Ecoulements et conditions aux limites particulières appliquées en hydrogéologie et théorie mathématique des processus de dissolution/précipitation en milieux poreux

Vincent Devigne

N^o d'ordre : 31 - 2006

Année 2005-2006

THESE

présentée

devant l'Université Lyon I - Claude Bernard

pour l'obtention

du DIPLOME DE DOCTORAT

(arrêté du 25 avril 2002)

spécialité : Mathématiques et Application des Mathématiques

présentée et soutenue publiquement le 30 mars 2006

par

Vincent DEVIGNE

TITRE : Ecoulements et Conditions aux Limites Particulières Appliquées en Hydrogéologie et Théorie Mathématique des Processus de Dissolution/Précipitation en Milieux Poreux - Flows and Particular Boundary Conditions applied in Hydrogeology and Mathematical Theory of Dissolution/Precipitation process in porous media

JURY :

Présidente/Rapporteur : Mme C. Rosier	Pr./LMPA-Université du littoral (Calais)
Rapporteur : M I.S. Pop	MCF/CASA-TUE (Eindhoven - Pays-Bas)
Rapporteur : M M. Kern	C.R./INRIA (Rocquencourt)
Co-Directeur : M T. Clopeau	MCF/UCBL-ICJ (Lyon)
Directeur : M D. Graillot	D.R./EMSE-SITE(Saint-Etienne)
Directeur : M A. Mikelić	Pr./UCBL-ICJ (Lyon)

♮ Centre SITE
Ecole Nationale Supérieure des Mines de Saint-Etienne
158, Cours Fauriel, 42027 St-Etienne cedex 02, FRANCE
e-mail : Vincent.Devigne@emse.fr

♭ Institut Camille Jordan
Université Claude Bernard Lyon I, site de Gerland
Bât A, bur. 1304
50 , Av. Tony Garnier, 69367 Lyon Cedex 07, FRANCE
e-mail : Vincent.Devigne@cdcsp.univ-lyon1.fr

Remerciements,

Cette thèse n'aurait pas pu être menée à bien sans l'orientation, l'aide et le soutien continu et inconditionnel de mes directeurs de thèse Andro Mikelić et Thierry Clopeau, je leur en suis profondément reconnaissant et puissent-ils trouver dans ces pages le témoignage de ma gratitude sans limite. Merci.

Ces trois ans ont passé vite, preuve du travail constant effectué de part et d'autre. La thèse est la conclusion d'une collaboration et donc d'un travail commun, j'espère que ce caractère transparaît dans ce manuscrit.

Merci à mes codirecteurs d'Eindhoven “Hans” van Duijn et “Sorin” Pop de m'avoir accueilli pendant ces 6 mois d'échange à TUE (Pays-Bas) pour leur gentillesse, leur “infinie” patience et leur fructueuse collaboration.

Merci au centre SITE de l'École Nationale Supérieure des Mines de Saint-Etienne de m'avoir hébergé et financé pendant ces trois ans, à mes directeurs Mireille et Didier ainsi qu'à Hervé et Igor dont j'ai partagé dans la bonne humeur le bureau. Merci à Christiane et Zahïa pour leur dévouement et leur gentillesse.

Merci au gdr MOMAS pour son soutien financier.

Merci à mes parents, à mes frère et soeur Jeff et Clémence et à mon Amie Aurélie qui m'ont supporté dans mon impatience, mes énervements et mon enthousiasme tout au long de ce projet.

Un merci tout spécial à la famille Lo, pour leur chaleureuse hospitalité et leur maison toujours ouverte pour l' ”étudiant sans logis”. A Ano, Ben, Pierre-François, Karim, et Marc pour leur amitié et leur complicité en musique. Et enfin à mes amis d'Europe rencontrés durant mon séjour à Eindhoven, Miguel, Rémo, Gert-Jan, Kamyar, Edi et Christina et bien sûr Dragan.

A vous tous merci.

Quelques citations,

Le commencement de toutes les sciences c'est l'étonnement de ce que les choses sont ce quelles sont.

Aristote.

La vérité scientifique est une proposition incontestable ou contestable énoncée en langage mathématiques. Une niaiserie est une proposition contestable ou incontestable énoncée en langage vulgaire.

Georges Elgozy.

La mathématique est une science dangereuse, elle dévoile les supercheries et les erreurs de calcul.

Galilée.

Ceux qui soutiennent que la science n'explique rien, l'on voudrait qu'ils nous expliquassent une bonne fois ce que serait pour eux que d'expliquer.

Jean Rostand.

et enfin,

La science ne cherche pas à énoncer des vérités éternelles ou des dogmes immuables ; loin de prétendre que chaque étape est définitive et qu'elle a dit son dernier mot, elle cherche à cerner la vérité par approximations successives.

Bertrand Russel.

Table des matières

1 Introduction **11**

I Écoulements et lois de paroi 15

2 Une approximation non conforme des équations de Darcy **17**

2.1 Hybrid and non-conforming methods applied to Darcy flow 18
2.2 The flow problem 20
2.2.1 Equation 20
2.2.2 The finite element spaces 21
2.3 Definitions 22
2.3.1 Interpolation operator and error estimate 22
2.3.2 Functional Framework 23
2.3.3 Discrete Green's formula 25
2.3.4 Discrete Poincaré inequality 26
2.4 Discrete problems 28
2.4.1 Pressure formulation 28
2.4.2 Velocity formulation 31
2.4.3 Dual problem 31
2.5 Numerical results 33
2.6 Conclusion 38
References 39

3 Écoulement tangentiel sur une surface rugueuse et loi de Navier **41**

3.1 Préambule 41
3.2 Introduction 43
3.3 Couche limite de Navier 45

3.4 Justification de la condition de glissement de Navier pour l'écoulement de Couette laminaire 3D 47
3.5 Conclusion et applications possibles en océanographie 52
References 55

II Ecoulements complexes, théorie mathématique des processus de dissolution/précipitation en milieux poreux 57

4 Rigorous upscaling of a reactive flow through a pore, under important Peclet's and Damkohler's numbers 61
4.1 Introduction 62
4.2 Study of the the upscaled diffusion-convection equation on the half-line 66
4.3 A simple L^2 error estimate 68
4.4 The formal 2-scale expansion leading to Taylor's dispersion 71
4.5 Boundary layer 73
4.6 First Correction 73
4.7 Error estimate involving the second order in expansion 83
References 89

5 L'écoulement de Stokes 91
5.1 Introduction 92
5.2 Problème continu, résultat de régularité et formulation mathématique 94
5.3 Approximation par élément fini du problème de Stokes 96
5.4 Approximation de type Petrov-Galerkin avec stabilisation bulle 97
5.4.1 condition inf-sup et estimations d'erreur 97
5.4.2 Fonctions de base 98
5.4.3 La bulle à résidu nulle (RFB) appliquée à l'équation de convection-diffusion 101
5.4.4 La fonction bulle 103
5.4.5 Formulation algébrique 106
5.5 Résultats Numériques 108
5.6 Conclusion 115
5.7 Introduction 116
5.8 Continuous problem, regularity result and mathematical formulation 117
5.9 Finite Element approximation of the Stokes problem 119

5.10 Petrov-Galerkin approximation of the problem with bubble stabilization 120
5.10.1 inf-sup condition and error estimates 120
5.10.2 Shape functions 122
5.10.3 Residual free Bubble applied to Advective-Diffusive equation 124
5.10.4 The bubble function 126
5.10.5 Algebraïc form 129
5.11 Results 131
5.12 Conclusion 133
References 135

6 Numerical analysis of flow, transport precipitation and dissolution in a porous medium 137
6.1 Introduction 138
6.2 The time discrete numerical scheme 142
6.2.1 Stability in L^∞ 143
6.2.2 A priori estimates 146
6.2.3 Convergence 149
6.3 A fixed point iteration for the time discrete problems 155
6.4 Numerical results 159
6.5 Conclusion 167
References 169

7 Questions ouvertes et travaux futurs - Open questions and future works 171

Chapitre 1

Introduction

En mécanique des fluides, la plupart des quantités étudiées sont subordonnées à la condition intrinsèque d'accessibilité à l'expérience, l'observation, le caractère mesurable et quantifiable. Ces grandeurs physiques sont par définition macroscopiques, la pression, la vitesse, la température, la concentration, la viscosité, la perméabilité ou la diffusion. Comment lier les caractéristiques macroscopiques et microscopiques de ces quantités ?

L'analyse des propriétés macroscopiques des matériaux composites a mobilisé les physiciens comme Rayleigh, Maxwell, Taylor. Dans les années 1970, la modélisation du problème physique des structures composites a élargi l'étude à l'analyse purement mathématique sous la forme de la théorie de l'homogénéisation. Nous citerons comme référence le livre "Asymptotic Analysis for Periodic Structures" de Lions, Bensoussan, Papanicolaou, North Holland, 1978. Les phénomènes naturels sont caractérisés par de multiples échelles, en espace et en temps. La complexité mathématique de leur description augmente avec le nombre d'échelles. Il apparaît que sous certaines hypothèses d'ergodicité, de périodicité et d'homogénéité du milieu au niveau microscopique, le niveau mésoscopique ou macroscopique puissent être décrit avec une bonne approximation par des équations plus simples que celles décrivant le comportement microscopique mais plus riches car les particularités microscopiques agissent au travers de leurs caractéristiques moyennées.

Les limites de validité de la théorie de l'homogénéisation sont contenues dans le développement asymptotique induit qui n'est exact que lorsque le rapport des échelles tend vers l'infini. La compréhension des conditions aux limites et leur modélisation reste une étape clef dans l'étude des phénomènes naturels. Au même titre qu'une analyse mathématique rigoureuse du problème dans sa forme continue est indispensable, une analyse numérique et une expérience solide des méthodes numériques est requise pour la compréhension des résultats aussi bien que pour leur mise en oeuvre.

La discussion s'articule autour de 6 chapitres que l'on peut regrouper autour de deux thématiques. Dans la première, les questions d'ordre numérique des problèmes de Darcy, Stokes et de convection-diffusion seront abordées. Ce sont les chapitres 2, 5 et 6. Dans la deuxième, nous étudierons la modélisation, les développements asymptotiques et l'homogénéisation appliquées aux lois de paroi, aux processus chimiques autrement dit à l'hydrogéologie et à la chimie dans les chapitres 3 et 4.

In fluid mechanics, most of the quantities are conditioned by their abilities to be measurable, evaluated, reached by experiments and observation. Those physical quantities are macroscopic, we mean the pressure, the velocity, the temperature, the concentration, the viscosity, the permeability and the diffusion. How is it possible to connect those macroscopic quantities to their microscopic behavior ?

The analysis of macroscopic properties concerning composite materials interested physicists like Rayleigh, Maxwell, and Taylor. In the 1970's, the modeling of the physical problem related to composite structure enlarged the horizons of pure mathematics analysis and provide an application field to the homogenization theory. We quote the reference book, "Asymptotic analysis for periodic structures" by Lions, Bensoussan and Papanicolaou, North Holland, 1978. Natural phenomena are dimensioned by several scales in space and time. The mathematical complexity of their modeling growth with the number of unknowns involved. It seems that the mesoscopic or macroscopic level may be reached with a good approximation by simplified equations compared to the microscopic description but enriched significantly, from the macroscopic point of view, with the averaged of microscopic behavior. It can be done under assumptions of ergodicity, periodicity, and homogeneity of the medium at the microscopic level.

The limit of homogenization theory (i.e. when its application doesn't hold anymore), is self contained in the asymptotic development. It is rigorously right when the ratio between scales tends to infinity. The understanding of the boundary conditions and their modeling is an important stage in the study of natural phenomenon. With the same meaning, a rigorous mathematical analysis of the continuous problem as well as a numerical analysis and experience with numerical simulations are required to understand the theoretical results and their use.

The discussion is grounded on six sections that can be brought together in two fields. In the first one, questions of numerical order for Darcy, Stokes and Advection-diffusion problems are treated. These are chapters 2, 5 and 6. In the second one, modeling, asymptotic expansion and homogenization theory will be applied to wall laws and chemical process in other word to hydro-geology and chemistry. These are chapters 3 and 4.

This thesis was the object of the following reports and articles :

Cette thèse a fait l'objet des rapports et articles suivant :

- *A. Mikelić, V. Devigne* "Tangential flows on rough surface and Navier's law", Annales Mathématiques Blaise Pascal vol.9, 313-327, (2002).
- *V. Devigne* "Simulating Groundwater and Verticals flow using a non-conforming approximation of the Darcy's and Stokes problem", Proceedings de la 4ème conférence internationale sur la méthode des éléments Analytiques des 20-21 Novembre 2003, à Saint-Etienne, ed. Ecole nationale Supérieure des Mines de Saint-Etienne.
- *V. Devigne, C.J. van Duijn, I.S. Pop, T. Clopeau* "Numerical analysis of flow, transport and chemical processes in a porous medium", preprint (2005).
- *V. Devigne, T. Clopeau* "Flow, transport and crystal dissolution in a porous medium" procceding de la conférence Fourth Conference on Applied Mathematics and Scientific Computing, 19-24 Juin, 2005 île de Brijuni, Croatie.
- *V. Devigne, I.S. Pop* "A numerical scheme for the micro scale dissolution and precipitation in porous media" proceeding de la conférence ENUMATH 2005, July 18-22 2005, Santiago de Compostela (Spain).
- *A. Mikelić, V. Devigne, C.J. van Duijn* "Rigorous upscaling of a reactive flow through a pore, under important Peclet's and Damkholer's number", CASA report no. 19, May 2005,
  `http ://www.win.tue.nl/casa/research/casareports/2005.html`,
  soumis pour publication.
- *V. Devigne, T. Clopeau* "Numerical analysis of a non-conforming approximation of the Darcy's equation", preprint (2005).
- *C. Rosier, V. Devigne, A. Mikelić* "Rigorous upscaling of the reactive flow through a pore, under dominant Peclet number and infinite adsorption rate" proceeding de la conférence Fourth Conference on Applied Mathematics and Scientific Computing, June 19-24, 2005 Brijuni island, Croatia.

and the following communications and participations to conference,

et les communications et participations aux conférences suivantes,

- 2002 "Applied non-linear problems : Flows with free boundaries" compte rendu des cours donnés par CEA-EDF-INRIA 7-10th October 2002, Rocquencourt, présentation donnée le 11 Octobre (2002), à l'Ecole Nationale Supérieur des Mines de Saint-Etienne.

– 2003 ”Analytic free divergence solution for Darcy’s flow” workshop on analytic element method 3,4 avril 2003, Ecole Nationale Supérieur des Mines de Saint-Etienne.
– 2004 ”Numerical analysis of flow, transport and chemistry around grains” mai 2004 Technische Universiteit Eindhoven.
– 2004 ”Crystal dissolution and precipitation in porous media : pore scale analysis” 14 Octobere 2004, Ecole Nationale Supérieur des Mines de Saint-Etienne.
– 2005 ”La dispersion de Taylor dans la modélisation du transport réactif par des méthodes asymptotiques - Simulations numériques” 3 mars 2005, Ecole Nationale Supérieur des Mines de Saint-Etienne.
– 2005 ”Taylor’s dispersion for reactive transport model using asymptotic methods-Numerical simulations” 7 avril 2005, Ecole Nationale Supérieur des Mines de Saint-Etienne.
– 2005 2ème Congrès National de Mathématiques Appliquées et Industrielles SMAI 2005 Evian 23-27 mai 2005, Talk : ”Homogénéisation d’un écoulement réactif à travers un pore, en présence de grands nombres de Peclet et Damkholer”, Workshop ”Problème réactif en milieu poreux” organisé par A. Michel (institut français du pétrole).
– 2005 SIAM Conference on Mathematical & Computational Issues in the Geosciences, June 7-10, 2005, Palais des Papes, the International Conference Center, Avignon, FRANCE, Co-author for two talks, ”Modeling reactive flows in Porous Media under important Peclet’s and Damkholer’s numbers, using Homogenization” avec A. Mikelić (UCBLI), et ”Dissolution and Precipitation in Porous media” avec I.Pop et C.J. van Duijn (TU Eindhoven).
– 2005 Fourth Conference on Applied Mathematics and Scientific Computing, June 19-24, 2005 Brijuni island, Croatia Présentation : ”Taylor’s dispersion for reactive model using asymptotic methods” co-auteur de deux présentations, ”Numerical Analysis of Flow, transport and Chemical Processes in a Porous medium” avec T.Clopeau (UCBLI) ”Method of Homogenization Applied to Dispersion, Convection and Reaction in Porous Media” par A.Mikelić.
– 2005 “Reactive Flow and transport Through Complex Systems” organized by C.J. van Duijn (Eindhoven), A. Mikelić (Villeurbanne) and C. Schwab (Zurich), du 30 Octobre au 5 novembre, Mathematisches Forschungsinstitut Oberwolfach (Allemagne).

Première partie

Écoulements et lois de paroi

Chapitre 2

Une approximation non conforme des équations de Darcy

Sommaire

2.1 Hybrid and non-conforming methods applied to Darcy flow . . 18

2.2 The flow problem 20

2.2.1 Equation 20

2.2.2 The finite element spaces 21

2.3 Definitions 22

2.3.1 Interpolation operator and error estimate 22

2.3.2 Functional Framework 23

2.3.3 Discrete Green's formula 25

2.3.4 Discrete Poincaré inequality 26

2.4 Discrete problems 28

2.4.1 Pressure formulation 28

2.4.2 Velocity formulation 31

2.4.3 Dual problem 31

2.5 Numerical results 33

2.6 Conclusion 38

Cette section a fait l'objet d'un proceeding dont les références sont les suivantes :

V. Devigne "Simulating Groundwater and Vorticals flow using a non-conforming approximation of the Darcy's and Stokes problem", Proceedings de la 4ème conférence internationale sur la méthode des Éléments Analytiques des 20-21 Novembre 2003 à Saint-Etienne, ed. Ecole Nationale Supérieure des Mines de Saint-Etienne, FRANCE, pp 47-59 (2003).

A l'état embryonnaire lors de l'impression du proceeding, remanié par la suite le titre final sera le suivant : "Numerical analysis of a non-conforming approximation of the Darcy's equation"

résumé-abstract : Nous proposons une étude des équations de Darcy avec une approximation non conforme du problème. Nous introduisons rigoureusement les espaces d'éléments finis. Cette formulation permet de découpler naturellement la vitesse et la pression. Nous donnons une estimation de l'erreur d'interpolation et montrons l'existence et l'unicité des solutions des problèmes discrets. Pour conclure, nous présenterons des résultats de simulation.

We study herein the Darcy's equation. We propose an original non-conforming approximation of the problem, introducing rigorous finite element spaces. This formulation gives a natural velocity-pressure uncoupling. We give an error estimate for the interpolation operator and show existence and uniqueness of solutions for the discrete problems. To conclude, we will illustrate with computational results.

2.1 Hybrid and non-conforming methods applied to Darcy flow

Let be Ω a bounded open subset of $\mathbb{R}^n$, with a continuous Lipschitzian boundary Γ. We consider the second order elliptic problem .

$$\begin{cases} -\triangle u = f \quad \text{in } \Omega \\ u = 0 \quad \text{on } \Gamma \end{cases} \tag{1.1}$$

where f is a given function belonging to the space $L^2(\Omega)$. The usual variational formulation of problem (1.1) can be expressed as follows, find $u \in H_0^1(\Omega)$ that minimize the energetic functional :

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega f v dx \quad \text{over all the space } H_0^1(\Omega). \tag{1.2}$$

Standard finite element methods to solve numerically the problem (1.1) are based on the following variational principle :

First of all, we must build a subspace of finite dimension V_h from the space $H_0^1(\Omega)$ made of regular functions continuous across each common boundary between two elements and that minimize the energetic functional $J(v)$ over the space V_h.

This type of conforming method has been extensively studied and convergence results are now classical.

See Ciarlet [6], Strang and Fix [8].

From the other side, we may weaken the continuity condition required across the elements for the functions of the space V_h and still obtain a finite element method that is converging. This is the so-called non-conforming method for which the space V_h is not necessarily contained in $H_0^1(\Omega)$.

For a precise analysis of particular non-conforming methods to solve second order elliptic problem, we refer the reader to Crouzeix et Raviart [9], Irons et Razzaque [10], Lesaint [12], Strang [11].

The first foundations of the theory made upon hybrid Finite Element Methods may be found in an article published in 1977 by P.A. Raviart and J.M. Thomas [2]. We may also quote the reference book by F. Brezzi and M. Fortin [4] where $ 2.2 is devoted to non-conforming method and $3 to the approximation of $H(\text{div}, \Omega)$ the continuous space that will be of interest for the forthcoming discussion.

Here we will define the space of solution for the Darcy's equations as subspace of $H(\text{div}, \Omega)$ and propose a numerical analysis of non-conforming finite element method restricted to regular geometry, i.e. hexaedra for a quasi-uniform decomposition of a convex domain $\Omega \subset \mathbb{R}^3$. We give exact definition in the form of polynoms for the interpolation spaces and we give an analysis of the method essentially built on a property of both the Darcy's equation itself and the space $H(\text{div}, \Omega)$. The discontinuous jump of the velocity between each element vanishes with an appropriate interpolation space definition.

We know that we could weaken the requirement of continuity for the functions of the approximation space and still obtain a convergent finite element method. We will use a detailed approach proposed in [2] based on the primal hybrid principle in which the constraint of inter-element continuity has been removed at the expense of introducing a Lagrange multiplier, the jump of velocity in the inter-element.

In [4], non-conforming approximations are seen as external approximation and this adds a term in the error estimate called the consistency term. It measures how the exact solution satisfies the discrete equation.

More recently, and more specifically, Finite Volume Element applied to Darcy's equation was discussed in a note by Y. Achdou and C. Bernardi [5]. The permeability is not constant. The total flux of the velocity between elements is equal to zero. But they consider a decomposition of the domain $\Omega \subset \mathbb{R}^2$ into triangles. They gave also an a posteriori error estimate for both unknowns, and recover the consistency term for Finite Volume Element equivalent to the one described in [4]. For more details, we quote [4] by the same authors where the equivalence between two Finite Volume Schemes and two Finite Element Methods is shown and the proof of the a posteriori error estimate is given.

Instead of a primal single field formulation for the pressure, A. Masud and al. considered in [3] a mixed two field formulation in which pressure and velocity are variables. Finite Element Approximation of $H(\text{div}, \Omega)$ is developed and satisfies the Babuska-Brezzi inf-sup condition where $\Omega \subset \mathbb{R}^2$ is decomposed into triangles. A mixed formulation with continuous velocity and discontinuous pressure is detailed together with many numerical examples. The convergence rate is studied.

We present a numerical approach of the Darcy's problem based on non-conforming finite element spaces.

This article is organized as follows. First, we introduce the flow problem, the finite element spaces and give its main properties. In the second part we study the interpolation operator and recall some known results for the continuous case but in their discrete forms, Green's formula and Poincaré 's inequality due to the fact that the space of approximation does not belong to the continuous one. We proceed differently from the classical study of the simplex. In the third part, after writing a pressure formulation of the discrete problem as a direct consequence of the equation of continuity for incompressible fluid, we give an a priori error estimate for the solution. A velocity formulation via a projection operator is proposed. We obtain error estimate for the velocity as a solution of Darcy's dual problem.

To finish with the analysis, we give some numerical results in the form of convergence rate and error curve drawing.

2.2 The flow problem

2.2.1 Equation

We consider the Darcy's equation in a regular domain $\Omega \subset \mathbb{R}^3$, with a Lipschitz-continuous boundary $\partial\Omega$. Let be $\Gamma_0 \cup \Gamma_I = \partial\Omega$ a partition of it.

The unknowns are the pressure p and the velocity v, f represents a density of body forces (for instance, the gravity g).

We have,

$$\left\{\begin{array}{rcl} u + A\nabla p & = & f \\ \mathrm{div}(u) & = & 0 \quad \text{in } \Omega \\ u.n & = & 0 \quad \text{on } \Gamma_N \\ p & = & 0 \quad \text{on } \Gamma_0 \end{array}\right. \tag{2.3}$$

where n is the outward normal unit vector of Ω and A is the symmetric positive defined tensor of permeability.

Remark 1 *We note that there is no boundary conditions imposed for the velocity and the pressure respectively on Γ_0 and Γ_N. One can add homogeneous boundary conditions without loss of generality.*

2.2.2 The finite element spaces

We assume Ω to be polyhedrical.

Let $\mathcal{K}$ be a quasi-uniform decomposition of Ω into hexaedras (or bricks). We set $\mathcal{N}_I$ the internal faces, $\mathcal{N}_{\Gamma_0}$ the faces belonging to Γ_0, K an hexaedra of $\mathcal{K}$, $\{e_i\}_{i=1}^{6}$ the faces of K.

Let us define

$$\begin{aligned} \Phi_1 = x^2 - y^2,\ \Phi_2 = y^2 - z^2 \\ \forall K \in \mathcal{K},\ Q(K) = Q = P_1 \oplus span(\Phi_1, \Phi_2) \\ \forall K \in \mathcal{K},\ RT_0(K) = \{v \in RT_1(K) \mid \mathrm{div}(v) = 0\} \end{aligned}$$

.

where $RT_1(K)$ denotes the Raviart-Thomas space defined on K.

Let be $\Sigma = \{q \in Q \to \int_{e_i} q \, d\sigma\}_{i=1}^{6}$ a finite set of linearly independent linear forms.

The triple (K, Q, Σ) defines a finite element in $\mathbb{R}^3$ after [7].

Remark 2 *The choice of Φ_1 and Φ_2 is invariant, it is equal to say that we could choose $x^2 - z^2$ and $z^2 - y^2$ instead of $x^2 - y^2$ and $y^2 - z^2$ generating the same vectorial subspace.*

We need to ensure that Φ_1 and Φ_2 are divergence free, this is shown by the following argument,

$Q = \mathbb{P}_1(K) + span(\Phi_1, \Phi_2)$

$\{\Phi_1, \Phi_2\}$ is free, $\triangle\Phi_i = 0$ for $i = 1, 2$, $\partial\Phi_i/\partial n = C$ where n denotes the outward normal unit vector and C is a constant.

for $p \in Q, \quad \int_{e_i} p \, d\sigma = 0$

$$\int_K \nabla p.\nabla p \, dx + \int_K \triangle p.p \, dx = \int_{\partial K} \frac{\partial p}{\partial n} p \, d\sigma = 0, \qquad (\triangle p = 0) \text{ so } \nabla p = 0,$$

consequently $p = C \Rightarrow p = 0.$

Remark 3 *Note that Q does not contain $Q_1(K)$, but just $P_1(K)$, dim $\Sigma = 6$.*

Let be,

$$Q_h = \Big\{ q_h \in L^2(\Omega), \quad \forall K \in \mathcal{K}, q_h|_K \in Q, \quad \forall e \in \mathcal{N}_I, \int_e [q_h] \, d\sigma = 0$$
$$\forall e \in \mathcal{N}_{\Gamma_0}, \int_e q_h \, d\sigma = 0 \Big\},$$

$$V_{h,-1} = \left\{ v_h \in L^2(\Omega)^3 | \, \forall K \in \mathcal{K}, \; v_h|_K \in RT_0(K) \right\},$$
$$V_h = V_{h,-1} \cap H^0_{\Gamma_I}(\text{div}, \Omega)$$
$$H^0_{\Gamma_I}(\text{div}, \Omega) = \{ v | v \in L^2(\Omega)^3, \quad \text{div}(v) \in L^2(\Omega)^3, \quad v = 0 \text{ on } \Gamma_0, \quad v.n = 0 \text{ on } \Gamma_I \}.$$

2.3 Definitions

In what follows, we underline an important assumption. The analysis holds in the case of non-uniform meshes. There is no reference element.

2.3.1 Interpolation operator and error estimate

After introducing the projection operator $\mathcal{I}_{h,K}$ from $H^1(K)$ into Q, and its characterization in the form of three lemmas we will give the interpolation error as a theorem.

Let be,

$$\mathcal{I}_{h,K} \; : \; \begin{array}{ccc} H^1(K) & \longrightarrow & Q \\ v & \longmapsto & \mathcal{I}_h v \end{array}$$

the interpolation operator defined as the unique element of Q such as $\int_e (v - \mathcal{I}_{h,K} v) \, d\sigma = 0$, $\forall e$ face of K.

The interpolation error in flux is orthogonal to Q,

Lemma 2.3.1 *For all $v \in H^1(K)$, we have :*

$$\forall q_h \in Q, \qquad \int_K \nabla q_h . \nabla (v - \mathcal{I}_{h,K} v)\, dx = 0.$$

Proof. After integrating by parts and applying definition of both Q and $P_0(e)$ we immediately see that,

$$\int_K \nabla q_h . \nabla (v - \mathcal{I}_{h,K} v)\, dx \quad = \quad 0,$$

■

The interpolation is optimal in Q. This is shown in the lemma below,

Lemma 2.3.2 *For all $v \in H^1(K)$, we have*

$$|v - \mathcal{I}_{h,K} v|_{1,K} = \inf_{q_h \in Q} |v - q_h|_{1,K}.$$

Proof.

$$\begin{aligned} |v - \mathcal{I}_{h,K} v|^2_{1,K} &= \int_K \nabla (v - \mathcal{I}_{h,K} v) \nabla (v - \mathcal{I}_{h,K} v) \\ &= \int_K \nabla (v - q_h) \nabla (v - \mathcal{I}_{h,K} v) + \int_K \nabla (q_h - \mathcal{I}_{h,K} v) \nabla (v - \mathcal{I}_{h,K} v). \end{aligned}$$

We immediately notice that $q_h - \mathcal{I}_{h,K} v \in Q$ implying that the second term on the right is equal to zero. Then the Cauchy-Schwartz inequality gives the result,

$$\begin{aligned} |v - \mathcal{I}_{h,K} v|^2_{1,K} &\leq |v - q_h|_{1,K} \, |v - \mathcal{I}_{h,K}|_{1,K}\,, \\ \Rightarrow |v - \mathcal{I}_{h,K} v|_{1,K} &\leq |v - q_h|_{1,K} \qquad \forall q_h \in Q, \\ \Leftrightarrow |v - \mathcal{I}_{h,K} v|_{1,K} &\leq \inf_{q_h \in Q} |v - q_h|_{1,K}\,. \end{aligned}$$

We just mention that $\mathcal{I}_{h,K} v \in Q$ gives the equality. ■

2.3.2 Functional Framework

Let Ω be a convex polygonal domain of $\mathbb{R}^3$ and Γ its boundary. We denote by $W^{s,p}(\Omega)$ and $W^{s,p}(\Gamma)$, $0 \leq s$ and $1 \leq p \leq +\infty$, the usual Sobolev space endowed with the norms $||.||_{s,p,\Omega}$ and $|.|_{s,p,\Gamma}$ where

$$\text{if p} < \infty, \qquad |v|_{s,p,\Omega} = \int \int_{\Omega \times \Omega} \frac{||D^{[s]} v(x) - D^{[s]} v(y)||^p}{|x - y|^{3 + p\sigma}}\, dxdy$$

and

$$|v|_{s,p,\Gamma} = \int\int_{\Gamma\times\Gamma} \frac{||D^{[s]}v(x) - D^{[s]}v(y)||^p}{|x-y|^{2+p\sigma}}\, dxdy$$

$$\text{if p} = \infty, \qquad |V|_{s,+\infty,\Omega} = \sup_{\Omega\times\Omega} \frac{||D^{[s]}v(x) - D^{[s]}v(y)||}{|x-y|^{\sigma}}$$

$[s]$ being the integer part of s and $\sigma = s - [s]$

$H^s(\Omega)$ is the usual space $W^{s,2}$ and $H_0^s(\Omega)$ the closure of $\mathcal{D}(\Omega)$ in $H^s(\Omega)$. For $s < 0$, $H^s(\Omega)$ is the topological dual space of $H_0^{-s}(\Omega)$ and $W^{s,p}(\Gamma)$ the topological dual space of $W^{-s,\frac{p}{(p-1)}}(\Gamma)$.

Lemma 2.3.3 *If $v \in H^{1+\sigma}(K)$, $0 < \sigma \leq 1$ we have :*

$$|v - \mathcal{I}_{h,K}v|_{1,K} \leq Ch_k^{\sigma}\, |v|_{1+\sigma,K}$$

Proof.

$$|v - \mathcal{I}_{h,K}v|_{1,K} = \inf_{q_h\in Q} |v - q_h|_{1,K} \leq \inf_{p\in P_1} |v-p|_{1,K} \qquad (P_1 \subset Q)$$

$$\text{We note : } |v|^2_{1+\sigma} = \int\int_{K\times K} \frac{|v(x)-v(y)|^2}{|x-y|^{3+2\sigma}}\, dx\, dy$$

$$\forall v \in H^1(K),\ \exists\, p \in P_1,\ \int_K \partial^{\alpha} v = \int_K \partial^{\alpha} p, \qquad |\alpha| \leq 1$$

$$\begin{aligned}
\text{let be } u &= \partial^{\alpha} v, \qquad \alpha = (1,0),\ (0,1) \text{ and } \bar{u} \text{ the mean of } u,\\
u - \bar{u} &= \frac{1}{|K|}\int_K (u(x) - u(y))\ dy,\\
&\leq \frac{1}{|K|^{1/2}}\left(\int_K |u(x)-u(y)|^2\right)^{1/2},\\
|u-\bar{u}|^2_{L^2(K)} &\leq \frac{1}{|K|}\int\int_{K\times K} \frac{|u(x)-u(y)|^2}{|x-y|^{3+2\sigma}}\, dx\, dy\, |x-y|^{3+2\sigma},\\
&\leq \frac{h^{1+2\sigma}}{|K|}|u|^2_{\sigma,K} \leq Ch^{2\sigma}|u|^2_{\sigma,K} \qquad (\frac{h}{|K|} \leq C),\\
\Leftrightarrow |\partial^{\alpha}v - \overline{\partial^{\alpha}v}|_{L^2(K)} &\leq Ch^{\sigma}|\partial^{\alpha}v|_{\sigma,K},\\
\Leftrightarrow |v - \bar{v}|_{1,K} &\leq Ch^{\sigma}|v|_{1+\sigma,K}.
\end{aligned}$$

$$
\begin{aligned}
\exists p_0 \in P_1,\ \bar{v} &= \frac{1}{|K|}\int_K v = \frac{1}{|K|}\int_K p_0,\\
\Rightarrow |v-p_0|_{1,K} &\le Ch^{\sigma}\,|v|_{1+\sigma,K},\\
\Rightarrow \inf_{p\in P_1}|v-p|_{1,K} &\le Ch^{\sigma}\,|v|_{1+\sigma,K},\\
\Rightarrow |v-\mathcal{I}_{h,K}v|_{1,K} &\le Ch^{\sigma}\,|v|_{1+\sigma,K}.
\end{aligned}
$$

■

And the a priori error estimate for the interpolation operator follows,

Theorem 2.3.4 *If* $v \in H^{1+\sigma}(\Omega)$, $0<\sigma\le 1$, $\exists\,\mathcal{I}_h : H^1(K) \longrightarrow Q_h$,

$$
\left(\sum_{K\in\mathcal{K}_h}|v-\mathcal{I}_hv|^2_{1,K}\right)^{1/2} \le Ch^{\sigma}\,|v|_{1+\sigma,\Omega}.
$$

Proof. Let's take $\mathcal{I}_h = \mathcal{I}_{h,K}$,

$$
\begin{aligned}
|v-\mathcal{I}_{h,K}|_{1,K} &\le Ch_K^{\sigma}\,|v|_{1+\sigma,K},\\
\sum_K|v-\mathcal{I}_{h,K}|^2_{1,K} &\le C_2\sum_K h_K^{2\sigma}|v|^2_{1+\sigma,K},\\
\text{let be } h = \max_K(h_K),
\end{aligned}
$$

$$
\begin{aligned}
\sum_K|v-\mathcal{I}_{h,K}|^2_{1,K} &\le C_2h_K^{2\sigma}\sum_K|v|^2_{1+\sigma,K},\\
\sum_K|v-\mathcal{I}_{h,K}|^2_{1,K} &\le C_2h_K^{2\sigma}|v|^2_{1+\sigma,\Omega},\\
\left(\sum_K|v-\mathcal{I}_{h,K}|^2_{1,K}\right)^{1/2} &\le C_3h^{\sigma}|v|_{1+\sigma,\Omega}.
\end{aligned}
$$

■

In other words, the ratio between the interpolation error for $v\in H^{1+\sigma}(\Omega)$ in the suitable discrete norm with respect to the decomposition and the continuous norm of $H^{1+\sigma}(\Omega)$ behaves like $\sim o(h^{\sigma})$. The next part is the study of the pressure formulation, and more specifically Q_h that is not contained in $H^1(\Omega)$. Its analysis requires the discrete Green's formula and the discrete Poincaré inequality . This is the subject of sections 2.3.3 and 2.3.4.

2.3.3 Discrete Green's formula

The discrete Green's formula is given by the following theorem,

Theorem 2.3.5 *For all $q_h \in Q_h$ and $v_h \in V_h$*

$$\sum_{K\in\mathcal{K}} \int_K (\nabla q_h v_h + q_h div\ v_h)\ dx \quad = \quad \int_{\Gamma_N} v_h.n\ q_h\ d\sigma.$$

Proof.

$$\begin{aligned}\sum_{K\in\mathcal{K}} \int_K (\nabla q_h v_h + q_h \text{div}\ v_h)\ dx \quad &= \quad \sum_{K\in\mathcal{K}} \int_{\partial K} v_h.n_K\ q_h\ d\sigma, \\ &= \quad \sum_{e\in\mathcal{N}_I} \int_e v_h.n\ [q_h]\ d\sigma + \int_\Gamma v_h.n\ q_h\ d\sigma.\end{aligned}$$

We notice that $v_h.n|_e \in P_0(e)$ and mention that the mean of q_h across each element is zero giving the result.

