K)$ is defined by

$$\eta(K)^2 = \|\eta_{K,1}\|_{0,K}^2 + \|\eta_{K,2}\|_{0,K}^2 + h_K^{2s} \|\eta_{K,3}\|_{0,K}^2 + h_K^{2s-2} \sum_{e \in \partial K} \frac{h_e}{2} (\|\eta_{e,1}\|_{0,e}^2 + \|\eta_{e,2}\|_{0,e}^2).$$

2.1.4. Numerical results. The previous results have been illustrated numerically in 2D ([A7], [A4], **[4]**) but also in 3D (**[5]**, [C5]). The convergence rates were computed, in the case of an exact solution with $\Gamma_i \neq \emptyset$ for $1 \leq i \leq 3$, for different values of the parameters h and β and for Reynolds numbers varying from 1 to 10000. For Re = 100, we have obtained $O(h)$ for the L^2-errors on p and ω and the H^1-error on $\mathbf{u}$, and $O(h)^2$ for the L^4-error on $\mathbf{u}$ in 2D.

The numerical tests were carried out by David Trujillo. Classical examples such as the cavity test, the step test and the T-shaped domain test were treated. The *a posteriori* error estimators were employed in order to improve the solution, and also to optimally choose the stabilization parameter β when the Reynolds number is fixed.

Fig. 2.1.2 shows the velocity for Re = 5000 and Re = 20000 in the driven cavity test on the rectangle $]0, 1[\times]0, 2[$. A second vortex can be observed for a large Reynolds number.

We also show the step test with the pressure given on the inlet and outlet boundaries and a zero velocity imposed elsewhere; similar results were obtained when taking $\omega = 0$ and $\mathbf{u} \cdot \mathbf{n} = 0$ on the upper boundary. For small Reynolds numbers (for instance Re = 10), we retrieve a linear pressure and a laminar flow, which is no longer the case for Re = 1000. These results are illustrated in Fig. 2.1.3 and Fig. 2.1.4.

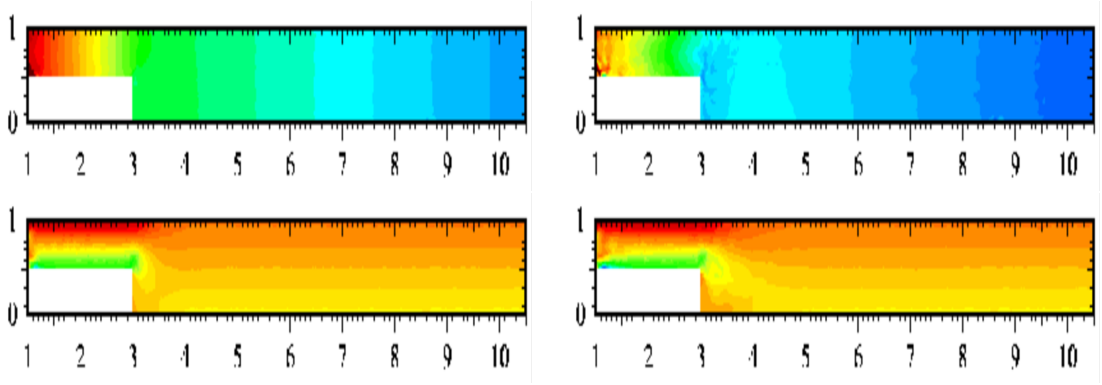


FIGURE 2.1.4. Pressure and vorticity for Re = 10 (left) and Re = 1000 (right)

2.2. Hierarchical modeling in fluvial hydrodynamics

This section is devoted to the derivation, analysis and numerical approximation of some multi-dimensional hydrodynamic models. A 2D horizontal, a 2D vertical and a 1D model are obtained from the weak form of a 3D time-discretized model as conforming approximations on adapted subspaces. For each derived model, a discretization based on conforming classical finite elements is proposed and studied. Our choices of projection subspaces are *inf-sup* stable at both the continuous and discrete levels and yield *a priori* and *a posteriori* error estimates between the physical model and any of its lower-dimensional approximations. Moreover, the deduced models are hierarchical, which allows for a unified analysis and which alleviates their adaptive coupling by means of *a posteriori* error estimators. Finally, some numerical tests are presented.

The following results can be mainly found in [B2], [B3] and [B4], which are submitted or will be soon submitted, and in [C8]. A quasi 3D model, obtained by combining the 2D horizontal and 2D vertical models, is described in [C9], see the PhD thesis of Agnès Petrau [**147**] for more details. Other variants of hydrodynamic models were considered in [A3], [C2], [C3], [C6], see also the technical reports [**2**], [**3**]; they are simpler since linear at each time-step, due to the use of the characteristics method for the time discretization, but also non hierarchical.

2.2.1. Problem setting. We consider the estuarian basin of a river (without islands), characterized by the following geometrical and bathymetrical data, see Fig. 2.2.1 (a). Let $\Sigma \subset \mathbb{R}^2$ be a bounded, Lipschitz continuous domain included in the plane $z = H_{\max}$ with $H_{\max}$ a sufficiently large constant; Σ represents the projection of the riverbed. We denote by $Z_B(x,y)$ the elevation of the bottom, with $Z_B \in W^{1,\infty}(\Sigma)$ and bounded from above by $H_{\max}$, and we introduce the 3D fixed domain

$$\Omega = \{(x,y,z);\ (x,y) \in \Sigma,\ \ Z_B(x,y) < z < H_{\max}\}\,.$$

We put $\partial\Sigma = \bar{\Upsilon}_I \cup \bar{\Upsilon}_{\rm lat}$ with $\Upsilon_{\rm lat}$ the lateral boundary, and $\partial\Omega = \bar{\Gamma}_B \cup \bar{\Sigma} \cup \bar{\Gamma}_I$ where

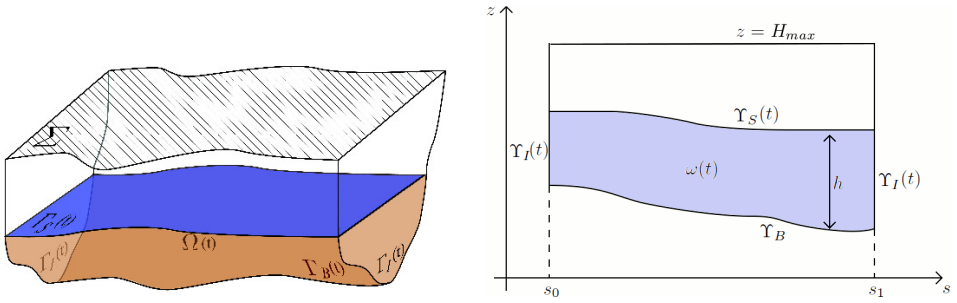
$$\begin{aligned} \Gamma_B &= \{(x,y,z);\ (x,y) \in \Sigma,\ \ z = Z_B(x,y)\}\,, \\ \Gamma_I &= \{(x,y,z);\ (x,y) \in \Upsilon_I,\ Z_B(x,y)\ < z < H_{\max}\}\,. \end{aligned}$$

We introduce similar notation at a given time t. Let the water depth $h(t;x,y)$ satisfy, for all $(x,y) \in \Sigma$ and $t > 0$, the bound $0 \le h(t;x,y) \le H_{\max} - Z_B(x,y)$, let $H(t;x,y) = Z_B(x,y) + h(t;x,y)$ and let the 3D domain occupied by the fluid:

$$\Omega(t) = \{(x,y,z);\ (x,y) \in \Sigma,\ \ Z_B(x,y) < z < H(t;x,y)\}\,.$$

It is useful to denote

$$\Sigma(t) = \{(x,y) \in \Sigma;\ h(t;x,y) > 0\}$$



a) 3D domain b) 2D-vertical domain

FIGURE 2.2.1. Geometrical framework

and to put $\partial\Sigma(t) = \bar{\Upsilon}_I(t) \cup \bar{\Upsilon}_{\rm lat}(t)$ with $\Upsilon_I(t) = \partial\Sigma(t) \cap \Upsilon_I$. We define the riverbed $\Gamma_B(t)$, the free surface $\Gamma_S(t)$ and the inflow/outflow boundary $\Gamma_I(t)$ of $\Omega(t)$ as follows:

$$\begin{aligned}
\Gamma_B(t) &= \partial\Omega(t) \cap \Gamma_B = \{(x,y,z);\ (x,y) \in \Sigma(t),\ \ z = Z_B(x,y)\}\,,\\
\Gamma_S(t) &= \{(x,y,z);\ (x,y) \in \Sigma(t),\ \ z = H(t;x,y)\}\,,\\
\Gamma_I(t) &= \partial\Omega(t) \cap \Gamma_I = \{(x,y,z);\ (x,y) \in \Upsilon_I,\ Z_B(x,y)\ < z < H(t;x,y)\}
\end{aligned}$$

such that $\partial\Omega(t) = \bar{\Gamma}_B(t) \cup \bar{\Gamma}_S(t) \cup \bar{\Gamma}_I(t)$.

The physical problem is described by the time-dependent incompressible Navier-Stokes equations in $\Omega(t)$, with constant density ρ and with gravity and Coriolis forces:

$$\left\{\begin{array}{l} {\rm div}\mathbf{u} = 0 \\ \dfrac{\partial \mathbf{u}}{\partial t} + {\rm curl}\mathbf{u} \times \mathbf{u} + \nu {\rm curl}({\rm curl}\mathbf{u}) + \nabla p - \mathbf{f} \times \mathbf{u} = \mathbf{g} \end{array}\right.$$

where the kinematic viscosity $\nu = \dfrac{\mu}{\rho}$ and $\mathbf{f} = (0,0,f)$ are constant, and where $p = \dfrac{1}{\rho}\widetilde{p} + \dfrac{1}{2}\left|\mathbf{u}\right|^2$ is the dynamic pressure.

Initial and boundary conditions are added. It is worth noting the originality of the latter on the free surface, in contrast with the usual shallow water system where the vertical velocity is set to zero. Thanks to the curl(curl) formulation, we can impose (see also Section 2.1) pure Neumann conditions, that is the pressure p and the tangential vorticity $\nu{\rm curl}\mathbf{u} \times \mathbf{n}$. We also impose a friction and an impermeability condition on the bottom, which finally yields:

$$\left\{\begin{array}{lll} \mathbf{u}\cdot\mathbf{n} = 0\ , & \nu{\rm curl}\mathbf{u} \times \mathbf{n} = -c_B\mathbf{u} & \text{on } \Gamma_B(t) \\ p = p_S, & \nu{\rm curl}\mathbf{u} \times \mathbf{n} = \mathbf{w} & \text{on } \Gamma_S(t) \\ \mathbf{u}\cdot\mathbf{n} = k, & \nu{\rm curl}\mathbf{u} \times \mathbf{n} = \mathbf{w} & \text{on } \Gamma_I(t) \end{array}\right. . \tag{2.2.1}$$

For simplicity of presentation, the surface pressure p_S, the flowrate k, the tangential data $\mathbf{w}$ (related to the wind or to the tide) and the friction coefficient $c_B \geq 0$ are given constants.

REMARK 2.2.1. The classical formulation leads to stronger conditions on the free surface. They were considered in **[84]**, where the authors set $(p - \nu\frac{\partial \mathbf{u}}{\partial \mathbf{n}})_{/z} = p_S$ and $(\nu\frac{\partial \mathbf{u}}{\partial \mathbf{n}})_{/xy} = \mathbf{w}$.

Finally, we close the system by adding the free surface equation, cf. for instance **[89]**:

$$\frac{\partial h}{\partial t} + \sum_{i=1}^{2} u_i^S \partial_i H = u_3^S \quad \text{on } \Sigma \tag{2.2.2}$$

where $\mathbf{u}^S(t;x,y) = \mathbf{u}(t;x,y,H)$. An inflow condition for h is also added.

For the time-discretization, we have chosen to employ the semi-implicit Euler scheme:

$$\frac{h-h^{n-1}}{\Delta t}+\sum_{i=1}^{2}u_i^{S,n-1}\partial_i H=u_3^{S,n-1}\quad\text{on }\Sigma,$$

$$\left\{\begin{array}{c}\mathrm{div}\mathbf{u}=0\\ \dfrac{1}{\Delta t}(\mathbf{u}-\mathbf{u}^{n-1})+\mathrm{curl}\mathbf{u}\times\mathbf{u}+\nu\mathrm{curl}(\mathrm{curl}\mathbf{u})-\mathbf{f}\times\mathbf{u}+\nabla p=\mathbf{g}\end{array}\right.\quad\text{in }\Omega(t^n),\tag{2.2.3}$$

where $\mathbf{u}^{n-1}$, defined on $\Omega(t^{n-1})$, is $\mathbf{H}(\mathrm{div},\mathrm{curl})$-continuously extended on $\Omega(t^n)$ whenever necessary. For simplicity of writing, we denote next $\Omega(t^n)$ by Ω^n and its boundary by $\partial\Omega^n=\bar\Gamma_B^n\cup\bar\Gamma_S^n\cup\bar\Gamma_I^n$. The letter c denotes any constant independent of $\triangle t$, of Ω^n and, whenever relevant, of the space discretization. For the simplicity of presentation, we assume in what follows that $\triangle t\leq 1$ and we take, without any loss of generality, $k=0$ and $\mathbf{w}=\mathbf{0}$.

2.2.2. 3D weak formulation. Let the Hilbert spaces:

$$\begin{aligned}\mathbf{X}&=\{\mathbf{v}\in\mathbf{H}(\mathrm{div},\mathrm{curl};\Omega);\quad\mathbf{v}\cdot\mathbf{n}=0\text{ on }\partial\Omega\},\\ \mathbf{X}(t)&=\left\{\mathbf{v};\ \exists\tilde{\mathbf{v}}\in\mathbf{X},\ \mathbf{v}=\tilde{\mathbf{v}}_{|\Omega(t)}\right\}\\ &=\{\mathbf{v}\in\mathbf{H}(\mathrm{div},\mathrm{curl};\Omega(t));\quad\mathbf{v}\cdot\mathbf{n}=0\text{ on }\Gamma_B(t)\cup\Gamma_I(t)\}.\end{aligned}$$

From now on, we assume for the fixed domain Ω that there exists $s\geq 3/4$ such that

$$(\mathbf{X},|||\cdot|||_{\mathbf{X}})\subset(\mathbf{H}^s(\Omega),\|\cdot\|_{s,\Omega})\tag{2.2.4}$$

where $|||\mathbf{v}|||_{\mathbf{X}}^2=\|\mathbf{v}\|_{0,\Omega}^2+\|\mathrm{div}\mathbf{v}\|_{0,\Omega}^2+\|\mathrm{curl}\mathbf{v}\|_{0,\Omega}^2$.

REMARK 2.2.2. It is known (see **[65]** or **[91]**) that if Ω is a convex polyhedron or if it has a $C^{1,1}$ boundary, then $\mathbf{X}$ is continuously embedded in $\mathbf{H}^1(\Omega)$, so (2.2.4) is fulfilled with optimal $s=1$. This assumption ensures that $\mathbf{X}(t)$ is compactly embedded in $\mathbf{L}^2(\Omega(t))$, and that the space of traces of its functions is compactly embedded in $\mathbf{L}^2(\partial\Omega(t))$. Without assuming (2.2.4), a function $\mathbf{v}$ of $\mathbf{X}(t)$ does not necessarily satisfy the condition $\mathbf{v}\in\mathbf{L}^2(\Gamma_B(t))$ but only $\mathbf{v}\in\mathbf{L}^2_{\mathrm{loc}}(\Gamma_B(t)\cup\Gamma_I(t))$, cf. **[65]**. The result becomes true if the boundary condition $\mathbf{v}\cdot\mathbf{n}=0$ or $\mathbf{v}\wedge\mathbf{n}=\mathbf{0}$ on the free surface $\Gamma_S(t)$ is satisfied, but we don't have this property here.

We next define on $\mathbf{X}(t)$ the following semi-norm and norms

$$\begin{aligned}|\mathbf{v}|_{\mathbf{X}(t)}&=\left(\|\mathrm{div}\mathbf{v}\|_{0,\Omega(t)}^2+\|\mathrm{curl}\mathbf{v}\|_{0,\Omega(t)}^2+c_B\|\mathbf{v}\|_{0,\Gamma_B(t)}^2\right)^{1/2},\\ \|\mathbf{v}\|_{\mathbf{X}(t)}&=\left(\|\mathbf{v}\|_{0,\Omega(t)}^2+|\mathbf{v}|_{\mathbf{X}(t)}^2\right)^{1/2},\\ \|\mathbf{v}\|_{\mathbf{X}(t),\triangle t}&=\left(\frac{1}{\Delta t}\|\mathbf{v}\|_{0,\Omega(t)}^2+|\mathbf{v}|_{\mathbf{X}(t)}^2\right)^{1/2}.\end{aligned}$$

Clearly, one has

$$\|\mathbf{v}\|_{\mathbf{X}(t)}\leq\|\mathbf{v}\|_{\mathbf{X}(t),\triangle t},\quad\forall\mathbf{v}\in\mathbf{X}(t).\tag{2.2.5}$$

Thanks to Sobolev's theorem, assumption (2.2.4) implies that the injection operator

$$\mathcal{I}_1:(\mathbf{X}(t),\|\cdot\|_{\mathbf{X}(t)})\to(\mathbf{L}^4(\Omega(t)),\|\cdot\|_{\mathbf{L}^4(\Omega(t))})$$

is continuous, of norm denoted by $c_1(\Omega(t))$.

We also consider the space $M(t) = L^2(\Omega(t))$, endowed with the usual L^2-norm, and we write problem (2.2.3) in weak form as follows:

$$\left\{\begin{array}{l} (\mathbf{u},p)\in \mathbf{X}(t^n)\times M(t^n) \\ \forall \mathbf{v}\in \mathbf{X}(t^n),\quad A(\mathbf{u};\mathbf{u},\mathbf{v})+B(p,\mathbf{v})=F^{n-1}(\mathbf{v}) \\ \forall q\in M(t^n),\quad B(q,\mathbf{u})=0 \end{array}\right. \tag{2.2.6}$$

where:

$$\begin{aligned} A(\mathbf{w};\mathbf{u},\mathbf{v}) &= A_0(\mathbf{u},\mathbf{v})+A_1(\mathbf{w};\mathbf{u},\mathbf{v}), \\ A_0(\mathbf{u},\mathbf{v}) &= \int_{\Omega^n}\frac{1}{\Delta t}\mathbf{u}\cdot\mathbf{v}\,\mathrm{d}\Omega+\int_{\Omega^n}\nu \mathrm{curl}\mathbf{u}\cdot\mathrm{curl}\mathbf{v}\,\mathrm{d}\Omega+\int_{\Gamma_B^n}c_B\mathbf{u}\cdot\mathbf{v}\,\mathrm{d}\Gamma-\int_{\Omega^n}(\mathbf{f}\times\mathbf{u})\cdot\mathbf{v}\,\mathrm{d}\Omega, \\ A_1(\mathbf{w};\mathbf{u},\mathbf{v}) &= \int_{\Omega^n}(\mathrm{curl}\mathbf{u}\times\mathbf{w})\cdot\mathbf{v}\,\mathrm{d}\Omega, \\ B(p,\mathbf{v}) &= -\int_{\Omega^n}p\,\mathrm{div}\mathbf{v}\,\mathrm{d}\Omega, \\ F^{n-1}(\mathbf{v}) &= \int_{\Omega^n}(\frac{1}{\Delta t}\mathbf{u}^{n-1}+\mathbf{g})\cdot\mathbf{v}\,\mathrm{d}\Omega-\int_{\Gamma_S^n}p_S\mathbf{v}\cdot\mathbf{n}\mathrm{d}\Gamma. \end{aligned}$$

All these forms are clearly continuous with respect to the norms $\|\cdot\|_{\mathbf{X}(t^n),\Delta t}$ and $\|\cdot\|_{0,\Omega^n}$. The continuity constant of $A_1(;,)$ is $c_1^2(\Omega^n)$, and we denote by $c_2(\Delta t,\Omega^n)$ the constant of $F^{n-1}(\cdot)$. Let us also introduce $\mathbf{V}(t^n)=\mathrm{Ker}B$ and notice that

$$A(\mathbf{v};\mathbf{v},\mathbf{v})=A_0(\mathbf{v},\mathbf{v})\geq c_3\,\|\mathbf{v}\|^2_{\mathbf{X}(t^n),\triangle t}\,,\quad \forall\mathbf{v}\in\mathbf{V}(t^n) \tag{2.2.7}$$

with $c_3=\min\{1,\nu\}$. Then we have :

THEOREM 2.2.3. *Problem (2.2.6) has at least one solution. The uniqueness holds if*

$$c_2(\triangle t,\Omega^n)<\frac{c_3^2}{c_1^2(\Omega^n)}. \tag{2.2.8}$$

Proof. We apply a consequence of Brouwer's fixed point theorem (see **[91]**). We first show

$$\inf_{q\in M(t^n)}\sup_{\mathbf{v}\in\mathbf{X}(t^n)}\frac{B(q,\mathbf{v})}{\|\mathbf{v}\|_{\mathbf{X}(t^n)}\,\|q\|_{M(t^n)}}\geq c$$

with c depending only on Ω. The Babuška-Brezzi theorem ensures that for each $\mathbf{u}$ solution of

$$\left\{\begin{array}{c} \mathbf{u}\in\mathbf{V}(t^n) \\ \forall\mathbf{v}\in\mathbf{V}(t^n),\quad A(\mathbf{u};\mathbf{u},\mathbf{v})=F^{n-1}(\mathbf{v}) \end{array}\right.\,, \tag{2.2.9}$$

there exists a unique $p\in M$ such that $(\mathbf{u},p)$ is solution of the mixed problem (2.2.6).

In order to prove existence of a solution, we consider the nonlinear problem (2.2.9) and show, thanks to (2.2.4), that $A_1(\cdot;\cdot,\mathbf{v})$ is sequentially weakly-continuous on $\mathbf{V}(t^n)$ for all $\mathbf{v}\in\mathbf{V}(t^n)$.

The proof of the uniqueness is classical. More details are given in [B2]. ■

2.2.3. Derivation of lower-dimensional models. We have derived several semi-discretized models as conforming approximations of (2.2.6) on adapted subspaces $\mathbf{X}_d(t^n)\times M_d(t^n)$. More precisely, we have obtained two bi-dimensional models, called 2D horizontal and 2D vertical, whether they are written on the river's free surface or on its median longitudinal surface, and also a one-dimensional model, written on the median curve.

The methodology is the following. We first derive the new free surface equation from (2.2.2), by taking $\mathbf{u}^{n-1}$ in $\mathbf{X}_d(t^{n-1})$. We compute the water depth and then we solve the approximate

problem:

$$\text{(2.2.10)}\qquad \begin{cases} (\mathbf{u}_d, p_d) \in \mathbf{X}_d(t^n) \times M_d^*(t^n) \\ \forall \mathbf{v} \in \mathbf{X}_d(t^n), \quad A(\mathbf{u}_d; \mathbf{u}_d, \mathbf{v}) + B(p_d, \mathbf{v}) = F_d^{n-1}(\mathbf{v}) \\ \forall q \in M_d(t^n), \quad B(q, \mathbf{u}_d) = 0. \end{cases}$$

For the 2D horizontal and 1D models, the pressure p_d is looked for in the affine set $M_d^*(t^n) = p_S + M_d(t^n)$ while for the 2D vertical model one simply has $M_d^*(t^n) = M_d(t^n)$.

REMARK 2.2.4. In the non-homogeneous case $\mathbf{u} \cdot \mathbf{n} = k$ on Γ_I^n, the velocity is also looked for in an affine set.

In what follows, I describe the choices of the approximation subspaces for each model and show that they yield well-posed problems (2.2.10). The corresponding boundary value problems are obtained in a classical way, after integration by parts, and are not described here.

2.2.3.1. *2D horizontal model.* For a given function $\alpha \geq 0$ and a given domain ω, we consider the weighted Hilbert space

$$L^2(\omega, \alpha) = \left\{ q;\ \int_\omega q^2 \alpha \, \mathrm{d}\omega < \infty \right\},$$

with the norm $\|q\|_{0,\omega,\alpha} = (\int_\omega q^2 \alpha \, \mathrm{d}\omega)^{1/2}$. It goes the same way for $H^1(\omega, \alpha)$ and $\mathbf{H}(\mathrm{div}, curl; \omega, \alpha)$.

The 2D horizontal model is written on the 2D domain $\Sigma(t) \subset \Sigma$ and is obtained under the assumption that the riverbed is described by $z = Z_B(x, y)$ with $Z_B \in W^{2,\infty}(\Sigma)$. For simplicity, we suppose that $\Gamma_I(t)$ is vertical. The projection spaces are obtained by specifying the dependence on z as follows

$$\begin{aligned} M_H(t) &= \left\{ q;\ q(x,y,z) = (H-z)Q(x,y),\ Q \in L^2(\Sigma(t), h^3) \right\}, \\ \mathbf{X}_H(t) &= \Big\{ (\mathbf{v}_H, v_3)^t;\ \mathbf{v}_H \in \mathbf{H}(\mathrm{div}, curl; \Sigma(t), h),\ \mathbf{v}_H \cdot \nabla Z_B \in H^1(\Sigma(t), h),\ \mathbf{v}_H \cdot \mathbf{n}_H = 0 \text{ on } \Upsilon_I(t), \\ &\qquad v_3(x,y,z) = \mathbf{v}_H \cdot \nabla Z_B + (z - Z_B) V_3(x,y),\ V_3 \in H^1(\Sigma(t), h^3) \cap L^2(\Sigma(t), h) \Big\} \end{aligned}$$

where we have put $\mathbf{v}_H(x,y) = (v_1, v_2)^t$ and where $\mathbf{n}_H$ is the outward normal unit vector to $\Upsilon_I(t)$. Thus, the vertical velocity and the pressure are affine with respect to z.

This choice guarantees a conforming approximation with respect to the initial 3D model. In particular, the condition $\mathbf{v} \cdot \mathbf{n} = 0$ on $\Gamma_B(t)$ is satisfied by construction of v_3, since a normal vector to $\Gamma_B(t)$ is $(\partial_1 Z_B, \partial_2 Z_B, -1)^t$.

The unknowns of the 2D horizontal model are h, P_H, $\mathbf{u}_H = (u_1, u_2)^t$ and U_3, all independent of z. The water depth satisfies the approximated free surface equation:

$$\frac{\partial h}{\partial t} + \mathbf{u}_H \cdot \nabla h = h U_3 \quad \text{on } \Sigma,$$

which allows to define the computational domain $\Sigma^n = \Sigma(t^n)$.

2.2.3.2. *Curvilinear coordinates.* The 2D vertical and 1D models are written in curvilinear coordinates, in order to better take into account the geometry of the river. I present next the geometrical and physical framework.

Let $\mathcal{C}(t) \subset \mathbb{R}^3$ the median curve of the free surface $\Gamma_S(t)$ and let $\mathcal{C}$ its projection on $\Sigma \subset \mathbb{R}^2$. We admit that the curve $\mathcal{C}$ is independent of time, smooth and described by $\varphi : I = [s_0, s_1] \to \mathcal{C}$ with s the curvilinear abscissa. In each point $\varphi(s) \in \mathcal{C}$, there exists the Frenet orthonormal basis $\{\boldsymbol{\tau}(s), \boldsymbol{\nu}(s)\}$ in the plane Σ. In the sequel, we employ the three-dimensional orthonormal basis $\{\boldsymbol{\tau}(s), \boldsymbol{\nu}(s), \mathbf{e}_3\}$ and we denote the associated curvilinear coordinates by $\{s, l, z\}$. It is useful to introduce the curvature $r = r(s)$ of $\mathcal{C}$, as well as the mid-width of the river $L = L(s,t)$.

The derivation of the 2D vertical and 1D models is achieved under the assumptions:

(H1) $h = h(t; s)$ and $Z_B = Z_B(s)$.

This implies that $L = L(s)$ and hence, L is given by the bathymetry.

(H2) the data $r,\ Z_B,\ L$ satisfy:

$$0 < L_0 \leq L \leq L_1,$$
$$r \in W^{1,\infty}(I), \quad Z_B \in W^{1,\infty}(I), \quad L \in W^{2,\infty}(I).$$

(H3) $1 - lr$ is of constant sign for all $l \in [-L, L]$, let's say positive to fix the ideas. Moreover, D_1 and $\frac{D_1}{r^2}$ are bounded from above, where $D_1(s) = \frac{1}{r} \ln \left| \frac{1+Lr}{1-Lr} \right| - 2L$.

The first hypothesis means that we are working with approximations of the bathymetry Z_B and of the water depth h, such that the transversal section of the river is rectangular for any s. For an easier presentation, we also asume that $\Gamma_I(t)$ is vertical and orthogonal to $\boldsymbol{\tau}$.

The 3D domain $\Omega(t)$ can be characterized as follows

$$\Omega(t) = \left\{(s, l, z);\ s \in I,\ -L(s) < l < L(s),\ Z_B(s) < z < H(t; s)\right\}.$$

We finally recall some results concerning the expression of some differential operators in curvilinear coordinates. Let $\mathbf{M} \in \Omega(t)$ be an arbitrary point such that $\mathbf{M} = \varphi(s) + l\boldsymbol{\nu}(s) + z\mathbf{e}_3$. Then by means of the Frenet formulae it follows that $\mathrm{d}\Omega = ((1 - lr)\mathrm{ds}, \mathrm{dl}, \mathrm{dz})^t$. For a scalar function f and a vector function $\mathbf{v} = (v_1, v_2, v_3)^t$, one has in the local basis $\{\boldsymbol{\tau}(s), \boldsymbol{\nu}(s), \mathbf{e}_3\}$:

$$\begin{aligned}
\mathrm{grad} f &= \left(\frac{1}{1 - lr} \partial_s f,\ \partial_l f,\ \partial_z f \right)^t, \\
\mathrm{grad}\mathbf{v} &= \begin{pmatrix} \dfrac{\partial_s v_1 - r v_2}{1 - lr} & \partial_l v_1 & \partial_z v_1 \\ \dfrac{\partial_s v_2 + r v_1}{1 - lr} & \partial_l v_2 & \partial_z v_2 \\ \dfrac{\partial_s v_3}{1 - lr} & \partial_l v_3 & \partial_z v_3 \end{pmatrix}, \\
\mathrm{div}\mathbf{v} &= \frac{\partial_s v_1 - r v_2}{1 - lr} + \partial_l v_2 + \partial_z v_3.
\end{aligned}$$

2.2.3.3. *2D vertical model.* The 2D vertical model is written on the longitudinal median surface of the river, which is mapped *via* the application φ into the vertical plane domain

$$\omega(t) = \left\{(s, z);\ s \in I,\ Z_B(s) < z < H(t; s)\right\}.$$

Its boundaries are defined as follows:

$$\begin{aligned}
\gamma_B &= \left\{(s, z);\ s \in I,\ z = Z_B(s)\right\}, \\
\gamma_I(t) &= \left\{(s, z);\ s \in \partial I,\ Z_B(s)\ < z < H(t; s)\right\}, \\
\gamma_S(t) &= \left\{(s, z);\ s \in I,\ z = H(t; s)\right\}.
\end{aligned}$$

At each t^n, we denote the computational domain $\omega(t^n)$ by ω^n and we put $\partial\omega^n = \gamma_B \cup \gamma_I^n \cup \gamma_S^n$ with obvious notations. Similarly to the 3D case, it is useful to introduce a fixed 2D maximal domain, see Fig. 2.2.1 (b), containing $\omega(t)$ and defined by

$$\omega = \left\{(s, z);\ s \in I,\ Z_B(s) < z < H_{\max}\right\}.$$

We construct the following subspaces of $M(t)$, respectively $\mathbf{X}(t)$ by specifying the dependence of their functions on l:

$$\begin{aligned}
M_V(t) &= \left\{q(s, z);\ q \in L^2(\omega(t))\right\}, \\
\mathbf{X}_V(t) &= \left\{ \left((1 - lr)v_1, \frac{lL'}{L} v_1, v_3 \right)^t \in \mathbf{X}(t);\ \ \mathbf{v}_V(s, z) = (v_1, v_3)^t,\ v_3(\cdot, Z_B) = v_1(\cdot, Z_B) Z_B' \right\}.
\end{aligned}$$

They are completely characterized by q and $\mathbf{v}_V = (v_1, v_3)^t$ respectively, which depend only on the variables $(s, z) \in \omega(t)$.

By writing the free surface equation (2.2.2) in curvilinear coordinates and by taking the velocity field in $\mathbf{X}_V(t)$, we deduce thanks to **(H1)** that:

$$\frac{\partial h}{\partial t} + u_1 H' = u_3 \quad \text{on } I.$$

As regards the time-discretization, we have chosen to solve at each t^n:

$$\frac{h - h^{n-1}}{\Delta t} + u_1^{n-1} H' = u_3^{n-1} \quad \text{on } I. \tag{2.2.11}$$

Then we solve (2.2.10) in ω^n, the unknowns $\mathbf{u}_V = (u_1, u_3)^t$ and p_V being independent of l.

REMARK 2.2.5. In the particular case of a channel (i.e. null curvature and constant width), one can show that the 2D vertical model consists of the free boundary equation together with:

$$\begin{aligned} \mathrm{div}\mathbf{u}_V &= 0 && \text{in } \omega(t) \\ \frac{d\mathbf{u}_V}{dt} + \nu \mathrm{curl}(curl\mathbf{u}_V) + \frac{c_B}{L}\mathbf{u}_V + \nabla(p_V - \frac{1}{2}|\mathbf{u}_V|^2) &= \mathbf{g} && \text{in } \omega(t), \end{aligned}$$

so it is very close to the 2D incompressible Navier-Stokes equations with a friction term.

2.2.3.4. *1D model.* We now derive the 1D model from the 2D vertical one under the additional hypothesis $Z_B \in W^{2,\infty}(I)$, by specifying the dependence on z as follows:

$$\begin{aligned} M_{1D}(t) &= \left\{ q \in M_V(t);\ q(s,z) = (H-z)Q(s),\ Q \in L^2(I) \right\}, \\ \mathbf{X}_{1D}(t) &= \Big\{ (v_1(1-lr), \frac{lL'}{L} v_1, v_3)^t \in \mathbf{X}_V(t);\ \mathbf{v}_{1D}(s) = (v_1, V_3)^t, \\ &\qquad v_3(s,z) = v_1 Z_B' + (z - Z_B)V_3 \Big\}. \end{aligned}$$

We also need the affine set $M_{1D}^*(t) = p_S + M_{1D}(t)$; note that both $M_{1D}^*(t)$ and $M_V(t)$ contain the hydrostatic pressure $p_{\mathrm{hyd}} = p_S + g(H - z)$, which is usually taken as approximation for the pressure in the classical shallow water approach.

The third component of the velocity v_3 is taken affine with respect to z, such that the boundary condition $\mathbf{v} \cdot \mathbf{n} = 0$ on the bottom holds. Thus, the elements of $\mathbf{X}_{1D}(t)$ are determined by a vector function $\mathbf{v}_{1D}(s) = (v_1, V_3)^t$, contrarily to most of the existing 1D models where the velocity is a scalar function $v_1(s)$.

The unknowns of the 1D model are $\mathbf{u}_{1D} = (u_1, U_3)^t$, P_{1D} and h and they depend only of s. The equation satisfied by h is now:

$$\frac{\partial h}{\partial t} + u_1 H' = u_1 Z_B' + h U_3 \quad \text{on } I.$$

The main point is that the previous choices of projection spaces ensure that:

$$\begin{aligned} &M_{1D}(t) \subset M_V(t) \subset M(t), && \mathbf{X}_{1D}(t) \subset \mathbf{X}_V(t) \subset \mathbf{X}(t) \\ &M_{1D}(t) \subset M_H(t) \subset M(t), && \mathbf{X}_{1D}(t) \subset \mathbf{X}_H(t) \subset \mathbf{X}(t). \end{aligned} \tag{2.2.12}$$

This hierarchy of the hydrodynamic models is important for their adaptive coupling and also yields a unified framework for the error analysis.

2.2.4. Well-posedness of time-discretized models. For the three previous choices of $\mathbf{X}_d(t^n)$ and $M_d(t^n)$, I have established:

THEOREM 2.2.6. *The approximated problem (2.2.10) has at least one solution. The uniqueness holds under a hypothesis similar to the 3D case.*

Proof. We follow the proof of Theorem 2.2.3. For each derived model, the condition $A_1(\mathbf{v};\mathbf{v},\mathbf{v})=0$ on $\mathbf{X}_d(t^n)\subset\mathbf{X}(t^n)$ is trivial. Let

$$\mathbf{V}_d(t^n) \;=\; \mathrm{Ker_d}B=\{\mathbf{v}\in\mathbf{X}_d(t^n);\ B(q,\mathbf{v})=0,\ \forall q\in M_d(t^n)\}\,.$$

It is then sufficient to prove the weak-continuity of $A_1(\cdot;\cdot,\mathbf{v})$ on $\mathbf{V}_d(t^n)$, as well as:

$$\inf_{q\in M_d(t^n)}\sup_{\mathbf{v}\in\mathbf{X}_d(t^n)}\frac{B(q,\mathbf{v})}{\|\mathbf{v}\|_{\mathbf{X}(t^n)}\|q\|_{M(t^n)}}\geq c, \tag{2.2.13}$$

$$A_0(\mathbf{v},\mathbf{v})\geq c'\,\|\mathbf{v}\|^2_{\mathbf{X}(t^n),\triangle t}\,,\quad \forall\mathbf{v}\in\mathbf{V}_d(t^n), \tag{2.2.14}$$

which is done in the next paragraphs. ■

Next, in order to derive error bounds between the 3D model and any of its previous approximations, I have adapted a result of **[50]** based on the implicit function theorem. For simplicity of presentation, let $p_S=0$ for the error analysis and let the space $\mathbf{Y}(t^n)=\mathbf{X}(t^n)\times M(t^n)$, endowed with the norm

$$\|(\mathbf{v},q)\|_{\mathbf{Y}(t^n)}=\Big(\|\mathbf{v}\|^2_{\mathbf{X}(t^n),\triangle t}+\Delta t\,\|q\|^2_{M(t^n)}\Big)^{1/2}.$$

The 3D nonlinear problem (2.2.6) can be written under the following form:

$$\mathcal{F}(\mathbf{u},p)=\mathcal{I}(\mathbf{u},p)-\mathcal{L}(\mathbf{f}-\mathcal{G}(\mathbf{u}))=\mathbf{0},$$

where $\mathbf{f}\in\mathbf{L}^{4/3}(\Omega^n)$ with $\mathbf{f}=\mathbf{g}+\frac{1}{\Delta t}\mathbf{u}^{n-1}$ for us, $\mathcal{I}:\mathbf{Y}(t^n)\to\mathbf{Y}(t^n)$ is the identity operator, $\mathcal{G}:\mathbf{X}(t^n)\to\mathbf{L}^{4/3}(\Omega^n)$ is the nonlinear operator

$$\mathcal{G}(\mathbf{v}) \;=\; \mathrm{curl}\mathbf{v}\times\mathbf{v},\quad\forall\mathbf{v}\in\mathbf{X}(t^n)$$

and $\mathcal{L}:\mathbf{L}^{4/3}(\Omega^n)\to\mathbf{Y}(t^n)$ is the linear operator associating with any $\mathbf{b}$ the unique solution $(\bar{\mathbf{u}},\bar{p})$ of the mixed variational problem:

$$\left\{\begin{array}{l}(\bar{\mathbf{u}},\bar{p})\in\mathbf{X}(t^n)\times M(t^n)\\ \forall\mathbf{v}\in\mathbf{X}(t^n),\quad A_0(\bar{\mathbf{u}},\mathbf{v})+B(\bar{p},\mathbf{v})=\langle\mathbf{b},\mathbf{v}\rangle_{\mathbf{L}^{4/3}(\Omega^n),\mathbf{L}^4(\Omega^n)}\\ \forall q\in M(t^n),\quad B(q,\bar{\mathbf{u}})=0.\end{array}\right. \tag{2.2.15}$$

Similarly, let $\mathcal{L}_d:\mathbf{L}^{4/3}(\Omega^n)\to\mathbf{X}_d(t^n)\times M_d(t^n)$ the linear operator associated with the approximation of (2.2.15) on $\mathbf{X}_d(t^n)\times M_d(t^n)$ and let $(\bar{\mathbf{u}}_d,\bar{p}_d)=\mathcal{L}_d(\mathbf{b})$. Then the lower-dimensional problem (2.2.10) can be written as follows, with $\mathbf{f}_d=\mathbf{g}+\frac{1}{\Delta t}\mathbf{u}_d^{n-1}$:

$$\mathcal{F}_d(\mathbf{u}_d,p_d)=\mathcal{I}(\mathbf{u}_d,p_d)-\mathcal{L}_d(\mathbf{f}_d-\mathcal{G}(\mathbf{u}_d))=\mathbf{0}.$$

The Babuška-Brezzi theorem, whose hypotheses were checked in Theorems 2.2.3 and 2.2.6, yields:

THEOREM 2.2.7. *The linear problem (2.2.15) and its approximations on $\mathbf{X}_d(t^n)\times M_d(t^n)$ are well-posed. Moreover, the following stability properties hold for any $\mathbf{b}\in\mathbf{L}^{4/3}(\Omega^n)$:*

$$\|\mathcal{L}(\mathbf{b})\|_{\mathbf{Y}(t^n)}\leq c\,c_1(\Omega^n)\,\|\mathbf{b}\|_{\mathbf{L}^{4/3}(\Omega^n)}\,,\qquad \|\mathcal{L}_d(\mathbf{b})\|_{\mathbf{Y}(t^n)}\leq c\,c_1(\Omega^n)\,\|\mathbf{b}\|_{\mathbf{L}^{4/3}(\Omega^n)}\,. \tag{2.2.16}$$

If $\mathbf{b}\in\mathbf{L}^2(\Omega^n)$ then one also has:

$$\|\mathcal{L}(\mathbf{b})\|_{\mathbf{Y}(t^n)}\leq c\sqrt{\Delta t}\,\|\mathbf{b}\|_{0,\Omega^n}\,,\qquad \|\mathcal{L}_d(\mathbf{b})\|_{\mathbf{Y}(t^n)}\leq c\sqrt{\Delta t}\,\|\mathbf{b}\|_{0,\Omega^n}\,. \tag{2.2.17}$$

Following the classical proof of error estimates for mixed variational problems (see for instance **[47]**, p. 54), one can next establish, with c independent of $\triangle t$ and Ω^n, that:

$$\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_d\|_{\mathbf{X}(t^n),\triangle t} \leq c\left(\inf_{\mathbf{w}\in\mathbf{V}_d(t^n)}\|\bar{\mathbf{u}}-\mathbf{w}\|_{\mathbf{X}(t^n),\triangle t}+\inf_{q\in M_d(t^n)}\|\bar{p}-q\|_{M(t^n)}\right)$$

$$\leq c\left(\inf_{\mathbf{v}\in\mathbf{X}_d(t^n)}(\|\bar{\mathbf{u}}-\mathbf{v}\|_{\mathbf{X}(t^n),\triangle t}+\tfrac{1}{\sqrt{\Delta t}}\sup_{q\in M_d(t^n)}\tfrac{B(q,\bar{\mathbf{u}}-\mathbf{v})}{\|q\|_{M(t^n)}})+\inf_{q\in M_d(t^n)}\|\bar{p}-q\|_{M(t^n)}\right),$$

$$\|\bar{p}-\bar{p}_d\|_{M(t^n)} \leq \tfrac{c}{\sqrt{\Delta t}}\left(\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_d\|_{\mathbf{X}(t^n),\triangle t}+(1+\sqrt{\Delta t})\inf_{q\in M_d(t^n)}\|\bar{p}-q\|_{M(t^n)}\right).$$

We can summarize the previous error bounds in the next theorem:

THEOREM 2.2.8. *The following* a priori *error bound holds true for any* $\mathbf{b}\in\mathbf{L}^{4/3}(\Omega^n)$*:*

(2.2.18)

$$\|(\mathcal{L}-\mathcal{L}_d)(\mathbf{b})\|_{\mathbf{Y}(t^n)} \leq c(\inf_{\mathbf{v}\in\mathbf{X}_d(t^n)}(\|\bar{\mathbf{u}}-\mathbf{v}\|_{\mathbf{X}(t^n),\triangle t}+\frac{1}{\sqrt{\Delta t}}\sup_{q\in M_d(t^n)}\frac{B(q,\bar{\mathbf{u}}-\mathbf{v})}{\|q\|_{M(t^n)}})+\inf_{q\in M_d(t^n)}\|\bar{p}-q\|_{M(t^n)}).$$

REMARK 2.2.9. Due to the hierarchy of the hydrodynamic models, similar error bounds hold between any of the 2D and the 1D models.

In the sequel, I establish the conditions (2.2.13) and (2.2.14) for each model. Note that the norms on $\mathbf{X}_d(t^n)$ used for the uniform coercivity and the uniform *inf-sup* condition are different.

2.2.4.1. *2D horizontal model.* A simple computation gives

$$B(q,\mathbf{v})=-\int_{\Sigma^n}\frac{h^2}{2}Q(\mathrm{div}\mathbf{v}_H+V_3)\mathrm{dxdy},\quad \forall(\mathbf{v},q)\in\mathbf{X}_H(t^n)\times M_H(t^n),$$

therefore

$$\mathbf{V}_H(t^n)=\{\mathbf{v}\in\mathbf{X}_H(t^n);\ \mathrm{div}\mathbf{v}_H+V_3=0 \text{ in } \Sigma^n\}\subset\mathbf{V}(t^n).$$

This inclusion trivially yields (2.2.14), so we only have to check the *inf-sup* condition.

I treat here only the case where the water depth is strictly positive and satisfies $h\in W^{1,\infty}(\Sigma)$, which implies certain simplifications: the 2D computational domain is independent of time since $\Sigma(t)=\Sigma$, and it is not necessary to work in weighted spaces since the projection subspaces are also independent of time:

$$\begin{aligned} M_H &= \left\{q;\ q=(H-z)Q,\ Q\in L^2(\Sigma)\right\},\\ \mathbf{X}_H &= \left\{(\mathbf{v}_H,v_3)^t;\ \mathbf{v}_H\in\mathbf{H}(\mathrm{div},curl;\Sigma),\ \mathbf{v}_H\cdot\nabla Z_B\in H^1(\Sigma),\ \mathbf{v}_H\cdot\mathbf{n}_H=0 \text{ on } \Upsilon_I,\right.\\ &\quad \left. v_3=\mathbf{v}_H\cdot\nabla Z_B+(z-Z_B)V_3,\ V_3\in H^1(\Sigma)\right\}. \end{aligned}$$

REMARK 2.2.10. A proof for the degenerate case where h vanishes on the lateral boundary Υ_{lat} can be found in [B4], based on a technical result of **[44]** concerning the regularity of a degenerate elliptic problem in weighted Sobolev spaces. Nevertheless, the hypotheses on the domain Σ^n are too restrictive in view of the finite element discretization.