■

2.3.4 Discrete Poincaré inequality

Given,

$$P_h = \left\{q_h \in \mathcal{C}^0(\Omega) \cap H^1_{0,\Gamma_0}(\Omega),\ \forall K \in \mathcal{K}, q_h|_K \in Q_1(K)\right\} \subset H^1_{0,\Gamma_0}(\Omega)$$
$$\text{with } H^1_{0,\Gamma_0}(\Omega) = \left\{v \in H^1(\Omega) \mid v|_{\Gamma_0} = 0\right\}$$

Given a an edge of K, $K \in \mathcal{K}$, we define

$$\triangle_a = \left\{K \in \mathcal{K}\ /\ a \in \bar{K}\right\} \text{ and } \sharp\triangle_a \quad = \quad \text{card}\triangle_a$$

We consider,

$$\begin{aligned} R_h : Q_h \quad &\longrightarrow \quad P_h \\ q_h \quad &\longrightarrow \quad R_h q_h, \end{aligned}$$

the projection from Q_h to P_h such as $\forall a \in \overline{\Gamma_0},\ R_h q_h(a) = 0,$

$$\text{if } a \notin \overline{\Gamma_0},\ R_h q_h(a) = \frac{1}{\sharp\triangle_a} \sum_{K\in\triangle_a} q_h|_K(a).$$

It is easy to show (see [14]) that,

Theorem 2.3.6

$$\forall l \in \{0,1\},\ \forall q_h \in P_h,$$
$$\forall K \in \mathcal{K}, \quad |q_h - R_h q_h|_{l,K} \leq C h^{1-l} \sum_{\substack{\bar{K} \in \mathcal{K} \\ K \cap \bar{K} \neq 0}} |q - q_h|_{1,\bar{K}} \quad \forall q \in H^1_{0,\Gamma_0}(\Omega).$$

Remark 4 *For the particular case $l = 0$ and $q = 0$,*

$$|R_h q_h - q_h|_{0,\Omega} \leq Ch \left(\sum_{K \in \mathcal{K}} |q_h|^2_{1,K} \right)^{1/2}.$$

As a consequence we obtain the Poincaré's inequality in the form of a lemma,

Lemma 2.3.7

$\exists\, C$ *independent of h such as*
$$\forall q_h \in Q_h, \quad |q_h|_{0,\Omega} = \|q_h\|_{0,\Omega} \leq C \left(\sum_{K \in \mathcal{K}} |q_h|^2_{1,K} \right)^{1/2} = C\,|q_h|_{1,h}.$$

Proof.

$$\|q_h\|_{0,\Omega} \leq \|R_h q_h\|_{0,\Omega} + \|q_h - R_h q_h\|_{0,\Omega}.$$

We apply Poincaré 's inequality to obtain,

$$\begin{aligned} \|q_h\|_{0,\Omega} &\leq |R_h q_h|_{1,h} + Ch\,|q_h|_{1,\Omega}, \\ &\leq C_1 |R_h q_h - q_h|_{1,\Omega} + C_2(1+h)\,|q_h|_{1,h}. \end{aligned}$$

Applying theorem 2.3.6 with $l = q = 0$ gives,

$$\begin{aligned} \|q_h\|_{0,\Omega} &\leq C_3 |q_h|_{1,\Omega} + C_2(1+h)\,|q_h|_{1,h}, \\ \|q_h\|_{0,h} &\leq C_5\,|q_h|_{1,h}. \end{aligned}$$

■

2.4 Discrete problems

2.4.1 Pressure formulation

We consider the problem

$$\left\{\begin{array}{c} \text{Find } p_h \in Q_h \text{ such as} \\ \forall q_h \in Q_h, \quad \sum_{K\in\mathcal{K}} \int_K A\nabla p_h.\nabla q_h = \sum_{K\in\mathcal{K}} \int_K f.\nabla q_h \end{array}\right. . \tag{4.4}$$

(4.4) known as the Neumann Problem is treated in [5] and errors are compared numerically between two methods. (4.4) is approximated by the mixed Raviart-Thomas (RT) Finite Element and the mixed Brezzi-Douglas-Marini (BDM) Finite Element. These will be taken as references for the numerical part. After [5], the convergence is linear for the (RT) elements.

Existence and uniqueness of the solution

Theorem 2.4.1 *Problem (4.4) admits a unique solution.*

Proof. A is a positive symmetric defined tensor so, there are 2 positive eigen values λ_1, λ_2 such as $\lambda_1\, v_h^2 \leq v_h A v_h \leq \lambda_2\, v_h^2 \quad \forall v_h \in V_{h,-1}$.

$$\begin{aligned}
\left|\sum_{K\in\mathcal{K}} \int_K A\nabla q_h.\nabla q_h\right| &\geq C\left|\sum_{K\in\mathcal{K}} \int_K \nabla q_h.\nabla q_h\right|, \\
&\geq C\left|\sum_{K\in\mathcal{K}} \int_K (\nabla q_h)^2\right|, \\
&\geq C\left|\sum_{K\in\mathcal{K}} \int_K |q_h|^2_{1,K}\right|, \\
&\geq C\sum_{K\in\mathcal{K}} \int_K |q_h|^2_{1,K}.
\end{aligned}$$

The discrete Poincaré's inequality gives,

$$\begin{aligned}
\left|\sum_{K\in\mathcal{K}} \int_K A\nabla q_h.\nabla q_h\right| &\geq \frac{C}{2}\sum_{K\in\mathcal{K}} |q_h|^2_{1,K} + \frac{C}{2}\frac{1}{C_p}\|q_h\|^2_{0,\Omega}, \\
\left|\sum_{K\in\mathcal{K}} \int_K A\nabla q_h.\nabla q_h\right| &\geq \alpha\|q_h\|^2_{1,h}.
\end{aligned}$$

Lax-Milgram's lemma gives the result. ■

Convergence and error estimate

An a priori estimate for p_h proved by applying the discrete Green's formula in addition with basic properties of the discrete bilinear form a_h gives the convergence and the error estimate for the solution of (4.4).

Lemma 2.4.2 *We have,*

$$\|p - p_h\|_{1,h} \leq C \left(\inf_{q_h \in Q_h} \|p - q_h\|_{1,h} + \inf_{v_h \in V_h} \|u - v_h\|_{0,\Omega} \right).$$

Proof.

Let be $q_h \in Q_h$, then we may write,

$$a_h(p_h, q_h) = \sum_{K \in \mathcal{K}} \int_K A\nabla p_h.\nabla q_h = \sum_{K \in \mathcal{K}} \int_K f.\nabla q_h = b_h(q_h),$$

we have,

$$\begin{array}{rcl}
Q_h\text{-coercivity} \quad \|p_h - q_h\|^2_{1,h} & \leq & C a_h(p_h - q_h, p_h - q_h), \\
& \leq & C \left(a_h(p_h, p_h - q_h) - a_h(q_h, p_h - q_h) \right), \\
a_h(p_h, p_h - q_h) - a_h(q_h, p_h - q_h) & = & a_h(p_h - p, p_h - q_h) - a_h(q_h - p, p_h - q_h).
\end{array}$$

The first term on the right gives,

$$a_h(p_h - p, p_h - q_h) = a_h(p_h, p_h - q_h) - a_h(p, p_h - q_h),$$

$$\begin{array}{rcl}
= \displaystyle\sum_{K \in \mathcal{K}} \int_K f.\nabla(p_h - q_h) & - & \displaystyle\sum_{K \in \mathcal{K}} \int_K A\nabla p.\nabla(p_h - q_h), \\
\Rightarrow a_h(p_h - p, p_h - q_h) & = & \displaystyle\sum_{K \in \mathcal{K}} \int_K (f - A\nabla p).\nabla(p_h - q_h), \\
& = & \displaystyle\sum_{K \in \mathcal{K}} \int_K u.\nabla(p_h - q_h), \\
= \displaystyle\sum_{K \in \mathcal{K}} \int_K (u - v_h).\nabla(p_h - q_h) & + & \displaystyle\sum_{K \in \mathcal{K}} \int_K v_h.\nabla(p_h - q_h).
\end{array}$$

Now, we apply Green's formula to obtain,

$$\begin{array}{rcl}
\displaystyle\sum_{K \in \mathcal{K}} \int_K v_h.\nabla(p_h - q_h) & = & \displaystyle\int_{\Gamma_I} v_h.n\, (p_h - q_h)\, d\sigma - \sum_{K \in \mathcal{K}} \int_K (p_h - q_h)\, \mathrm{div}(v_h), \\
& = & 0
\end{array}$$

so, $a_h(p_h - p, p_h - q_h) \leq C_2 \|u - v_h\|_{0,\Omega} \|p_h - q_h\|_{1,h}$

The second term gives by linearity,

$$
\begin{aligned}
a_h(q_h - p, p_h - q_h) &\leq C_3 \, \|p - q_h\|_{1,h} \, \|p_h - q_h\|_{1,h}, \\
\text{and} \quad \|p_h - q_h\|_{1,h} &\leq C_4 \left(\|u - v_h\|_{0,\Omega} + \|p - q_h\|_{1,h} \right), \\
\|p - p_h\|_{1,h} &\leq \|p - q_h\|_{1,h} + \|q_h - p_h\|_{1,h}, \\
&\leq \|p - q_h\|_{1,h} + C_4 \left(\|u - v_h\|_{0,\Omega} + \|p - q_h\|_{1,h} \right), \\
&\leq C_5 \left(\|u - v_h\|_{0,\Omega} + \|p - q_h\|_{1,h} \right), \\
&\quad \forall q_h \in Q_h, \quad \forall v_h \in V_h, \\
\|p - p_h\|_{1,h} &\leq C_5 \left(\inf_{v_h \in V_h} \|u - v_h\|_{0,\Omega} + \inf_{q_h \in Q_h} \|p - q_h\|_{1,h} \right).
\end{aligned}
$$

■

As a consequence, the a priori estimate for the pressure reads,

Theorem 2.4.3 *If $p \in H^{1+\sigma}$ and $u \in (H^\sigma)^d \cap H(div, \Omega)$, $1/2 < \sigma \leq 1$, we have*

$$\|p - p_h\|_{1,h} \leq C h^\sigma \left(|p|_{1+\sigma,\Omega} + \|u\|_{\sigma,\Omega} \right).$$

Projection operator

Let us study,

$$
\begin{aligned}
Pr : (L^2)^d &\longrightarrow V_{h,-1} \\
g &\longmapsto Pr_h \, g / < g - Pr_h \, g, v_h >_{0,\Omega} = 0 \quad \forall v_h \in V_{h,-1}.
\end{aligned}
$$

We know that $\nabla_h Q_h \subset V_{h,-1}$,

We have,

$$\forall q_h \in Q, \int_K (g - Pr_h \, g).\nabla q_h = 0.$$

Let's take $u_h = Pr_h[f - A\nabla p_h]$,

Remark 5 *If A is diagonal, $u_h = Pr_h \, f - A\nabla p_h$ because $\nabla p_h \in RT_0(K)$ implies $div(\nabla p_h) = 0$, and $Pr_h \, (A\nabla p_h) = A\nabla p_h$.*

2.4.2 Velocity formulation

The velocity formulation takes the form (4.5)-(4.6),

Lemma 2.4.4

$$
\begin{aligned}
& u_h \in V_{h,-1} \textit{ and verifies },\\
< u_h, v_h >_{0,\Omega} + \sum_K \int_K A\nabla p_h.\nabla q_h &= \int_\Omega f.v_h, \qquad \forall v_h \in V_{h,-1}, \qquad (4.5)\\
\sum_K \int_K u_h.\nabla q_h &= 0, \qquad \forall q_h \in Q_h. \qquad (4.6)
\end{aligned}
$$

Proof. We have,

$$
\begin{aligned}
< u_h, v_h >_{0,\Omega} &= \sum_K \int_K Pr_h[f - A\nabla p_h].v_h\\
&= \sum_K \left(\int_K f.v_h - \int_K A\nabla p_h.v_h \right),\\
&= 0.
\end{aligned}
$$

and,

$$
\sum_K \int_K u_h.\nabla q_h = \sum_K \int_K (Pr_h[f - A\nabla p_h])\nabla q_h = 0.
$$

■

Existence and uniqueness of the solution

Theorem 2.4.5 *The problem (4.5)-(4.6) admits a unique solution (this is a square set of equations).*

Proof. For a square set of equations, uniqueness $\Leftrightarrow$ existence $\Leftrightarrow$ existence and uniqueness.

■

2.4.3 Dual problem

Lemma 2.4.6 $u_h \in V_h = V_{h,-1} \cap H^0_{\Gamma_0}(div, \Omega)$

a.e., u_h verifies the property of flux continuity ($[v_h.n]|_e = 0$).

Proof.

$$
\begin{aligned}
u_h = Pr_h\,(f - A\nabla p_h),\ u_h \in V_{h,-1}\\
\text{so } u_h.n_K|_e \in P_0(e),\ \forall e \text{ face in } \mathcal{K},\\
\text{let be } I = |e|\ [u_h.n_K]|_e = \int_e [u_h.n_K]\, d\sigma,\ \text{and}\\
q_h \in Q_h \ /\ \int_f q_h\, d\sigma = \delta_e^f, \quad \forall f \text{ face in } \mathcal{K}.
\end{aligned}
$$

$$\int_e [u_h.n]\, d\sigma \;=\; \int_e [u_h.n]\, q_h\, d\sigma,$$

We apply the discrete Green's formula,

$$\begin{aligned}
\int_e [u_h.n]\, d\sigma &= \sum_{i=0}^{1}\left(\int_{K_i} u_h.\nabla q_h\, d\sigma + \int_{K_i} \underbrace{\operatorname{div}(u_h)}_{=0}\, q_h\right) \qquad (e = \partial K_0 \cap \partial K_1),\\
&= \sum_{i=0}^{1}\int_{K_i} \left(Pr_h\left(f - A\nabla p_h\right)\right).\nabla q_h,\\
\int_e [u_h.n]\, d\sigma &= 0,\\
\Rightarrow [u_h.n_K]|_e &= 0.
\end{aligned}$$

■

We follow the discussion keeping in mind that $u_h \in H(\text{div}, \Omega)$,

If $e \in \Gamma_N$, the same arguments show that

$$u_h.n|_e = 0, \qquad \forall e \in \Gamma_N$$

Theorem 2.4.7 *We have, u_h is the unique solution of problem $(P)_h$:*

$$u_h \in V_h, \tag{4.7}$$

$$\forall v_h \in V_h, \qquad \int_\Omega A^{-1}u_h.\, v_h = \int_\Omega f.v_h. \tag{4.8}$$

Remark 6 *(4.7)-(4.8) is the dual Darcy's problem.*

Convergence and error estimate

Theorem 2.4.8 *We have the estimate :*

$$\|u - u_h\|_{0,\Omega} \leq \inf_{v_h \in V_h} \|u - v_h\|_{0,\Omega}.$$

Proof.

We start with the V_h-coercivity,

$$\begin{aligned}
\alpha\|u_h - v_h\|_{0,\Omega}^2 &\leq a_h(u_h - v_h, u_h - v_h)\\
&\leq a_h(u_h - u, u_h - v_h) + a_h(u - v_h, u_h - v_h),
\end{aligned}$$

The second term on the right gives,

$$a_h(u - v_h, u_h - v_h) \quad \leq \quad \beta \, \|u - v_h\|_{0,\Omega} \, \|u_h - v_h\|_{0,\Omega},$$

whereas the first term can be written as,

$$\begin{aligned}
a_h(u_h - u, u_h - v_h) &= \sum_{K \in \mathcal{K}} \int_K A^{-1}(u_h - u) \, (u_h - v_h), \\
&= \sum_{K \in \mathcal{K}} \int_K A^{-1} u_h \, (u_h - v_h) - \sum_{\mathcal{K}} \int_K A^{-1} v_h \, (u_h - v_h), \\
&= \sum_{K \in \mathcal{K}} \int_K f \, (u_h - v_h) - \sum_{K \in \mathcal{K}} \int_K A^{-1} u \, (u_h - v_h), \\
a_h(u_h - u, u_h - v_h) &= 0
\end{aligned}$$

We sum both results for the first and second term,

$$\begin{aligned}
& \alpha \, \|u_h - v_h\|_{0,\Omega}^2 \leq \beta \, \|u - v_h\|_{0,\Omega} \, \|u_h - v_h\|_{0,\Omega}, \\
\Rightarrow \quad & \|u_h - v_h\|_{0,\Omega} \leq C \, \|u - v_h\|_{0,\Omega} \\
& \|u - u_h\|_{0,\Omega} \leq \|u - v_h\|_{0,\Omega} + \|u_h - v_h\|_{0,\Omega}, \\
\Rightarrow \quad & \|u - u_h\|_{0,\Omega} \leq C_2 \, \|u - v_h\|_{0,\Omega} \qquad \forall v_h \in V_h, \\
\text{so,} \qquad & \|u - u_h\| \leq C_2 \inf_{v_h \in V_h} \|u - v_h\|_{0,\Omega}.
\end{aligned}$$

■

Acknowledgement. The authors are especially grateful to Dr A. Agouzal for his highly pertinent remarks and help leading the analysis toward this paper.

2.5 Numerical results

With the aim of illustrating the accuracy of the non-conforming method, let us consider the Neumann problem borrowed from [5] and adapted in three dimensions.

$$\begin{cases} v - A\nabla P = F & \text{in } \Omega = (0,1)^3, \\ \frac{\partial p}{\partial n} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.9}$$

where F is given by $(x - 3/2x^2; y - 3/2y^2; z - 3/2z^2)$ for which the exact solution is expressed by

$$
\begin{aligned}
p_{ex} &= \frac{1}{2}\left[x^2(x-1) + y^2(y-1) + z^2(z-1)\right] + C, \\
v_{ex} &= 0.
\end{aligned}
$$

Ω is decomposed uniformly in regular hexaedras of equal size. We report the respective error in $e_2(u) = |I_h(u_{ex}) - u_h|_{2,\Omega}$, $e_1(u) = |I_h(u_{ex}) - u_h|_{1,\Omega}$ and $e_{max} = \max_u |I_h(u_{ex}) - u_h|$ norms for $u = p,\ v$ and took $A = I$. We compare the errors obtained by Q_1 parallelotope finite element and Q_1 non conform finite element. We observe quadratic convergence for Q1 parallelotope method and linear convergence for Q1 non conform method. We mention also [15], where super convergence of the vector field along Gauss lines in a mixed finite element approximation for a second order elliptic problem on a rectangular grid was proven. In other words, super convergence exists for non conform finite element but in the case of Gauss-Lobato interpolation on the interface for error estimates.

We summarize our results below, N denotes the number of nodes on each edge of the cube $\Omega = (0,1)^3$, and s denotes the slop of the error-curve given by $log(e_i)$, $i = 1, 2$ or max with respect to $log(h) = log(1/N)$ obtained by linear regression, so in a certain manner we obtain a value of the numerical order of convergence.

Q_1-parall.						
N	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
10	0.0089124	0.0092593	0.0046652	0.0082305	0.0082305	0.0085768
20	0.0016784	0.0020776	0.0007075	0.0019682	0.0019682	0.0020147
40	0.0004791	0.0004931	0.0002444	0.0004805	0.0004805	0.0004863
60	0.0001544	0.0002155	0.0001213	0.0002118	0.0002118	0.0002135
80	0.0001201	0.0001202	0.0000635	0.0001187	0.0001187	0.0001194
N	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
10	0.0044352	0.0072016	0.0073160	0.0080018		
20	0.0006286	0.0018589	0.0018646	0.0019567		
40	0.0002305	0.0004678	0.0004681	0.0004798		
60	0.0001215	0.0002082	0.0002082	0.0002117		
80	0.0000603	0.0001171	0.0001171	0.0001186		

Q_1-parall.						
	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
s	2.01	2.08	1.98	2.04	2.04	2.05
	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
s	1.98	1.99	2.03			

Q_1 non-cf.						
N	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
10	0.0160639	0.0174254	0.0478395	0.0103351	0.0072842	0.0072842
20	0.0037858	0.0041114	0.0245845	0.0034324	0.0018023	0.0018023
40	0.0016169	0.0016966	0.0124096	0.0011766	0.0004470	0.0004470
60	0.0009436	0.0009788	0.0082950	0.0006339	0.0001981	0.0001981
80	0.0005698	0.0000540	0.0062290	0.0004096	0.0001112	0.0001112
N	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
10	0.0160537	0.0072842	0.0072842	0.0072842		
20	0.0037833	0.0018023	0.0018023	0.0018028		
40	0.0016165	0.0004470	0.0004470	0.0004470		
60	0.0009435	0.0001981	0.0001981	0.0001981		
80	0.0005697	0.0001112	0.0001112	0.0001112		

Q_1 non-cf.						
	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
s	1.57	2.32	1.54	1.55	2.01	2.01
	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
s	1.54	2.01	2.01	2.01		

We see clearly that both methods are converging in this case. It is important to say that the result for this sample are shown to illustrate the non-conforming method and not to compare one method versus the other one, but qualitatively to obtain almost the same result.

Remark 7 *All computations were made with the under development research library SciFEM*[1]

[1]Finite Element Library developped in a joint work by T. Clopeau and V. Devigne for research in

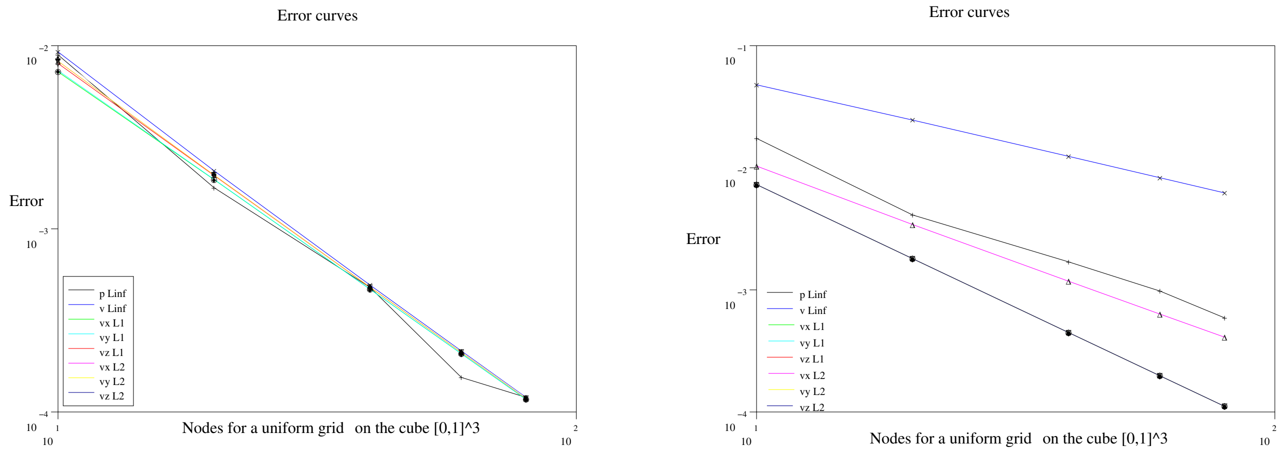


FIG. 2.1 – Error curves for Q1-parallelotope FEM (left) and for Q1-non conform FEM (right)

(for Scilab Finite Element Method) and pictures were made with the free mesh viewer paraview.

In the graphics 2.1 we show error curves in log-log scale.

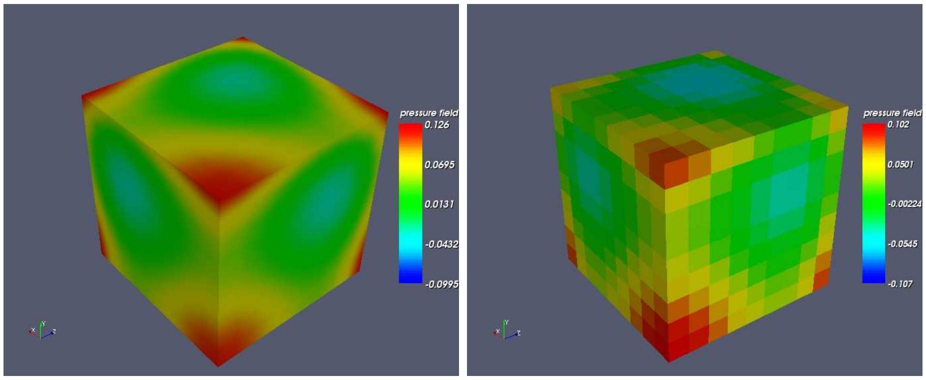


FIG. 2.2 – pressure field for N=10

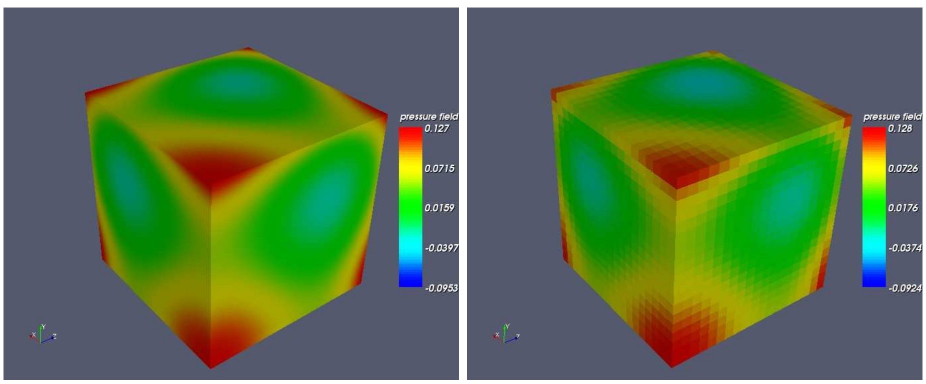


FIG. 2.3 – pressure field for N=20

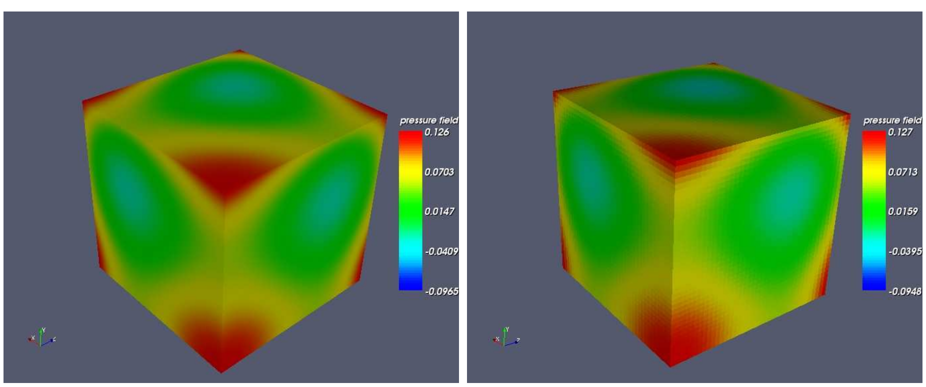


FIG. 2.4 – pressure field for N=40

2.6 Conclusion

The non-conforming element we propose may be seen as an alternative to standard finite element methods for a decompostion of the domain into hexaedras. With reasonable accuracy it provides another kind of approximation for Darcy flow, natural and easy to compute. Another interest is the low price in degrees of freedom. One per each face with polynoms of order one. Order of convergence for the velocities are qualitatively the same for both methods. We quote one advantage disregarding the non-conforming method that is not shown in this example, the convergence still holds for anisotropic or let's say "discontinuous" coefficients of the permeability tensor. This is not the case for conforming method.

L'élément non conforme que nous proposons peut être vu comme une alternative aux méthodes d'éléments finis standards pour une décomposition du domaine en hexaèdres. Avec une précision raisonnable nous obtenons un autre approximation de l'écoulement de Darcy, naturelle et simple à mettre en oeuvre. Un autre point positif est le coût moins important en degrés de liberté. Un pour chaque face pour des polynômes de degré un. L'ordre de convergence pour les vitesses est qualitativement le même pour les deux méthodes. Nous soulignons que dans le cas non conforme, l'approximation est convergente pour des coefficients anisotropiques ou "discontinus" du tenseur des perméabilités ce qui n'est pas vrai dans le cas conforme.

Bibliographie

[1] *P.A. Raviart, J.M. Thomas, Primal Hybrid Finite Element Methods for second Order Elliptic Equations , Mathematics of computation, vol. 31, n° 138, (1977), p 391-413.*

[2] F. Brezzi, M. Fortin, Mixed and hybrid Finite Element Methods, Springer-Verlag New-York inc., (1991).

[3] A. Masud, T. J.R. Hughes, A stabilized mixed finite element method for Darcy Flow, Computer Methods in applied Mechanics and Engineering, n° 191, (2002), p 4341-4370.

[4] Y. Achdou, C. Bernardi, F. Coquel, A priori and a posteriori analysis of finite volume discretizations of Darcy's equations, Université Pierre et Marie Curie, Université Paris VII, (2001).

[5] Y. Achdou, C. Bernardi, Un schéma de volumes ou éléments finis adaptatif pour les équations de Darcy à perméabilité variable, C.R. Acad. Sci. Paris, t. 333, Série I, (2001), p 693-698.

[6] P.G. Ciarlet Numerical analysis of the Finite Element Method, Séminaire de Mathématique Supérieures, Univ. de Montréal (1975).

[7] P.G. Ciarlet The finite element method for elliptic problems, North-Holland Publishing Company, (1978).

[8] G. Strang et G. Fix An analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs, N.J., (1973).

[9] M. Crouzeix, P.A. Raviart Conforming and nonconforming methods for solving the stationary Stokes equations , Rev. Française Autom. Inform. Recherche Opérationelle Sér. Anal. Numér., v. 7, (1973), no R.3,pp. 33-75.

[10] B.M. Irons, A. Razzaque, Experience with the patch test for convergence of finite elements, The mathematical foundation of the finite element method with applications to partial differential equations, Part II (A.K.Aziz editor), Academic press, New York, (1972), pp 557-587.

[11] G. Strang Variational crimes in the finite element methods, The mathematical foundation of the finite element method with applications to partial differential equations, Part II (A.K.Aziz editor), Academic press, New York, (1972), pp 689-710.

[12] P. Lesaint On the convergence of the Wilson's non conforming element for solving elastic problems, Comput. Methods Apll. Mech. Engrg., V. 7, (1976), pp. 1-16.

[13] A. Quarteroni, A. Valli, Numerical Approximation of Partial Differential Equations, Springer-Verlag, (1994).

[14] A.Agouzal, , CRAS, (2002)

[15] J. Jr. Douglas, J. Wang Superconvergence of mixed finite element methods on rectangular domains, Calcolo 26 (1989), no. 2-4, 121-133, (1990).

Chapitre 3

Écoulement tangentiel sur une surface rugueuse et loi de Navier

Sommaire

3.1 **Préambule** **41**

3.2 **Introduction** **43**

3.3 **Couche limite de Navier** **45**

3.4 **Justification de la condition de glissement de Navier pour l'écoulement de Couette laminaire 3D** **47**

3.5 **Conclusion et applications possibles en océanographie** **52**

Cette section a été publiée avec les références suivantes :

A. Mikelić, V. Devigne "Tangential flows on rough surface and Navier's law", Annales Mathématiques Blaise Pascal vol.9, 313-327, (2002).

3.1 Préambule

Nous considérons des écoulements visqueux tangents sur des surfaces rugueuses et obtenons la condition aux limites efficaces. C'est la condition de glissement de Navier, utilisée dans les calculs d'écoulement de fluide visqueux pour des géométries complexes. Les coefficients efficaces, la matrice de Navier, sont obtenus par changement d'échelle et en résolvant un problème de couche limite adéquat. Ensuite, nous appliquons ce résultat à la réduction de la trainée.

L'homogénéisation, couplée aux couches limites donnent des lois efficaces aux interfaces. Dans ce chapitre nous expliquerons comment ces résultats sont obtenus. D'un point de vue physique, la condition d'adhérence $v = 0$, sur une paroi solide immobile, n'est justifiée que lorsque la viscosité moléculaire est prise en compte. Sachant que le fluide ne peut pas pénétrer le solide, sa composante normale est égale à zéro. C'est la condition de non-pénétration. Au contraire, l'absence de glissement n'est pas très intuitive. Pour les fluides newtoniens cela a été établi expérimentalement mais contesté par Navier en personne (voir [16]). Il proclamait que la vitesse de glissement devait être proportionnelle à la contrainte de cisaillement. Les calculs de la théorie cinétique ont confirmé la condition aux limites de Navier, mais ils ont donné la longueur du glissement proportionnelle au chemin libre moyen divisé par la longueur du continuum (voir [17]). Pour des raisons pratiques, cela signifie une longueur de glissement nulle, justifiant l'emploi de la condition d'adhérence.

Dans beaucoup de cas pratiques significatifs la frontière est rugueuse. Un exemple sont les frontières complexes dans la dynamique des fluides géophysiques. Comparées avec la taille caractéristique d'un domaine simulable, de telles frontières ne peuvent être considérées dans leur complète géométrie. D'autres exemples décrivent les fonds marins comme des rugosités aléatoires ou comme des corps artificielles avec distribution périodique de petites bosses. Une simulation numérique de problèmes d'écoulement en présence de frontière rugueuse est très difficile car il demande beaucoup de degrés de liberté et beaucoup de données. Pour des raisons de calcul, une frontière artificielle régulière, approchant l'originale, est préférable et les équations sont résolues dans le nouveau domaine. Ainsi, les frontières rugueuses sont évitées, mais les conditions aux limites aux frontières artificielles ne sont pas données pour autant. Il est clair que la condition de non-pénétration $v \cdot n = 0$ doit être conservée. Habituellement, il est supposé que la contrainte de cisaillement est une fonction non-linéaire F de la vitesse tangentielle. F est déterminée empiriquement et sa forme varie suivant les problèmes. De telles relations sont appelées *lois de paroi*. Un exemple classique est la condition de Navier,

$$\frac{\partial v_\tau}{\partial \nu} = \frac{1}{\epsilon\beta} v_\tau + O(1)$$

où v_τ est la vitesse tangentielle, ϵ la taille caractéristique des pores et β est le coefficient de Navier.

L'article est dédié au 60ième anniversaire de Willi Jäger

résumé : Nous considérons un écoulement visqueux tangentiel à une surface rugueuse. Nous étudions le comportement efficace d'écoulements au voisinage de la frontière lorsque la taille caractéristique des rugosités est petite. Nous montrons que la frontière rugueuse peut

être approchée par une surface artificielle lisse (la paroi numérique) sur laquelle la condition d'adhérence à la paroi pour la vitesse tangentielle devient la condition de glissement de Navier. Pour une surface rugueuse périodique nous déterminons la matrice de frottement de Navier et montrons que le problème efficace, basé sur la loi de paroi, est une approximation du problème physique d'ordre la taille des rugosités puissance 3/2.

3.2 Introduction

En physique des fluides newtoniens, la condition d'adérence à la paroi est justifiée seulement en présence de viscosité. Comme le fluide ne peut pas traverser la paroi, sa vitesse normale est égale à zéro. C'est la condition de non-pénétration. En revanche, l'absence de glissement n'est pas très intuitive. Il s'agit d'une loi établie expérimentalement et contestée par Navier lui-même (voir [16]). Navier pensait que la vitesse tangentielle devait être proportionnelle au cisaillement (la loi de Navier). La théorie cinétique a confirmé la loi de Navier, mais avec le coefficient de frottement proportionnel au chemin moléculaire moyen libre (voir e.g. [17]). En conséquence pour les fluides réels ce coefficient vaut zéro et confirme la condition d'adhérence à la paroi.

Dans beaucoup d'applications la frontière est rugueuse comme dans le cas des frontières complexes en dynamique des écoulements géophysiques. La taille du domaine de calcul (mer, océan) est tellement grande que les détails de la côte peuvent être considérés comme des rugosités. Comme autres exemples on peut citer les fonds marins constitués de rugosités aléatoires et les corps artificiels avec une distribution périodique des rugosités. Une simulation numérique des problèmes d'écoulements en présence d'une géométrie rugueuse est très difficile pour deux raisons. Un grand nombre de noeuds de maillage est nécessaire ainsi que de multiples données. D'un point de vue numérique, une frontière rugueuse artificielle proche de celle d'origine est choisie et les équations sont résolues dans le "nouveau" domaine. La difficulté de la frontière rugueuse est contournée mais des conditions aux limites viennent à manquer. Il est clair que la condition de non-pénétration $v.n = 0$ doit être conservée mais il n'y a en outre aucune raison de garder la condition d'adhérence. Habituellement le cisaillement est supposé être une fonction non-linéaire F de la vitesse tangentielle. F est déterminée de façon empirique et varie suivant la nature du problème. De telles relations sont appelées *loi de paroi* et la condition classique de Navier en est un exemple. Un autre exemple est la modélisation de la couche limite turbulente à proximité de la surface rugueuse par un profil logarithmique des vitesses.

$$v_\tau = \sqrt{\frac{\tau_w}{\rho}}\Big(\frac{1}{\kappa}\ln\big(\frac{y}{\mu}\sqrt{\frac{\tau_w}{\rho}}\big) + C^+(k_s^+)\Big) \tag{1}$$

où v_τ est la vitesse tangentielle, y la coordonnée verticale et τ_w le cisaillement, ρ représente la densité et μ la viscosité, $\kappa \approx 0.41$ est la constante de von Kármán's et C^+ une fonction du rapport k_s^+ c.a.d de la hauteur des rugosités k_s sur l'épaisseur de la sous-couche mince de la paroi $\delta_v = \frac{\mu}{v_\tau}$. Pour plus de détail nous donnons la référence du livre de Shlichting [18] .

Justifier le profil logarithmique des vitesses dans la couche superposée est mathématiquement hors d'atteinte pour le moment. Cependant à partir de résultats récents, nous sommes capables de justifier la loi de Navier pour des écoulements laminaires visqueux incompressibles au dessus de frontières à rugosité périodique. Dans cet article nous allons donner un aperçu de résultats récents rigoureux sur la condition de Navier avec des application possibles aux écoulements géophysiques.

Nous nous concentrons sur les équations de Stokes et de Navier-Stokes incompressibles et nous allons présenter une dérivation de la loi de Navier en construisant des couches limites développées dans [9].

Les problèmes d'écoulement sur des surfaces rugueuses ont été traités par O.Pironneau et ses collaborateurs dans [15] , [1] et [2]. L'article [15] traite l'écoulement sur une surface rugueuse et l'écoulement à la surface d'une mer ondulant.Quelques problèmes sont discutés et un résultat rigoureux dans [2] est annoncé. Finalement dans l'article [2] l'écoulement stationnaire incompressible pour un grand nombre de Reynolds $Re \sim \frac{1}{\varepsilon}$ sur une frontière rugueuse périodique est traité avec ϵ la période des rugosités. Un développement asymptotique est construit et à partir de couches limites correctrices définies dans une cellule semi-infinie des lois de paroi effectives sont obtenues. Une validation numérique est jointe mais il n'y a pas de résultats de convergence mathématique rigoureux. L'estimation d'erreur annoncée dans [1] n'a pas été prouvée dans [2].

Nous mentionnons également les résultats de Y.Amirat, J.Simon et leurs collaborateurs de l'écoulement de Couette sur une plaque ondulée (voir [3] , [4] et [5]). Ils traitent principalement de la réduction de la trainée pour l'écoulement de Couette.

Notre approche sera différente des réferences mentionnées ci-dessus. Nous allons justifier la loi de glissement de Navier par la technique développée dans [8] pour l'opérateur de Laplace et par suite dans [9] pour le système de Stokes. Le résultat pour un écoulement 2D laminaire visqueux incompressible sur une surface rugueuse est dans [9]. Dans le texte qui suit, nous considérons un écoulement de Couette 3D sur une surface rugueuse. Dans le §2 nous introduisons le problème de couche limite correspondant et dans §3 nous obtenons la condition de glissement de Navier. La section §4 est consacrée aux applications possibles à la dynamique des écoulements géophysiques.

3.3 Couche limite de Navier

Comme il a été déjà observé en hydrodynamique, les phénomènes importants pour une frontière apparaissent dans une couche mince l'entourant . Nous ne nous intéressons pas aux couches limites, correspondantes à la limite de la zéro viscosité pour les équations de Navier-Stokes, mais plutôt à la couche limite visqueuse, décrivant des effets de la rugosité. Il y a une similarité avec les couches limites décrivant les effets interfaciaux entre un domaine rempli par un fluide et un milieu poreux saturé . De telles couches limites ont été développées pour l'opérateur de Laplace, avec des conditions de Dirichlet homogènes sur des perforations, dans [8] . La théorie correspondante pour le système de Stokes est dans [9] et [14] . Le but de cette section est de présenter la construction de la couche limite principale, qui sera utilisée pour calculer la matrice de frottement, figurant comme coefficient dans la loi de glissement de Navier. Il est naturel de l'appeler *la couche limite de Navier*. Dans [10] la 2D couche limite était construite. On poursuit ici la 3D construction entamée dans [12] qui généralise les résultats du [10] .

Soient b_j, $j = 1, 2, 3$ des constantes positives. Soit $Z = (0, b_1) \times (0, b_2) \times (0, b_3)$ et soit Υ une surface Lipschitzienne $y_3 = \Upsilon(y_1, y_2)$, prenant valeurs entre 0 et b_3. Nous supposons que la surface rugueuse $\cup_{k\in \mathbb{Z}^2}\big(\Upsilon + k\big)$ est aussi Lipschitzienne. Nous introduisons la cellule canonique de rugosité par $Y = \big\{y \in Z \mid b_3 > y_3 > \max\{0, \Upsilon(y_1, y_2)\}\big\}$.

Le rôle crucial appartient au problème auxiliaire suivant :

Pour un vecteur constant $\lambda \in \mathbb{R}^2$, trouver $\{\xi^\lambda, \omega^\lambda\}$ satisfaisant

$$-\triangle_y \xi^\lambda + \nabla_y \omega^\lambda = 0 \quad \text{dans } Z^+ \cup (Y - b_3\vec{e}_3) \tag{2}$$

$$\mathrm{div}_y \xi^\lambda = 0 \qquad \text{dans } Z_{bl} \tag{3}$$

$$\big[\xi^\lambda\big]_S(\cdot, 0) = 0 \quad \text{sur } S \tag{4}$$

$$\big[\{\nabla_y \xi^\lambda - \omega^\lambda I\}\vec{e}_3\big]_S(\cdot, 0) = \lambda \;\text{ sur } S \tag{5}$$

$$\xi^\lambda = 0 \quad \text{sur } (\Upsilon - b_3\vec{e}_3), \quad \{\xi^\lambda, \omega^\lambda\} \text{ est } y' = (y_1, y_2) - \text{périodique}, \tag{6}$$

où $S = (0, b_1) \times (0, b_2) \times \{0\}$, $Z^+ = (0, b_1) \times (0, b_2) \times (0, +\infty)$, et $Z_{bl} = Z^+ \cup S \cup (Y - b_3\vec{e}_3)$.

Soit $V = \{z \in L^2_{loc}(Z_{bl})^3 : \nabla_y z \in L^2(Z_{bl})^9;\ z = 0$ on $(\Upsilon - b_3\vec{e}_3)$; $\mathrm{div}_y z = 0$ dans Z_{bl} et z est $y' = (y_1, y_2)$-périodique $\}$. Le lemme de Lax-Milgram nous garantie l'existence d'un unique $\xi^\lambda \in V$ satisfaisant

$$\int_{Z_{bl}} \nabla\xi^\lambda \nabla\varphi \, dy = -\int_S \varphi\lambda \, dy_1 dy_2, \qquad \forall\varphi \in V. \tag{7}$$

En utilisant le théorême de De Rham, nous obtenons une fonction $\omega^\lambda \in L^2_{loc}(Z_{bl})$, unique modulo une constante et satisfaisant (2). La théorie elliptique nous assure que $\{\xi^\lambda, \omega^\lambda\} \in V \cap C^{\inf}(Z^+ \cup (Y - b_3\vec{e}_3))^3 \times C^{\inf}(Z^+ \cup (Y - b_3\vec{e}_3))$ to (2)-(6).

Alors on a

Lemme 3.3.1 *([9] ,[14]). Soient a, a_1 et a_2, $a_1 > a_2$, des constantes positives. Alors la solution $\{\xi^\lambda, \omega^\lambda\}$ satisfait*

$$\begin{cases} \int_0^{b_1}\int_0^{b_2} \xi_2^\lambda(y_1, y_2, a) \, dy_1 dy_2 = 0, \\ \int_0^{b_1}\int_0^{b_2} \omega^\lambda(y_1, y_2, a_1) \, dy_1 dy_2 = \int_0^{b_1}\int_0^{b_2} \omega^\lambda(y_1, y_2, a_2) \, dy_1 dy_2, \\ \int_0^{b_1}\int_0^{b_2} \xi_j^\lambda(y_1, y_2, a_1) dy_1 dy_2 = \int_0^{b_1}\int_0^{b_2} \xi_j^\lambda(y_1, y_2, a_2) dy_1 dy_2, j = 1, 2 \\ C_\lambda^{bl} = \sum_{j=1}^2 C_\lambda^{j,bl}\lambda_j = \int_S \xi^\lambda\lambda \, dy_1 dy_2 = -\int_{Z_{bl}} |\nabla\xi^\lambda(y)|^2 \, dy < 0 \end{cases} \tag{8}$$

Lemme 3.3.2 . *Soit $\lambda \in I\!R^2$ et soit $\{\xi^\lambda, \omega^\lambda\}$ la solution de $(2) - (6)$ satisfaisant $\int_S \omega^\lambda \, dy_1 dy_2 = 0$. Alors $\xi^\lambda = \sum_{j=1}^2 \xi^j\lambda_j$ et $\omega^\lambda = \sum_{j=1}^2 \omega^j\lambda_j$, où $\{\xi^j, \omega^j\} \in V \times L^2_{loc}(Z_{bl})$, $\int_S \omega^j \, dy_1 dy_2 = 0$, est la solution du $(2) - (6)$ pour $\lambda = \vec{e}_j$, $j = 1, 2$.*

Lemme 3.3.3 . *Soit $a > 0$ et soit $\xi^{a,\lambda}$ la solution du $(2) - (6)$ avec S remplacé par $S_a = (0, b_1) \times (0, b_2) \times \{a\}$ et Z^+ by $Z_a^+ = (0, b_1) \times (0, b_2) \times (a, +\infty)$. Alors nous avons*

$$C_\lambda^{a,bl} = \int_0^{b_1}\int_0^{b_2} \xi^{a,\lambda}(y_1, y_2, a)\lambda \, dy_1 = C_\lambda^{bl} - a \mid \lambda \mid^2 b_1 b_2 \tag{9}$$

Ce résultat simple va impliquer l'invariance de la loi obtenue par rapport à la position de la paroi artificielle.

Corollaire 3.3.4 . *$\mid C_\lambda^{a,bl} \mid$ est minimale pour $a = 0$.*

Remarque 1 . *Si la frontière est plane, i.e. $\Upsilon =$ cte., alors la constante minimale $\mid C_\lambda^{a,bl} \mid$ est égale à zéro. Alors dans ce cas la condition d'adhérence à la paroi va rester. Si b_3 est la hauteur maximale de Υ, alors le lemme 3 n'est pas valable pour $a < 0$.*

Lemme 3.3.5 . *Soit $\{\xi^j, \omega^j\}$ vérifiant le lemme 3.3.2 et soit $M_{ij} = \dfrac{1}{b_1 b_2}\int_S \xi_i^j \, dy_1 dy_2$ la matrice de Navier. Alors la matrice M est symétrique définie négative.*

Lemme 3.3.6 . *Soit* $\{\xi^j, \omega^j\}$, $j = 1$ *et* $j = 3$, *vérifiant le lemme 3.3.2. Alors on a*

$$\begin{cases} | D^\alpha\, rot_y \xi^j(y) | \leq C e^{-2\pi y_3 \min\{1/b_1, 1/b_2\}},\ y_3 > 0,\ \alpha \in I\!N^2 \cup (0,0) \\ | \xi^j(y) - (M_{1j}, M_{2j}, 0) | \leq C(\delta) e^{-\delta y_3},\ y_3 > 0,\ \forall \delta < \dfrac{2\pi}{\max\{b_1, b_2\}} \\ | D^\alpha \xi^j(y) | \leq C(\delta) e^{-\delta y_3},\ y_3 > 0,\ \alpha \in I\!N^2,\ \forall \delta < \dfrac{2\pi}{\max\{b_1, b_2\}} \\ | \omega^j(y) | \leq C e^{-2\pi y_3 \min\{1/b_1, 1/b_2\}}, \qquad y_3 > 0. \end{cases} \tag{10}$$

Corollaire 3.3.7 . *Le système* $(2) - (6)$ *définie une couche limite.*

Remarque 2 *Pour des simulation numériques de la couche limite de même type mais correspondant à la loi de Beavers et Joseph voir la référence [11] .*

3.4 Justification de la condition de glissement de Navier pour l'écoulement de Couette laminaire 3D

Une justification mathématiquement rigoureuse de la condition de glissement de Navier pour l'écoulement de Poiseuille 2D tangentiel à une frontière rugueuse est dans [10] . Un écoulement correspondant aux nombres de Reynolds modérés a été considéré et les résultats suivants ont été montrés : a) Un résultat de stabilité non-linéaire par rapport aux petites perturbations de la frontière régulière par une frontière rugueuse ; b) Un résultat d'approximation d'ordre $\varepsilon^{3/2}$; c) La justification rigoureuse de la condition de glissement de Navier.

Ici, nous allons présenter les résultats qui figurent dans la prépublication [12] et qui concernent un écoulement de Couette 3D.

Nous considérons un écoulement visqueux incompressible dans le domaine Ω^ε, contenant le parallélépipède $P = (0, L_1) \times (0, L_2) \times (0, L_3)$, l'interface $\Sigma = (0, L_1) \times (0, L_2) \times \{0\}$ et la couche rugueuse $R^\varepsilon = \Big(\cup_{\{k \in Z\!\!Z^2\}} \varepsilon \big(Y + (k_1, k_2, -b_3) \big) \Big) \cap \Big((0, L_1) \times (0, L_2) \times (-\varepsilon b_3, 0) \Big)$. La cellule canonique $Y \subset (0, b_1) \times (0, b_2) \times (0, b_3)$ est définie dans la section précédente . Pour la simplicité, nous supposons que $L_1/(\varepsilon b_1)$ et $L_2/(\varepsilon b_2)$ sont des entiers. Soit $\mathcal{I} = \{0 \leq k_1 \leq L_1/b_1;\ 0 \leq k_2 \leq L_2/b_2;\ k \in Z\!\!Z^2\}$. Alors, notre frontière rugueuse $\mathcal{B}^\varepsilon = \cup_{\{k \in \mathcal{I}\}} \varepsilon \big(\Upsilon + (k_1, k_2, -b_3) \big)$ contient un grand nombre de " bosses " , distribués périodiquement et ayant la longueur et l'amplitude caractéristique $\varepsilon << 1$.

Pour un $\varepsilon > 0$ et une vitesse $\vec{U} = (U_1, U_2, 0)$, donnés, l'écoulement de Couette est décrit par le système suivant

$$-\nu\triangle v^\varepsilon + (v^\varepsilon\nabla)v^\varepsilon + \nabla p^\varepsilon = 0 \quad \text{dans } \Omega^\varepsilon, \tag{11}$$

$$\operatorname{div} v^\varepsilon = 0 \qquad \text{dans } \Omega^\varepsilon, \tag{12}$$

$$v^\varepsilon = 0 \qquad \text{sur } \mathcal{B}^\varepsilon, \tag{13}$$

$$v^\varepsilon = \vec{U} \qquad \text{sur } \Sigma_2 = (0, L_1) \times (0, L_2) \times \{L_3\} \tag{14}$$

$$\{v^\varepsilon, p^\varepsilon\} \quad \text{est périodique en } (x_1, x_2) \text{ avec le période } (L_1, L_2) \tag{15}$$

où $\nu > 0$ est la viscosité cinématique et $\int_\Omega^\varepsilon p^\varepsilon \, dx = 0$.

Notons qu'un problème analogue a été considéré dans [5], mais dans une bande infinie ayant une frontière rugueuse. Dans [5] Y.Amirat et ses collaborateurs ont cherché une solution périodique en (x_1, x_2), avec la période $\varepsilon(b_1, b_2)$. Si notre solution est unique, elle coïncidera avec la solution construite dans [5] .

Comme nous avons besoin, non seulement de l'existence pour ε donné , mais aussi des estimations a priori indépendantes de ε, nous présentons un résultat de stabilité non-linéaire par rapport aux perturbations rugueuses de la frontière. Il va impliquer les estimations a priori uniformes.

Notons que l'écoulement de Couette dans P, satisfaisant les conditions d'adhérence à Σ, est donné par

$$v^0 = \frac{U_1 x_3}{L_3}\vec{e}_1 + \frac{U_2 x_3}{L_3}\vec{e}_2 = \vec{U}\frac{x_3}{L_3}, \qquad p^0 = 0. \tag{16}$$

Soit $|U| = \sqrt{U_1^2 + U_2^2}$. On montre aisément l'unicité si $|U|L_3 < 2\nu$ i.e. si le nombre de Reynolds est modéré .

Nous prolongeons la vitesse sur $\Omega^\varepsilon \setminus P$ par zéro.

On cherche une solution pour $(11) - (15)$ comme une petite perturbation de l'écoulement de Couette (16).

Théorème 3.4.1 *([12]). Soit $|U|L_3 \leq \nu$. Alors il existe des constantes $C_0 = C_0(b_1, b_2, b_3, L_1, L_2)$ telles que pour $\varepsilon \leq C_0(\frac{L_3}{|U|})^{3/4}\nu^{3/4}$ le problème $(11) - (15)$ admette une solution unique $\{v^\varepsilon, p^\varepsilon\} \in H^2(\Omega^\varepsilon)^3 \times H^1(\Omega^\varepsilon)$, $\int_\Omega^\varepsilon p^\varepsilon \, dx = 0$, satisfaisant*

$$\|\nabla(v^\varepsilon - v^0)\|_{L^2(\Omega^\varepsilon)^9} \leq C\sqrt{\varepsilon}\frac{|U|}{L_3}. \tag{19}$$

De plus,

$$\|v^\varepsilon\|_{L^2(\Omega^\varepsilon\setminus P)^3} \leq C\varepsilon\sqrt{\varepsilon}\frac{|U|}{L_3} \tag{20}$$

$$\|v^\varepsilon\|_{L^2(\Sigma)^3} + \|v^\varepsilon - v^0\|_{L^2(P)^3} \leq C\varepsilon\frac{|U|}{L_3} \tag{21}$$

$$\|p^\varepsilon - p^0\|_{L^2(P)} \leq C\frac{|U|}{L_3}\sqrt{\varepsilon}, \tag{22}$$

où $C = C(b_1, b_2, b_3, L_1, L_2)$.

Maintenant nous avons les estimations à priori uniformes par rapport à ε. Malheureusement l'approximation obtenue n'est pas suffisamment bonne. Il nous faut ajouter la correction d'ordre $O(\varepsilon)$.

En suivant l'approche de [10] et [12], la condition de glissement de Navier correspond à la correction de la vitesse d'ordre $O(\varepsilon)$ sur la frontière. Le calcul non-rigoureux donne

$$v^\varepsilon = v^0 - \frac{\varepsilon}{L_3}\sum_{j=1}^{2}U_j\big(\xi^j(\frac{x}{\varepsilon}) - (M_{j1}, M_{j2}, 0)H(x_3)\big) -$$
$$\frac{\varepsilon}{L_3}\sum_{j=1}^{2}U_j(1-\frac{x_3}{L_3})(M_{j1}, M_{j2}, 0)H(x_3) + O(\varepsilon^2)$$

où v^0 est la vitesse de Couette dans P et le dernier terme correspond au contre-écoulement créé par la stabilisation de la vitesse, qui correspond à la couche limite, vers une vitesse constante non nulle.En conséquence sur l'interface Σ

$$\frac{\partial v_j^\varepsilon}{\partial x_3} = \frac{U_j}{L_3} - \frac{1}{L_3}\sum_{i=1}^{2}U_i\frac{\partial \xi_j^i}{\partial y_3} + O(\varepsilon) \quad \text{et} \quad \frac{1}{\varepsilon}v_j^\varepsilon = -\frac{1}{L_3}\sum_{i=1}^{2}U_i\xi_j^i(\frac{x}{\varepsilon}) + O(\varepsilon).$$

Par passage à la moyenne nous trouvons la loi de glissement de Navier

$$u_j^{eff} = -\varepsilon\sum_{i=1}^{2}M_{ji}\frac{\partial u_i^{eff}}{\partial x_3} \quad \text{sur} \quad \Sigma, \tag{NFC}$$

où u^{eff} est la moyenne de v^ε sur une rugosité représentative et la matrice M est définie dans le lemme 2.5. On néglige les termes d'ordre plus élevé.

Nous allons maintenant donner une démonstration rigoureuse.

Il est clair que dans P l'écoulement va continuer à être régi par le système de Navier-Stokes. La présence des irrégularités va contribuer seulement aux conditions efficaces sur la paroi artificielle . La contribution dominante pour l'estimation (19) vient de l'intégrale de surface $\int_\Sigma \varphi_j$. En suivant [12], on l'élimine en utilisant les couches limites

$$\xi^{j,\varepsilon}(x) = \varepsilon\xi^j(\frac{x}{\varepsilon}) \quad \text{et} \quad \omega^{j,\varepsilon}(x) = \omega^j(\frac{x}{\varepsilon}), \quad x \in \Omega^\varepsilon,\ j = 1, 2, \tag{23}$$

où $\{\xi^j, \omega^j\}$ est définie dans le lemme 3.3.2. Nous avons, pour tout $q \geq 1$ et $j = 1, 2$,

$$\frac{1}{\varepsilon}\|\xi^{j,\varepsilon} - \varepsilon(M_{1j}, M_{2j}, 0)\|_{L^q(P)^3} + \|\omega^{j,\varepsilon}\|_{L^q(P)} + \|\nabla\xi^{j,\varepsilon}\|_{L^q(\Omega)^9} = C\varepsilon^{1/q} \tag{24}$$

et

$$-\triangle\xi^{j,\varepsilon} + \nabla\omega^{j,\varepsilon} = 0 \qquad \text{dans } \Omega^\varepsilon \setminus \Sigma \tag{25}$$

$$\text{div } \xi^{j,\varepsilon} = 0 \quad \text{dans } \Omega^\varepsilon \tag{26}$$

$$\big[\xi^{j,\varepsilon}\big]_\Sigma(\cdot, 0) = 0 \quad \text{sur } \Sigma \tag{27}$$

$$\big[\{\nabla\xi^{j,\varepsilon} - \omega^{j,\varepsilon}I\}e_3\big]_\Sigma(\cdot, 0) = e_j \quad \text{sur } \Sigma, \tag{28}$$

i.e. notre couche limite correspond à la création de petits tourbillons par des rugosités.

Comme dans [10] la stabilisation de $\xi^{j,\varepsilon}$ vers une vitesse constante non-nulle $\varepsilon\big(M_{1j}, M_{2j}, 0\big)$, sur la frontière supérieure, crée un contre-écoulement. Il est donné par $d^i = (1 - \dfrac{x_3}{L_3})\vec{e}_i$ et $g^i = 0$.

On veut montrer maintenant que les quantités suivantes sont $o(\varepsilon)$ pour la vitesse et $O(\varepsilon)$ pour la pression :

$$\mathcal{U}^\varepsilon(x) = v^\varepsilon - \frac{1}{L_3}\left(x_3^+\vec{U} - \varepsilon\sum_{j=1}^{2} U_j\xi^j(\frac{x}{\varepsilon}) + \varepsilon\frac{x_3^+}{L_3}M\vec{U}\right) \tag{29}$$

$$\mathcal{P}^\varepsilon = p^\varepsilon + \frac{\nu}{L_3}\sum_{j=1}^{2} U_j\omega^{j,\varepsilon}. \tag{30}$$

On a le résultat suivant

Théorème 3.4.2 *([12]). Soit $\mathcal{U}^\varepsilon$ donnée par (29) et $\mathcal{P}^\varepsilon$ par (30). Alors $\mathcal{U}^\varepsilon \in H^1(\Omega^\varepsilon)^3, \mathcal{U}^\varepsilon = 0$ sur Σ, elle est périodique en (x_1, x_2), exponentiellement petite sur Σ_2 et div $\mathcal{U}^\varepsilon = 0$ dans Ω^ε. De plus, $\forall\varphi$ satisfaisant les mêmes conditions aux limites, nous avons l'estimation suivante*

$$\begin{aligned}|\nu\int_{\Omega^\varepsilon}\nabla\mathcal{U}^\varepsilon\nabla\varphi - \int_{\Omega^\varepsilon}\mathcal{P}^\varepsilon\text{div}\varphi + \int_{\Omega^\varepsilon}\frac{x_3^+}{L_3}\sum_{j=1}^{2}U_j\frac{\partial\mathcal{U}^\varepsilon}{\partial x_j}\varphi + \int_{\Omega^\varepsilon}\mathcal{U}_3^\varepsilon\frac{\vec{U}}{L_3}\varphi \\ +|\int_{\Omega^\varepsilon}\big((v^\varepsilon - v^0)\nabla\big)(v^\varepsilon - v^0)\varphi| \leq C\varepsilon^{3/2}\|\nabla\varphi\|_{L^2(\Omega^\varepsilon)^9}\frac{|U|^2}{L_3}.\end{aligned} \tag{31}$$

Corollaire 3.4.1 *([12]). Soit $\mathcal{U}^\varepsilon(x)$ et $\mathcal{P}^\varepsilon$ définie par $(29) - (30)$ et soit*

$$\varepsilon \leq \frac{\nu^{6/7}}{|U|}\min\big\{\frac{\nu^{1/7}}{4(|M| + \|\xi\|_{L^\infty})}, C(b_1, b_2, b_3, L_1, L_2)L_3^{3/7}|U|^{1/7}\big\}. \tag{32}$$

Alors v^ε, construite dans le théorème 3.1, est une solution unique pour $(11) - (15)$ et

$$\|\nabla\mathcal{U}^\varepsilon\|_{L^2(\Omega^\varepsilon)^9} + \|\mathcal{P}^\varepsilon\|_{L^2(P)} \leq C\varepsilon^{3/2}\frac{|U|^2}{\nu L_3} \tag{33}$$

$$\|\mathcal{U}^\varepsilon\|_{L^2(P)^3} + \|\mathcal{U}^\varepsilon\|_{L^2(\Sigma)^3} \leq C\varepsilon^2\frac{|U|^2}{\nu L_3} \tag{34}$$

Les estimations $(33)-(34)$ permettent de justifier la condition de glissement de Navier.

Remarque 3 *Il est possible d'ajouter des termes correctifs additionnels et alors notre problème aura une erreur qui décroît exponentiellement. Pour des résultats de ce type on peut consulter [5], [3] et [4] . Le cas de frontières rugueuses ayant la hauteur et la longueur d'ordre différent est considéré dans la thèse doctorale de I. Cotoi [7]. L'avantage de notre méthode est que nous trouvons la loi de glissement de Navier avec un coefficient matriciel défini négatif, ce qui n'a pas été le cas dans les publications antérieures.*

Maintenant nous introduisons l'écoulement de Couette-Navier efficace par

Trouver la vitesse u^{eff} et la pression p^{eff} telles que

$$-\nu\triangle u^{eff}+(u^{eff}\nabla)u^{eff}+\nabla p^{eff}=0 \quad \text{dans } P, \tag{35}$$

$$\text{div } u^{eff}=0 \quad \text{dans } P, \quad u^{eff}=(U_1,U_2,0) \text{ sur } \Sigma_2,\ u_3^{eff}=0 \text{ sur } \Sigma \tag{36}$$

$$u_j^{eff}=-\varepsilon\sum_{i=1}^{2}M_{ji}\frac{\partial u_i^{eff}}{\partial x_3},\ j=1,2 \text{ sur } \Sigma \tag{37}$$

$$\{u^{eff},p^{eff}\} \text{ est périodique en } (x_1,x_2) \text{ avec la période } (L_1,L_2) \tag{38}$$

Soit $|U|L_3\le\nu$, alors le problème $(35)-(38)$ admet une solution unique

$$\begin{cases} u^{eff}=(\tilde{u}^{eff},0), \tilde{u}^{eff}=\vec{U}+(\dfrac{x_3}{L_3}-1)\Big(I-\dfrac{\varepsilon}{L_3}M\Big)^{-1}\vec{U} & \text{pour } x\in P \\ p^{eff}=0 & \text{pour } x\in P. \end{cases} \tag{39}$$

Nous estimons l'erreur faite en remplaçant $\{v^\varepsilon,p^\varepsilon,\mathcal{M}^\varepsilon\}$ par $\{u^{eff},p^{eff},\mathcal{M}^{eff}\}$.

Théorème 3.4.3 *([12]). Sous les hypothèses du théorème 3.1 nous avons*

$$\|\nabla(v^\varepsilon-u^{eff})\|_{L^1(P)^9}\le C\varepsilon, \tag{40}$$

$$\sqrt{\varepsilon}\|v^\varepsilon-u^{eff}\|_{L^2(P)^3}+\|v^\varepsilon-u^{eff}\|_{L^1(P)^3}\le C\varepsilon^2\frac{|U|}{L_3}, \tag{41}$$

L'etape suivante est de calculer la *traînée*

$$\mathcal{F}_{t,j}^\varepsilon=\frac{1}{L_1L_2}\int_\Sigma \nu\frac{\partial v_j^\varepsilon}{\partial x_3}(x_1,x_2,0)\,dx_1dx_2,\ j=1,2. \tag{42}$$

Théorème 3.4.4 *([12]). Soit la traînée $\mathcal{F}_t^\varepsilon$ définie par (42). Alors on a*

$$|\mathcal{F}_t^\varepsilon-\nu\frac{1}{L_3}\big(\vec{U}+\frac{\varepsilon}{L_3}M\vec{U}\big)|\le C\varepsilon^2\frac{|U|^2}{\nu L_3}(1+\frac{\nu}{L_3|U|}) \tag{43}$$

Remarque 4 *On tire la conclusion que la présence des rugosités périodiques peut diminuer la traînée. La contribution est linéaire en ε, et en conséquence assez petite. D'après [4] pour* *un écoulement laminaire, la réduction de la traînée causée par la rugosité est négligeable. Néanmoins, pour un écoulement turbulent la présence des rugosités réduit la traînée d'une manière significative. Le phénomène a été observé sur les Avians et les Nektons (la peau de requin) et utilisé pour les bateaux et les avions (e.g. pour le yacht " Stars and Strips " dans la finale de l'America's Cup). Une référence classique est l'article de Bushnell, Moore [6] . Pour une explication mathématique on peut consulter l'article [12] .*

Remarque 5 *Comme dans [10] nous montrons qu'une perturbation de position de la paroi artificielle d'ordre $O(\varepsilon)$ implique une perturbation de la solution de $O(\varepsilon^2)$. Ce résultat est une conséquence du Lemme 2.3. En effet la matrice M change, mais sa perturbation est compensée par le changement de la position de Σ. Ainsi, nous avons la liberté de fixer la position de Σ. L'influence sur le résultat est seulement d'un ordre plus élevé dans le développement asymptotique.*

3.5 Conclusion et applications possibles en océanographie

Parmi les applications aux écoulements géophysiques on trouvera des écoulements turbulents en présence des frontières rugueuses. On peut se poser la question de l'utilité des estimations laminaires obtenues dans le §3. L'application est basée sur la théorie de la couche limite turbulente présentée dans le livre [18].

On sait que l'écoulement de Couette turbulent a une structure bi-couche : une couche turbulente où la viscosité moléculaire est négligeable comparée aux effets turbulents et une sous-couche visqueuse où on est obligé de tenir compte simultanément des effets turbulents et visqueux. L'écoulement dans cette sous-couche est régi par le cisaillement turbulent τ_w, qui dépend seulement du temps t. τ_w définie la vitesse de frottement $v = \sqrt{\frac{\tau_w}{\rho}}$, où ρ est la densité . Alors l'épaisseur de la couche est $\delta_v = \frac{\nu}{v}$ et la théorie de l'écoulement de Couette turbulent (voir e.g. [18]) donne

$$u^+ = f(y^+),\ y^+ = \frac{x_3}{\delta_v} \quad \text{et} \quad u^+ = \frac{\bar{u}}{v}$$

où $\bar{u}$ est la vitesse moyenne, comme la loi universelle pour la distribution de la vitesse dans la couche visqueuse. y^+ est la coordonnée caractéristique de la paroi et $u^+ = \frac{\bar{u}}{v}$ est la vitesse renormalisé . Cette loi est valable pour $0 \leq y^+ < 5$, pour $5 < y^+ < 70$ on a la zone tampon et pour $y^+ > 70$ la couche de chevauchement , où on a un profil logarithmique de la vitesse $u^+ = \frac{1}{\kappa} \ln y^+ + C^+$. Pour plus de détails voir [18] . Pour une paroi lisse $C^+ = 5$. Cette théorie s'applique aussi aux parois rugueuses et on les traite en ajustant la constante C^+. C^+ depend du rapport $k_s^+ = \dfrac{k_s}{\delta_v}$, où k_s est la hauteur des rugosités. Alors $\lim\limits_{k_s^+ \to 0} C^+(k_s^+) = 5$ (paroi lisse) et $\lim\limits_{k_s^+ \to +\infty} \big(C^+(k_s^+) + \dfrac{1}{\kappa} \ln k_s^+\big) = 8$ (paroi très rugueuse).

L'analyse mathématique de l'écoulement dans la zone tampon et dans la couche de chevauchement est hors de portée pour le moment. Néanmoins, si nous supposons que les petites rugosités restent en permanence dans la sous-couche visqueuse nous pouvons appliquer la théorie de la section §3.

Les équations correspondantes sont $(11) - (15)$ avec $L_3 = \delta_v$ et la vitesse de la paroi supérieure $v = \sqrt{\dfrac{\tau_w}{\rho}} = (v_1, v_2, 0)$ en $x_3 = \delta_v$. Comme $\delta_v \sqrt{\dfrac{\tau_w}{\rho}} = \nu < 2\nu$, nos résultats de §3 peuvent être appliqués et on obtient la loi de glissement de Navier si

$$\varepsilon \leq \frac{\nu^{6/7}}{|U|} \min\big\{\frac{\nu^{1/7}}{4(|M| + \|\xi\|_{L^\infty})}, C(b_1, b_2, b_3, L_1, L_2) L_3^{3/7} |U|^{1/7}\big\}. \tag{32}$$

Cette théorie était applicable aux Nektons (voir [12]).

Si cette condition est violée mais avec $\varepsilon < \sqrt{\nu}$ alors on a des variantes non-linéaires de la loi de glissement de Navier. Cette variante non-linéaire est en effet non-locale elle aussi et on a

$$u_j^{eff} = -\varepsilon \sum_{i=1}^{2} M_{ji}(\zeta(y_1, 0, \frac{\partial u_k^{eff}}{\partial x_2})) \frac{\partial u_i^{eff}}{\partial x_3} \quad \text{sur} \quad \Sigma, \tag{$NFCN$}$$

En général M_{ji} est une fonction bornée du gradient de u^{eff}, étant lui-même une fonction non-locale de ζ. ζ satisfait des équations de la couche limite pour un système de Navier-Stokes non-stationnaire. Pour plus de détails voir [13] .

En addition aux effets non-linéaires, les frontières rugueuses sont où fractales où aléatoires. On peut envisager le travail dans cette direction.

Pour conclure, nous pensons qu'il est tout à fait possible de développer une théorie des lois de parois, appliquer en Océanographie et fondée sur des résultats mathématiques rigoureux.

Bibliographie

[1] Y. Achdou, O. Pironneau, F. Valentin, *Shape control versus boundary control*, eds F.Murat et al , Equations aux dérivées partielles et applications. Articles dédiés à J.L.Lions, Elsevier, Paris, 1998, p. 1-18.

[2] Y. Achdou, O. Pironneau, F. Valentin, *Effective Boundary Conditions for Laminar Flows over Periodic Rough Boundaries*, J. Comp. Phys. , 147 (1998), p. 187-218.

[3] Y. Amirat, J. Simon, *Influence de la rugosité en hydrodynamique laminaire*, C. R. Acad. Sci. Paris, Série I, 323 (1996), p. 313-318.

[4] Y. Amirat, J. Simon, *Riblet and Drag Minimization*, in Cox, S (ed) et al., Optimization methods in PDEs, Contemp. Math, 209, p.9-17, American Math. Soc., Providence, 1997.

[5] Y. Amirat, D. Bresch, J. Lemoine, J. Simon, *Effect of rugosity on a flow governed by Navier-Stokes equations* , à paraître dans Quaterly of Appl. Maths 2001.

[6] D.M. Bushnell, K.J. Moore, *Drag reduction in nature* , Ann. Rev. Fluid Mech. 23(1991), p. 65-79.

[7] I. Cotoi, *Etude asymptotique de l'écoulement d'un fluide visqueux incompressible entre une plaque lisse et une paroi rugueuse* , thèse doctorale , Université Blaise Pascal, Clermont-Ferrand, janvier 2000.

[8] W. Jäger, A. Mikelić, *Homogenization of the Laplace equation in a partially perforated domain* , prépublication no. 157, Equipe d'Analyse Numérique Lyon-St-Etienne, septembre 1993, publiée dans " *Homogenization, In Memory of Serguei Kozlov* " , eds. V. Berdichevsky, V. Jikov and G. Papanicolaou, p. 259-284, Word Scientific, Singapore, 1999.

[9] W. Jäger, A. Mikelić , *On the Boundary Conditions at the Contact Interface between a Porous Medium and a Free Fluid*, Annali della Scuola Normale Superiore di Pisa, Classe Fisiche e Matematiche - Serie IV 23 (1996), Fasc. 3, p. 403 - 465.

[10] W. Jäger, A. Mikelić , *On the roughness-induced effective boundary conditions for a viscous flow*, J. of Differential Equations, 170(2001), p. 96-122.

[11] W. Jäger, A. Mikelić , N. Neuß, *Asymptotic analysis of the laminar viscous flow over a porous bed*, SIAM J. on Scientific and Statistical Computing, 22(2001), p. 2006-2028.

[12] W. Jäger, A. Mikelić , *Turbulent Couette Flows over a Rough Boundary and Drag Reduction*, prépublication, IWR/ SFB 359, Universität Heidelberg, Allemagne, 2001-44, novembre 2001.

[13] W. Jäger, A. Mikelić , *Derivation of nonlinear wall laws and applications*, en préparation.

[14] A. Mikelić , *Homogenization theory and applications to filtration through porous media*, chapitre dans *Filtration in Porous Media and Industrial Applications* , by M. Espedal, A.Fasano and A. Mikelić, Lecture Notes in Mathematics Vol. 1734, Springer-Verlag, 2000, p. 127-214.

[15] B. Mohammadi, O. Pironneau, F. Valentin, *Rough Boundaries and Wall Laws*, Int. J. Numer. Meth. Fluids, 27 (1998), p. 169-177.

[16] C. L. M. H. Navier, *Sur les lois de l'équilibre et du mouvement des corps élastiques*, Mem. Acad. R. Sci. Inst. France , 369 (1827).

[17] R. L. Panton, *Incompressible Flow* , John Wiley and Sons, New York, 1984.

[18] H. Schlichting, K. Gersten, *Boundary-Layer Theory*, 8th Revised and Enlarged Edition, Springer-Verlag, Berlin, 2000.

Deuxième partie

Ecoulements complexes, théorie mathématique des processus de dissolution/précipitation en milieux poreux

En toile de fond de la discussion, et partiellement abordée dans les chapitres précédents, le problème des écoulements complexes, c.a.d., le couplage de différents phénomènes physiques (au sens large) sera le sujet de notre propos dans cette partie. Nous avons traité précédemment des développements multi-échelles appliqués à la loi de Navier et ici nous nous intéresserons aux écoulements et aux interactions chimiques. L'argumentation sera divisée en deux chapitres. Le premier verra l'homogénéisation de l'écoulement d'un soluté dans un tube, et le second l'étude numérique à l'échelle microscopique du modèle de Duijn-Knabner[1]. Dans les deux cas la réaction chimique sera localisée sur la surface de la matrice poreuse, les parois internes du tube pour le premier et un grain avec une géométrie spécifique pour le second.

Les deux approches sont motivées par le programme sur le stockage des déchets nucléaires MOMAS[2].

La première discussion porte sur la dérivation rigoureuse par une technique de perturbation singulière du modèle efficace pour une diffusion améliorée à travers un étroit et long tube en deux dimensions. Nous commençons par un modèle à l'échelle du pore pour le transport d'un soluté réactif dans l'espace du pore par convection-diffusion. Les pores contiennent initialement une substance soluble à une concentration donnée et cette même substance à une concentration différente est injectée en $x = 0$. Les particules du soluté subissent une réaction chimique du premier ordre à la surface des pores. Nous nous plaçons dans les conditions de l'étude de Taylor en présence (cette fois-ci en plus) de réaction chimique. Le comportement à grande échelle pour des nombres de Péclet et Damkohler importants, en utilisant le rapport entre les temps caractéristiques transversal et longitudinal comme petit paramètre est donné. De plus, nous donnons une justification mathématique rigoureuse du comportement efficace comme une approximation physique du problème. L'estimation d'erreur est obtenue, tout d'abord dans la norme d'énergie, et enfin dans les normes L^1 et L^2. Elles garantissent la validité du modèle homogénéisé. Comme cas particulier nous retrouvons la formule très connue de la dispersion de Taylor. À notre connaissance, c'est la première fois que la formule de la dispersion de Taylor est justifiée d'une façon mathématique rigoureuse.

Dans la seconde discussion, nous poursuivons les investigations de C.J. van Duijn et al., où l'analyse d'un schéma numérique du modèle microscopique décrivant le transport d'ions

[1]P. Knabner, C. J. van Duijn, S. Hengst, An analysis of crystal dissolution fronts in flows through porous media. Part 1 : Compatible boundary conditions, Adv. Water Res. **18** (1995), 171–185.

[2]Gdr MOMAS (Modélisation Mathématique et Simulations liées aux problèmes de gestion des déchets nucléaires : 2439 - ANDRA, BRGM, CEA, EDF, CNRS) comme une partie du projet "Modélisation micro-macro des phénomènes couplés de transport-chimie-déformation en milieux argileux"

par un écoulement de soluté dans un milieu poreux est présenté. Ce soluté subit une réaction de précipitation/dissolution à la surface du pore. Nous discutons la méthode numérique. Nous nous intéressons ici à la chimie, qui est modélisée à travers une condition aux limites sous la forme d'une inclusion différentielle ordinaire. Une méthode semi-implicite en temps combinée à une approche régularisante est construite pour donner un schéma numérique stable et convergent. Pour se départir de l'apparition de la discontinuité non-linéaire en temps nous proposons une procédure itérative de point fixe. Pour achever cette étude, nous montrerons des résultats de simulations numériques.