LEMMA 2.2.11. *Suppose that* $0<h_{\min}\leq h\leq h_{\max}$*. Then (2.2.13) holds on* $\mathbf{X}_H\times M_H$*, with a constant* c *proportional to* $\sqrt{\frac{h_{\min}}{h_{\max}}}$*.*

Proof. For a given $q=(H-z)Q\in M_H$, we consider $\mathbf{v}_H\in\mathbf{H}^1_0(\Sigma)$ such that

$$-\mathrm{div}\mathbf{v}_H=hQ-\frac{1}{|\Sigma|}\int_\Sigma hQ\,\mathrm{dxdy}\quad\text{in }\Sigma.$$

We put $V_3=-\frac{1}{|\Sigma|}\int_\Sigma hQ\,\mathrm{dxdy}$, we define the operator

$$\mathcal{R}:M_H\to\mathbf{X}_H,\quad \mathcal{R}q=(\mathbf{v}_H,\,\mathbf{v}_H\cdot\nabla Z_B+(z-Z_B)V_3)^t$$

and we check that

$$B(q,\mathcal{R}q)=\frac{3}{2}\left\|q\right\|^2_{M(t^n)}\,,\qquad \left\|\mathcal{R}q\right\|_{\mathbf{X}(t^n)}\leq c\sqrt{h_{\max}}\left\|hQ\right\|_{0,\Sigma}\leq c\sqrt{\frac{h_{\max}}{h_{\min}}}\left\|q\right\|_{M(t^n)}$$

which implies the desired statement. For more details, see [B4]. ■

2.2.4.2. *2D vertical model.* In the next two paragraphs, the differential operators applied to $\mathbf{v}_V$ are the classical ones with respect to the variables s and z, that is:

$$\mathrm{div}\mathbf{v}_V=\partial_s v_1+\partial_z v_3,\quad curl\mathbf{v}_V=\partial_s v_3-\partial_z v_1.$$

For simplicity of writing, it is useful to introduce the coefficients:

$$D_1(s)=\int_{-L}^{L}\frac{lr}{1-lr}\,\mathrm{dl}=\frac{1}{r}\ln\left|\frac{1+Lr}{1-Lr}\right|-2L,\qquad D_2(s)=r^2L^2+\frac{(L')^2}{3}. \tag{2.2.19}$$

Remark 2.2.12. In the particular case of null curvature, one has:

$$D_1(s)=0,\quad D_2(s)=\frac{(L')^2}{3},\quad \lim_{r\to 0}\frac{D_1}{r^2}=\frac{2L^3}{3}.$$

In order to understand which constraints the inclusion $\mathbf{X}_V(t^n)\subset\mathbf{X}(t^n)$ imposes on the 2D function $\mathbf{v}_V$ associated with a 3D test-function $\mathbf{v}\in\mathbf{X}_V(t^n)$, we compute:

$$\begin{aligned}
\int_{\Omega^n}|\mathbf{v}|^2\,\mathrm{d}\Omega &= 2\int_{\omega^n}L\left((1+D_2)v_1^2+v_3^2\right)\mathrm{dz}\,\mathrm{ds},\\
\int_{\Omega^n}|\mathrm{curl}\mathbf{v}|^2\,\mathrm{d}\Omega &= \int_{\omega^n}\left(2L(curl\mathbf{v})^2+2LD_2(\partial_z v_1)^2+8Lr^2v_1^2+D_1(\partial_s v_3)^2\right)\mathrm{dz}\,\mathrm{ds}\\
&\quad+\int_{\omega^n}\frac{D_1}{r^2}(\partial_s(\frac{L'v_1}{L}))^2\mathrm{dz}\,\mathrm{ds},\\
\int_{\Omega^n}|\mathrm{div}\mathbf{v}|^2\,\mathrm{d}\Omega &= \int_{\omega^n}\left(2L(\mathrm{div}\mathbf{v}_V+\frac{L'}{L}v_1)^2+D_1\frac{((Lr)')^2}{(Lr)^2}v_1^2\right)\mathrm{dz}\,\mathrm{ds}.
\end{aligned}$$

I have then shown in [B3], thanks to **(H2)** and **(H3),** that $\mathbf{v}\in\mathbf{X}(t^n)$ if and only if $\mathbf{v}_V$ satisfies

$$\mathbf{v}_V\in\mathbf{H}(\mathrm{div},curl;\omega^n)\text{ and }\sqrt{D_2}\partial_z v_1,\ \sqrt{D_1}\partial_s v_3,\ \frac{\sqrt{D_1}}{r}\frac{(Lr)'}{Lr}v_1,\ \frac{\sqrt{D_1}}{r}\partial_s(\frac{L'v_1}{L})\in L^2(\omega^n).$$

A sufficient condition is $\mathbf{v}_V\in\mathbf{H}^1(\omega^n)$. Then one has that $\left\|\mathbf{v}\right\|_{\mathbf{X}(t^n)}\leq c\left\|\mathbf{v}_V\right\|_{1,\omega^n}$ with c depending on the data r, L and Z_B. It is useful to note that:

$$\begin{aligned}
B\left(q,\mathbf{v}\right) &= -2\int_{\omega^n}q\mathrm{div}(L\mathbf{v}_V)\,\mathrm{dz}\,\mathrm{ds},\quad\forall q\in M_V(t^n),\ \forall\mathbf{v}\in\mathbf{X}_V(t^n)\\
\mathbf{V}_V(t^n) &= \left\{\mathbf{v}\in\mathbf{X}_V(t^n);\ \mathrm{div}(L\mathbf{v}_V)=0\text{ in }\omega^n\right\}.
\end{aligned}$$

I have then shown the following two lemmas.

Lemma 2.2.13. *The nonlinear form $A_1(\cdot;\cdot,\mathbf{v})$ is sequentially weakly-continuous on $\mathbf{V}_V(t^n)$, for all $\mathbf{v}\in\mathbf{V}_V(t^n)$. Moreover, one has:*

$$A(\mathbf{v};\mathbf{v},\mathbf{v})=A_0(\mathbf{v},\mathbf{v})\geq c\left\|\mathbf{v}\right\|^2_{\mathbf{X}(t^n),\triangle t},\quad\forall\mathbf{v}\in\mathbf{V}_V(t^n).$$

Proof. The proof of the first assertion is similar to the 3D case, except that now the elements of $\mathbf{V}_V(t^n)$ are not divergence free; we use then their particular form as elements of $\mathbf{X}_V(t^n)$. As regards the coercivity, we note that

$$\mathrm{div}\mathbf{v}=-\frac{l(Lr)'}{L(1-lr)}v_1\quad\text{in }\Omega^n,\quad\forall\mathbf{v}\in\mathbf{V}_V(t^n)$$

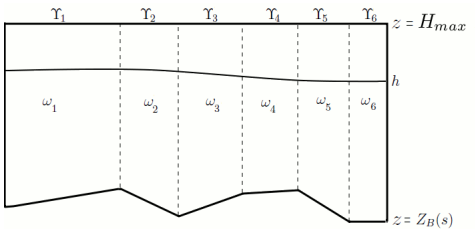


FIGURE 2.2.2. Example of partition of ω in convex subdomains

so one can finally bound the missing divergence term in $A_0(\mathbf{v},\mathbf{v})$ as follows:

$$\|\mathrm{div}\mathbf{v}\|^2_{0,\Omega^n} \leq c\,\|v_1\|^2_{0,\omega^n} \leq cA_0(\mathbf{v},\mathbf{v})$$

with c depending only on $L,\ r$. This yields the uniform coercivity with respect to Δt and ω^n. ■

LEMMA 2.2.14. *Assume that ω admits a finite partition in convex subdomains $\omega = \cup_{i=1}^N \omega_i$ as in Fig.2.2.2. Then (2.2.13) holds on $\mathbf{X}_V(t^n) \times M_V(t^n)$.*

Proof. We make the change of variables $\bar{\mathbf{v}}_V = L\mathbf{v}_V$. Thanks to **(H2)**, $\mathbf{v}_V \in \mathbf{H}^1(\omega^n)$ is equivalent to $\bar{\mathbf{v}}_V \in \mathbf{H}^1(\omega^n)$ and their $\mathbf{H}^1(\omega^n)$-norms are equivalent. We denote by γ_i the top boundary of ω_i, for all $i = 1, ..., N$. For any $q \in M_V(t^n)$, we denote by $\tilde{q}$ its extension by zero on the whole domain ω.

On each subdomain ω_i, we consider the auxiliary problems:

$$\left\{\begin{array}{ll} \Delta\phi_i = -\tilde{q} & \text{in } \omega_i \\ \partial_n\phi_i = 0 & \text{on } \partial\omega_i \setminus \gamma_i \\ \phi_i = 0 & \text{on } \gamma_i \end{array}\right. , \qquad \left\{\begin{array}{ll} \Delta^2\psi_i = 0 & \text{in } \omega_i \\ \partial_n\psi_i = -\partial_t\phi_i & \text{on } \partial\omega_i \\ \psi_i = 0 & \text{on } \partial\omega_i \end{array}\right. .$$

The regularity of the Laplace operator in ω_i ensures that $\phi_i \in H^2(\omega_i)$ and $\|\phi_i\|_{2,\omega_i} \leq c\,\|\tilde{q}\|_{0,\omega_i}$. One next deduces from the biharmonic problem that $\psi_i \in H^2(\omega_i)$ and $\|\psi_i\|_{2,\omega_i} \leq c\,\|\partial_t\phi_i\|_{1/2,\partial\omega_i} \leq c\,\|\tilde{q}\|_{0,\omega_i}$. By putting $\bar{\mathbf{v}}^i_V = \mathrm{grad}\phi_i + \mathrm{curl}\psi_i$, one next has that:

$$\begin{array}{rl} \bar{\mathbf{v}}^i_V \in \mathbf{H}^1(\omega_i), & \left\|\bar{\mathbf{v}}^i_V\right\|_{1,\omega_i} \leq c\,\|\tilde{q}\|_{0,\omega_i}\,, \\ \mathrm{div}\bar{\mathbf{v}}^i_V = -\tilde{q} \text{ in } \omega_i, & \bar{\mathbf{v}}^i_V \cdot \mathbf{n} = \bar{\mathbf{v}}^i_V \cdot \mathbf{t} = 0 \text{ on } \partial\omega_i \setminus \gamma_i. \end{array}$$

Next, we define $\tilde{\mathbf{v}}_V = (\tilde{v}_1, \tilde{v}_3)^t \in \mathbf{H}^1(\omega)$ by its restriction to each subdomain $(\tilde{\mathbf{v}}_V)_{/\omega_i} = \bar{\mathbf{v}}^i_V$, we put $\tilde{\mathbf{v}} = \frac{1}{L}\left((1-lr)\tilde{v}_1, \frac{lL'}{L}\tilde{v}_1, \tilde{v}_3\right)^t$ and we finally consider its restriction $\mathbf{v}$ to ω^n. It is then obvious that:

$$\begin{array}{rcl} \|\mathbf{v}\|_{\mathbf{X}(t^n)} & \leq & c_1\,\|\tilde{\mathbf{v}}_V\|_{1,\omega} \leq c_2\,\|\tilde{q}\|_{0,\omega} \leq c\,\|q\|_{M(t^n)}\,, \\ B(q,\mathbf{v}) & = & \|q\|^2_{0,\omega^n} \geq c\,\|q\|^2_{M(t^n)}\,, \end{array}$$

with a constant c independent of $\triangle t$ and ω^n. This ends the proof. ■

2.2.4.3. *1D model.* Thanks to the hypothesis **(H1)**, the water depth h is bounded by strictly positive constants. Then one easily deduces from the study of the 2D vertical model that $(\mathbf{v}, q) \in \mathbf{X}_{1D}(t^n) \times M_{1D}(t^n)$ if and only if $\mathbf{v}_{1D} = (v_1, V_3)^t \in H^1_0(I) \times H^1(I)$ and $Q \in L^2(I)$. A simple computation yields that

$$\begin{array}{rcl} B(q,\mathbf{v}) & = & -\displaystyle\int_I h^2 Q\left((Lv_1)' + LV_3\right)\mathrm{ds}, \\ \mathbf{V}_{1D}(t^n) & = & \left\{\mathbf{v} \in \mathbf{X}_{1D}(t^n);\ (Lv_1)' + LV_3 = 0 \text{ in } I\right\} \subset \mathbf{V}_V(t^n), \end{array}$$

so Lemma 2.2.13 obviously holds on $\mathbf{V}_{1D}(t^n)$.

LEMMA 2.2.15. *The inf-sup condition (2.2.13) holds on* $\mathbf{X}_{1D}(t^n) \times M_{1D}(t^n)$.

Proof. Let any $q \in M_{1D}(t^n)$. We define the function $\bar{\mathbf{v}}_{1D} = (\bar{v}_1, \bar{V}_3) \in \mathbf{H}^1(I)$ by

$$\bar{V}_3 \;=\; -\frac{1}{|I|}\int_I hQ\,\mathrm{ds}, \quad \bar{v}_1(\theta) = -\int_{s_0}^{\theta}(hQ + \bar{V}_3)\mathrm{ds}, \quad \forall \theta \in I.$$

We conclude by considering next the 3D vector function of $\mathbf{X}_{1D}(t^n)$ associated with $\frac{1}{L}\bar{\mathbf{v}}_{1D}$. More details are given in [B3]. ■

2.2.5. Finite element approximation. The space discretization of each lower-dimensional model is achieved by conforming finite elements. A generic discrete model is simply written as follows:

$$\left\{\begin{array}{l} (\mathbf{u}_a, p_a) \in \mathbf{X}_a(t^n) \times M_a(t^n) \\ \forall \mathbf{v} \in \mathbf{X}_a(t^n), \quad A(\mathbf{u}_a; \mathbf{u}_a, \mathbf{v}) + B(p_a, \mathbf{v}) = F_a^{n-1}(\mathbf{v}) \\ \forall q \in M_a(t^n), \quad B(q, \mathbf{u}_a) = 0 \end{array}\right. \tag{2.2.20}$$

where $\mathbf{X}_a(t^n)$ and $M_a(t^n)$ are finite dimensional spaces satisfying $\mathbf{X}_a(t^n) \times M_a(t^n) \subset \mathbf{X}_d(t^n) \times M_d(t^n)$. The weak formulation (2.2.20) is equivalent to

$$\mathcal{F}_a(\mathbf{u}_a, p_a) = \mathcal{I}(\mathbf{u}_a, p_a) - \mathcal{L}_a(\mathbf{f}_a - \mathcal{G}(\mathbf{u}_a)) = \mathbf{0},$$

with $\mathcal{L}_a : \mathbf{L}^{4/3}(\Omega^n) \to \mathbf{X}_a(t^n) \times M_a(t^n)$ the linear operator associated with the discrete version of (2.2.15).

Then one can carry out the same analysis as previously, under the sole conditions (2.2.13) and (2.2.14), with constants which are now independent of both the space and time discretization. Error estimates for $\mathcal{L} - \mathcal{L}_a$ and $\mathcal{L}_d - \mathcal{L}_a$ are derived exactly as for $\mathcal{L} - \mathcal{L}_d$.

In what follows, I present a choice of finite dimensional spaces for each model.

2.2.5.1. *2D horizontal model.* We consider only the non-degenerate case and assume moreover that Σ is polygonal. Let $(\mathcal{T}_a)_{a>0}$ be a regular family of triangulations of Σ consisting of triangles. The continuity equation is written in conservative form and its space discretization is achieved by a vertex-centered finite volume scheme, combined with a mass lumping technique. The water depth is approximated by P_1-continuous elements on each $K \in \mathcal{T}_a$. We introduce the spaces:

$$\begin{array}{rcl} \mathbf{X}_H^a & = & \left\{\mathbf{v} \in \mathbf{X}_H;\ (\mathbf{v}_H, V_3)^t \in \mathbf{H}^1(\Sigma),\ \forall K \in \mathcal{T}_a,\ (\mathbf{v}_H)_{/K} \in \mathbf{P}_1,\ (V_3)_{/K} \in P_1 \oplus \mathcal{B}_K\right\}, \\ M_H^a & = & \left\{q \in M_H;\ \forall K \in \mathcal{T}_a,\ Q_{/K} \in P_0\right\} \end{array}$$

with $\mathcal{B}_K = \mathrm{span}\{b_K\}$ the space of bubble functions on K. We also replace $B(\cdot,\cdot)$ by

$$B_a(q, \mathbf{v}) = -\frac{1}{2}\int_\Sigma (\pi^a h^a)^2 Q(\mathrm{div}\mathbf{v}_H + V_3)\mathrm{dxdy}, \quad \forall (q, \mathbf{v}) \in M_H^a \times \mathbf{X}_H^a$$

where the discrete water depth h^a is substituted by its L^2-orthogonal projection $\pi^a h^a$ on M_H^a.

I have then established in [B4] a preliminary result which allows to replace the weight h^a by $\pi^a h^a$ in the norms $\|\cdot\|_{\mathbf{X}(t^n)}$ and $\|\cdot\|_{M(t^n)}$, on the discrete spaces $\mathbf{X}_H^a$ and M_H^a.

LEMMA 2.2.16. *Let any $K \in \mathcal{T}_a$, $\mathcal{P}$ a polynomial space and ζ a linear and strictly positive function on K. Then there exist $c_1,\ c_2 > 0$ independent of ζ and K such that:*

$$\forall v \in \mathcal{P}, \quad c_1 \int_K \zeta\,|v|\ \mathrm{dxdy} \le \int_K \pi_0^K \zeta\,|v|\ \mathrm{dxdy} \le c_2 \int_K \zeta\,|v|\ \mathrm{dxdy}.$$

A similar result holds when replacing ζ, $\pi_0^K\zeta$ by ζ^m, $(\pi_0^K\zeta)^m$ respectively, for given $m \in \mathbb{N}^*$.

In order to prove the well-posedness of the discrete problem, we assume in what follows that

$$\exists \sigma_0, \sigma_1 > 0 \text{ such that } \forall K \in \mathcal{T}_a, \quad \sigma_0 \leq \frac{\pi_0^K h^a}{|K|^{1/2}} \leq \sigma_1. \tag{2.2.21}$$

REMARK 2.2.17. One can associate with $\mathcal{T}_a$ a 3D triangulation $\mathcal{T}_{3D}$, consisting of one layer of prisms (of basis $K \in \mathcal{T}_a$ and height $\pi_0^K h^a$). Then the assumption (2.2.21) translates the fact that $\mathcal{T}_{3D}$ is regular cf. **[58]**, and is not so restrictive under the shallow water assumption.

THEOREM 2.2.18. *Under the hypothesis (2.2.21), the conditions (2.2.13) and (2.2.14) hold on $\mathbf{X}_H^a \times M_H^a$, with c proportional to $\sqrt{\frac{h^a_{\min}}{h^a_{\max}}}$ and c' independent of the water depth h^a.*

Proof. In order to prove the coercivity, it suffices to bound $\int_\Sigma h^a (\mathrm{div}\mathbf{v}_H + V_3)^2 \mathrm{dxdy}$ for $\mathbf{v}$ in

$$\mathbf{V}_H^a = \left\{ \mathbf{v} \in \mathbf{X}_H^a;\ \forall K \in \mathcal{T}_a,\ \int_K (\mathrm{div}\mathbf{v}_H + V_3)\mathrm{dxdy} = 0 \right\} \not\subseteq \mathbf{V}_H.$$

Since $\mathrm{div}\mathbf{v}_H = -\pi_0^K V_3$ on any $K \in \mathcal{T}_a$, it follows that

$$\int_K (\pi_0^K h^a)(\mathrm{div}\mathbf{v}_H + V_3)^2 \mathrm{dxdy} \leq c\sigma_0 (\pi_0^K h^a)^3 \left|V_3\right|^2_{1,K}$$

which leads to

$$\int_\Sigma h^a (\mathrm{div}\mathbf{v}_H + V_3)^2 \mathrm{dxdy} \leq c \left\|\mathrm{curl}\mathbf{v}\right\|^2_{0,\Omega^n}, \quad \forall \mathbf{v} \in \mathbf{V}_H^a.$$

Concerning the uniform *inf-sup* condition for $B_a(\cdot,\cdot)$, let any $q = (Z_B + h^a - z)Q \in M_H^a$. Then, according to Lemma 2.2.16, one has that

$$c_1 \int_\Sigma (h^a)^3 Q^2 \mathrm{dxdy} \leq \int_\Sigma (\pi^a h^a)^3 Q^2 \mathrm{dxdy} \leq c_2 \int_\Sigma (h^a)^3 Q^2 \mathrm{dxdy},$$

hence $(\pi^a h^a)Q \in L^2(\Sigma, h)$ and $\|(\pi^a h^a)Q\|_{0,\Sigma,h}$ and $\|q\|_{M(t^n)}$ are equivalent. Following the proof of Lemma 2.2.11, we associate with $(\pi^a h^a)Q$ a function $\mathbf{v} \in \mathbf{X}_H$ satisfying $\mathrm{div}\mathbf{v}_H + V_3 = -(\pi^a h^a)Q$ and

$$\|\mathbf{v}\|_{0,\Omega^n} + \|\mathrm{div}\mathbf{v}\|_{0,\Omega^n} + \|\mathrm{curl}\mathbf{v}\|_{0,\Omega^n} + \|\mathbf{v}\|_{0,\Gamma_B^n} \leq c\sqrt{\tfrac{h_{\max}}{h_{min}}} \|q\|_{M(t^n)},$$
$$B_a(q,\mathbf{v}) = \tfrac{1}{2}\int_\Sigma (\pi^a h^a)^3 Q^2 \mathrm{dxdy} = \tfrac{3}{2}\|q\|^2_{M(t^n)}.$$

We next construct the discrete functions $(\mathbf{v}_H^a, V_3^a)$ by taking $\mathbf{v}_H^a$ as the Clément interpolate (cf. **[59]**) of $\mathbf{v}_H$ and

$$\forall K \in \mathcal{T}_a, \quad (V_3^a)_{/K} = \alpha_K b_K \quad \text{with} \quad \alpha_K = \frac{\int_K \mathrm{div}(\mathbf{v}_H - \mathbf{v}_H^a)\mathrm{dxdy}}{\int_K b_K \mathrm{dxdy}}.$$

This choice ensures that the corresponding 3D function $\mathbf{v}^a$ belongs to $\mathbf{X}_H^a$ and $B_a(q,\mathbf{v}^a) = B_a(q,\mathbf{v})$. I have shown in [B4], using Lemma 2.2.16, hypothesis (2.2.21) and interpolation error bounds, the remaining estimate $\|\mathbf{v}^a\|_{\mathbf{X}(t^n)} \leq c\|q\|_{M(t^n)}$ with c proportional to $h^a_{\max}/h^a_{\min}$. ■

2.2.5.2. *2D vertical model.* We suppose here that ω^n is polygonal, which means that Z_B and $Z_B + h^a$ are piecewise linear and continuous. Let $(\mathcal{T}_a)_{a>0}$ be a regular family of triangulations of ω^n consisting of triangles. The free surface equation (2.2.11) is discretized as previously and we introduce the spaces

$$\begin{aligned} M_V^a(t^n) &= \left\{ q \in M_V(t^n);\ q \in H^1(\omega^n),\ q_{/\Upsilon_S^n} = 0,\ q_{/K} \in P_1\ \forall K \in \mathcal{T}_a \right\}, \\ \mathbf{X}_V^a(t^n) &= \left\{ \mathbf{v} \in \mathbf{X}_V(t^n);\ \mathbf{v}_V \in \mathbf{H}^1(\omega^n),\ (L\mathbf{v}_V)_{/K} \in (P_1 \oplus \mathcal{B}_K)^2\ \forall K \in \mathcal{T}_a \right\}. \end{aligned}$$

For any $\mathcal{T}_a$, it is useful to introduce a triangulation $\tilde{\mathcal{T}}_a$ of the whole domain ω such that $\tilde{\mathcal{T}}_a \supset \mathcal{T}_a$ and to consider the following finite element spaces on ω:

$$\begin{aligned} \tilde{M}^a(\omega) &= \left\{ q \in H^1(\omega);\ q_{/\Upsilon_S} = 0,\ q_{/K} \in P_1\ \forall K \in \tilde{\mathcal{T}}_a \right\}, \\ \tilde{\mathbf{X}}^a(\omega) &= \left\{ \mathbf{w} \in \mathbf{H}^1_{0,\Upsilon}(\omega);\ \mathbf{w}_{/K} \in (P_1 \oplus \mathcal{B}_K)^2\ \forall K \in \tilde{\mathcal{T}}_a \right\} \end{aligned}$$

where $\mathbf{H}^1_{0,\gamma}(\omega) = \left\{ \mathbf{w} \in \mathbf{H}^1(\omega);\ \mathbf{w} \cdot \mathbf{n}_{/\gamma_I \cup \gamma_B} = 0 \right\}$. Then I have established:

LEMMA 2.2.19. *Condition (2.2.13) holds uniformly on* $\mathbf{X}_V^a(t^n) \times M_V^a(t^n)$.

Proof. The proof uses Lemma 2.2.14 together with the fact that the above MINI finite elements are *inf-sup* stable for the 2D Stokes problem. Nevertheless, special care has to be taken of the non-standard boundary conditions and of the independence of the domain ω^n. With $q \in M_V^a(t^n)$ we associate $\tilde{q}$ and $\tilde{\mathbf{v}}$ as in Lemma 2.2.14, and we note that $\tilde{q} \in \tilde{M}^a(\omega)$ thanks to the boundary condition on the free surface Υ_S^n imposed in $M_V^a(t^n)$. We consider next (see for instance **[91]**, p. 175) the interpolation operator $\mathcal{I}_a : \mathbf{H}^1_{0,\gamma}(\omega) \to \tilde{\mathbf{X}}^a(\omega)$, continuous with respect to the $H^1(\omega)$-norm, defined by the relations:

$$\begin{aligned} \int_K \mathcal{I}_a \mathbf{w}\, \mathrm{dzds} &= \int_K \mathbf{w}\, \mathrm{dzds}, \quad \forall K \in \tilde{\mathcal{T}}_a \\ (\mathcal{I}_a \mathbf{w})(N) &= (\mathcal{R}_a \mathbf{w})(N), \quad \forall \text{ node } N \text{ of } \tilde{\mathcal{T}}_a \end{aligned}$$

where $\mathcal{R}_a$ is a local regularization operator of Clément type (see **[59]** or **[31]**).

Then one obtains, by using the $H^1(\omega)$-conformity of both $\tilde{M}^a(\omega)$ and $\tilde{\mathbf{X}}^a(\omega)$, the boundary conditions and the fact that gradξ is piecewise constant, that

$$\int_\omega \xi \mathrm{div}\,(\mathbf{w} - \mathcal{I}_a \mathbf{w})\, \mathrm{dz\,ds} = 0, \quad \forall \xi \in \tilde{M}^a(\omega).$$

Starting from $\tilde{\mathbf{v}}_V^a = \mathcal{I}_a \tilde{\mathbf{v}}_V \in \tilde{\mathbf{X}}^a(\omega)$, we construct the corresponding 3D vector field $\tilde{\mathbf{v}}^a$ and finally we consider its restriction $\mathbf{v}^a$ to ω^n. We clearly have $\mathbf{v}^a \in \mathbf{X}_V^a(t^n)$, $B(q, \mathbf{v}^a) = \|q\|^2_{0,\omega^n}$ and

$$\|\mathbf{v}^a\|_{\mathbf{X}(t^n)} \le c_1 \|\mathcal{I}_a \tilde{\mathbf{v}}_V\|_{1,\omega} \le c_2 \|\tilde{\mathbf{v}}_V\|_{1,\omega} \le c_3 \|q\|_{0,\omega^n}$$

which yields the announced statement. ∎

In order to obtain the uniform coercivity on

$$\mathbf{V}_V^a(t^n) = \left\{ \mathbf{v} \in \mathbf{X}_V^a(t^n);\ \int_{\omega^n} q \mathrm{div}(L\mathbf{v}_V)\, \mathrm{dzds} = 0,\ \forall q \in M_V^a(t^n) \right\},$$

we have replaced in the discrete problem $A_0(\cdot,\cdot)$ by $A_\beta(\cdot,\cdot) = A_0(\cdot,\cdot) + \beta A_2(\cdot,\cdot)$, where β is a stabilization parameter independent of the discretization and

$$A_2(\mathbf{u}, \mathbf{v}) = \int_{\omega^n} \frac{1}{L} \mathrm{div}(L\mathbf{u}_V) \mathrm{div}(L\mathbf{v}_V)\, \mathrm{dz\,ds}.$$

Then one has for any $\mathbf{v} \in \mathbf{V}_V^a(t^n)$ that

$$\begin{aligned}\int_{\Omega^n} |\mathrm{div}\mathbf{v}|^2 \,\mathrm{d}\Omega &= \int_{\omega^n} \tfrac{2}{L} \left(\mathrm{div} L\mathbf{v}_V\right)^2 \mathrm{d}z\,\mathrm{d}s + \int_{\omega^n} D_1 \tfrac{\left((Lr)'\right)^2}{L^2 r^2} \left(v_1\right)^2 \mathrm{d}z\,\mathrm{d}s \\ &\leq c\left(A_2\left(\mathbf{v},\mathbf{v}\right) + \|\mathbf{v}\|_{0,\Omega^n}^2\right),\end{aligned}$$

which leads to

$$A_\beta\left(\mathbf{v},\mathbf{v}\right) \geq c\,\|\mathbf{v}\|_{\mathbf{X}(t^n),\triangle t}^2, \quad \forall \mathbf{v} \in \mathbf{V}_V^a(t^n).$$

2.2.5.3. *1D model.* Let now $\left(\mathcal{T}_a\right)_{a>0}$ a regular family of subdivisions of the interval I and let

$$\begin{aligned}M_{1D}^a(t^n) &= \left\{q \in M_{1D}(t^n);\ \forall K \in \mathcal{T}_a,\ Q_{/K} \in P_0\right\}, \\ \mathbf{X}_{1D}^a(t^n) &= \left\{\mathbf{v} \in \mathbf{X}_{1D}(t^n);\ \forall K \in \mathcal{T}_a,\ \left(L\mathbf{v}_{1D}\right)_{/K} \in \left(P_1\right)^2\right\}.\end{aligned}$$

It is important to note that h^a is strictly positive, and is assumed to be uniformly bounded with respect to the discretization.

Concerning the discrete coercivity, one can then show that

$$\int_{\Omega^n} |\mathrm{div}\mathbf{v}|^2 \,\mathrm{d}\Omega \leq c A_0(\mathbf{v},\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_{1D}^a(t^n).$$

The discrete *inf-sup* condition is established by following the proof of the continuous case and by applying the 1D version of Lemma 2.2.16. More details can be found in [B3].

2.2.6. *A priori* and *a posteriori* error analysis. In order to study the error between the time-discretized models (2.2.6) and (2.2.20), we assume in what follows that the exact solution $(\mathbf{u}, p)$ of the 3D model (2.2.6) is well approximated in the finite dimensional spaces $\mathbf{X}_a(t^n) \times M_a(t^n)$, that is:

at each time-step t^n, there exists $\delta^n > 0$ sufficiently small (independent of Δt) such that

$$(\mathbf{A}) \qquad \inf_{\mathbf{v}\in\mathbf{X}_a(t^n)} \left(\|\mathbf{u}-\mathbf{v}\|_{\mathbf{X}(t^n),\triangle t} + \frac{1}{\sqrt{\Delta t}} \sup_{q\in M_a(t^n)} \frac{B(q,\mathbf{u}-\mathbf{v})}{\|q\|_{M(t^n)}}\right) + \inf_{q\in M_a(t^n)} \|p-q\|_{M(t^n)} \leq \delta^n.$$

Remark 2.2.20. By changing correspondingly the previous assumption, a similar *a priori* and *a posteriori* analysis can be carried out between (2.2.6) and any hydrodynamic model (2.2.10), or still between (2.2.10) and its finite element approximation (2.2.20). In the first case, only the modeling error is considered while in the second one, only the classical discretization error is studied.

Let the error related to the time-discretization $\epsilon^{n-1} = \frac{1}{\sqrt{\Delta t}} \left\|\mathbf{u}^{n-1} - \mathbf{u}_a^{n-1}\right\|_{0,\Omega^n}$.

2.2.6.1. *Error bounds.* In order to derive an *a priori* error estimate, we adapt a result of Brezzi, Rappaz and Raviart **[50]** (see also **[150]**) to our multiscale approximation. We first prove:

Theorem 2.2.21. *Assume **(A)** and that $D\mathcal{F}(\mathbf{u},p) \in \mathrm{Isom}(\mathbf{Y}(t^n))$. Then the nonlinear mapping $\mathcal{F}_a$ fulfils the conditions:*

***(C1)** there exists $c > 0$ such that for any $(\mathbf{v},q) \in \mathbf{Y}(t^n)$:*

$$\|D\mathcal{F}_a(\mathbf{u},p) - D\mathcal{F}_a(\mathbf{v},q)\|_{\mathcal{L}(\mathbf{Y}(t^n))} \leq c\,c_1^2(\Omega^n)\,\|(\mathbf{u},p)-(\mathbf{v},q)\|_{\mathbf{Y}(t^n)}$$

***(C2)** there exists $c > 0$ such that*

$$\|\mathcal{F}_a(\mathbf{u},p)\|_{\mathbf{Y}(t^n)} \leq c\left(\delta^n + \epsilon^{n-1}\right).$$

Then according to **[150]** the next statement holds true:

THEOREM 2.2.22. *Assume* ***(A)*** *and that* $D\mathcal{F}(\mathbf{u},p) \in \mathrm{Isom}(\mathbf{Y}(t^n))$. *For* δ^n *and* ϵ^{n-1} *small enough, the following a priori estimate holds, with a constant* c *independent of* Δt*:*

$$\|(\mathbf{u},p)-(\mathbf{u}_a,p_a)\|_{\mathbf{Y}(t^n)} \leq c\,\|\mathcal{F}_a(\mathbf{u},p)\|_{\mathbf{Y}(t^n)} \leq c\left(\delta^n+\epsilon^{n-1}\right).$$

We next establish *a posteriori* error bounds. This can be achieved following **[150]**, under the assumption that **(A)** holds true for any element $(\mathbf{w},r)$ of $\mathbf{Y}(t^n)$. This hypothesis seems to be rather strong in view of the dependence on Δt of the $\mathbf{Y}(t^n)$-norm. In order to avoid it, we apply a result of Verfürth **[164]**.

THEOREM 2.2.23. *Assume* ***(A)****,* $D\mathcal{F}(\mathbf{u},p) \in \mathrm{Isom}(\mathbf{Y}(t^n))$ *and* $\mathbf{u}$, $\mathrm{curl}\mathbf{u} \in \mathbf{L}^\infty(\Omega^n)$. *If* δ^n, ϵ^{n-1} *and* Δt *are small enough, then there exist two constants* c_1, c_2 *such that:*

$$\frac{1}{c_1}\,\|\mathcal{F}(\mathbf{u}_a,p_a)\|_{\mathbf{Y}(t^n)} \leq \|(\mathbf{u},p)-(\mathbf{u}_a,p_a)\|_{\mathbf{Y}(t^n)} \leq c_2\,\|\mathcal{F}(\mathbf{u}_a,p_a)\|_{\mathbf{Y}(t^n)}.$$

2.2.6.2. *Residual-based a posteriori error estimators.* Thanks to the variational framework and to the hierarchy of models, we can now define at each t^n generic error estimators between the 3D model (2.2.6) and any of its lower-dimensional finite element approximations (2.2.20). They indicate the validity domain of the employed model, from a qualitative point of view, and they measure both the discretization and the modeling errors. We have already applied this idea in [A3], in order to couple some linearized 1D and 2D vertical models.

We consider three sections Φ of Σ^n, S of I and $\Psi = S\times(Z_B,H)$ of ω^n and we construct a 3D subdomain $\Theta \subset \Omega^n$ by putting

$$\Theta = \{(x,y,z);\ (x,y)\in\Phi,\ Z_B(x,y)<z<H(x,y)\}$$

for the 2D horizontal model, respectively

$$\Theta = \{(s,l,z);\ s\in S, -L(s)<l<L(s),\ Z_B(s)<z<H(s)\}$$

for the 2D vertical and 1D models. We denote by $\mathcal{T}_{3D}$ the 3D triangulation of Θ induced by the respective triangulations of Φ, Ψ or S, and by $\mathcal{E}_{3D}$ the set of faces.

It is useful to recall that thanks to the hypothesis (2.2.4), the injection operator $\mathcal{I}_0$: $(\mathbf{X}(t),\|\cdot\|_{\mathbf{X}(t)}) \to (\mathbf{H}^s(\Omega(t)),\|\cdot\|_{s,\Omega(t)})$ with $s\geq 3/4$ is continuous, of norm $c_0(\Omega(t))$. We assume moreover that there exists a projection operator $\mathcal{P}_a : \mathbf{X}(t^n)\to\mathbf{X}_a(t^n)$ which satisfies:

$$\sum_{T\in\mathcal{T}_{3D}} w_T\,\|\mathbf{v}-\mathcal{P}_a\mathbf{v}\|_{0,T} + \sum_{\gamma\in\mathcal{E}_{3D}} w_\gamma\,\|\mathbf{v}-\mathcal{P}_a\mathbf{v}\|_{0,\gamma} \leq C\,|\mathbf{v}|_{s,\Omega^n},\quad \forall\mathbf{v}\in\mathbf{X}(t^n). \tag{2.2.22}$$

The weights w_T and w_γ are model-dependent due to the anisotropy of the 3D domain Θ and their computation is described in [B2]. For instance, for the 1D model the 3D cell $T\in\mathcal{T}_{3D}$ is a hexahedron of dimensions h, $2L$ and d_T, which yields $w_T=(d_T+2L+h)^{-s}$.

Next, we introduce the residuals on a given cell $T\in\mathcal{T}_a$:

$$\begin{aligned}\eta_1 &= -\mathrm{div}\mathbf{u}_a,\\ \eta_2 &= \frac{1}{\Delta t}(\mathbf{u}_a-\mathbf{u}_a^n)+(\mathrm{curl}\mathbf{u}_a\times\mathbf{u}_a)+\nu\mathrm{curl}(\mathrm{curl}\mathbf{u}_a)-\mathbf{f}\times\mathbf{u}_a+\nabla p_a-\mathbf{g},\end{aligned}$$

and η_γ on a given face $\gamma\in\mathcal{E}_{3D}$. On an internal face, we simply have $\eta_\gamma=\nu\,[\mathrm{curl}\mathbf{u}_a]\times\mathbf{n}-[p_a]\,\mathbf{n}$ whereas on $\partial\Theta$ we take into account the boundary conditions. We define the error estimator on Θ by:

$$\eta(\Theta)^2 = \sum_{T\in\mathcal{T}_{3D}}\Big(\frac{1}{\Delta t}\int_T\eta_1^2\,\mathrm{d}\Omega+\frac{1}{w_T^2}\int_T\eta_2^2\,\mathrm{d}\Omega\Big)+\sum_{\gamma\in\mathcal{E}_{3D}}\frac{1}{w_\gamma^2}\int_\gamma\eta_\gamma^2\,\mathrm{d}\Gamma \tag{2.2.23}$$

and we prove in [B2], using classical tools and Theorem 2.2.23, its reliability on Ω^n.

THEOREM 2.2.24. *Assume (2.2.22) and the hypotheses of Theorem 2.2.23. Then*

$$\|(\mathbf{u},p)-(\mathbf{u}_a,p_a)\|_{\mathbf{Y}(t^n)} \leq c\left((1+c_0(\Omega^n))\,\eta(\Omega^n)+\epsilon^{n-1}\right).$$

2.2.7. Numerical validation. The previous hydrodynamic models have been implemented in the C++ library LIBMESH. The numerical tests were mainly carried out by David Trujillo, and some of them by Agnès Petrau. Realistic simulations of the Adour river using data provided by IFREMER and comparisons with measured velocities were also carried out. For the sake of completeness, I show next some results in order to validate the proposed models.

2.2.7.1. *1D model.* We compare the 1D model and the shallow water equations in an academic configuration. Then in order to show the effects of the hydrodynamic pressure on the circulation pattern, we compare both models with an analytical solution in the small amplitude wave test.

Comparison with 1D shallow water equations

The classical 1D shallow water equations are written in a rectilinear channel in the Ox direction and obtained by integrating the 3D Navier-Stokes equations over a transversal section. They are deduced under the hydrostatic pressure assumption and by neglecting the vertical and the transversal velocities, and they are usually written under the following form:

$$\begin{aligned}\frac{\partial\sigma}{\partial t}+\operatorname{div}(\sigma u_m) &= 0\\ \frac{\partial(\sigma u_m)}{\partial t}+(\sigma u_m^2)'+J &= -\sigma g(Z_B+h)'\end{aligned} \tag{2.2.24}$$

where J is a friction term, modeled in practice by empirical formulae (Manning-Strickler, Chézy etc.). The unknowns are the mean velocity $u_m(t;x)$ and $\sigma(t;x)$.

In order to compare the two 1D models, we consider a rectilinear channel with flat bottom and constant width. After some computations, we obtain exactly the same continuity equation, while the momentum equation of u_1 in our model is:

$$\frac{\partial(\sigma u_1)}{\partial t}+(\sigma u_1^2)'+\frac{2c_B}{\rho}(h+L)\,u_1+\frac{\sigma h'}{\rho}(P-\rho g)+\frac{\sigma}{2\rho}\left(hP'-\rho(u_1^2)'-\rho\frac{h^2}{3}(U_3^2)'\right)=-\sigma g h'.$$

In this simple geometrical configuration, the differences between the two models are only related to the friction term and to the non-hydrostatic pressure, whose influence will be discussed further.

We illustrate next the similarity of the two models in the case of a flat bottom rectilinear channel. The details of the test are given in [B3]. One can see in Fig. 2.2.3 that the evolution in time of the water depth is very similar for both models. A flat bottom meander was also considered, in order to see the influence of the curvature; as expected, a slight difference appeared now between the two water depths.

Comparison with analytical solution

We perform the small amplitude wave test (cf. **[105]** or **[111]**), for which an analytical solution is known with non-hydrostatic presure. Water is confined in a closed basin with flat bottom, with a square base of length 10 m and with an equilibrium depth of $H = 5\,\mathrm{m}$. The friction coefficient c_B is null and the viscosity is $\mu = 10^{-3}\,\mathrm{Pa}\cdot\mathrm{s}$. A zero initial velocity is imposed and the initial free surface elevation is given by $h(x) = A\cos(kx)$ for $0 < x < 10$, where $k = \frac{\pi}{10}$ and where $A = 0.1\,\mathrm{m}$ is the wave amplitude (1% of the water depth, such that small amplitude wave theory applies). The wave celerity c is then computed according to the relationship $c = [(g/k)\tanh(kH)]^{\frac{1}{2}}$ and equals 5.35 m/s, while the period of oscillation is 3.74 s. In order to highlight the influence of the hydrostatic pressure, we plot in Fig. 2.2.4 the water elevation at $x = 0$ during several periods of oscillation, computed with the 1D model, the shallow water model and analytically. We note that our wave speed is in very good agreement with the

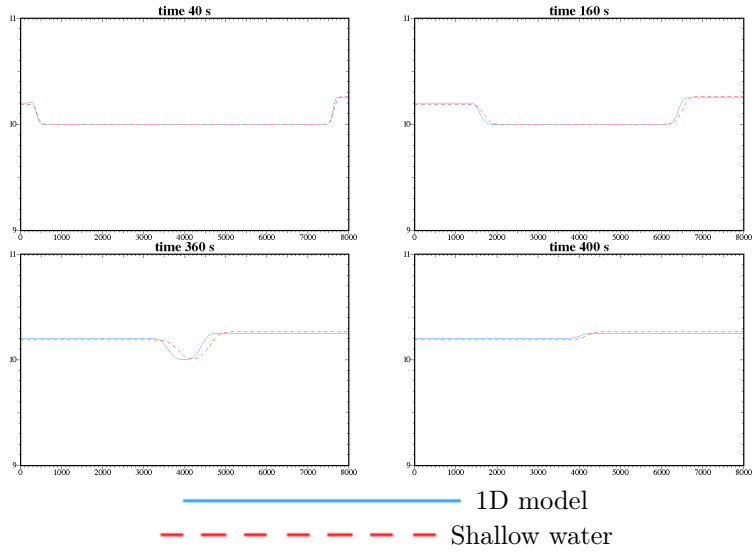


FIGURE 2.2.3. Water depth: 1D comparison in a rectilinear channel

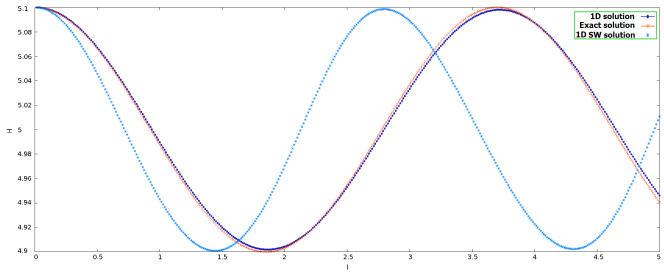


FIGURE 2.2.4. Comparison between 1D model, shallow water and analytical solution

analytical one. This test shows the superiority of our approach with respect to the shallow water model, as regards the propagation of small amplitude waves.

2.2.7.2. *2D vertical model.* We test the 2D vertical model in a rectilinear channel with an irregular bottom, in order to highlight the effect of the vertical velocity. Then, as for the 1D case, we compare it with the shallow water model and with the analytical solution in the small amplitude wave test.

Influence of the vertical velocity

The curvature r is null and the width is constant, 10 m; the length is 80 m while the initial free surface elevation is about 4 m. A null discharge is imposed downstream and an inflow condition is given upstream, sufficiently large with respect to σ such that the velocities are quite important and hence, the flow is not laminar. The velocity profiles obtained at different t are shown in Fig. 2.2.5, scaled by a factor 10 in the vertical direction. The recirculation zone appearing downstream the first step is well represented by the model, which justifies the interest of using the 2D vertical model in such a configuration. We also show in Fig. 2.2.6 the difference between our computed pressure and the hydrostatic one, and see that it is not negligeable.

Comparison with shallow water and analytical solution

We carry out the previous small amplitude wave test but we now employ the 2D vertical model. We plot in Fig. 2.2.7 the water depth at $x = 0$ computed by the three models. Again, our wave speed is in very good agreement with the analytical one, but this is not the case for the shallow water model. Our velocity and the analytical one also have almost identical profiles.

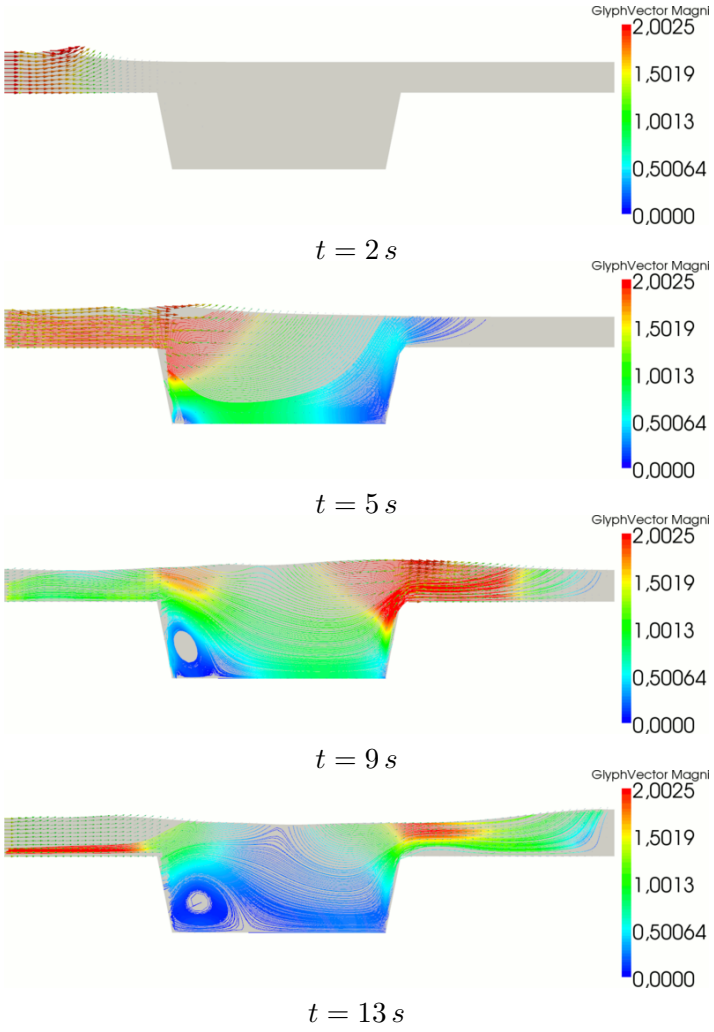


FIGURE 2.2.5. Scaled velocity profiles in a channel with irregular bottom

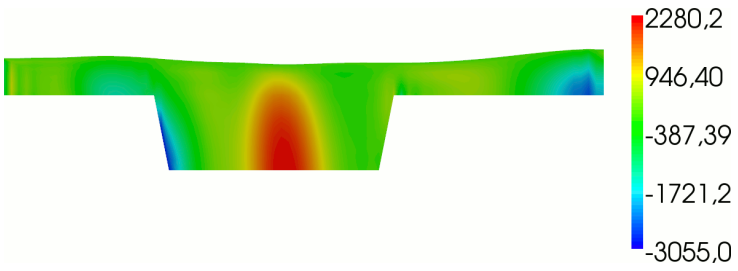


FIGURE 2.2.6. Difference between the pressure given by the 2D vertical model and the hydrostatic pressure at $t = 13\,\mathrm{s}$ (scaled by 10 in the Oz direction)

2.2.7.3. *2D horizontal model.* We are interested here in the comparison of the 2D horizontal model with the classical shallow water equations, which are:

$$\begin{aligned} \frac{\partial h}{\partial t} + \operatorname{div}(h\mathbf{u}_m) &= 0, \\ \frac{\partial (h\mathbf{u}_m)}{\partial t} + \operatorname{div}(h\mathbf{u}_m \otimes \mathbf{u}_m) - \gamma\Delta(h\mathbf{u}_m) + gh\mathbf{J} - hf\mathbf{u}_m^{\perp} &= -gh\nabla(Z_B + h). \end{aligned}$$

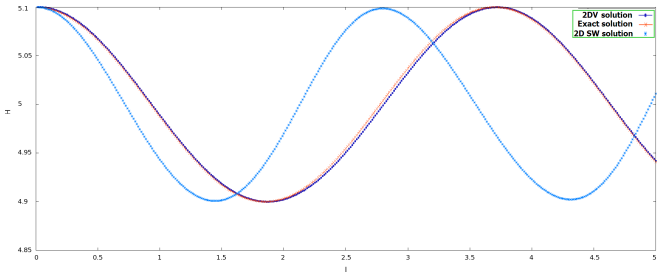
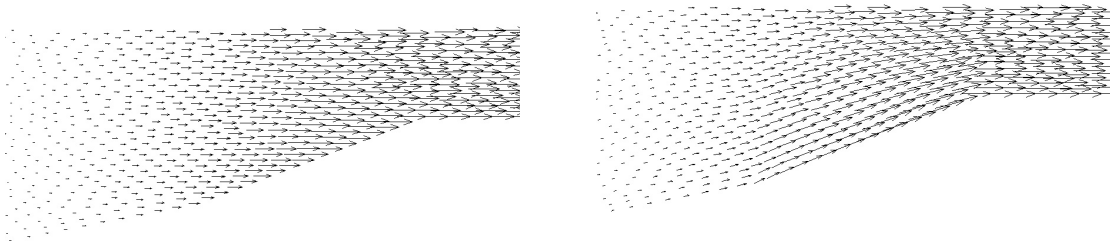


FIGURE 2.2.7. Comparison between 2D vertical model, 2D shallow water and analytical solution



a) Shallow water model b) Our model

FIGURE 2.2.8. Zoom of longitudinal velocity in a rectilinear channel with varying bottom

Here above, γ is a viscosity coefficient (that certain models do not take into account), $\mathbf{J}$ a friction term and $\mathbf{u}_m = (u_{1m}, u_{2m})^t$ represents an averaged 2D velocity. Concerning our model, we obtain the same continuity equation, whereas the momentum equation can be written, in the case of a flat bottom, as below:

$$\frac{\partial(h\mathbf{u}_H)}{\partial t} + \text{div}\,(h\mathbf{u}_H \otimes \mathbf{u}_H) + \nu \text{curl}(h\, curl\mathbf{u}_H) + c_B\mathbf{u}_H - hf\mathbf{u}_H^{\perp}$$
$$-\frac{h^3}{3}U_3\nabla U_3 + h\nabla h(P-g) + \frac{h}{2}(h\nabla P - \nabla\,|\mathbf{u}_H|^2) = -gh\nabla h.$$

As in 1D, the differences with the shallow water system result from the friction modeling, the non-hydrostatic pressure and the non-zero vertical velocity.

Numerical comparisons between the 2D horizontal and the shallow water (with Manning-Strickler friction) models were carried out in three academic configurations: a flat bottom rectilinear channel, a flat bottom nonrectilinear channel and a rectilinear channel with varying topography. We have observed very similar evolution in time of the water depth in the three cases, and also of the horizontal velocities in the case of flat bottom channels. However, our model provides a more accurate representation of the 3D velocity. We have reconstructed in Fig. 2.2.7.3 the velocity computed by the two models in the longitudinal plane (u_1, u_3). One can see that the vertical component u_3 (null in the shallow water model) is actually non negligeable, especially near the variations of the bottom and of the free surface.

2.2.8. Coupling of hydrodynamic models. We discuss here how to employ the *a posteriori* error estimators in order to develop an adaptive coupling strategy of the models, and we show an academic test which illustrates the coupling possibilities offered by this approach.

2.2.8.1. *Adaptive coupling strategy.* The idea of coupling multiscale models in computational fluid dynamics is not new. In the past years, it has been widely employed in blood flow simulations (see for instance **[86]** for the coupling of a 3D model with a 1D nonlinear hyperbolic one or **[151]** for a geometrical multiscale approach based on the coupling of ODE's and PDE's). The approach provided in **[86]** was extended to free surface flows, in particular to the coupling of 2D and 1D shallow water systems in a simple hydrodynamic configuration (see **[133]**). In these papers, the coupling is achieved through adequate matching conditions, which are not obvious to prescribe as dimensionally different quantities have to be related, and usually a domain decomposition technique is carried out to solve the coupled problem. Moreover, this coupling strategy is based on the *a priori* knowledge of the regions where different models have to be employed, which are chosen once for all at the beginning of the approximation procedure. However, there exist many physical configurations where the previous choice is not immediate or feasible.

An alternative coupling strategy based on *a posteriori* modeling error indicators is proposed in **[139]**, **[38]**. Even though the underlying idea is similar, the approach developed in **[38]** is different from ours since they are considering a dimensionally homogeneous - physically heterogeneous coupling, and the goal-oriented method cf. **[26]** is used for model adaptation. The authors obtained a coarse model from a fine one by dropping a part, possibly nonlinear, which is usually expensive to compute. Meanwhile, our "coarse" models are obtained by a projection method on adapted subspaces.

Concerning now the adaptive modeling of free-surface flows, Perotto **[146]** extended to time-dependent problems the steady-case analysis provided in **[38]** and studied the coupling of two 2D simplified Saint-Venant models, where the adapted model is obtained by neglecting the nonlinear convective term in the fine model.

Instead, we propose to achieve the coupling of the hydrodynamic models by means of the *a posteriori* error estimators defined in (2.2.23). Thanks to the hierarchy of the models (2.2.12), the transmission conditions are implicitly contained in the formulations. As a future work, the definition of the computing zones could be done automatically.

2.2.8.2. *Numerical results.* The 3D geometry considered is a section of river of 850 m length. The mid-width $L(s)$ varies between 10 and 30 m, while the bathymetry $Z_B(s)$ varies between -10 and 2 m. Both are shown in Fig. 2.2.9. As regards the curvature, it is equal to $1/400$ on the first 630 m and to $-1/400$ on the rest. The free surface level $Z_B + h$ is initially set at 4 m. We impose an inflow rate $q = 100$ on the left and $q = 50$ on the right boundary.

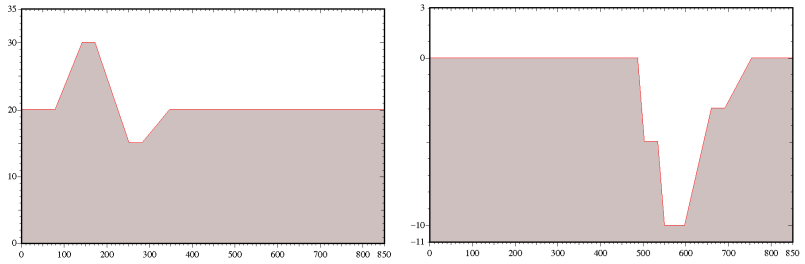


FIGURE 2.2.9. Mid-width (left) and bathymetry (right) versus curvilinear abscissa

In this example, we have first employed the 1D model on the whole domain and then computed the corresponding *a posteriori* error estimator. In order to diminish this estimator, and implicitly the modeling error, we have decomposed the initial domain in several regions, to which different models are associated cf. Fig. 2.2.10 (a). We have decided to employ a 2D horizontal model in the regions with varying width while a 2D vertical model is used in the regions with

variations of the bathymetry. This choice is next validated by the computation of the error estimators on the coupled model. We have represented in Fig. 2.2.11 (a) the velocity field coloured by the water depth in the 2D horizontal region, as well as the corresponding mesh. One can also see in Fig. 2.2.11 (b) the streamlines computed in the 2D vertical region, showing vertical effects than cannot be captured by the 1D or the 2D horizontal model. Finally, we show in Fig. 2.2.10 (b) the reconstructed 3D velocity on the free surface of the river. We have also observed that the error estimator computed with the coupled model decreased significantly and, as expected, it mainly took into account the wave motion.

For more complex examples, the 2D error estimators might be used to choose the most appropriate 2D model. In regions where both the bathymetry and the width have important variations, it is possible to couple the two 2D models in order to get a quasi 3D model, which was developed in [**147**].

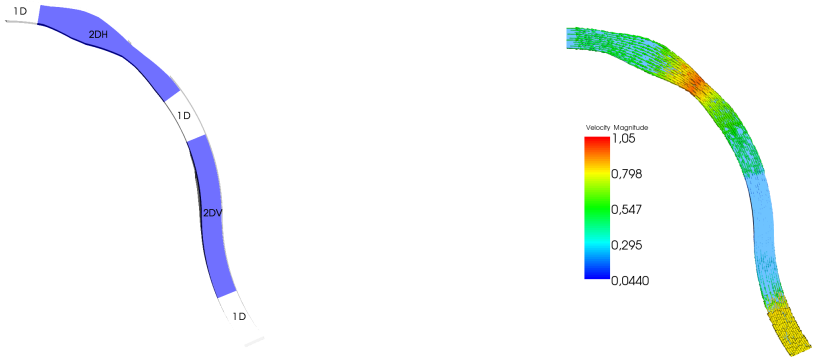


FIGURE 2.2.10. Definition of 2D regions and top view of velocity after coupling

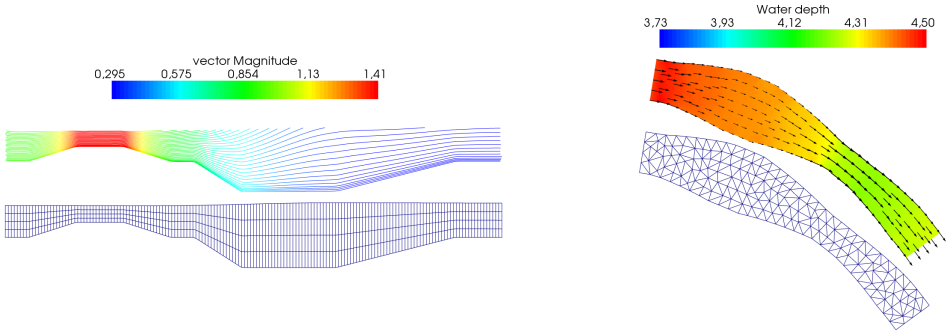


FIGURE 2.2.11. Velocity computed by 2D vertical (left) and 2D horizontal (right) models

2.3. Discontinuous Galerkin approximation of Stokes equations

The results of this section can be found in [A11] and [**24**], see also the PhD thesis of Julie Joie [**107**].

To summarize, a discontinuous Galerkin method for the Stokes equations with a different stabilization of the viscous term is analyzed. It yields the same *a priori* error estimates as the classical interior penalty method. Its main advantage is the robustness with respect to the stabilization parameter γ, since it allows us to recover as γ tends towards infinity some stable and well-known nonconforming approximations of the Stokes problem. Besides, one can define a

simple *a posteriori* error estimator, based on the reconstruction of a locally conservative $H(\mathrm{div})$-tensor. The accuracy and the robustness of the scheme are illustrated by numerical tests.

The extension to the strain-rate formulation of the Stokes problem is given in [C14] and is not detailed here.

2.3.1. Discrete formulation. We consider the 2D stationary Stokes equations

$$-\mu\Delta\mathbf{u}+\nabla p=\mathbf{f},\quad \mathrm{div}\mathbf{u}=0$$

endowed here, for simplicity of presentation, with a homogeneous Dirichlet boundary condition $\mathbf{u}=\mathbf{0}$ on $\partial\Omega$.

The classical velocity-pressure formulation is:

$$\left\{\begin{array}{lll}(\mathbf{u},p)\in\mathbf{H}^1_0(\Omega)\times L^2_0(\Omega)&&\\ a(\mathbf{u},\mathbf{v})+b(p,\mathbf{v})&=l(\mathbf{v})&\forall\mathbf{v}\in\mathbf{H}^1_0(\Omega)\\ b(q,\mathbf{u})&=0&\forall q\in L^2_0(\Omega)\end{array}\right. \tag{2.3.1}$$

where the bilinear and the linear forms are defined by

$$a(\mathbf{u},\mathbf{v})=\mu\int_\Omega \underline{\nabla}\mathbf{u}:\underline{\nabla}\mathbf{v}\,\mathrm{dx},\qquad b(p,\mathbf{v})=-\int_\Omega p\mathrm{div}\mathbf{v}\,\mathrm{dx},\qquad l(\mathbf{v})=\int_\Omega\mathbf{f}\cdot\mathbf{v}\,\mathrm{dx}.$$

In what follows, we take $k=1$, 2 or 3 and we introduce the finite dimensional spaces:

$$\begin{aligned}\mathbf{V}_h&=\left\{\mathbf{v}_h\in\mathbf{L}^2(\Omega);\ (\mathbf{v}_h)_{/T}\in\mathbf{P}_k,\ \forall T\in\mathcal{T}_h\right\},\\ Q_h&=\left\{q_h\in L^2_0(\Omega);\ (q_h)_{/T}\in P_{k-1},\ \forall T\in\mathcal{T}_h\right\}.\end{aligned}$$

A review of approximation results for $\mathbf{V}_h$ and Q_h can be found, for instance, in **[93]**. The case $k=1$ follows from **[66]**, $k=2$ from **[88]** and $k=3$ from **[67]**. The interpolation operators on $\mathbf{V}_h$ and Q_h are denoted by $\boldsymbol{I}_h$ and i_h respectively.

We introduce our new stabilization term on $(\mathbf{H}^1(\Omega)+\mathbf{V}_h)\times(\mathbf{H}^1(\Omega)+\mathbf{V}_h)$:

$$J(\mathbf{u},\mathbf{v})=\mu\sum_{e\in\varepsilon_h}\frac{1}{|e|}\int_e[\boldsymbol{\pi}_{k-1}\mathbf{u}]\cdot[\boldsymbol{\pi}_{k-1}\mathbf{v}]\mathrm{ds}. \tag{2.3.2}$$

Note that $J(\mathbf{u},\mathbf{v})=0$ for any $\mathbf{u}\in\mathbf{H}^1_0(\Omega)$.

We consider the next discontinuous Galerkin formulation of (2.3.1):

$$\left\{\begin{array}{lll}(\mathbf{u}_h^\gamma,p_h^\gamma)\in\mathbf{V}_h\times Q_h&&\\ a_h(\mathbf{u}_h^\gamma,\mathbf{v}_h)+b_h(p_h^\gamma,\mathbf{v}_h)&=l(\mathbf{v}_h)&\forall\mathbf{v}_h\in\mathbf{V}_h\\ b_h(q_h,\mathbf{u}_h^\gamma)&=0&\forall q_h\in Q_h\end{array}\right. \tag{2.3.3}$$

where the bilinear forms are defined as follows:

$$\begin{aligned}a_h(\cdot,\cdot)&=A_0(\cdot,\cdot)+A_1(\cdot,\cdot)+\gamma J(\cdot,\cdot)\\ A_0(\mathbf{u}_h,\mathbf{v}_h)&=\mu\sum_{T\in\mathcal{T}_h}\int_T\underline{\nabla}\mathbf{u}_h:\underline{\nabla}\mathbf{v}_h\mathrm{dx}\\ A_1(\mathbf{u}_h,\mathbf{v}_h)&=-\mu\sum_{e\in\varepsilon_h}\left(\int_e\{\partial_n\mathbf{u}_h\}\cdot[\mathbf{v}_h]\,\mathrm{ds}+\int_e\{\partial_n\mathbf{v}_h\}\cdot[\mathbf{u}_h]\,\mathrm{ds}\right)\\ b_h(q_h,\mathbf{v}_h)&=-\sum_{T\in\mathcal{T}_h}\int_T q_h\mathrm{div}\mathbf{v}_h\mathrm{dx}+\sum_{e\in\varepsilon_h}\int_e\{q_h\}[\mathbf{v}_h\cdot\mathbf{n}_e]\mathrm{ds}\end{aligned}$$

and where $\gamma>0$ is a stabilization parameter. We denote by $[\cdot]$ and $\{\cdot\}$ the jump, respectively the average of a piecewise continuous function on an arbitrary edge of the triangulation.

REMARK 2.3.1. The non-homogeneous case is treated by Nitsche's method [**138**], implying a modification of the righthand side terms (cf. [**24**]) and does not raise any particular difficulty.

The present stabilization can be linked to other discontinuous Galerkin methods for the Stokes equations. In [**93**], the authors applied the well-known interior penalty stabilization to the Stokes and Navier-Stokes equations:

$$J^*(\mathbf{u}_h, \mathbf{v}_h) = \mu \sum_{e \in \varepsilon_h} \frac{1}{\mid e \mid} \int_e [\mathbf{u}_h] \cdot [\mathbf{v}_h] \mathrm{ds}.$$

Another stabilization term $\int_\Omega \boldsymbol{R}([\mathbf{u}_h]) \cdot \boldsymbol{R}([\mathbf{v}_h])\mathrm{dx}$ was proposed by Bassi and Rebay in [**22**], where $\boldsymbol{R}$ is a lifting of the jumps across the edges in $\mathbf{P}_k^{disc}$. In order to improve the computational efficiency, they replace the contributions from the global operator $\boldsymbol{R}$ with a local lifting operator $\boldsymbol{R}_e$, defined by

$$\sum_{\bar{K} \supset e} \int_K \boldsymbol{R}_e(\mathbf{w}) \cdot \mathbf{v}_h \mathrm{dx} \quad = \quad \int_e \mathbf{w} \cdot \{\mathbf{v}_h\} \, \mathrm{ds}, \quad \forall \mathbf{v}_h \in \mathbf{P}_k^{disc}.$$

So finally the stabilization term is approximated by

$$J^{\#}(\mathbf{u}_h, \mathbf{v}_h) = \sum_{e \in \varepsilon_h} \sum_{\bar{K} \supset e} \gamma_K \int_K \boldsymbol{R}_e([\mathbf{u}_h]) \cdot \boldsymbol{R}_e([\mathbf{v}_h]) \mathrm{dx}.$$

By taking $\gamma_K = \gamma$ and a regular mesh, a simple computation yields that $J^{\#}(\cdot,\cdot) \simeq J^*(\cdot,\cdot) + 2J(\cdot,\cdot)$ for $k = 1$. If we look for the lifting operator in $\boldsymbol{P}_0$ instead of $\boldsymbol{P}_1$, then $J^{\#}(\cdot,\cdot) \simeq \frac{1}{2} J(\cdot,\cdot)$.

2.3.1.1. *Well-posedness.* In what follows, the constants are independent of the discretization parameter h, the viscosity μ and the stabilization parameter γ, unless it is specified otherwise. Let $|\cdot|_{1,h}$ denote the $\mathbf{H}^1(\Omega)$- broken semi-norm and let the semi-norm on $\mathbf{H}^1(\Omega) + \mathbf{V}_h$:

$$|||\mathbf{v}||| = \left(\mu |\mathbf{v}|_{1,h}^2 + \gamma J(\mathbf{v}, \mathbf{v})\right)^{1/2}.$$

Then it can be proved that $|||\cdot|||$ is a norm on $\mathbf{V}_h$ and that for γ large enough, $a_h(\cdot,\cdot)$ is uniformly coercive:

$$\forall \mathbf{v} \in \mathbf{V}_h, \quad a_h(\mathbf{v}, \mathbf{v}) \geq \alpha \, ||| \, \mathbf{v} \, |||^2 .$$

As regards the *inf-sup* condition for $b_h(\cdot,\cdot)$, the important point is that it now holds with a constant independent of γ.

LEMMA 2.3.2. *There exists a constant $\beta > 0$ such that*

$$\inf_{q \in Q_h} \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{b_h(q, \mathbf{v})}{\|q\|_{0,\Omega} \, ||| \, \mathbf{v} \, |||} \geq \frac{\beta}{\sqrt{\mu}}.$$

Proof. The idea is classical: with any $q \in Q_h \subset L_0^2(\Omega)$ we associate, cf. for instance [**91**], $\mathbf{z} \in \mathbf{H}_0^1(\Omega)$ such that $\mathrm{div}\mathbf{z} = q$ and $\|\mathbf{z}\|_{1,\Omega} \leq c\|q\|_{0,\Omega}$. Then we put $\mathbf{w} = \boldsymbol{I}_h \mathbf{z} \in \mathbf{V}_h$ and we have, thanks to the properties of $\boldsymbol{I}_h$, that $J(\mathbf{w}, \mathbf{w}) = 0$, $b_h(q, \mathbf{w}) = \|q\|_{0,\Omega}^2$ and $||| \, \mathbf{w} \, ||| \leq c\sqrt{\mu}\|q\|_{0,\Omega}$, which yields the desired statement. ∎

REMARK 2.3.3. When employing the stabilization term $J^*(\cdot,\cdot)$ of the classical Interior Penalty method, then $J^*(\mathbf{w}, \mathbf{w}) \neq 0$ and the *inf-sup* constant β behaves like $O(1/\sqrt{\gamma})$. Thus, this method is less robust for large values of γ, which has been highlighted by numerical tests.

Thanks to Babuška-Brezzi theorem (cf. [**47**]), one deduces now that the mixed problem (2.3.3) is well-posed for γ sufficiently large.

2.3.1.2. *A priori error bounds.* In order to derive optimal *a priori* error estimates, we have established the following auxiliary results. The details of the proofs can be found in [A11] and, as regards the improved convergence rate for the $L^2(\Omega)$-norm of the velocity error, in **[24]**.

LEMMA 2.3.4. *The solution $(\mathbf{u}, p)$ of (2.3.1) satisfies the consistency properties:*

$$\begin{aligned} a_h(\mathbf{u}, \mathbf{v}_h) + b_h(p, \mathbf{v}_h) &= l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in \mathbf{V}_h \\ b_h(q_h, \mathbf{u}) &= 0, \quad \forall q_h \in Q_h. \end{aligned}$$

LEMMA 2.3.5. *There exists $c > 0$ such that:*

$$\left(\sum_{e \in \varepsilon_h} \frac{1}{|e|} \|[\mathbf{v}_h]\|_{0,e}^2 \right)^{1/2} \leq \frac{c \max\{1, 1/\gamma\}}{\sqrt{\mu}} |||\mathbf{v}_h|||, \quad \forall \mathbf{v}_h \in \mathbf{V}_h.$$

LEMMA 2.3.6. *Suppose $(\mathbf{u}, p) \in \mathbf{H}^{k+1}(\Omega) \times H^k(\Omega)$ and let $\gamma \geq 1$. Then there exists $c > 0$ such that for all $\mathbf{v}_h \in \mathbf{V}_h$,*

$$\begin{aligned} |a_h(\mathbf{u} - \boldsymbol{I}_h \mathbf{u}, \mathbf{v}_h)| &\leq c\sqrt{\mu} h^k |||\mathbf{v}_h||| |\mathbf{u}|_{k+1,\Omega} \\ |b_h(p - i_h p, \mathbf{v}_h)| &\leq \frac{c}{\sqrt{\mu}} h^k |||\mathbf{v}_h||| |p|_{k,\Omega}. \end{aligned}$$

THEOREM 2.3.7. *Let $(\mathbf{u}, p) \in \mathbf{H}^{k+1}(\Omega) \times H^k(\Omega)$ and let γ be sufficiently large. Then the solution $(\mathbf{u}_h^\gamma, p_h^\gamma)$ of (2.3.3) satisfies:*

$$\begin{aligned} |||\, \mathbf{u} - \mathbf{u}_h^\gamma \,||| &\leq ch^k (\sqrt{\mu} |\mathbf{u}|_{k+1,\Omega} + \frac{1}{\sqrt{\mu}} |p|_{k,\Omega}) \\ \left\| p - p_h^\gamma \right\|_{0,\Omega} &\leq ch^k (\mu |\mathbf{u}|_{k+1,\Omega} + |p|_{k,\Omega}). \end{aligned}$$

If, moreover, Ω is convex then there exists c such that

$$\left\| \mathbf{u} - \mathbf{u}_h^\gamma \right\|_{0,\Omega} \leq ch^{k+1} (|\mathbf{u}|_{k+1,\Omega} + \frac{1}{\mu} |p|_{k,\Omega}).$$

2.3.2. Robustness with respect to the stabilization parameter. We have also studied the behaviour of the proposed dG method when γ tends towards infinity. I have first shown that the dG solution converges towards the solution of the $\mathbf{P}_k \times P_{k-1}$ nonconforming approximation of the Stokes problem, given by

$$\left\{ \begin{array}{l} (\mathbf{u}_h^*, p_h^*) \in \mathbf{H}_h \times Q_h, \\ A_0(\mathbf{u}_h^*, \mathbf{v}_h) + b(p_h^*, \mathbf{v}_h) = l(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{H}_h \\ b(q_h, \mathbf{u}_h^*) = 0 \quad \forall q_h \in Q_h \end{array} \right. \tag{2.3.4}$$

where

$$\begin{aligned} \mathbf{H}_h = \Big\{ & \mathbf{v}_h \in \mathbf{L}^2(\Omega);\ (\mathbf{v}_h)_{/T} \in \mathbf{P}_k,\ \forall T \in \mathcal{T}_h, \\ & \mathbf{v}_h \text{ continuous (resp. } \mathbf{0}) \text{ at the } k \text{ Gauss points of } e \in \varepsilon_h^{int} \text{ (resp. } \varepsilon_h^\partial) \Big\}. \end{aligned}$$

For $k = 1$, $\mathbf{H}_h$ is the well-known Crouzeix-Raviart space **[66]**; $k = 2$ corresponds to the Fortin-Soulie element **[88]**, whereas for $k = 3$ we retrieve the element of Crouzeix-Falk **[67]**. It is well-known that (2.3.4) is well-posed for $k = 1,\ 2,\ 3$, thanks to a discrete Poincaré inequality on $\mathbf{H}_h$. Next, it is important to notice that our choice of stabilization yields

$$\mathrm{Ker}_h J = \left\{ \mathbf{v}_h \in \mathbf{V}_h;\ [\boldsymbol{\pi}_{k-1} \mathbf{v}_h]_{/e} = 0,\ \forall e \in \varepsilon_h \right\} = \mathbf{H}_h.$$

THEOREM 2.3.8. *Let $(\mathbf{u}_h^\gamma, p_h^\gamma)$ the solution of (2.3.3) and $(\mathbf{u}_h^*, p_h^*)$ the solution of (2.3.4). Then*

$$\lim_{\gamma\mapsto\infty} \left(||| \mathbf{u}_h^\gamma - \mathbf{u}_h^* ||| + \|p_h^\gamma - p_h^*\|_{0,\Omega} \right) = 0.$$

Proof. Let us recall the following Poincaré-Friedrichs inequality for discontinuous finite element spaces, cf. Brenner **[40]**:

$$\|\mathbf{v}\|_{0,\Omega}^2 \leq c(|\mathbf{v}|_{1,h}^2 + \sum_{e\in\varepsilon_h^{int}} \frac{1}{|e|}\|[\pi_0\mathbf{v}]\|_{0,e}^2 + \phi(\mathbf{v})^2), \quad \forall \mathbf{v}\in\mathbf{V}_h$$

where $\phi : \mathbf{H}^1(\Omega) \to \mathbb{R}$ is a continuous semi-norm such that for a constant function $\mathbf{c}$, $\phi(\mathbf{c}) = 0$ if and only if $\mathbf{c} = \mathbf{0}$.

By choosing $\phi(\mathbf{v}) = (\sum_{e\in\varepsilon_h^\partial} \|\boldsymbol{\pi}_0\mathbf{v}\|_{0,e}^2)^{1/2}$, one deduces a slightly different Poincaré-Friedrichs inequality:

$$\|\mathbf{v}\|_{0,\Omega} \leq c\left(|\mathbf{v}|_{1,h}^2 + \frac{1}{\mu}J(\mathbf{v},\mathbf{v})\right)^{1/2}, \quad \forall \mathbf{v}\in\mathbf{V}_h. \tag{2.3.5}$$

By using that

$$\alpha\ ||| \mathbf{u}_h^\gamma |||^2 \leq \|\mathbf{f}\|_{0,\Omega}\|\mathbf{u}_h^\gamma\|_{0,\Omega}$$

and Lemma 2.3.2, one finally obtains $\sqrt{\mu}\ ||| \boldsymbol{u}_h^\gamma ||| + \|p_h^\gamma\|_{0,\Omega} \leq C$ for γ large enough.

The rest of the proof is standard: there exists a subsequence which converges as $\gamma\to\infty$ towards $(\mathbf{u}_h^\infty, p_h^\infty) \in \mathbf{H}_h \times Q_h$, solution of the limit problem (2.3.4). The well-posedness of (2.3.4) implies that the whole sequence $(\mathbf{u}_h^\gamma, p_h^\gamma)_\gamma$ is convergent towards $(\mathbf{u}_h^*, p_h^*)$. ■

REMARK 2.3.9. If the stabilization term $J(\cdot,\cdot)$ is replaced by $J^*(\cdot,\cdot)$, we can only conclude that the limit $\mathbf{u}_h^\infty$ belongs to $\mathbf{P}_k^{cont}$. Since $\mathbf{P}_k^{cont} \times P_{k-1}^{disc}$ is not a stable pair of spaces for the Stokes problem, we cannot even deduce that $(p_h^\gamma)_\gamma$ is bounded.

The previous result can be now improved by establishing its convergence rate. The main tool is hybridization following **[155]**, together with a known result in standard optimization. For this purpose, let us introduce the Lagrange multiplier $\boldsymbol{\lambda}_h^\gamma$ belonging to:

$$\mathbf{L}_h = \left\{\theta_h \in \Pi_{e\in\varepsilon_h}\mathbf{L}^2(e);\ \forall e\in\varepsilon_h,\ (\boldsymbol{\theta}_h)_{/e} \in \mathbf{P}_{k-1}\right\}$$

and defined by its restriction on any edge $e\in\varepsilon_h$ as follows:

$$\boldsymbol{\lambda}_h^\gamma = \frac{\gamma\sqrt{\mu}}{|e|}[\boldsymbol{\pi}_{k-1}\mathbf{u}_h^\gamma]. \tag{2.3.6}$$

The space $\mathbf{L}_h$ is endowed with the mesh-dependent norm $\|\cdot\|_{0,\varepsilon_h}$ associated with the scalar product:

$$\langle\boldsymbol{\vartheta}_h, \boldsymbol{\theta}_h\rangle_{0,\varepsilon_h} = \sum_{e\in\varepsilon_h} |e| \int_e \boldsymbol{\vartheta}_h\cdot\boldsymbol{\theta}_h \mathrm{ds}, \quad \forall\boldsymbol{\vartheta}_h, \boldsymbol{\theta}_h \in \mathbf{L}_h.$$

Let us put $x_h^\gamma = (\mathbf{u}_h^\gamma, p_h^\gamma)$ and the space $X_h = \mathbf{V}_h \times Q_h$, endowed with the product norm

$$[\chi_h]^2 = \mu\ |\ \mathbf{v}_h\ |_{1,h}^2 + J(\mathbf{v}_h,\mathbf{v}_h) + \frac{1}{\mu}\|q_h\|_{0,\Omega}^2, \quad \forall\chi_h = (\mathbf{v}_h, q_h) \in X_h.$$

We also introduce the continuous bilinear forms on $X_h\times X_h$, respectively $\mathbf{L}_h\times X_h$:

$$\begin{aligned}
\Xi(x_h^\gamma, \chi_h) &= A_0(\mathbf{u}_h^\gamma, \mathbf{v}_h) + A_1(\mathbf{u}_h^\gamma, \mathbf{v}_h) + b_h(p_h^\gamma, \mathbf{v}_h) - b_h(q_h, \mathbf{u}_h^\gamma),\\
\Lambda_h(\boldsymbol{\theta}_h, \chi_h) &= \sum_{e\in\varepsilon_h} \sqrt{\mu}\int_e \boldsymbol{\theta}_h\cdot[\boldsymbol{\pi}_{k-1}\mathbf{v}_h]\mathrm{ds}.
\end{aligned}$$

Then the discrete problem (2.3.3) can be equivalently written as follows:

$$\left\{\begin{array}{ll} (x_h^\gamma, \boldsymbol{\lambda}_h^\gamma) \in X_h \times \mathbf{L}_h & \\ \Xi(x_h^\gamma, \chi_h) + \Lambda_h(\boldsymbol{\lambda}_h^\gamma, \chi_h) = \int_\Omega \mathbf{f}\cdot\mathbf{v}_h \mathrm{d}\mathbf{x} & \forall \chi_h \in X_h \\ \Lambda_h(\boldsymbol{\theta}_h, x_h^\gamma) - \frac{1}{\gamma}\left\langle \boldsymbol{\lambda}_h^\gamma, \boldsymbol{\theta}_h\right\rangle_{0,\varepsilon_h} = 0 & \forall \boldsymbol{\theta}_h \in \mathbf{L}_h \end{array}\right. . \tag{2.3.7}$$

I have shown the next result, which is the main ingredient in the proof of the convergence rate.

LEMMA 2.3.10. *For γ sufficiently large, $\Lambda_h(\cdot,\cdot)$ satisfies:*

$$\inf_{\boldsymbol{\theta}_h \in \mathbf{L}_h} \sup_{\chi_h \in X_h} \frac{\Lambda_h(\boldsymbol{\theta}_h, \chi_h)}{\|\boldsymbol{\theta}_h\|_{0,\varepsilon_h}\,[\chi_h]} \geq \delta.$$

Proof. With any $\boldsymbol{\theta}_h = (\boldsymbol{\theta}_h^e)_{e\in\varepsilon_h} \in \mathbf{L}_h$, we associate $\chi_h = (\mathbf{v}_h, 0) \in X_h$ as follows. For any edge $e \in \varepsilon_h$, let T_e denote the triangle such that $e \subseteq \partial T_e$ and $\mathbf{n}_e$ is exterior to T_e. Then we consider the following auxiliary problem:

$$\left\{\begin{array}{lll} -\Delta \mathbf{v}^e &= 0 & \text{in } T_e \\ \partial_n \mathbf{v}^e &= \boldsymbol{\theta}_h^e & \text{on } e \\ \mathbf{v}^e &= 0 & \text{on } \partial T_e \setminus \{e\} \end{array}\right.$$

whose solution satisfies $|\mathbf{v}^e|_{1,T_e} = \|\boldsymbol{\theta}_h^e\|_{-1/2,e}$. Let $\mathbf{v}$ be defined on Ω by $(\mathbf{v})_{/T_e} = \mathbf{v}^e$ for all $e \in \varepsilon_h$.

Next, we consider on T_e the discrete function $\mathbf{v}_h^e = \mathbf{I}_h \mathbf{v}^e$ and we finally define $\mathbf{v}_h \in \mathbf{V}_h$ by $(\mathbf{v}_h)_{/T_e} = \mathbf{v}_h^e$ for all $e \in \varepsilon_h$. Thanks to the properties of $\mathbf{I}_h$, we get that

$$\sup_{\chi_h \in X_h} \frac{\Lambda_h(\boldsymbol{\theta}_h, \chi_h)}{\|\boldsymbol{\theta}_h\|_{0,\varepsilon_h}\,[\chi_h]} \geq C\,|\mathbf{v}|_{1,h} = C\left(\sum_{e\in\varepsilon_h} \|\boldsymbol{\theta}_h^e\|_{-1/2,e}^2\right)^{1/2}.$$

The desired statement follows by introducing $\underline{q}^e = \nabla \mathbf{v}^e \in \underline{H}(\mathrm{div}, T_e)$ as well as its Raviart-Thomas interpolation on T_e , $\underline{q}_h^e \in \underline{RT}_{k-1}$ and by noting that $\underline{q}_h^e \mathbf{n}_e = \underline{q}^e \mathbf{n}_e = \boldsymbol{\theta}_h^e$ on e. Then we obtain

$$\sqrt{|e|}\,\|\boldsymbol{\theta}_h^e\|_{0,e} \leq C\left\|\underline{q}^e\right\|_{0,T_e} = C\,\|\boldsymbol{\theta}_h^e\|_{-1/2,e}$$

thanks to the equivalence of norms in finite dimensional spaces, the normal trace theorem in $\underline{H}(\mathrm{div}, \hat{T}_e)$ and a classical scaling argument based on Piola's transformation (cf. **[158]**). ■

From Theorem 2.3.8 and Lemma 2.3.10, we know that $(x_h^\gamma, \boldsymbol{\lambda}_h^\gamma)_\gamma$ is uniformly bounded in $X_h \times \mathbf{L}_h$ and convergent as $\gamma \to \infty$ towards $(x_h^*, \boldsymbol{\lambda}_h^*)$, solution of the following well-posed problem:

$$\left\{\begin{array}{lll} (x_h^*, \boldsymbol{\lambda}_h^*) \in X_h \times \mathbf{L}_h & & \\ \Xi(x_h^*, \chi_h) + \Lambda_h(\boldsymbol{\lambda}_h^*, \chi_h) &= \int_\Omega \mathbf{f}\cdot\mathbf{v}_h \mathrm{d}\mathbf{x} & \forall \chi_h \in X_h \\ \Lambda_h(\boldsymbol{\theta}_h, x_h^*) &= 0 & \forall \boldsymbol{\theta}_h \in \mathbf{L}_h \end{array}\right. . \tag{2.3.8}$$

We have now assembled the ingredients for the main result of this section.

THEOREM 2.3.11. *Let γ be sufficiently large. Then one has:*

$$\left|\left|\left|\mathbf{u}_h^\gamma - \mathbf{u}_h^*\right|\right|\right| + \frac{1}{\sqrt{\mu}}\left\|p_h^\gamma - p_h^*\right\|_{0,\Omega} + \|\boldsymbol{\lambda}_h^\gamma - \boldsymbol{\lambda}_h^*\|_{0,\varepsilon_h} \leq \frac{C}{\gamma\sqrt{\mu}}.$$

Proof. We remark that $\mathrm{Ker}\Lambda_h = \mathbf{H}_h \times M_h$ and that $\Xi(\cdot,\cdot)$ is coercive on $\mathbf{H}_h$. Together with Lemma 2.3.10, the statement follows with the help of well-known results concerning penalty methods from the literature, see for instance Proposition II.4.1 from **[47]**. ■

REMARK 2.3.12. One can study in a similar way the behaviour of Nitsche's method [**138**] when γ tends towards infinity. As an example, I have considered a second order elliptic problem discretised by usual conforming or nonconforming elements, and where the boundary conditions are treated weakly. Thus, the stabilisation term only takes into account the edges situated on the Dirichlet boundary. We employ $J(\cdot,\cdot)$ for the nonconforming approximation, respectively $J^*(\cdot,\cdot)$ for the conforming one. By adapting the proof of Lemma 2.3.10, one can then easily show that the solutions respectively converge with $O(1/\gamma)$ towards the solutions of the nonconforming and conforming problems with strongly imposed boundary conditions. This result seems to be new.

2.3.3. Robust *a posteriori* error analysis based on $H(div)$ - fluxes. Our analysis follows the idea of Kim, who mentioned in [**110**] how to construct an *a posteriori* error indicator for the dG approximation of the Laplace equation based on the reconstruction of a locally conservative $H(\mathrm{div},\Omega)$-conforming vector approximation.

We introduce the Raviart-Thomas finite element space (cf. [**158**])

$$\underline{\Sigma}_h = \left\{ \underline{\theta}_h \in \underline{H}(\mathrm{div},\Omega);\ (\underline{\theta}_h)_{/T} \in \underline{RT}_{k-1},\ \ \forall T \in \mathcal{T}_h \right\}$$

where $\underline{RT}_{k-1} = \underline{P}_{k-1} + (\mathbf{x} \otimes \mathbf{P}_{k-1})$. Then we construct a tensor $\underline{\sigma}_h^\gamma \in \underline{\Sigma}_h$ from the solution $(\mathbf{u}_h^\gamma, p_h^\gamma)$ of (2.3.3) by specifying its degrees of freedom as follows:

$$\underline{\sigma}_h^\gamma \mathbf{n}_e = \mu \left\{ \partial_n \mathbf{u}_h^\gamma \right\} - \frac{\mu\gamma}{|e|} \left[\boldsymbol{\pi}_{k-1} \mathbf{u}_h^\gamma \right] - \left\{ p_h^\gamma \right\} \mathbf{n}_e, \quad \forall e \in \varepsilon_h$$

and for $k = 2$ or 3,

$$\int_T \underline{\sigma}_h^\gamma : \underline{r} \mathrm{dx} = \int_T (\mu \underline{\nabla} \mathbf{u}_h^\gamma - p_h^\gamma \underline{I}) : \underline{r} \mathrm{dx}, \quad \forall T \in \mathcal{T}_h,\ \ \forall \underline{r} \in \underline{P}_{k-2}.$$

One then has that $\underline{\sigma}_h^\gamma$ is locally conservative, that is $\int_T (\mathrm{div}\underline{\sigma}_h^\gamma + \mathbf{f}) \mathrm{dx} = \mathbf{0}$ on every triangle $T \in \mathcal{T}_h$. It is useful to introduce $\underline{\sigma} = \mu \underline{\nabla} \mathbf{u} - p \underline{I}$ which clearly belongs to $\underline{H}(\mathrm{div},\Omega)$.

We next define following [**110**] a residual-type error estimator η^γ by

$$\begin{aligned} \eta_1^\gamma &= \Big(\frac{1}{\mu} \sum_{T \in \mathcal{T}_h} \left\| \underline{\sigma}_h^\gamma - \mu \underline{\nabla} \mathbf{u}_h^\gamma + p_h^\gamma \underline{I} \right\|_{0,T}^2 \Big)^{1/2}, \\ \eta^\gamma &= \left((\eta_1^\gamma)^2 + J^*(\mathbf{u}_h^\gamma, \mathbf{u}_h^\gamma) \right)^{1/2} \end{aligned}$$

and we also introduce $\eta_{osc}^2 = \frac{1}{\mu} \sum_{T \in \mathcal{T}_h} h_T^2 \left\| \mathbf{f} - \boldsymbol{\pi}_{k-1}^T \mathbf{f} \right\|_{0,T}^2$.

2.3.3.1. *Reliability and efficiency of the error indicator.* Let $S(\cdot,\cdot)$ the bilinear form of the continuous Stokes problem (2.3.1), which we extend on $(\mathbf{H}_0^1(\Omega) + \mathbf{V}_h) \times L_0^2(\Omega)$ as follows:

$$S((\mathbf{u},p),(\mathbf{v},q)) = \mu \sum_{T \in \mathcal{T}_h} \int_T \underline{\nabla} \mathbf{u} : \underline{\nabla} \mathbf{v} \, \mathrm{dx} - \sum_{T \in \mathcal{T}_h} \int_T p \mathrm{div} \mathbf{v} \, \mathrm{dx} + \sum_{T \in \mathcal{T}_h} \int_T q \mathrm{div} \mathbf{v} \, \mathrm{dx}$$

and let $(\boldsymbol{\phi}, \xi) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega)$ the unique solution of

$$S((\boldsymbol{\phi},\xi),(\mathbf{v},q)) = S((\mathbf{u}_h^\gamma, p_h^\gamma),(\mathbf{v},q)), \quad \forall (\mathbf{v},q) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega).$$

It is useful to introduce the *inf-sup* constant $\tilde{\beta} = \tilde{\beta}(\Omega)$ for the continuous Stokes problem:

$$\tilde{\beta} \left\| r \right\|_{0,\Omega} \leq \sup_{\mathbf{v} \in \mathbf{H}_0^1(\Omega)} \frac{b(r,\mathbf{v})}{|\mathbf{v}|_{1,\Omega}}.$$

We split the error by means of the triangle inequality and we bound the consistency and the nonconformity error (cf. [**108**]) separately. We agree to denote by $c(k)$ any constant depending only on the degree k.

We have proved in [A11] that:

LEMMA 2.3.13. *There exists $c(k)$ such that*

$$\begin{aligned}\sqrt{\mu}\left|\mathbf{u}-\boldsymbol{\phi}\right|_{1,\Omega}+\frac{\tilde{\beta}}{2\sqrt{\mu}}\left\|p-\xi\right\|_{0,\Omega} &\leq \eta_1^\gamma + c(k)\left(\frac{1}{\tilde{\beta}}\sqrt{J^*(\mathbf{u}_h^\gamma,\mathbf{u}_h^\gamma)}+\eta_{osc}\right),\\ \sqrt{\mu}\left|\mathbf{u}_h^\gamma-\boldsymbol{\phi}\right|_{1,h}+\frac{\tilde{\beta}}{\sqrt{\mu}}\left\|p_h^\gamma-\xi\right\|_{0,\Omega} &\leq \frac{c(k)(1+\sqrt{1+\tilde{\beta}^2})}{\tilde{\beta}}\sqrt{J^*(\mathbf{u}_h^\gamma,\mathbf{u}_h^\gamma)}.\end{aligned}$$

The previous lemma ensures the reliability of η^γ, with a factor 1 in front of η_1^γ. Concerning the efficiency of the error estimator, let us introduce for any $T\in\mathcal{T}_h$ the local contributions

$$\begin{aligned}\eta_{1,T}^\gamma &= \frac{1}{\sqrt{\mu}}\left\|\underline{\sigma}_h^\gamma-\mu\underline{\nabla}\mathbf{u}_h^\gamma+p_h^\gamma\underline{I}\right\|_{0,T},\\ (\eta_{2,T}^\gamma)^2 &= \sum_{e\subset(\partial T\setminus\partial\Omega)}\frac{\mu}{2\left|e\right|}\left\|\left[\mathbf{u}_h^\gamma\right]\right\|_{0,e}^2+\sum_{e\subset(\partial T\cap\partial\Omega)}\frac{\mu}{\left|e\right|}\left\|\left[\mathbf{u}_h^\gamma\right]\right\|_{0,e}^2\end{aligned}$$

such that $(\eta^\gamma)^2=\sum_{T\in\mathcal{T}_h}(\eta_{1,T}^\gamma)^2+(\eta_{2,T}^\gamma)^2$.

THEOREM 2.3.14. *There exist a constant $c(k,\mathcal{T}_h)$ depending on the minimum angle of $\mathcal{T}_h$ and on k and a constant $c(k)$ such that, for any $T\in\mathcal{T}_h$, one has*

$$(\eta_{1,T}^\gamma)^2\leq c(k,\mathcal{T}_h)\left(\mu\left|\mathbf{u}-\mathbf{u}_h^\gamma\right|_{1,\omega_T}^2+\tfrac{1}{\mu}\left\|p-p_h^\gamma\right\|_{0,\omega_T}^2\right)+c(k)\gamma^2\sum_{e\subset\partial T}\frac{\mu}{\left|e\right|}\left\|\left[\boldsymbol{\pi}_{k-1}\mathbf{u}_h^\gamma\right]\right\|_{0,e}^2,$$

$$(\eta_{2,T}^\gamma)^2\leq c(k)\left|\mathbf{u}-\mathbf{u}_h^\gamma\right|_{1,\omega_T}^2+\sum_{e\subset\partial T}\frac{\mu}{\left|e\right|}\left\|\left[\boldsymbol{\pi}_{k-1}\mathbf{u}_h^\gamma\right]\right\|_{0,e}^2$$

where ω_T is the set of all elements sharing an edge with T. Consequently,

$$(\eta^\gamma)^2\leq c_1(k,\mathcal{T}_h)\left(\mu\left|\mathbf{u}-\mathbf{u}_h^\gamma\right|_{1,h}^2+\frac{1}{\mu}\left\|p-p_h^\gamma\right\|_{0,\Omega}^2+(1+\gamma^2)J(\mathbf{u}_h^\gamma,\mathbf{u}_h^\gamma)\right).$$

Proof. The proof is rather technical and can be found in [A11]. It uses the argument of Verfürth **[164]**, based on weighted norms by bubble functions and on inverse inequalities. ■

2.3.3.2. *Behaviour of the error indicator for large γ.* We are now interested in the limit of the previous estimator η^γ as $\gamma\to\infty$. According to paragraph 2.3.2, we define $\underline{\sigma}_h^*\in\underline{\Sigma}_h$ as follows:

$$\underline{\sigma}_h^*\mathbf{n}_e=\mu\left\{\partial_n\mathbf{u}_h^*\right\}-\sqrt{\mu}\lambda_h^*-\left\{p_h^*\right\}\mathbf{n}_e,\quad\forall e\in\varepsilon_h$$

and for $k=2$ or 3,

$$\int_T\underline{\sigma}_h^*:\underline{r}\mathrm{dx}=\int_T(\mu\underline{\nabla}\mathbf{u}_h^*-p_h^*\underline{I}):\underline{r}\mathrm{dx},\quad\forall T\in\mathcal{T}_h,\ \ \forall\underline{r}\in\underline{P}_{k-2}.$$

Then we introduce

$$\begin{aligned}(\eta_1^*)^2 &= \frac{1}{\mu}\sum_{T\in\mathcal{T}_h}\left\|\underline{\sigma}_h^*-\mu\underline{\nabla}\mathbf{u}_h^*+p_h^*\underline{I}\right\|_{0,T}^2,\\ (\eta^*)^2 &= (\eta_1^*)^2+J^*(\mathbf{u}_h^*,\mathbf{u}_h^*).\end{aligned}$$

It follows, thanks to Theorem 2.3.11, that there exists a constant c such that

$$\left|\eta^\gamma-\eta^*\right|\leq\frac{c}{\gamma\sqrt{\mu}}.$$

In the case $k=1$, I have shown in [A11] that η_1^* can be written in a simpler way.