Chapitre 4

Rigorous upscaling of a reactive flow through a pore, under important Peclet's and Damkohler's numbers

Sommaire

4.1 Introduction 62

4.2 Study of the the upscaled diffusion-convection equation on the half-line 66

4.3 A simple L^2 error estimate 68

4.4 The formal 2-scale expansion leading to Taylor's dispersion . . . 71

4.5 Boundary layer 73

4.6 First Correction 73

4.7 Error estimate involving the second order in expansion 83

Cette section a été l'objet de la publication référencée suivante :

A. Mikelić, V. Devigne, C.J. van Duijn "Rigorous upscaling of a reactive flow through a pore, under important Peclet's and Damkholer's number", CASA report no. 19, May 2005,

`http ://www.win.tue.nl/casa/research/casareports/2005.html`, submitted for publication.

et du proceeding

C. Rosier, V. Devigne, A. Mikelić "Rigorous upscaling of the reactive flow through a pore, under dominant Peclet number and infinite adsorption rate" proceeding for the conference Fourth Conference on Applied Mathematics and Scientific Computing, June 19-24, 2005 Brijuni island, Croatia.

4.1 Introduction

We study the diffusion of the solute particles transported by the Poiseuille velocity profile in a semi-infinite 2D channel. Solute particles are participants in a first-order chemical reaction with the boundary of the channel. They don't interact between them. The simplest example, borrowed from [8], is described by the following model for the solute concentration c^*

$$\frac{\partial c^*}{\partial t^*} + q(z)\frac{\partial c^*}{\partial x^*} - D^*\frac{\partial^2 c^*}{\partial (x^*)^2} - D^*\frac{\partial^2 c^*}{\partial z^2} = 0 \quad \text{in } \Omega^* = (0,+\infty)\times(-H,H), \tag{1.1}$$

where $q(z) = Q^*(1-(z/H)^2)$ and Q^* (velocity) and D^* (molecular diffusion) are positive constants. At the lateral boundaries $z = \pm H$ the first-order chemical reaction with the solute particles is modeled through the following boundary condition :

$$D^*\partial_z c^* + k^* c^* = 0 \quad \text{on } z = \pm H, \tag{1.2}$$

where k^* is the surface reactivity coefficient.

The natural way of analyzing this problem is to introduce the appropriate scales. They would come from the characteristic concentration $\hat{c}$, the characteristic length L_R, the characteristic velocity Q_R, the characteristic diffusivity D_R and the characteristic time T_c. The characteristic length L_R coincides in fact with the " observation distance ". Set now

$$c = \frac{c^*}{\hat{c}},\ x = \frac{x^*}{L_R},\ y = \frac{z}{H},\ t = \frac{t^*}{T_c},\ Q = \frac{Q^*}{Q_R},\ D = \frac{D^*}{D_R},\ k_0 = \frac{k^*}{k_R} \tag{1.3}$$

Then

$$\Omega = (0,+\infty)\times(-1,1), \quad \Gamma_w = (0,+\infty)\times\{-1,1\}. \tag{1.4}$$

Then the dimensionless form of (1.2) reads

$$\frac{\partial c}{\partial t} + \frac{Q_R T_c}{L_R} Q(1-y^2)\frac{\partial c}{\partial x} - \frac{D_R T_c}{L_R^2} D\partial_{xx} c - \frac{D_R T_c}{H^2} D\partial_{yy} c = 0 \quad \text{in } \Omega\times(0,T) \tag{1.5}$$

On Γ_w we impose the condition (1.2)

$$-\frac{D D_R T_c}{H^2}\frac{\partial c}{\partial y} = k\frac{T_c}{H}c \quad \text{on } \Gamma_w\times(0,T). \tag{1.6}$$

Problem involves the following time scales :

T_L = characteristic longitudinal time scale = $\dfrac{L_R}{Q_R}$

T_T = characteristic transversal time scale = $\dfrac{H^2}{D_R}$

T_R = superficial chemical reaction time scale = $\dfrac{H}{k_R}$

and the following characteristic non-dimensional numbers

Pe $= \dfrac{L_R Q_R}{D_R}$ (Peclet's number)

Da $= \dfrac{L_R^2 k_R}{H D_R}$ (Damkohler's number)

Further we set $\varepsilon = \dfrac{H}{L_R} << 1$ and choose $T_c = T_L$. Solving the full problem for arbitrary values of coefficients is costly and practically impossible. Consequently, one would like to find the effective (or averaged) values of the dispersion coefficient and the transport velocity and an effective corresponding 1D parabolic equation for the effective concentration.

In the absence of the chemical reaction, in [15] Taylor obtained, for the equation (1.1) describing only a diffusive transport of a passive scalar, an explicit effective expression for the enhanced diffusion coefficient. It is called in literature *Taylor's dispersion formula*.

The problem studied by Sir Taylor could be easily embedded in our setting. We choose $Q = \mathcal{O}(1)$, and $\dfrac{T_T}{T_L} = \dfrac{HQ_R}{D_R}\varepsilon = \mathcal{O}(\varepsilon^{2-\alpha}) = \varepsilon^2$ **Pe**. Then the situation from Taylor's article [15] corresponds to the case when $\alpha = 1$, i.e. Peclet's number is equal to $\dfrac{1}{\varepsilon}$, and $k_0 = 0$. Our equations in their non-dimensional form are

$$\frac{\partial c}{\partial t} + Q(1-y^2)\frac{\partial c}{\partial x} = D\varepsilon^{\alpha}\partial_{xx}c + D\varepsilon^{\alpha-2}\partial_{yy}c \quad \text{in } I\!R_+ \times (0,1) \times (0,T) \tag{1.7}$$

$$c(x,y,0) = 1, \quad (x,y) \in I\!R_+ \times (0,1), \tag{1.8}$$

$$-D\varepsilon^{\alpha-2}\partial_y c|_{y=1} = -D\frac{1}{\varepsilon^2 \mathbf{Pe}}\partial_y c|_{y=1} = k_0 \frac{\mathbf{Da}}{\mathbf{Pe}} c|_{y=1} = k_0 \varepsilon^{\alpha+\beta} c|_{y=1} \tag{1.9}$$

$$\partial_y c(x,0,t) = 0, \ (x,t) \in I\!R_+ \times (0,T) \tag{1.10}$$

$$\text{and} \quad c(0,y,t) = 0, \ (y,t) \in (0,1) \times (0,T), \tag{1.11}$$

where it was used that c is antisymmetric in y and Damkohler's number was set to ε^{β}. Our domain is now the infinite strip $Z^+ = I\!R_+ \times (0,1)$. We study the behavior of the solution to (1.7) -(1.11), with square integrable gradient in x and y, when $\varepsilon \to 0$. Clearly, the most interesting case is $\boxed{\beta = -\alpha}$ and $\boxed{0 \le \alpha < 2}$ and we restrict our considerations to this situation.

In this paper we prove that the correct upscaling of the problem (1.7)-(1.11) gives the following 1D parabolic problem

$$\begin{cases} \partial_t c^{Mau} + Q\big(\frac{2}{3} + \frac{4k_0}{45D}\varepsilon^{2-\alpha}\big)\partial_x c^{Mau} + k_0\big(1 - \frac{k_0}{3D}\varepsilon^{2-\alpha}\big)c^{Mau} = \\ (D\varepsilon^{\alpha} + \frac{8}{945}\frac{Q^2}{D}\varepsilon^{2-\alpha})\partial_{xx}c^{Mau} \ \text{in } I\!R_+ \times (0,T) \\ c^{Mau}|_{x=0} = 0, \ c^{Mau}|_{t=0} = 1, \ \partial_x c^{Mau} \in L^2(I\!R_+ \times (0,T)). \end{cases} \tag{EFF}$$

We note that for $k_0 = 0$ and $\alpha = 1$ this is exactly the effective model of Sir Taylor.

What is known concerning the derivation of the effective problem (EFF), with or without chemical reactions ?

$\diamond$ In the absence of the chemical reactions, there is a formal derivation by R. Aris, using the method of moments. For more details see [1].

$\diamond$ There have been numerous attempts at providing a rigorous justification for the approximation in absence of the chemical reactions. The most convincing is the " *near rigorous* " derivation using the centre manifold theory by G.N. Mercer and A.J. Roberts. For details see [9] , where the initial value problem is studied and the Fourier transform with respect to x is applied. The resulting PDE is written in the form $\dot{u} = \mathcal{A}u + F(u)$, with $u = (k, \hat{c})$. Then the centre manifold theory is applied to obtain effective equations at various orders. Since the corresponding centre manifold isn't finite dimensional, the results aren't rigorous.

$\diamond$ When the chemistry is added (e.g. having an irreversible, 1st order, chemical reaction with equilibrium at $y = 1$, as we have), then there is a paper [11] by M.A. Paine, R.G. Carbonell and S. Whitaker. The authors use the "single-point" closure schemes of turbulence modeling by Launder to obtain a closed model for the averaged concentration.

Hence the mentioned analysis don't provide a rigorous mathematical derivation of the Taylor's dispersion formula and in the presence of the chemical reactions it is even not clear how to average the starting problem.

It should be noted that the real interest is in obtaining *dispersion equations* for reactive flows through porous media. If we consider a porous medium comprised of a bundle of capillary tubes, then we come to our problem. The disadvantage is that a bundle of capillary tubes represents a geometrically oversimplified model of a porous medium. Nevertheless, there is considerable insight to be gained from the analysis of our toy problem.

Our technique is strongly motivated by the paper [14] by J. Rubinstein and R. Mauri, where effective dispersion and convection in porous media is studied, using the homogenization technique. Their analysis is based on the hierarchy of time scales and in getting the dimensionless equations we follow their approach. In our knowledge the only rigorous result concerning the effective dispersion in porous media, in the presence of high Peclet's numbers (and no chemistry), is in the recent paper [2] by A. Bourgeat, M. Jurak and A.L. Piatnitski. Nevertheless, their approach is based on the regular solutions for compatible data for the underlying linear transport equation. They suppose a high order compatibility between the initial and boundary data, involving derivatives up to order five, construct a smooth solution to the linear transport equation and then add the appropriate boundary layer, initial layer and the correction due to the perturbation of the mean flow. The effective solution obtained on this way is an

H^1-approximation of order ε and an L^2-approximation of order ε^2. Nevertheless, in problems involving chemistry, it is important to have a jump between the initial values of the concentration and the values imposed at the injection boundary $x = 0$. This is the situation from [15] and simply the compatibility of the data isn't acceptable for the reactive transport.

Homogenization of a problem with dissolution/precipitation at the grain boundaries in porous media, for small Peclet's number, ($\alpha = 0$) is in [3].

For the bounds on convection enhanced diffusion in porous media we refer to papers by Fannjiang, Papanicolaou, Zhikov, Kozlov, Piatnitskii ....

Plan of the paper is the following : In the section 2 we study the homogenized problem. It turns out that it has an explicit solution having the form of moving Gaussian as the 1D boundary layers of parabolic equations, when viscosity goes to zero (see [6]). Its behavior with respect to ε and t is very singular.

Then in section 3 we give a justification of a lower order approximation, using a simple energy argument. In fact such approximation doesn't use Taylor's dispersion formula and gives an error of the same order in $L^\infty(L^2)$ as the solution to the linear transport equation. Furthermore, when $\alpha > 4/3$ this approach doesn't give an approximation any more !

In the section 4 we give a formal derivation of the upscaled problem (EFF), using the approach from [14].

Construction of the spatial boundary layer taking care of the injection boundary is in Section 5.

Then in sections 6 and 7 we prove that the effective concentration satisfying the corresponding 1D parabolic problem, with Taylor's diffusion coefficient and the reactive correction, is an approximation in $L^\infty(L^2)$ and in $L^\infty(L^\infty)$ for the physical concentration.

To satisfy the curiosity of the reader not familiar with the singular perturbation techniques, we give here the simplified version of the results stated in Theorems 4.7.1, 4.7.2 and 4.7.3 in Section 7. For simplicity, we compare only the physical concentration c^ε with c^{Mau}. Keeping the correction terms is necessary in order to have the precision reached in Theorems 4.7.1, 4.7.2 and 4.7.3, Section 7. Throughout the paper $H(x)$ is Heaviside's function.

Theorem 4.1.1 *Let c^{Mau} be given by (EFF). Then we have*

$$\|t^3(c^\varepsilon - c^{Mau})\|_{L^\infty((0,T)\times Z^+)} \leq \begin{cases} C\varepsilon^{2-3\alpha/2}, & \textit{if } \alpha < 1, \\ C\varepsilon^{3/2-\alpha-\delta}, \ \forall \delta > 0, & \textit{if } 2 > \alpha \geq 1. \end{cases} \tag{1.12}$$

$$\|t^3\big(c^\varepsilon - c^{Mau}\big)\|_{L^2(0,T;L^1_{loc}(Z^+))} \leq C\varepsilon^{2-\alpha} \tag{1.13}$$

$$\|t^3\big(c^\varepsilon - c^{Mau}\big)\|_{L^2(0,T;L^2_{loc}(Z^+)} \leq C\big(\varepsilon^{2-5\alpha/4}H(1-\alpha) + \varepsilon^{3/2-3\alpha/4}H(\alpha-1)\big) \tag{1.14}$$

$$\|t^3\partial_y c^\varepsilon\|_{L^2(0,T;L^2_{loc}(Z^+))} \leq C\big(\varepsilon^{2-5\alpha/4}H(1-\alpha) + \varepsilon^{3/2-3\alpha/4}H(\alpha-1)\big) \tag{1.15}$$

$$\|t^3\partial_x\big(c^\varepsilon - c^{Mau}\big)\|_{L^2(0,T;L^2_{loc}(Z^+))} \leq C\big(\varepsilon^{2-7\alpha/4}H(1-\alpha) + \varepsilon^{3/2-5\alpha/4}H(\alpha-1)\big) \tag{1.16}$$

Corollary 1 *In the conditions of Taylor's article [15], $\alpha = 1$ and $k_0 = 0$, we have*

$$\|t^3(c^\varepsilon - c^{Mau})\|_{L^\infty((0,T)\times Z^+)} \leq C\varepsilon^{1/2-\delta}, \quad \forall \delta > 0, \tag{1.17}$$

$$\|t^3\big(c^\varepsilon - c^{Mau}\big)\|_{L^2(0,T;L^1_{loc}(Z^+))} \leq C\varepsilon \tag{1.18}$$

Our result could be restated in dimensional form :

Theorem 4.1.2 *Let us suppose that $L_R > \max\{D_R/Q_R, Q_R H^2/D_R, H\}$. Then the upscaled dimensional approximation for (1.1) reads*

$$\frac{\partial c^{*,eff}}{\partial t^*} + \big(\frac{2}{3} + \frac{4}{45}\mathbf{Da}_T\big)Q^*\frac{\partial c^{*,eff}}{\partial x^*} + \frac{k^*}{H}\big(1 - \frac{1}{3}\mathbf{Da}_T\big)c^{*,eff} = D^*\Big(1 + \frac{8}{945}\mathbf{Pe}_T^2\Big)\frac{\partial^2 c^{*,eff}}{\partial (x^*)^2}, \tag{1.19}$$

where $\mathbf{Pe}_T = \dfrac{Q^* H}{D^*}$ is the transversal Peclet's number and $\mathbf{Da}_T = \dfrac{k^* H}{D^*}$ is the transversal Damkohler's number.

Finally, let us note that in the known literature on boundary layers for parabolic regularization, the transport velocity is supposed to be zero at the injection boundary (see [5]) and our result doesn't enter into existing framework.

One could try to get even higher order approximations. Unfortunately, our procedure from Section 4 then leads to higher order differential operators and it is not clear if they are easy to handle. In the absence of the boundaries, determination of the higher order terms using the computer program REDUCE was undertaken in [9].

4.2 Study of the the upscaled diffusion-convection equation on the half-line

For $\bar{Q}, \bar{D}, \varepsilon > 0$ and $\bar{k} \geq 0$, we consider the problem

$$\begin{cases} \partial_t u + \bar{Q}\partial_x u + \bar{k}u = \gamma\bar{D}\partial_{xx}u \ \text{ in } \mathbb{R}_+ \times (0,T), & \partial_x u \in L^2(\mathbb{R}_+ \times (0,T)) \\ u(x,0) = 1 \ \text{ in } \mathbb{R}_+, & u(0,t) = 0 \ \text{ at } x = 0. \end{cases} \tag{2.20}$$

The unique solution is given by the following explicit formula

$$u(x,t) = e^{-\bar{k}t}\Big(1 - \frac{1}{\sqrt{\pi}}\left[e^{\frac{\bar{Q}x}{\gamma \bar{D}}} \int_{\frac{x+t\bar{Q}}{2\sqrt{\gamma \bar{D} t}}}^{+\infty} e^{-\eta^2}\, d\eta + \int_{\frac{x-t\bar{Q}}{2\sqrt{\gamma \bar{D} t}}}^{+\infty} e^{-\eta^2}\, d\eta \right]\Big) \tag{2.21}$$

The explicit formula allows us to find the exact behavior of u with respect to γ. We note that for $\alpha \in [0,1]$, we will set $\gamma = \varepsilon^{\alpha}$. For $\alpha \in [1,2)$, we choose $\gamma = \varepsilon^{2-\alpha}$. The derivatives of u are found using the computer program MAPLE and then their norms are estimated. Since the procedure is standard, we don't give the details. In more general situations there are no explicit solutions and these estimates could be obtained using the technique and results from [6].

$\boxed{\text{1. STEP}}$ First, by the maximum principle we have

$$0 \le u(x,t;\gamma) = u(x,t) \le 1 \tag{2.22}$$

$\boxed{\text{2. STEP}}$ Next we estimate the difference between $\chi_{x<\bar{Q}t}$ and u. We have

Lemma 4.2.1 *Function u satisfies the estimates*

$$\int_0^{\infty} |e^{-\bar{k}t}\chi_{\{x>\bar{Q}t\}} - u(t,x)|\, dx = 3\sqrt{\gamma \bar{D} t}e^{-\bar{k}t} + C\gamma \tag{2.23}$$

$$\|e^{-\bar{k}t}\chi_{\{x>\bar{Q}t\}} - u\|_{L^{\infty}(0,T;L^p(I\!R_+))} \le C\gamma^{1/(2p)}, \quad \forall p \in (1,+\infty). \tag{2.24}$$

$\boxed{\text{3. STEP}}$ For the derivatives of u the following estimates hold

Lemma 4.2.2 *Let ζ be defined by*

$$\zeta(t) = \begin{cases} (\dfrac{t}{\bar{D}\gamma})^r & \textit{for } 0 \le t \le \bar{D}\gamma, \\ 1 & \textit{otherwise}, \end{cases} \tag{2.25}$$

with $r \ge q \ge 1$. Then we have

$$\|\zeta(t)\partial_t u\|_{L^q((0,T)\times I\!R_+)} + \|\zeta(t)\partial_x u\|_{L^q((0,T)\times I\!R_+)} \le C(\gamma\bar{D})^{\min\{1/(2q)-1/2,2/q-1\}},\ q \neq 3 \tag{2.26}$$

$$\|\zeta(t)\partial_t u\|_{L^3((0,T)\times I\!R_+)} + \|\zeta(t)\partial_x u\|_{L^3((0,T)\times I\!R_+)} \le C\big((\gamma\bar{D})^{-1}\log(\frac{1}{\gamma\bar{D}})\big)^{1/3} \tag{2.27}$$

$\boxed{\text{4. STEP}}$

Now we estimate the second derivatives :

Lemma 4.2.3 *Let ζ be defined by (2.25). Then the second derivatives of u satisfy the estimates*

$$\begin{aligned} &\|\zeta(t)u_{tt}\|_{L^q((0,T)\times I\!R_+)} + \|\zeta(t)u_{tx}\|_{L^q((0,T)\times I\!R_+)} + \|\zeta(t)u_{xx}\|_{L^q((0,T)\times I\!R_+)} \\ &\qquad \le C_q(\gamma\bar{D})^{\min\{1/(2q)-1,2/q-2\}},\ q \neq 3/2 \end{aligned} \tag{2.28}$$

$$\begin{aligned} &\|\zeta(t)u_{tt}\|_{L^{3/2}((0,T)\times I\!R_+)} + \|\zeta(t)u_{tx}\|_{L^{3/2}((0,T)\times I\!R_+)} + \|\zeta(t)u_{xx}\|_{L^{3/2}((0,T)\times I\!R_+)} \\ &\qquad \le C\big((\gamma\bar{D})^{-1}\log(\frac{1}{\gamma\bar{D}})\big)^{2/3} \end{aligned} \tag{2.29}$$

$\boxed{\text{5. STEP}}$

For the 3rd order derivatives we have :

Lemma 4.2.4 *Let ζ be defined by (2.25) . Then*

$$\|\partial_{xxx}(\zeta(t)u)\|_{L^q((0,T)\times I\!R_+)} + \|\zeta(t)\partial_{xxt}u\|_{L^q((0,T)\times I\!R_+)}$$
$$+\|\zeta(t)\partial_{xtt}u\|_{L^q((0,T)\times I\!R_+)} \leq C_q(\gamma\bar{D})^{2/q-3},\ q>1 \tag{2.30}$$
$$\|\partial_{xxx}(\zeta(t)u)\|_{L^1((0,T)\times I\!R_+)} + \|\zeta(t)\partial_{xxt}u\|_{L^1((0,T)\times I\!R_+)}$$
$$+\|\zeta(t)\partial_{xtt}u\|_{L^1((0,T)\times I\!R_+)} \leq C_1(\gamma\bar{D})^{-1}\log\frac{1}{\gamma\bar{D}} \tag{2.31}$$

4.3 A simple L^2 error estimate

The simplest way to average the problem (1.7)-(1.11) is to take the mean value with respect to y. Supposing that the mean of the product is the product of the means, which is in general wrong, we get the following problem for the " averaged " concentration $c_0^{eff}(x,t)$:

$$\left\{\begin{array}{l} \dfrac{\partial c_0^{eff}}{\partial t} + \dfrac{2Q}{3}\dfrac{\partial c_0^{eff}}{\partial x} + k_0 c_0^{eff} = \varepsilon^\alpha D \dfrac{\partial c_0^{eff}}{\partial x^2}\ \text{ in } I\!R_+\times(0,T), \\ \\ \partial_x c_0^{eff}\in L^2(I\!R_+\times(0,T)),\quad c_0^{eff}|_{t=0}=1,\quad c_0^{eff}|_{x=0}=0. \end{array}\right. \tag{3.32}$$

This is the problem (2.20) with $\tilde{Q}=\dfrac{2}{3}Q$, $\bar{k}=k_0$ and $\bar{D}=D$. The small parameter γ is equal to ε^α. Let the operator $\mathcal{L}^\varepsilon$ be given by

$$\mathcal{L}^\varepsilon\zeta = \frac{\partial\zeta}{\partial t} + Q(1-y^2)\frac{\partial\zeta}{\partial x} - D\varepsilon^\alpha\Big(\partial_{xx}\zeta + \varepsilon^{-2}\partial_{yy}\zeta\Big) \tag{3.33}$$

The non-dimensional physical concentration c^ε satisfies (1.7)-(1.11), i.e.

$$\mathcal{L}^\varepsilon c^\varepsilon = 0 \quad \text{in } I\!R_+\times(0,1)\times(0,T) \tag{3.34}$$
$$c^\varepsilon(0,y,t)=0 \quad \text{on } (0,1)\times(0,T) \tag{3.35}$$
$$\partial_y c^\varepsilon(x,0,t)=0 \quad \text{on } I\!R_+\times(0,T) \tag{3.36}$$
$$-D\varepsilon^{\alpha-2}\partial_y c^\varepsilon(x,1,t) = k_0 c^\varepsilon(x,1,t) \quad \text{on } I\!R_+\times(0,T) \tag{3.37}$$
$$c^\varepsilon(x,y,0)=1 \quad \text{on } I\!R_+\times(0,1) \tag{3.38}$$

We want to approximate c^ε by c_0^{eff}. Then

$$\mathcal{L}^\varepsilon(c_0^{eff}) = -k_0 c_0^{eff} + Q\partial_x c_0^{eff}(1/3-y^2) = R^\varepsilon$$
$$\mathcal{L}^\varepsilon(c^\varepsilon - c_0^{eff}) = -R^\varepsilon \ \text{ in } I\!R_+\times(0,1)\times(0,T) \quad \text{and} \tag{3.39}$$
$$-D\varepsilon^{\alpha-2}\partial_y(c^\varepsilon(x,1,t)-c_0^{eff}) = k_0 c^\varepsilon(x,1,t) \quad \text{on } I\!R_+\times(0,T) \tag{3.40}$$

Let $\Psi(x) = 1/(x+1)$. Then $(\partial_x \Psi^2)^2/\Psi^2 \leq 4\Psi^2$. We have the following proposition, which will be useful in getting the estimates :

Proposition 4.3.1 *Let g^ε, ξ_0^ε and R^ε be such that $\Psi g^\varepsilon \in H^1(Z^+ \times (0,T))$, $\Psi \xi_0^\varepsilon \in L^2(Z^+)$ and $\Psi R^\varepsilon \in L^2(Z^+ \times (0,T))$. Let ξ, $\Psi\xi \in C([0,T]; L^2(Z^+))$, $\Psi \nabla_{x,y}\xi \in L^2(Z^+ \times (0,T))$, be a bounded function which satisfies the system*

$$\mathcal{L}^\varepsilon(\xi) = -R^\varepsilon \ \ in \ \ Z^+ \times (0,T) \tag{3.41}$$

$$-D\varepsilon^{\alpha-2}\partial_y \xi|_{y=1} = k_0 \xi|_{y=1} + g^\varepsilon|_{y=1} \ and \ \partial_y \xi|_{y=0} = 0 \ on \ I\!R_+ \times (0,T) \tag{3.42}$$

$$\xi|_{t=0} = \xi_0^\varepsilon \quad on \ \ Z^+ \quad and \ \xi|_{x=0} = 0 \ on \ \ (0,1)\times(0,T). \tag{3.43}$$

Then we have the following energy estimate

$$\mathcal{E}(\xi,t) = \frac{1}{2}\int_{Z^+} \Psi(x)^2 \xi^2(t)\, dxdy + \frac{D}{2}\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi(x)^2 \Big\{ \varepsilon^{-2}|\partial_y \xi|^2 +$$
$$|\partial_x \xi|^2 \Big\}\, dxdyd\tau + k_0 \int_0^t \int_{I\!R_+} \xi^2|_{y=1} \Psi^2(x)\, dxd\tau \leq -\int_0^t \int_{Z^+} \Psi(x)^2 R^\varepsilon \xi\, dxdyd\tau -$$
$$\int_0^t \int_{I\!R_+} g^\varepsilon|_{y=1} \xi|_{y=1} \Psi^2(x)\, dxd\tau + 2D\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi(x)^2 \xi^2\, dxdyd\tau. \tag{3.44}$$

Proof. We test (3.41)-(3.43) by $\Psi^2(x)\xi$ and get

$$\frac{1}{2}\int_{Z^+} \xi^2(t)\Psi^2(x)\, dxdy + D\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi^2(x) \Big\{ \varepsilon^{-2}|\partial_y \xi|^2 + |\partial_x \xi|^2 \Big\}\, dxdyd\tau +$$
$$k_0 \int_0^t \int_0^{+\infty} \xi^2|_{y=1} \Psi^2\, dxd\tau \leq \frac{1}{2}\int_{Z^+} (\xi_0^\varepsilon)^2 \Psi^2(x)\, dxdy -$$
$$k_0 \int_0^t \int_0^{+\infty} (g^\varepsilon \xi)|_{y=1} \Psi^2\, dxd\tau - D\varepsilon^\alpha \int_0^t \int_{Z^+} \partial_x \xi \xi \partial_x \Psi^2(x)\, dxdyd\tau. \tag{3.45}$$

Next, we use that

$$D\varepsilon^\alpha \int_0^t \int_{Z^+} \partial_x \xi \xi \partial_x \Psi^2(x)\, dxd\tau \leq \frac{D}{2}\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi^2(x) |\partial_x \xi|^2\, dxdyd\tau$$
$$+2D\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi^2(x)|\xi|^2\, dxdyd\tau \tag{3.46}$$

and get (3.44). ■ This simple proposition allows us to prove

Proposition 4.3.2 *In the setting of this section we have*

$$\|\Psi(x)(c^\varepsilon - c_0^{eff})\|_{L^\infty(0,T;L^2(I\!R_+\times(0,1))} \leq \varepsilon^{1-\alpha/2}\frac{F^0}{\sqrt{D}} \tag{3.47}$$

$$\|\Psi(x)\partial_x(c^\varepsilon - c_0^{eff})\|_{L^2(0,T;L^2(I\!R_+\times(0,1))} \leq \varepsilon^{1-\alpha}\frac{F^0}{D} \tag{3.48}$$

$$\|\Psi(x)\partial_y(c^\varepsilon - c_0^{eff})\|_{L^2(0,T;L^2(I\!R_+\times(0,1))} \leq \varepsilon^{2-\alpha}\frac{F^0}{D}, \tag{3.49}$$

$$where \quad F^0 = C_1^F \|\partial_x c_0^{eff}\|_{L^2(O_T)} + C_2^F k_0 \leq C_3^F \varepsilon^{-\alpha/4} \tag{3.50}$$

Proof. We are in the situation of Proposition 4.3.1 with $\xi_0^\varepsilon = 0$ and $g^\varepsilon = k_0 c_0^{eff}$. Consequently, for $\xi = c^\varepsilon - c_0^{eff}$ we have

$$\mathcal{E}(\xi, t) \leq k_0 \int_0^t \int_0^{+\infty} c_0^{eff} (\int_0^1 c^\varepsilon \, dy - c^\varepsilon|_{y=1}) \Psi^2 \, dx d\tau + 2D\varepsilon^\alpha \cdot \\ \int_0^t \int_{Z^+} |\xi|^2 \Psi^2(x) \, dx dy d\tau - \int_0^t \int_{Z^+} Q(1/3 - y^2) \xi \partial_x c_0^{eff} \Psi^2 \, dx dy d\tau. \tag{3.51}$$

It remains to estimate the first and the third term at the right hand side of (3.51). We have

$$|\int_0^t \int_{Z^+} Q \partial_x c_0^{eff} (1/3 - y^2) \xi \Psi^2(x) \, dx dy d\tau| = \\ |\int_0^t \int_{Z^+} Q \partial_x c_0^{eff} (y/3 - y^3/3) \partial_y \xi \Psi^2(x) \, dx dy d\tau| \tag{3.52}$$

$$\text{and } \ k_0 |\int_0^t \int_0^{+\infty} c_0^{eff} (\int_0^1 \xi \, dy - \xi|_{y=1}) \Psi^2 \, dx d\tau| \leq \frac{D}{8} \varepsilon^\alpha \int_0^t \int_{Z^+} \\ \Psi^2(x) |\partial_y \xi|^2 \, dx dy d\tau + \frac{k_0^2}{D} \varepsilon^{2-\alpha} \int_0^t \int_0^{+\infty} (c_0^{eff})^2 \Psi^2 \, dx d\tau. \tag{3.53}$$

After inserting (3.52)-(3.53) into (3.51) we get

$$\mathcal{E}(c^\varepsilon - c_0^{eff}, t) \leq \varepsilon^{2-\alpha} \int_0^t \int_0^{+\infty} \left\{ \frac{2k_0^2}{D} (c_0^{eff})^2 + \frac{32}{315} \frac{Q^2}{D} (\partial_x c_0^{eff})^2 \right\} \Psi^2 \, dx d\tau \\ + \int_0^t \int_0^1 \int_0^{+\infty} 2D\varepsilon^\alpha \Psi^2(x) (c^\varepsilon - c_0^{eff})^2 \, dx dy d\tau, \tag{3.54}$$

and after applying Gronwall's inequality, we obtain (3.47)-(3.49). ■

Corollary 2

$$\|c^\varepsilon - c_0^{eff}\|_{L^\infty(0,T;L^2_{loc}(I\!R_+ \times (0,1))} \leq C\varepsilon^{1-3\alpha/4} \tag{3.55}$$

Remark 8 *It is reasonable to expect some L^1 estimates with better powers for ε. Unfortunately, testing the equation (3.39) by the regularized sign $(c^\varepsilon - c_0^{eff})$, doesn't lead to anything useful. Hence at this stage claiming a $\sqrt{\varepsilon}$ estimate in L^1 is not justified.*

Remark 9 *There are recent papers by Grenier and Gues on singular perturbation problems. In [5] Grenier supposes that Q is zero as x at $x = 0$, together with its derivatives. Such hypothesis allows better estimates.*

Remark 10 *The estimate (2.23) implies that* $\exp\{-k_0 t\} \chi_{\{x > Qt\}}$ *is an approximation for the physical concentration which is of the same order in $L^\infty(L^2)$ as c_0^{eff}.*

Remark 11 *For $\alpha > 4/3$ the estimate (3.55) is without any value.*

4.4 The formal 2-scale expansion leading to Taylor's dispersion

The estimate obtained in the previous section isn't satisfactory. At the other hand, it is known that the Taylor dispersion model gives a very good 1D approximation. With this motivation we briefly explain how to obtain formally the higher precision effective models and notably the variant of Taylor's dispersion formula, by the 2-scale asymptotic expansion.

We start with the problem (3.34)-(3.38) and search for c^ε in the form

$$c^\varepsilon = c^0(x,t;\varepsilon) + \varepsilon c^1(x,y,t) + \varepsilon^2 c^2(x,y,t) + \dots \tag{4.56}$$

After introducing (4.56) into the equation (3.34) we get

$$\begin{aligned}&\varepsilon^0\Big\{\partial_t c^0 + Q(1-y^2)\partial_x c^0 - D\varepsilon^{\alpha-1}\partial_{yy}c^1\Big\} + \varepsilon\Big\{\partial_t c^1 + \\ &Q(1-y^2)\partial_x c^1 - D\varepsilon^{\alpha-1}\partial_{xx}c^0 - D\varepsilon^{\alpha-1}\partial_{yy}c^2\Big\} = \mathcal{O}(\varepsilon^2)\end{aligned} \tag{4.57}$$

In order to have (4.57) for every $\varepsilon \in (0,\varepsilon_0)$, all coefficients in front of the powers of ε should be zero.

The problem corresponding to the order ε^0 is

$$\begin{cases} -D\partial_{yy}c^1 = -\varepsilon^{1-\alpha}Q(1/3-y^2)\partial_x c^0 - \varepsilon^{1-\alpha}\big(\partial_t c^0 + 2Q\partial_x c^0/3\big) \ \text{ on } \ (0,1), \\ \partial_y c^1 = 0 \text{ on } \ y=0 \ \text{ and } \ -D\partial_y c^1 = k_0\varepsilon^{1-\alpha}c^0 \text{ on } \ y=1 \end{cases} \tag{4.58}$$

for every $(x,t) \in (0,+\infty)\times(0,T)$. By the Fredholm's alternative, the problem (4.58) has a solution if and only if

$$\partial_t c^0 + 2Q\partial_x c^0/3 + k_0 c^0 = 0 \quad \text{in } \ (0,L)\times(0,T). \tag{4.59}$$

Unfortunately our initial and boundary data are incompatible and the hyperbolic equation (4.59) has a discontinuous solution. Since the asymptotic expansion for c^ε involves derivatives of c^0, the equation (4.59) doesn't suit our needs. In [2] the difficulty was overcome by supposing compatible initial and boundary data. We proceed by following an idea from [14] and suppose that

$$\partial_t c^0 + 2Q\partial_x c^0/3 + k_0 c^0 = \mathcal{O}(\varepsilon) \quad \text{in } \ (0,+\infty)\times(0,T). \tag{4.60}$$

The hypothesis (4.60) will be justified *a posteriori*, after getting an equation for c^0.

Hence (4.58) reduces to

$$\begin{cases} -D\partial_{yy}c^1 = -\varepsilon^{1-\alpha}Q(1/3-y^2)\partial_x c^0 + \varepsilon^{1-\alpha}k_0 c^0 \text{ on } \ (0,1), \\ \partial_y c^1 = 0 \text{ on } \ y=0 \ \text{ and } \ -D\partial_y c^1 = k_0\varepsilon^{1-\alpha}c^0 \text{ on } \ y=1 \end{cases} \tag{4.61}$$

for every $(x,t) \in (0,+\infty) \times (0,T)$, and we have

$$c^1(x,y,t) = \varepsilon^{1-\alpha}\Big(\frac{Q}{D}(\frac{y^2}{6} - \frac{y^4}{12})\partial_x c^0 + \frac{k_0}{D}(\frac{1}{6} - \frac{y^2}{2})c^0 + C_0(x,t)\Big), \tag{4.62}$$

where C_0 is an arbitrary function.

Let us go to the next order. Then we have

$$\begin{cases} -D\partial_{yy}c^2 = -\varepsilon^{1-\alpha}Q(1-y^2)\partial_x c^1 + D\partial_{xx}c^0 - \varepsilon^{1-\alpha}\partial_t c^1 + D\varepsilon\partial_{xx}c^1 \\ -\varepsilon^{-\alpha}\big(\partial_t c^0 + 2Q\partial_x c^0/3 + k_0 c^0\big) \quad \text{on } (0,1), \\ \partial_y c^2 = 0 \text{ on } y=0 \text{ and } -D\partial_y c^2 = k_0\varepsilon^{1-\alpha}c^1 \text{ on } y=1 \end{cases} \tag{4.63}$$

for every $(x,t) \in (0,+\infty) \times (0,T)$. The problem (4.63) has a solution if and only if

$$\partial_t c^0 + 2Q\partial_x c^0/3 + k_0(c^0 + \varepsilon c^1|_{y=1}) + \varepsilon\partial_t(\int_0^1 c^1\,dy) - \varepsilon^\alpha D\partial_{xx}c^0 + \\ \varepsilon\partial_x(\int_0^1 (1-y^2)c^1\,dy) = 0 \quad \text{in } (0,+\infty)\times(0,T). \tag{4.64}$$

(4.64) is the equation for c^0 . In order to get the simplest possible equation for c^0 we choose C_0 giving $\int_0^1 c^1\,dy = 0$. Now c^1 takes the form

$$c^1(x,y,t) = \varepsilon^{1-\alpha}\Big(\frac{Q}{D}(\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180})\partial_x c^0 + \frac{k_0}{D}(\frac{1}{6} - \frac{y^2}{2})c^0\Big) \tag{4.65}$$

and the equation (4.64) becomes

$$\partial_t c^0 + Q\big(\frac{2}{3} + \frac{4k_0}{45D}\varepsilon^{2-\alpha}\big)\partial_x c^0 + k_0\big(1 - \frac{k_0}{3D}\varepsilon^{2-\alpha}\big)c^0 = \varepsilon^\alpha\tilde{D}\partial_{xx}c^0 \quad \text{in } (0,+\infty)\times(0,T). \tag{4.66}$$

with

$$\tilde{D} = D + \frac{8}{945}\frac{Q^2}{D}\varepsilon^{2(1-\alpha)} \tag{4.67}$$

Now the problem (4.63) becomes

$$\begin{cases} -D\partial_{yy}c^2 = \varepsilon^{2-2\alpha}\Big\{-\frac{Q^2}{D}\partial_{xx}c^0\Big\{\frac{8}{945} + (1-y^2)(\frac{y^2}{6} - \frac{y^4}{12} \\ -\frac{7}{180})\Big\} + \partial_x c^0\frac{Qk_0}{D}\Big\{\frac{2}{45} - (1-y^2)(\frac{1}{6} - \frac{y^2}{2})\Big\} + \frac{2k_0Q}{45D}\partial_x c^0 - \\ \frac{k_0^2}{3D}c^0 - (\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180})(\partial_{xt}c^0\frac{Q}{D} - \varepsilon^\alpha Q\partial_{xxx}c^0) - \\ (\frac{1}{6} - \frac{y^2}{2})(\partial_t c^0\frac{k_0}{D} - \varepsilon^\alpha k_0\partial_{xx}c^0)\Big\} \text{ on } (0,1), \quad \partial_y c^2 = 0 \text{ on } y=0 \\ \text{and } -D\partial_y c^2 = \frac{Qk_0}{D}\varepsilon^{2-2\alpha}\partial_x c^0\frac{2}{45} - \frac{k_0^2}{3D}\varepsilon^{2-2\alpha}c^0 \text{ on } y=1. \end{cases} \tag{4.68}$$

If we choose c^2 such that $\int_0^1 c^2\,dy = 0$, then

$$c^2(x,y,t) = \varepsilon^{2-2\alpha}\Big\{ -\frac{Q^2}{D^2}\partial_{xx}c^0\Big(\frac{281}{453600} + \frac{23}{1512}y^2 - \frac{37}{2160}y^4 + \frac{1}{120}y^6 - \frac{1}{672}y^8\Big) + (\frac{Q}{D^2}\partial_{xt}c^0 - \varepsilon^\alpha\frac{Q}{D}\partial_{xxx}c^0)\Big(\frac{31}{7560} - \frac{7}{360}y^2 + \frac{y^4}{72} - \frac{y^6}{360}\Big) + \frac{Qk_0}{D^2}\partial_x c^0(\frac{y^6}{60} - \frac{y^4}{18} + \frac{11y^2}{180} - \frac{11}{810}) + (\frac{k_0}{2D^2}\partial_t c^0 - \frac{k_0}{2D}\varepsilon^\alpha\partial_{xx}c^0)(- \frac{y^4}{12} + \frac{y^2}{6} - \frac{7}{180}) + (\frac{Qk_0}{45D^2}\partial_x c^0 - \frac{k_0^2}{6D^2}c^0)(\frac{1}{3} - y^2)\Big\} \tag{4.69}$$

4.5 Boundary layer

If we add corrections to c^0, the obtained function doesn't satisfy any more the boundary conditions. We correct the new values using the appropriate boundary layer.

Let $Z^+ = (0,+\infty)\times(0,1)$.

$$\begin{cases} -\Delta_{y,z}\beta = 0 \quad \text{for } (z,y)\in Z^+. \\ -\partial_y\beta = 0 \ \text{ for } y=1, \quad \text{and for } y=0, \\ \beta = \dfrac{y^2}{6} - \dfrac{y^4}{12} - \dfrac{7}{180} \ \text{ for } z=0. \end{cases} \tag{5.70}$$

Using the elementary variational theory for PDEs, we get the existence of a unique solution $\beta \in L^2_{loc}(Z^+)$ such that $\nabla\beta \in L^2_{loc}(Z^+)^2$. Next, we note that the average of the boundary value at $z=0$ is zero. This implies that $\int_0^1 \beta(z,y)\,dy = 0$, for every $z\in(0,+\infty)$. Now we can apply the Poincaré's inequality in H^1 :

$$\int_0^1 \varphi^2\,dy \le \frac{1}{\pi^2}\int_0^1 |\partial_y\varphi|^2\,dy,\ \forall\varphi\in H^1(0,1),\ \int_0^1 \varphi\,dy = 0, \tag{5.71}$$

and conclude that in fact $\beta\in H^1(Z^+)$. In order to prove that β represents a boundary layer, one should prove the exponential decay. We apply the theory from [10] and get the following result describing the decay of β as $z\to+\infty$:

Proposition 4.5.1 *There exists a constant $\gamma_0 > 0$ such that the solution β for (5.70) satisfies the estimates*

$$\int_z^{+\infty}\int_0^1 |\nabla_{y,z}\beta|^2\,dydz \le c_0 e^{-\gamma_0 z}, \quad z>0 \tag{5.72}$$

$$|\beta(y,z)| \le c_0 e^{-\gamma_0 z}, \quad \forall (y,z)\in Z^+ \tag{5.73}$$

4.6 First Correction

The estimate (3.55) isn't satisfactory. In order to get a better approximation we take the correction constructed using the formal 2-scale expansion in Section 4.

Let $0 \leq \alpha < 2$. We start by the $\mathcal{O}(\varepsilon^2)$ approximation and consider the function

$$c_1^{eff}(x,y,t;\varepsilon) = c^{Mau}(x,t;\varepsilon) + \varepsilon^{2-\alpha}\zeta(t)\Big(\frac{Q}{D}(\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180}) \cdot \frac{\partial c^{Mau}}{\partial x} + \frac{k_0}{D}(\frac{1}{6} - \frac{y^2}{2})c^{Mau}(x,t;\varepsilon)\Big) \tag{6.74}$$

where c^{Mau} is the solution to the effective problem with Taylor's dispersion coefficient and reaction terms :

$$\begin{cases} \partial_t c^{Mau} + Q\big(\frac{2}{3} + \frac{4k_0}{45D}\varepsilon^{2-\alpha}\big)\partial_x c^{Mau} + k_0\big(1 - \frac{k_0}{3D}\varepsilon^{2-\alpha}\big)c^{Mau} = \\ (D\varepsilon^\alpha + \frac{8}{945}\frac{Q^2}{D}\varepsilon^{2-\alpha})\partial_{xx}c^{Mau}, \text{ in } I\!R_+ \times (0,T) \\ c^{Mau}|_{x=0} = 0,\ c^{Mau}|_{t=0} = 1,\ \partial_x c^{Mau} \in L^2(I\!R_+ \times (0,T)), \end{cases} \tag{6.75}$$

$\tilde{D} = D\varepsilon^\alpha + \frac{8}{945}\frac{Q^2}{D}\varepsilon^{2-\alpha}$ is Taylor's dispersion coefficient. The cut-off in time ζ is given by (2.25) and we use to eliminate the time-like boundary layer appearing at $t = 0$. These effects are not visible in the formal expansion.

Let $\mathcal{L}^\varepsilon$ be the differential operator given by (3.33). Following the formal expansion from Section 4, we know that $\mathcal{L}^\varepsilon$ applied to the correction without boundary layer functions and cut-offs would give $F_1^\varepsilon + F_2^\varepsilon + F_3^\varepsilon + F_4^\varepsilon + F_5^\varepsilon$, where

$$\begin{cases} F_1^\varepsilon = \partial_{xx}c^{Mau}\frac{Q^2}{D}\varepsilon^{2-\alpha}\Big\{\frac{8}{945} + (1-y^2)(\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180})\Big\} \\ F_2^\varepsilon = \partial_x c^{Mau}\frac{Qk_0}{D}\varepsilon^{2-\alpha}\Big\{-\frac{2}{45} + (1-y^2)(\frac{1}{6} - \frac{y^2}{2})\Big\} \\ F_3^\varepsilon = \varepsilon^{2-\alpha}(\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180})\big\{\partial_{xt}c^{Mau}\frac{Q}{D} - \varepsilon^\alpha\partial_{xxx}c^{Mau}Q\big\} \\ F_4^\varepsilon = \varepsilon^{2-\alpha}(\frac{1}{6} - \frac{y^2}{2})\big\{\partial_t c^{Mau}\frac{k_0}{D} - \varepsilon^\alpha\partial_{xx}c^{Mau}k_0\big\} \\ F_5^\varepsilon = \varepsilon^{2-\alpha}\{-\frac{2}{45}\partial_x c^{Mau}\frac{Qk_0}{D} + \frac{k_0^2}{3D}c^{Mau}\} \end{cases} \tag{6.76}$$

These functions aren't integrable up to $t = 0$ and we need a cut off ζ in order to deal with them.

After applying $\mathcal{L}^\varepsilon$ to c_1^{eff}, we find out that

$$\mathcal{L}^\varepsilon(c_1^{eff}) = \zeta(t)\sum_{j=1}^{5} F_j^\varepsilon + \Big(\varepsilon^{2-\alpha}\partial_{xx}c^{Mau}\frac{Q^2}{D}\frac{8}{945} + Q(1/3-y^2)\partial_x c^{Mau} - k_0 c^{Mau}\Big)(1-\zeta(t)) + \zeta'(t)\varepsilon^{2-\alpha}\Big(\partial_x c^{Mau}\frac{Q}{D}\{\frac{y^2}{6}-\frac{y^4}{12}-\frac{7}{180}\} + \frac{k_0}{2D}(\frac{1}{3}-y^2)c^{Mau}\Big) \equiv \Phi_1^\varepsilon \text{ and } -\mathcal{L}^\varepsilon(c_1^{eff}) = \mathcal{L}^\varepsilon(c^\varepsilon - c_1^{eff}) = -\Phi_1^\varepsilon \tag{6.77}$$

At the lateral boundary $y=1$ we have

$$-D\varepsilon^{\alpha-2}\partial_y c_1^{eff}|_{y=1} = \zeta(t)k_0 c^{Mau} \tag{6.78}$$

$$k_0 c_1^{eff}|_{y=1} = k_0\big(c^{Mau} + \varepsilon^{2-\alpha}\frac{Q}{D}\zeta(t)\frac{2}{45}\partial_x c^{Mau} - \varepsilon^{2-\alpha}\frac{k_0}{3D}c^{Mau}\zeta(t)\big) \tag{6.79}$$

Now $c^\varepsilon - c_1^{eff}$ satisfies the system

$$\mathcal{L}^\varepsilon(c^\varepsilon - c_1^{eff}) = -\Phi_1^\varepsilon \text{ in } Z^+\times(0,T) \tag{6.80}$$

$$-D\varepsilon^{\alpha-2}\partial_y(c^\varepsilon - c_1^{eff})|_{y=1} = k_0(c^\varepsilon - c_1^{eff})|_{y=1} + g^\varepsilon|_{y=1} \text{ on } I\!R_+\times(0,T) \tag{6.81}$$

$$\partial_y(c^\varepsilon - c_1^{eff})|_{y=0} = 0 \text{ on } I\!R_+\times(0,T) \tag{6.82}$$

$$(c^\varepsilon - c_1^{eff})|_{t=0} = 0 \text{ on } Z^+ \text{ and } (c^\varepsilon - c_1^{eff})|_{x=0} = \eta_0^\varepsilon \text{ on } (0,1)\times(0,T). \tag{6.83}$$

with

$$g^\varepsilon = k_0\zeta(t)\varepsilon^{2-\alpha}\big(\partial_x c^{Mau}\frac{2Q}{45D} - c^{Mau}\frac{k_0}{3D}\big) + (1-\zeta)k_0 c^{Mau} \tag{6.84}$$

$$\text{and}\quad \eta_0^\varepsilon = -\varepsilon^{2-\alpha}\zeta(t)\partial_x c^{Mau}|_{x=0}(\frac{y^2}{6}-\frac{y^4}{12}-\frac{7}{180})\frac{Q}{D}. \tag{6.85}$$

Now we should estimate Φ_1^ε to see if the right hand side is smaller than in Section 3. We have

Proposition 4.6.1 *Let $O_T = I\!R_+\times(0,1)\times(0,T)$. Let $\varphi\in H^1(O_T)$, $\varphi=0$ at $x=0$. Then we have*

$$|\int_0^t\int_{Z^+}\zeta F_1^\varepsilon\varphi\, dxdyd\tau| \le C\varepsilon^{3(2-\alpha)/2}\|\zeta(\tau)\partial_{xx}c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \le C\big(\varepsilon^{3-5\alpha/2}H(1-\alpha)+\varepsilon^{1-\alpha/2}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.86}$$

$$|\int_0^t\int_{Z^+}\zeta(\tau)F_3^\varepsilon\varphi\, dxdyd\tau| \le C\varepsilon^{3(2-\alpha)/2}\Big(\|\zeta(\tau)\partial_{xt}c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} + \|\zeta(\tau)\partial_{xx}c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\Big)\cdot\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \le C\big(\varepsilon^{3-5\alpha/2}H(1-\alpha)+\varepsilon^{1-\alpha/2}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.87}$$

$$|\int_0^t\int_{Z^+}(1-\zeta)\partial_{xx}c^{Mau}\varepsilon^{2-\alpha}\frac{Q^2}{D}\varphi\, dxdyd\tau| \leq C\varepsilon^{2-3\alpha/2}\|\varepsilon^{\alpha/2}\partial_x\varphi\|_{L^2(O_t)}\cdot$$
$$\|(1-\zeta)\partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \leq C\varepsilon^{2-3\alpha/2}\|\varepsilon^{\alpha/2}\partial_x\varphi\|_{L^2(O_t)} \tag{6.88}$$

$$|\int_0^t\int_{Z^+}(1-\zeta)Q(1/3-y^2)\partial_x c^{Mau}\varphi\, dxdyd\tau| \leq C\varepsilon^{1-\alpha/2}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)}\cdot$$
$$\|(1-\zeta)\partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \leq C\varepsilon^{1-\alpha/2}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.89}$$

$$|\int_0^t\int_{Z^+}\zeta'(\frac{t}{D\varepsilon})\varepsilon^{2-\alpha}\{\partial_x c^{Mau}\frac{Q}{D}\{\frac{y^2}{6}-\frac{y^4}{12}-\frac{7}{180}\}-\frac{k_0}{2D}(\frac{1}{3}-y^2)c^{Mau}\}\cdot$$
$$\varphi\, dxdyd\tau| \leq C\varepsilon^{3-3\alpha/2}\|\zeta'\partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \leq$$
$$C\big(\varepsilon^{3-5\alpha/2}H(1-\alpha)+\varepsilon^{1-\alpha/2}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.90}$$

Proof. Let us note that in (6.86)-(6.87) and (6.89)-(6.90) the averages of the polynomials in y are zero. We write them in the form $P(y)=\partial_y P_1(y)$, where P_1 has zero traces at $y=0,1$, and after partial integration and applying the results from Section 2, giving us the precise regularity, obtain the estimates. Since $(1-\zeta)\partial_{xx}c^{Mau}$ isn't square integrable, we use the x-derivative in order to obtain (6.88). ■

Proposition 4.6.2 *Let $O_T = I\!R_+\times(0,1)\times(0,T)$. Let $\varphi\in H^1(O_T)$, $\varphi=0$ at $x=0$. Then we have*

$$|\int_0^t\int_{Z^+}\zeta F_2^\varepsilon\varphi\, dxdyd\tau| \leq C\varepsilon^{3(1-\alpha/2)}\|\zeta\partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{3-7\alpha/4}H(1-\alpha)+\varepsilon^{5/2-5\alpha/4}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.91}$$

$$|\int_0^t\int_{Z^+}\zeta F_4^\varepsilon\varphi\, dxdyd\tau| \leq C\varepsilon^{3-3\alpha/2}\Big(\|\zeta\partial_t c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}+$$
$$\varepsilon^{\alpha}\|\zeta\partial_{xx}c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\Big)\cdot\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \leq$$
$$C\big(\varepsilon^{3-7\alpha/4)}H(1-\alpha)+\varepsilon^{5(2-\alpha)/4}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.92}$$

$$|\int_0^t\int_0^{+\infty}\zeta\partial_x c^{Mau}\varepsilon^{2-\alpha}(\int_0^1\varphi\, dy-\varphi|_{y=1})\, dxd\tau| \leq$$
$$C\varepsilon^{2-\alpha}\|\partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))}\|\int_0^1\varphi\, dy-\varphi|_{y=1}\|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{3-7\alpha/4}H(1-\alpha)+\varepsilon^{5(2-\alpha)/4}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.93}$$

$$|\int_0^t\int_0^{+\infty}\zeta(t)c^{Mau}\varepsilon^{2-\alpha}(\int_0^1\varphi\, dy-\varphi|_{y=1})\, dxd\tau| \leq$$
$$C\varepsilon^{3(1-\alpha/2)}\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.94}$$

$$|\int_0^t\int_0^{+\infty}(1-\zeta(t))c^{Mau}(\int_0^1\varphi\, dy-\varphi|_{y=1})\, dxd\tau| \leq$$
$$C\big(\varepsilon H(1-\alpha)+\varepsilon^{2-\alpha}H(\alpha-1)\big)\|\varepsilon^{\alpha/2-1}\partial_y\varphi\|_{L^2(O_t)} \tag{6.95}$$

Corollary 3 *Let $\varphi \in H^1(O_T)$, $\varphi = 0$ at $x = 0$. Let Φ_1^ε be given by (6.77) and g^ε by (6.84). Then we have*

$$|\int_0^t \int_{Z^+} \Phi_1^\varepsilon \varphi \, dxdyd\tau + \int_0^t \int_{I\!R_+} g^\varepsilon|_{y=1} \varphi|_{y=1} \, dxd\tau| \leq C\big(\varepsilon^{1-\alpha/2} H(1-\alpha)$$

$$+\varepsilon^{2-3\alpha/2} H(\alpha - 1)\big) \Big\{ \|\varepsilon^{\alpha/2-1} \partial_y \varphi\|_{L^2(O_t)} + \|\varepsilon^{\alpha/2} \partial_x \varphi\|_{L^2(O_t)} \Big\} \tag{6.96}$$

Next we should correct the values at $x = 0$ and apply Proposition 4.3.1. Due to the presence of the term containing the first order derivative in x, the boundary layer corresponding to our problem doesn't enter into the theory from [10] and one should generalize it. The generalization in the case of the periodic boundary conditions at the lateral boundary is in the paper [12]. In our knowledge, the generalization to the case of Neumann's boundary conditions at the lateral boundary, was never published. It seems that the results from [12] apply also to this case ([13]). In order to avoid developing the new theory for the boundary layer, we simply use the boundary layer for the Neumann problem for Laplace operator (5.70). Then the transport term is ignored and a large error in the forcing term is created. The error is concentrated at small times and by eliminating them we would obtain a good estimate.

In order to use this particular point, we prove the following proposition :

Proposition 4.6.3 *Let $\Psi(x) = 1/(1+x)$. Let g^ε and Φ^ε be bounded functions such that $\Psi g^\varepsilon \in H^1(Z^+ \times (0,T))$ and $\Psi \Phi^\varepsilon \in L^2(Z^+ \times (0,T))$. Let ξ, $\Psi \xi \in C^{0,\alpha_0}([0,T]; L^2(Z^+))$, $\Psi \nabla_{x,y} \xi \in L^2(Z^+ \times (0,T))$, be a bounded function which satisfies the system*

$$\mathcal{L}^\varepsilon(\xi) = -\Phi^\varepsilon \quad in \quad Z^+ \times (0,T) \tag{6.97}$$

$$-D\varepsilon^{\alpha-2} \partial_y \xi|_{y=1} = k_0 \xi|_{y=1} + g^\varepsilon|_{y=1} \;\; and \;\; \partial_y \xi|_{y=0} = 0 \;\; on \;\; I\!R_+ \times (0,T) \tag{6.98}$$

$$\xi|_{t=0} = 0 \quad on \quad Z^+ \quad and \; \xi|_{x=0} = 0 \; on \;\; (0,1) \times (0,T). \tag{6.99}$$

Then we have the following energy estimate

$$\mathcal{E}(t^k \xi, t) = t^{2k} \int_{Z^+} \Psi(x)^2 \xi^2(t) \, dxdy + D\varepsilon^\alpha \int_0^t \int_{Z^+} \Psi(x)^2 \tau^{2k} \Big\{ \varepsilon^{-2} |\partial_y \xi|^2 +$$

$$|\partial_x \xi|^2 \Big\} \, dxdyd\tau + k_0 \int_0^t \int_{I\!R_+} \tau^{2k} \xi^2|_{y=1} \Psi^2(x) \, dxd\tau \leq C_1 | \int_0^t \int_{Z^+} \tau^{2k} \Psi(x)^2 \Phi^\varepsilon \xi \, dxdyd\tau$$

$$+ \int_0^t \int_{I\!R_+} \tau^{2k} g^\varepsilon|_{y=1} \xi|_{y=1} \Psi^2(x) \, dxd\tau| + C_2 D\varepsilon^\alpha \int_0^t \int_{Z^+} \tau^{2k} \Psi(x)^2 \xi^2 \, dxdyd\tau, \; \forall k \geq 1. \tag{6.100}$$

Remark 12 *Clearly we have in our mind $\xi = c^\varepsilon - c_1^{eff}$. Then $\zeta(t) \partial_x c^{Mau}$ has the required regularity, since the cut-off erases the singularity. With c^{Mau} things are more complicated. By a* *direct calculation we have $\partial_t c^{Mau} \in L^q(0,T;L^2(I\!R_+))$, $\forall q \in [1,4/3)$ and we get the required Hölder regularity by the Sobolev embedding. $\int_0^A \int_0^1 |\xi(x,y,t)|^2 \, dxdy$ is Hölder-continuous with some exponent $\alpha_0 > 0$, $\forall A < +\infty$, which is* **independent of** *ε. In complete analogy, c_0^{eff} defined by (3.32) has also the required regularity. Finally, the difference $c^\varepsilon - c_0^{eff}$ satisfies the equations (3.39) and (3.40) and it is zero at $x = 0$ and at $t = 0$. Then the classical parabolic regularity theory (see e.g. [7]) implies the Hölder regularity in time of the L^2-norm with respect to x, y. After putting all these results together, we get the required regularity of ξ.*

Proof. By the supposed Hölder continuity, there is $t_M \in [0,T]$, $t_M > 0$, such that

$$\frac{1}{t_M^{\alpha_0}} \int_0^{+\infty} \int_0^1 |\xi(x,y,t_M)|^2 \Psi^2(x) \, dxdy = \max_{t\in[0,T]} \frac{1}{t^{\alpha_0}} \int_0^{+\infty} \int_0^1 |\xi(x,y,t)|^2 \Psi^2(x) \, dx \tag{6.101}$$

Then we have

$$\begin{aligned} \int_0^{t_M} k\tau^{2k-1} \int_{Z^+} |\xi|^2 \Psi^2(x) \, dxdyd\tau &\le \int_{Z^+} \frac{|\xi|^2(t_M)}{t_M^{\alpha_0}} \Psi^2(x) \int_0^{t_M} k\tau^{2k-1+\alpha_0} \, d\tau \\ &= \frac{k}{2k+\alpha_0} t_M^{2k} \int_{Z^+} |\xi|^2(t_M) \Psi^2(x) \, dxdy \end{aligned} \tag{6.102}$$

and

$$\begin{aligned} &\frac{1}{2} t_M^{2k} \int_{Z^+} |\xi|^2(t_M)\Psi^2(x) \, dxdy + k_0 \int_0^t \int_0^{+\infty} \tau^{2k} \xi^2|_{y=1} \Psi^2(x) \, dxd\tau + \\ &\int_0^{t_M} D\Big(\varepsilon^\alpha \int_{Z^+} \tau^{2k} |\partial_x \xi|^2(\tau)\Psi^2(x) \, dxdy + \varepsilon^{\alpha-2} \int_{Z^+} \tau^{2k} |\partial_y \xi|^2(\tau)\Psi^2(x) \, dxdy\Big) \, d\tau \\ &\le - \int_0^{t_M} \int_{Z^+} \tau^{2k} \Phi^\varepsilon \xi \, dxdyd\tau - k_0 \int_0^t \int_0^{+\infty} \tau^{2k} \xi|_{y=1} g^\varepsilon|_{y=1} \Psi^2(x) \, dxd\tau + \\ &D\varepsilon^\alpha \int_0^{t_M} \int_{Z^+} \tau^{2k} \Psi^2(x) \xi^2 \, dxdyd\tau + k \int_0^{t_M} \int_{Z^+} \tau^{2k-1} |\xi|^2 \Psi^2 \, dxdyd\tau \end{aligned} \tag{6.103}$$

Using (6.102) we get (6.100) for $t = t_M$ and with $C_2 = 0$. Getting the estimates (6.100) for general $t \in (0,T)$ is now straightforward. ■ Next, in order to use this estimate we should refine the estimates from Propositions 4.6.1 and 4.6.2 . First we note that the estimate (2.28) changes to

$$\begin{aligned} &\|t^k \partial_{tt} c^{Mau}\|_{L^q((0,T)\times I\!R_+)} + \|t^k \partial_{tx} c^{Mau}\|_{L^q((0,T)\times I\!R_+)} + \|t^k \partial_{xx} c^{Mau}\|_{L^q((0,T)\times I\!R_+)} \\ &\qquad \le C_q(k)(\gamma \bar{D})^{1/(2q)-1}. \end{aligned} \tag{6.104}$$

Hence one gains $\varepsilon^{\alpha/4}$ (respectively $\varepsilon^{1/2-\alpha/4}$) for the L^2-norm. In analogy with Propositions 4.6.1 and 4.6.2 we have

Proposition 4.6.4 *Let $O_T = I\!R_+ \times (0,1) \times (0,T)$. Let $\varphi \in H^1(O_T)$, $\varphi = 0$ at $x = 0$ and $k > 1$. Then we have*

$$|\int_0^t \int_0^\infty \int_0^1 \tau^k \zeta F_1^\varepsilon \varphi \, dxdyd\tau| \leq C\varepsilon^{3(2-\alpha)/2} \|\tau^k \partial_{xx} c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{3-9\alpha/4} H(1-\alpha) + \varepsilon^{3/2-3\alpha/4} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.105)$$

$$|\int_0^t \int_0^\infty \int_0^1 \tau^k \zeta F_3^\varepsilon \varphi \, dxdyd\tau| \leq C\varepsilon^{3(2-\alpha)/2} \Big(\|\tau^k \partial_{xt} c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} + \|\tau^k \partial_{xx} c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \Big) \cdot \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \leq C\big(\varepsilon^{3-9\alpha/4} H(1-\alpha) + \varepsilon^{3/2-3\alpha/4} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.106)$$

$$|\int_0^t \int_{Z^+} \zeta \tau^k F_2^\varepsilon \varphi \, dxdyd\tau| \leq C\varepsilon^{3(1-\alpha/2)} \|\tau^k \zeta \partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{3-7\alpha/4} H(1-\alpha) + \varepsilon^{5/2-5\alpha/4} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.107)$$

$$|\int_0^t \int_{Z^+} \zeta \tau^k F_4^\varepsilon \varphi \, dxdyd\tau| \leq C\varepsilon^{3-3\alpha/2} \Big(\|\zeta \tau^k \partial_t c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} + \varepsilon^\alpha \|\zeta \tau^k \partial_{xx} c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \Big) \cdot \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \leq C\big(\varepsilon^{3-7\alpha/4)} H(1-\alpha) + \varepsilon^{5(2-\alpha)/4} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.108)$$

$$|\int_0^t \int_0^{+\infty} \zeta \tau^k \partial_x c^{Mau} \varepsilon^{2-\alpha} (\int_0^1 \varphi \, dy - \varphi|_{y=1}) \, dxd\tau| \leq C\varepsilon^{2-\alpha} \|\tau^k \partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \| \int_0^1 \varphi \, dy - \varphi|_{y=1} \|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{3-7\alpha/4} H(1-\alpha) + \varepsilon^{5(2-\alpha)/4} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.109)$$

$$|\int_0^t \int_0^{+\infty} \zeta(t) \tau^k c^{Mau} \varepsilon^{2-\alpha} (\int_0^1 \varphi \, dy - \varphi|_{y=1}) \, dxd\tau| \leq C\varepsilon^{3(1-\alpha/2)} \|\varepsilon^{\alpha/2-1}\partial_y \varphi\|_{L^2(O_t)} \quad (6.110)$$

Proof. These estimates are straightforward consequences of Propositions 4.6.1 and 4.6.2 . ■We gain more with other terms :

Proposition 4.6.5 *Let $O_T = I\!R_+ \times (0,1) \times (0,T)$. Let $\varphi \in H^1(O_T)$, $\varphi = 0$ at $x = 0$. Then* *we have*

$$|\int_0^t \int_0^\infty \int_0^1 (1-\zeta)\tau^k \partial_{xx} c^{Mau} \frac{Q^2}{D} \varepsilon^{2-\alpha} \varphi \, dxdyd\tau|$$
$$\leq C\varepsilon^{2-3\alpha/2} \|(1-\zeta)\tau^k \partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \|\varepsilon^{\alpha/2} \partial_x \varphi\|_{L^2(O_t)}$$
$$\leq C\big(\varepsilon^{k\alpha+2-3\alpha/2} H(1-\alpha) + \varepsilon^{k(2-\alpha)+2-3\alpha/2} H(\alpha-1)\big) \|\varepsilon^{\alpha/2} \partial_x \varphi\|_{L^2(O_t)} \tag{6.111}$$
$$|\int_0^t \int_0^\infty \int_0^1 (1-\zeta)\tau^k Q(1/3-y^2) \partial_x c^{Mau} \varphi \, dxdyd\tau| \leq$$
$$C\varepsilon^{1-\alpha/2} \|(1-\zeta)\tau^k \partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \|\varepsilon^{\alpha/2-1} \partial_y \varphi\|_{L^2(O_t)} \leq$$
$$C\big(\varepsilon^{k\alpha+1-\alpha/2} H(1-\alpha) + \varepsilon^{k(2-\alpha)+1-\alpha/2} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1} \partial_y \varphi\|_{L^2(O_t)} \tag{6.112}$$
$$|\int_0^t \int_{Z^+} \zeta'(\frac{t}{D\varepsilon}) \tau^k \varepsilon^{2-\alpha} \big\{\partial_x c^{Mau} \frac{Q}{D}\{\frac{y^2}{6} - \frac{y^4}{12} - \frac{7}{180}\} - \frac{k_0}{2D}(\frac{1}{3} - y^2) c^{Mau}\big\}.$$
$$\varphi \, dxdyd\tau| \leq C\varepsilon^{3-3\alpha/2} \|\zeta' \tau^k \partial_x c^{Mau}\|_{L^2(0,t;L^2(I\!R_+))} \|\varepsilon^{\alpha/2-1} \partial_y \varphi\|_{L^2(O_t)} \leq$$
$$C\big(\varepsilon^{3-3\alpha/2+\alpha(k-1)} H(1-\alpha) + \varepsilon^{3-3\alpha/2+(2-\alpha)(k-1)} H(\alpha-1)\big) \|\varepsilon^{\alpha/2-1} \partial_y \varphi\|_{L^2(O_t)} \tag{6.113}$$

Before applying Proposition 4.6.3 and getting the final estimate, we should correct the trace at $x = 0$. It is done by adding

$$\bar{c}_1^{eff} = -\varepsilon^{2-\alpha} \zeta(t) \beta^\varepsilon \partial_x c^{Mau} \frac{Q}{D}, \tag{6.114}$$

where $\beta^\varepsilon(x, y) = \beta(x/\varepsilon, y)$ is the boundary layer function given by (5.70). Then for $\xi^\varepsilon = c^\varepsilon - c_1^{eff} - \bar{c}_1^{eff}$ we have

$$\mathcal{L}^\varepsilon(\xi) = -\Phi^\varepsilon = -\Phi_1^\varepsilon + \partial_t \zeta \varepsilon^{2-\alpha} \partial_x c^{Mau} \frac{Q}{D} \beta^\varepsilon + \varepsilon^{2-\alpha} \beta^\varepsilon \zeta(t) \big\{\partial_{xt} c^{Mau} \frac{Q}{D} -$$

$$\varepsilon^\alpha \partial_{xxx} c^{Mau} Q\big\} + \partial_x \beta^\varepsilon \frac{Q^2}{D} (1-y^2) \zeta \varepsilon^{2-\alpha} \partial_x c^{Mau} - \varepsilon^{2-\alpha} Q \partial_{xx} c^{Mau} \zeta(t) \big(2\varepsilon^\alpha \partial_x \beta^\varepsilon -$$
$$\beta^\varepsilon (1-y^2) \frac{Q}{D}\big) \text{ in } Z^+ \times (0,T) \tag{6.115}$$
$$-D\varepsilon^{\alpha-2} \partial_y \xi^\varepsilon|_{y=1} = k_0 \xi|_{y=1} + g^\varepsilon|_{y=1} - k_0 \varepsilon^{2-\alpha} \zeta \frac{Q}{D} \partial_x c^{Mau} \beta^\varepsilon|_{y=1} \text{ on } I\!R_+ \times (0,T) \tag{6.116}$$

$$\text{and } \partial_y \xi^\varepsilon|_{y=0} = 0 \text{ on } I\!R_+ \times (0,T) \tag{6.117}$$

$$\xi^\varepsilon|_{t=0} = 0 \quad \text{on } Z^+ \quad \text{and } \xi^\varepsilon|_{x=0} = 0 \text{ on } (0,1) \times (0,T). \tag{6.118}$$

We need an estimate for new terms. The estimates will follow from the following auxiliary result

Lemma 4.6.6 *Let β be defined by (5.70), let $k \geq 1$ and c^{Mau} the solution for (6.75). Then we* *have*

$$\|\tau^k\zeta'\beta^\varepsilon\partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k-3/4}\Big\{\varepsilon^{-\alpha/4}H(1-\alpha)+ \varepsilon^{\alpha/4-1/2}H(\alpha-1)\Big\} \le C\varepsilon^{k-1} \tag{6.119}$$

$$\|\tau^k\zeta\beta^\varepsilon|_{y=1}\partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k+1/4}\Big\{\varepsilon^{-\alpha/4}H(1-\alpha)+ \varepsilon^{\alpha/4-1/2}H(\alpha-1)\Big\} \le C\varepsilon^{k} \tag{6.120}$$

$$\|\tau^k\zeta\partial_x\beta^\varepsilon\partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k-3/4}\Big\{\varepsilon^{-\alpha/4}H(1-\alpha)+ \varepsilon^{\alpha/4-1/2}H(\alpha-1)\Big\} \le C\varepsilon^{k-1} \tag{6.121}$$

$$\|\tau^k\zeta\partial_x\beta^\varepsilon\partial_t c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k-5/4}\Big\{\varepsilon^{\alpha/2}H(1-\alpha)+ \varepsilon^{1-\alpha/2}H(\alpha-1)\Big\} \le C\varepsilon^{k-5/4} \tag{6.122}$$

$$\|\tau^k\zeta\beta^\varepsilon\partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k}\Big\{(\varepsilon^{-1/4-\alpha/2}+\varepsilon^{1/4-3\alpha/4})H(1-\alpha)+ (\varepsilon^{\alpha/2-5/4}+\varepsilon^{-5/2+3\alpha/4})H(\alpha-1)\Big\} \le C\varepsilon^{k-7/4} \tag{6.123}$$

$$\|\tau^k\zeta\partial_x\beta^\varepsilon\partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C\varepsilon^{k-1}\Big\{(\varepsilon^{-1/4-\alpha/2}+\varepsilon^{1/4-3\alpha/4})H(1-\alpha)+ (\varepsilon^{\alpha/2-5/4}+\varepsilon^{-5/2+3\alpha/4})H(\alpha-1)\Big\} \le C\varepsilon^{k-7/4} \tag{6.124}$$

Proof. We have

$$\int_0^{+\infty} |\partial_x c^{Mau}\beta^\varepsilon|^2\,dx \le C\int_0^{+\infty} \exp\{-\frac{2\gamma_0 x}{\varepsilon}\}\exp\{-\frac{(x-\tau\bar Q)^2}{2\gamma\bar D\tau}\}\,\frac{dx}{\gamma\tau\bar D} \le C(\varepsilon D\tau)^{-1/2}\exp\{-C_0\tau/\varepsilon\}\,dxd\tau \tag{6.125}$$

Now (6.119) , (6.120) and (6.121) follow by integration with respect to τ. Next,

$$\int_0^{+\infty} |\partial_t c^{Mau}\beta^\varepsilon|^2\,dx \le C\int_0^{+\infty} x^2\exp\{-\frac{2\gamma_0 x}{\varepsilon}\}\exp\{-\frac{(x-\tau\bar Q)^2}{2\gamma\bar D\tau^3}\}\,\frac{dx}{\gamma\tau\bar D} \le C(\varepsilon D\tau^3)^{-1/2}\exp\{-C_0\tau/\varepsilon\}\,dxd\tau \tag{6.126}$$

and (6.122) follows. Since

$$\|\tau^k\zeta\beta^\varepsilon\partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \le C(\|\tau^k\zeta\beta^\varepsilon\partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)}+ \|\tau^k\zeta\beta^\varepsilon\partial_t c^{Mau}\|_{L^2((0,t)\times Z^+)})(\varepsilon^{-\alpha}H(1-\alpha)+\varepsilon^{\alpha-2}H(\alpha-1)) \tag{6.127}$$

we get (6.122) and (6.123). ■

Proposition 4.6.7 *Let $\varphi \in H^1(O_T)$, $\varphi = 0$ at $x = 0$. Then we have*

$$|\int_0^t \int_{Z^+} \varepsilon^{2-\alpha} \tau^k \zeta(\tau) \beta^\varepsilon \{\partial_{xt} c^{Mau} \frac{Q}{D} - \varepsilon^\alpha \partial_{xxx} c^{Mau} Q\} \varphi \, dxdyd\tau|$$
$$\leq C\varepsilon^{2-\alpha} \Big(\{ \|\zeta \tau^k \partial_t c^{Mau} \partial_x \beta^\varepsilon\|_{L^2((0,t)\times Z^+)} + \varepsilon^\alpha \|\tau^k \zeta \partial_x \beta^\varepsilon \partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \} \cdot$$
$$\|\varphi\|_{L^2((0,t)\times Z^+)} + \varepsilon^{-\alpha/2} \{ \|\zeta \tau^k \partial_t c^{Mau} \beta^\varepsilon\|_{L^2((0,t)\times Z^+)} + \varepsilon^\alpha \|\tau^k \zeta \partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \} \cdot$$
$$\|\varepsilon^{\alpha/2} \partial_x \varphi\|_{L^2((0,t)\times Z^+)} \Big) \leq C\varepsilon^{k+1/4-\alpha} \big(\|\varphi\|_{L^2((0,t)\times Z^+)} + \|\varepsilon^{\alpha/2} \partial_x \varphi\|_{L^2((0,t)\times Z^+)} \big) \quad (6.128)$$

$$|\int_0^t \int_{Z^+} \varepsilon^{2-\alpha} \zeta \tau^k \partial_{xx} c^{Mau} \varphi \Big(-\beta^\varepsilon \frac{Q}{D}(1-y^2) + 2\varepsilon^\alpha \partial_x \beta^\varepsilon \Big) \, dxdyd\tau|$$
$$\leq C\varepsilon^{2-\alpha} \big(\|\tau^k \zeta \partial_x \beta^\varepsilon \partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} + \|\tau^k \zeta \partial_x \beta^\varepsilon \partial_{xx} c^{Mau}\|_{L^2((0,t)\times Z^+)} \big) \cdot$$
$$\|\varphi\|_{L^2(O_t)} \leq C\varepsilon^{k-\alpha+1/4} \|\varphi\|_{L^2(O_t)} \quad (6.129)$$

$$|\int_0^t \int_{Z^+} \varepsilon^{2-\alpha} \zeta \tau^k \partial_x c^{Mau} \partial_x \beta^\varepsilon (1-y^2) \varphi \, dxdyd\tau|$$
$$\leq C\varepsilon^{2-\alpha} \|\tau^k \zeta \partial_x \beta^\varepsilon \partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)} \|\varphi\|_{L^2(O_t)} \leq C\varepsilon^{k-\alpha+1} \|\varphi\|_{L^2(O_t)} \quad (6.130)$$

$$|\int_0^t \int_0^{+\infty} \varepsilon^{2-\alpha} \zeta \tau^k \partial_x c^{Mau} \varphi|_{y=1} \beta^\varepsilon|_{y=1} \, dxd\tau|$$
$$\leq C\varepsilon^{2-\alpha} \|\tau^k \zeta \partial_x \beta^\varepsilon_{y=1} \partial_x c^{Mau}\|_{L^2((0,t)\times I\!R_+)} \|\varphi\|_{L^2(O_t)} \leq C\varepsilon^{k-\alpha+1} \|\varphi|_{y=1}\|_{L^2((0,t)\times I\!R_+))} \quad (6.131)$$

$$|\int_0^t \int_{Z^+} \varepsilon^{2-\alpha} \zeta'(\tau) \tau^k \partial_x c^{Mau} \varphi \beta^\varepsilon \, dxdyd\tau|$$
$$\leq C\varepsilon^{2-\alpha} \|\tau^k \zeta' \beta^\varepsilon \partial_x c^{Mau}\|_{L^2((0,t)\times Z^+)} \|\varphi\|_{L^2(O_t)} \leq C\varepsilon^{k-\alpha+3/4} \|\varphi\|_{L^2(O_t)} \quad (6.132)$$

Now the application of Proposition 4.6.3 is straightforward and after considering various powers we get

Theorem 4.6.8 *Let c^{Mau} be given by (6.75), let c_1^{eff} be given by (6.74) and $\bar{c}_1^{eff}$ by (6.114). Then we have*

$$\|t^3 (c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff})\|_{L^\infty(0,T;L^2_{loc}(I\!R_+\times(0,1))} \leq C\big(\varepsilon^{3-9\alpha/4} H(1-\alpha) +$$
$$\varepsilon^{3(1-\alpha/2)/2} H(\alpha-1)\big) \quad (6.133)$$

$$\|t^3 \partial_y \big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff}\big)\|_{L^2(0,T;L^2_{loc}(I\!R_+\times(0,1))} \leq$$
$$C\varepsilon^{1-\alpha/2} \big(\varepsilon^{3-9\alpha/4} H(1-\alpha) + \varepsilon^{3(1-\alpha/2)/2} H(\alpha-1)\big) \quad (6.134)$$

$$\|t^2 \partial_x \big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff}\big)\|_{L^2(0,T;L^2_{loc}(I\!R_+\times(0,1))} \leq$$
$$C\varepsilon^{-\alpha/2} \big(\varepsilon^{3-9\alpha/4} H(1-\alpha) + \varepsilon^{3(1-\alpha/2)/2} H(\alpha-1)\big) \quad (6.135)$$

4.7 Error estimate involving the second order in expansion

The most important power of α is $\alpha = 1$, which describes Taylor's scaling. In this case our approximation is of order $\varepsilon^{3/4}$. Nevertheless, it is interesting to reach the order ε at least in this case. Also, it could be of interest to get the higher order estimates, which can be useful for ε which is not very small.