LEMMA 2.3.15. *Let $k = 1$ and $\mathbf{f}$ piecewise constant with respect to $\mathcal{T}_h$. Then:*

$$\eta_1^* = \left(\frac{1}{4\mu} \sum_{T \in T_h} \|\mathbf{f} \otimes (\mathbf{x} - \mathbf{x}_T)\|_{0,T}^2 \right)^{1/2}.$$

Proof. Following the idea of Marini **[130]**, we define on any T a tensor in $\underline{RT}_0$ by

$$\underline{\sigma}'_h = \mu \underline{\nabla} \mathbf{u}_h^* - p_h^* \underline{I} - \frac{1}{2} \mathbf{f} \otimes (\mathbf{x} - \mathbf{x}_T).$$

Then we obtain from the limit problem (2.3.8) that

$$\sum_{e \in \varepsilon_h} \int_e \underline{\sigma}'_h \mathbf{n}_e \cdot [\mathbf{v}_h] \, \mathrm{ds} \quad = \quad \sum_{e \in \varepsilon_h} \int_e \underline{\sigma}_h^* \mathbf{n}_e \cdot [\mathbf{v}_h] \, \mathrm{ds}, \quad \forall \mathbf{v}_h \in \mathbf{V}_h$$

which finally implies $\underline{\sigma}'_h = \underline{\sigma}_h^*$. ∎

Moreover, the *a posteriori* error estimator η^* can now be easily related to the one introduced by Dari, Duran and Padra in **[69]** for the P_1-nonconforming discretization (2.3.4), that is:

$$(\eta^{DDP})^2 = \frac{1}{\mu} \sum_{T \in T_h} |T| \, \|\mathbf{f}\|_{0,T}^2 + \mu \sum_{e \in \varepsilon_h} |e| \int_e [\partial_t \mathbf{u}_h^*]^2 \, \mathrm{ds}. \tag{2.3.9}$$

Indeed, there exist two numeric constants c_1, c_2 such that

$$c_1 \eta^{DDP} \leq \eta^* \leq c_2 \eta^{DDP}.$$

2.3.4. Numerical tests. I present next some numerical experiments illustrating the influence of the stabilization parameter and the adaptive mesh refinement. The convergence rate $\mathcal{O}(h^k)$ with respect to mesh refinement has also been tested for $1 \leq k \leq 3$, cf. [A11] and [C14]. The developed codes are based on the C++ library CONCHA.

2.3.4.1. *Behaviour with respect to the stabilization parameter.* We vary γ and compare with the Interior Penalty (IP) scheme on a fixed mesh. We first consider an exact solution and compare the errors computed by our method (continuous lines) with the IP method (dotted lines). The results for $k = 1$ and $k = 2$ are shown in Fig. 2.3.1. One observes the robustness of the analyzed method with respect to γ, in contrast to the standard interior penalty stabilization, which blows up with γ.

We next consider a Poiseuille flow in the domain $\Omega = [0; 0.06] \times [-0.01; 0.01]$ with Dirichlet inflow conditions and a Neumann type condition on the outflow, $\mu(\underline{\nabla} \mathbf{u})\mathbf{n} - p\mathbf{n} = \mathbf{0}$. A parabolic inflow leads to the analytical solution $\mathbf{u} = (a(0.01^2 - y^2), 0)$, $p = bx + c$. For $k \geq 2$ both dG codes give the exact solution, as expected. We now vary γ and compare the pressures for $k = 1$ in Fig. 2.3.2. Clearly, the IP method is less accurate and completely loses stability.

We have also tested the two dG methods for non-smooth solutions, by imposing in the previous test-case $\mathbf{u} \cdot \mathbf{n} = 1$ on the inflow. To dispose of a reference solution, we have computed it with P_1-nonconforming elements, see Fig. 2.3.3 a), on a fine mesh. One notices again a lack of accuracy of the IP method, which becomes visible at rather small values of the stabilization parameter.

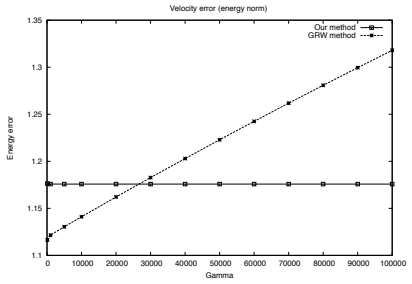
2.3.4.2. *Adaptive mesh refinement.* We now consider the *a posteriori* error estimators, using the following adaptive algorithm **[25]**. Let $\gamma_a \in [0, 1]$, $\theta_a \in]0, 1]$ be given. At each iteration k, we determine a set $\mathcal{M}_k$ of marked cells by the criterion :

IF $\eta^2 > \gamma_a \eta_{osc}^2$

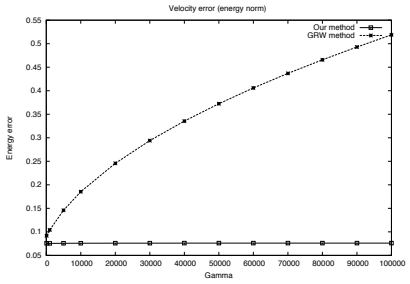
Choose $\mathcal{M}_k$ of minimal cardinal such that $\eta^2(\mathcal{M}_k) \geq \theta_a \eta^2$

ELSE

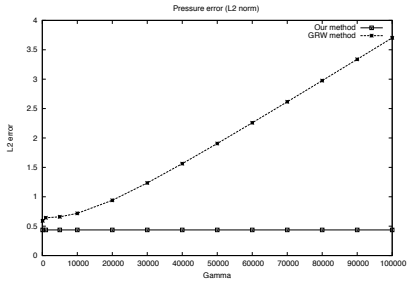
Choose $\mathcal{M}_k$ of minimal cardinal such that $\eta_{osc}^2(\mathcal{M}_k) \geq \theta_a \eta_{osc}^2$



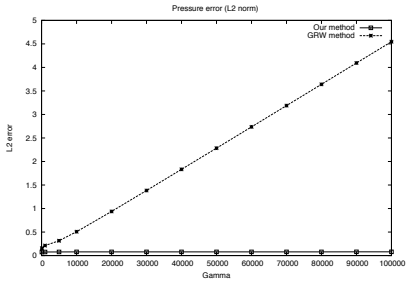
a) Velocity errors for $k = 1$



b) Velocity errors for $k = 2$

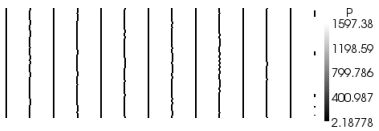


c) Pressure errors for $k = 1$

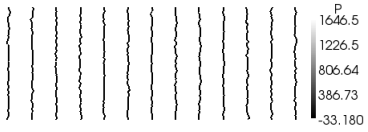


d) Pressure errors for $k = 2$

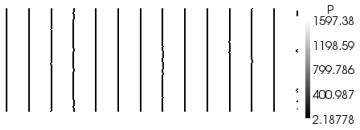
FIGURE 2.3.1. Behaviour of the errors with respect to γ



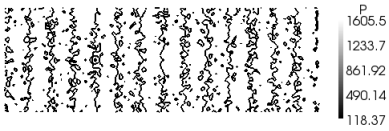
a) Our method for $\gamma = 100$



b) IP method for $\gamma = 100$



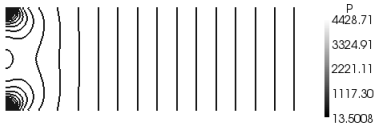
c) Our method for $\gamma = 1000$



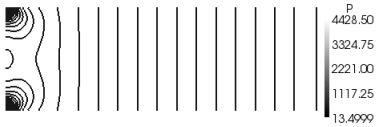
d) IP method for $\gamma = 1000$

FIGURE 2.3.2. Poiseuille flow: comparison of pressures for different γ

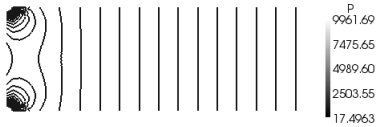
Following **[69]**, we consider the Stokes equations on the slit domain $\Omega =]-1,1[^2 \backslash [0,1] \times \{0\}$ with zero righthand-side and non-homogeneous boundary conditions. The exact solution in polar coordinates is given in [A11]. We take $\gamma = 10$, $k = 1$, denote by N the number of cells and



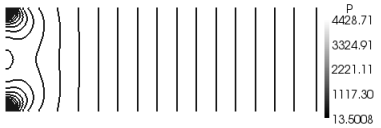
a) $\mathbf{P}_1 \times P_0$ nonconforming reference solution



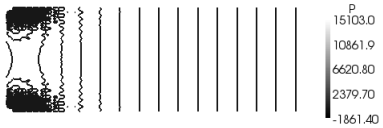
b) Our method for $\gamma = 100$



c) IP method for $\gamma = 100$



d) Our method for $\gamma = 1000$



e) IP method for $\gamma = 1000$

FIGURE 2.3.3. Constant inflow: comparison of pressures for different γ

compute the ratios :

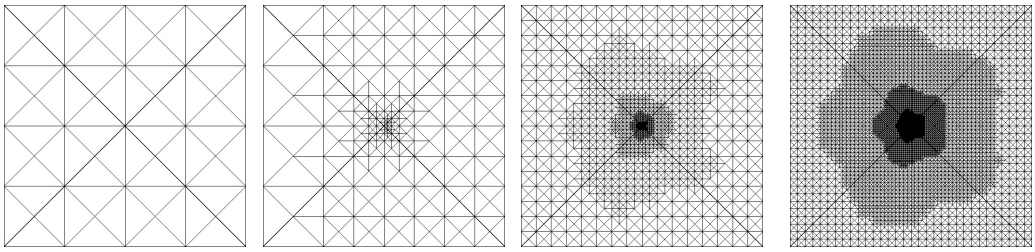
$$\mathcal{K}_1 = \left(\frac{N_{fine}}{N_{coarse}}\right)^{-1/2}, \qquad \mathcal{K}_2 = \frac{\eta_{fine}}{\eta_{coarse}}, \qquad \mathcal{K}_3 = \frac{(|\mathbf{u} - \mathbf{u}_h|_{1,h})_{fine}}{(|\mathbf{u} - \mathbf{u}_h|_{1,h})_{coarse}}.$$

When considering a uniform refinement, we approximately get $\mathcal{K}_2 \approx \mathcal{K}_3 \approx \frac{1}{\sqrt{2}}$, so the convergence rate for this singular solution is $\mathcal{O}(h^{1/2}) = \mathcal{O}(N^{-1/4})$. We now consider an adaptive mesh refinement with $\theta_a = 0.5$ and we show in Table 1 the values of $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_3$ computed on succesive meshes. We get an improved convergence rate $\mathcal{O}(N^{-1/2})$, similarly to the one reported in the literature, cf. **[69]** for the nonconforming solution with the error estimator (2.3.9). We show in Fig. 2.3.4 the meshes obtained at different steps of the refinement procedure. As expected, the refinement takes well into account the singularity of the solution at the origin.

Finally, we are interested in the behaviour of η^γ as $\gamma \to \infty$. We consider a polynomial exact solution and we present in Table 2 the values of η_1^γ and $J^*(\mathbf{u}_h^\gamma, \mathbf{u}_h^\gamma)$ computed by the dG code for different γ, as well as η_1^* and $J^*(\mathbf{u}_h^*, \mathbf{u}_h^*)$ computed by the P_1 -nonconforming code (on the same mesh). We observe the expected convergence as $\gamma \to \infty$.

N	$\mathcal{K}_1$	η	$\mathcal{K}_2$	$\|\mathbf{u}-\mathbf{u}_h\|_{1,h}$	$\mathcal{K}_3$
64	-	1.550	-	3.001	-
94	0.825	1.471	0.949	2.537	0.845
150	0.791	1.352	0.919	2.166	0.841
206	0.853	1.254	0.927	1.882	0.869
354	0.763	1.086	0.866	1.526	0.811
649	0.738	0.8695	0.800	1.182	0.774
1177	0.742	0.6822	0.784	0.9104	0.770
1998	0767	0.5334	0.782	0.6959	0.7643
3616	0.743	0.4026	0.754	0.5204	0.7478
6544	0.743	0.3046	0.7565	0.3911	0.751
11372	0.758	0.231	0.758	0.2942	0.7522
19502	0.764	0.1778	0.769	0.2252	0.7654

TABLE 1. H^1 -error and estimator for adaptive mesh refinement



a) $N = 64$ b) $N = 354$ c) $N = 3616$ d) $N = 19502$

FIGURE 2.3.4. Sequence of locally refined meshes for the singular solution

γ	5	10	20	100	1000	$\gamma \to \infty$
η_1^γ	0.251	0.1687	0.1681	0.1680	0.1680	η_1^*=0.1680
$J^*(\boldsymbol{u}_h^\gamma, \boldsymbol{u}_h^\gamma)^{1/2}$	0.2106	0.1920	0.1918	0.1917	0.1917	$J^*(\boldsymbol{u}_h^*, \boldsymbol{u}_h^*)^{1/2}$ =0.1917

TABLE 2. Behaviour of η^γ for large γ

CHAPTER 3

APPLICATIONS IN POROUS MEDIA. COUPLING WITH FLUID FLOW

CHAPTER 3

Applications in porous media. Coupling with fluid flow

In this chapter, I describe the work carried out within the project MoTher financed by Total. The goal is to develop a numerical model describing the flow of a compressible fluid in porous and fluid media, by taking into account specific thermodynamical aspects.

I mainly present here the case of single-phase flows. Section 3.1 is devoted to the numerical modeling of a 2D axisymmetric petroleum reservoir, which stands apart from the classical models due to its complex energy equation. A Forchheimer term is also added to the classical Darcy law in order to improve the physical modeling. Section 3.2 deals with the study and numerical approximation of a 1.5D vertical wellbore model, derived as a conforming approximation of a 2D axisymmetric one. The coupling of the previous models is presented in Section 3.3.

Finally, in Section 3.4 I briefly describe multi-component multi-phase 3D flows in porous media.

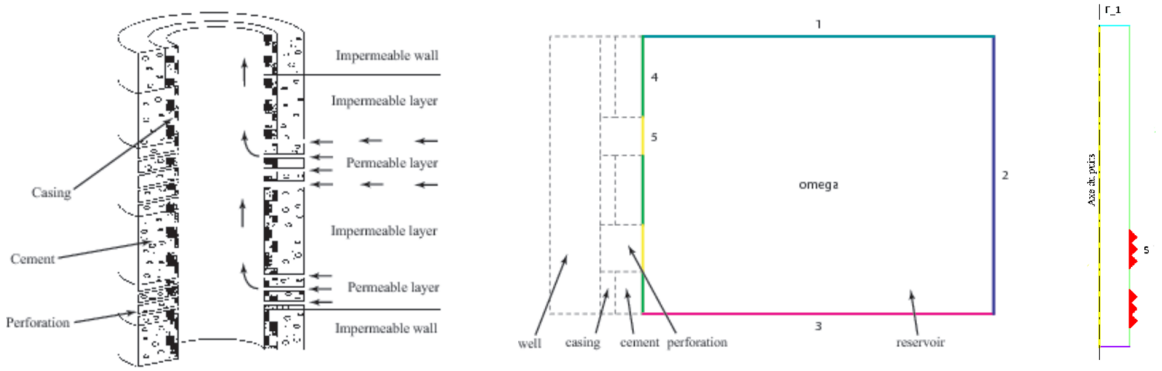
3.1. Darcy-Forchheimer equations with heat transfer

The axisymmetric reservoir model was studied in [A5] and [A6], see also the thesis of Bertrand Denel [**71**]. The well-posedness of the time-discretized system was established and its finite element approximation, based on Raviart-Thomas elements for the mass and heat fluxes and on piecewise constant elements for the pressure and temperature, was analyzed. I have proved that the non-standard mixed variational formulation thus obtained is well-posed. An *a posteriori* error analysis was also carried out. Finally, the developed code was validated by means of several numerical tests.

3.1.1. Physical modeling in axisymmetric framework. We consider the anisothermal flow of a monophasic compressible fluid in a petroleum reservoir. We have represented in Fig. 3.1.1 (a) a cylindrical petroleum well, delimited by a casing and surrounded by a cement layer and by a reservoir, assumed to be a porous medium with an axisymmetric geometry. The communication between the well and the reservoir is achieved through perforations.

The reservoir is treated as a porous medium divided into several geological layers, characterized by their own dip and physical properties. Each layer is made of a porous rock, characterized by vertical and horizontal permeabilities, and saturated with both a mobile monophasic fluid and a residual formation water. We use the following notations: ρ is the density of the fluid, μ its viscosity, $\mathbf{g} = -g\mathbf{e}_3$ the gravitational acceleration, ϕ the porosity of the medium and $\underline{K} = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$ its permeability, with ϕ and $\underline{K}$ depending on the geological layers. We also denote by $\mathbf{v}$ the Darcy velocity and we introduce the specific flux $\mathbf{G} = \rho\mathbf{v}$.

3.1.1.1. *Conservation laws.* The fluid flow is modeled by the Darcy-Forchheimer equation coupled with a non-standard energy balance (cf. [**131**]) which takes into account, besides the convection and the diffusion, the Joule-Thomson compressibility effect S_p and the frictional heating S_μ.



(a) 3D axisymmetric geometry (b) Boundaries of 2D reservoir and wellbore domains

FIGURE 3.1.1. Geometry of a wellbore surrounded by a reservoir

The problem is described by the following conservation laws:

$$\begin{aligned} \phi\frac{\partial \rho}{\partial t} + \mathrm{div}\mathbf{G} &= 0, \\ \rho^{-1}\left(\mu \underline{K}^{-1}\mathbf{G} + F|\mathbf{G}|\mathbf{G}\right) + \nabla p &= \rho\mathbf{g}, \\ (\rho c)_* \frac{\partial T}{\partial t} + \rho^{-1}(\rho c)_f \mathbf{G}\cdot\nabla T - \mathrm{div}\mathbf{q} - S_p - S_\mu &= 0, \\ \rho &= \rho(p,T). \end{aligned}$$

Due to the high filtration velocity which can arise around gas wells, a quadratic term is introduced (cf. **[85]**) in the standard Darcy equation to take into account the kinematic energy losses; F denotes the Forchheimer coefficient. In the energy equation, β is the expansion coefficient, $(\rho c)_*$ characterizes the heat capacity of a virtual medium, equivalent to the fluid and the porous matrix, while $(\rho c)_f$ only symbolizes the fluid properties. The coefficient λ is the equivalent thermal conductivity, T is the temperature, $\mathbf{q} = \lambda\nabla T$ represents the heat flux and

$$S_p = \beta T\left(\phi\frac{\partial p}{\partial t} + \rho^{-1}\mathbf{G}\cdot\nabla p\right), \qquad S_\mu = -\rho^{-1}\mathbf{G}\cdot\nabla p.$$

As usually when modeling petroleum fluids, we use the Peng-Robinson cubic state equation cf. **[145]**. The system is closed by adequate initial and boundary conditions.

3.1.1.2. *Problem in cylindrical coordinates.* Due to the geometry of the domain, it is natural to write the problem in 2D axisymmetric form. The flow is supposed to be radial, the pressure and the temperature independent of the cylindrical coordinate θ, and the 2D domain is supposed to be rectangular, defined by:

$$\Omega_1 = \{(r,z);\ R \le r \le R_\infty,\ z \in [z_{min}, z_{max}]\}$$

where R and R_∞ are the well's and the reservoir's radius. The (r,z) formulation of the problem is:

$$\text{(3.1.1)}\qquad \begin{cases} r\phi\dfrac{\partial\rho}{\partial t}+\operatorname{div}(r\mathbf{G})=0 \\ \rho^{-1}\left(\mu\underline{K}^{-1}+F|\mathbf{G}|\underline{I}\right)\mathbf{G}+\nabla p=\rho\mathbf{g} \\ \dfrac{1}{\lambda}\mathbf{q}-\nabla T=0 \\ r\left(\rho c\right)_*\dfrac{\partial T}{\partial t}+\dfrac{r}{\rho}\left(\rho c\right)_f\mathbf{G}\cdot\nabla T-\operatorname{div}(r\mathbf{q})-r\phi\beta T\dfrac{\partial p}{\partial t}-\dfrac{r}{\rho}(\beta T-1)\mathbf{G}\cdot\nabla p=0 \\ \rho=\rho(p,T) \end{cases}$$

where $\mathbf{G}$ now refers to $(G_r,G_z)^t$, $\mathbf{q}=(q_r,q_z)^t$ and $\nabla=(\frac{\partial}{\partial r},\frac{\partial}{\partial z})^t$, $\operatorname{div}\mathbf{v}=\nabla\cdot\mathbf{v}$. Problem (3.1.1) is a coupled nonlinear system whose unknowns are $\mathbf{G}$, $\mathbf{q}$, p, T and ρ.

3.1.1.3. *Boundary conditions.* The boundary $\Upsilon=\partial\Omega_1$ is divided into five parts as in Fig. 3.1.1 (b). The boundary conditions apply to $\mathbf{G}$ and its dual variable p, as well as to $\mathbf{q}$ and T. Concerning the specific flux, an impermeability condition $\mathbf{G}\cdot\mathbf{n}=0$ is imposed on the top, on the botttom and on the common boundary with the wellbore. On the external boundary, either a normal specific flux or a pressure can be set; this notably allows us to treat the standard cases of a closed reservoir (no-flow condition $\mathbf{G}\cdot\mathbf{n}=0$) and of a reservoir fed at constant pressure. Concerning the temperature, the geothermal gradient is imposed on the top and the bottom of the reservoir, whereas a normal flux condition or a temperature can be set on the lateral boundaries. On the perforations Σ, one can impose $\mathbf{G}\cdot\mathbf{n}$ or p , respectively $\mathbf{q}\cdot\mathbf{n}$ or T.

In what follows, we respectively denote by Υ_p, Υ_T, Υ_G and Υ_q the union of the boundaries where a pressure p^*, a temperature T^*, a normal specific flux G^* and a normal heat flux q^* are given. Thus, we have $\overline{\Upsilon}=\bar{\Upsilon}_G\cup\bar{\Upsilon}_p=\bar{\Upsilon}_q\cup\bar{\Upsilon}_T$ and for the sake of simplicity, we assume that $\Upsilon_p\neq\emptyset$ and $\Upsilon_T\neq\emptyset$.

3.1.2. Analysis of the semi-discretized problem.

3.1.2.1. *Weak formulation.* The time discretization is based on Euler's implicit scheme. At each time step, we determine $\mathbf{G}$, $\mathbf{q}$, p, T and then we update ρ by means of a thermodynamic module. With this aim in view, we first replace in the mass conservation law the time derivative of ρ as follows:

$$\frac{\partial\rho}{\partial t}=\chi\rho\frac{\partial p}{\partial t}-\beta\rho\frac{\partial T}{\partial t},$$

with the compressibility coefficient χ and the expansion coefficient β defined by

$$\chi=\frac{1}{\rho}\left(\frac{\partial\rho}{\partial p}\right)_T,\quad \beta=\frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p=-\frac{1}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_p.$$

By linearizing the convective terms, we obtain at each t^n the following linear system:

$$\text{(3.1.2)}\qquad \begin{cases} \frac{1}{r}\underline{M}\mathbf{G}+\nabla p=-\rho^{n-1}\mathbf{g} \\ \frac{1}{r\lambda}\mathbf{q}-\nabla T=0 \\ r\frac{a}{\Delta t}p-r\frac{b}{\Delta t}T+\operatorname{div}\mathbf{G}=r\frac{a}{\Delta t}p^{n-1}-r\frac{b}{\Delta t}T^{n-1} \\ r\frac{d}{\Delta t}T+\kappa\mathbf{G}^{n-1}\cdot\nabla T-r\frac{f}{\Delta t}p+l\mathbf{G}^{n-1}\cdot\nabla p-\operatorname{div}\mathbf{q}=r\frac{d}{\Delta t}T^{n-1}-r\frac{f}{\Delta t}p^{n-1} \end{cases}$$

where the thermodynamic coefficients a, b, d, f, k, l are computed at t^{n-1} and are all positive, except l which is of variable sign. The tensor

$$\underline{M} = \frac{1}{\rho^{n-1}}(\mu \underline{K}^{-1} + \frac{F}{r}\left|\mathbf{G}^{n-1}\right|\underline{I})$$

is bounded and positive definite and the thermal conductivity λ satisfies $\lambda_1 \geq \lambda \geq \lambda_0 > 0$. For simplicity of writing, we drop the index $n-1$ on the previous coefficients.

From now on, we make the following assumptions on the thermodynamic coefficients, which are justified in practice by all the available experimental data:

(**A1**) a, d, $\frac{1}{\lambda}$ are uniformly bounded from below and $\underline{M}$ is uniformly positive definite;
(**A2**) a, b, d, f, k, l, $\frac{1}{\lambda}$ are bounded a.e. in Ω_1;
(**A3**) $\exists c > 0$ such that $4ad - (b+f)^2 \geq c$ a.e. in Ω_1.

We agree to make the change of variables $\widetilde{\mathbf{G}} = r\mathbf{G}$, $\widetilde{\mathbf{q}} = r\mathbf{q}$ and to denote from now on $\widetilde{\mathbf{G}}$ and $\widetilde{\mathbf{q}}$ by $\mathbf{G}$ and $\mathbf{q}$ respectively. We denote by $\mathbf{V} = (\mathbf{G}, \mathbf{q})$ the vector unknowns, respectively by $\mathbf{s} = (p, T)$ the scalar ones and we introduce the spaces:

$$\begin{aligned}
\mathbf{L}^2(\Omega_1) &= L^2(\Omega_1) \times L^2(\Omega_1),\\
\mathbf{H}(\mathrm{div}, \Omega_1) &= H(\mathrm{div}, \Omega_1) \times H(\mathrm{div}, \Omega_1),\\
\mathbf{H}^0(\mathrm{div}, \Omega_1) &= \left\{\mathbf{V}' = (\mathbf{G}', \mathbf{q}') \in \mathbf{H}(\mathrm{div}, \Omega_1);\ \mathbf{G}' \cdot \mathbf{n} = 0 \text{ on } \Upsilon_G,\ \mathbf{q}' \cdot \mathbf{n} = 0 \text{ on } \Upsilon_q\right\},\\
\mathbf{H}^*(\mathrm{div}, \Omega_1) &= \left\{\mathbf{V}' = (\mathbf{G}', \mathbf{q}') \in \mathbf{H}(\mathrm{div}, \Omega_1);\ \mathbf{G}' \cdot \mathbf{n} = G^* \text{ on } \Upsilon_G,\ \mathbf{q}' \cdot \mathbf{n} = q^* \text{ on } \Upsilon_q\right\}
\end{aligned}$$

endowed with their natural norms $\|\cdot\|_{0,\Omega_1}$ and $|||\cdot|||_{\Omega_1}$.

Then the time-discretized problem can be put in the following mixed weak form:

$$\left\{\begin{array}{ll}
(\mathbf{V}, \mathbf{s}) \in \mathbf{H}^*(\mathrm{div}, \Omega_1) \times \mathbf{L}^2(\Omega_1) & \\
A(\mathbf{V}, \mathbf{V}') + B(\mathbf{s}, \mathbf{V}') = F_1(\mathbf{V}'), & \forall\ \mathbf{V}' \in \mathbf{H}^0(\mathrm{div}, \Omega_1)\\
B(\mathbf{s}', \mathbf{V}) - C(\mathbf{s}, \mathbf{s}') - \alpha D(\mathbf{s}, \mathbf{s}') = F_2(\mathbf{s}'), & \forall\ \mathbf{s}' \in \mathbf{L}^2(\Omega_1)
\end{array}\right. \tag{3.1.3}$$

where the bilinear forms are defined by:

$$A(\mathbf{V}, \mathbf{V}') = \int_{\Omega_1} \frac{1}{r}\underline{M}\mathbf{G} \cdot \mathbf{G}' \mathrm{d}x + \int_{\Omega_1} \frac{1}{r\lambda}\mathbf{q} \cdot \mathbf{q}' \mathrm{d}x,$$

$$B(\mathbf{s}, \mathbf{V}') = -\int_{\Omega_1} p\mathrm{div}\mathbf{G}' \mathrm{d}x + \int_{\Omega_1} T\mathrm{div}\mathbf{q}' \mathrm{d}x,$$

$$C(\mathbf{s}, \mathbf{s}') = \int_{\Omega_1} r\frac{a}{\Delta t}pp' \mathrm{d}x - \int_{\Omega_1} r\frac{b}{\Delta t}Tp' \mathrm{d}x + \int_{\Omega_1} r\frac{d}{\Delta t}TT' \mathrm{d}x - \int_{\Omega_1} r\frac{f}{\Delta t}pT' \mathrm{d}x,$$

$$D(\mathbf{s}, \mathbf{s}') = \textstyle\int_{\Omega_1} \kappa\mathbf{G}^{n-1} \cdot \nabla TT' \mathrm{d}x + \int_{\Omega_1} l\mathbf{G}^{n-1} \cdot \nabla pT' \mathrm{d}x.$$

The parameter α equals 1 for the complete problem (3.1.2), respectively 0 for the problem without convection in the energy equation.

Problem (3.1.3) can be equivalently written as follows:

$$\left\{\begin{array}{l}
x_1 \in X_1^*\\
\mathcal{A}_1(x_1, x_1') = \mathcal{F}_1(x_1'),\ \forall x_1' \in X_1^0
\end{array}\right. \tag{3.1.4}$$

where $x_1 = (\mathbf{V}, \mathbf{s})$, $X_1 = \mathbf{H}(\mathrm{div}, \Omega_1) \times \mathbf{L}^2(\Omega_1)$ and:

$$\mathcal{A}_1 = \begin{pmatrix} A & B \\ B^T & -C - \alpha D \end{pmatrix}, \qquad \mathcal{F}_1 = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$$
$$X_1^0 = \mathbf{H}^0(\mathrm{div}, \Omega_1) \times \mathbf{L}^2(\Omega_1), \qquad X_1^* = \mathbf{H}^*(\mathrm{div}, \Omega_1) \times \mathbf{L}^2(\Omega_1).$$

3.1.2.2. *Problem without convection.* I have first considered the case $\alpha = 0$ and I have shown that the problem has a unique solution. Note that the bilinear form $C(\cdot, \cdot)$ being non-symmetric, one cannot use the results of Brezzi and Fortin [**47**] in order to prove well-posedness.

THEOREM 3.1.1. *Assume **(A1)** to **(A3)** and that ρ^{n-1}, p^{n-1}, $T^{n-1} \in L^2(\Omega)$. Then the mixed problem (3.1.3) with $\alpha = 0$ has a unique solution.*

Proof. I have used an extension of the Babuška-Brezzi theorem (cf. [**158**]) to the case $A(\cdot, \cdot)$ positive and elliptic on Ker$B$, $C(\cdot, \cdot)$ positive, $B(\cdot, \cdot)$ satisfying an *inf-sup* condition, and one of the two forms $A(\cdot, \cdot)$ and $C(\cdot, \cdot)$ is symmetric. These conditions were checked in [A5], in particular the positivity of $C(\cdot, \cdot)$ which is ensured by the hypothesis **(A3)**. ■

Next, let us denote the data of the initial problem by $\mathbf{f} = (f_\Omega, f_\Upsilon) \in \mathcal{X}_\Omega \times \mathcal{X}_\Upsilon$ where:

$$f_\Omega = (\mathbf{f}_1, \mathbf{f}_2, f_3, f_4) \in \mathcal{X}_\Omega = \mathbf{L}^2(\Omega_1) \times \mathbf{L}^2(\Omega_1) \times L^2(\Omega_1) \times L^2(\Omega_1),$$

$$f_\Upsilon = (p^*, T^*, G^*, q^*) \in \mathcal{X}_\Upsilon = H^{1/2}(\Upsilon_p) \times H^{1/2}(\Upsilon_T) \times H^{-1/2}(\Upsilon_G) \times H^{-1/2}(\Upsilon_q).$$

In our case, we have

$$\mathbf{f}_1 = \rho^{n-1}\mathbf{g}, \quad \mathbf{f}_2 = \mathbf{0}, \quad f_3 = \frac{r}{\Delta t}(ap^{n-1} - bT^{n-1}), \quad f_4 = \frac{r}{\Delta t}(dT^{n-1} - fp^{n-1}).$$

Then the right-hand side term of (3.1.3) can be written in terms of $\mathbf{f}$ and, thanks to Theorem 3.1.1, we can define a linear continuous operator

$$\mathcal{L} : \mathcal{X}_\Omega \times \mathcal{X}_\Upsilon \longrightarrow X_1^*$$

which associates with any $\mathbf{f}$ the unique solution $\boldsymbol{\sigma} = (\mathbf{V}, \mathbf{s})$ of (3.1.3). So the variational problem without convection is equivalent to $\mathcal{L}\boldsymbol{\sigma} = \mathbf{f}$.

It is important to note that for sufficiently smooth boundary conditions and thermodynamic coefficients, the solution of (3.1.3) is smoother on each layer Ω_1^i, where $\bar{\Omega}_1 = \cup_{i=1}^N \bar{\Omega}_1^i$. More precisely, one can prove that the operator $\mathcal{L}$ is well-defined from $\mathcal{X}_\Omega \times \mathcal{Y}_\Upsilon$ to $\mathbf{H}^*(\mathrm{div}, \Omega_1) \times \boldsymbol{\mathcal{Z}}$, where now

$$\mathcal{Y}_\Upsilon = \prod_{i=1}^N \left[H^{3/2}(\Upsilon_p^i) \times H^{3/2}(\Upsilon_T^i) \times H^{1/2}(\Upsilon_G^i) \times H^{1/2}(\Upsilon_q^i) \right], \quad \boldsymbol{\mathcal{Z}} = \prod_{i=1}^N \mathbf{H}^2(\Omega_1^i).$$

Here above, we have used the additional notation $H^s(\Upsilon^i) = H^s(\Upsilon \cap \partial\Omega_1^i)$.

THEOREM 3.1.2. *Suppose that $\rho^{n-1} \in H^1(\Omega_1)$, $\nabla\lambda \in \mathbf{L}^\infty(\Omega_1^i)$ and $\underline{M}^{-1} \in \underline{\mathcal{C}}^{0,1}(\bar{\Omega}_1^i)$, for each $i = 1, .., N$. Then for any $f_\Upsilon \in \mathcal{Y}_\Upsilon$ one has that $\mathbf{s} \in \boldsymbol{\mathcal{Z}}$.*

Proof. The proof is based on the interpretation of the variational problem as a set of boundary value problems in each geological layer, with transmission conditions at the interfaces between the subdomains. The conclusion follows thanks to the regularity of an elliptic problem with discontinuous coefficients on a convex polygon (cf. [**96**]). ■

3.1.2.3. *Problem with convection.* Let us now take into account the convective terms and define therefore the linear continuous operator

$$\mathcal{D} : \boldsymbol{\mathcal{Z}} \longrightarrow L^2(\Omega_1), \quad \mathcal{D}(\mathbf{s}) = k\mathbf{G}^{n-1} \cdot \nabla T + l\mathbf{G}^{n-1} \cdot \nabla p,$$

such that $D(\mathbf{s}, \mathbf{s}') = \int_{\Omega_1} \mathcal{D}(\mathbf{s}) T' \mathrm{dx}$. The main point is that $\mathcal{D}$ is compact thanks to the compact embedding $H^1(\Omega_1^i) \hookrightarrow L^2(\Omega_1^i)$, for $i = 1, .., N$. We also need to introduce

$$\mathcal{K} : \mathbf{H}(\mathrm{div}, \Omega_1) \times \boldsymbol{\mathcal{Z}} \longrightarrow \mathcal{X}_\Omega \times \mathcal{Y}_\Upsilon, \qquad \mathcal{K}(\boldsymbol{\sigma}) = (f_\Omega, f_\Upsilon)$$

with $f_\Omega = (\mathbf{0}, \mathbf{0}, 0, \mathcal{D}(\mathbf{s}))$ and $f_\Upsilon = 0$.

Then problem (3.1.3) with $\alpha = 1$ can be written as:

$$\boldsymbol{\sigma} = \mathcal{L}(\mathbf{f} + \mathcal{K}(\boldsymbol{\sigma})) \Longleftrightarrow (\mathcal{I} - \mathcal{L} \circ \mathcal{K})\boldsymbol{\sigma} = \mathcal{L}\mathbf{f}, \tag{3.1.5}$$

where $\mathcal{L} \circ \mathcal{K}$ is now a compact operator from $\mathbf{H}(\mathrm{div}, \Omega_1) \times \boldsymbol{\mathcal{Z}}$ to itself.

In order to prove the well-posedness of problem (3.1.5), I have applied Fredholm's theory.

THEOREM 3.1.3. *Assume **(A1)** to **(A3)**. For Δt sufficiently small, one has* $\mathrm{Ker}(\mathcal{I} - \mathcal{L} \circ \mathcal{K}) = \{\mathbf{0}\}$*, so problem (3.1.5) has a unique solution for any right-hand side term.*

Proof. The solution of $(\mathcal{I} - \mathcal{L} \circ \mathcal{K})\boldsymbol{\sigma} = \mathbf{0}$ clearly satisfies:

$$A(\mathbf{V}, \mathbf{V}) + C(\mathbf{s}, \mathbf{s}) + \int_{\Omega_1} \mathcal{D}(\mathbf{s}) T \mathrm{dx} = 0.$$

By replacing $\nabla T = \dfrac{1}{r\lambda}\mathbf{q}$, $\nabla p = -\dfrac{1}{r}\underline{M}\mathbf{G}$ and by means of Gauss reduction, the previous relation finally leads to $\boldsymbol{\sigma} = \mathbf{0}$ for $\triangle t$ small enough. ■

In conclusion, the time-discretized problem (3.1.3) is well-posed, under non-restrictive regularity assumptions on the data but for a sufficiently small time step.

3.1.3. Finite element approximation.

3.1.3.1. *The discrete problem.* Let a regular family $(\mathcal{T}_h^1)_h$ of triangulations of Ω_1 consisting of triangles, matching at the interfaces between the geological layers, and let $h_1 = \max_{K \in \mathcal{T}_h^1} h_K$. The space discretization is achieved by low-order conforming finite elements. We define:

$$\begin{aligned} L_h &= \{p' \in L^2(\Omega_1);\ p'_{|K} \in P_0,\ \forall K \in \mathcal{T}_h^1\}, \\ V_h &= \{\mathbf{G}' \in H(\mathrm{div}, \Omega_1);\ \mathbf{G}'_{|K} \in RT_0,\ \forall K \in \mathcal{T}_h^1\}, \end{aligned}$$

then we put

$$\mathbf{L}_h = L_h \times L_h, \quad \mathbf{V}_h^0 = (V_h \times V_h) \cap \mathbf{H}^0(\mathrm{div}, \Omega_1)$$

and we also introduce the affine set

$$\mathbf{V}_h^* = \{(\mathbf{G}', \mathbf{q}') \in V_h \times V_h;\ \mathbf{G}' \cdot \mathbf{n} = \mathcal{I}_h G^* \text{ on } \Upsilon_G,\ \mathbf{q}' \cdot \mathbf{n} = \mathcal{I}_h q^* \text{ on } \Upsilon_q\}$$

where $\mathcal{I}_h G^*$ and $\mathcal{I}_h q^*$ are piecewise constant approximations of the boundary data.

The convective term $\mathcal{D}(\mathbf{s})$ is treated by an upwind scheme, similarly to Lesaint and Raviart **[119]**. More precisely, $D(\cdot, \cdot)$ is approximated on $\mathbf{L}_h \times \mathbf{L}_h$ by

$$D_h(\mathbf{s}, \mathbf{s}') = I_h(T, T') + J_h(p, T'),$$

where:

$$\begin{aligned} I_h(T, T') &= \sum_{e \in \varepsilon_h} \int_e F_e(T', -\mathbf{G}_h^{n-1}, \mathbf{n}_e)[kT]\mathrm{ds} = \sum_{e \in \varepsilon_h} \int_e F_e(T, \mathbf{G}_h^{n-1}, \mathbf{n}_e)[kT']\mathrm{ds}, \\ J_h(p, T') &= \sum_{e \in \varepsilon_h} \int_e F_e(T', -\mathbf{G}_h^{n-1}, \mathbf{n}_e)[lp]\mathrm{ds} = \sum_{e \in \varepsilon_h} \int_e F_e(p, \mathbf{G}_h^{n-1}, \mathbf{n}_e)[lT']\mathrm{ds}, \end{aligned}$$

the second expressions being obtained after integration by parts. The numerical flux is given by:

$$F_e(T, \mathbf{G}_h^{n-1}, \mathbf{n}_e) = (\mathbf{G}_h^{n-1} \cdot \mathbf{n}_e)^+ T^{\text{in}} + (\mathbf{G}_h^{n-1} \cdot \mathbf{n}_e)^- T^{\text{ex}}.$$

We recall that $[T] = T^{\text{in}} - T^{\text{ex}}$ on an internal edge; on a boundary edge, we take $T^{\text{ex}} = T^{\text{in}}$ if e belongs to $\partial\Omega_1^+$ and $T^{\text{ex}} = T^*$ otherwise, which yields the additionnal right-hand side term

$$F_{3h}(\mathbf{s}') = \sum_{e \in \partial\Omega_1^-} \left(\int_e k \mathbf{G}_h^{n-1} \cdot \mathbf{n} T^* T' \mathrm{ds} + \int_e l \mathbf{G}_h^{n-1} \cdot \mathbf{n} p^* T'_K \mathrm{ds} \right).$$

We can now write the discrete problem as below:

$$\text{(3.1.6)} \qquad \begin{cases} \mathbf{V}_h \in \mathbf{V}_h^*, \ \mathbf{s}_h \in \mathbf{L}_h \\ A(\mathbf{V}_h, \mathbf{V}') + B(\mathbf{s}_h, \mathbf{V}') = F_{1h}(\mathbf{V}') \quad \forall \mathbf{V}' \in \mathbf{V}_h^0 \\ -B(\mathbf{s}', \mathbf{V}_h) + (C + D_h)(\mathbf{s}_h, \mathbf{s}') = F_{2h}(\mathbf{s}') + F_{3h}(\mathbf{s}') \quad \forall \mathbf{s}' \in \mathbf{L}_h \end{cases}.$$

Concerning the continuity of $I_h(\cdot,\cdot)$ and $J_h(\cdot,\cdot)$, it has been proved in [A5] that:

LEMMA 3.1.4. *Suppose that $\mathcal{T}_h^1$ satisfies the inverse hypothesis $h_1 \leq c h_K$. Then there exist positive constants c_1 and c_2 independent of h_1 such that for any p, T, $T' \in L_h$ one has:*

$$\begin{aligned} \left| I_h(T, T') \right| &\leq \frac{c_1}{h^2} \| \boldsymbol{G}_h^{n-1} \|_{0,\Omega_1} \|T\|_{0,\Omega_1} \|T'\|_{0,\Omega_1}, \\ \left| J_h(p, T') \right| &\leq \frac{c_2}{h^2} \| \boldsymbol{G}_h^{n-1} \|_{0,\Omega_1} \|p\|_{0,\Omega_1} \|T'\|_{0,\Omega_1}. \end{aligned}$$

3.1.3.2. *Existence and uniqueness of a solution.* The well-posedness of the discrete problem (3.1.6) follows by applying the same variant of the Babuška-Brezzi theory as in the continuous case. Since $\mathrm{Ker_h} B \subset \mathrm{Ker} B$, one has that $A(\cdot,\cdot)$ is uniformly $\mathbf{H}(\mathrm{div}, \Omega_1)$-elliptic on the discrete kernel of $B(\cdot,\cdot)$. The discrete *inf-sup* condition on $B(\cdot,\cdot)$ is also uniformly satisfied.

LEMMA 3.1.5. *For Δt sufficiently small, one has that:*

$$(C + D_h)(\mathbf{s}, \mathbf{s}) \geq 0, \quad \forall \mathbf{s} \in \mathbf{L}_h.$$

Proof. It is known that $I_h(\cdot,\cdot)$ is positive (see for instance **[71]** for a proof). By means of Lemma 3.1.4 and of **(A3)**, one gets:

$$(C + D_h)(\mathbf{s}, \mathbf{s}) \geq \frac{c}{\Delta t} (\|p\|_{0,\Omega_1}^2 + \|T\|_{0,\Omega_1}^2) - \frac{c_2}{h_1^2} \|\mathbf{G}_h^{n-1}\|_{0,\Omega_1} \|p\|_{0,\Omega_1} \|T\|_{0,\Omega_1}.$$

So, for

$$\text{(3.1.7)} \qquad \Delta t \leq \frac{2 c h_1^2}{c_2 \|\mathbf{G}_h^{n-1}\|_{0,\Omega_1}}$$

one deduces the announced statement. ∎

REMARK 3.1.6. If one considers p^{n-1} instead of p^n in the energy equation, then $D_h(\cdot,\cdot)$ is positive without any condition on Δt.

We can now deduce the invertibility of the matrix $\mathcal{A}_h = \begin{pmatrix} A & B \\ -B^T & (C + D_h) \end{pmatrix}$, and thus the existence and uniqueness of the solution to (3.1.6).

REMARK 3.1.7. Concerning the convergence of the approximation method, one cannot directly apply the classical error estimates for mixed formulations, since the continuous problem (3.1.1) does not satisfy the Babuška-Brezzi conditions. Nevertheless, one has for any $\boldsymbol{\tau}_h \in \mathbf{V}_h^0 \times \mathbf{L}_h$:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\Omega_1} \leq \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\Omega_1} + c \sup_{\boldsymbol{\tau}'_h \in \mathbf{V}_h^0 \times \mathbf{L}_h} \frac{\mathcal{A}_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{\tau}'_h)}{\|\boldsymbol{\tau}'_h\|_{\Omega_1}}.$$

3.1.3.3. A posteriori *error analysis.* We have also defined and implemented residual-based *a posteriori* error estimators, following Verfürth and Braess [**165**]. An error analysis was carried out with respect to the mesh-dependent norms on $\mathbf{V}_h^0$ and $\mathbf{L}_h$:

$$\begin{aligned} |||\mathbf{V}|||_h^2 &= \|\mathbf{G}\|_{0,\Omega_1}^2 + \|\mathbf{q}\|_{0,\Omega_1}^2 + \sum_{e \in \mathcal{E}_h} h_e \|\mathbf{G} \cdot \mathbf{n}\|_{0,e}^2 + \sum_{e \in \mathcal{E}_h} h_e \|\mathbf{q} \cdot \mathbf{n}\|_{0,e}^2, \\ |\mathbf{s}|_{1,h}^2 &= \sum_{K \in \mathcal{T}_h^1} |p|_{1,K}^2 + \sum_{K \in \mathcal{T}^\infty{}_h} |T|_{1,K}^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[p]\|_{0,e}^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[T]\|_{0,e}^2. \end{aligned}$$

$A(\cdot,\cdot)$ and $B(\cdot,\cdot)$ uniformly satisfy the Babuška-Brezzi conditions with respect to $|||\cdot|||_h$ and $|\cdot|_{1,h}$. Although the continuity constant of $(C + D_h)(\cdot,\cdot)$ depends on h_1 and on $\triangle t$, a uniform stability property still holds true because the norm of $\mathcal{A}_h^{-1}$ is independent of the norm of $C + D_h$.

We have defined in [A5] residuals $\eta_1(K)$, $\eta_2(K)$ on any triangle $K \in \mathcal{T}_h^1$ and $\eta(e)$ on any edge $e \in \mathcal{E}_h$, as well as a local error indicator:

$$\eta^2(K) = \eta_1^2(K) + h_K^2 \eta_2^2(K) + \frac{1}{2} \sum_{e \in \partial K \cap \mathcal{E}_h^{int}} h_e^{-1} \eta^2(e) + \sum_{e \in \partial K \cap \partial\Omega_1} h_e^{-1} \eta^2(e)$$

and a global one, $\eta^2 = \sum_{K \in \mathcal{T}_h^1} \eta^2(K)$. Under the saturation assumption:

$$\exists \beta < 1, \quad |||\mathbf{V} - \mathbf{V}_{h/2}|||_{h/2} + |\mathbf{s} - \mathbf{s}_{h/2}|_{1,h/2} \leq \beta \left(|||\mathbf{V} - \mathbf{V}_h|||_{h/2} + |\mathbf{s} - \mathbf{s}_h|_{1,h/2} \right),$$

we have proved similarly to [**165**] the following upper bound for the error:

THEOREM 3.1.8. *One has that:*

$$|||\boldsymbol{V}_{h/2} - \boldsymbol{V}_h|||_{h/2} + |s_{h/2} - s_h|_{1,h/2} \leq c\eta$$

with c independent of h and Δt. Hence, the following error bound holds:

$$|||\boldsymbol{V} - \boldsymbol{V}_h|||_h + |\mathbf{s} - \mathbf{s}_h|_{1,h} \leq \frac{c}{1-\beta}\eta.$$

The *a posteriori* error estimator also yields a local lower bound.