Clearly, the estimate isn't sufficiently good due to the terms ζF_1^ε and ζF_3^ε. When deriving formally the effective equation, we have seen that they could be eliminated by introducing the next order correction. Following the formal expansion we find out that c_1^{eff} should be replaced by $c_1^{eff} + c_2^{eff}$, where

$$c_2^{eff} = -\varepsilon^{4-2\alpha}\frac{Q}{D^2}\zeta(t)\Big\{Q\partial_{xx}c^{Mau}\Big(\frac{281}{453600} + \frac{23}{1512}y^2 - \frac{37}{2160}y^4 + \frac{1}{120}y^6 - \frac{1}{672}y^8 - \tilde\beta_1\Big) - (\partial_{xt}c^{Mau} - D\varepsilon^\alpha\partial_{xxx}c^{Mau})\Big(-\frac{1}{360}y^6 + \frac{1}{72}y^4 - \frac{7}{360}y^2 - \frac{31}{7560} - \tilde\beta_2\Big)\Big\} + \varepsilon^{4-2\alpha}\frac{k_0}{D^2}\zeta(t)\Big\{Q\partial_x c^{Mau}\Big(\frac{1}{60}y^6 - \frac{1}{18}y^4 + \frac{11}{180}y^2 - \frac{11}{810} - \tilde\beta_3\Big) + \frac{1}{2}(\partial_t c^{Mau} - D\varepsilon^\alpha\partial_{xx}c^{Mau})\Big(-\frac{1}{12}y^4 + \frac{1}{6}y^2 - \frac{7}{180} - \tilde\beta_5\Big) + \frac{Q}{45}\partial_x c^{Mau}\Big(\frac{1}{3} - y^2 - \tilde\beta_4\Big) - \frac{k_0}{6}c^{Mau}\Big(\frac{1}{3} - y^2\Big)\Big\}, \tag{7.136}$$

where $\tilde\beta_j$, $j = 1, \dots, 5$, are solutions to the boundary layers analogous to (5.70) which correct those new values at $x = 0$.

Using this additional correction term we have

Theorem 4.7.1 *Let c^{Mau} be given by (6.75), let c_1^{eff} be given by (6.74), $\bar c_1^{eff}$ by (6.114) and c_2^{eff} by (7.136). Then we have*

$$\|t^5(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar c_1^{eff})\|_{L^\infty(0,T;L^2_{loc}(I\!R_+\times(0,1)))} \leq C\big(\varepsilon^{4-13\alpha/4}H(1-\alpha) + \varepsilon^{3(1-\alpha/2)/2}H(\alpha-1)\big) \tag{7.137}$$

$$\|t^5\partial_y\big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar c_1^{eff}\big)\|_{L^2(0,T;L^2_{loc}(I\!R_+\times(0,1)))} \leq C\varepsilon^{1-\alpha/2}\big(\varepsilon^{4-13\alpha/4}H(1-\alpha) + \varepsilon^{3(1-\alpha/2)/2}H(\alpha-1)\big) \tag{7.138}$$

$$\|t^5\partial_x\big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar c_1^{eff}\big)\|_{L^2(0,T;L^2_{loc}(I\!R_+\times(0,1)))} \leq C\varepsilon^{-\alpha/2}\big(\varepsilon^{4-13\alpha/4}H(1-\alpha) + \varepsilon^{3(1-\alpha/2)/2}H(\alpha-1)\big) \tag{7.139}$$

Proof. After applying the operator $\mathcal{L}^\varepsilon$, given by (3.33), to $c^\varepsilon - c_1^{eff} - \bar c_1^{eff} - c_2^{eff}$ we obtain a forcing term Φ_2^ε, analogous to (6.115). Let us study it. In fact it is enough to study what happened with $\zeta\sum_{j=1}^{5} F_j$. As we have seen in Proposition 4.6.5, Lemma 4.6.6 and Proposition 4.6.7, other terms are small. We have

– F_1^ε and F_3^ε are replaced by

$$\begin{cases} \tilde F_1^\varepsilon = (1-y^2)\dfrac{Q^2\varepsilon^{4-2\alpha}}{D^2}\Big\{ -\partial_{xxx}c^{Mau}P_8(y)Q + (\partial_{xxt}c^{Mau} - \\ D\varepsilon^\alpha \partial_{xxxx}c^{Mau})P_6(y)\Big\} \\ \tilde F_3^\varepsilon = -\varepsilon^{4-2\alpha}P_8(y)\dfrac{Q^2}{D^2}\Big\{\partial_{xxt}c^{Mau} - \varepsilon^\alpha\partial_{xxxx}c^{Mau}D\Big\} + \\ \varepsilon^{4-2\alpha}P_6(y)\dfrac{Q}{D^2}\Big\{\partial_{xtt}c^{Mau} - 2D\varepsilon^\alpha\partial_{xxxt}c^{Mau} + \varepsilon^{2\alpha}\partial_{xxxxx}c^{Mau}D^2\Big\} \\ \\ P_8(y) = \dfrac{281}{453600} + \dfrac{23}{1512}y^2 - \dfrac{37}{2160}y^4 + \dfrac{1}{120}y^6 - \dfrac{1}{672}y^8, \\ \\ P_4(y) = \dfrac{y^2}{6} - \dfrac{y^4}{12} - \dfrac{7}{180};\ P_6(y) = -\dfrac{1}{60}y^5 + \dfrac{1}{18}y^3 - \dfrac{7}{180}y - \dfrac{31}{7560}. \end{cases} \tag{7.140}$$

Using (2.31) we find out, in analogy with (6.105)-(6.106), that

$$\int_0^t\int_0^\infty\int_0^1 \tau^k\zeta(|\tilde F_1^\varepsilon| + |\tilde F_3^\varepsilon|)|\varphi|\ dxdyd\tau \leq \\ C\Big(\varepsilon^{4-13\alpha/4}H(1-\alpha) + \varepsilon^{3/2-3\alpha/4}H(\alpha-1)\Big)\|\varphi\|_{L^2(O_t)}, \tag{7.141}$$

$\forall\varphi\in H^1(O_T),\ \varphi=0$ at $x=0$ and $k>2$.

– F_2^ε and F_4^ε are replaced by

$$\begin{cases} \tilde F_2^\varepsilon = (1-y^2)\dfrac{Qk_0\varepsilon^{4-2\alpha}}{D}\Big\{\partial_{xx}c^{Mau}\dfrac{Q}{D}\tilde P_6(y) + (\partial_{xt}c^{Mau}\dfrac{1}{2D} - \\ \varepsilon^\alpha\partial_{xxx}c^{Mau}\dfrac{1}{2})P_4(y) + (\dfrac{Q}{45D}\partial_x c^{Mau} - \dfrac{k_0}{6D}c^{Mau})P_2(y)\Big\} \\ \tilde F_4^\varepsilon = -\varepsilon^{4-2\alpha}\tilde P_6(y)\dfrac{Qk_0}{D^2}\Big\{\partial_{xt}c^{Mau} - \varepsilon^\alpha\partial_{xxx}c^{Mau}D\Big\} + \\ \varepsilon^{4-2\alpha}P_4(y)\dfrac{k_0}{2D^2}\Big\{\partial_{tt}c^{Mau} - 2D\varepsilon^\alpha\partial_{xxt}c^{Mau} + D^2\varepsilon^{2\alpha}\partial_{xxxx}c^{Mau}\Big\} \\ +\varepsilon^{4-2\alpha}P_2(y)\dfrac{k_0}{3D^2}\Big\{\dfrac{Q}{15}\partial_{xt}c^{Mau} - \dfrac{k_0}{2}\partial_t c^{Mau} - \\ \dfrac{DQ\varepsilon^\alpha}{15}\partial_{xxx}c^{Mau} + \dfrac{Dk_0\varepsilon^\alpha}{2}\partial_{xx}c^{Mau}\Big\}, \end{cases} \tag{7.142}$$

where $P_2(y) = 1/3 - y^2$ and $\tilde P_6 = \dfrac{y^6}{60} - \dfrac{y^4}{18} + \dfrac{11y^2}{180} - \dfrac{11}{810}$. Using (2.31) we find out, in analogy with (6.107)-(6.108), that

$$\int_0^t\int_0^\infty\int_0^1 \tau^k\zeta(|\tilde F_2^\varepsilon| + |\tilde F_4^\varepsilon|)|\varphi|\ dxdyd\tau \leq \\ C\Big(\varepsilon^{4-11\alpha/4}H(1-\alpha) + \varepsilon^{5/2-5\alpha/4}H(\alpha-1)\Big)\|\varphi\|_{L^2(O_t)}, \tag{7.143}$$

$\forall\varphi\in H^1(O_T),\ \varphi=0$ at $x=0$ and $k>2$.

– It should be noted that the means of the polynomials in y, contained in $\tilde{F}_1$ and $\tilde{F}_3$ aren't zero any more. Hence we can't gain some powers of ε using the derivative with respect to y of the test function.

– F_5 and the boundary term $k_0\zeta(t)\varepsilon^{2-\alpha}\big(\partial_x c^{Mau}\dfrac{2Q}{45D} - c^{Mau}\dfrac{k_0}{3D}\big)$ are canceled. At the boundary $y=1$ we have a new non-homogeneous term

$$\hat{g}^\varepsilon = (1-\zeta)k_0 c^{Mau} - \zeta\varepsilon^{4-2\alpha}\big(\frac{2Qk_0^2}{45D^2}\partial_x c^{Mau}\tilde{P}_6|_{y=1}+ \\ (\frac{k_0}{2D^2}\partial_t c^{Mau} - \varepsilon^\alpha\frac{k_0}{2D}\partial_{xx}c^{Mau})P_4|_{y=1}\big) \tag{7.144}$$

and the principal boundary contribution is given by

$$|\int_0^t\int_0^\infty\int_0^1 \tau^k\zeta\varepsilon^{4-2\alpha}\big(\frac{2Qk_0^2}{45D^2}\partial_x c^{Mau}\tilde{P}_6|_{y=1} + (\frac{k_0}{2D^2}\partial_t c^{Mau} \\ -\varepsilon^\alpha\frac{k_0}{2D}\partial_{xx}c^{Mau})P_4|_{y=1}\big)\varphi|_{y=1}\,dxdyd\tau| \le C\bigg(\varepsilon^{4-9\alpha/4}H(1-\alpha)+ \\ \varepsilon^{7/2-7\alpha/4}H(\alpha-1)\bigg)\|\varphi|_{y=1}\|_{L^2((0,t)\times I\!R_+)}, \tag{7.145}$$

– Other terms are much smaller and don't have to be discussed.

After collecting the powers of ε and applying Proposition 4.6.3 we obtain the estimates (7.137)-(7.139). ■

Theorem 4.7.2 *Let c^{Mau} be given by (6.75), let c_1^{eff} be given by (6.74), $\bar{c}_1^{eff}$ by (6.114) and c_2^{eff} by (7.136). Then we have*

$$\|t^5\big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff} - c_2^{eff}\big)\|_{L^2(0,T;L^1_{loc}(I\!R_+\times(0,1))} \le \\ C\big(\varepsilon^{4-3\alpha}H(1-\alpha) + \varepsilon^{2-\alpha}H(\alpha-1)\big) \tag{7.146}$$

$$\|t^5\big(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff} - c_2^{eff}\big)\|_{L^2(0,T;L^2_{loc}(I\!R_+\times(0,1))} \le \\ C\big(\varepsilon^{4-3\alpha}H(1-\alpha) + \varepsilon^{2-\alpha}H(\alpha-1)\big) \tag{7.147}$$

Proof. First we prove the $L^\infty(L^1)$-estimates (7.146). We test the equation for $\xi = c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff} - c_2^{eff}$ with regularized sign of ξ multiplied by Ψ^2 and get

$$t^{2k}\int_{Z^+}\Psi(x)^2|\xi|(t)\,dxdy + k_0\int_0^t\int_{I\!R_+}\tau^{2k}|\xi|_{y=1}|\Psi^2(x)\,dxd\tau \le \\ C_1\int_0^t\int_{Z^+}\tau^{2k}\Psi(x)^2|\Phi_2^\varepsilon|\,dxdyd\tau + \int_0^t\int_{I\!R_+}\tau^{2k}|\hat{g}^\varepsilon|_{y=1}|\Psi^2(x)\,dxd\tau|+ \\ C_2\varepsilon^\alpha\int_0^t\int_{Z^+}\tau^{2k}\Psi(x)^2|\xi|\,dxdyd\tau + k\int_0^t\int_{Z^+}\tau^{2k-1}|\xi|\Psi^2\,dxdyd\tau, \tag{7.148}$$

$\forall k \ge 3$. As before, the L^1-norm of $\Psi^2\xi$ is Hölder continuous in time with some exponent $\alpha_0 > 0$. Consequently, arguing as in Proposition 4.6.3, we obtain

$$\sup_{0\le t\le T}\|t^k\Psi^2\xi(t)\|_{L^1(Z^+)} \le C(\|\Psi^2\Phi_2^\varepsilon\|_{L^1(Z^+\times(0,T))} + \|\Psi^2\hat{g}^\varepsilon|_{y=1}\|_{L^1(I\!R_+\times(0,T))}) \tag{7.149}$$

and (7.146) is proved.

The improved $L^2(L^2)$-estimate (7.147) follows from (7.146), (7.138) and the Poincaré's inequality in H^1 (see e.g. [4]). ■ Next we prove the corresponding $L^\infty(L^\infty)$-estimate. We have

Theorem 4.7.3 *Let c^{Mau} be given by (6.75), let c_1^{eff} be given by (6.74), $\bar{c}_1^{eff}$ by (6.114) and c_2^{eff} by (7.136). Then we have*

$$\|t^5(c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff} - c_2^{eff})\|_{L^\infty((0,T)\times(I\!R_+\times(0,1))} \leq C(\delta)\big(\varepsilon^{4-7\alpha/2-\delta}H(1-\alpha)$$
$$+\varepsilon^{3/2-\alpha-\delta}H(\alpha-1)\big), \quad \forall\delta>0. \tag{7.150}$$

Remark 13 *From the proof we see that $C(\delta)$ has an exponential growth when $\delta\to 0$.*

Proof. Let $M>0$, $\xi = c^\varepsilon - c_1^{eff}(x,t;\varepsilon) - \bar{c}_1^{eff} - c_2^{eff}$ and $\xi_M = \sup\{t^k\xi - M, 0\}$.We test the equation for ξ with $\Psi^2\xi_M$ and get

$$\frac{1}{2}\int_{Z^+}\Psi(x)^2\xi_M^2(t)\,dxdy + D\varepsilon^\alpha\int_0^t\int_{Z^+}\Psi(x)^2|\partial_x\xi_M(\tau)|^2\,dxdyd\tau +$$
$$D\varepsilon^{\alpha-2}\int_0^t\int_{Z^+}\Psi(x)^2|\partial_y\xi_M(\tau)|^2\,dxdyd\tau + k_0\int_0^t\int_{I\!R_+}(\xi_M|_{y=1}+$$
$$M\tau^k)\xi_M|_{y=1}\Psi^2(x)\,dxd\tau \leq C_1\Big|\int_0^t\int_{Z^+}\tau^k\Psi(x)^2|\Phi_3^\varepsilon|\xi_M\,dxdyd\tau$$
$$+\int_0^t\int_{I\!R_+}\tau^k|\hat{g}^\varepsilon|_{y=1}\xi_M|_{y=1}\Psi^2(x)\,dxd\tau\Big| + C_2\varepsilon^\alpha\int_0^t\int_{Z^+}\tau^{2k}\Psi(x)^2\xi_M^2\,dxdyd\tau \tag{7.151}$$

$\forall k\geq 3$, where $\tau^k\Phi_3^\varepsilon = -\tau^k\Phi_2^\varepsilon + k\tau^{k-1}\xi$. We suppose that

$$k_0 M \geq \sup_{0\leq\tau\leq T}\tau^k\|\Psi\hat{g}^\varepsilon(\tau)|_{y=1}\|_{L^\infty(I\!R_+)} = c_0\big(\varepsilon^{4-5\alpha/2}H(1-\alpha) + \varepsilon^{3(1-\alpha/2)}H(\alpha-1)\big) \tag{7.152}$$

As in the classical derivation of the Nash-Moser estimate (see [7], pages 181-186)) we introduce

$$\mu(M) = \int_0^T\int_{Z^+\cap\{t^k\xi - M>0\}}\Psi^2\,dxdydt \tag{7.153}$$

Now in exactly the same way as in [7], pages 181-186, on a time interval which could be smaller than $[0,T]$, but suppose equal to it without loosing the generality, we get

$$\|\xi_M\|_{V_2}^2 = \sup_{0\leq t\leq T}\int_{Z^+}\Psi(x)^2\xi_M^2(t)\,dxdy + D\varepsilon^\alpha\int_0^T\int_{Z^+}\Psi(x)^2|\partial_x\xi_M(\tau)|^2\,dxdyd\tau$$
$$+D\varepsilon^{\alpha-2}\int_0^T\int_{Z^+}\Psi(x)^2|\partial_y\xi_M(\tau)|^2\,dxdyd\tau \leq \beta_0^2\|\tau^k\Phi_3^\varepsilon\Psi\|^2_{L^q(Z^+\times(0,T))}\mu(M)^{1-2/q},$$
$$q>2. \tag{7.154}$$

Next, the estimate (7.154) is iterated in order to conclude that $\xi_M = 0$. Here we should modify the classical argument from [7], pages 102-103, and adapt it to our situation.

We note that, after making appropriate extensions,

$$\|\Psi\varphi\|_{L^4(Z^+\times(0,T))} \leq c_0\|\Psi\varphi\|^{1/2}_{L^2(Z^+\times(0,T))}\|\Psi\varphi\|^{1/2}_{H^1(Z^+\times(0,T))} \leq c_0\varepsilon^{-\alpha/4}\|\varphi\|_{V_2}, \quad (7.155)$$

$\forall\varphi \in V_2$, $\varphi|_{x=0} = 0$. As in [7], page 102, now we take the sequence of levels $k_h = M(2 - 2^{-h}),\ h = 0, 1, \ldots$. Then

$$(k_{h+1} - k_h)\mu^{1/4}(k_{h+1}) \leq \|\Psi\xi_{k_h}\|_{L^4(Z^+\times(0,T))} \leq \frac{\bar{\beta}\varepsilon^{-\alpha/4}}{k_{h+1} - k_h}\|\xi_{k_h}\|_{V_2} \quad (7.156)$$

and

$$\mu^{1/4}(k_{h+1}) \leq 2^h\frac{2\bar{\beta}\beta_0\|\tau^k\Phi^\varepsilon_3\Psi\|_{L^q(Z^+\times(0,T))}\varepsilon^{-\alpha/4}}{M}\mu^{(1+\kappa)/4}(k_h),\ \kappa = 1 - 2/q > 0. \quad (7.157)$$

$\mu^{1/4}(k_{h+1})$ will tend to zero for $h \to \infty$ if $\mu^{1/4}(M)$ satisfies

$$\mu^{1/4}(M) \leq \left(\frac{2\bar{\beta}\beta_0\|\tau^k\Phi^\varepsilon_3\Psi\|_{L^q(Z^+\times(0,T))}\varepsilon^{-\alpha/4}}{M}\right)^{-1/\kappa} 2^{-1/\kappa^2} \quad (7.158)$$

(7.158) is satisfied if M equals the right hand side of the estimate (7.150). ■Next result concern higher order norms. It it not very satisfactory for large α and we state it without giving a proof, which follows from the demonstrations given above.

Theorem 4.7.4 *Let c^{Mau} be given by (6.75), let c^{eff}_1 be given by (6.74), $\bar{c}^{eff}_1$ by (6.114) and c^{eff}_2 by (7.136). Then we have*

$$\|t^5\partial_x(c^\varepsilon - c^{eff}_1(x,t;\varepsilon) - \bar{c}^{eff}_1)\|_{L^\infty(0,T;L^2_{loc}(\mathbb{R}_+\times(0,1))} \leq C\big(\varepsilon^{4-15\alpha/4}H(1-\alpha)+ \varepsilon^{(1-\alpha/2)/2}H(\alpha-1)\big) \quad (7.159)$$

$$\|t^5\partial_t\big(c^\varepsilon - c^{eff}_1(x,t;\varepsilon) - \bar{c}^{eff}_1\big)\|_{L^2(0,T;L^2_{loc}(\mathbb{R}_+\times(0,1))} \leq C\big(\varepsilon^{4-15\alpha/4}H(1-\alpha) + \varepsilon^{(1-\alpha/2)/2}H(\alpha-1)\big) \quad (7.160)$$

Final improvement concerns the $L^\infty(L^2)$-norma for small values of α. As mentioned in the proof of Theorem 4.7.1, the reason was that $\tilde{F}^\varepsilon_1$ and $\tilde{F}^\varepsilon_3$ didn't have zero means with respect to y. Nevertheless, when computing the term c^2 in the asymptotic expansion, there was a liberty in adding an arbitrary function C_2 of x and t. This function can be chosen such that the appropriate means are zero and estimates (7.137)-(7.139) are multiplied by $\varepsilon^{1-\alpha/2}$. Unfortunately, there is a new contribution of the form $QP_2(y)\partial_x C_2$. Its norm destroys the estimate for $\alpha \geq 4/5$. Since this amelioration isn't of real importance we just give it as a result. Proof is completely analogous to the preceding ones.

Corollary 4 *Let c^{Mau} be given by (6.75), let c_1^{eff} be given by (6.74), $\bar{c}_1^{eff}$ by (6.114) and c_2^{eff} by (7.136). Let the polynomials $P_j(y)$ be defined by (7.140) and after (7.142). Finally, let C_2 be given by*

$$\frac{\partial C_2}{\partial t} + \frac{2Q}{3}\frac{\partial C_2}{\partial x} - \varepsilon^\alpha D \frac{\partial^2 C_2}{\partial x^2} = -\frac{Qk_0}{D}\zeta(t)\Big\{ \partial_{xx} c^{Mau} \frac{Q}{D} \int_0^1 (1-y^2)\tilde{P}_6(y)\, dy$$
$$+(\partial_{xt} c^{Mau} \frac{1}{2D} - \varepsilon^\alpha \partial_{xxx} c^{Mau} \frac{1}{2}) \int_0^1 (1-y^2) P_4(y)\, dy + (\frac{Q}{45D}\partial_x c^{Mau} -$$

$$\frac{k_0}{6D} c^{Mau}) \int_0^1 (1-y^2) P_2(y)\, dy \Big\} - \frac{Q^2}{D^2}\zeta(t)\Big\{ -\partial_{xxx} c^{Mau} Q \int_0^1 (1-y^2) P_8(y)\, dy$$
$$+(\partial_{xxt} c^{Mau} - D\varepsilon^\alpha \partial_{xxxx} c^{Mau}) \int_0^1 (1-y^2) P_6(y)\, dy \Big\} \ \ in \ I\!R_+ \times (0,T), \tag{7.161}$$

$$\partial_x C_2 \in L^2(I\!R_+ \times (0,T)), \quad C_2|_{t=0} = 0, \quad C_2|_{x=0} = 0. \tag{7.162}$$

Then for $\alpha \in [0, 4/5]$ we have

$$\|t^5 (c^\varepsilon - c_1^{eff} - \bar{c}_1^{eff} - c_2^{eff} - C_2)\|_{L^\infty(0,T; L^2_{loc}(I\!R_+ \times (0,1))} \le C\varepsilon^{5-17\alpha/4} \tag{7.163}$$
$$\|t^5 \partial_y \big(c^\varepsilon - -c_1^{eff} - \bar{c}_1^{eff} - c_2^{eff}\big)\|_{L^2(0,T; L^2_{loc}(I\!R_+ \times (0,1))} \le C\varepsilon^{6-19\alpha/4} \tag{7.164}$$
$$\|t^5 \partial_x \big(c^\varepsilon - c_1^{eff} - \bar{c}_1^{eff} - c_2^{eff} - C_2\big)\|_{L^2(0,T; L^2_{loc}(I\!R_+ \times (0,1))} \le \varepsilon^{5-19\alpha/4} \tag{7.165}$$

Acknowledgement. The authors are grateful to Dr. I.S. Pop for help with Section 2.

Bibliographie

[1] R. Aris, *On the dispersion of a solute in a fluid flowing through a tube*, Proc. Roy. Soc. London Sect A., 235 (1956), pp. 67-77.

[2] A. Bourgeat, M. Jurak, A.L. Piatnitski, *Averaging a transport equation with small diffusion and oscillating velocity* , Math. Meth. Appl. Sci., Vol. 26 (2003), pp. 95-117.

[3] C.J. van Duijn , I.S. Pop, *Crystal dissolution and precipitation in porous media : pore scale analysis* , preprint RANA 03-27, TU Eindhoven, December 2003.

[4] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, Providence, 1998.

[5] E. Grenier, *Boundary Layers for Parabolic Regularization of Totally Characteristic Quasilinear Parabolic Equations* , J. Math. Pures Appl. , Vol. 76 (1997), p. 965-990.

[6] E. Grenier, F. Rousset, *Stability of One-Dimensional Boundary Layers by Using Green's Functions* , Comm. on Pure and Applied Mathematics, Vol. LIV (2001), p. 1343-1385.

[7] O.A. Ladyzhenskaya, V.A. Solonnikov and N.N. Uralceva, *Linear and Quasi-Linear Equations of Parabolic Type*, American Mathematical Society, Providence, 1968.

[8] R. Mauri, *Dispersion, convection and reaction in porous media* , Phys. Fluids A (1991), p. 743-755.

[9] G.N. Mercer, A.J. Roberts, *A centre manifold description of contaminant dispersion in channels with varying flow profiles* , SIAM J. Appl. Math. , Vol. 50 (1990), p. 1547-1565.

[10] O.A. Oleinik and G.A. Iosif'jan, *On the behavior at infinity of solutions of second order elliptic equations in domains with noncompact boundary* , Math. USSR Sbornik, Vol. 40 (1981), p. 527 - 548.

[11] M.A. Paine, R.G. Carbonell, S. Whitaker, *Dispersion in pulsed systems – I, Heterogeneous reaction and reversible adsorption in capillary tubes*, Chemical Engineering Science, Vol. 38 (1983), p. 1781-1793.

[12] A.I. Pyatniskii, *Averaging singularly perturbed equation with rapidly oscillating coefficients in a layer* , Math. USSR Sbornik, Vol. 49 (1984), p. 19 - 40.

[13] A.I. Pyatniskii, personal communication.

[14] J. Rubinstein, R. Mauri, *Dispersion and convection in porous media* , SIAM J. Appl. Math. , Vol. 46 (1986), p. 1018 - 1023.

[15] G.I. Taylor, *Dispersion of soluble matter in solvent flowing slowly through a tube* , Proc. Royal Soc. A , Vol. 219 (1953), p. 186-203.

Chapitre 5

L'écoulement de Stokes

Sommaire

5.1 Introduction **92**

5.2 Problème continu, résultat de régularité et formulation mathématique **94**

5.3 Approximation par élément fini du problème de Stokes **96**

5.4 Approximation de type Petrov-Galerkin avec stabilisation bulle . **97**

5.4.1 condition inf-sup et estimations d'erreur 97

5.4.2 Fonctions de base 98

5.4.3 La bulle à résidu nulle (RFB) appliquée à l'équation de convection-diffusion 101

5.4.4 La fonction bulle 103

5.4.5 Formulation algébrique 106

5.5 Résultats Numériques **108**

5.6 Conclusion **115**

5.7 Introduction **116**

5.8 Continuous problem, regularity result and mathematical formulation **117**

5.9 Finite Element approximation of the Stokes problem **119**

5.10 Petrov-Galerkin approximation of the problem with bubble stabilization **120**

5.10.1 inf-sup condition and error estimates 120

5.10.2 Shape functions 122

5.10.3 Residual free Bubble applied to Advective-Diffusive equation 124

5.10.4 The bubble function 126

5.10.5 Algebraïc form 129

5.11 Results 131

5.12 Conclusion 133

---

résumé-abstract : Nous étudions dans cette partie les équations de Stokes. Nous nous plaçons dans le contexte de la mécanique des fluides et rappelons les résultats les plus importants de l'analyse du problème. Après une explication générale de la méthode des éléments finis, nous donnons une analyse de l'élément $\mathbb{P}_1$+bubble-$\mathbb{P}_1$ appliqué au problème de Stokes, en donnant les détails des calculs des foncions de base dans le cas tridimensionnelle et des cas tests variés pour valider l'approximation.

5.1 Introduction

Les premiers modèles mathématiques décrivant le mouvement d'un fluide sont alloués historiquement à Euler et Lagrange et datent du XIIIème siècle. La description eulérienne du mouvement d'un fluide donne une distribution spatiale des vitesses tandis que la description lagrangienne donne la vitesse en chacun des points du fluide. Dans la première le repère est fixe alors que dans la seconde il est attaché à la particule. En fait les deux descriptions proviennent de la loi de Newton qui lie l'accélération aux forces qui agissent sur la particule. En mécanique des fluides on ne considère que la description eulérienne. Au niveau le plus élémentaire le comportement d'un fluide laminaire est décrit par un ensemble d'équations fondamentales. Parmi elles on trouve,

– la continuité du fluide,
– l'équation du mouvement,

En plus de deux relations qui donnent la déformation du fluide et le taux de déformation. En relaxant les deux premières et en utilisant les deux relations pour éliminer la déformation et le taux de déformation on obtient les équations de Navier-Stokes (7.28) pour un fluide incompressible trouvées en 1821 par C.L. Navier[1] et G.S. Stokes[2].

$$u.\nabla u - \nu\triangle u + \nabla p = f \tag{1.1}$$

Au niveau microscopique , et dans le cas d'un écoulement lent on peut négliger les termes d'inertie dans l'équation, le fluide est alors considéré comme visqueux et les équations de Navier-Stokes se réduisent à l'équation de Stokes (7.29).

$$-\nu\triangle u + \nabla p = f \tag{1.2}$$

Le but de cette discussion n'est pas d'entamer une révolution dans l' abondante littérature qui traite du sujet, mais de fournir une revue du champ mathématique aussi complète que possible dans les références de la théorie aux applications. Notre intérêt serait plus de souligner une méthode numérique pour résoudre les équations, que nous utiliserons dans le chapitre suivant, basée sur une méthode spécifique de type élément fini, efficace, précise et qui donne des résultats aussi bien en deux dimensions qu'en trois dimensions sous des hypothèse raisonnables et des conditions aux limites choisies. Nous voulons obtenir une bonne approximation de la vitesse dans un polygone convexe de la même façon qu'une approximation de la pression satisfaisante.

Ce travail est motivé par l'étude d'écoulement complexe que nous rencontrerons plus tard dans le document, la convection diffusion de concentrations de substance chimique, à l'échelle du pore.

---

[1]Claude Louis Navier (1785-1836), étudiant de Fourier à l' École polytechnique, Navier entre à l'École des Ponts et Chaussées en 1804. Navier prend en charge les cours de mécanique appliquée à l'École des Ponts et Chaussées en 1819, devenant professeur en 1830. Il a changé la faọn traditionnelle d'enseigner afin d'insister plus sur les aspects physique et mathématique de la discipline. Il trouva l'équation qui porte son nom en 1821 pour l'écoulement d'un fluide incompressible. Jusqu'à aujourd'hui l'étude des équations de Navier-Stokes reste encore inachevée à cause de leur complexité.

[2]Georges Gabriel Stokes (1819-1903), étudiant au college de Bristol à Londres à l'âge de 16 ans, Stokes entre à Cambridge en 1837 . Encadré par William Hopkins, il est diplômé en tant que Senior Wrangler (le premier dans les plus titrés) dans le Tripos mathématique en 1841, et il a été le premier titulaire du prix Smith . Un de ces articles les plus célèbres est intitulé “On the steady motion of incompressible fluids” publié en 1842. Après qu'il ait trouvé ces fameuses équations, Stokes découvrit qu'il n'était pas le premier à l'obtenir. En 1849 Stokes est nommé Professeur Lucasien de Mathématiques à Cambridge (la chair de Newton). Il obtint la médaille Rumford de la société royale en 1852.

La structure de notre discussion sera la suivante, en partant de la forme continue du problème, les espaces de solution H^2 et H^1 conduiront à des résultats de régularité importants analysés dans la première partie. La seconde traitera des propriétés et de l'approximation par élément fini de ces espaces. Nous rappellerons les résultats les plus importants de l'analyse numérique du problème. Après une explication générale de la méthode des éléments finis, nous donnons une analyse de l'élément $\mathbb{P}_1$+bubble-$\mathbb{P}_1$ appliqué au problème de Stokes, en donnant les détails des calculs des fonctions de base dans le cas tridimensionnelle et des cas tests variés pour valider l'approximation.

5.2 Problème continu, résultat de régularité et formulation mathématique

La solution du problème de Stokes admet des dérivées secondes de racine intégrables. Un tel résultat de régularité est important pour l'analyse des méthodes numériques de résolution de ces équations. Ce résultat est du à R.B. Kellog et J.E. Osborn dans [1]. Introduisons la formulation du problème. Le système d'équations différentielles de Stokes s'écrit :

$$\left\{\begin{array}{rcll} -\nu\triangle u+\nabla p & = & f & \text{in } \Omega\subset\mathbb{R}^d, d=2,3 \\ \text{div}u & = & 0 & \text{sur } \Omega \\ u & = & 0 & \text{ou des conditions aux limites sur } \partial\Omega \end{array}\right. \tag{2.3}$$

$f:\Omega\to\mathbb{R}^d$ est une fonction donnée tandis que $u:\Omega\to\mathbb{R}^d$ et $p:\Omega\to\mathbb{R}$ sont des inconnues. C'est le cas de l'écoulement de fluide avec une très importante viscosité (un nombre de Reynolds petit $Re=\mu\rho L/U$). Ici, nous considérons une condition aux limites de type Dirichlet mais aussi d'autres conditions impliquant les dérivées normales de u ou des combinaisons linéaires entre u et p sur $\partial\Omega$ sont aussi possibles.

Nous avons les théorèmes suivants,

Théorème 5.2.1 *Si $f\in L^2(\Omega)$, et soient u et p des solutions générales de (8.30). Nous supposons que $\partial\Omega$ est régulière, alors $u\in H^2(\Omega)$ et $p\in H^1(\Omega)$ et il existe une constante C dépendant seulement de Ω telle que,*

$$\|u\|_{H^2(\Omega)}+\|\nabla p\|_{H^1(\Omega)}\leq C\left(\|f\|_{L^2(\Omega)}+\|p\|_{L^2(\Omega)}\right). \tag{2.4}$$

Théorème 5.2.2 *Si Ω est un polygone convexe alors, sous les mêmes hypothèses du théorème 5.8.1,*

$$\|u\|_{H^2(\Omega)} + \|\nabla p\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)}\right). \tag{2.5}$$

Cela nous donne des espaces continus $H^2(\Omega)$ et $H^1(\Omega)$ auxquels appartiennent les solutions continues du problème de Stokes. Les méthodes numériques sont basées sur ce principe : Trouver des espaces discrets pour la vitesse et la pression qui réalisent des approximations correctes des espaces continus, et l'analyse numérique se concentre sur les critères de distance entre les solutions discrètes et les solutions continues.