3.1.4. Numerical simulations. The reservoir model has been thoroughly validated from a numerical point of view in [A5] and [**71**]. Mesh convergence was studied, influence of the Joule-Thomson effect (which tends to cool a gas and to warm an oil) was illustrated, comparisons with measured data were carried out and *a posteriori* estimators were implemented. A realistic reservoir was also treated and we have obtained physically acceptable results. We present here only one test, a comparison with the well-test software PIE [1].

Well-testing consists in varying the flow rate in the well and then in measuring and interpreting the variations in pressure versus time, to get more information about the reservoir. Different softwares such as PIE are employed in the petroleum engineering community. For a given flowrate history, they are able to analytically evaluate the distribution of pressure in the reservoir (cf. [**36**]) by using Fourier and Laplace transforms. These analytical simulators only work in simplified frameworks and do not take into account the energetic aspect.

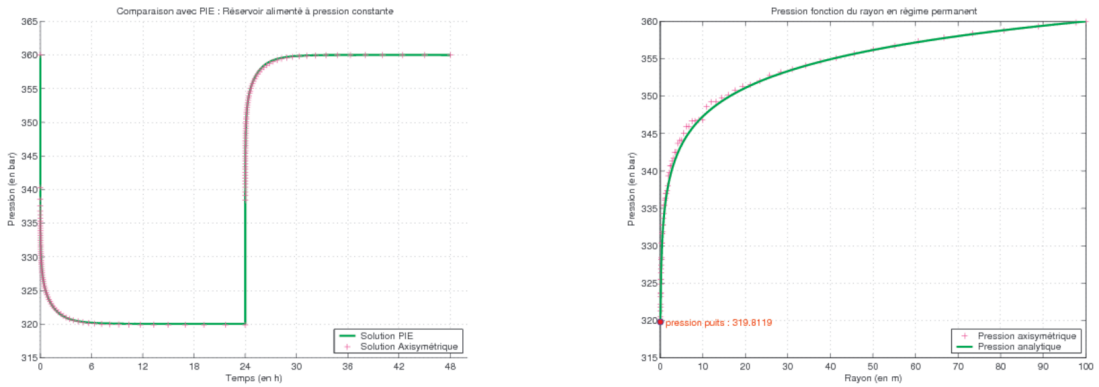
[1] *www.welltestsolutions.com*

We show next a comparison between our computed pressure and the one given by PIE. We consider the case of a mono-layer reservoir with constant physical and thermodynamic coefficients and with horizontal and vertical permeabilities $k_1 = 100\,\text{mD}$ and $k_2 = 1\,\text{mD}$; we take $F = 0$, $g = 0$. The considered reservoir is characterized by a constant pressure $p^* = 360\,\text{bar}$ on its external boundary. A production at constant flow rate $G^* = 150\,\text{m}^3/\text{day}$ is simulated during the first 24 hours and is followed by a shut-in period ($G^* = 0$) during the next 24 hours.

As expected, we observe in Fig. 3.1.2 (a) that during the draw-down period, the flow regime goes through a transitory state to reach a permanent one. This permanent state is characterized by a constant wellbore pressure given by

$$p_{\text{well}} = p^* - \alpha \frac{G^* B \mu}{K h} \ln \frac{R_\infty}{R}, \tag{3.1.8}$$

where α is a conversion factor and B the volume factor. Still in permanent regime, we can modify (3.1.8) to estimate, at a given z, the pressure at any r. Fig. 3.1.2 (b) shows that our solution (in blue) and the analytical one (in red) are very closed. Similar results were obtained for a closed reservoir, a situation which oftenly arises in practice.



(a) Pressure evolution during 24 h (b) Pressure at a given z in a permanent regime

FIGURE 3.1.2. Reservoir fed at constant pressure: comparison with PIE

3.2. Quasi-1D anisothermal Navier-Stokes equations

The wellbore model was studied in [B1] and briefly recalled in [A9]; I also refer to the PhD thesis of Bertrand Denel [**71**] for another variant and for further numerical validation. We have studied an axisymmetric vertical wellbore model, based on compressible Navier-Stokes type equations coupled with an energy equation. Due to the particular geometry and flow, we have derived and analyzed a 1.5D model, by constructing an explicit solution in terms of the radial coordinate and by treating carefully the horizontal inflow/outflow boundary condition at the perforations. We have then proposed a well-posed conforming finite element approximation. The developed code was validated numerically.

3.2.1. Physical problem in axisymmetric framework. The governing kinematic equations are the mass conservation law and the Navier-Stokes equations with a source term which takes into account the friction at the pipe's surface. We also consider the energy equation and we close the system by the Peng-Robinson state equation. As for the reservoir, the problem is written in 2D axisymmetric form. The 2D domain merely consists of:

$$\Omega_2 = \{(r, z);\ 0 \le r \le R,\ z \in I\}$$

where $I = [z_1, z_2]$. In practice, $R \simeq 4\,\text{in}$ while the length of the pipe can attend several thousands meters.

The problem in cylindrical coordinates (r, z) is described by:

$$\begin{cases} \frac{\partial}{\partial t}(r\rho) + \nabla \cdot (r\rho\mathbf{u}) = 0 \\ \frac{\partial}{\partial t}(r\rho u_r) + \nabla \cdot (ru_r\rho\mathbf{u}) + r\frac{\partial p}{\partial r} - \frac{\partial}{\partial r}(r\tau_{rr}) - \frac{\partial}{\partial z}(r\tau_{zr}) + \tau_{\theta\theta} + r\kappa\rho|\mathbf{u}|u_r = 0 \\ \frac{\partial}{\partial t}(r\rho u_z) + \nabla \cdot (ru_z\rho\mathbf{u}) + r\frac{\partial p}{\partial z} - \frac{\partial}{\partial r}(r\tau_{rz}) - \frac{\partial}{\partial z}(r\tau_{zz}) + r\rho g + r\kappa\rho|\mathbf{u}|u_z = 0 \\ \frac{\partial}{\partial t}(r\rho E) + \nabla \cdot (r(\rho E + p)\mathbf{u}) - \nabla \cdot (r\underline{\tau}\mathbf{u}) - \nabla \cdot (r\lambda\nabla T) + r\rho g u_z = 0 \\ \rho = \rho(p, T) \end{cases}$$

where $\mathbf{u} = (u_r, u_z)^t$ and where the tensor $\underline{\tau}$ is defined by

$$\tau_{rr} = 2\mu\frac{\partial u_r}{\partial r} - \frac{2}{3}\mu\left(\frac{1}{r}\frac{\partial}{\partial r}(ru_r) + \frac{\partial u_z}{\partial z}\right), \qquad \tau_{rz} = \tau_{zr} = \mu\left(\frac{\partial u_z}{\partial r} + \frac{\partial u_r}{\partial z}\right),$$
$$\tau_{zz} = 2\mu\frac{\partial u_z}{\partial z} - \frac{2}{3}\mu\left(\frac{1}{r}\frac{\partial}{\partial r}(ru_r) + \frac{\partial u_z}{\partial z}\right), \qquad \tau_{\theta\theta} = 2\mu\frac{u_r}{r} - \frac{2}{3}\mu\left(\frac{1}{r}\frac{\partial}{\partial r}(ru_r) + \frac{\partial u_z}{\partial z}\right).$$

Here above, $E = c_v T + \frac{|\mathbf{u}|^2}{2}$ is the total energy, c_v the specific heat and κ a positive coefficient depending on the diameter of the pipe. We assume in what follows that $\rho_1 \geq \rho(z) \geq \rho_0 > 0$ a.e. on Σ and $\lambda_1 \geq \lambda \geq \lambda_0 \geq 0$ a.e. in Ω_2.

3.2.2. Derivation of the 1.5D wellbore model. The flow in the wellbore is essentially vertical. In order to take into account this privileged direction, the particular geometry of the domain as well as the supply at the perforations, we proposed a 1.5D model. Thus, the computational cost is reduced and moreover, one avoids the numerical instabilities due to the large aspect ratio of a 2D grid on Ω_2.

In what follows, I briefly present the derivation of the simplified wellbore model. One first introduces the specific flux $\mathbf{G} = \rho\mathbf{u}$, the heat flux $\mathbf{q} = \lambda\nabla T$ and a time discretization which yields, at each time step, a nonlinear system. A fixed point method with respect to the density is then applied and the proposed algorithm consists in solving, for a given ρ, three decoupled problems:

$$\text{div}(r\mathbf{G}) = -r\frac{\rho - \rho^{n-1}}{\Delta t}, \tag{3.2.1}$$

$$\begin{cases} \text{div}(r\mathbf{u}) = \frac{1}{\rho}(\text{div}(r\mathbf{G}) - \frac{r}{\rho}\mathbf{G}\cdot\nabla\rho) \\ r\rho\frac{\mathbf{u}}{\Delta t} + r\mathbf{G}\cdot\nabla\mathbf{u} + r\nabla p - \text{div}(r\underline{\tau}) + \tau_{\theta\theta}\mathbf{e}_r + r\kappa|\mathbf{G}|\mathbf{u} = r\rho\mathbf{g} + r\rho\frac{\mathbf{u}^{n-1}}{\Delta t} \end{cases} \tag{3.2.2}$$

$$\begin{cases} rc_v\left(\rho\frac{T}{\Delta t} + \mathbf{G}\cdot\nabla T\right) - \text{div}(r\mathbf{q}) \\ = r\rho c_v\frac{T^{n-1}}{\Delta t} - \frac{1}{2}r\left(\rho\frac{|\mathbf{u}|^2 - |\mathbf{u}^{n-1}|^2}{\Delta t} + \mathbf{G}\cdot\nabla(|\mathbf{u}|^2)\right) - \text{div}(rp\mathbf{u}) + \text{div}(r\underline{\tau}\mathbf{u}) + r\mathbf{g}\cdot\mathbf{G} \\ \mathbf{q} = \lambda\nabla T. \end{cases} \tag{3.2.3}$$

Finally, the density is updated by means of a thermodynamic module and one loops until convergence is achieved. The first equation of (3.2.2) translates the fact that $\text{div}(r\mathbf{u}) = \text{div}(\frac{r}{\rho}\mathbf{G})$ while in the other equations we have simply substituted $\rho\mathbf{u}$ by $\mathbf{G}$. So, at this stage, the system (3.2.1)-(3.2.3) is deduced from but not equivalent to the initial one.

Next, in order to specify the boundary conditions, $\partial\Omega_2$ is divided into several parts as shown in Fig. 3.1.1 (b). We impose:

$$\left\{\begin{array}{l}\mathbf{G}\cdot\mathbf{n}=G_\Sigma \text{ on } \Sigma, \quad \mathbf{G}\cdot\mathbf{n}=0 \text{ on } \Gamma_2\cup\Gamma_3\cup\Gamma_4,\\ \mathbf{u}\cdot\mathbf{t}=0 \text{ on } \Sigma, \quad \underline{\tau}\mathbf{n}\cdot\mathbf{t}=0 \text{ on } \partial\Omega_2\setminus\Sigma,\\ \mathbf{u}\cdot\mathbf{n}=\frac{\mathbf{G}\cdot\mathbf{n}}{\rho} \text{ on } \partial\Omega_2\setminus\Sigma,\\ T=T_\Sigma \text{ on } \Sigma, \quad \mathbf{q}\cdot\mathbf{n}=0 \text{ on } \partial\Omega_2\setminus\Sigma.\end{array}\right.$$

We still have to prescribe a boundary condition on Σ, which we take of Neumann's type : $p-\underline{\tau}\mathbf{n}\cdot\mathbf{n}=p_\Sigma$.

REMARK 3.2.1. If one rather imposes a Dirichlet condition $\mathbf{u}\cdot\mathbf{n}=\frac{G_\Sigma}{\rho}$ on Σ, then one can show that the relation $\mathrm{div}(r\rho\mathbf{u})=\mathrm{div}(r\mathbf{G})$ implies $\rho\mathbf{u}=\mathbf{G}$ in Ω_2, which justifies the proposed algorithm. In this case, the radial velocity is completely determined, $u_r=\frac{G_r}{\rho}$ and the corresponding momentum equation is just neglected. The corresponding wellbore model was analyzed in [**71**]. Here, in view of the coupling with the reservoir, we impose a Neumann condition on Σ and we use later the relation $\rho\mathbf{u}\cdot\mathbf{n}=\mathbf{G}\cdot\mathbf{n}$ as an additional transmission condition.

A relevant issue concerns the boundary condition on the top of the wellbore. Let us note that, even if the flowrate $Q=\mathbf{G}\cdot\mathbf{n}$ is known thanks to recorded data, one cannot impose it on the outflow boundary Γ_1 since Q and G_Σ are related, according to (3.2.1), by the compatibility condition

$$\int_{\Omega_2} r\frac{\rho-\rho^{n-1}}{\Delta t}\mathrm{dx}+\int_{\Gamma_1} r\rho Q\mathrm{ds}+\int_\Sigma rG_\Sigma\mathrm{ds}=0.$$

Next, the 1.5D model is obtained as a conforming approximation of the 2D semi-discretized problem, by considering an explicit dependence of the unknowns on the radial coordinate. The velocity is taken affine with respect to r whereas the scalar unknowns only depend on z:

$$\begin{array}{l}\mathbf{u}=\begin{pmatrix}u_r\\ u_z\end{pmatrix}=\begin{pmatrix}\frac{r}{R}\overline{u}_r(z)\\ \frac{r}{R}\overline{u}_z(z)+\frac{R-r}{R}\widehat{u}_z(z)\end{pmatrix},\\ \mathbf{G}=\begin{pmatrix}G_r\\ G_z\end{pmatrix}=\begin{pmatrix}\frac{r}{R}\bar{G}_r(z)\\ G_z(z)\end{pmatrix}, \quad \mathbf{q}=\begin{pmatrix}q_r\\ q_z\end{pmatrix}=\begin{pmatrix}\frac{r}{R}\bar{q}_r(z)\\ q_z(z)\end{pmatrix},\\ \rho=\rho(z), \quad p=p(z), \quad T=T(z).\end{array} \tag{3.2.4}$$

Thanks to the boundary conditions, one further has $\overline{u}_r=0$ on Γ_2 and $\overline{u}_z=0$ on Σ.

3.2.3. Weak formulation. In order to write the time-discretized problem in weak form, we introduce the spaces:

$$\begin{array}{rcl}\mathcal{W} &=& \left\{\mathbf{w}=\left(\frac{r}{R}\bar{w}_r(z),w_z(z)\right)^t;\ \bar{w}_r\in L^2(I),\ w_z\in H^1(I)\right\}\subset H(\mathrm{div},\Omega_2),\\ \boldsymbol{H} &=& \left\{\mathbf{v}=\left(\frac{r}{R}\bar{v}_r(z),v_z(r,z)\right)^t;\ v_z=\frac{r}{R}\overline{v}_z(z)+\frac{R-r}{R}\widehat{v}_z(z),\ \ \bar{v}_r,\overline{v}_z,\widehat{v}_z\in H^1(I)\right\}\subset\mathbf{H}^1(\Omega_2)\\ M &=& \{q=q(z);\ q\in L^2(I)\}\subset L^2(\Omega_2)\end{array}$$

and we further consider

$$\begin{array}{rcl}W &=& \{\mathbf{w}\in\mathcal{W};\ \mathbf{w}\cdot\mathbf{n}=0 \ \text{ on } \partial\Omega_2\setminus\Sigma\},\\ \mathbf{H}^0 &=& \{\mathbf{v}\in\boldsymbol{H};\ \mathbf{v}\cdot\mathbf{n}=0 \ \text{ on } \partial\Omega_2\setminus\Sigma,\ \mathbf{v}\cdot\mathbf{t}=0 \ \text{ on } \Sigma\}\end{array}$$

as well as the affine sets:

$$\begin{array}{rcl}W^* &=& \{\mathbf{w}\in\mathcal{W};\ \mathbf{w}\cdot\mathbf{n}=G_\Sigma \text{ on } \Sigma,\ \mathbf{w}\cdot\mathbf{n}=0 \ \text{ on } \Gamma_2\cup\Gamma_3\cup\Gamma_4\},\\ \mathbf{H}^* &=& \{\mathbf{v}\in\boldsymbol{H};\ \mathbf{v}\cdot\mathbf{n}=Q \text{ on } \Gamma_1,\ \mathbf{v}\cdot\mathbf{n}=0 \ \text{ on } \Gamma_2\cup\Gamma_3\cup\Gamma_4,\ \mathbf{v}\cdot\mathbf{t}=0 \ \text{ on } \Sigma\}\end{array}$$

where Q denotes now $\frac{\mathbf{G}\cdot\mathbf{n}}{\rho}$ and is assumed to be constant.

We consider the following Petrov-Galerkin formulation of (3.2.1), respectively the mixed variational formulations of (3.2.2) and (3.2.3):

$$\left\{\begin{array}{l} \mathbf{G}\in W^* \\ \displaystyle\int_{\Omega_2}\mathrm{div}(r\mathbf{G})\chi\mathrm{dx} = -\int_{\Omega_2} r\frac{\rho-\rho^{n-1}}{\Delta t}\chi\mathrm{dx} \quad \forall\chi\in M, \end{array}\right. \tag{3.2.5}$$

$$\left\{\begin{array}{lcl} \mathbf{u}\in\mathbf{H}^*,\ p\in M & & \\ m(\mathbf{u},\mathbf{v}) + n(p,\mathbf{v}) = l_1(\mathbf{v}) & \forall\mathbf{v}\in\mathbf{H}^0 & \\ n(q,\mathbf{u}) = l_2(q) & \forall q\in M, & \end{array}\right. \tag{3.2.6}$$

$$\left\{\begin{array}{lcl} \mathbf{q}\in W,\ T\in M & & \\ a(\mathbf{q},\mathbf{w}) + b(T,\mathbf{w}) = f_1(\mathbf{w}) & \forall\mathbf{w}\in W & \\ b(S,\mathbf{q}) - c(T,S)-\alpha d(T,S) = f_2(S) & \forall S\in M. & \end{array}\right. \tag{3.2.7}$$

The parameter α is equal to 0 if one neglects the convective term in the energy equation, and to 1 for the full problem. The definition of the bilinear forms can be found in [B1], where the well-posedness of each of the previous problems has been established. In the case $\alpha = 0$, we have applied Babuška's theorem for (3.2.5) and Babuška-Brezzi theorem for (3.2.6) and (3.2.7), for Δt sufficiently small. Finally, the well-posedness of the complete problem with convective terms was proved similarly to the reservoir case, thanks to Fredholm's alternative, under a smoothness hypothesis for T_Σ and λ.

In view of its coupling with the reservoir model, we denote the unknowns of the wellbore model by $x_2 = (\mathbf{G}_2, \mathbf{u}_2, \mathbf{q}_2, p_2, T_2)$ and the associated test-functions by $x_2' = (\chi, \mathbf{u}_2', \mathbf{q}_2', p_2', T_2')$. They respectively belong to the following product spaces:

$$X_2 = \mathcal{W}\times\boldsymbol{H}\times W\times M\times M, \quad Y_2 = M\times\mathbf{H}^0\times W\times M\times M.$$

It is also useful to introduce the affine set $X_2^* = W^*\times\mathbf{H}^*\times W\times M\times M$. Then the nonlinear time-discretized 1.5D wellbore problem can be written as follows:

$$\left\{\begin{array}{l} x_2\in X_2^* \\ \tilde{\mathcal{A}}_2(x_2,x_2') = \mathcal{F}_2(x_2'),\ \forall x_2'\in Y_2. \end{array}\right. \tag{3.2.8}$$

3.2.4. Finite element approximation. We consider a specific 2D grid $\mathcal{T}_h^2$, consisting of only one cell in the radial direction and of a regular mesh in the z direction. We put $\overline{\Omega}_2 = \cup_{K\in\mathcal{T}_h^2}K$ with K rectangle of width R and of height h_K, and let $h_2 = \max_{K\in\mathcal{T}_h^2} h_K$. $\mathcal{E}_h^\Sigma$ denotes the set of edges situated on Σ. The pressure, the temperature (and hence, the density) as well as the multiplier λ are approached by P_0- elements, the specific and the heat fluxes by RT_0 elements and the velocity by $\mathbf{Q}_1$- continuous elements. Therefore, we define:

$$\begin{array}{rcl} \mathcal{W}_h &=& \left\{\mathbf{G}\in\mathcal{W};\ \mathbf{G}_{|K}\in RT_0,\ \forall K\in\mathcal{T}_h^2\right\}, \\ \boldsymbol{H}_h &=& \left\{v\in\boldsymbol{H};\ \mathbf{v}_{|K}\in\mathbf{Q}_1,\ \forall K\in\mathcal{T}_h^2\right\}, \\ M_h &=& \left\{\lambda\in M;\ \lambda_{|K}\in Q_0,\ \forall K\in\mathcal{T}_h^2\right\} \end{array}$$

and then we put:

$$W_h = \mathcal{W}_h \cap W, \quad W_h^0 = \mathcal{W}_h \cap W^0, \quad \mathbf{H}_h^0 = \boldsymbol{H}_h \cap \mathbf{H}^0.$$

The above choice of 2D finite dimensional spaces is compatible with the dependence in r prescribed in (3.2.4). For $\mathbf{G}$ and q, one can equivalently take $\overline{G}_r$, $\bar{q}_r$ piecewise constant on I and G_z, q_z P_1-continuous on I, whereas $\mathbf{u} \in \boldsymbol{H}_h$ is equivalent to $\bar{u}_r$, $\bar{u}_z$, $\hat{u}_z$ P_1- continuous on I.

The convective terms are treated by upwinding. The first one, $\int_{\Omega_2} r(\mathbf{G}\cdot\nabla)\mathbf{u}\cdot\mathbf{v}\mathrm{dx}$, is approximated on $\mathbf{H}_h^0 \times \mathbf{H}_h^0$ by

$$\tilde{m}_h(\mathbf{u},\mathbf{v}) = \int_{\Omega_2} r(\mathbf{G}\cdot\nabla)\mathbf{u}\cdot\mathbf{v}\mathrm{dx} + \sum_{K\in\mathcal{T}_h^2}\sum_{e\in\partial K^-}\int_e r\mathbf{G}_h\cdot\mathbf{n}\,[\pi_0\mathbf{u}]\cdot\mathbf{v}\,\mathrm{ds},$$

where π_0 is the piecewise constant $L^2(\Omega_2)$-projection. One has for all $\mathbf{u},\mathbf{v}\in\mathbf{H}_h^0$ that

$$|\tilde{m}_h(\mathbf{u},\mathbf{v})| \le \frac{c}{h_{2,min}}\|\mathbf{G}_h\|_{0,\Omega_2}\,\|\mathbf{u}\|_{1,\Omega_2}\,\|\mathbf{v}\|_{1,\Omega_2}\,. \tag{3.2.9}$$

The second one, $\int_{\Omega_2} rc_v(\mathbf{G}\cdot\nabla)T\,S\mathrm{dx}$, is approximated on $M_h\times M_h$ as for the reservoir model by a positive form $d_h(\cdot,\cdot)$.

Then I have established in [B1] the well-posedness of the discrete versions of (3.2.5), (3.2.6) and (3.2.7). The first problem is obviously well-posed, since one can locally compute $\mathbf{G}_h$ on every rectangle $K\in\mathcal{T}_h^2$. For Δt sufficiently small, the discrete velocity-pressure formulation (3.2.6) is shown to satisfy the Babuška-Brezzi theorem uniformly with respect to h_2 (see [**122**]). The condition on Δt is sufficient in order to ensure the coercivity of $m_h(\cdot,\cdot)$ on the discrete kernel of $n(\cdot,\cdot)$. So, taking into account (3.2.9) and the Friedrichs-Poincaré inequality, we deduce the desired coercivity for

$$\frac{1}{\Delta t} \ge \frac{c\|\mathbf{G}_h\|_{0,\Omega_2}}{\rho h_{2,min}} - \frac{\kappa\,|\mathbf{G}_h|}{\rho}. \tag{3.2.10}$$

Concerning the discretization of (3.2.7), the analysis is based on the same result of [**158**] as for the reservoir model. In particular, we have checked that $(c+d_h)(\cdot,\cdot)$ is positive without any additional condition, $a(\cdot,\cdot)$ is uniformly coercive on $\mathrm{Ker}_h b$, and $b(\cdot,\cdot)$ satisfies a uniform *inf-sup* condition.

The two wellbore models, with Dirichlet and Neumann boundary conditions on Σ, have been validated numerically. The behaviour of the solution with respect to mesh refinement was studied, and a more realistic case was also treated, were the wellbore model was interfaced with the previous reservoir simulator. We have shown that, near the perforations, the wellbore temperature and pressure are in good agreement with those computed by the reservoir code. Moreover, the behaviour of the pressure clearly agrees with the one given by the software Pie.

3.3. Coupling of the previous models

The coupling of the fluid and porous media has been treated in [A8] and [A9], see also the PhD thesis of Layal Lizaik [**122**]. To summarize, adequate transmission conditions at the perforations were imposed and dualized by means of Lagrange multipliers. In order to take into account recorded flowrates at the pipe's surface, we have turned towards a global resolution of the coupled problem at each time step. Thus, one ends up with a non-standard mixed formulation for which I managed to show the uniqueness of the solution. The well-posedness of the discrete problem was established by means of a technical analysis and consequently, the existence of a solution for the continuous problem has also been deduced. The coupled code was validated by numerical experiments.

3.3.1. Transmission conditions. We agree to denote by $\mathbf{n}$ the normal unit vector to the interface Σ, oriented from the reservoir towards the wellbore, and to index by 1 and 2 the unknowns related to the reservoir and the wellbore, respectively. The interface terms that have to be matched are those appearing by integration by parts in the 2D axisymmetric models, that is for the reservoir:

$$\int_\Sigma p_1 \mathbf{G}_1' \cdot \mathbf{n} \mathrm{ds} - \int_\Sigma T_1 \mathbf{q}_1' \cdot \mathbf{n} \mathrm{ds},$$

respectively for the wellbore:

$$\int_\Sigma R(p_2 - \underline{\tau_2}\mathbf{n}\cdot\mathbf{n})\mathbf{u}_2' \cdot \mathbf{n} \mathrm{ds} - \int_\Sigma R T_2 \mathbf{q}_2' \cdot \mathbf{n} \mathrm{ds} - \int_\Sigma R(\underline{\tau_2}\mathbf{n}\cdot\mathbf{t})\mathbf{u}_2' \cdot \mathbf{t} \mathrm{ds}.$$

We classically impose the mass conservation and the balance of normal forces on the interface:

$$\mathbf{G}_1 \cdot \mathbf{n} = \mathbf{G}_2 \cdot \mathbf{n}, \quad -p_1 = -p_2 + \underline{\tau_2}\mathbf{n}\cdot\mathbf{n}. \tag{3.3.1}$$

Due to the viscous context, one also has to prescribe a condition on the tangential velocity of the fluid. The one which seems to be in best agreement with experimental evidence is the Beavers-Joseph-Saffman law and it reads $\mathbf{u}_2 \cdot \mathbf{t} = -\frac{\sqrt{k}}{\delta}\underline{\sigma_2}\mathbf{n}\cdot\mathbf{t}$, with $\delta > 0$ a parameter experimentally determined (see for instance **[114]** and references therein). However, the mathematical analysis doesn't lose in generality if one simply takes (as in **[33]** or **[75]**) $\mathbf{u}_2 \cdot \mathbf{t} = 0$, since the Beavers-Joseph-Saffman condition only enhances the coercivity of the main operator. In what follows, we impose in agreement with the wellbore model that

$$\mathbf{u}_2 \cdot \mathbf{t} = 0 \quad \text{on } \Sigma. \tag{3.3.2}$$

The energetic aspect implies the continuity of the temperature and of the normal heat flux across Σ:

$$T_1 = T_2, \quad \mathbf{q}_1 \cdot \mathbf{n} = \mathbf{q}_2 \cdot \mathbf{n}. \tag{3.3.3}$$

Furthermore, we bind together the unknowns on Σ by imposing:

$$\rho_2 \mathbf{u_2} \cdot \mathbf{n} = \mathbf{G_2} \cdot \mathbf{n}. \tag{3.3.4}$$

In conclusion, the set of transmission conditions consists of (3.3.1) - (3.3.4).

3.3.2. Analysis of the coupled problem. Similarly to Layton *et al.* **[114]** or to **[75]**, we write a mixed weak formulation linking together the two formulations (3.1.4) and (3.2.8). For the sake of simplicity, we linearize the wellbore model and we replace at each t^n, $\mathbf{G}_2$ by $\mathbf{G}_2^{n-1}$ in the corresponding momentum and energy equations.

We next dualize the transmission conditions on Σ by means of Lagrange multipliers. Let us first introduce the following spaces, obtained by removing the boundary conditions on Σ and by adding more regularity on the normal traces of $\mathbf{G}_1$, $\mathbf{q}_1$ on Σ:

$$\begin{aligned}
\mathbb{X} &= \{x = (x_1, x_2) \in X_1 \times X_2 \,;\; \mathbf{G}_1 \cdot \mathbf{n},\, \mathbf{q}_1 \cdot \mathbf{n} \in L^2(\Sigma)\},\\
\mathbb{Y} &= \{x' = (x_1', x_2') \in X_1 \times Y_2 \,;\; \mathbf{G}_1 \cdot \mathbf{n},\, \mathbf{q}_1 \cdot \mathbf{n} \in L^2(\Sigma)\},\\
\mathbb{Y}^0 &= \{x' \in \mathbb{Y} \,;\; \mathbf{G}_1' \cdot \mathbf{n} = 0 \text{ on } \Upsilon_{\mathbf{G}} \setminus \Sigma,\; \mathbf{q}_1' \cdot \mathbf{n} = 0 \text{ on } \Upsilon_{\mathbf{q}} \setminus \Sigma,\; \mathbf{u}_2' \cdot \mathbf{n} = 0 \text{ on } \Gamma_1\},\\
\mathbb{X}^* &= \{x \in \mathbb{X} \,;\; \mathbf{G}_1 \cdot \mathbf{n} = 0 \text{ on } \Upsilon_{\mathbf{G}} \setminus \Sigma,\; \mathbf{q}_1 \cdot \mathbf{n} = q^* \text{ on } \Upsilon_{\mathbf{q}} \setminus \Sigma,\; \mathbf{u}_2 \cdot \mathbf{n} = Q \text{ on } \Gamma_1\}.
\end{aligned}$$

The Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ are endowed with the graph norms.

Let the multipliers' spaces:

$$\mathbb{L} = (L^2(\Sigma))^2, \quad \mathbb{K} = (L^2(\Sigma))^3$$

and the bilinear forms on $\mathbb{L}\times\mathbb{Y}$, respectively $\mathbb{K}\times\mathbb{X}$:

$$\begin{aligned}\mathcal{I}(\Lambda,x') &= \int_\Sigma(\mathbf{G}_1'\cdot\mathbf{n}-R\mathbf{u}_2'\cdot\mathbf{n})\theta\mathrm{ds}-\int_\Sigma(\mathbf{q}_1'\cdot\mathbf{n}-R\mathbf{q}_2'\cdot\mathbf{n})\mu\mathrm{ds},\\ \mathcal{J}(\Lambda',x) &= \int_\Sigma(\mathbf{G}_1\cdot\mathbf{n}-R\rho_2\mathbf{u}_2\cdot\mathbf{n})\theta'\mathrm{ds}+\int_\Sigma(\mathbf{G}_1\cdot\mathbf{n}-R\mathbf{G}_2\cdot\mathbf{n})\zeta'\mathrm{ds}-\int_\Sigma(\mathbf{q}_1\cdot\mathbf{n}-R\mathbf{q}_2\cdot\mathbf{n})\mu'\mathrm{ds}\end{aligned}$$

for any $\Lambda=(\theta,\mu)\in\mathbb{L}$ and $\Lambda'=(\zeta',\theta',\mu')\in\mathbb{K}$. Then, putting

$$\begin{aligned}\mathcal{A}(x,x') &= \mathcal{A}_1(x_1,x_1')+\mathcal{A}_2(x_2,x_2'),\quad\forall x\in\mathbb{X},\forall x'\in\mathbb{Y},\\ \mathcal{F}(x') &= \mathcal{F}_1(x_1')+\mathcal{F}_2(x_2'),\quad\forall x'\in\mathbb{Y},\end{aligned}$$

the coupled problem can be written as follows:

$$\left\{\begin{array}{l} x\in\mathbb{X}^*,\Lambda\in\mathbb{L}\\ \\ \mathcal{A}(x,x') + \mathcal{I}(\Lambda,x') = \mathcal{F}(x'),\ \forall x'\in\mathbb{Y}^0\\ \\ \mathcal{J}(\Lambda',x) = 0,\ \forall\Lambda'\in\mathbb{K}.\end{array}\right. \tag{3.3.5}$$

The multiplier $\Lambda=(\theta,\mu)$ can be interpreted as (p_1,T_1), or still as $(p_2-\underline{\tau}_2\mathbf{n}\cdot\mathbf{n},T_2)$.

I have first established that $\mathcal{I}$ and $\mathcal{J}$ satisfy each an *inf-sup* condition.

Lemma 3.3.1. *The following conditions hold:*

$$\begin{aligned}\exists b_1 &> 0,\ \forall\Lambda\in\mathbb{L},\quad \sup_{x'\in\mathbb{Y}^0}\frac{\mathcal{I}(\Lambda,x')}{\|x'\|_{\mathbb{Y}}}\geq b_1\|\Lambda\|_{0,\Sigma},\\ \exists b_2 &> 0,\ \forall\Lambda'\in\mathbb{K},\quad \sup_{x\in\mathbb{X}^0}\frac{\mathcal{J}(\Lambda',x)}{\|x\|_{\mathbb{X}}}\geq b_2\left\|\Lambda'\right\|_{0,\Sigma}.\end{aligned}$$

Therefore, according to the general theory of saddle point problems, it is sufficient to study:

$$\left\{\begin{array}{l} x\in\mathbb{J}^*\\ \\ \mathcal{A}(x,x')=\mathcal{F}(x'),\ \forall x'\in\mathbb{I}\end{array}\right. \tag{3.3.6}$$

where :

$$\mathbb{J}^*=\left\{x\in\mathbb{X}^*\,;\,\mathcal{J}(\Lambda',x)=0,\ \forall\Lambda'\in\mathbb{K}\right\},\quad \mathbb{I}=\left\{x'\in\mathbb{Y}^0\,;\,\mathcal{I}(\Lambda,x')=0,\ \forall\Lambda\in\mathbb{L}\right\}.$$

By separating the vector functions from the scalar ones and by consequently putting $\mathbb{J}^*=\mathbb{U}^*\times\mathbb{S}$ and $\mathbb{I}=\mathbb{T}\times\mathbb{S}$, one can still write (3.3.6) as follows:

$$\left\{\begin{array}{ll}(\mathbf{U},s)\in\mathbb{U}^*\times\mathbb{S}\\ \\ \mathtt{A}(\mathbf{U},\mathbf{U}')+\mathtt{B}(s,\mathbf{U}')=\mathtt{F}_1(\mathbf{U}'), & \forall\mathbf{U}'\in\mathbb{T}^0\\ \\ \mathtt{B}(s',\mathbf{U})-\mathtt{C}(s,s')-\alpha\mathtt{D}(s,s')=\mathtt{F}_2(s'), & \forall s'\in\mathbb{S}\end{array}\right. \tag{3.3.7}$$

where $\mathbf{U} = (\mathbf{G}_1, \mathbf{q}_1, \mathbf{G}_2, \mathbf{u}_2, \mathbf{q}_2)$, the test-function $\mathbf{U}'$ stands for $(\mathbf{G}'_1, \mathbf{q}'_1, \chi, \mathbf{u}'_2, \mathbf{q}'_2)$ and $s = (p_1, T_1, p_2, T_2)$. Here above, we have put:

$$\mathtt{A}(\mathbf{U}, \mathbf{U}') = \int_{\Omega_1} \frac{1}{r} \underline{M} \mathbf{G}_1 \cdot \mathbf{G}'_1 \mathrm{d}x + \int_{\Omega_1} \frac{1}{r\lambda_1} \mathbf{q}_1 \cdot \mathbf{q}'_1 \mathrm{d}x$$
$$+ \int_{\Omega_2} \chi \mathrm{div}(r\mathbf{G}_2) \mathrm{d}x + \int_{\Omega_2} \frac{r}{\lambda_1} \mathbf{q}_2 \cdot \mathbf{q}'_2 \mathrm{d}x + m(\mathbf{u}_2, \mathbf{u}'_2),$$

$$\mathtt{B}(s, \mathbf{U}') = -\int_{\Omega_1} p_1 \mathrm{div} \mathbf{G}'_1 \mathrm{d}x + \int_{\Omega_1} T_1 \mathrm{div} \mathbf{q}'_1 \mathrm{d}x - \int_{\Omega_2} p_2 \mathrm{div}(r\mathbf{u}'_2) \mathrm{d}x + \int_{\Omega_2} T_2 \mathrm{div}(r\mathbf{q}'_2) \mathrm{d}x,$$

$$\mathtt{C}(s, s') = \int_{\Omega_1} r \frac{a}{\Delta t} p_1 p'_1 \mathrm{d}x - \int_{\Omega_1} r \frac{b}{\Delta t} T_1 p'_1 \mathrm{d}x$$
$$+ \int_{\Omega_1} r \frac{d}{\Delta t} T_1 T'_1 \mathrm{d}x - \int_{\Omega_1} r \frac{f}{\Delta t} p_1 T'_1 \mathrm{d}x + \int_{\Omega_2} r \frac{c_v \rho_2}{\Delta t} T_2 T'_2 \mathrm{d}x,$$

$$\mathtt{D}(s, s') = \int_{\Omega_1} \kappa \mathbf{G}_1^{n-1} \cdot \nabla T_1 T'_1 \mathrm{d}x + \int_{\Omega_1} l \mathbf{G}_1^{n-1} \cdot \nabla p_1 T'_1 \mathrm{d}x + \int_{\Omega_2} r\ c_v \mathbf{G}_2^{n-1} \cdot \nabla T_2 S_2 \mathrm{d}x.$$

Note that neither $\mathtt{A}(\cdot,\cdot)$ nor $\mathtt{C}(\cdot,\cdot)$ are symmetric and moreover, the spaces employed for the solution and the test-functions are different. Hence, one cannot apply the existing generalizations of the Babuška-Brezzi theorem ([**47**], [**137**] or [**158**]) in the case $\alpha = 0$. I have then established the following preliminary results, which allow to prove that the operator $\mathcal{A} = \begin{pmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{B}^T & -\mathtt{C} \end{pmatrix}$ is injective.

LEMMA 3.3.2. *There exist two constants β_1 and β_2 independent of Δt such that:*

$$\forall s \in \mathbb{S}, \quad \sup_{\mathbf{U} \in \mathbb{U}^0} \frac{\mathtt{B}(s, \mathbf{U})}{\|\mathbf{U}\|} \geq \beta_1 \|s\|, \qquad \sup_{\mathbf{U}' \in \mathbb{T}^0} \frac{\mathtt{B}(s, \mathbf{U}')}{\|\mathbf{U}'\|} \geq \beta_2 \|s\|.$$

We follow the proofs of the *inf-sup* conditions related to the wellbore and the reservoir models.

LEMMA 3.3.3. *There exists a positive constant γ, depending on Δt, such that:*

$$\forall s \in \mathbb{S}, \ \ \mathtt{C}(s, s) \geq \gamma (\|p_1\|^2_{0,\Omega_1} + \|T_1\|^2_{0,\Omega_1} + \|T_2\|^2_{0,\Omega_2}).$$

The proof follows from the study of the reservoir model. Note that $\mathtt{C}$ is not positive definite, since the norm of p_2 is missing from the previous estimate.

LEMMA 3.3.4. *For Δt sufficiently small, the following statement holds:*

$$\forall \mathbf{U} \in \mathbb{U}^0 \setminus \{0\}, \qquad \sup_{\substack{\mathbf{U}' \in \mathbb{T}^0 \\ \mathbf{U} - \mathbf{U}' \in \mathrm{Ker}\mathtt{B}}} \frac{\mathtt{A}(\mathbf{U}, \mathbf{U}')}{\|\mathbf{U}\| \|\mathbf{U}'\|} > 0.$$

Proof. It is sufficient to construct a linear continuous operator $\mathcal{R} : \mathbb{U}^0 \to \mathbb{T}^0$ satisfying:

$$\mathtt{B}(s, \mathbf{U}) = \mathtt{B}(s, \mathcal{R}\mathbf{U}), \quad \forall s \in \mathbb{S},$$
$$\mathtt{A}(\mathbf{U}, \mathcal{R}\mathbf{U}) > 0, \quad \forall \mathbb{U}^0 \setminus \{0\}.$$

Any $\mathbf{U} = (\mathbf{G}_1, \mathbf{q}_1, \mathbf{G}_2, \mathbf{u}_2, \mathbf{q}_2) \in \mathbb{U}^0$ satisfies $\mathbf{G}_1 \cdot \mathbf{n} = R\mathbf{G}_2 \cdot \mathbf{n} = R\rho_2 \mathbf{u}_2 \cdot \mathbf{n}$ and $\mathbf{q}_1 \cdot \mathbf{n} = R\mathbf{q}_2 \cdot \mathbf{n}$ on Σ. Then we take $\mathcal{R}\mathbf{U} = \mathbf{U}' = (\mathbf{G}'_1, \mathbf{q}_1, \chi, \mathbf{u}_2, \mathbf{q}_2)$, where $\mathbf{G}'_1$ and χ are chosen such that:

$$\mathbf{G}'_1 \cdot \mathbf{n} = \frac{1}{\rho_2} \mathbf{G}_1 \cdot \mathbf{n} \text{ on } \Sigma, \quad \mathrm{div} \mathbf{G}'_1 = \mathrm{div} \mathbf{G}_1 \text{ in } \Omega_1, \quad \left\| \mathbf{G}'_1 \right\|_{0,\Omega_1} + \|\chi\|_{0,\Omega_2} \leq c \|\mathbf{U}\|.$$

Then obviously $\mathbf{U}'$ belongs to $\mathbb{T}^0$, satisfies $\|\mathbf{U}'\| \leq c\|\mathbf{U}\|$ and $\mathbf{U} - \mathbf{U}' \in \mathrm{Ker}\mathtt{B}$. Moreover, one has:

$$\mathtt{A}(\mathbf{U}, \mathbf{U}') \geq c \left(\|\mathbf{q}_1\|^2_{0,\Omega_1} + \|\mathbf{q}_2\|^2_{0,\Omega_2} \right) + \int_{\Omega_1} \frac{1}{r} \underline{M} \mathbf{G}_1 \cdot \mathbf{G}'_1 \mathrm{d}x + \int_{\Omega_2} \chi \mathrm{div}(r\mathbf{G}_2) \mathrm{d}x + m(\mathbf{u}_2, \mathbf{u}_2).$$

The construction of $\mathbf{G}_1'$ and χ is quite technical, we refer to [A9] for details. Finally, by applying several times Young's inequality (see [A9] and [B1]), it is possible to choose Δt such that

$$\mathtt{A}(\mathbf{U},\mathbf{U}') \geq \alpha \left(\|\mathbf{G}_1\|^2_{0,\Omega_1} + \|\mathbf{q}_1\|^2_{0,\Omega_1} + \|\mathbf{u}_2\|^2_{1,\Omega_2} + \|\mathbf{q}_2\|^2_{0,\Omega_2} + \|\mathbf{G}_2\|^2_{H(\mathrm{div},\Omega_2)} \right) \tag{3.3.8}$$

with $\alpha > 0$, which ends the proof. ■

THEOREM 3.3.5. *For Δt sufficiently small, the following statement is true:*

$$\forall x \in \mathbb{J}^0 \setminus \{0\}, \quad \sup_{x' \in \mathbb{I}} \frac{\mathcal{A}(x,x')}{\|x'\|_{\mathbb{Y}}} > 0. \tag{3.3.9}$$

Therefore, problems (3.3.6) and (3.3.5) have at most one solution for $\alpha = 0$.

Proof. We focus on problem (3.3.6) and we prove that the homogeneous problem admits only the trivial solution. So, let $(\mathbf{U}, s) \in \mathbb{U}^0 \times \mathbb{S}$ satisfy:

$$\begin{cases} \mathtt{A}(\mathbf{U},\mathbf{U}') + \mathtt{B}(s,\mathbf{U}') = 0, & \forall \mathbf{U}' \in \mathbb{T}^0 \\ -\mathtt{B}(s',\mathbf{U}) + \mathtt{C}(s,s') = 0, & \forall s' \in \mathbb{S} \end{cases}$$

and let $s' = s$ and $\mathbf{U}' = \mathcal{R}\mathbf{U}$, where $\mathcal{R}$ is the operator introduced in Lemma 3.3.4. Then the positivity of $\mathtt{A}(\cdot,\cdot)$ and $\mathtt{C}(\cdot,\cdot)$ imply that $\mathbf{U} = \mathbf{0}$ and $(p_1, T_1, T_2) = \mathbf{0}$, and the second *inf-sup* condition of Lemma 3.3.2 yields that $p_2 = 0$. Finally, the uniqueness of the solution of (3.3.5) holds thanks to Lemma 3.3.1. ■

Note that the L^2-norms of div$\mathbf{G}_1$, div$\mathbf{q}_1$ and div$\mathbf{q}_2$ are missing from the estimate (3.3.8). At this stage, I couldn't establish the second *inf-sup* condition for $\mathcal{A}$:

$$\exists c > 0, \quad \forall x' \in \mathbb{I}, \quad \sup_{x \in \mathbb{J}^0} \frac{\mathcal{A}(x,x')}{\|x\|_{\mathbb{X}}} \geq c\|x'\|_{\mathbb{Y}}$$

and therefore, I couldn't apply Babuška's theorem in order to get the existence. This will be proved in the next section, by a Galerkin method.

3.3.3. Finite element approximation. From now on, we suppose that the two meshes are matching on the perforations Σ and we agree to denote by $\mathcal{E}_h^\Sigma$ the set of edges situated on Σ. We shall use the notation $h_{min,\Sigma} = \min_{e \in \mathcal{E}_h^\Sigma} h_e$. We also assume that:

$$(\mathbf{H}) \quad \overline{\rho} \geq \rho_{2h}(z) \quad \geq \quad \underline{\rho} > 0 \quad \text{a.e. on } \Sigma$$

where ρ_{2h} is a piecewise constant approximation of ρ_2 on $\mathcal{T}_h^2$.

We next write a conforming approximation of problem (3.3.5) based on the finite element spaces already used for the separate reservoir and wellbore models. Concerning the Lagrange multipliers on the interface, we introduce

$$K_h = \{\mu \in L^2(\Sigma);\ \mu_{/e} \in P_0,\ \forall e \in \mathcal{E}_h^\Sigma\}$$

and we put $\mathbb{L}_h = (K_h)^2 \subset \mathbb{L}$, $\mathbb{K}_h = (K_h)^3 \subset \mathbb{K}$. We consider the following discrete version of (3.3.5):

$$\begin{cases} x_h \in \mathbb{X}_h^*, \Lambda_h \in \mathbb{L}_h \\ \mathcal{A}_h(x_h,x') \ + \ \mathcal{I}(\Lambda_h,x') \ = \ \mathcal{F}_h(x'),\ \forall x' \in \mathbb{Y}_h \\ \mathcal{J}_h(\Lambda',x_h) \ = \ 0,\ \forall \Lambda' \in \mathbb{K}_h, \end{cases} \tag{3.3.10}$$

where $\mathcal{A}_h(\cdot,\cdot)$ and $\mathcal{F}_h(\cdot)$ are obtained after upwinding and where $\mathcal{J}_h(\cdot,\cdot)$ is obtained from $\mathcal{J}(\cdot,\cdot)$ by replacing ρ_2 by ρ_{2h}.

3.3.3.1. *Well-posedness of discrete problem.* Due to the finite dimensional framework, it is sufficient to show the uniqueness of the solution of (3.3.10). I have followed the analysis of the continuous coupled problem and I have established the discrete versions of Lemmas 3.3.1-3.3.4, uniformly with respect to the discretisation parameter h. The key ingredient in their proofs is the following auxiliary result, that we have established under a hypothesis on the mesh size.