Remarque 6 *Dans la plupart des livres d'analyse numérique le problème de Stokes s'exprime légèrement différemment de (8.30) voir [3]. Soit f* $f \in L^2(\Omega)$ *ou* $H^{-1}(\Omega)$*, trouver* $u \in H^1(\Omega)$ *et* $p \in L^2(\Omega)$ *tels que,*

$$\begin{cases} -\nu\triangle u + \nabla p &= f \quad \textit{dans } \Omega \\ \textit{div}(u) &= 0 \quad \textit{dans } \Omega \end{cases} \tag{2.6}$$

par définition de $H^m(\Omega)$*, si* $f \in L^2(\Omega)$ *alors* $u \in H^2(\Omega)$ *et* $p \in H^1(\Omega)$*, si* $f \in H^{-1}(\Omega)$ *alors* $u \in H^1(\Omega)$ *et* $p \in L^2(\Omega)$*.*

Maintenant définissons rigoureusement la formulation variationnelle du problème (8.33). En s'inspirant de [5] nous supposons $f \in L^2(\Omega)$ et $v \in H_0^1(\Omega)$. Nous appliquons la formule de Green à la première équations de (8.33).

$$\nu(\nabla u, \nabla v) - (p, \mathrm{div} v) = (f, v) \quad \forall v \in \mathcal{D}(\Omega), \tag{2.7}$$

et ainsi,

$$\nu(\nabla u, \nabla, v) = (f, v) \quad \forall v \in \{v \in \mathcal{D}(\Omega) | \mathrm{div} v = 0\}.$$

Définissons les espaces $V_{div} = \{v \in V | \mathrm{div} v = 0\}$ un sous ensemble fermé de V.

La forme bilinéaire $a(w, v) = \nu(\nabla w, \nabla v)$ $w, v \in V$ est coercive sur $V_{div \times V_{div}}$ sachant que l'application $v \to (f, v)$ est linéaire et continue sur V_{div}, le lemme de Lax-Milgram nous assure que le problème

$$\text{find } u \in V_{div} : a(u, v) = (f, v) \quad \forall v \in V_{div} \tag{2.8}$$

admet une solution unique.

Théorème 5.2.3 *Soit Ω un domaine borné de $\mathbb{R}^d$ de frontière lipschitzienne continue , et pour tout $f \in L^2(\Omega)$, soit u la solution de (8.35). Alors il existe une fonction $p \in L^2(\Omega)$, qui est unique à une constante additive près, telle que,*

$$a(u,v) - (p, divv) = (f,v) \quad \forall v \in V. \tag{2.9}$$

C'est une des formulations variationnelles les plus couramment utilisées. Une différente approche est de définir l'espace $Q = L_0^2(\Omega)$,

$$L_0^2(\Omega) = \{q \in L^2(\Omega) | \int_\Omega q = 0\},$$

et la forme bilinéaire

$$b(v,q) = -(q, \mathrm{div} v), \quad v \in V, q \in Q,$$

du théorème ci-dessus nous déduisons que le problème,

$$\begin{cases} \text{find } u \in V, p \in Q : & & & \\ a(u,v) + b(v,p) & = & (f,v) & \forall v \in V \\ b(u,q) & = & 0 & \forall q \in Q \end{cases} \tag{2.10}$$

admet une solution unique, la pression définie à une constante additive près. La solution de (8.37) est aussi une solution de (8.34). Cette équation est valable presque partout pour $f \in L^2(\Omega)$. (8.37) est la formulation faible du problème de Stokes.

5.3 Approximation par élément fini du problème de Stokes

Pour plus de détails nous citons comme référence le livre [4]. Les méthodes non-conformes sont les méthodes d'éléments finis les plus naturelles au sens qu'elles permettent de construire un sous-ensemble de dimension fini de l'espace des fonctions que l'on veut approcher. Soit $\mathcal{T}_h$ une partition de Ω en triangles et K, $\hat{K}$ deux triangles. Un approximation non-conforme de $H^1(\Omega)$ est un espace de fonctions continues définies par un nombre fini de degrés de libertés. Pour des éléments triangulaires il est habituel d'utiliser des éléments finis linéaires par morceaux sur K. La condition de continuité entre K et $\hat{k}$ est obtenue par une transformation affine. Soit $\mathbb{P}_1(K)$ l'espace des polynôme de degré ≤ 1 de dimension 3 dans $\mathbb{R}^2$ et 4 dans $\mathbb{R}^3$. Pour définir un élément fini on doit spécifier trois choses conf. [2].

- La géométrie basée sur un élément de référence $\hat{K}$ et un changement de variables $F(\hat{x})$ tel que $K = F(\hat{K})$.
- Un ensemble $\hat{P}_1$ de polynômes sur $\hat{K}$. Pour $\hat{p} \in \hat{P}_1$ on définit $p = \hat{p}^\circ F^{-1}$
- Un ensemble de degrés de liberté $\hat{\Sigma}$, qui est une ensemble de formes linéaires $\{\hat{l}_i\}_{1 \leq i \leq 3}$ sur $\hat{P}_1$.

L'ensemble est unisolvant quand les formes linéaires sont linéairement indépendantes. Un élément fini est de type Lagrange si ses degrés de liberté sont des noeuds. Soit un ensemble $\{\hat{a}_i\}_{1 \leq i \leq 3}$ de points dans $\hat{K}$ et soit $\hat{l}_i(\hat{p}) = \hat{p}(a_i) \quad 1 \leq i \leq 3$. Approcher $H^1(\Omega)$ avec un élément de type Lagrange est suffisant. Cependant le choix d'un élément fini $\mathbb{P}_1$ pour le problème de Stokes n'est pas suffisant car il ne satisfait pas la condition inf-sup qui assure que le problème est bien posé (en particulier que l'opérateur linéaire associé au système linéaire total est inversible). Cependant nous pouvons enrichir cet espace $\mathbb{P}_1$ avec des fonctions bulles de degré deux qui s'annulent sur ∂K. Cela donne une approximation stable et convergente du problème de Stokes.

5.4 Approximation de type Petrov-Galerkin avec stabilisation bulle

5.4.1 condition inf-sup et estimations d'erreur

Nous procédons de la même façon que dans [6]. Soit une triangulation $\mathcal{T}_h$ de $\bar{\Omega}$ et nous approchons la vitesse sur chaque élément K par un polynôme

$$\mathcal{P}_1(K) = [\mathbb{P}_1 + span\{\lambda_1, \lambda_2, \lambda_3\, \lambda_4]^3, \tag{4.11}$$

et la pression par un polynôme $\mathbb{P}_1$. Les espaces d'éléments finis sont définis par

$$V_h = \left\{ v \in \mathcal{C}^0(\bar{\Omega})^3,\ v|_K \in \mathcal{P}_1(K),\ v|_\Gamma = 0 \right\},$$
$$Q_h = \left\{ q \in \mathcal{C}^0(\bar{\Omega})^3, q|_K \in \mathbb{P}_1(K),\ \forall K \in \mathcal{T}_h \right\}$$
$$P_h = Q_h \cap L^2_0(\Omega).$$

Les degrés de liberté sont les valeurs de la vitesse au centre et aux sommets de K et le valeurs de la pression aux sommets de K. Et le problème discret s'énonce :

$$\text{Trouver une paire } (u_h, v_h) \in V_h \times P_h \text{ satisfaisant :}$$
$$\begin{cases} a(u_h, v_h) - (p_h, \text{div} v_h) & = & (f, v_h) & \forall v_h \in V_h \\ (\text{div} u_h, q_h) & = & 0 & \forall q_h \in P_h \end{cases} \tag{4.12}$$

où a est donné par $a(u, v) = \nu(\nabla u, \nabla v)$.

Lemme 5.4.1 *Si la triangulation $\mathcal{T}_h$ est régulière, le doublet d'espaces (V_h, P_h) satisfait la condition inf-sup :*

$$\sup_{v_h \in V_h} \frac{(q_h, div v_h)}{\|v_h\|_{H^1(\Omega)}} \geq \beta \|q_h\|_{L_0^2(\Omega)} \qquad \forall q_h \in P_h$$

voir [6] pour les détails de la preuve.

Nous rappelons le théorème de [6],

Théorème 5.4.1 *Soit Ω un polygone borné et soit la solution (u, p) du problème de Stokes satisfaisant $u \in [H^2(\Omega) \cap H_0^1(\Omega)]^d,\ p \in H^1(\Omega) \cap L_0^2(\Omega)$. Si la triangulation $\mathcal{T}_h$ est régulière, la solution (u_h, p_h) du problème (4.12) satisfait les majorations d'erreur :*

$$\|u - u_h\|_{H^1(\Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq Ch \left\{ \|u\|_{H^2(\Omega)} + \|p\|_{H^1(\Omega)} \right\}. \tag{4.13}$$

Si Ω est convexe, nous avons l'estimation L^2 :

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \left\{ \|u\|_{H^2(\Omega)} + \|p\|_{H^1(\Omega)} \right\}. \tag{4.14}$$

5.4.2 Fonctions de base

Cette partie sera traitée en détails, en partant de la formulation variationnelle du problème de Stokes, mais découplée entre la partie linéaire provenant de la discrétisation $\mathbb{P}_1$ et le terme quadratique provenant de la fonction bulle, en finissant par la formulation algébrique du problème de Stokes. Nous procédons comme dans [18] et donnons la valeur exacte des fonctions de bases de chacun des termes de l'équation dans le cas d'un domaine $\Omega \cup \mathbb{R}^3$. Nous considérons la formulation variationnelle du problème de Stokes tirée de (4.12) où $\Omega \subset \mathbb{R}^3$ est un domaine polygonal. Si (u_h, p_h), $(v_h, q_h) \in V_h \times P_h$, on peut définir :

$$\begin{aligned} B(u_h,p_h;v_h,q_h) &= \nu\int_\Omega \nabla u_h : \nabla v_h dx - \int_\Omega \mathrm{div} v_h\, p_h\, dx && (4.15)\\ &- \int_\Omega \mathrm{div} u_h\, q_h\, dx \\ \text{and } F(v_h,q_h) &= \int_\Omega f v_h\, dx && (4.16)\end{aligned}$$

Le problème (4.12) peut s'écrire

$$\text{Trouver } u_h \in V_h,\ p_h \in P_h \text{ tel que,}$$

$$B(u_h,p_h;v_h,q_h) = F(v_h,q_h) \quad \forall v_h \in V_h\ \forall q_h \in P_h. \tag{4.17}$$

Nous considérons maintenant les espaces :

$$V_h^L = \left\{ v_L = (v_L^1, v_L^2, v_L^3) \in [\mathcal{C}^0(\bar\Omega)]^3 : v_L^i|_T \in \mathbb{P}_1, i=1,2,3, T \in \mathcal{T}_h \right\}$$

et $\boldsymbol{b}_T = (b_T^1, b_T^2, b_T^3) \in [H_0^1(T)]^3$ b_T^i prenant ses valeurs sur T, et nous définissons $\boldsymbol{b}_T^1 = (b_T^1,0,0)$, $\boldsymbol{b}_T^2 = (0,b_T^2,0)$, $\boldsymbol{b}_T^3 = (0,0,b_T^3)$ et

$$\boldsymbol{B} = vect\{b_T^1, b_T^2, b_T^3, T \in \mathcal{T}_h\}$$

L'espace d'approximation des vitesses est défini par $V_h = V_h^L \oplus \boldsymbol{B}$. Et définissons

$$P_h = \left\{ q_h \in C^0(\overline{\Omega}) : \int_\Omega q_h = 0, q_{h|T} \in \mathbb{P}_1, T \in \mathcal{T}_h \right\}$$

Nous avons vu dans le paragraphe précédent que ces fonction vérifient la condition inf-sup et conduisent à une approximation stable et convergente du problème. Nous exprimons v_h sous la forme d'une partie linéaire v_L et d'une partie "bulle" :

$$\boldsymbol{v}_h = \boldsymbol{v}_L + \sum_{T\in\mathcal{T}_h} (c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) \tag{4.18}$$

où $\boldsymbol{v}_L \in V_h^L$ et $c_T^1, c_T^2, c_T^3 \in \mathbb{R}\ \forall T \in \mathcal{T}_h$. Le problème (4.17) prend la forme :

$$\begin{cases} Trouver\ \boldsymbol{u}_h = \boldsymbol{u}_L + \sum_{T\in\mathcal{T}_h} (c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) \in V_h, p_h \in P_h\ tel\ que \\ B(\boldsymbol{u}_h, p_h; \boldsymbol{v}_L, q_h) = F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h \\ B(\boldsymbol{u}_h, p_h; \boldsymbol{b}_T^i, q_h) = F(\boldsymbol{b}_T^i, q_h)\ \forall T \in \mathcal{T}_h, q_h \in P_h, i = 1\dots 3 \end{cases} \tag{4.19}$$

Maintenant nous éliminons le terme "bulle" de l'équation. pour $i = 1$, $q_h = 0$ nous obtenons :

$$\nu \int_T \nabla(\boldsymbol{u}_L + c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) : \nabla \boldsymbol{b}_T^1 d\tau - \int_T div\ \boldsymbol{b}_T^1 p_h d\tau = \int_T \boldsymbol{f}.\boldsymbol{b}_T^1 d\tau$$

u_L est linéaire par morceaux, f est constante impliquant,

$$\nu \int_T \nabla \boldsymbol{u}_L : \nabla \boldsymbol{b}_T^1 d\tau = 0, \ \ \nu \int_T \nabla \boldsymbol{b}_T^1 : \nabla \boldsymbol{b}_T^1 d\tau = \nu \int_T |\nabla b_T^1|^2 d\tau,$$

$$\int_T \text{div } \boldsymbol{b}_T^1 p_h d\tau = \int_T \frac{\partial b_T^1}{\partial x} p_h d\tau = -\int_T b_T^1 \frac{\partial p_h}{\partial x} d\tau = -\left(\frac{\partial p_h}{\partial x}\right)_{|T} \int_T b_T^1 d\tau$$

L'équation pour $\boldsymbol{b}_T^1$ devient :

$$\nu c_T^1 \int_T |\nabla b_T^1|^2 d\tau + \left[\frac{\partial p_h}{\partial x}\right]_{|T} \int_T b_T^1 d\tau = f^1_{|T} \int_T b_T^1 d\tau$$

c_T^1 peut s'exprimer par :

$$c_T^1 = -\frac{\int_T b_T^1 d\tau}{\int_T |\nabla b_T^1|^2 d\tau} \left[\frac{\partial p_h}{\partial x} - f^1\right]_{|T}$$

De la même façon nous obtenons :

$$c_T^2 = -\frac{\int_T b_T^2 d\tau}{\int_T |\nabla b_T^2|^2 d\tau} \left[\frac{\partial p_h}{\partial y} - f^2\right]_{|T}$$

et

$$c_T^3 = -\frac{\int_T b_T^3 d\tau}{\int_T |\nabla b_T^3|^2 d\tau} \left[\frac{\partial p_h}{\partial z} - f^3\right]_{|T}$$

Maintenant, la première équation de (4.19) donne,

$$\nu \int_\Omega \nabla(\boldsymbol{u}_L + \boldsymbol{u}_B) : \nabla \boldsymbol{v}_L d\tau - \int_\Omega div(\boldsymbol{v}_L) p_h d\tau$$

$$-\int_\Omega div\ (\boldsymbol{u}_L + \boldsymbol{u}_B) q_h d\tau = \int_\Omega \boldsymbol{f}.\boldsymbol{v}_L d\tau \ \ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h$$

Nous remarquons que

$$\nu \int_\Omega \nabla \boldsymbol{u}_B : \nabla \boldsymbol{v}_L d\tau = 0,$$

$$\begin{aligned}
\int_\Omega div(\boldsymbol{u}_B) q_h d\tau &= \sum_{T\in\mathcal{T}_h}\left[\sum_{i=1}^{3} c_T^i \int_T div\ \boldsymbol{b}_T^i q_h d\tau\right] \\
&= \sum_{T\in\mathcal{T}_h}\left[\sum_{i=1}^{3} c_T^i \int_T \frac{\partial b_T^i}{\partial x_i} q_h d\tau\right] \\
&= -\sum_{T\in\mathcal{T}_h}\left[\sum_{i=1}^{3} c_T^i \int_T b_T^i \frac{\partial q_h}{\partial x_i} d\tau\right] \\
&= -\sum_{T\in\mathcal{T}_h}\left[\sum_{i=1}^{3} c_T^i \int_T \left(\frac{\partial q_h}{\partial x_i}\right)_{|T} b_T^1 d\tau\right] \\
&= \sum_{T\in\mathcal{T}_h}\left[\sum_{i=1}^{3} \frac{\left(\int_T b_T^i d\tau\right)^2}{\int_T |\nabla b_T^i|^2 d\tau}\left[\frac{\partial p_h}{\partial x_i} - f^i\right]_{|T}\left[\frac{\partial q_h}{\partial x_i}\right]_{|T}\right]
\end{aligned}$$

Pour obtenir finalement,

$$\begin{cases}
Trouver\ \boldsymbol{u}_L \in V_h^L, p_h \in P_h\ tels\ que \\
B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) + \sum_{T\in\mathcal{T}_h} \frac{1}{|T|}\left[\sum_{i=1}^{3} \frac{\left(\int_T b_T^i d\tau\right)^2}{\int_T |\nabla b_T^i|^2 d\tau} \int_T \left[\frac{\partial p_h}{\partial x_i} - f^i\right]\left[\frac{\partial q_h}{\partial x_i}\right] d\tau\right] \\
= F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h
\end{cases} \tag{4.20}$$

qui peut s'exprimer par

$$\begin{cases}
Trouver\ \boldsymbol{u}_L \in V_h^L, p_h \in P_h\ tels\ que \\
B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) \\
+ \sum_{T\in\mathcal{T}_h} \frac{1}{|T|} \frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\int_T |\nabla \boldsymbol{b}_T|^2 d\tau} \int_T \left[grad(p_h) - \boldsymbol{f}\right]\left[-grad(q_h)\right] d\tau \\
= F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h
\end{cases} \tag{4.21}$$

5.4.3 La bulle à résidu nulle (RFB) appliquée à l'équation de convection-diffusion

Nous faisons ici une petite digression pour souligner l'effet stabilisant de l'ajout de la contribution "bulle" pour une équation de convection-diffusion. Dans certains cas la méthode RFB (de la bulle à résidu nulle) est similaire à la méthode SUPG. Les simulations indiquent une amélioration de la précision des résultats. D'après [8] si nous utilisons une interpolation bilinéaire de l'espace de réduction dans une méthode de Galerkin et que nous l'enrichissons avec des bulles de résidu nulle, nous obtenons une stabilisation différente de la diffusion des caractéristiques. Et dans le cas où l'espace de réduction est linéaire nous retrouvons la méthode SUPG avec des paramètres optimaux. Nous décrivons brièvement la méthode.

Nous prenons l'opérateur différentiel linéaire suivant,

$$L = -D\triangle + \nabla.(q.)$$

$D > 0$ est la diffusion, $\vec{q}$ la vitesse, u la concentration

$u = 0$ sur $\partial\Omega$.

La méthode de Galerkin consiste à trouver $u_h \in V_h$ tel que,

$$a(u_h, v_h) = (Lu_h, v_h) = (f, v_h) \quad \forall v_h \in V_h \tag{4.22}$$

Avec V_h défini précédemment. Chaque membre u_h et v_h de l'espace des fonctions V_h est exprimé par par des polynômes continus linéaires ou bilinéaires par morceaux plus des fonctions bulles défini ci-dessous, i.e.,

$$u_h = u_1 + u_b \tag{4.23}$$

La partie bulle de l'espace admet une condition de Dirichlet homogène sur la frontière de chaque élément K, i.e.,

$$u_b = 0 \quad \text{sur } \partial K \tag{4.24}$$

En prenant v_1 comme fonction test on a,

$$D(\nabla u_1, \nabla v_1) + D(\nabla u_b, \nabla v_1) + (\nabla.(qu_1), v_1) + \sum_K (\nabla.(qu_b), v_1)_K = (f, v_1)$$

En intégrant par parties,

$$D(\nabla u_1, \nabla v_1) + (\nabla.(qu_1), v_1) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_1)_K = (f, v_1)$$

La partie à résidu nulle de la solution, u_b peut être calculée comme solution de

$$-D\triangle u_b + \nabla.(qu_b) = -\nabla.(qu_1) + D\triangle u_1 + f$$

En testant avec v_b on a,

$$D(\nabla u_1, \nabla v_b) + D(\nabla u_b, \nabla v_b) + (\nabla.(qu_1), v_b) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_b)_K = (f, v_b)$$

En intégrant par parties,

$$D(\nabla u_1, \nabla v_b) + (\nabla.(qu_1), v_b) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_b)_K = (f, v_b)$$

Et la formulation variationnelle stabilisée de l'équation de convection-diffusion s'écrit,

$$\begin{aligned} & D(\nabla u_1, \nabla v_1) + D(\nabla u_1, \nabla v_b) + D(\nabla u_b, \nabla v_1) + D(\nabla u_b, \nabla v_b) \\ + \quad & (\nabla.(qu_1), v_1) + (\nabla.(qu_1), v_b) + \sum_K ((\nabla.(qu_b), v_1) + (\nabla.(qu_b), v_b))_K \\ = \quad & (f, v_1) + (f, v_b) \end{aligned} \tag{4.25}$$

5.4.4 La fonction bulle

La fonction bulle sur le tétraèdre est définie par,

$$\boldsymbol{b}_T = 4^4 \phi_1 \phi_2 \phi_3 \phi_4 \;\; avec$$

$$\phi_i = \frac{\begin{vmatrix} 1 & x & y & z \\ 1 & x_i & y_i & z_i \\ 1 & x_{i+1} & y_{i+1} & z_{i+1} \\ 1 & x_{i+2} & y_{i+2} & z_{i+2} \end{vmatrix}}{\begin{vmatrix} 1 & x_i & y_i & z_i \\ 1 & x_{i+1} & y_{i+1} & z_{i+1} \\ 1 & x_{i+2} & y_{i+2} & z_{i+2} \\ 1 & x_{i+3} & y_{i+3} & z_{i+3} \end{vmatrix}} \;\; et \;\; \phi_4 = 1 - \phi_1 - \phi_2 - \phi_3$$

où $p_1 = (x_1, y_1, z_1)$ $p_2 = (x_2, y_2, z_2)$ $p_3 = (x_3, y_3, z_3)$ $p_4 = (x_4, y_4, z_4)$ sont les noeuds du tétraèdre.

Nous rappelons la formule suivante pour les coordonnées barycentriques,

$$\int_T \phi_1^{\alpha_1} \phi_2^{\alpha_2} \phi_3^{\alpha_3} = d!|T| \frac{\alpha_1! \alpha_2! \alpha_3!}{(d + \alpha_1 + \alpha_2 + \alpha_3)!}$$

$$\frac{|T|}{d!} = \begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}$$

est la mesure de l'aire du tétraèdre.

Nous devons déterminer les coefficients sur chaque tétraèdre :

$$\frac{1}{|T|} \frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\nu \int_T |\nabla \boldsymbol{b}_T|^2 d\tau}$$

Nous vérifions que $\int_T \boldsymbol{b}_T d\tau = \frac{32}{105}|T|$ et $\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = \frac{4096S}{8505|T|}$ où S est la somme de la racine de l'aire de chaque face du tètraèdre. Ainsi $\frac{1}{|T|} \frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\nu \int_T |\nabla \boldsymbol{b}_T|^2 d\tau} = \frac{27|T|^2}{\nu 140 S}$.

Calculs

$$\begin{aligned}
\int_T \boldsymbol{b}_T \, d\tau &= \int_T 256 \, \phi_1\phi_2\phi_3\phi_4 \, d\tau = 256 \int_T \phi_1\phi_2\phi_3(1 - \phi_1 - \phi_2 - \phi_3) \, d\tau \\
&= 256 \left(\int_T \phi_1\phi_2\phi_3 \, d\tau - \int_T \phi_1^2\phi_2\phi_3 \, d\tau - \int_T \phi_1\phi_2^2\phi_3 \, d\tau - \int_T \phi_1\phi_2\phi_3^2 \, d\tau \right) \\
&= 256 \left(3!|T|\frac{1}{6!} - 3 \times 3!|T|\frac{2}{7!} \right) = \frac{32}{105}|T|
\end{aligned}$$

$$\nabla \boldsymbol{b}_T = \nabla(256 \, \phi_1\phi_2\phi_3\phi_4) = 256 \sum_{i=1}^{4} \nabla\phi_i \prod_{\substack{j\neq i \\ j=1}}^{4} \phi_j$$

$$\begin{aligned}
|\nabla \boldsymbol{b}_T|^2 &= \left(256 \sum_{i=1}^{4} \nabla\phi_i \prod_{\substack{j\neq i \\ j=1}}^{4} \phi_j \right)^2 \\
&= 256^2 \left(\sum_{i=1}^{4} |\nabla\phi_i|^2 \prod_{j\neq i}^{4} \phi_j^2 + 2 \sum_{\substack{i<j \\ i,j=1}}^{4} \nabla\phi_i.\nabla\phi_j\phi_i\phi_j \prod_{\substack{k\neq i,j \\ k=1}}^{4} \phi_k^2 \right)
\end{aligned}$$

$$\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = 256^2 \left(\sum_{i=1}^{4} |\nabla\phi_i|^2 \int_T \prod_{j\neq i}^{4} \phi_j^2 \, d\tau + 2 \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j \int_T \phi_i\phi_j \prod_{\substack{k\neq i,j \\ k=1}}^{4} \phi_k^2 \, d\tau \right)$$

$$\int_T \phi_i^2\phi_j^2\phi_k^2 \, d\tau = 3!|T|\frac{8}{9!} = \frac{1}{7560}|T|$$

$$\int_T \phi_i\phi_j^2\phi_k^2\phi_l \, d\tau \;\; avec \;\; \phi_i = 1 - \phi_j - \phi_k - \phi_l$$

donne,

$$\int_T \phi_j^2\phi_k^2\phi_l \, d\tau - \int_T \phi_j^3\phi_k^2\phi_l \, d\tau - \int_T \phi_j^2\phi_k^3\phi_l \, d\tau - \int_T \phi_j^2\phi_k^2\phi_l^2 \, d\tau$$
$$= 3!|T|\left(\frac{4}{8!} - \frac{12}{9!} - \frac{12}{9!} - \frac{8}{9!}\right) = \frac{1}{15120}|T|$$

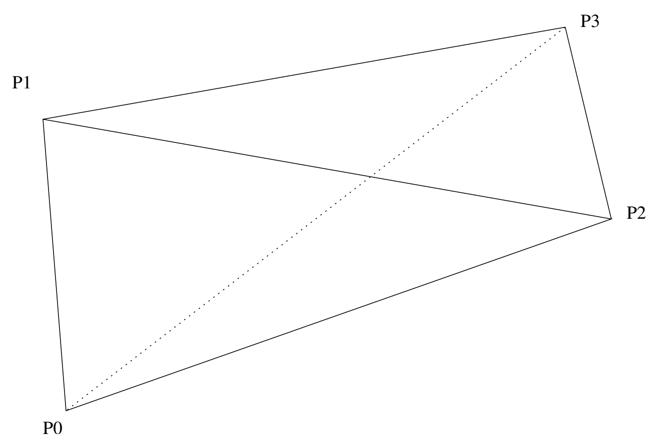
ainsi,

$$\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = 256^2 \left(\frac{1}{7560}|T| \sum_{i=1}^{4} |\nabla\phi_i|^2 + 2\frac{1}{15120}|T| \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j \right)$$
$$= \frac{8192}{945}|T| \left(\sum_{i=1}^{4} |\nabla\phi_i|^2 + \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j \right)$$

Il reste à montrer,

$$\sum_{i=1}^{4} |\nabla\phi_i|^2 + \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j = \frac{S}{18|T|^2}$$

où S est la somme de la racine de l'aire de chaque face du tétraèdre. Remarquons que $p_0 = (x_0, y_0, z_0)$ $p_1 = (x_1, y_1, z_1)$ $p_2 = (x_2, y_2, z_2)$ $p_3 = (x_3, y_3, z_3)$ sont les noeuds du tétraèdre.



Nous obtenons l'aire des carrés comme suit,

$$\begin{cases} \frac{||\overrightarrow{p_0p_1} \wedge \overrightarrow{p_0p_2}||^2}{4} = S_1^2 \\ \frac{||\overrightarrow{p_0p_3} \wedge \overrightarrow{p_0p_2}||^2}{4} = S_2^2 \\ \frac{||\overrightarrow{p_0p_1} \wedge \overrightarrow{p_0p_3}||^2}{4} = S_3^2 \\ \frac{||\overrightarrow{p_1p_3} \wedge \overrightarrow{p_1p_2}||^2}{4} = S_4^2 \end{cases} \tag{4.26}$$

S_1 est l'aire du triangle $\triangle p_0p_1p_2$, S_2 l'aire du triangle $\triangle p_0p_2p_3$, S_3 l'aire du triangle $\triangle p_0p_1p_3$, S_4 l'aire du triangle $\triangle p_1p_2p_3$.

Nous remarquons que $S = S_1^2 + S_2^2 + S_3^2 + S_4^2$.

Nous obtenons,

$$\sum_{i=1}^{4} |\nabla\phi_i|^2 + \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j = \frac{2\ S}{36|T|^2} = \frac{S}{18|T|^2}.$$

5.4.5 Formulation algébrique

Nous souhaitons réécrire (4.19) en terme de tenseurs ou d'opérateur linéaires.

Nous partons du problème :

$$\begin{cases} -\nu\Delta\boldsymbol{u} + \nabla p = f \ \ sur \ \ \Omega \\ div\ \boldsymbol{u} = 0 \ \ sur \ \ \Omega \\ \boldsymbol{u} = 0 \ \ sur \ \ \partial\Omega \end{cases} \tag{4.27}$$

$$B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) = \nu \int_\Omega \nabla\boldsymbol{u}_L : \nabla\boldsymbol{v}_L\ d\tau - \int_\Omega div\ \boldsymbol{v}_L\ p_h\ d\tau - \int_\Omega div\ \boldsymbol{u}_L\ q_h\ d\tau = 0$$

Mais $\boldsymbol{u}_h$, $\boldsymbol{v}_h \in V_h^L$, ainsi, $\boldsymbol{u}_h = \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} \sum_i u_i^1\phi_i \\ \sum_i u_i^2\phi_i \\ \sum_i u_i^3\phi_i \end{pmatrix}$ et $\boldsymbol{v}_h = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} \sum_j v_j^1\phi_j \\ \sum_j v_j^2\phi_j \\ \sum_j v_j^3\phi_j \end{pmatrix}$

Le terme $\int_\Omega \nabla u^s : \nabla v^s\ d\tau$ pour $s = 1..3$ devient

$$\sum_{k=1}^{3}\sum_{i,j} u_i^s v_j^s \int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k} = \sum_{i,j}\sum_{k=1}^{3} u_i^s v_j^s \int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k}$$

En termes d'opérateur cela revient à déterminer pour $s = 1..3$

$$A_s = \left(\sum_{k=1}^{3}\int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k} d\tau\right)_{i,j}$$

Avec $p_h = \sum_j p^j\phi_j$ et $q_h = \sum_i p^i\phi_i$, les termes $\int_\Omega div\ \boldsymbol{u}\ q_h\ d\tau$ pour $s = 1..3$ deviennent

$$\sum_{k=1}^{3}\sum_{i,j} q^j u_i^k \int_\Omega \frac{\partial\phi_i}{\partial x_k}\phi_j d\tau$$

Et cela revient è déterminer pour $k = 1..3$:

$$D_{x_k}(p) = \left(\int_\Omega \frac{\partial \phi_i}{\partial x_k} \phi_j d\tau \right)_{i,j}$$

Le terme $\int_\Omega div\, \boldsymbol{v}\, p_h\, d\tau$ devient :

$$\sum_{k=1}^{3} \sum_{i,j} p^i v_j^k \int_\Omega \frac{\partial \phi_j}{\partial x_k} \phi_i d\tau$$

Cela revient à déterminer :

$$D_{x_k}(p) = \left(\int_\Omega \frac{\partial \phi_j}{\partial x_k} \phi_i d\tau \right)_{i,j}$$

Finalement le terme pour $s = 1, 2, 3$,

$$\sum_{T \in \mathcal{T}_h} \alpha_T \int_T \frac{\partial p_h}{\partial x_s} \frac{\partial q_h}{\partial x_s} d\tau$$

devient

$$\sum_{T \in \mathcal{T}_h} \alpha_T \sum_{i,j} \int_T p^i q^j \frac{\partial \phi_i}{\partial x_s} \frac{\partial \phi_j}{\partial x_s} d\tau$$

qui est équivalent à :

$$Kl_T = \left(\sum_{T \in \mathcal{T}_h} \alpha_T \int_T \frac{\partial \phi_i}{\partial x_s} \frac{\partial \phi_j}{\partial x_s} d\tau \right)_{i,j}$$

Et le second membre avec $\boldsymbol{f} = \begin{pmatrix} f^1 \phi_j \\ f^2 \phi_j \\ f^3 \phi_j \end{pmatrix}$ for $k = 1, 2, 3$,

$$\sum_j \sum_i f^k q^i \frac{\partial \phi_i}{\partial x_k} \int_T \phi_j d\tau$$

est équivalent à,

$$CD_k = \left(\sum_i \alpha_T \frac{\partial \phi_i}{\partial x_k} \int_T \phi_j d\tau \right)_j .f^k$$

Donc le système peut être écrit sous sa forme algébrique,

$$\left(\begin{array}{c|c|c|c} A_1 & 0 & 0 & -D_x(p) \\ \hline 0 & A_2 & 0 & -D_y(p) \\ \hline 0 & 0 & A_3 & -D_z(p) \\ \hline -D_x(p) & -D_y(p) & -D_z(p) & -Kl_T \end{array} \right) \left(\begin{array}{c} u^1 \\ \hline u^2 \\ \hline u^3 \\ \hline p_h \end{array} \right) = \left(\begin{array}{c} b_1 - CD_1 \\ b_2 - CD_2 \\ b_3 - CD_3 \\ 0 \end{array} \right)$$

5.5 Résultats Numériques

Nous commençons par emprunter un cas test présenté dans [9] mais adapté à la situation en trois dimensions. Nous comparons l'erreur donnée dans les normes relatives $e_2(u) = |I_h(u_{ex}) - u_h|_{2,\Omega}$, $e_1(u) = |I_h(u_{ex}) - u_h|_{1,\Omega}$ et $e_{max} = \max_u |I_h(u_{ex}) - u_h|$ pour la vitesse et la pression $(u = p, v)$ par comparaison de la solution exacte à la solution approchée.

Soient $u = (u_1, u_2, u_3)$ et p les solutions exactes données par

$$\begin{aligned}
u_i &= 100\, x_i^2\, (1-x_i)^2\, x_j\, (1-x_j)(1-2x_j) \quad i \neq j \\
u_j &= -100\, x_j^2\, (1-x_j)^2\, x_i\, (1-x_i)(1-2x_i) \\
u_k &= 0 \quad k \neq i,j \quad i,j,k \in \{1,2,3\} \\
p &= x_i^3 + x_j^3 + cste
\end{aligned}$$

Nous résolvons le poblème de Stokes dans un domaine cubique $\Omega \subset \mathbb{R}^3$, avec les conditions tangentielles de Dirichlet suivantes. Nous fixons $\Gamma_W, \Gamma_E, \Gamma_N, \Gamma_S, \Gamma_U, \Gamma_D$ les frontières Ouest-Est-Nord-Sud-Supérieure-Inférieure de ce cube. Pour les deux faces opposées correspondantes,

$$\begin{aligned}
u_i &= 100\, x_i^2 (1-x_i)\, x_j\, (1-x_j)(1-2x_j) \\
u_j &= 100\, x_j (1-x_j)^2\, x_i\, (1-x_i)(1-2x_i) \\
u_k &= 0
\end{aligned}$$

Nous résumons les résultats de convergence dans le tableau qui suit et par la figure 5.1,

N représente le nombre de noeuds par arête du cube s représente la pente des droites d'erreur données par $log(e_i)$, $i = 1,2$ or max en fonction de $log(h) = log(1/N)$ obtenues par régression linéaire c.a.d. une évaluation de l'ordre de convergence numérique .

$\mathbb{P}_1$-bubble $\mathbb{P}_1$	e_{max}		e_2			
N	p	v	p	v_x	v_y	v_z
10	0.3966957	0.3529597	0.0533052	0.0812666	0.0796666	0.0122089
20	0.2048136	0.0836469	0.0151948	0.0189809	0.0189060	0.0030173
30	0.1380296	0.0370121	0.0076317	0.0082081	0.0081998	0.0013435
40	0.1041125	0.0208432	0.0047935	0.0045508	0.0045510	0.0007567
	e_1					
N	p	v_x	v_y	v_z		
10	0.0690151	0.1397889	0.1478699	0.0216202		
20	0.0209554	0.0334138	0.0359207	0.0053798		
30	0.0114437	0.0145506	0.0156653	0.0024133		
40	0.0076799	0.0080925	0.0087163	0.0013635		

$\mathbb{P}_1$-bubble $\mathbb{P}_1$						
	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
s	0.96	2.04	1.74	2.08	2.07	2.01
	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
s	1.59	2.02	2.04	1.99		

Dans le premier tableau nous voyons clairement que la méthode est convergente.

Le deuxième tableau ne fait que confirmer numériquement l'ordre de convergence formulé par (4.13) et (4.14).

La méthode étant validée numériquement sur un cas test nous illustront l'approximation par éléments finis $\mathbb{P}_1$-bulle $\mathbb{P}_1$ sur des exemples connus pour le problème de Stokes.

Les figures 5.2, 5.3, 5.4, 5.5 montrent des champs de vitesse et de pression obtenues pour des cas connus, la cavité entraînée, l'écoulement de Poiseuille. Les légendes sont les suivantes : les grandeurs scalaires sont repésentées par un dégradé de couleur allant du bleu pour les valeurs minimales au rouge pour les valeurs maximales. Pour les vitesses, grandeurs vectorielles bien-sûr cette couleur représente la norme.

Dans les exemples suivants, différentes conditions aux limites et différents domaines pour des écoulements basiques où plus complexes ont été testés et les figures résument les résultats. Un grand soin a été pris pour vérifier que la vitesse est à divergence nulle y compris pour les conditions aux limites, afin de respecter l'équation de continuité. La cavité entraînée a été

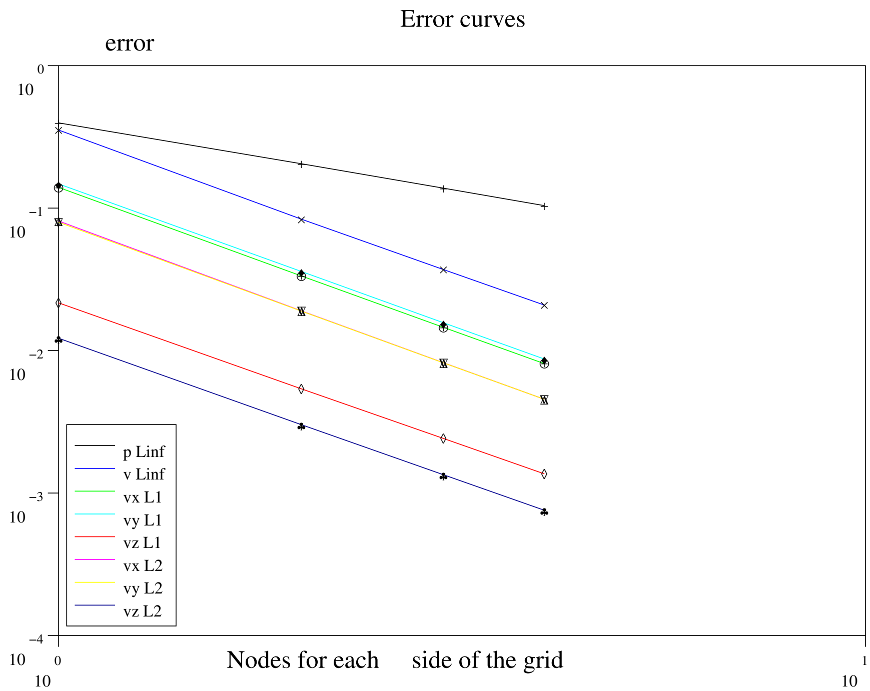


FIG. 5.1 – Courbes d'erreur pour le problème de Stokes-3D

obtenue avec la condition aux limites suivantes imposées à la frontière supérieure de chacun des domaines,

$$\left\{\begin{array}{rcl} u_1(x) & = & 4x(1-x) \quad \text{sur } \Gamma_U \\ u_2 & = & 0 \end{array}\right.$$

et une condition de Dirichlet homogène imposée sur les autres frontières.