LEMMA 3.3.6. *Assume that there exists* $\epsilon \in]0, \frac{1}{2}]$ *such that any* $\mathcal{T}_h^1$ *satisfies the property:*

$$h_1^{\epsilon+\frac{1}{2}} \leq c\ (h_{min,\Sigma})^{\epsilon}. \tag{3.3.11}$$

Then, for any $p \in M_h$ *and* $\theta \in K_h$, *there exists* $\mathbf{G} \in V_h$ *satisfying:*

$$\begin{cases} \mathbf{G}\cdot\mathbf{n} = \theta \ \text{on}\ \Sigma, \quad \mathbf{G}\cdot\mathbf{n} = 0 \ \text{on}\ \Upsilon_{\mathbf{G}} \setminus \Sigma \\ \text{div}\mathbf{G} = p \quad \text{in}\ \Omega_1. \end{cases} \tag{3.3.12}$$

Moreover, the next bound holds with c *independent of* h*:*

$$\|\mathbf{G}\|_{H(\text{div},\Omega_1)} + \|\mathbf{G}\cdot\mathbf{n}\|_{0,\Sigma} \leq c(\|p\|_{0,\Omega_1} + \|\theta\|_{0,\Sigma}). \tag{3.3.13}$$

Proof. The idea is to define $\mathbf{G}$ as the Raviart-Thomas interpolate of a function satisfying the above properties.

Let us first note that θ belongs to $H^{\frac{1}{2}-\epsilon}(\Sigma)$ only, for any $0 < \epsilon \leq 1/2$. We regularize θ and we define $\tilde{\theta} \in H_0^1(\Sigma)$ by $\tilde{\theta}_{/e} = \theta\,\chi_e$, where χ_e is the bubble-function associated with the edge $e \in \mathcal{E}_h^{\Sigma}$ satisfying $\chi_e \in P_2$ and $\int_e \chi_e \text{ds} = h_e$. It is useful to note that:

$$\|\chi_e\|_{0,e} = c_0\, h_e^{1/2}, \quad |\chi_e|_{1,e} = c_1\, h_e^{-1/2}.$$

Then $\tilde{\theta} \in H_0^1(e)$ and $\int_e \tilde{\theta}\text{ds} = \int_e \theta\text{ds}$.

We consider the auxiliary problem in the rectangle Ω_1:

$$\begin{cases} \Delta\phi = p & \text{in}\ \Omega_1 \\ \frac{\partial\phi}{\partial n} = \tilde{\theta} & \text{on}\ \Sigma \\ \frac{\partial\phi}{\partial n} = 0 & \text{on}\ \Upsilon_{\mathbf{G}} \setminus \Sigma \\ \phi = 0 & \text{on}\ \Upsilon_p \end{cases}, \tag{3.3.14}$$

whose unique solution belongs to $H^2(\Omega_1)$ (cf. **[96]**) and satisfies for any $0 < \epsilon \leq \frac{1}{2}$:

$$|\phi|_{\frac{3}{2}+\epsilon,\Omega_1} \leq c(\epsilon)\left(\|\Delta\phi\|_{-\frac{1}{2}+\epsilon,\Omega_1} + \|\partial_n\phi\|_{\epsilon,\Sigma}\right) \leq c(\epsilon)\left(\|p\|_{0,\Omega_1} + \left\|\tilde{\theta}\right\|_{\epsilon,\Sigma}\right).$$

Then $\mathbf{G} = E_h(\nabla\phi)$, with E_h the Raviart-Thomas interpolation operator, obviously satisfies the relations (3.3.12) cf. **[158]** and, since $\mathbf{G}\cdot\mathbf{n}$ and θ are piecewise constant, we also get $\|\mathbf{G}\cdot\mathbf{n}\|_{0,\Sigma} = \|\theta\|_{0,\Sigma}$.

We still have to bound $\|\mathbf{G}\|_{0,\Omega_1}$. By classical tools, we first get

$$\|\mathbf{G}\|_{0,\Omega_1} \leq c\left(|\phi|_{1,\Omega_1} + h_1^{\epsilon+\frac{1}{2}}\,|\phi|_{\frac{3}{2}+\epsilon,\Omega_1}\right).$$

Since $H^{\epsilon}(\Sigma)$ is the interpolate space of $L^2(\Sigma)$ and $H^1(\Sigma)$ (cf. **[121]**), we have:

$$\left\|\tilde{\theta}\right\|_{\epsilon,\Sigma} \leq c\left\|\tilde{\theta}\right\|_{0,\Sigma}^{1-\epsilon}\left\|\tilde{\theta}\right\|_{1,\Sigma}^{\epsilon} \leq c\left(\left\|\tilde{\theta}\right\|_{0,\Sigma} + \left\|\tilde{\theta}\right\|_{0,\Sigma}^{1-\epsilon}\left|\tilde{\theta}\right|_{1,\Sigma}^{\epsilon}\right).$$

By using that

$$\left\|\tilde{\theta}\right\|_{0,\Sigma} \leq c\,\|\theta\|_{0,\Sigma}, \qquad \left|\tilde{\theta}\right|_{1,\Sigma} \leq \frac{c}{h_{min,\Sigma}}\,\|\theta\|_{0,\Sigma},$$

we finally obtain the desired estimate (3.3.13), under the condition (3.3.11). ■

Let us recall here that both the discrete reservoir and wellbore models have unique solutions if

$$\Delta t \leq \min(C_1 h^2_{min,\Omega_1}, C_2 h_{min,\Omega_2}) \tag{3.3.15}$$

with $h_{min,\Omega_1} = \min_{T\in\mathcal{T}_h^1} h_T$, $h_{min,\Omega_2} = \min_{T\in\mathcal{T}_h^2} h_T$ and with C_1, C_2 independent of the discretisation. Then we immediately get, thanks to the discrete versions of Lemmas 3.3.2, 3.3.3 and 3.3.4 :

THEOREM 3.3.7. *Assume (3.3.11) and (3.3.15). Then problem (3.3.10) has a unique solution.*

3.3.3.2. *Existence of a solution for the continuous problem.* Finally, we can now prove the existence of a solution in the continuous case.

THEOREM 3.3.8. *Assume (3.3.11) and (3.3.15). The continuous coupled problem (3.3.5) with* $\alpha = 0$ *has at least one solution.*

Proof. We apply a Galerkin method. We first consider a sequence of approximated problems of (3.3.5), written on the finite dimensional spaces previously introduced:

$$\left\{\begin{array}{lcl} \widetilde{x}_h \in \mathbb{X}_h^*, \widetilde{\Lambda}_h \in \mathbb{L}_h & & \\ \mathcal{A}(\widetilde{x}_h, x') + \mathcal{I}(\widetilde{\Lambda}_h, x') = \mathcal{F}(x'),\ \forall x' \in \mathbb{Y}_h & & \\ \mathcal{J}_h(\Lambda', \widetilde{x}_h) = 0,\ \forall \Lambda' \in \mathbb{K}_h & & \end{array}\right. \tag{3.3.16}$$

where in the definition of $\mathcal{J}_h(\cdot,\cdot)$, $\rho_{\tilde{2}h}$ now stands for the piecewise constant $L^2(\Sigma)$-orthogonal projection of ρ_2. We already know that each discrete problem (3.3.16) has a unique solution $(\widetilde{x}_h, \widetilde{\Lambda}_h)$, and the sequence $(\widetilde{x}_h, \widetilde{\Lambda}_h)_h$ is bounded in the $\mathbb{X}\times\mathbb{L}$-norm. Therefore, we can extract a subsequence weakly convergent in $\mathbb{X}\times\mathbb{L}$ towards $(\widetilde{x}, \widetilde{\Lambda})$. A classical passage to the limit in (3.3.16) yields that the weak limit $(\widetilde{x}, \widetilde{\Lambda})$ is in fact a solution of problem (3.3.5). ■

We can also prove that the problem with convection ($\alpha = 1$) has a unique solution for $\triangle t$ small enough, by using the regularity of the solution of (3.3.5) together with Fredholm's alternative.

3.3.4. Numerical results. I present here a comparison between the coupled code and the separate reservoir and wellbore simulators, in the realistic case of a seven-layer reservoir. Each layer is characterized by high heterogeneities cf. Table 1. The producing layers (the 2nd, 4th and 6th from the top) have high permeabilities and can be separated by quasi-walls with low porosity and permeability. The reservoir is 50 m large and 20 m high, and the wellbore is only 0.15 m large but 70 m high. We simulate the production of a light oil during 28 days for the three models. The reservoir is fed by imposing a constant pressure 400 bar on its external boundary, and a difference of pressure $\Delta p = 10$ bar between the perforations and the external boundary. For the coupled code, we impose a constant flowrate $Q = 6500\,\mathrm{m}^3/\mathrm{day}$ at the pipe's surface. For the sole wellbore model, we impose as boundary conditions on the perforations the values given by the reservoir code.

Concerning the comparison with the reservoir simulator, one can see in Fig. 3.3.1 that the flowrate imposed at the top of the well in the coupled model yields a difference of pressure $\Delta p \backsimeq 10$ bar which coincides with that imposed as boundary condition in the reservoir. The two simulators also give very similar results for the temperature. As regards the comparison with the wellbore, Fig. 3.3.2 shows very similar results for G_z, from which one computes the production flowrate in the well. Thus, we obtain a flowrate Q for the wellbore close to that imposed as boundary condition in the coupled problem.

$k_1 = 7000\,\mathrm{mD}$	$k_2 = 350\,\mathrm{mD}$	$\phi = 0.20$	$s_w = 0.15$
$k_1 = 7000\,\mathrm{mD}$	$k_2 = 350\,\mathrm{mD}$	$\phi = 0.28$	$s_w = 0.15$
$k_1 = 10\,\mathrm{mD}$	$k_2 = 1\,\mathrm{mD}$	$\phi = 0.08$	$s_w = 0.90$
$k_1 = 1000\,\mathrm{mD}$	$k_2 = 15\,\mathrm{mD}$	$\phi = 0.24$	$s_w = 0.42$
$k_1 = 1000\,\mathrm{mD}$	$k_2 = 15\,\mathrm{mD}$	$\phi = 0.26$	$s_w = 0.30$
$k_1 = 1000\,\mathrm{mD}$	$k_2 = 15\,\mathrm{mD}$	$\phi = 0.22$	$s_w = 0.38$
$k_1 = 1000\,\mathrm{mD}$	$k_2 = 15\,\mathrm{mD}$	$\phi = 0.24$	$s_w = 0.40$

TABLE 1. Characteristics of a realistic reservoir

We show in Fig. 3.3.3 the tevolution of the temperature computed by the coupled code during a one month production. Besides the initial and final time-steps, we have represented the maps at $t = 2\,\mathrm{days}$ and $t = 7\,\mathrm{days}$ since afterwards the flow almost reaches the steady state. The figures focus on the neighbourhood of the perforations since due to the large aspect ratio between the reservoir and the wellbore, we only visualise 8m in the radial direction. The transmission conditions at the interface are satisfied, and the results correspond to the physical behaviour expected by petroleum engineers.

We have also looked at the specific flux $\mathbf{G}$. As one can see in Fig. 3.3.4 (a), the velocity in the wellbore is much more important than the velocity in the reservoir, since the flux on a given cell in the wellbore is obtained by summing up the contributions of all the lower perforations. In order to better visualise the flow near the perforations, we have applied in Fig. 3.3.4 (b) different scalings in the two domains (of ratio equal to 10).

As regards the wellbore results, we recover the well-known fact that the pressure is primarily influenced by the gravity; it goes the same way for the temperature above perforations.

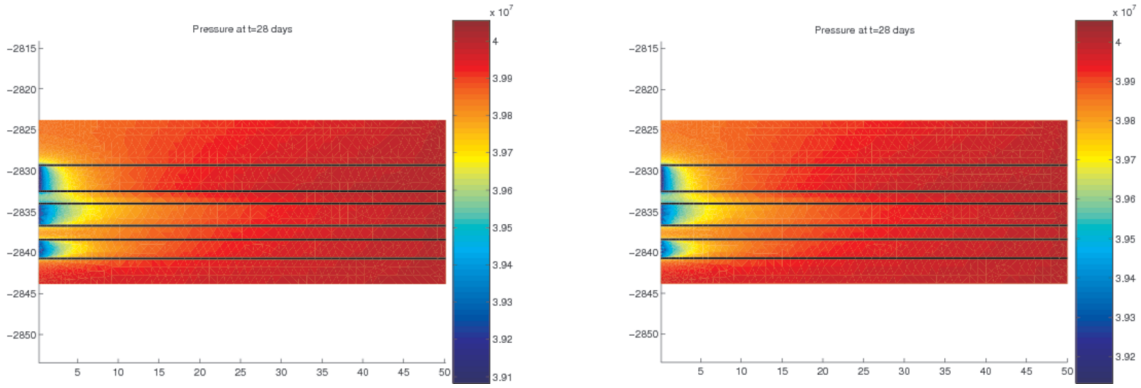


FIGURE 3.3.1. Pressure given by reservoir (left) and coupled (right) codes

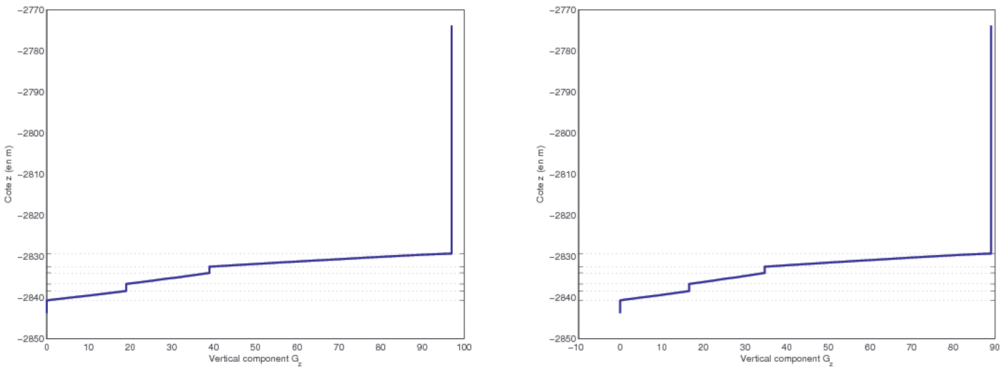


FIGURE 3.3.2. G_z given by wellbore (left) and coupled (right) codes

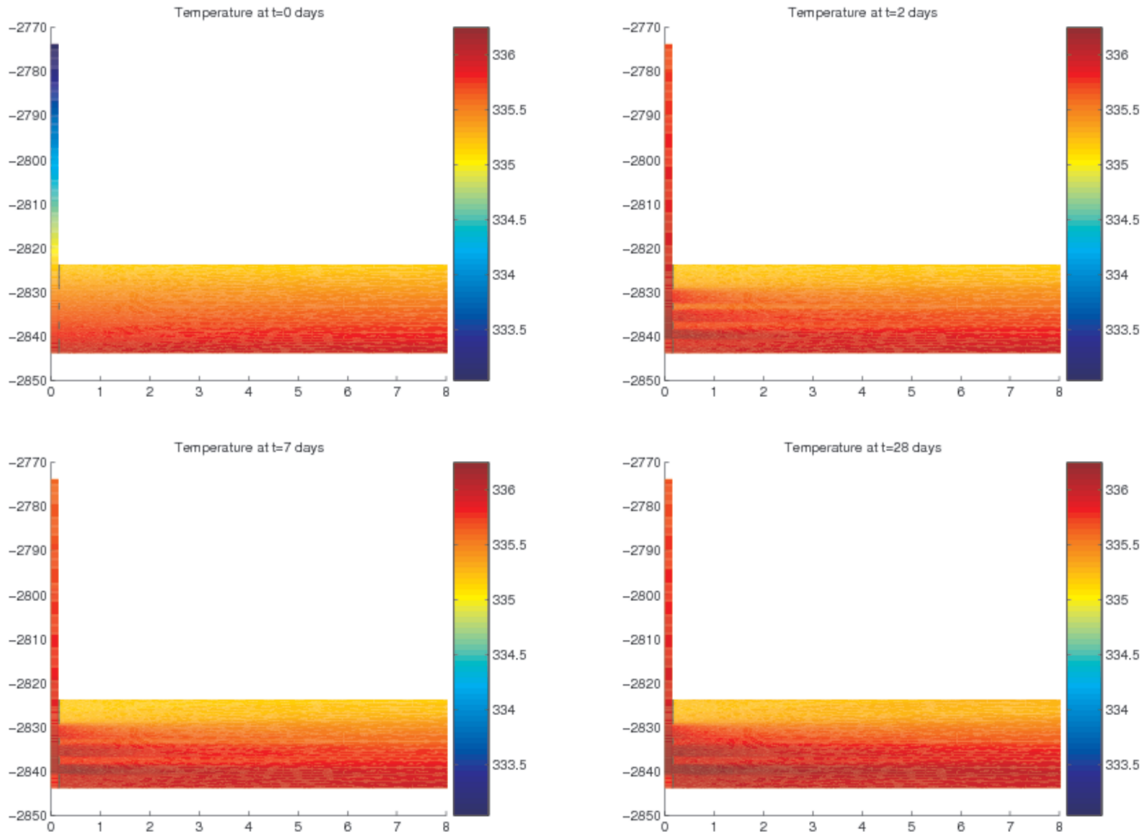
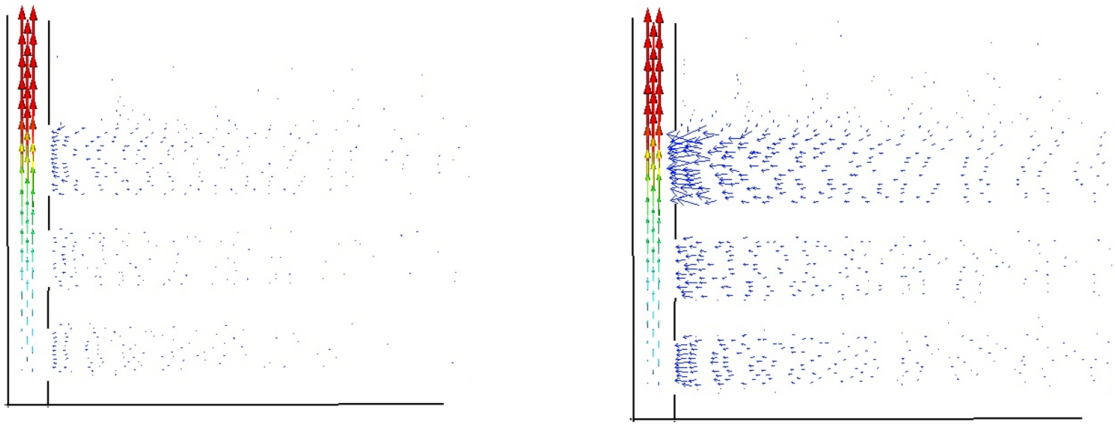


FIGURE 3.3.3. Evolution of the temperature in the coupled model



a) Same scale in the 2 domains b) Different scales in each domain

FIGURE 3.3.4. Specific flux at $t = 28$ days

3.4. Multi-phase multi-component reservoir model

This work was published in [A10]; more details can be found in the PhD thesis of Layal Lizaik [**122**]. We have mainly addressed the thermodynamic modeling, which is of paramount importance for realistic simulations of flows in porous media. We have introduced an energy equation, including the Joule-Thomson effect and the energy dissipation as for single phase flows, and the corresponding thermodynamics. They were next integrated in the isothermal simulator GPRS and first numerical tests were carried out.

3.4.1. Physical modeling.

3.4.1.1. *Conservation laws.* Petroleum fluids contain several chemical components (such as hydrogen, hydrogenated and nitrogenous components, hydrocarbon molecules). We consider a system composed of n_c components, three phases : water(w), oil (o), gas (g), and the rock (s). We assume that there is no mass transfer between water and hydrocarbon phases. The number of hydrocarbon components is equal to $n_h = n_c - 1$.

We denote by $\mathbf{u}_p$ the Darcy velocity of the phase p, by p_p its pressure, by S_p its saturation, by k_{rp} its relative permeability, by ρ_p its density and by μ_p its viscosity. The diagonal tensor $\underline{\mathbf{K}}$ represents the permeability of the medium and ϕ its porosity.

The governing equations (cf. [**14**], [**55**]) are the mass conservation law for each hydrocarbon component c in oil and gas phases, and for the water:

$$\begin{aligned} \sum_{p=o,g} \left(\frac{\partial}{\partial t}(\phi S_p \rho_p y_{c,p}) + \mathrm{div}(\rho_p y_{c,p} \mathbf{u}_p) \right) &= 0, \\ \frac{\partial}{\partial t}(\phi S_w \rho_w) + \mathrm{div}(\rho_w \mathbf{u}_w) &= 0, \end{aligned}$$

coupled with the extended Darcy law for each phase:

$$\mathbf{u}_p = -k_{rp} \mu_p^{-1} \underline{\mathbf{K}} (\nabla p_p - \gamma_p \nabla Z).$$

$y_{c,p}$ is the molar fraction of component c in the phase p and γ_p is equal to $\rho_p g$. We consider the following energy balance:

$$\frac{\partial}{\partial t}\left[\sum_{p=o,g,w}(\phi S_p\rho_p\mathcal{H}_p - p_p) + (1-\phi)\,\rho_s\mathcal{H}_s\right] + \sum_{p=o,w,g}\mathrm{div}(\phi S_p\rho_p\mathcal{H}_p\mathbf{u}_p) \\ -\mathrm{div}\,(\lambda\nabla T) + \sum_{p=o,g,w}\mathbf{u}_p\cdot\nabla p_p \;=\; 0,$$

where $\mathcal{H}_p$ represents the enthalpy of phase p, λ denotes the equivalent thermal conductivity and T is the temperature. By substituting the velocity in the mass and the energy balances, we obtain

$$\begin{cases} \sum\limits_p \left(\frac{\partial}{\partial t}(\phi S_p\rho_p y_{c,p}) - \mathrm{div}(\rho_p k_{rp}\mu_p^{-1}\underline{\mathbf{K}}(\nabla p_p - \gamma_p\nabla Z)y_{c,p})\right) = 0, \quad c = 1,\cdots,n_c \\ \frac{\partial}{\partial t}\left[\sum\limits_p(\phi S_p\rho_p\mathcal{H}_p - p_p) + (1-\phi)\,\rho_s\mathcal{H}_s\right] - \sum\limits_p \mathrm{div}(\rho_p k_{rp}\mu_p^{-1}\underline{\mathbf{K}}(\nabla p_p - \gamma_p\nabla Z)\mathcal{H}_p) \\ \qquad -\mathrm{div}\,(\lambda\nabla T) - \sum\limits_p(\phi^{-1}S_p^{-1}\mu_p^{-1}k_{rp}\underline{\mathbf{K}}(\nabla p_p - \gamma_p\nabla Z)\cdot\nabla p_p) = 0. \end{cases} \tag{3.4.1}$$

Besides conservation laws, we also need phase equilibrium relations for each component in oil and gas phases, since water and hydrocarbon components are totally separated. This equilibrium is illustrated by the equality of fugacities $f_{c,p}$ of each component in the two phases,

$$f_{c,o} = f_{c,g}, \qquad c = 1,\cdots,n_h. \tag{3.4.2}$$

Finally, in order to close the system, some linear constraints must be satisfied. We respectively have the saturation constraint, the component mole fraction constraints and the capillary pressure constraints (with $p_{c,ow}$ and $p_{c,go}$ denoting the oil-water, respectively the gas-oil capillary pressures):

$$\begin{gathered} \sum_{p=1}^{n_p} S_p = 1, \\ \sum_{c=1}^{n_h} y_{c,p} = 1, \qquad p = o,g, \\ p_{c,ow} = p_o - p_w, \qquad p_{c,go} = p_g - p_o. \end{gathered} \tag{3.4.3}$$

In conclusion, the full system consists of $2n_h + 7$ equations, (3.4.1), (3.4.2) and (3.4.3), to which initial and boundary conditions are added.

The boundary conditions refer either to the well or to an external boundary of the reservoir. On the lateral boundary, one can impose a null mass flux or a constant pressure. Note that in the isothermal GPRS code, only a constant pressure can be set. Concerning the well, two types of controls can be imposed: a well bottom pressure p_f or a constant phase volumetric phase (cf. **[55]**). In the latter case, p_f is an extra unknown and an extra equation in the well (the component mass balance within the wellbore) is then added. As regards the energetic aspect, a null heat flux or a temperature can be set.

3.4.1.2. *Thermodynamic properties and flash calculations.* For the calculation of the thermodynamic coefficients of the fluid (such as the density, the enthalpy, the viscosity and the fugacity) and the representation of fluid phase equilibriums, we use the cubic equation of state of Peng-Robinson (cf. **[145]**), because of its capability to represent both the liquid and the gas phases. For a mixture of n_c components, the needed parameters are described in **[135]**.

For given pressure, temperature and overall composition, the flash is the mechanism by means of which one computes the molar composition of hydrocarbon phases at equilibrium. It is based on the resolution of the Rachford-Rice equation (cf. **[117]**) and it leads to different variables in

different gridblocks. Thus, one has to check the appearence or disappearence of hydrocarbon phases in all gridblocks, after every Newton's iteration.

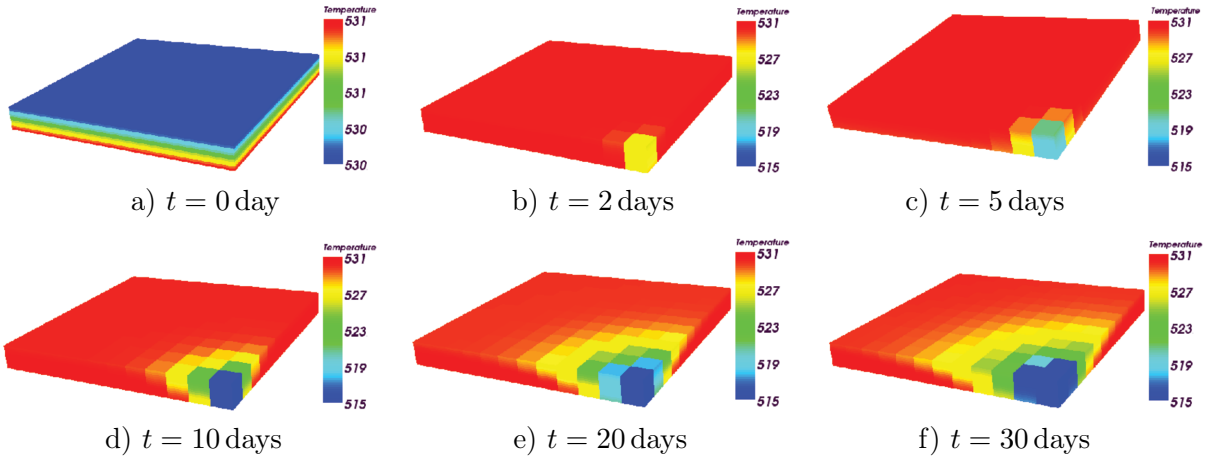
3.4.2. Numerical resolution. According to Gibbs phase rule (cf. [**14**]), the number of degrees of freedom is

$$(n_c + 2 - n_p) + (n_p - 1) = n_c + 1,$$

where n_p is the number of phases. One then has to solve $(n_c + 1)$ equations, called primary equations. By using the linear equations (3.4.3) to remove two pressures, one saturation and two molar fractions, only $(2n_h+2)$ nonlinear equations are left. Multiple choices of primary equations and variables can be made, leading to different models (see [**57**], [**170**]). Here, we consider the Coats model (cf. [**14**], [**60**]), where the primary equations are the mass and the energy balance equations. The primary variables are one pressure, the temperature, (n_p-1) saturations and (n_c-n_p) molar fractions. More precisely, these variables are $p_g,\ T,\ S_g,\ S_o,\ y_{c,g}$ with $c = 3,\ \cdots,\ n_h$ when both the gas and oil phases are present, and $p_p,\ T,\ S_p,\ y_{c,p}$ with $c = 1,\ \cdots,\ n_h - 1$ when one of the phase (gas for $p = o$, oil for $p = g$) disappears. The secondary equations are the phase equilibrium relations.

The time discretization is based on Euler's implicit scheme, and the time step is calculated by a given formula (cf. [**14**], [**55**]). The mesh implemented in GPRS is a cartesian one, where the 3D domain is a parallelepiped. The space discretization is achieved by classical cell-centered finite volumes cf. [**81**]. The nonlinear system is solved by Newton's method. As regards the initial state of the reservoir, an equilibrium between the coexisting fluids is imposed. The temperature is initialized thanks to the geothermal gradient, and phase saturations are assigned according to the positioning of the fluid contacts. Initial overall compositions are calculated and a flash calculation is performed at constant pressure, in order to assign the molar fractions of the components. Finally, we compute the densities of gas and oil phases and we reinitialize the saturations and distribute the pressure.

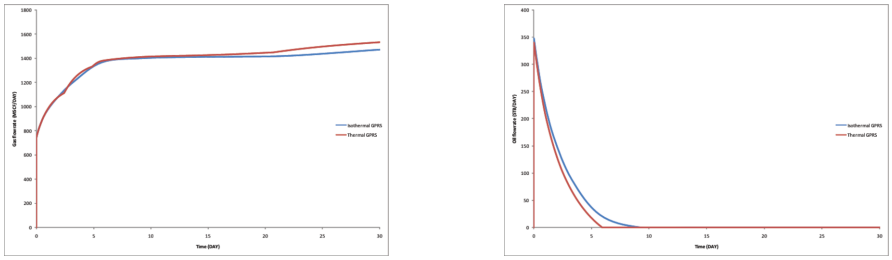
3.4.2.1. *Numerical examples.* We consider here a three-component (methan, butan and heptan), two-phase (gas and oil) compositional simulation in a reservoir of dimensions 6000 ft × 6000 ft × 60 ft. A producing well is located at the gridblock $(0, 0, 0)$ and is under the bottom hole pressure control of 300 psi. The simulation was run during 30 days. We have obtained physically acceptable results: the values of the gas saturation confirmed the tendency of the gas to go up to the top of the reservoir, and the pressure increase first concerns the well's gridblock, then it fastly extends to the other gridbloks. One can see in Fig. 3.4.1 the evolution of the temperature which is, as expected, of only a few degrees.



a) $t = 0$ day b) $t = 2$ days c) $t = 5$ days

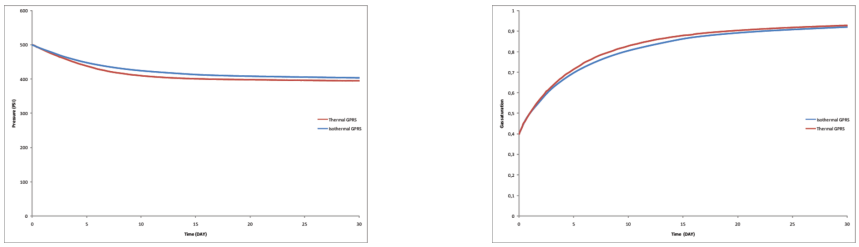
d) $t = 10$ days e) $t = 20$ days f) $t = 30$ days

FIGURE 3.4.1. Evolution of the temperature during 30 days of production.



a) Gas flowrates b) Oil flowrates

FIGURE 3.4.2. Comparison of gas and oil flowrates at the well head



a) Pressure b) Gas saturation

FIGURE 3.4.3. Comparison of pressures and gas saturation in the well block

We have also compared our results with those given by the isothermal GPRS code. In Fig. 3.4.2, one can see the gas and oil flowrates at the well head given by the two simulators while in Fig. 3.4.3, we show the pressures and the saturations at the well block. The results are similar, the small differences can be explained by the varying temperature in our code. One can also notice that the production of gas is dominant, which justifies the cooling around the well. This phenomenon is caused by the expansion of the gas, known as the Joule-Thomson effect. More tests can be found in [**122**].

CHAPTER 4

APPLICATIONS TO NON-NEWTONIAN FLUIDS

CHAPTER 4

Applications to non-Newtonian fluids

This chapter deals with the numerical approximation of certain viscoelastic non-Newtonian fluids, which is one of the research topics that I developed within the INRIA team Concha.

In Section 4.1, a realistic model for polymer flows, namely the Giesekus model, is described and a nonconforming finite element discretization is proposed. The developed code is validated by numerical tests, highlighting its robustness with respect to physical parameters such that the Weissenberg number and the pertinence of the chosen model.

In the last Section, a more general matrix-valued transport equation is considered and the positivity of its discrete solution is studied; applications to the nonlinear Giesekus law and to the quasi-linear Oldroyd-B law are presented, allowing to explain the better behaviour of the first model compared to the latter.

4.1. Numerical simulation of polymer flows

I introduce next the Giesekus model and then I describe its numerical approximation, based on nonconforming finite elements combined with an upwind scheme à la Lesaint and Raviart. This discretization yields well-posedness and optimal error estimates for the underlying Stokes problem; in particular, the influence of regularization terms is discussed. Finally, numerical tests including comparisons on typical geometries are presented.

I present here only the case of quadrilateral mesh, published in [A12] and [C11]; the triangular case is described in [C12], see also [C13]. More details can be found in the PhD thesis of Julie Joie [**107**].

4.1.1. Giesekus model. Polymeric liquids are, from a rheological point of view, viscoelastic non-Newtonian fluids. Their non-Newtonian behavior can be observed in a variety of physical phenomena (rod climbing effect, die swell, extrusion instabilities) which are unseen with Newtonian liquids and which cannot be predicted by the Navier-Stokes equations. The rheological behavior of polymers is so complex that many different constitutive equations have been proposed in the literature in order to describe these phenomena, see for instance [**142**] for a review.

We choose here to study the nonlinear differential model of Giesekus introduced in [**90**], whose constitutive law involves a quadratic term in the stress tensor and is given by:

$$\lambda\left(\overset{\triangledown}{\underline{\tau}}+\frac{\bar{\alpha}}{\eta}\underline{\tau}^2\right)+\underline{\tau}=2\eta\underline{\varepsilon}(\mathbf{u}) \tag{4.1.1}$$

with $\underline{\tau}$ the viscous stress tensor, $\underline{\varepsilon}(\mathbf{u})=\frac{1}{2}(\nabla\mathbf{u}+\nabla\mathbf{u}^T)$ the strain rate tensor, λ the relaxation time of the fluid, η its viscosity and $\bar{\alpha}\in]0,1[$ a parameter. We take next $\bar{\alpha}=0.5$ which seems to be an appropriate value from a physical point of view.

Here above, $\overset{\triangledown}{\underline{\tau}}$ denotes the upper convective derivative and is defined by:

$$\overset{\triangledown}{\underline{\tau}}=\partial_t\underline{\tau}+(\mathbf{u}\cdot\nabla)\underline{\tau}-(\nabla\mathbf{u}\underline{\tau}+\underline{\tau}\nabla\mathbf{u}^T).$$

This model presents two main advantages. First, it yields a realistic behaviour for shear flows, elongational flows and mixed flows. Second, only two material parameters (η and λ), which are moreover easily measurable, are needed to describe it.

In what follows, we focus on the steady case and in order to fix the ideas, we consider a Dirichlet boundary condition for the velocity, $\mathbf{u} = \mathbf{g}$ on $\partial\Omega$, and $\underline{\tau} = \underline{\tau}^D$ on the inflow boundary $\partial\Omega^-$. The complete Giesekus model is obtained by adding the mass and the momentum conservation laws, where the density ρ is supposed to be constant:

$$\begin{aligned} \mathrm{div}\mathbf{u} &= 0, \\ \rho\left(\mathbf{u}\cdot\nabla\right)\mathbf{u} - \mathrm{div}\underline{\tau} + \nabla p &= \mathbf{f}. \end{aligned}$$

We take $\mathbf{f} \in \mathbf{L}^2(\Omega)$, $\mathbf{g} \in \mathbf{H}^{1/2}(\partial\Omega)$ and $\underline{\tau}^D \in \underline{L}^2_{\mathrm{sym}}(\partial\Omega^-)$.

4.1.2. Discrete nonlinear formulation. Let $(\mathcal{T}_h)_{h>0}$ be a family of regular meshes of the polygonal domain $\Omega \subset \mathbb{R}^2$ consisting of quadrilaterals, $\Omega = \cup_{K\in\mathcal{T}_h} K$. We approach the velocity by nonconforming finite elements of Rannacher-Turek **[154]**, whose degrees of freedom are the mean values across the edges, and the pressure and the stress tensor by piecewise constant functions.

Let $\hat{K} = [-1,1]\times[-1,1]$ and $\Psi_K : \hat{K} \to K$ the bilinear one-to-one transformation; let also $\hat{Q}_1^{\mathrm{rot}} = \mathrm{vect}\{1, \hat{x}, \hat{y}, \hat{x}^2 - \hat{y}^2\}$ and $Q_K = \{v;\ v\circ\Psi_K \in \hat{Q}_1^{\mathrm{rot}}\}$.

Then we introduce the discrete spaces:

$$\begin{aligned} \mathbf{V}_h &= \left\{\mathbf{v}_h \in \mathbf{L}^2(\Omega);\ \mathbf{v}_h \in Q_K\ \forall K\in\mathcal{T}_h,\ \frac{1}{|e|}\int_e [\mathbf{v}_h]\mathrm{ds} = \mathbf{0}\ \ \forall e \in \varepsilon_h^{int}\right\}, \\ \mathbf{V}_h^g &= \{\mathbf{v}_h \in \mathbf{V}_h;\ \int_e \mathbf{v}_h\mathrm{ds} = \int_e \mathbf{g}\mathrm{ds}\ \ \forall e\in\varepsilon_h^{\partial}\}, \\ Q_h &= \{q_h \in L^2_0(\Omega);\ q_h \in P_0\ \forall K\in\mathcal{T}_h\}, \\ \underline{X}_h &= \left\{\underline{\theta}_h \in \underline{L}^2_{\mathrm{sym}}(\Omega);\ \underline{\theta}_h \in \underline{P}_0\ \forall K\in\mathcal{T}_h\right\} \end{aligned}$$

and we consider the following discrete formulation:

$$\left\{\begin{array}{lll} (\mathbf{u}_h, p_h, \underline{\tau}_h) \in \mathbf{V}_h^g \times Q_h \times \underline{X}_h & & \\ & & \\ a^\gamma(\mathbf{u}_h,\mathbf{v}_h) + b(p_h,\mathbf{v}_h) + c_0(\mathbf{v}_h,\underline{\tau}_h) & = f(\mathbf{v}_h) & \forall \mathbf{v}_h \in \mathbf{V}_h^0 \\ b(q_h,\mathbf{u}_h) & = 0 & \forall q_h \in Q_h \\ c(\mathbf{u}_h,\underline{\tau}_h;\underline{\theta}_h) + d(\underline{\tau}_h,\underline{\theta}_h) & = l(\underline{\theta}_h) & \forall \underline{\theta}_h \in \underline{X}_h. \end{array}\right. \tag{4.1.2}$$

The previous forms are defined by:

$$\begin{aligned} a^\gamma(\cdot,\cdot) &= a_0(\cdot,\cdot) + \gamma\, J(\cdot,\cdot) + R(\cdot,\cdot), \\ b(q_h,\mathbf{v}_h) &= -\sum_{K\in\mathcal{T}_h}\int_K q_h \mathrm{div}\mathbf{v}_h \mathrm{dx}, \\ c(\cdot,\cdot;\cdot) &= -2\eta\, c_0(\cdot,\cdot) + c_1(\cdot,\cdot;\cdot) - c_2(\cdot,\cdot;\cdot), \\ d(\cdot,\cdot) &= d_0(\cdot,\cdot) + d_1(\cdot,\cdot), \\ f(\mathbf{v}_h) &= \sum_{K\in\mathcal{T}_h}\int_K \mathbf{f}\cdot\mathbf{v}_h \mathrm{dx}, \\ l(\underline{\theta}_h) &= -\sum_{e\in\varepsilon_h^{\partial}\cap\partial\Omega^-}\int_e \{\mathbf{u}_h\cdot\mathbf{n}\}^-\underline{\tau}^D : \underline{\theta}_h \mathrm{ds}, \end{aligned}$$

where

$$a_0(\mathbf{u}_h,\mathbf{v}_h)=\sum_{K\in\mathcal{T}_h}\int_K\frac{\rho}{2}\left((\mathbf{u}_h\cdot\underline{\nabla})\mathbf{u}_h\cdot\mathbf{v}_h-(\mathbf{u}_h\cdot\underline{\nabla})\mathbf{v}_h\cdot\mathbf{u}_h\right)\mathrm{dx},$$

$$c_0(\underline{\tau}_h,\mathbf{v}_h)=\sum_{K\in\mathcal{T}_h}\int_K\underline{\tau}_h:\underline{\varepsilon}(\mathbf{v}_h)\mathrm{dx},$$

$$c_2(\mathbf{u}_h,\underline{\tau}_h;\underline{\theta}_h)=\lambda\sum_{K\in\mathcal{T}_h}\int_K(\underline{\tau}_h\underline{\nabla}\mathbf{u}_h^T+\underline{\nabla}\mathbf{u}_h\underline{\tau}_h):\underline{\theta}_h\mathrm{dx},$$

$$d_0(\underline{\theta}_h,\underline{\tau}_h)=\sum_{K\in\mathcal{T}_h}\int_K\underline{\theta}_h:\underline{\tau}_h\mathrm{dx},$$

$$d_1(\underline{\tau}_h,\underline{\theta}_h)=\frac{\lambda}{2\eta}\sum_{K\in\mathcal{T}_h}\int_K(\underline{\tau}_h\underline{\tau}_h):\underline{\theta}_h\mathrm{dx}.$$

The analysis of the underlying Stokes problem has highlighted the necessity of adding to $a_0(\cdot,\cdot)$ two stabilization terms, $J(\cdot,\cdot)$ in order to recover a Korn inequality on nonconforming spaces and $R(\cdot,\cdot)$ to attain optimal convergence:

$$J(\mathbf{u}_h,\mathbf{v}_h)=\eta\sum_{e\in\varepsilon_h^{\mathrm{int}}}\frac{1}{|e|}\int_e[\pi_1^e(\mathbf{u}_h\cdot\mathbf{n}_e)][\pi_1^e(\mathbf{v}_h\cdot\mathbf{n}_e)]\mathrm{ds},$$

$$R(\mathbf{u}_h,\mathbf{v}_h)=2\eta\sum_{K\in\mathcal{T}_h}\int_K(\underline{\varepsilon}(\mathbf{u}_h)-\pi_0^K\underline{\varepsilon}(\mathbf{u}_h)):\underline{\varepsilon}(\mathbf{v}_h)\mathrm{dx}.$$

This aspect will be developed in the next paragraph. The parameter γ is independent of h.

The convective term $\displaystyle\int_\Omega\mathbf{u}\cdot\nabla\underline{\tau}:\underline{\theta}\mathrm{dx}$ of the constitutive law is treated by an upwind scheme, which extends the well-known Lesaint-Raviart scheme **[119]** for constant vectors $\mathbf{u}$ to the present nonconforming approximation of the velocity. Finally, we take

$$c_1(\mathbf{u}_h,\underline{\tau}_h;\underline{\theta}_h)=\lambda\sum_{e\in\varepsilon_h}\int_e F_e(\underline{\tau}_h,\mathbf{u}_h,n_e):[\underline{\theta}_h]\,\mathrm{ds},$$

where $F_e(\underline{\tau}_h,\mathbf{u}_h,\mathbf{n}_e)=\{\mathbf{u}_h\cdot\mathbf{n}_e\}^+\underline{\tau}_h^{\mathrm{in}}+\{\mathbf{u}_h\cdot\mathbf{n}_e\}^-\underline{\tau}_h^{\mathrm{ex}}$ is the numerical flux. The nonlinear problem (4.1.2) is solved by Newton's method.

REMARK 4.1.1. Another well-known possibility for the approximation of viscoelastic flows is to introduce the strain rate tensor $\underline{e}=\underline{\varepsilon}(\mathbf{u})$ as a fourth unknown of the problemn and to split the stress tensor $\underline{\tau}$ (see for instance **[97]** for the so-called DEVSS method). It is then interesting to note that the elimination of $\underline{e}$ at the discrete level yields a three-fields formulation with an additional term similar to our regularization $R(\cdot,\cdot)$.

4.1.3. Influence of the stabilization terms. Let here $\lambda=0$ and $\rho=0$, such that (4.1.2) is now a three-fields formulation of the Stokes problem. We can then recover locally the stress tensor by the relation $\underline{\tau_h}=2\eta\pi_0^K\underline{\varepsilon}(\mathbf{u}_h)$ and obtain the following equivalent two-fields formulation:

$$\left\{\begin{array}{lll}(\mathbf{u}_h,p_h)\in\mathbf{V}_h^g\times Q_h & & \\ \tilde{a}^\gamma(\mathbf{u}_h,\mathbf{v}_h)+b(p_h,\mathbf{v}_h) & =l(\mathbf{v}_h) & \forall\mathbf{v}_h\in\mathbf{V}_h^0\\ b(q_h,\mathbf{u}_h) & =0 & \forall q_h\in Q_h\end{array}\right. \tag{4.1.3}$$

where now $\tilde{a}^\gamma(\cdot,\cdot) = e(\cdot,\cdot) + \gamma J(\cdot,\cdot) + R(\cdot,\cdot)$ and

$$e(\mathbf{u}_h, \mathbf{v}_h) = 2\eta \sum_{K\in\mathcal{T}_h} \int_K \pi_0^K \underline{\varepsilon}(\mathbf{u}_h) : \pi_0^K \underline{\varepsilon}(\mathbf{v}_h) \mathrm{dx}.$$

I have proved in [C14] that the stabilization $J(\cdot,\cdot)$ yields the following inequality:

$$\sum_{K\in\mathcal{T}_h} \|\mathbf{v}\|_{1,K}^2 \leq c \left(\sum_{K\in\mathcal{T}_h} \|\underline{\varepsilon}(\mathbf{v})\|_{0,K}^2 + \frac{1}{\eta} J(\mathbf{v},\mathbf{v}) \right), \ \forall \mathbf{v} \in \mathbf{V}_h^0$$

with a constant c independent of h, η and γ. The main tool is a result of Brenner [**41**] for piecewise H^1-functions, which was later improved in [**129**]. More precisely, the authors showed in [**129**] that it is sufficient to consider $[\pi_1^e(\mathbf{u}_h \cdot \mathbf{n}_e)]$ instead of $[\pi_1^e \mathbf{u}_h]$ in the definition of $J(\cdot,\cdot)$, as initially proposed in [**41**].

For the error analysis, it is useful to introduce the following semi-norm on $\mathbf{H}^1(\Omega)+\mathbf{V}_h$, which is a norm on $\mathbf{V}_h^0$:

$$[[\mathbf{v}]]^2 = 2\eta \sum_{K\in\mathcal{T}_h} \|\underline{\varepsilon}(\mathbf{v})\|_{0,K}^2 + \gamma J(\mathbf{v},\mathbf{v}).$$

We have then checked the discrete hypotheses of the Babuška-Brezzi theorem with respect to the norms $[[\cdot]]$ and $\|\cdot\|_{0,\Omega}$ on $\mathbf{V}_h^0$ and Q_h, and we have deduced the well-posedness of problem (4.1.3) as well as the following *a priori* error bounds, cf. [A12] or [**107**].

THEOREM 4.1.2. *Let $(\mathbf{u},p) \in \mathbf{H}^2(\Omega)\times H^1(\Omega)$ be the solution of the continuous Stokes problem. Then the solution $(\mathbf{u}_h, p_h)$ of (4.1.3) satisfies:*

$$[[\mathbf{u}-\mathbf{u}_h]] + \frac{1}{\sqrt{\eta}} \|p-p_h\|_{0,\Omega} \leq ch(\sqrt{\eta}|\mathbf{u}|_{2,\Omega} + \frac{1}{\sqrt{\eta}}|p|_{1,\Omega}).$$

If one omits the term $R(\cdot,\cdot)$, which may seem natural at a first glance, then the corresponding two-fields formulation (4.1.3) has a unique solution but is not consistent. Indeed, the norm $[[\cdot]]$ is now replaced by:

$$[[[\mathbf{v}]]] = \left(2\eta \sum_{K\in\mathcal{T}_h} \|\pi_0^K \underline{\varepsilon}(\mathbf{v})\|_{0,K}^2 + \gamma J(\mathbf{v},\mathbf{v}) \right)^{1/2}.$$

In order to bound the consistency error with respect to $[[[\cdot]]]$, one needs the following uniform estimates:

$$|\mathbf{w}_h|_{1,h} \leq c_1 [[[\mathbf{w}_h]]], \quad \|\underline{\varepsilon}(\mathbf{w}_h)\|_{0,\Omega} \leq c_2 [[[\mathbf{w}_h]]],$$

which do not hold on $\mathbf{V}_h^0$. To illustrate numerically this phenomenon, we show in Fig.4.1.1 the results obtained for u_2 and τ_{12} for the driven cavity test, with and without regularization. We have considered a triangular mesh with a Crouzeix-Raviart nonconforming approximation as a reference solution. One may clearly see that u_2 computed without $R(\cdot,\cdot)$ is not correct and that τ_{12} is constant.

REMARK 4.1.3. Note that in the triangular case, $R(\cdot,\cdot)$ vanishes since $\mathbf{u}_h$ is piecewise linear.

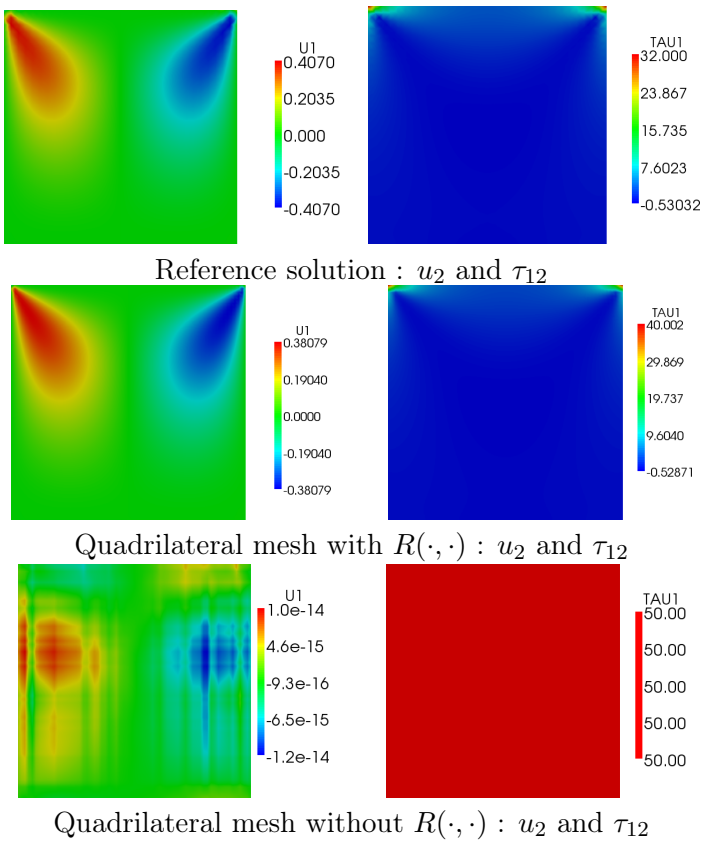


FIGURE 4.1.1. Influence of the regularization term $R(\cdot,\cdot)$

4.1.4. Numerical results. We have implemented both the triangular and the quadrilateral numerical schemes in the library CONCHA, for the Giesekus law but also for other differential models of polymers such as upper convected Maxwell, the Oldroyd-B and Phan-Thien Tanner (PTT) models. The behaviour of the errors with respect to mesh refinement was studied, and several academic test-cases such as channel flow, 4:1 and 4:1:4 contractions, flow past a cylinder were treated.

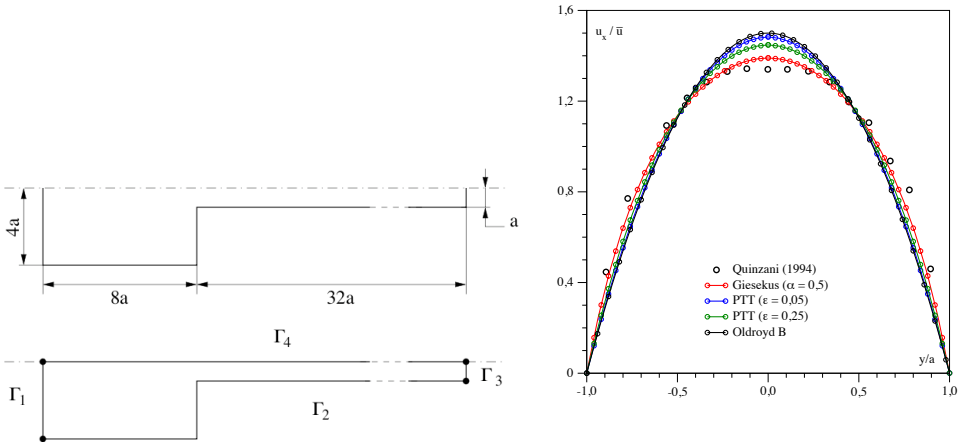
In order to validate the approximation of the Giesekus model, several comparisons were carried out: comparison with a semi-analytical solution of [**120**], comparison with measures given in [**152**], comparison with POLYFLOW which is the most popular commercial code for the simulation of polymer flows. We were able to perform simulations for high Weissenberg numbers and to obtain physically acceptable results, which we have tried to justify from a mathematical point of view in Section 4.2.

Finally, comparisons with drag values found in the literature for the Oldroyd-B model were also carried out for a benchmark problem. More tests can be found in [**107**].

4.1.4.1. *Pertinence of rheological models.* We first validate the choice of the Giesekus model by comparison between different models and measures. For the sake of clarity, it is useful to recall the Oldroyd-B constitutive law:

$$\lambda \overset{\triangledown}{\underline{\tau}} + \underline{\tau} = 2\eta \left(\underline{\varepsilon}(\mathbf{u}) + \lambda_r \overset{\triangledown}{\underline{\varepsilon}}(\mathbf{u}) \right)$$

with λ_r the retardation time, and the Phan-Thien Tanner law:



a) Geometry of the 4:1 contraction b) Comparison of u_1 in the channel

FIGURE 4.1.2. Giesekus, PTT and Oldroyd-B models vs. experimental data

$$\lambda \overset{\nabla}{\underline{\tau}} + \exp(\frac{\varepsilon\lambda}{\eta}\mathrm{tr}\underline{\tau})\underline{\tau} = 2\eta\underline{\varepsilon}(\mathbf{u})$$

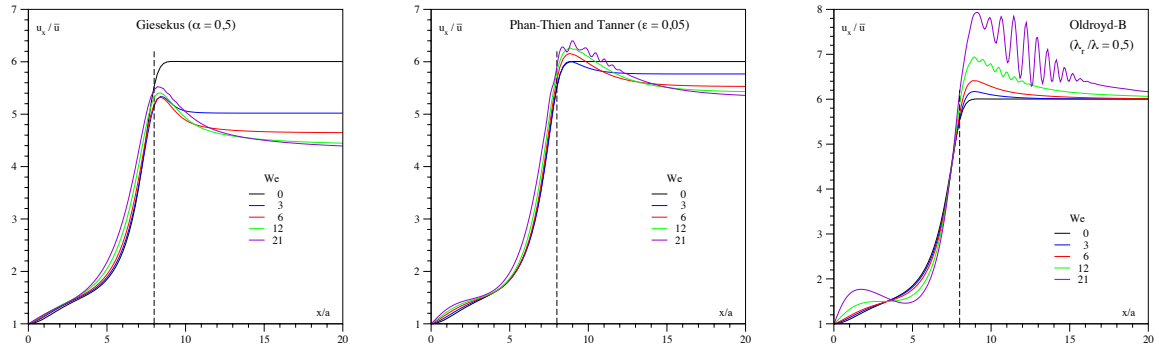
with $\varepsilon > 0$ an additional parameter.

We consider a 4:1 contraction cf. Fig. 4.1.2 a) and compare in Fig. 4.1.2 b) the first component of the velocity obtained experimentally in **[152]** with the one computed on a triangular mesh of 32748 cells by the following models : Giesekus, PTT with two choices of parameter ($\varepsilon = 0.05$ and $\varepsilon = 0.25$) and Oldroyd-B. One can observe the shear thinning effect (implying the flatenning of the velocity profile) with both the Giesekus and PTT nonlinear models. The Giesekus models yields a velocity very close to the measured one while for PTT, it is possible to optimally choose ε in order to get closer to the desired values. However, the value of ε would be then inappropriate to describe the elongational flow in the channel. Meanwhile, the Oldroyd-B model yields a parabolic profile, characteristic for a Newtonian fluid.

The next test also points out the more realistic behaviour of the Giesekus model. We compare in Fig. 4.1.3 the velocity u_1 along the symmetry axis for different Weissenberg numbers, for the three previous models. On the inflow boundary Γ_1, we set $\bar{u} = 0.1$m/s and $\underline{\tau}^D = \underline{0}$, on Γ_2 and Γ_3 we impose homogeneous Dirichlet and Neumann conditions respectively, and Γ_4 is a symmetry axis. We compute the Weissenberg number associated with a Newtonian liquid by the formula $We = \frac{12\lambda\bar{u}}{a}$ where $a = 0.001$m. For more details on the computation of the Weissenberg number in different geometries, see **[107]**.

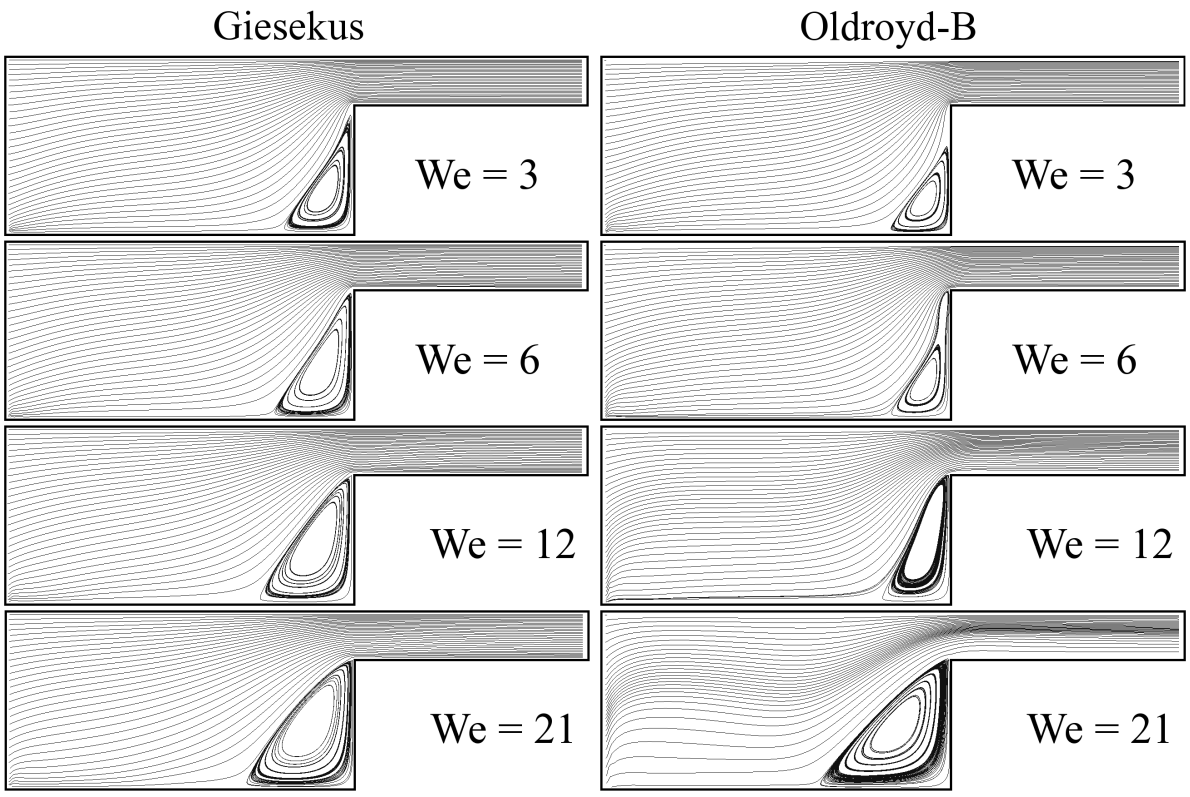
One can see that the velocity computed with the Oldroyd-B model is larger than the Newtonian one, which illustrates again the fact that this model is not realistic. The peak of u_1 near the contraction is explained by the appearance of normal stresses. Moreover, the PTT and Oldroyd-B models present oscillations of increasing amplitude with respect to the Weissenberg number. We have also shown in Fig. the streamlines for the Giesekus and Oldroyd-B models; the behaviour is typical for polymer flows, in particular the recirculation zone before the contraction.

4.1.4.2. *Comparison with Polyflow.* We consider here a popular benchmark in the computational rheology community, the 2D flow past a cylinder. The geometry is described in Fig. 4.1.5 (a) where R, the radius of the cylinder, is equal to 1m. We impose the same inflow conditions as in **[68]**, in particular a parabolic velocity profile with $u_{\mathrm{mean}} = 1$m/s. On the outflow we impose a homogeneous Neumann condition, and on the other boundaries, including the cylinder, no-slip conditions.



a) Giesekus model b) PTT model c) Oldroyd-B model

FIGURE 4.1.3. Velocity profiles along the symmetry axis in a 4:1 contraction



a) Giesekus model b) Oldroyd-B model

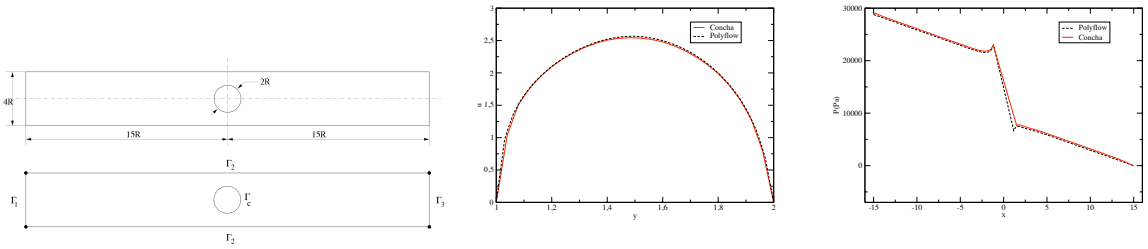
FIGURE 4.1.4. Streamlines for different Weissenberg numbers

We simulate the Giesekus model for which we compute the Weissenberg number by the relation $We = \lambda\dot{\gamma} = 12\lambda\bar{u}$ where $\bar{u}$ is the mean inflow velocity and where the shear-rate $\dot{\gamma}$ is computed for an equivalent Newtonian liquid. We take $\eta = 1000$Pa.s and $\rho = 1000$kg/m^3.

We compare our results with those obtained with POLYFLOW, with a discretization using Q_2-continuous elements for the velocity, Q_1-continuous for the pressure and the EVSS method with streamline-upwind for the stress tensor. The velocity profiles along the vertical axis passing through the centre of the cylinder, in the half domain, are shown in Fig.4.1.5 (b), while in Fig.4.1.5

(c) we present the pressures obtained along the horizontal symmetry axis of the channel. These results are obtained for $We = 6$ since for higher values, POLYFLOW has difficulties to converge. One observes a good agreement between the two approaches, the differences can be explained by the different meshes used (16 384 cells with CONCHA and 16 000 with POLYFLOW).

Similar conclusions were obtained when considering a $4 : 1 : 4$ contraction/expansion.



a) Geometry b) Velocity c) Pressure

FIGURE 4.1.5. Flow past a cylinder : comparison with POLYFLOW

4.1.4.3. *Simulations for high Weissenberg numbers.* The loss of convergence of the algorithms for high Weissenberg numbers is a major issue in computational rheology and is associated with the loss of the positivity of the so-called conformation tensor at the discrete level. We show in Section 4.2, based on the theory of algebraic Riccati equations, that our discretisation of the Giesekus model ensures the positivity of the discrete conformation tensor, under some mild conditions. We thus justify the good behaviour of the scheme which has been noticed in the numerical experiments for large Weissenberg numbers, contrarily to the Oldroyd-B model.

In this paragraph, we illustrate the stability and the robustness of the scheme with respect to the Weissenberg number in the academic test-case of flow past a cylinder. Moreover, the simulations exhibit the specific behaviour of polymers flows which are related to their elastic character and which increases with the relaxation time. In Fig.4.1.6, one can observe two swellings after the cylinder, explained by the emergence of important normal stresses above and below the cylinder, and also by the memory effect. An asymmetric velocity profile was obtained, typical for a polymeric liquid and due to the memory effect.

A similar bahaviour has been remarked in other test-cases.

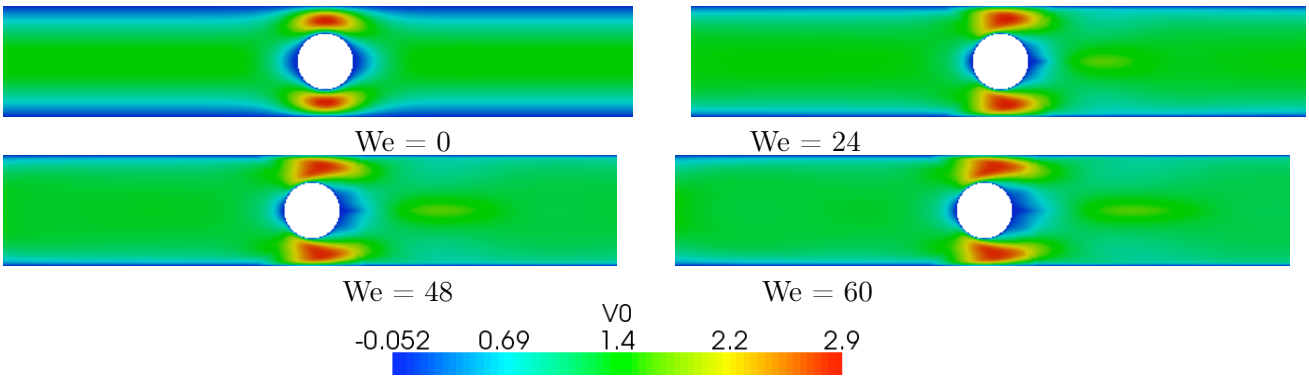


FIGURE 4.1.6. First component of the velocity for different Weissenberg numbers

4.1.4.4. *Drag comparison for Oldroyd-B model.* We now carry out some comparisons with results of the literature, in particular with [**141**], [**104**], [**80**] and [**68**] where reference drag values for the Oldroyd-B model are given. First, we write the model as in the previous papers; by means of the decomposition $\underline{\tau} = \underline{\tau}_p + 2\eta_n \underline{\varepsilon}(\mathbf{u})$ and of the relations $\lambda_r = \lambda \frac{\eta_n}{\eta_n + \eta_p}$ and $\eta = \eta_n + \eta_p$, we get the following equivalent form:

$$\left\{ \begin{array}{rl} -2\eta_n \mathrm{div}(\underline{\varepsilon}(\mathbf{u})) + \nabla p & = \mathrm{div}\underline{\tau}_p \\ \mathrm{div}\mathbf{u} & = 0 \\ \underline{\tau}_p + \lambda \overset{\triangledown}{\underline{\tau}_p} & = 2\eta_p \underline{\varepsilon}(\mathbf{u}). \end{array} \right.$$

Here above, η_n and η_p are the Newtonian and the polymeric viscosity respectively, and τ_p represents the polymeric stress tensor.

As in the cited papers, we take $\eta = 1$, $\eta_p = 0.41$ and the Weissenberg number defined by $We = \dfrac{\lambda u_{\mathrm{mean}}}{R}$, with u_{mean} the mean inflow velocity.

The drag along the cylinder Γ_c is given by the relation $\mathcal{C} = \int_{\Gamma_c} (1,0)^T \cdot \underline{\Pi}\mathbf{n}\mathrm{ds}$, with $\underline{\Pi} = \underline{\tau}_p - p\underline{I} + 2\eta_n \underline{\varepsilon}(\mathbf{u})$ the total stress tensor. We have implemented the same numerical scheme as for the Giesekus model.

In order to obtain accurate drag values, we have solved the linear system on a very fine mesh thanks to a multigrid method based on Vanka's smoother [**163**]. The values of $\mathcal{C}$ on successive meshes, for $\lambda = 0.6$, are given in Table 1. The linear convergence obtained yields extrapolated values $\mathcal{C}^*$ which are more accurate. n_N denotes the number of Newton iterates whereas n_M is the sum of the multigrid iterates.

One may see in Table 2 that the drag values $\mathcal{C}^*$ obtained with Concha on a mesh consisting of 1 048 576 elements, for different values of λ, are quite close to those of the literature, in particular with [**104**] and [**68**].

N	n_N	n_M	$\mathcal{C}$	$\Delta\mathcal{C}$	$\mathcal{C}^*$	
1024	7	19	118.081	-	-	
4096	6	12	118.421	0.340	-	
16384	6	18	118.349	0.072	-	
65536	6	24	118.085	0.264	117.821	
262144	5	20	117.936	0.149	117.787	
1048576	5	31	117.858	0.078	117.780	

Table 1. Drag coefficient on succesive meshes for λ=0.6

λ	0.0	0.3	0.6	0.7
Concha	132.357	123.190	117.780	117.321
Ref. [**141**]	132.357	-	117.775	-
Ref. [**104**]	132.358	123.193	117.792	117.29
Ref. [**80**]	132.33	123.41	-	-
Ref. [**68**]	-	123.194	117.779	117.321

Table 2. Comparison of drag coefficient for different λ

4.2. Positivity preserving scheme for a matrix-valued transport equation

Some of the next results can be found in [C13], and in the paper [B5] which will be soon submitted for publication.

To summarize, we consider the discontinuous Galerkin discretization of a general matrix-valued nonlinear transport with applications in non-Newtonian flows. Based on known results for algebraic Lyapunov equations, the convergence of a modified Newton method towards a positive definite solution is established, under a mild condition on the initialization of the iteration. Applications to Giesekus and Oldroyd-B models for polymer flows are discussed and finally, numerical simulations are presented.

4.2.1. Problem setting. Algebraic Riccati and Lyapunov equations. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a bounded domain with polygonal boundary and $\mathbf{v} \in \mathbf{W}^{1,\infty}(\Omega)$ a divergence free vector field. We consider the nonlinear first-order partial differential equation for a matrix-valued function $X = X^T$:

$$\begin{aligned} \mathbf{v} \cdot \nabla X - XG - G^T X + XDX + \alpha X = F \quad &\text{in } \Omega, \\ X = X^D \quad &\text{on } \partial\Omega^-, \end{aligned} \tag{4.2.1}$$

where G, D, F, X^D are matrix-valued functions with bounded, piecewise continuous coefficients and α is a bounded, piecewise continuous function. Assume moreover that F and X^D are symmetric positive definite, whereas D is only symmetric positive semidefinite; no sign hypothesis is necessary for α.

(4.2.1) is a system of transport equations for the components X_{ij}, $1 \leq i, j \leq d$, coupled by the zero-order terms. For $\mathbf{v} = \mathbf{0}$, it can be interpreted as a spatial or a time-discretized (for $\alpha = \frac{1}{\triangle t}$) Riccati equation.

Our motivation to study it stems from the modeling of non-Newtonian flows, where (4.2.1) describes the constitutive law of the conformation tensor as we shall see next.

One can then prove that (4.2.1) has a positive definite solution.

THEOREM 4.2.1. *Let $X^D \in \mathcal{C}(\partial\Omega^-, \mathcal{M}^d)$. If the solution of (4.2.1) is continuous on $\overline{\Omega}$, then $X > 0$.*

Proof. The main ingredient is the closed-form solution of differential Riccati equations, see for instance **[73]** where the authors analyzed equations of the type

$$\partial_t X - XA - A^T X + XDX = F.$$

I have adapted the proof to the steady case, by using the characteristics method. ■

REMARK 4.2.2. The same argument was applied in **[116]** to unsteady constitutive laws of viscoelastic fluids, which were writen as above by means of a Lie derivative instead of ∂_t.

A more challenging question is the positivity of the solution at the discrete level, which we address in what follows. It is useful to recall first some known results for Riccati and Lyapunov equations (see **[113]** or **[132]**), well-studied and used in optimal control.

A matrix A is said to be stable if its eigenvalues are in the open left half-plane; we write then $\Re\lambda(A) < 0$. A pair (A, D) is said to be stabilizable if there exists a (feedback) matrix T such that $A + DT$ is stable. Let us now consider an algebraic Riccati equation:

$$XDX - XA - A^T X = C \tag{4.2.2}$$

under the assumptions $D \geq 0$, (A, D) stabilizable and $C = C^T$.

The existence of a symmetric solution of (4.2.2) can be characterized using the spectral properties of the matrix $M = \begin{pmatrix} A & D \\ C & -A^T \end{pmatrix}$, cf. for instance **[143]** or **[113]**. In particular, existence is ensured if $D \geq 0$, $C \geq 0$, (A, D) and (A^T, C^T) stabilizable. A simpler but more restrictive condition is $D > 0$ and $C > 0$. The quadratic equation (4.2.2) does not have a unique solution. It is known that if (4.2.2) has a symmetric solution, then it admits a maximal symmetric solution X^* (i.e. $X^* \geq X$ for any other symmetric solution X). If $C \geq 0$ ($C > 0$), then (4.2.2) has symmetric solutions, and moreover the maximal solution X^* is nonnegative (positive).

The special case $D = 0$ is known as the Lyapunov equation. Let us recall the next key result.

LEMMA 4.2.3. *If A is stable, the Lyapunov equation*

$$XA + A^T X + C = 0 \tag{4.2.3}$$

admits a unique solution given by:

$$X = \int_0^\infty e^{At} C e^{A^T t} dt.$$

So if $C \geq 0$ ($C > 0$), then $X \geq 0$ ($X > 0$). Reciprocally, if $C > 0$, $X > 0$ and they satisfy (4.2.3), then A is necessarily stable.

4.2.2. Discretization scheme. Existence of a positive solution. We consider next $d = 2$ and we discretize (4.2.1) by a discontinuous Galerkin method. We follow the scheme of Lesaint and Raviart **[119]** introduced in the scalar case for a constant velocity field. The following results hold for both triangular and quadrilateral regular meshes h. We use piecewise constant elements for the tensor X and we define

$$V_h = \left\{ X_h \in L^2(\Omega, \mathcal{M}_{\rm sym});\ X_h|_K \in P_0\ \forall K \in \mathcal{T}_h \right\}.$$

As regards the discretization of the velocity field $\mathbf{v}$, we assume that we dispose of a finite element approximation $\mathbf{v}_h$ such that:

$$\pi_e \left[\mathbf{v}_h \cdot \mathbf{n}_e \right] = 0, \quad \forall e \in \mathcal{E}_h \quad \text{and} \quad \pi_K {\rm div} \mathbf{v}_h = 0, \quad \forall K \in \mathcal{T}_h$$

where π_ω is the $L^2(\omega)$-orthogonal projection on P_0.

REMARK 4.2.4. The applications we are having in mind are related to polymer flows. X is then the stress tensor while $\mathbf{v}$ is the fluid's velocity. It is easy to see that the previous properties are satisfied if we employ, for instance, nonconforming finite elements of Crouzeix-Raviart **[66]** in the triangular case, or of Rannacher-Turek **[154]** in the quadrilateral case.

The discontinuous finite element discretization of (4.2.1) reads:

$$\begin{cases} X_h \in V_h \\ a_h(X_h, Y_h) = f(Y_h), \quad \forall Y_h \in V_h \end{cases} \tag{4.2.4}$$

where:

$$\begin{aligned} a_h(X_h, Y_h) &= -\sum_{K \in \mathcal{T}_h} \int_K \left(X_h G + G^T X_h - \alpha X_h \right) : Y_h \, {\rm dx} + \sum_{K \in \mathcal{T}_h} \int_K X_h D X_h : Y_h \, {\rm dx} \\ &\quad - \sum_{e \in \mathcal{E}_h} \int_e \mathcal{F}_e(X_h, \mathbf{v}_h, \mathbf{n}_e) : [Y_h] \, {\rm ds}, \\ f(Y_h) &= \sum_{K \in \mathcal{T}_h} \int_K F : Y_h \, {\rm dx} + \sum_{e \in \mathcal{E}_h \cap \partial\Omega^-} \int_e |\mathbf{v}_h \cdot \mathbf{n}_e| \, X^D : Y_h \, {\rm ds} \end{aligned}$$

and where $\mathcal{F}_e(X_h, \mathbf{v}_h, \mathbf{n}_e) = \{\mathbf{v}_h \cdot \mathbf{n}_e\}^+ X_h^{\text{in}} + \{\mathbf{v}_h \cdot \mathbf{n}_e\}^- X_h^{\text{ex}}$ is the numerical flux on the edge e.

Let us denote by X_i the cell-wise values of X_h on the cell K_i, depending on the numbering of the cells. Similar notations are used for the data X^D and α, G, D, F which are taken, for the sake of simplicity, piecewise constant (otherwise, we replace them by their piecewise constant L^2-projection).

The discrete system (4.2.4) can then be written as follows:

$$b_{ii}X_i + \sum_{j\neq i} b_{ij}X_j - X_i(G_i - \frac{1}{2}\alpha_i I) - (G_i - \frac{1}{2}\alpha_i I)^T X_i + X_i D_i X_i = L_i, \quad \forall i \tag{4.2.5}$$

where $L_i = b_i^D X_i^D + F_i > 0$ and

$$\begin{aligned} b_{ii} &= \frac{1}{|K_i|}\int_{\partial K_i^+} \mathbf{v}_h \cdot \mathbf{n}_e \mathrm{ds} \geq 0, \\ b_{ij} &= \frac{1}{|K_i|}\int_{\partial K_i^- \cap \partial K_j^+} \mathbf{v}_h \cdot \mathbf{n}_e \mathrm{ds} \leq 0, \quad j \neq i \\ b_i^D &= -\frac{1}{|K_i|}\int_{\partial K_i \cap \partial\Omega^-} \mathbf{v}_h \cdot \mathbf{n} \mathrm{ds} \geq 0. \end{aligned} \tag{4.2.6}$$

One has $b_{ii} + \sum_{j\neq i} b_{ij} - b_i^D = 0$ for all i. Note that (4.2.5) is not yet an algebraic Riccati equation (4.2.2), due to the term $\sum_{j\neq i} b_{ij}X_j$ which couples with the cells K_j such that $\partial K_i^- \cap \partial K_j^+ \neq \emptyset$.

By applying Newton's method to the previous nonlinear system, the iterate X_i^n satisfies on the cell K_i the linear equation:

$$-X_i^n A_i^{n-1} - (A_i^{n-1})^T X_i^n + \sum_{j\neq i} b_{ij}X_j^n = L_i^{n-1}$$

with $A_i^{n-1} = G_i - \frac{1}{2}(\alpha_i + b_{ii})I - D_i X_i^{n-1}$ and $L_i^{n-1} = L_i + X_i^{n-1} D_i X_i^{n-1}$.

In order to establish the positive definiteness of the iterates, we propose to modify Newton's method by means of a Gauss-Seidel splitting of the transport operator B. We write that $B = B_1 + B_2$ with

$$(B_1 X)_i = \sum_{j\leq i} b_{ij}X_j, \quad (B_2 X)_i = \sum_{j>i} b_{ij}X_j, \quad \forall i.$$

The previous splitting depends on the numbering of cells, and such strategies are known in computational fluid dynamics, see for example **[98]**. It is well-known cf. **[119]** that in the case where $\mathbf{v}$ is a constant vector on Ω, a numbering can be found such that

$$\partial K_i^- \subset (\cup_{j<i} \partial K_j^+) \cup \partial\Omega^- \quad \forall i, \tag{4.2.7}$$

and therefore the transport operator B becomes lower triangular ($B_2 = 0$). This result can be generalized to non-recirculating flows.

We now consider the algorithm

$$-X_i^n(G_i - \frac{1}{2}\alpha_i I - D_i X_i^{n-1}) - (G_i - \frac{1}{2}\alpha_i I - D_i X_i^{n-1})^T X_i^n + (B_1 X^n)_i = L_i^{n-1} - (B_2 X^{n-1})_i$$

which is equivalent to

$$-X_i^n A_i^{n-1} - (A_i^{n-1})^T X_i^n + \sum_{j<i} b_{ij}X_j^n = L_i^{n-1} - \sum_{j>i} b_{ij}X_j^{n-1}, \quad \forall i. \tag{4.2.8}$$

In order to prove the main result of this section, we assume next that

$$(\mathbf{H1}) \qquad \left(G_i - \frac{1}{2}(\alpha_i + b_{ii})I,\ D_i\right) \text{ is stabilizable}, \quad \forall i$$

to which we add another two standard assumptions (see **[132]**, **[113]**) on the starting value X^0:

$$(\mathbf{H2}) \qquad X_i^0 > 0$$
$$(\mathbf{H3}) \qquad \Re\lambda(A_i^0) < 0.$$

THEOREM 4.2.5. *Under the assumptions **(H1)** to **(H3)**, the algorithm (4.2.8) converges monotonically towards a positive solution X^* of (4.2.4).*

Proof. We first show by induction that the sequence of Newton's iterates $(X_i^n)_{n\in\mathbb{N}}$ is nonincreasing and positive, on every cell K_i. According to (4.2.8), X_i^n satisfies an algebraic Lyapunov equation:

$$X_i^n A_i^{n-1} + (A_i^{n-1})^T X_i^n + C_i^n = 0 \tag{4.2.9}$$

where

$$C_i^n = L_i^{n-1} - \sum_{j>i} b_{ij} X_j^{n-1} - \sum_{j<i} b_{ij} X_j^n.$$

We also have that

$$X_i^n A_i^n + (A_i^n)^T X_i^n + L_i^n - \sum_{j>i} b_{ij} X_j^{n-1} - \sum_{j<i} b_{ij} X_j^n + (X_i^n - X_i^{n-1}) D_i (X_i^n - X_i^{n-1}) = 0 \tag{4.2.10}$$

and, by combining the previous relations,

$$\begin{aligned}&(X_i^n - X_i^{n+1})A_i^n + (A_i^n)^T(X_i^n - X_i^{n+1}) - \textstyle\sum_{j<i} b_{ij}(X_j^n - X_j^{n+1})\\ &- \textstyle\sum_{j>i} b_{ij}(X_j^{n-1} - X_j^n) + (X_i^n - X_i^{n-1})D_i(X_i^n - X_i^{n-1}) = 0, \qquad \forall n \geq 1.\end{aligned} \tag{4.2.11}$$

Thanks to Lemma 4.2.3, we obtain inductively on n and i from (4.2.9) that X_i^n is positive since C_i^n is positive, from (4.2.10) that A_i^n is stable and from (4.2.11) that $X_i^n - X_i^{n+1}$ is nonnegative. Hence, for all i the limit $X_i^* = \lim_{n\to\infty} X_i^n$ exists, is nonnegative and satisfies the equation (4.2.5).

In order to prove its positivity, we introduce the positive matrix $C_i^* = L_i - \sum_{j\neq i} b_{ij} X_j^*$ and note that X_i^* satisfies the Riccati equation (4.2.2), with matrices D_i, $A_i = G_i - \frac{1}{2}(\alpha_i + b_{ii})I$ and C_i^*. Then by adapting some results from **[113]** (Theorems 7.2.8 and 7.9.4), we can deduce that $A_i^* = \lim_{n\to\infty} A_i^n$ is stable and $X_i^* > 0$, since it satisfies a Lyapunov equation of matrices A_i^* and $C_i^* + X_i^* D_i X_i^*$. ∎

REMARK 4.2.6. At this stage, I could not apply any known result to conclude that X_i^* is maximal, because the right-hand side C_i^* of the Lyapunov equation mentioned above depends on the solution itself. I could only establish the maximality of X_i^* in the case where a numbering (4.2.7) exists, by showing inductively on n and i that $X_i^n - Y_i \geq 0$, for any symmetric solution Y of (4.2.5). A similar remark holds as regards the convergence rate of Newton's method applied to (4.2.4). We could establish the quadratic order under the hypothesis (4.2.7); the general case is ongoing work.

Finally, it is important to note that we can also prove the previous theorem in a variational framework, by choosing in (4.2.4) appropriate test-functions $Y_i = y_i^n \otimes y_i^n$ with y_i^n an eigenvector of X_i^n. This is encouraging for the extension to other discretizations than discontinuous Galerkin.

4.2.3. Application to polymer flows. We have next applied the previous result to the Giesekus and the Oldroyd-B models for polymer flows.

4.2.3.1. *The Giesekus model.* Its constitutive law is given by (4.1.1). Let the conformation tensor

$$X = \frac{\lambda}{\eta}\tau + I. \tag{4.2.12}$$

REMARK 4.2.7. The conformation tensor can be seen, in the case of an elastic solid, as the Cauchy tensor $\mathcal{X}$. Indeed, thanks to the generalized Hooke's law one has $\tau = 2G\,E$ with $E = \frac{1}{2}(\mathcal{X} - I)$ the Green-Lagrange tensor and $G = \frac{\eta}{\lambda}$ the elastic modulus. Since (4.2.12) is equivalent to $\tau = G(X - I)$, it follows that $X = \mathcal{X}$ if one considers the viscoelastic liquid as an elastic solid.

We mainly focus on the steady case. Then (4.1.1) can be written in terms of X as follows:

$$\mathbf{v}\cdot\nabla X - \nabla\mathbf{v}X - X\nabla\mathbf{v}^T + \frac{1-2\bar{\alpha}}{\lambda}X + \frac{\bar{\alpha}}{\lambda}X^2 = \frac{1-\bar{\alpha}}{\lambda}I,$$

to which we add the boundary condition $X = X^D$ on $\partial\Omega^-$, with $X^D > 0$ by hypothesis. By putting

$$G = (\nabla\mathbf{v})^T, \quad D = \frac{\bar{\alpha}}{\lambda}I, \quad F = \frac{1-\bar{\alpha}}{\lambda}I, \quad \alpha = \frac{1-2\bar{\alpha}}{\lambda},$$

it can be obviously recasted into the general form (4.2.1). Since $\bar{\alpha} \in]0,1[$, it is obvious that $F > 0$ and $D > 0$. So the conformation tensor is positive definite, according to Theorem 4.2.1.

REMARK 4.2.8. A different proof for the positivity of X was given by Hulsen in **[103]** for the Giesekus and Leonov constitutive laws, in the case without convection. He analyzed the sign of $\partial_t(\det X)$, and implicitly of the eigenvalues of X, by means of an ordinary differential equation satisfied by $\det X$. More recently, a similar idea was used in **[37]** for the unsteady Oldroyd-B model with convection.

Let us now check the hypotheses of Theorem 4.2.5. Since $\left((\pi_{K_i}\nabla\mathbf{v}_h)^T - \frac{1-2\bar{\alpha}+\lambda b_{ii}}{2\lambda}I,\, I\right)$ is trivially stabilizable for all i, **(H1)** is satisfied. As regards **(H2)** and **(H3),** it is sufficient to consider $X_i^0 = \delta_i I$ with

$$\delta_i > \frac{\lambda}{2\bar{\alpha}}\max\left\{0,\, 2\Re\lambda(\pi_{K_i}\nabla\mathbf{v}_h) - (1-2\bar{\alpha}+b_{ii})\right\}, \quad \forall i.$$

In conclusion, the discontinuous Galerkin method combined with a modified Newton algorithm yield a positive solution X_h, under the sole assumption on the choice of the initial iterate. Even if this result can still be improved (as mentioned in the perspectives), it is at our knowledge the first of this type which holds for both the steady and the unsteady cases, independently of the time step.

4.2.3.2. *The Oldroyd-B model.* We recall that the constitutive law of the Oldroyd-B model **[140]** can be written in the steady case as follows:

$$\lambda(\mathbf{v}\cdot\nabla)\tau_p - \lambda(\nabla\mathbf{v}\tau_p + \tau_p\nabla\mathbf{v}^T) + \tau_p = 2\eta_p\varepsilon(\mathbf{v}).$$

By means of the conformation tensor $X = \frac{\lambda}{\eta_p}\tau_p + I$, it becomes:

$$\mathbf{v}\cdot\nabla X - \nabla\mathbf{v}X - X\nabla\mathbf{v}^T + \frac{1}{\lambda}X = \frac{1}{\lambda}I$$

which can be put under the form (4.2.1) with obvious notations:

$$G = \nabla\mathbf{v}^T, \quad D = 0, \quad F = \frac{1}{\lambda}I, \quad \alpha = \frac{1}{\lambda}.$$

Therefore, the matrix A_i^n does not depend on the iteration since $A_i^n = A_i = (\pi_{K_i}\nabla\mathbf{v}_h)^T - \dfrac{1+\lambda b_{ii}}{2\lambda}I$.

The hypothesis **(H1)** translates into A_i stable for all i. Clearly, it may occur that this condition is violated for a large relaxation time λ, such that $\frac{1}{\lambda} \leq 2\Re\lambda(\pi_{K_i}\nabla\mathbf{v}_h) - b_{ii}$, in which case one cannot apply Theorem 4.2.5 in order to guarantee the positivity of the solution in the steady case.

Note that the instationary case is easier to treat since A_i is now replaced by $A_i - \frac{1}{2\Delta t}I$, which can be rendered stable for Δt small enough.

Finally, let us comment the results of Lee and Xu [**116**] related to the instationary Oldroyd-B model. They employed the characteristics method together with a positivity preserving projection operator π_h . In the case of a piecewise constant approximation, their discrete equation at t_n is

$$-X_i^n\tilde{A}_i - (\tilde{A}_i)^T X_i^n = \frac{1}{\lambda}I + \frac{1}{\Delta t}\pi_h(X_i^{n-1}\circ\chi^{n-1})$$

with $\tilde{A}_i = (\pi_{K_i}\nabla\mathbf{v}_h)^T - (\frac{1}{2\lambda} + \frac{1}{2\Delta t})I$.

This is obviously a Lyapunov and not a Riccati equation, since the quadratic term is missing. Therefore, the stability of the matrix $\tilde{A}_i$ is necessary in order to ensure the positivity of the solution, according to Lemma 4.2.3. This aspect seems to have been neglected in [**116**], where the authors claimed the positivity of X_h without any restriction on Δt or on the relaxation time λ.

So as regards the Oldroyd-B model, the positivity can be deduced only for sufficiently small Weissenberg numbers in the steady case, or for sufficiently small time steps in the unsteady case.

4.2.4. Numerical results. I present next some tests related to the positivity of the solution; the developed code is integrated in the library Concha.

4.2.4.1. *Constant velocity field.* We first consider the general equation (4.2.1) with constant velocity field $\mathbf{v}$ and constant input G, D and α, in order to compute the coefficients b_{ii} and to ascertain the importance of hypothesis **(H1).**

We take $\Omega =]-1,1[\times]-1,1[$, a velocity field $\mathbf{v} = (1,\, 0.25)$, $\alpha = 0$ and a fixed quadrilateral uniform mesh.

We first treat the linear case $D = 0$ and we take a constant matrix $G = \begin{pmatrix} -g & 1 \\ -2 & g \end{pmatrix}$ with $\mathrm{tr}G = 0$. The hypothesis **(H1)** translates into either $g^2 \leq 2$, or $g^2 \geq 2$ if $b_{ii} \geq 2\sqrt{g^2-2}$. For a mesh with $|e| = 2^{-2}$, we get after computation that $b_{ii} = 5$ for all i and hence, the critical value for g is $\sqrt{8.25} \approx 2.872$. We vary g and show in Table 3 (a) the minimum and maximum values of the eigenvalues. As expected, for g close or larger than the critical value the system is ill-conditioned and one completely loses stability.

We next consider the nonlinear case $D = I$ and keep the same mesh. The hypothesis **(H1)** is then checked and we now obtain in Table 3 (b) positive solutions for values of g larger than in the previous case.

4.2.4.2. *Computed velocity field.* In what follows, we take a variable velocity field $\mathbf{v}$, computed by solving Stokes or Navier-Stokes equations with nonconforming quadrilateral elements, and we take $G = (\nabla\mathbf{v})^T$.

We first treat the driven cavity with Stokes velocity on a mesh consisting of 1024 cells. We fix $F = I$, $X^D = 0$, we take $D = d\,I$ and we vary d. The minimum eigenvalues of X are shown in Table 4 and they are positive for all $d > 0$. Note that we couldn't get convergence of Newton's algorithm for $d = 0$.

TABLE 3. Constant velocity: minimum and maximum eigenvalues for different G

g	λ_{min}	λ_{max}
2	0.0444611	33754.3
2.5	0.0247383	5.2105×10^{9}
2.6	0.0230103	5.2235×10^{11}
2.7	0	4.689×10^{14}
2.8	0	2.02551×10^{20}
2.87	-2.25×10^{15}	6.39×10^{42}
2.873	-2.13×10^{50}	1.76×10^{47}

(a) Linear case

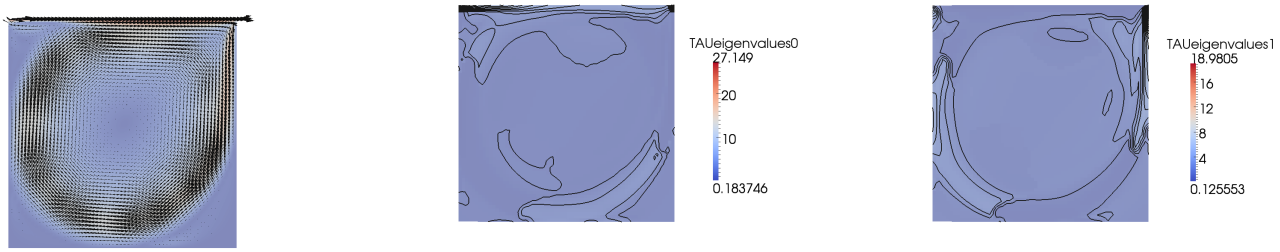
g	λ_{min}	λ_{max}
2	0.0349058	3.20822
2.5	0.0233634	4.21265
2.6	0.0219904	4.42953
2.7	0.0207962	4.64724
2.8	0.0197479	4.8669
2.87	0.0190868	5.02236
2.873	0.0190597	5.02901
3	0.0179912	5.31022
5	0.0100943	9.59359
10	0.00500466	19.7996

(b) Nonlinear case

We have also considered the case with recirculation. The velocity is now computed by the Navier-Stokes equations cf. Fig. 4.2.1 (a), with Reynolds number $Re = 4000$ on a mesh of 4096 cells. We still get positive eigenvalues; for $D = I$ they are shown in Fig. 4.2.1 (b) and (c).

TABLE 4. Driven cavity test with Stokes velocity: eigenvalues for different D

d	λ_{min}	λ_{max}
1	0.1296	12.8518
0.5	0.186456	25.0031
0.25	0.284703	48.8578
0.1	0.52891	119.027
0.05	0.850888	234.13
0.025	1.06101	461.579
0.0125	1.35125	911.945
0.00625	1.77429	1805.5
0.003125	2.42633	3581.72
0.001	4.41702	11087.6
0.0001	11.6608	110185

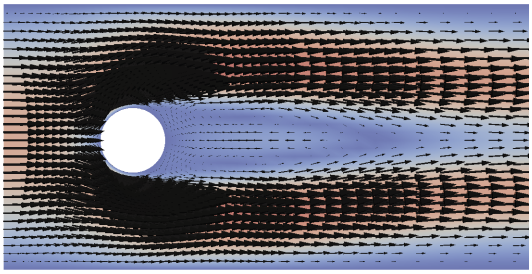


a) Velocity field b) First eigenvalue c) Second eigenvalue

FIGURE 4.2.1. Isolines of the eigenvalues with Navier-Stokes velocity

4.2.4.3. *Giesekus constitutive law.* We now consider the flow around a cylinder on a mesh of 2560 cells. We show in Fig. 4.2.2 (a) the velocity computed by Navier-Stokes equations with Reynolds number $Re = 80$; one can see the recirculation zones after the cylinder. We take $X^D = 0$ and $F = D = d\,I$, which corresponds to the Giesekus constitutive law with $\bar{\alpha} = 1/2$ and with different relaxation times $\lambda = \frac{1}{2d}$. The solution at $d = 1$ is represented in Fig. 4.2.2 (b), (c), (d). We have computed the Weissenberg number equivalent to a Newtonian liquid and obtained $We = 1.023933\lambda$. We show in Table 5 the minimum and maximum eigenvalues obtained for different d. One may note that the solution remains positive for very large Weissenberg numbers.

4.2.4.4. *Oldroyd-B constitutive law.* We change now F and D in order to obtain the Oldroyd-B constitutive law and we take $D = 0$, $F = \frac{1}{\lambda}I$ and $\alpha = \frac{1}{\lambda}$. For $\lambda = 1$ one still gets a positive solution cf. Fig. 4.2.3 but, contrarily to the nonlinear Giesekus law, this is not the case for any λ. Indeed, one can see in Fig. 4.2.4 that for $\lambda = 2$ the eigenvalues become negative. It is interesting to note that the corresponding Weissenberg number $We = 2.047866$ is approximately equal to the critical value reported in the literature for the Oldroyd-B flow past a cylinder.



a) Velocity : zoom around the cylinder

b) X_{12}

c) X_{22}

FIGURE 4.2.2. Solution of Giesekus law for $\lambda = 0.5$

TABLE 5. Giesekus flow: minimum and maximum eigenvalues for different We

d	We	λ_{min}	λ_{max}
1	0.511967	0.360205	2.77298
0.8	0.639958	0.301105	3.10585
0.6	0.853278	0.234669	3.60822
0.4	1.279918	0.161258	4.49192
0.2	2.559835	0.0822221	6.6333
0.1	5.11967	0.0413202	10.3087
0.01	51.1967	0.00413901	73.454
0.001	511.967	0.000413908	720.706

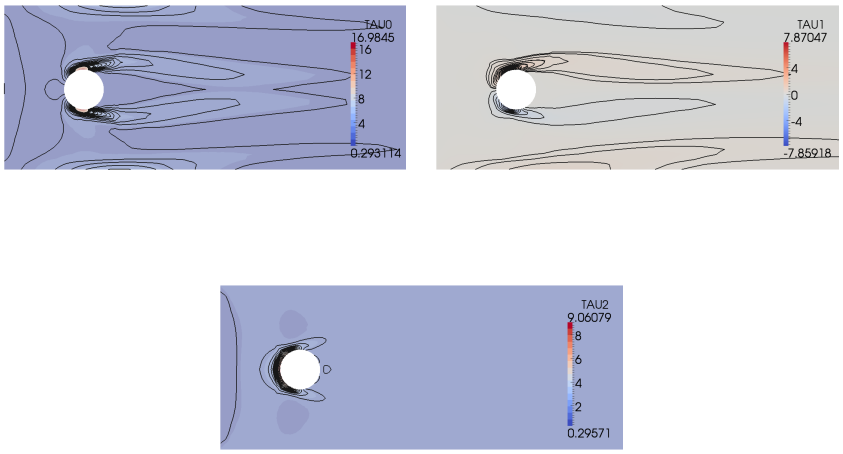


FIGURE 4.2.3. Solution of Oldroyd-B law for $\lambda = 1$

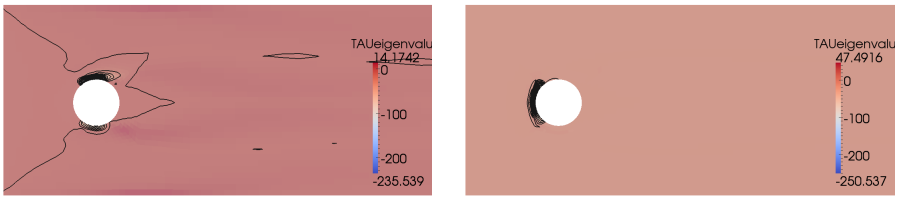


FIGURE 4.2.4. Eigenvalues of the Oldroyd-B conformation tensor for $\lambda = 2$

PERSPECTIVES

Perspectives

I describe next some short-term and mid-term perspectives that I see for my research. I am mainly interested in the analysis and implementation of numerical methods in fluid mechanics. Some of my projects are already ongoing works.

1. Ongoing projects

1.1. Robust discretization of polymer flows. I intend to continue my work on the numerical approximation of polymers and I would like to address the specific topics below.

Positivity preserving schemes

In Chapter 4, I have presented such a scheme for a general matrix-valued transport equation, which models certain constitutive laws of polymeric liquids. Our result was based on a lowest-order discontinuous Galerkin method together with a modification of Newton's algorithm.

In collaboration with Roland Becker, we are working on several challenging questions in order to improve the obtained result.

We can already recast it in a variational framework by choosing appropriate test-functions, which is a good starting point for the generalization to other discretizations than dG. Nevertheless, the treatment of the Lyapunov term $XG+G^TX$ in a general variational framework is still an open question, as far as positivity is concerned. Our further goal is the extension to higher-order conforming or nonconforming discretizations. We are also working on the convergence rate of the iterative method and on another variant of Newton's method.

Another topic of interest is the comparison with the log-transformation proposed in **[83]** and described in the Introduction. Besides the comparison of the numerical results obtained with the two methods, it could be also interesting to carry on an error analysis of the log-transformation for a simpler but still relevant model problem, such as the scalar diffusion-convection-reaction equation.

Finally, a related but very challenging question is the design of a monotone scheme for the transport operator, based on P_1-continuous elements with an adequate stabilization.

Energy estimates. Existence results for Giesekus model

It has been pointed out that the positivity of the discrete conformation tensor associated with a viscoelastic fluid often allows to show that the discrete energy of the system decreases, see for instance **[124]**, **[116]**, **[37]** for the Oldroyd-B model. Such numerical schemes seem to be more robust with respect to the Weissenberg number and moreover, the free energy estimates can be employed in order to prove existence of continuous and discrete solutions, as in **[37]** or **[18]**.

In **[37]**, the log-transformation is employed and a restriction on the time step was required to ensure that the approximation of the conformation tensor for the Oldroyd-B model remained positive definite. The results of **[37]** were improved by Barrett and Boyaval in **[18]**, where the authors considered a regularized Oldroyd-B model, with an additional dissipative term in the stress equation, and proved the positivity of the discrete conformation tensor without any constraint on the time step. For this purpose, they used the technique of cut-off functions

introduced by Barrett and Süli in [**21**] for the microscopic- macroscopic FENE model of a dilute polymeric fluid and they managed to show existence of global-in-time weak solutions to this regularized model (see also [**20**], [**19**] for similar results on other polymer models). Although no log-transformation is necessary in [**18**], one ends up with a discrete model depending on 6 parameters.

My objective is to obtain a similar result for the Giesekus model without any additional regularization and any restriction on the time step. I can already establish discrete free energy estimates under the hypothesis of a positive conformation tensor which is, as far as I know, new for the Giesekus model. In Section 4.2, we have shown this positivity when considering only the constitutive law, decoupled of the other equations. Now the remaining question is how to deduce it for the coupled problem; an idea could be to employ a fixed point method.

1.2. *A posteriori* error estimators based on $H(\mathrm{div})$- reconstructed fluxes. In collaboration with Robert Luce and Roland Becker, we are interested in *a posteriori* error estimators obtained by reconstructing locally conservative fluxes which belong to the Raviart-Thomas space on each cell. Our aim is to propose a unified framework for several finite element approximations (conforming, nonconforming and discontinuous Galerkin) and several model problems.

For the moment, we are focusing on the convection-diffusion-reaction equation but in the near future, we intend to consider other equations such as Stokes, Oseen, Navier-Stokes and viscoelastic constitutive laws. A first work on adaptive finite element methods for viscoelastic fluids was initiated in [C10].

There already exist references devoted to the design of unified theories for *a posteriori* error analysis without any $H(\mathrm{div})$-reconstruction, such as [**1**] or [**56**]. As regards the use of $H(\mathrm{div})$-reconstructed fluxes, we can cite for instance [**125**] for the mixed finite element method and [**78**] for the discontinuous Galerkin method applied to elliptic problems. Concerning now the diffusion-convection-reaction equation, the works [**79**] for the dG method, [**166**] for the finite volume method and [**167**] for the mixed finite element method yield a unified approach, in which the fluxes are constructed on a dual mesh formed by dual volumes around each vertex of the primal mesh.

Instead, we propose to use only the primal mesh, which presents certain advantages from a computational point of view. For this purpose, the construction of the $H(\mathrm{div})$-vector involved in the error estimator is inspired by the hypercircle method cf. [**39**] and is achieved on patches, which may overlap. A patch depends on the type of the employed finite elements and is defined as the support of a basis function. Thus, for the dG method the patch is reduced to the element itself, for the nonconforming method it is composed of two elements sharing a given edge whereas for the conforming method it consists of the elements sharing a given node. We finally obtain an *a posteriori* error estimator consisting only of the L^2-norm of a piecewise $H(\mathrm{div})$-vector ; the latter actually represents the correction of the discrete flux to a globally $H(\mathrm{div})$ one.

Note that with this approach, one can carry out exactly the same error analysis for all considered discretizations.

Our first theoretical and numerical results were presented in a mini-syposium at the Enumath conference in September 2011, and a paper is in preparation. For the moment, only triangular meshes have been treated and P_k discontinuous, P_1-nonconforming and P_1-continuous elements have been considered, the latter combined with the SUPG method [**51**]. Also, only upper error bounds have been established.

We are now working on the extension to other stabilizations, such as edge stabilization cf. [**52**] or local projection stabilization cf. [**23**], and to higher-order approximations. The next step is to prove lower error bounds as well as the convergence and optimality of the adaptive

algorithm. The treatment of quadrilateral meshes is another point of interest, for which specific technical questions need to be addressed.

Finally, let us point out that as regards the transport equation, the (weak) norm that is usually employed for the error analysis does not allow to carry out a goal-oriented error control. The choice of a more adequate norm is another open question.

1.3. Sensitivity analysis. Very recently, with my colleagues Roland Becker, Robert Luce and David Trujillo we have started to work on the *a posteriori* error estimation of sensitivities which arise in parameter-dependent problems in continuum mechanics.

Most practical applications involve parameters $q = (q_i)_{1\leq i\leq N}$ of different origins: physical (viscosity, heat conduction etc.), modeling (computational domain, boundary conditions etc.) and numerical (mesh, stabilization parameters, stopping criteria, values of a turbulence model). Numerical simulations can provide information related to the dependence of a quantity of physical interest $I(q)$ with respect to different parameters. Well-known examples of such functionals are the drag in fluid flow and the Nusselt number in heat transfer.

Our motivation for this study is that the computation of such sensitivities can help to validate the physical model, to explain unexpected behaviour and also to guide efforts to improve both the physical and the computational models. First and second order sensitivities $(\partial I/\partial q_i)_{1\leq i\leq N}$ and $(\partial^2 I/\partial q_i \partial q_j)_{1\leq i,j\leq N}$ could also be used in order to predict the change of the functional under parameter changes.

A posteriori error estimates for the quantity of physical interest, $I(q) - I_h(q)$, for fixed parameters q are well-known, see for example **[26]** where a goal-oriented error control is achieved by introducing an adjoint problem. It has been pointed out in **[27]** that the information provided by the adjoint problem can also be used to compute the discrete first order sensitivities. However, no error estimator for the error in the sensitivity $\partial I/\partial q_i - \partial I_h/\partial q_i$ has been given so far.

Our goal is to provide a general framework for the *a posteriori* error estimation of sensitivities. Following the approach of **[26]**, we have derived abstract error representations for these quantities involving interpolation errors which have to be further approximated, in order to obtain computable error estimates. Note that besides the state and adjoint equations, a tangent equation as well as a second adjoint equation need to be solved.

So far, we have considered in order to illustrate the theory the computation of the Nusselt number measuring the efficiency of a cooling process. A cold liquid is injected in a annular domain through several inlets in order to cool a heated interior stator. For the sake of simplicity, we have considered the dependence of the Nusselt number with respect to only one parameter, the inflow speed. First numerical results, including adaptation with respect to the functional and to the sensitivity, have been carried out with the library Concha. They have been presented in a mini-symposium at US National Congress on Computational Mechanics in July 2011 and a paper is in preparation.

Several important aspects related to the adaptive method are still to be investigated (design of an appropriate adaptive algorithm, proof of its convergence and optimality). Generalizations to nonconforming and stabilized methods should also be envisaged, and an *a posteriori* error analysis for the second order sensitivities should be performed. Finally, it would be interesting to carry out simulations at hand of concrete applications.

2. Future works

2.1. Anisothermal flows. A realistic description of complex flows in fluid mechanics implies to take into account the thermo-mechanical coupling. I have considered this aspect in my work but mostly in axisymmetric geometries in porous media (see Chapter 3).

In the future, I intend to study anisothermal Newtonian and non-Newtonian flows in a more general framework. The collaboration with the physicists Didier Graebling and Eric Schall, members of the INRIA team Concha, could be an advantage as regards the physical relevance of the considered models and of the obtained numerical results.

I have already supervised two trainees in the Master Degree on the simulation of fluid flows with heat transfer. The first subject, proposed by Didier Graebling, dealt with the axisymmetric flow of a generalized Newtonian fluid; the second one concerned the simulation with the library Concha of anisothermal Newtonian flows by using the mass flux $\rho\mathbf{v}$ as an additional unknown, see [C15]. I think that these previous works could serve as a basis for future developments.

There are many challenging questions related to the numerical approximation of compressible flows, both at theoretical and numerical levels. Some of them, to cite only a few, are: choice of boundary conditions and of functional spaces, choice of variables (primitive or conservative) and of discretization spaces, stabilization of the convective terms $\rho\mathbf{v}\cdot\nabla\mathbf{v}$ and $\rho\mathbf{v}\cdot\nabla T$ where the density ρ, the velocity $\mathbf{v}$ and the temperature T are unknown, robustness with respect to parameters such as Reynolds, Mach or Rayleigh numbers, development of efficient solvers etc.

Other questions are specific to the considered model. For instance, as regards polymeric liquids characterized by very high viscosity and very low thermal conductivity, it is known that the viscous dissipation term shouldn't be neglected in the energy equation since it plays an important role; moreover, the dependence of the viscosity on the temperature should also be taken into account. Meanwhile, the state equation could be simplified and an affine dependence of the density with respect to the temperature as in [C15] could be envisaged.

A related question is how to take into account the thermal exchange with a solid wall, in a finer way than imposing boundary conditions which are not well known from a physical point of view. An idea could be to solve the heat equation in the wall and to impose transmission conditions at the interface, which yields a fluid-structure interaction problem.

Another possible extension concerns the mixed formulation of the energy equation in an axisymmetric geometry, which naturally leads to finite elements of Raviart-Thomas type in an axisymmetric framework. There exist references on axisymmetric approximations of Stokes and Navier-Stokes equations **[28]**, **[29]** but not, as far as I know, for the space $H(\mathrm{div})$.

2.2. 3D approximations. Realistic applications require 3D simulations. The extension from 2D to 3D generates a large computational cost and thus includes non-trivial aspects at the theoretical, algorithmic and computational levels.

This implies the development of stable schemes but also of robust and efficient iterative solvers. In order to gain computing time and memory, modern numerical tools such as multigrid methods, adaptivity and parallelization of the code should also be employed.

In the near future, the library Concha will be enriched with the tools necessary for 3D simulations and several members of the team including myself will be involved in this task. Note that the parallelization of the library is already ongoing work within the team.

2.3. Higher-order nonconforming elements on quadrilaterals. I am also interested in higher-order approximations of fluid flows on quadrilateral cells, achieved with nonconforming finite elements. This presents certain advantages as regards the stencil of the matrix, the adaptivity and the generalization to 3D computations. For the moment, we are employing a first order method based on Rannacher-Turek **[154]** or Han **[99]** elements for the velocity and pressure in the Newtonian case; for non-Newtonian fluids, the additional unknown which is the stress tensor is approximated by piecewise constant elements. A cheaper non-conforming quadrilateral element was introduced in **[144]**, but only for elliptic problems.

To begin with, I would like to investigate how to obtain a second-order nonconforming method for Stokes equations, eventually by adding suitable stabilization terms and/or by enriching the pressure space. The final goal is to propose a second-order method for non-Newtonian flows, cheaper than those of [**68**] or [**104**] where Q_2 -continuous elements for the velocity, P_1-discontinuous elements for the pressure and Q_1 or Q_2 elements for the stress tensor are employed, with adequate edge stabilizations.

2.4. Applications of viscoelastic flows. So far, I have addressed rather academic questions related to the numerical simulation of polymers flows in some well-known benchmark problems. They are prerequisite to further industrial developments, which necessitate more sophisticated numerical tools as those mentioned here above in the paragraph 2.2.

An obvious field of application of our code could be the optimization of polymer processing techniques such as extrusion, injection moulding, film-blowing and mixing.

But there are many other application domains where the numerical simulation of viscoelastic fluids plays an important role, such as food-processing industry, cosmetics, medicine, biology etc.

In the future, I would like to treat more concrete problems in biomedicine. An example is to take into account the viscoelastic character of biological fluids, which allows a better modeling of the motion of cells in these liquids. The numerical simulations would lead to a better understanding of certain phenomena such as the asthenospermia, that is the male infertility due to a lack of mobility of sperm, or the movement of microorganisms in the mucus present in the respiratory system.

2.5. Free surface flows. I have considered the modeling of such flows in fluvial hydrodynamics, see Chapter 2. The focus in this work was on the derivation of hierarchical 2D and 1D models.

One of the topics which I am interested in is related to the non-standard boundary conditions on the free surface. Similar conditions to those imposed in Section 2.2 could be employed in other applications like ocean modeling, coastal flows, viscoelastic flows etc. Different weak formulations could be obtained and adequate discretizations (not necessarily conforming or *inf-sup* stable) could be proposed and studied.

Bibliography

[1] M. Ainsworth, J. T. Oden : *A unified approach to a posteriori error estimation using element residual methods*, Numer. Math., Vol. 65, n. 1, p. 23–50, 1993

[2] M. Amara, D. Capatina, D. Trujillo : *Modélisation 2D-horizontale de l'écoulement d'un fleuve*, 17 p, Préprint LMA UPPA n°23, 2003, *http://lma.univ-pau.fr/data/pub/pub_pdf2003/0323.pdf*

[3] M. Amara, D. Capatina, D. Trujillo : *Modélisation 2D-verticale et 1D de l'écoulement d'un fleuve*, 23 p, Préprint LMA UPPA n°24, 2003, *http://lma.univ-pau.fr/data/pub/pub_pdf2003/0324.pdf*

[4] M. Amara, D. Capatina-Papaghiuc, D. Trujillo : *Stabilized finite element method for Navier-Stokes equations with non-standard boundary conditions*, Preprint LMA UPPA n° 0325, 29 p, 2003*, http://lma.univ-pau.fr/data/pub/pub_pdf2003/0325.pdf*

[5] M. Amara, D. Capatina-Papaghiuc, D. Trujillo : *A 3D numerical model for the vorticity-velocity-pressure formulation of the Navier-Stokes problem*, Preprint LMA UPPA n° 0510, 11 p, 2005*, http://lma.univ-pau.fr/data/pub/pub_pdf2005/0510.pdf*

[6] M. Amara, E. Chacon Vera, D. Trujillo : *Vorticity-velocity-pressure formulation for Stokes problem*, Math. of Comp., Vol. 73, n. 248, p. 1673-1697, 2004

[7] D.N. Arnold : *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., Vol. 19, n. 4, p. 742-760, 1982

[8] D.N. Arnold, F. Brezzi, B. Cockburn, L.D. Marini : *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., Vol. 39, n. 5, p. 1749-1779, 2002

[9] D. N. Arnold, F. Brezzi, R.S. Falk, L.D. Marini : *Locking-free Reissner-Mindlin elements without reduced integration*, Comput. Methods Appl. Mech. Engrg., Vol. 96, p. 3660-3671, 2007

[10] D.N. Arnold, R.S. Falk : *A uniformly accurate finite element method for the Reissner-Mindlin plate,* SIAM J. Numer. Anal., Vol. 26, p. 1276-1290, 1989

[11] D.N. Arnold, R.S. Falk : *Asymptotic analysis of the boundary layer for the Reissner-Mindlin plate model*, SIAM J. Math. Anal., Vol. 27, no. 2, p. 486-514, 1996

[12] E. Audusse, M.-O. Bristeau, A. Decoene : *Numerical simulations of 3D free surface flows by a multilayer Saint-Venant model*, Int. J. Num. Meth. Fluids, Vol. 56, n. 3, p. 331-350, 2008

[13] M. Azaïez, C. Bernardi, N. Chorfi : *Spectral discretization of the vorticity, velocity and pressure formulation of the Navier–Stokes equations*, Numer. Math., Vol. 104, p. 1–26, 2006

[14] K. Aziz, A. Settari : *Petroleum Reservoir Simulation*, Applied Science Publishers, London, 1979

[15] F. Baaijens, M. Hulsen, P. Anderson : *The Use of Mixed Finite Element Methods for Viscoelastic Fluid Flow Analysis*, Encyclopedia of Computational Mechanics, Vol. 3 Fluids, p. 481, John Wiley & Sons Ltd., Chichester, 2004

[16] I. Babuška, M. Suri : *On locking and robustness in the finite element method*, SIAM J. Numer. Anal., Vol. 29, p. 1261-12931, 1992

[17] S. Badia, R. Codina : *Unified stabilized finite element formulations for the Stokes and the Darcy problems*, SIAM J. Numer. Anal., Vol. 47, n. 3, p. 1977-2000, 2009

[18] J. Barrett, S. Boyaval : *Existence and approximation of a (regularized) Oldroyd-B model*, Technical Report, http://hal.archives-ouvertes.fr/hal-00409594/en/, 2009

[19] J.W. Barrett, C. Schwab, E. Süli : *Existence of global weak solutions for some polymeric flow models,* Math. Models Methods Appl. Sci., Vol. 15, p. 939-983, 2005

[20] J.W. Barrett, E. Süli : *Existence of global weak solutions to some regularized kinetic models for dilute polymers*, Multiscale Model. Simul., Vol. 6, n. 2, p. 506-546 (electronic), 2007

[21] J.W. Barrett, E. Süli : *Existence of global weak solutions to dumbbell models for dilute polymers with microscopic cut-off*, Math. Models Methods Appl. Sci., Vol. 18, n. 6, p. 935-971, 2008

[22] F. Bassi, S. Rebay : *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys., Vol. 131, n. 2, p. 267-279, 1997

[23] R. Becker, M. Braack : *A finite element pressure gradient stabilization for the Stokes equations based on local projections*, Calcolo, Vol. 38, n. 4, p. 173–199, 2001

[24] R. Becker, D. Capatina, J. Joie : *A DG method for the Stokes equations related to nonconforming finite element methods*, 35 p, Rapport de Recherche INRIA, 2009, *http://hal.inria.fr/inria-00380772*

[25] R. Becker, S. Mao : *Quasi-Optimality of Adaptive Nonconforming Finite Element Methods for the Stokes Equations*, SIAM J. Numer. Anal., Vol. 49, n. 3, p. 970-991, 2011

[26] R. Becker, R. Rannacher : *An optimal control approach to a posteriori error estimation in finite element methods*, in A. Iserles (Ed.), Acta Numerica, Vol. 10, p. 1-102, Cambridge University Press, 2001

[27] R. Becker, B. Vexler : *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations*, J. Comp. Phs., Vol. 206, p. 95–110, 2005

[28] Z. Belhachmi, C. Bernardi, S. Deparis : *Weighted Clément operator and application to the finite element discretization of the axisymmetric Stokes problem*, Numer. Math., Vol. 105, p. 217–247, 2006

[29] Z. Belhachmi, C. Bernardi, S. Deparis, F. Hecht : *An efficient discretization of the Navier–Stokes equations in an axisymmetric domain, Part 1: The discrete problem and its numerical analysis*, J. Sci. Comp., Vol. 27, p. 97–110, 2006

[30] C. Bernardi, N. Chorfi : *Spectral discretization of the vorticity, velocity and pressure formulation of the Stokes problem*, SIAM J. Numer. Anal., Vol. 44, p. 826–850, 2006

[31] C. Bernardi, V. Girault : *A local regularization operator for triangular and quadrilateral finite elements*, SIAM J. Numer. Anal., Vol. 35, n. 5, p. 1893-1916, 1998

[32] C. Bernardi, F. Hecht, F.Z. Nouri : *A new finite element discretization of the Stokes problem coupled with Darcy equations*, IMA J. Numer. Anal., Vol. 30, p. 61–93, 2010

[33] C. Bernardi, F. Hecht, O. Pironneau : *Coupling Darcy and Stokes equations for porous media with cracks*, M2AN, Vol. 39, No. 1, p. 7–35, 2005

[34] C. Bernardi, F. Hecht, R. Verfürth : *A finite element discretization of the three-dimensional Navier-Stokes equations with mixed boundary conditions*, Math. Model. Numer. Anal., Vol. 43, p. 1185–1201, 2009

[35] F. Bouchut, M. Westdiskenberg : *Gravity driven shallow water models for arbitrary topography*, Comm. in Math. Sci., Vol. 2, n.3, p. 359-389, 2004

[36] G. Bourdarot : *Well testing : Interpretation methods*, Editions Technip, Paris, 1998

[37] S. Boyaval, T. Lelièvre, C. Mangoubi : *Free-energy-dissipative schemes for the Oldroyd-B model*, M2AN, Vol. 43, n. 3, p. 523-561, 2009

[38] M. Braack, A. Ern : *A posteriori control of modeling errors and discretization errors*, SIAM J. on MMS, Vol. 1, n. 2, p. 221-238, 2003

[39] D. Braess, R.H. Hoppe, J. Schoberl : *A posteriori estimators for obstacle problems by the hypercircle method*, Comput. Vis. Sci., Vol. 11, n. 4-6, p. 351–362, 2008

[40] S. Brenner : *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, J. Numerical Analysis, Vol. 41, n. 1, p. 306-324, 2003

[41] S. Brenner : *Korn's inequalities for piecewise H^1 vector fields*, Math. Comp., Vol. 73, n. 247, p. 1067-1087, 2004

[42] S. Brenner, R. Scott : *The Mathematical Theory of Finite Element Methods*, Springer Verlag, New York, 1994

[43] J. Bramble, T. Sun : *A negative-norm least squares method for Reissner-Mindlin plates*, Math. Comp., Vol. 67, p. 901-916, 1998

[44] D. Bresch, F. Guillen, J. Lemoine : *A note on a degenerate elliptic equation with applications for seas and lakes,* Elect. J. Diff. Eqs., Vol. 42, p. 1-13, 2004

[45] F. Brezzi : *On the existence, uniqueness, and approximation of saddle point problems arising from Lagrangian multipliers*, R.A.I.R.O. Anal. Numer., Vol. 8, n. 32, p. 129-151, 1974

[46] F. Brezzi, B. Cockburn, L.D. Marini, E. Süli : *Stabilization mechanisms in discontinuous Galerkin finite element methods*, Comput. Methods Appl. Mech. Engrg., Vol. 195, Issues 25-28, p. 3293-3310, 2006

[47] F. Brezzi, M. Fortin : *Mixed and Hybrid Finite Element Methods*, Springer Verlag, New York, 1991

[48] F. Brezzi, L.D. Marini, E. Süli : *Discontinuous Galerkin methods for first-order hyperbolic problems*, M3AS, Vol. 14, n. 12, p. 1893-1903, 2004

[49] F. Brezzi, J. Pitkäranta : *On the stabilization of finite element approximations of the Stokes equations*, in W. Hackbusch (Ed.), Efficient Solution of Elliptic Systems, Vieweg, 1984

[50] F. Brezzi, J. Rappaz, P.-A. Raviart : *Finite Dimensional Approximation of Nonlinear Problems. Branches of Nonsingular Solutions*, Numer. Math., Vol. 36, p. 1-25, 1980

[51] A. Brooks, T. Hughes : *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., Vol. 31, p. 199-259, 1982

[52] E. Burman, P. Hansbo : *Edge stabilizations for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg., Vol. 193, p. 1437-1453, 2004

[53] E. Burman, P. Hansbo : *Edge stabilization for the generalized Stokes problem : A continuous interior penalty method*, Comput. Methods Appl. Mech. Engrg., Vol. 195, n. 19-22, p. 2393-2410, 2006

[54] G. Caloz, J. Rappaz : *Numerical Analysis for Nonlinear and Bifurcation Problems,* Handbook of Numerical Analysis, Vol. V, P.G. Ciarlet and J.L. Lions eds, North-Holland, Amsterdam, 1997

[55] H. Cao : *Development of techniques for general purpose simulators*, Ph.D. Thesis, University of Stanford, 2002

[56] C. Carstensen, T. Gudi, M. Jensen : *A unifying theory of a posteriori error control for discontinuous Galerkin FEM*, Numer. Math., Vol. 112, n. 3, p. 363–379, 2009

[57] M. C. H. Chien, S. T. Lee, W. H. Chen : *A new fully implicit compositional simulator*, SPE 13385, Proceedings of the 8th SPE symposium on reservoir simulation, Dallas, 1985

[58] P.G. Ciarlet : *The Finite Element Method for Elliptic Problems*, North Holland, Amsterdam, 1978

[59] P. Clément : *Approximation by finite element functions using local regularization*, R.A.I.R.O. Anal. Numer. , Vol. 9 R2, p. 77-84, 1975

[60] K. H. Coats : *An equation of state compositional model*, SPE J., Vol. 38, p. 363-376, 1980

[61] B. Cockburn, G. Kanschat, D. Schötzau, C. Schwab : *Local discontinuous Galerkin methods for the Stokes system*, SIAM J. Numer. Anal., Vol. 40, n.1, p. 319–343, 2002

[62] R. Codina : *Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods*, Comput. Methods Appl. Mech. Engrg., Vol. 190, pp. 1579-1599, 2000

[63] R. Codina, J. Blasco : *Analysis of a pressure-stabilized finite element approximation of the stationary Navier-Stokes equations,* Numer. Math., Vol. 87, p. 59-81, 2000

[64] C. Conca, C. Pares, O. Pironneau, M. Thiriet : *Navier-Stokes Equations with imposed pressure and velocity fluxes,* Int. J. Num. Meth. Fluids, Vol. 20, p. 267-287,1995

[65] M. Costabel : *A remark on the regularity of solutions of Maxwell's equations on Lipschitz domains*, Math. Meth. Appl. Sci., Vol. 12, p. 365-368, 1990

[66] M. Crouzeix, P.-A. Raviart : *Conforming and nonconforming finite element methods for solving the stationary Stokes equations*, R.A.I.R.O., Vol. 7, p. 33-75, 1973

[67] M. Crouzeix, R. Falk : *Nonconforming finite elements for the Stokes problem*, Math. Comp., Vol. 52, n. 186, p. 437-456, 1989

[68] H. Damanik, J. Hron, A. Ouazzi, S. Turek : *A monolithic FEM approach for the log–conformation reformulation (LCR) of viscoelastic flow problems*, J. Non-Newtonian Fluid Mech., 2010

[69] E. Dari, R. Duran, C. Padra : *Error estimators for nonconforming finite element approximations of the Stokes problem*, Math. Comp., Vol. 64, n. 211, p. 1017-1033, 1995

[70] A. Decoene, J.-F. Gerbeau : *Sigma transformation and ALE formulation for three-dimensional free surface flows*, Int. J. Num. Meth. Fluids, Vol. 59, n. 4, p. 357-386, 2009

[71] B. Denel : *Simulation numérique et couplage de modèles thermomécaniques puits-milieux poreux*, Thèse de doctorat, Université de Pau, 2004

[72] P. Destuynder, M. Salaün : *Mathematical Analysis of Thin Plate Models*, Springer, 1996

[73] L. Dieci, T. Eirola : *Positive definiteness in the numerical solution of Riccati differential equations*, Numer. Mathematik, Vol. 67, p. 303-313, 1994

[74] D.A. Di Pietro, A. Ern : *Mathematical Aspects of Discontinuous Galerkin Methods*, Mathématiques et Applications, Vol. 69, Springer-Verlag, Berlin, 2011

[75] M. Discacciati, A. Quarteroni : *Convergence analysis of a subdomain iterative method for the finite element approximation of the coupling of Stokes and Darcy equations*, Comput. Visual. Sci., Vol. 6, Numbers 2-3, p. 93-103, 2004

[76] M. Discacciati, A. Quarteroni : *Navier-Stokes / Darcy Coupling: Modeling, analysis, and numerical approximation*, Rev. Mat. Complut., Vol. 22, n. 2, p. 315-426, 2009

[77] F. Dubois, M. Salaün, S. Salmon : *Vorticity-velocity-pressure and stream function-vorticity formulations for the Stokes problem,* J. Math. Pures Appl., Vol. 82, p. 1395-1451, 2003

[78] A. Ern, S. Nicaise, M. Vohralik : *An accurate $H(\mathrm{div})$ flux reconstruction for discontinuous Galerkin approximations of elliptic problems*, C.R. Acad. Sci. Paris, Series I, Vol. 34, n. 12, p. 709-712, 2007

[79] A. Ern, A.F. Stephansen, M. Vohralik : *Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems*, J. Comput. Appl. Math., Vol. 234, n. 1, p. 114-130, 2010

[80] J. Etienne, E.J. Hinch, J. Li : *A Lagrangian–Eulerian approach for the numerical simulation of free-surface flow of a viscoelastic material*, J.Non-Newtonian Fluid Mech., Vol. 136, p. 136-157, 2006

[81] R. Eymard, T. Gallouët, R. Herbin : *Finite Volume Methods*, in Handbook of Numerical Analysis, vol. VII, P.G. Ciarlet & J.L. Lions, ed., North Holland, Amsterdam, p. 713-1020, 2000

[82] R. Falk, T. Tu : *Locking-free finite elements for the Reissner-Mindlin plate*, Math. Comp., Vol. 69, p. 911-928, 2000

[83] R. Fattal, R. Kupferman : *Constitutive laws of the matrix-logarithm of the conformation tensor*, J.Non-Newtonian Fluid Mech., Vol. 123, p. 281-285, 2004
[84] S. Ferrari, F. Saleri : *A new two-dimensional shallow-water model including pressure effects and slow varying bottom topography*, M2AN, Vol. 38, n. 2, p. 211-234, 2004
[85] P. Forchheimer : *Wasserbewegung durch Boden*, Z. Ver. Deutsh. Ing., Vol. 45, p. 1782–1788, 1901
[86] L. Formaggia, J.-F. Gerbeau, F. Nobile, A. Quarteroni : *On the coupling of 3D and 1D Navier-Stokes equations for Flow Problems in Compliant Vessels*, Comp. Methods Appl. Mech. Engnrg., Vol. 191 (6-7), p. 561-582, 2001
[87] M. Fortin, A. Fortin : *A new approach for the FEM simulation of viscoelastic flows*, J.Non-Newtonian Fluid Mech., Vol. 32, p. 295-310, 1989
[88] M. Fortin, M. Soulie : *A nonconforming piecewise quadratic finite element on triangles*, Int. J. Numer. Methods Engrg., Vol. 19, n. 4, p. 505-520,1983
[89] J.-F. Gerbeau, B. Perthame : *Derivation of Viscous Saint-Venant System for Laminar Shallow Water. Numerical Validation,* Discrete and Continuous Dynamical Systems, Ser. B, Vol. I, n. 1, p. 89-102, 2001
[90] H. Giesekus : *A simple constitutive equation for polymer fluids based on the concept of deformation-dependent tensorial mobility*, J.Non-Newtonian Fluid Mech., Vol. 11, p. 69-109, 1982
[91] V. Girault, P.A. Raviart : *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*, Springer Verlag, Berlin, 1986
[92] V. Girault, B. Rivière : *DG approximation of coupled Navier-Stokes and Darcy equations by Beaver-Joseph-Saffman interface condition*, SIAM J. Numer. Anal., Vol. 47, p. 2052-2089, 2009
[93] V. Girault, B. Rivière, M. Wheeler : *A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems*, Math. Comp., Vol. 74, n. 249, p. 53-84, 2005
[94] V. Girault, M. F. Wheeler : *Numerical Discretization of a Darcy-Forchheimer Model*, Numer. Mathematik, Vol. 110, n. 2, p. 161–198, 2008
[95] W. H. Graf : *Fluvial Hydraulics*, John Wiley & Sons, Chichester, 1998
[96] P. Grisvard : *Elliptic Problems in Non-Smooth Domains*, Pitman, Boston, 1985
[97] R. Guénette, M. Fortin : *A new mixed finite element method for computing viscoelastic flows*, J.Non-Newtonian Fluid Mech., Vol. 60, n.1, p. 27-52, 1995
[98] W. Hackbusch, T. Probst : *Downwind Gauss-Seidel Smoothing for Convection Dominated Problems*, Numerical Linear Algebra with Applications, Vol. 4, p. 85-102, 1997
[99] H. Han : *Nonconforming elements in the mixed finite element method*, J. Comput. Math., Vol. 2, p. 223–233, 1984
[100] P. Hansbo, M. Larson : *Discontinuous Galerkin and the Crouzeix-Raviart element: application to elasticity*, M2AN, Vol. 37, n. 1, p. 63-72, 2003
[101] P. Houston, E. Süli : *A note on the design of hp-adaptive finite element methods for elliptic partial differential equations*, Comput. Methods Appl. Mech. Engrg., Vol. 194, n. 2-5, p. 229-243, 2005
[102] D. Hu, T. Lelièvre : *New entropy estimates for the Oldroyd-B model and related models*, Commun. Math. Sci., Vol. 5, no. 4, p. 909-916, 2007
[103] M. Hulsen : *A sufficient condition for a positive definite configuration tensor in differential models*, J.Non-Newtonian Fluid Mech., Vol. 38, p. 93–100, 1990
[104] M. Hulsen, R. Fattal, R. Kupferman : *Flow of viscoelastic fluids past a cylinder at high Weissenberg number: Stabilized simulations using matrix logarithms*, J.Non-Newtonian Fluid Mech., Vol. 127, p. 27–39, 2005
[105] J.A. Jankowski : *A non-hydrostatic model for free surface flows*, Ph.D. Thesis, Hannover University, Germany, 1999
[106] C. Johnson, U. Nävert, J. Pitkäranta : *Finite element methods for linear hyperbolic problems*, Comput. Methods Appl. Mech. Engrg., Vol. 45, p. 285-312, 1984
[107] J. Joie : *Simulation numérique des écoulements de liquides polymères,* PhD Thesis, University of Pau, 2010
[108] O. Karakashian, F. Pascal : *A posteriori error estimates for a discontinuous Galerkin approximation of a second-order elliptic problem*, SIAM J. Numer. Anal., Vol. 41, n. 6, p. 2374-2399, 2003
[109] R. Keunings : *On the high Weissenberg number problem*, J.Non-Newtonian Fluid Mech., Vol. 20, p. 209-226, 1986
[110] K. Kim : *A posteriori error analysis for locally conservative mixed methods*, Math. Comp., Vol. 76, n. 257, p. 43-66, 2007
[111] M.B. Koçyigit, R.A. Falconer, B. Lin : *Three-dimensional numerical modeling of free surface flows with non-hydrostatic pressure*, Int. J. Numer. Meth. Fluids, Vol. 40, p. 1145-1162, 2002
[112] V.A. Kondratiev, O.A. Oleinik : *Boundary value problems for partial differential equations in non smooth domains*, Russian Mathematical Surveys, Vol. 38, n. 2, p. 1-86, 1983
[113] P. Lancaster, L. Rodman : *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995

[114] W. Layton, F. Schiewek, I. Yotov : *Coupling fluid flow with porous media flow*, SIAM J. Numer. Anal., Vol. 40, No. 6, p. 2195-2218, 2003
[115] J.W. Lee, M.D. Teubner, J.B. Nixon, P.M. Gill : *A 3D non-hydrostatic pressure model for small amplitude free surface flows*, Int. J. Numer. Meth. Fluids, Vol. 50, n. 6, p. 649-672, 2006
[116] Y. Lee, J. Xu : *New formulations, positivity preserving discretizations and stability analysis for non-Newtonian flow models*, Comput. Methods Appl. Mech. Engrg., Vol. 195, p. 1180-1206, 2006
[117] C. F. Leibovici, J. Neoschil : *A solution for Rachford-Rice equations for multiphase systems*, Fluid Phase Equilib., Vol. 112, p. 217-221, 1995
[118] V. Le Manchec : *Modélisation de la thermométrie dans un réservoir*, Document Total, 2004
[119] P. Lesaint, P.-A. Raviart : *On a finite element method for solving the neutron transport equation*, Mathematical Aspects of Finite Element Methods in Partial Differential Equations, C.A. de Boor (ed.), Academic Press, p. 89-123, 1974
[120] F. J. Lim, W.R. Schowalter : *Pseudo-spectral analysis of the stability of pressure-driven flow of a Giesekus fluid between parallel planes*, J.Non-Newtonian Fluid Mech., Vol. 26, p. 135-142, 1987
[121] J.L. Lions, E. Magenes : *Problèmes aux limites non homogènes et applications*, Ed. Dunod, Vol. 1, 1968
[122] L. Lizaik : *Modélisation numérique de modèles thermomécaniques polyphasiques puits-milieux poreux*, Thèse de doctorat, Université de Pau, 2008
[123] C. Lovadina, L.D. Marini : *A Posteriori Error Estimates for Discontinuous Galerkin Approximations of Second Order Elliptic Problems*, J. Sci. Comput., Vol. 40, p. 340-359, 2009
[124] A. Lozinski, R. Owens : *An energy estimate for the Oldroyd-B model: theory and applications*, J.Non-Newtonian Fluid Mech., Vol. 112, p. 161-176, 2003
[125] R. Luce, B. Wohlmuth : *A local a posteriori error estimator based on equilibrated fluxes*, SIAM J. Numer. Anal., Vol. 42, n. 4, p. 1394-1414, 2004
[126] J. M. Marchal, M. J. Crochet : *A new mixed finite element for calculating viscoelastic flow*, J.Non-Newtonian Fluid Mech., Vol. 26, p. 77-114, 1987
[127] F. Marche : *Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects*, Eur. J. Mech. B Fluidss, Vol. 26, n. 1, p. 49-63, 2007
[128] K.A. Mardal, X.C. Tai, R. Winther : *A robust finite element method for Darcy-Stokes flow*, SIAM J. Numer. Anal., Vol. 40, p. 1605–1631, 2002
[129] K.-A. Mardal, R. Winther : *An observation on Korn's inequality for nonconforming finite element methods*, Math. Comp., Vol. 75, n. 253, p. 1–6, 2005
[130] L.D. Marini : *An inexpensive method for the evaluation of the solution of the lowest order Raviart-Thomas mixed method*, SIAM J. Numer. Anal., Vol. 22, No. 3, p. 493-496, 1985
[131] F. Maubeuge, M. Didek, E. Arquis, O. Bertrand, J.-P. Caltagirone : *MoTher : A model for interpreting thermometrics*, SPE 28588, 1994
[132] V.L Mehrmann : *The autonomous linear quadratic control problem*, Lecture Notes in Control and Information Sciences, **163**, Springer Verlag, Berlin, 1991
[133] E. Miglio, S. Perotto, F. Saleri : *Model coupling techniques for free-surface flow problems : Part I*, Nonlinear Analysis, n. 63, p. 1885-1896, 2005
[134] E. Miglio, A. Quarteroni, F. Saleri : *Finite element approximation of quasi-3D shallow water equations*, Comput. Methods Appl. Mech. Engrg., Vol. 174, n. 3-4, p. 355-369, 1999
[135] F. Montel : *Petroleum thermodynamics*, Document Elf Aquitaine, 1994
[136] J.-C. Nédélec : *Schémas d'approximation pour des équations intégro-différentielles de Riccati*, Thèse d'Etat, Université Paris IV - Sorbonne, 1970
[137] R.A. Nicolaides : *Existence, uniqueness and approximation for generalized saddle point problems*, SIAM J. Numer. Anal., Vol. 19, No. 2, p. 349–357, 1982
[138] J. Nitsche : *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, Collection of articles dedicated to Lothar Collatz on his 60th birthday, Vol. 36, p. 9-15, 1971
[139] J.T. Oden, S. Prudhomme : *Estimation of modeling error in computational mechanics*, J. Comput. Phys., Vol. 182, p. 469-515, 2002
[140] J. G. Oldroyd : *On the Formulation of Rheological Equations of State*, Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences, Vol. 200, p. 523-541, 1950
[141] R.G. Owens, C. Chauvire, T.N. Philips : *A locally-upwinded spectral technique (LUST) for viscoelastic flows*, J.Non-Newtonian Fluid Mech., Vol. 49, p. 108, 2002
[142] R. G. Owens, T. N. Phillips : *Computational Rheology*, Imperial College Press, London, 2002
[143] C. Paige, C. Van Loan : *A Schur decomposition for Hamiltonian matrices*, Linear Algebra and its Applications, Vol. 41, p. 11-32, 1981

[144] C. Park, D. Sheen : *P1-Nonconforming Quadrilateral Finite Element Methods for Second-Order Elliptic Problems*, SIAM J. Numer. Anal., Vol. 41, n. 2, p. 624-640, 2003

[145] D.Y. Peng, D.B. Robinson : *A new two-constant equation of state*, Ind. Eng. Chem. Fundam., Vol. 15, p. 59-64, 1976

[146] S. Perotto : *Adaptive modeling for free-surface flows*, M2AN, Vol. 40, n. 3, p. 469-500, 2006

[147] A. Petrau : *Simulation numérique multidimensionnelle d'écoulements estuariens,* PhD Thesis, University of Pau, 2009

[148] N. Phan-Thien, R. Tanner : *A new constitutive equation derived from network theory*, J.Non-Newtonian Fluid Mech., Vol. 2, p. 353-365, 1977

[149] O. Pironneau : *Finite Element Methods for Fluids*, John Wiley & Sons and Masson, Paris, 1989

[150] J. Pousin, J. Rappaz: *Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problems*, Numer. Math., Vol. 69, p. 213-231, 1994

[151] A. Quarteroni, A. Veneziani : *Analysis of a geometrical multiscale model based on the coupling of PDE's and ODE's for Blood Flow Simulations*, SIAM J. on MMS, Vol. 1, no. 2, p. 173-195, 2003

[152] L. Quinzani, R. Armstrong, R. Brown : *Birefringence and laser-Doppler velocimetry (LDV) studies of viscoelastic flow through a planar contraction*, J.Non-Newtonian Fluid Mech., Vol. 52, n. 1, p. 1-36, 1994

[153] D. Rajagopalan, R. Armstrong, R. Brown : *Finite element methods for calculation of steady, viscoelastic flow using constitutive equations with a Newtonian viscosity*, J.Non-Newtonian Fluid Mech., Vol. 36, p. 159-192, 1990

[154] R. Rannacher, S. Turek : *Simple nonconforming quadrilateral Stokes element*, Numer. Methods Partial Differential Equations, Vol. 8, n. 2, p. 97-111, 1992

[155] P.-A. Raviart, J.-M. Thomas : *Primal hybrid finite element methods for 2nd order elliptic equations*, Math. of Comp., Vol. 31, n. 138, p. 391-413, 1977

[156] B. Rivière : *Analysis of a discontinuous finite element method for the coupled Stokes and Darcy problems*, Journal of Scientific Computing, Vol. 22, no. 1, p. 479–500, 2005

[157] B. Rivière, I. Yotov : *Locally conservative coupling of Stokes and Darcy flows*, SIAM J. Numer. Anal., Vol. 42, No. 5, p. 1959–1977, 2005

[158] J.E. Roberts, J.-M. Thomas: *Mixed and Hybrid Methods*, in Handbook of Numerical Analysis, vol. II : Finite Element Methods, P.G. Ciarlet & J.L. Lions, ed., North Holland, Amsterdam, p. 523-639,1991

[159] J. Sun, N. Phan-Thien, R.I. Tanner : *An adaptive viscoelastic stress splitting scheme and its applications : AVSS/SI and AVSS/SUPG*, J.Non-Newtonian Fluid Mech., Vol. 65, p. 75-91, 1996

[160] M. Suri, I. Babuška, C. Schwab : *Locking effects in the finite element approximation of plate models*, Math. Comp., Vol. 64, p. 461-482, 1995

[161] R. Temam : *Navier-Stokes Equations. Theory and Numerical Analysis*, North Holland, Amsterdam, 1979

[162] J.-M. Thomas : *Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes*, Thèse de doctorat d'état, Université de Paris VI, 1977

[163] S. Vanka : *Block-implicit multigrid calculation of two-dimensional recirculating flows*, Comput. Methods Appl. Mech. Engrg., Vol. 59, n. 1, p. 29 - 48, 1986

[164] R. Verfürth : *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley &Teubner, Chichester, 1996

[165] R. Verfürth, D. Braess : *A posteriori error estimator for the Raviart-Thomas element*, SIAM J. Numer. Anal., Vol. 33, No. 6, p. 2431-2444, 1996

[166] M. Vohralik : *Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods*, Numer. Math. , Vol. 111, n. 1, p. 121-158, 2008

[167] M. Vohralik : *Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods*, Math. Comp., Vol. 79, n. 272, p. 2001–2032, 2010

[168] K. Walters, M. Webster : *The distinctive CFD challenges of computational rheology*, Int. J. Numer. Meth. Fluids, Vol. 43, p. 577-596, 2003

[169] M.F. Wheeler : *An elliptic collocation finite element method with interior penalties*, SIAM J. Numer. Anal., Vol. 15, p. 152–161, 1978

[170] L. C. Young, R. E. Stephenson : *A generalized compositional approach for reservoir simulation*, SPE J., p. 727-742, 1983