Pour l'écoulement de Poiseuille, un profil parabolique à l'injection et en sortie de la forme suivante a été choisie pour le carré,

$$\left\{\begin{array}{rcl} u_2(x) & = & 4x(1-x) \\ u_1 & = & 0 \end{array}\right.$$

et pour le cube,

$$\left\{\begin{array}{rcl} u_1 & = & 16y(1-y)z(1-z) \\ u_2 & = & 0 \\ u_3 & = & 0 \end{array}\right.$$

Pour les écoulement complexes, un profil des vitesses constant à l'injection comme en sortie et une condition de type glissement sur les faces supérieures et inférieures , une condition de Dirichlet homogène sur la surface circulaire ou sphérique ont été ajoutés. La solution numérique n'est pas évidemment comparable à une solution analytique dans cette géométrie générale mais les résultats sont qualitativement satisfaisants.

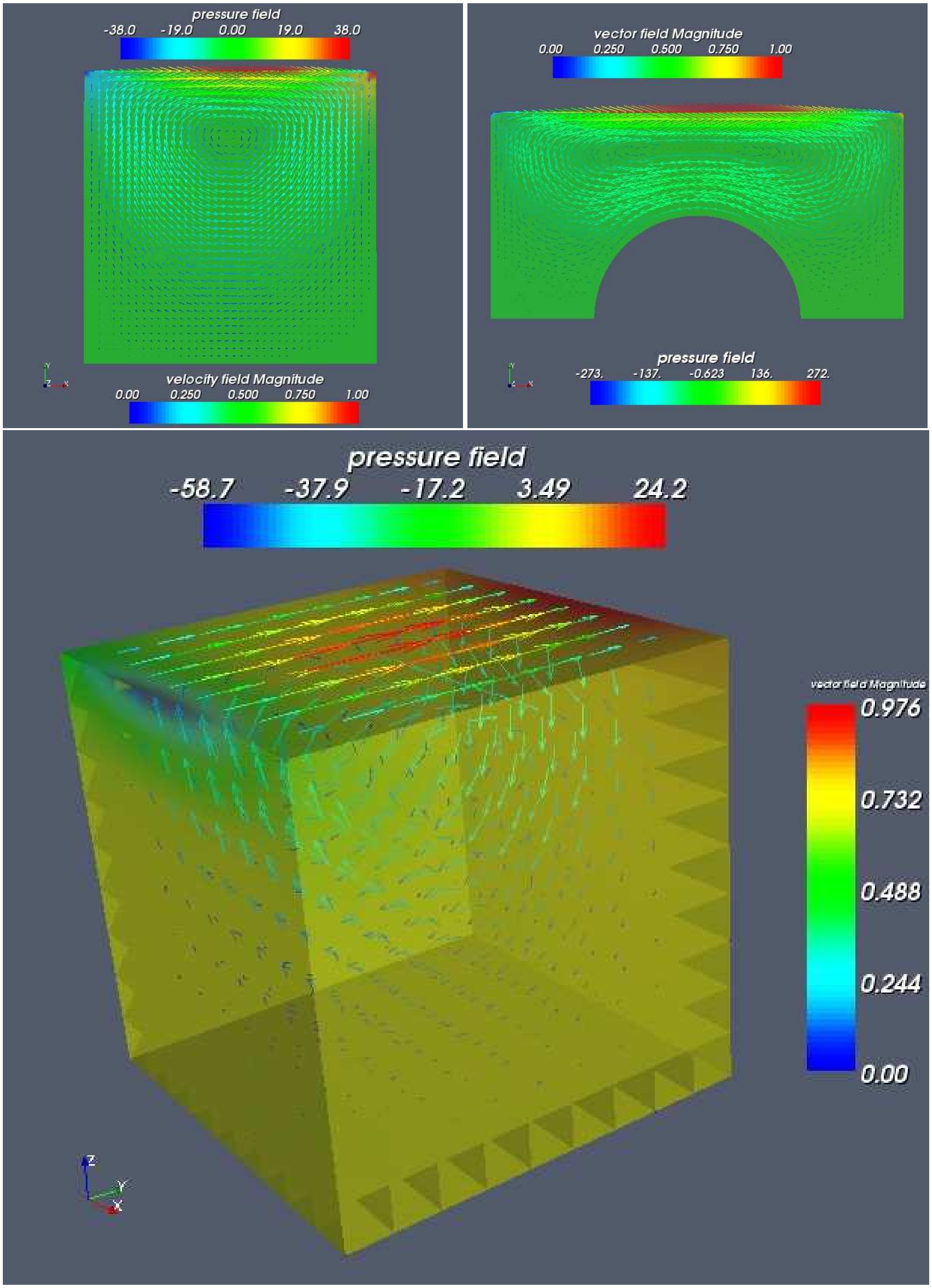


FIG. 5.2 – Cavités entraînées

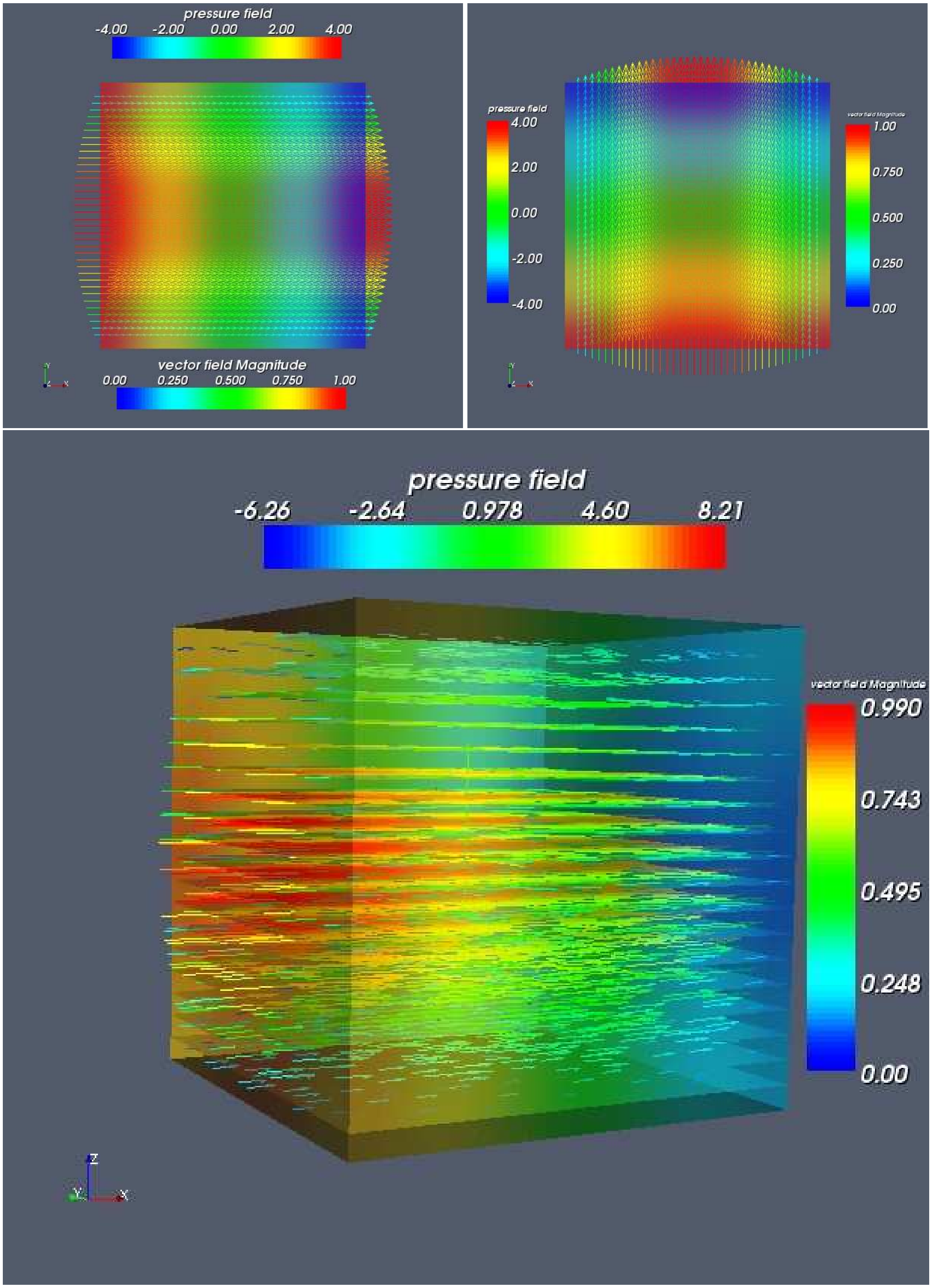


FIG. 5.3 – Ecoulements de Poiseuille

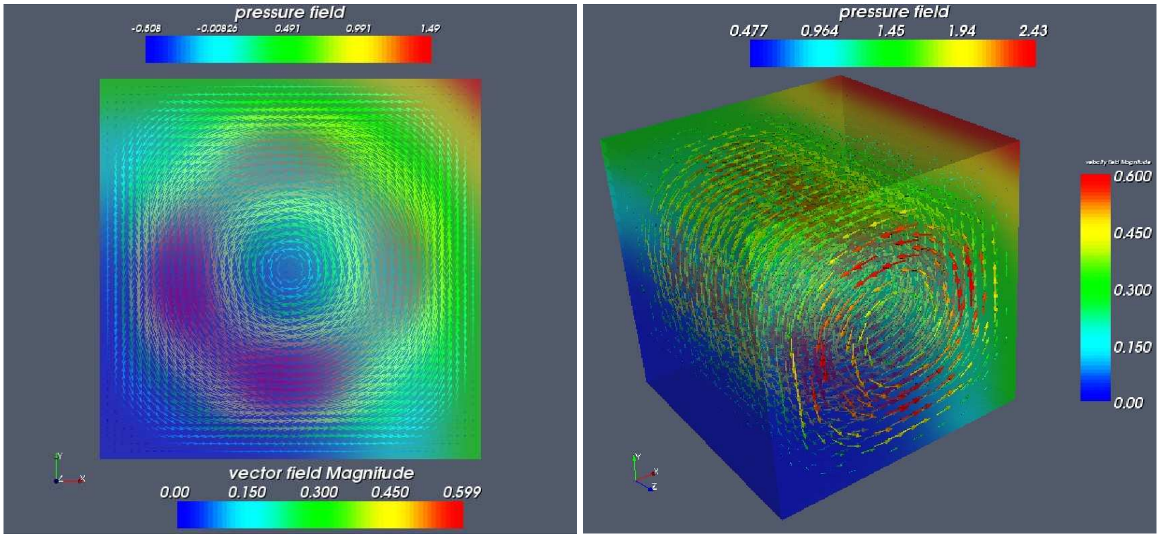


FIG. 5.4 – Problème emprunté à Knobloch [9]

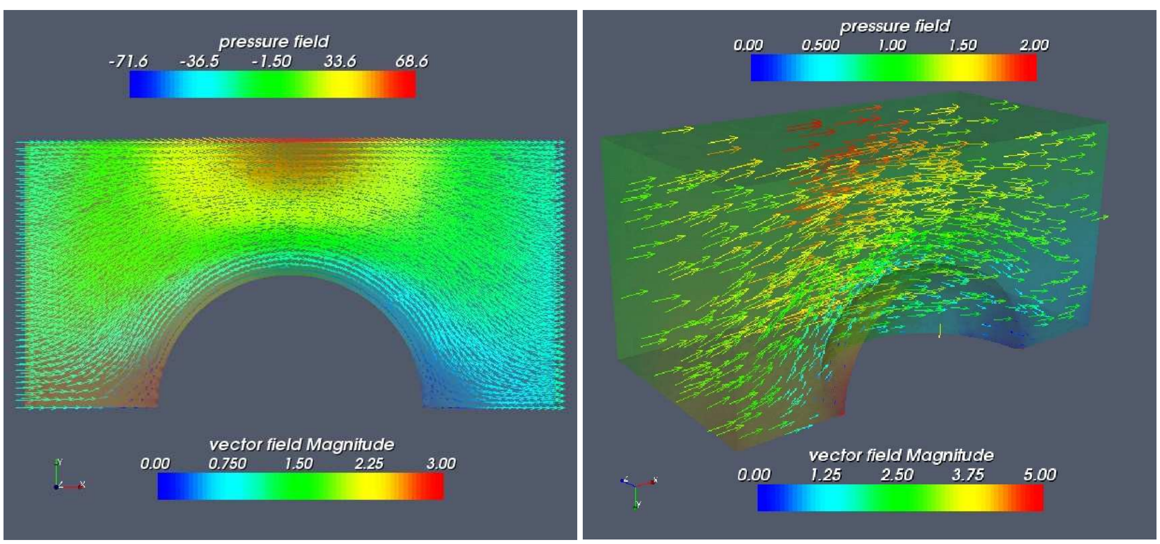


FIG. 5.5 – Problème d'écoulement complexe

5.6 Conclusion

Les équations de Stokes décrivent l'écoulement d'un fluide newtonien incompressible caractérisé par un petit nombre de Reynolds. C'est le cas lorsque le fluide est laminaire. Une description plus complète serait celle où les termes d'inertie ne sont plus négligeables. C'est le cas lorsque le nombre de Reynolds est grand, et décrit par les équations de Navier-Stokes. La résolution des équations de Navier-Stokes revient à considérer de façon itérative plusieurs problèmes de Stokes. Le probème principal est la transition entre un écoulement laminaire et un écoulement turbulent, problème que l'on rencontre plus volontiers dans les applications réalistes en aéronautique ou en météorologie par exemple. Nous avons expliqué en détail l'approximation $\mathbb{P}_1$+bubble-$\mathbb{P}_1$ du problème de Stokes. Celle-ci donne une approximation d'ordre h^2 pour la vitesse et h pour la pression, avec peu de degrés de liberté, donc intéressante pour les calculs en dimension 3, efficace et précise. Un autre avantage est son application à l'équation de convection-diffusion qui a pour effet de stabiliser le schéma numérique par addition des contributions non linéaires dues à la fonction bulle. Même si la méthode et le problème ne sont pas nouveaux, l'étude d'écoulement plus complexe comme nous allons le voir dans la section suivante n'aurait pas été envisageable sans une bonne connaissance de ce cas d'école.

version anglaise

5.7 Introduction

The first mathematical models of fluid motion dated from the XVIIIth century are historically allocated to Euler and Lagrange. The Eulerian description of fluid motion gives a spatial distribution of the velocity whereas the Lagrangian description gives the velocity on each point of the fluid. In the first one, the location is fixed, in the second one the location moves with the particle. In fact both descriptions are coming form the Newton law that connects the force to the acceleration of the particle. In fluid mechanics we consider the Eulerian description. On the most basic level, laminar fluid behavior is described by a set of fundamental equations. Amongst them, we have :

– the fluid continuity,

– the equation of motion,

in addition with two relations that gives the fluid stress and the strain rate. Combining the two first equations and using the two relations to eliminate the fluid stress and the strain rate gives the Navier-Stokes equation (7.28) for an incompressible fluid found in 1821 by C.L. Navier[3] , and G.G. Stokes[4].

$$u.\nabla u - \nu \triangle u + \nabla p = f \tag{7.28}$$

At the microscopic level, and in the case of slow motion one can neglect the inertial terms in the equation, the fluid is then considered viscous and the Navier-Stokes equation reduces to the Stokes equation (7.29).

---

[3]Claude Louis Navier (1785-1836), Fourier's student at the École polytechnique, Navier entered the École des Ponts et Chaussées in 1804. Navier took charge of the applied mechanics courses at the École des Ponts et Chaussées in 1819, becoming professor in 1830. He changed the traditional way of teaching to put much more emphasis on physics and mathematical analysis. He gave the well known equation in 1821 for an incompressible fluid. Till now the study of the Navier-Stokes is far from being achieved due to its complexity.

[4]Georges Gabriel Stokes (1819-1903), student at the Bristol college in London at the age of 16, Stokes entered Cambridge in 1837 . Coached by William Hopkins, he graduated as Senior Wrangler (top First Class Degree) in the mathematical Tripos in 1841, and he was the first Smith's prizeman. One famous article was "On the steady motion of incompressible fluids" published in 1842. After he had deduced the correct equations of motion, Stokes discovered that he was not the first to obtain this equation. In 1849 Stokes was appointed Lucasian Professor of Mathematics at Cambridge (the one from Newton). He awarded the Rumford medal of the Royal Society in 1852.

$$-\nu\triangle u + \nabla p = f \tag{7.29}$$

The aim of our discussion is not to start a revolution in the abundant literature regarding this problem, but to provide an overview of the field with references from the theory to the application. Our interest is more to emphasize a numerical method in order to solve the equations that will be used in the next chapter. This method is one specific, efficient, useful, accurate, Finite Element Method used in two as well as in three dimensions under reasonable assumptions and selected boundary conditions. We want to obtain a good approximation of the velocity in a convex polygon as well as a not too bad approximation of the pressure.

This work is motivated by the study of more complex flows coming later on in this document, advection-diffusion of chemical concentrations at the pore scale.

The structure of our discussion will be the following, starting from the continuous form of the problem, the spaces of the solution H^2 and H^1 will lead to important regularity results analyzed in the first part. The second part will deal with the properties and the Finite Element approximation of those spaces.

5.8 Continuous problem, regularity result and mathematical formulation

The solution of the Stokes problem has square integrable second derivatives. Such a regularity result is of importance in the analysis of numerical methods for solving the Stokes equations. This result is due to R.B. Kellog and J.E. Osborn in [1]. Let introduce the formulation of the problem. The Stokes system of differential equations may be written as :

$$\begin{cases} -\nu\triangle u + \nabla p &= f & \text{in } \Omega \subset \mathbb{R}^d, d = 2, 3 \\ \text{div} u &= 0 & \text{in } \Omega \\ u &= 0 & \text{or boundary conditions on } \partial\Omega \end{cases} \tag{8.30}$$

$f : \Omega \to \mathbb{R}^d$ is a given function while $u : \Omega \to \mathbb{R}^d$ and $p : \Omega \to \mathbb{R}$ are unknowns. This is the case of a slow motion of fluids with very high viscosity (low Reynolds number $Re = \mu\rho L/U$). Here we consider an homogeneous Dirichlet boundary condition but other conditions involving the normal derivative of u or linear combination between u and p on $\partial\Omega$ are also possible. We have the following theorem,

Theorem 5.8.1 *If $f \in L^2(\Omega)$, and let u and p be a generalized solution of (8.30). We assume $\partial\Omega$ is smooth, then $u \in H^2(\Omega)$ and $p \in H^1(\Omega)$ and there exists a constant C only depending on Ω such that,*

$$\|u\|_{H^2(\Omega)} + \|\nabla p\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)} + \|p\|_{L^2(\Omega)}\right). \tag{8.31}$$

Theorem 5.8.2 *If Ω is a convex polygon then, under the same assumptions as in theorem 5.8.1,*

$$\|u\|_{H^2(\Omega)} + \|\nabla p\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)}\right). \tag{8.32}$$

It gives us the continuous space $H^2(\Omega)$ and $H^1(\Omega)$ in which belong the continuous solutions of the Stokes problem. Numerical methods will be built on this basis : finding discrete spaces for both discrete velocity and pressure that approximate correctly the continuous spaces, and numerical analysis will focus on criteria of distance between the discrete solution and the continuous one.

Remark 14 *In most of books of numerical analysis the Stokes problem is expressed slightly differently from (8.30) see [3]. Given $f \in L^2(\Omega)$ or $H^{-1}(\Omega)$ find $u \in H^1(\Omega)$ and $p \in L^2(\Omega)$ such that,*

$$\left\{\begin{array}{rcl} -\nu\triangle u + \nabla p &=& f \quad \mathit{in}\ \Omega \\ \mathit{div}(u) &=& 0 \quad \mathit{in}\ \Omega \end{array}\right. \tag{8.33}$$

by definition of $H^m(\Omega)$, if $f \in L^2(\Omega)$ then $u \in H^2(\Omega)$ and $p \in H^1(\Omega)$, if $f \in H^{-1}(\Omega)$ then $u \in H^1(\Omega)$ and $p \in L^2(\Omega)$.

Now let define rigorously a variational formulation of problem (8.33). Taking inspiration from [5] we assume $f \in L^2(\Omega)$ and $v \in H_0^1(\Omega)$. We apply Green formulae to the first equation of (8.33).

$$\nu(\nabla u, \nabla v) - (p, \mathrm{div} v) = (f, v) \quad \forall v \in \mathcal{D}(\Omega), \tag{8.34}$$

and therefore,

$$\nu(\nabla u, \nabla, v) = (f, v) \quad \forall v \in \{v \in \mathcal{D}(\Omega) | \mathrm{div} v = 0\}.$$

Let define the space $V_{div} = \{v \in V | \mathrm{div} v = 0\}$ a closed subset of V.

The bilinear form $a(w, v) = \nu(\nabla w, \nabla v)$ $w, v \in V$ is coercive over $V_{div \times V_{div}}$ since the application $v \to (f, v)$ is linear and continuous over V_{div}, Lax-Milgram lemma ensures that the problem

$$\text{find } u \in V_{div} : a(u, v) = (f, v) \quad \forall v \in V_{div} \tag{8.35}$$

admits a unique solution.

Theorem 5.8.3 *Let Ω be a bounded domain in $\mathbb{R}^d$ with a Lipschitz continuous boundary, and for each $f \in L^2(\Omega)$, let u be the solution of (8.35). Then there exists a function $p \in L^2(\Omega)$, which is unique up to an additive constant, such that,*

$$a(u, v) - (p, divv) = (f, v) \quad \forall v \in V. \tag{8.36}$$

This is one of the most common used variational formulation. A different approach is to define the space $Q = L^2_0(\Omega)$,

$$L^2_0(\Omega) = \{q \in L^2(\Omega) | \int_\Omega q = 0\},$$

and the bilinear form

$$b(v, q) = -(q, \mathrm{div} v), \qquad v \in V, q \in Q,$$

from the theorem above we deduced that the problem,

$$\left\{ \begin{array}{lcll} \text{find } u \in V, p \in Q : & & & \\ a(u, v) + b(v, p) & = & (f, v) & \forall v \in V \\ b(u, q) & = & 0 & \forall q \in Q \end{array} \right. \tag{8.37}$$

admits a unique solution for whom the pressure is defined up to an additive constant. The solution to (8.37) is also a solution to (8.34). This equation holds in the almost everywhere sense since $f \in L^2(\Omega)$. (8.37) is the weak formulation of the Stokes equation.

5.9 Finite Element approximation of the Stokes problem

For more details we refer the reader to [4]. Conforming methods are the most natural finite element methods essentially because they allow us to build finite dimensional subspaces of the functions' space we want to approximate. Let be $\mathcal{T}_h$ a partition of Ω into triangles and K and $\hat{K}$ two triangles. A conforming approximation of $H^1(\Omega)$ is a space of continuous functions defined by a finite number of degrees of freedom. For triangular element it is usual to use piecewise linear finite element on K. The continuous condition between K and $\hat{K}$ is obtained by using affine transformation. Let be $\mathbb{P}_1(K)$ the polynomials' space of degree ≤ 1 which dimension equals 3 in $\mathbb{R}^2$ and 4 in $\mathbb{R}^3$. In order to define a finite element we must specify three things conf. [2].

- The geometry, based on a reference element $\hat{K}$ and a change of variables $F(\hat{x})$ such as $K = F(\hat{K})$.
- A set $\hat{P}_1$ of polynomials on $\hat{K}$. For $\hat{p} \in \hat{P}_1$ we define $p = \hat{p}^\circ F^{-1}$
- A set of degrees of freedom $\hat{\Sigma}$, that is a set of linear forms $\{\hat{l}_i\}_{1\leq i\leq 3}$ on $\hat{P}_1$.

The set is unisolvent if the linear forms are linearly independent. A finite element is of Lagrange type if its degrees of freedom are point values. Let be a set $\{\hat{a}_i\}_{1\leq i\leq 3}$ of points in $\hat{K}$ and let be $\hat{l}_i(\hat{p}) = \hat{p}(a_i) \quad 1 \leq i \leq 3$. Approximating $H^1(\Omega)$ with Lagrange type element is possible. However, the choice of $\mathbb{P}_1$ finite element for the Stokes problem is not relevant because it does not satisfy the inf-sup condition (ensuring indeed, that the linear operator of the full linear system is invertible). Nevertheless, we can enrich the space $\mathbb{P}_1$ with a bubble function whose main property is to vanish on ∂K. This gives a stable, convergent approximation of the Stokes problem.

5.10 Petrov-Galerkin approximation of the problem with bubble stabilization

5.10.1 inf-sup condition and error estimates

We proceed as in [6]. Let be a triangulation $\mathcal{T}_h$ of $\bar{\Omega}$ and we approximate the velocity on each element K by a polynomial in

$$\mathcal{P}_1(K) = [\mathbb{P}_1 + span\{\lambda_1, \lambda_2, \lambda_3\, \lambda_4]^3, \tag{10.38}$$

and the pressure by a polynomial in $\mathbb{P}_1$.

The finite element spaces will be

$$V_h = \left\{v \in \mathcal{C}^0(\bar{\Omega})^3,\ v|_K \in \mathcal{P}_1(K),\ v|_\Gamma = 0\right\},$$
$$Q_h = \left\{q \in \mathcal{C}^0(\bar{\Omega})^3, q|_K \in \mathbb{P}_1(K),\ \forall K \in \mathcal{T}_h\right\}$$
$$P_h = Q_h \cap L_0^2(\Omega).$$

The degrees of freedom are the values of the velocity at the center and the vertices of K and the values of the pressure at the vertices of K. The discrete problem reads :

$$\text{Find a pair } (u_h, v_h) \in V_h \times P_h \text{ satisfying :}$$
$$\begin{cases} a(u_h, v_h) - (p_h, \mathrm{div} v_h) &= (f, v_h) \quad \forall v_h \in V_h \\ (\mathrm{div} u_h, q_h) &= 0 \qquad \forall q_h \in P_h \end{cases} \tag{10.39}$$

where a is given by, $a(u, v) = \nu(\nabla u, \nabla v)$.

Lemma 5.10.1 *If the triangulation $\mathcal{T}_h$ is regular, the pair of spaces (V_h, P_h) satisfies the inf-sup condition :*

$$\sup_{v_h \in V_h} \frac{(q_h, div v_h)}{\|v_h\|_{H^1(\Omega)}} \geq \beta \|q_h\|_{L_0^2(\Omega)} \qquad \forall q_h \in P_h$$

see [6] for the details of the proof.

We recall also the following theorem from [6],

Theorem 5.10.2 *Let Ω be a bounded polygon and let the solution (u, p) of the Stokes problem satisfy $u \in [H^2(\Omega) \cap H_0^1(\Omega)]^d$, $p \in H^1(\Omega) \cap L_0^2(\Omega)$. If the triangulation $\mathcal{T}_h$ is regular, the solution (u_h, p_h) of problem (10.39) satisfies the error bound :*

$$\|u - u_h\|_{H^1(\Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq Ch \left\{\|u\|_{H^2(\Omega)} + \|p\|_{H^1(\Omega)}\right\}. \tag{10.40}$$

When Ω is convex, we have the L^2-estimate :

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \left\{\|u\|_{H^2(\Omega) + \|p\|_{H^1(\Omega)}}\right\}. \tag{10.41}$$

5.10.2 Shape functions

This part will be treated in details, starting from the variational discrete form of the Stokes problem, but decoupled between the linear term coming from the $\mathbb{P}_1$ discretization and the quadratic term coming from the bubble-function, ending with the algebraïc form of the Stokes problem. We proceed as in [18] and give the exact value of the shape functions for each term of the equation in the case of a domain $\Omega \cup \mathbb{R}^3$. We consider the variational form of the Stokes problem (10.39) where $\Omega \subset \mathbb{R}^3$ is a polygonal domain. If $(u_h, p_h), (v_h, q_h) \in V_h \times P_h$, we may define :

$$\begin{aligned} B(u_h, p_h; v_h, q_h) &= \nu \int_\Omega \nabla u_h : \nabla v_h dx - \int_\Omega \mathrm{div} v_h \, p_h \, dx \\ &- \int_\Omega \mathrm{div} u_h \, q_h \, dx \end{aligned} \tag{10.42}$$

$$\text{and } F(v_h, q_h) = \int_\Omega f v_h \, dx \tag{10.43}$$

Problem (10.39) may be written as

$$\text{Find } u_h \in V_h,\ p_h \in P_h \text{ such as,}$$
$$B(u_h, p_h; v_h, q_h) = F(v_h, q_h) \quad \forall v_h \in V_h \, \forall q_h \in P_h. \tag{10.44}$$

We consider now the following spaces :

$$V_h^L = \left\{ v_L = (v_L^1, v_L^2, v_L^3) \in [\mathcal{C}^0(\bar{\Omega})]^3 : v_L^i|_T \in \mathbb{P}_1, i = 1, 2, 3, T \in \mathcal{T}_h \right\}$$

and $\boldsymbol{b}_T = (b_T^1, b_T^2, b_T^3) \in [H_0^1(T)]^3$ b_T^i taking its value on T, and we define $\boldsymbol{b}_T^1 = (b_T^1, 0, 0)$, $\boldsymbol{b}_T^2 = (0, b_T^2, 0)$, $\boldsymbol{b}_T^3 = (0, 0, b_T^3)$ and

$$\boldsymbol{B} = vect\{b_T^1, b_T^2, b_T^3, T \in \mathcal{T}_h\}$$

The approximation space for the velocity is defined by $V_h = V_h^L \oplus \boldsymbol{B}$. And let define

$$P_h = \left\{ q_h \in C^0(\overline{\Omega}) : \int_\Omega q_h = 0, q_{h|T} \in \mathbb{P}_1, T \in \mathcal{T}_h \right\}$$

We saw in the previous chapter that this couple of spaces verify the inf-sup condition and lead to a stable convergent approximation of the problem. We express v_h in the form of a linear part v_L and a bubble part :

$$\boldsymbol{v}_h = \boldsymbol{v}_L + \sum_{T\in\mathcal{T}_h} (c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) \tag{10.45}$$

where $\boldsymbol{v}_L \in V_h^L$ et $c_T^1, c_T^2, c_T^3 \in \mathbb{R}\ \ \forall T \in \mathcal{T}_h$. Problem (10.44) takes the form :

$$\begin{cases} Find\ \boldsymbol{u}_h = \boldsymbol{u}_L + \sum_{T\in\mathcal{T}_h} (c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) \in V_h, p_h \in P_h\ such\ that \\ B(\boldsymbol{u}_h, p_h; \boldsymbol{v}_L, q_h) = F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h \\ B(\boldsymbol{u}_h, p_h; \boldsymbol{b}_T^i, q_h) = F(\boldsymbol{b}_T^i, q_h)\ \forall T \in \mathcal{T}_h, q_h \in P_h, i = 1 \ldots 3 \end{cases} \tag{10.46}$$

We now eliminate the bubble term from the second member of the equation. for $i = 1$, $q_h = 0$ we get :

$$\nu \int_T \nabla(\boldsymbol{u}_L + c_T^1 \boldsymbol{b}_T^1 + c_T^2 \boldsymbol{b}_T^2 + c_T^3 \boldsymbol{b}_T^3) : \nabla \boldsymbol{b}_T^1 d\tau - \int_T div\ \boldsymbol{b}_T^1 p_h d\tau = \int_T \boldsymbol{f}.\boldsymbol{b}_T^1 d\tau$$

u_L is piecewise linear, f is constant implying,

$$\nu \int_T \nabla \boldsymbol{u}_L : \nabla \boldsymbol{b}_T^1 d\tau = 0,\ \ \nu \int_T \nabla \boldsymbol{b}_T^1 : \nabla \boldsymbol{b}_T^1 d\tau = \nu \int_T |\nabla b_T^1|^2 d\tau,$$

$$\int_T \text{div}\ \boldsymbol{b}_T^1 p_h d\tau = \int_T \frac{\partial b_T^1}{\partial x} p_h d\tau = -\int_T b_T^1 \frac{\partial p_h}{\partial x} d\tau = -\left(\frac{\partial p_h}{\partial x}\right)_{|T} \int_T b_T^1 d\tau$$

The equation for $\boldsymbol{b}_T^1$ becomes :

$$\nu c_T^1 \int_T |\nabla b_T^1|^2 d\tau + \left[\frac{\partial p_h}{\partial x}\right]_{|T} \int_T b_T^1 d\tau = f_{|T}^1 \int_T b_T^1 d\tau$$

c_T^1 may be expressed by :

$$c_T^1 = -\frac{\int_T b_T^1 d\tau}{\int_T |\nabla b_T^1|^2 d\tau} \left[\frac{\partial p_h}{\partial x} - f^1\right]_{|T}$$

In the same manner we obtain :

$$c_T^2 = -\frac{\int_T b_T^2 d\tau}{\int_T |\nabla b_T^2|^2 d\tau} \left[\frac{\partial p_h}{\partial y} - f^2\right]_{|T}$$

and

$$c_T^3 = -\frac{\int_T b_T^3 d\tau}{\int_T |\nabla b_T^3|^2 d\tau} \left[\frac{\partial p_h}{\partial z} - f^3\right]_{|T}$$

Now, the first equation of (10.46) gives,

$$\nu \int_\Omega \nabla(\boldsymbol{u}_L + \boldsymbol{u}_B) : \nabla \boldsymbol{v}_L d\tau - \int_\Omega div(\boldsymbol{v}_L) p_h d\tau$$
$$- \int_\Omega div\ (\boldsymbol{u}_L + \boldsymbol{u}_B) q_h d\tau = \int_\Omega \boldsymbol{f}.\boldsymbol{v}_L d\tau \ \ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h$$

We notice that

$$\nu \int_\Omega \nabla \boldsymbol{u}_B : \nabla \boldsymbol{v}_L d\tau = 0,$$

$$\begin{aligned}
\int_\Omega div(\boldsymbol{u}_B) q_h d\tau &= \sum_{T\in\mathcal{T}_h} \left[\sum_{i=1}^3 c_T^i \int_T div\ \boldsymbol{b}_T^i q_h d\tau \right] \\
&= \sum_{T\in\mathcal{T}_h} \left[\sum_{i=1}^3 c_T^i \int_T \frac{\partial b_T^i}{\partial x_i} q_h d\tau \right] \\
&= - \sum_{T\in\mathcal{T}_h} \left[\sum_{i=1}^3 c_T^i \int_T b_T^i \frac{\partial q_h}{\partial x_i} d\tau \right] \\
&= - \sum_{T\in\mathcal{T}_h} \left[\sum_{i=1}^3 c_T^i \int_T \left(\frac{\partial q_h}{\partial x_i} \right)_{|T} b_T^1 d\tau \right] \\
&= \sum_{T\in\mathcal{T}_h} \left[\sum_{i=1}^3 \frac{\left(\int_T b_T^i d\tau\right)^2}{\int_T |\nabla b_T^i|^2 d\tau} \left[\frac{\partial p_h}{\partial x_i} - f^i \right]_{|T} \left[\frac{\partial q_h}{\partial x_i} \right]_{|T} \right]
\end{aligned}$$

To obtain finally,

$$\begin{cases}
Find\ \boldsymbol{u}_L \in V_h^L, p_h \in P_h\ such\ that \\
B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) + \sum_{T\in\mathcal{T}_h} \frac{1}{|T|} \left[\sum_{i=1}^3 \frac{\left(\int_T b_T^i d\tau\right)^2}{\int_T |\nabla b_T^i|^2 d\tau} \int_T \left[\frac{\partial p_h}{\partial x_i} - f^i \right] \left[\frac{\partial q_h}{\partial x_i} \right] d\tau \right] \\
= F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h
\end{cases} \quad (10.47)$$

which could be expressed as

$$\begin{cases}
Find\ \boldsymbol{u}_L \in V_h^L, p_h \in P_h\ such\ that \\
B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) \\
+ \sum_{T\in\mathcal{T}_h} \frac{1}{|T|} \frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\int_T |\nabla \boldsymbol{b}_T|^2 d\tau} \int_T \left[grad(p_h) - \boldsymbol{f} \right] \left[-grad(q_h) \right] d\tau \\
= F(\boldsymbol{v}_L, q_h)\ \forall \boldsymbol{v}_L \in V_h^L, q_h \in P_h
\end{cases} \quad (10.48)$$

5.10.3 Residual free Bubble applied to Advective-Diffusive equation

We make here a short deflection to emphasize the stability impact of the bubble stabilization for an advective-diffusive equation. In some cases the RFB (Residual Free Bubble method) is similar to SUPG method. Computations indicate improvement in the accuracy of the results. From [8] if we use bilinear interpolation as the reduced space in a Galerkin method and enrich with residual-free bubbles, we obtain a stabilization which is not equal to the streamline diffusive one. Only if the reduced space is constituted with linear functions we recover the SUPG with optimal parameters. We briefly describe the method.

We take the following linear differential operator,

$$L = -D\triangle + \nabla.(q.)$$

$D > 0$ the diffusivity, $\vec{q}$ the velocity, u the concentration

$u = 0$ on $\partial\Omega$.

The Galerkin method consists of finding $u_h \in V_h$ such as,

$$a(u_h, v_h) = (Lu_h, v_h) = (f, v_h) \quad \forall v_h \in V_h \tag{10.49}$$

With V_h as defined previously. Each member u_h and v_h of the space of functions V_h is spanned by continuous piecewise linear or bilinear polynomials plus bubble functions to be defined below, i.e.,

$$u_h = u_1 + u_b \tag{10.50}$$

The bubble part of this space is subject to zero Dirichlet boundary condition on each element K, i.e.,

$$u_b = 0 \quad \text{on } \partial K \tag{10.51}$$

Testing with v_1 as a test function gives,

$$D(\nabla u_1, \nabla v_1) + D(\nabla u_b, \nabla v_1) + (\nabla.(qu_1), v_1) + \sum_K (\nabla.(qu_b), v_1)_K = (f, v_1)$$

Integrating by parts,

$$D(\nabla u_1, \nabla v_1) + (\nabla.(qu_1), v_1) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_1)_K = (f, v_1)$$

The residual free part of the solution, u_b can be computed as a solution to

$$-D\triangle u_b + \nabla.(qu_b) = -\nabla.(qu_1) + D\triangle u_1 + f$$

Testing now with v_b gives,

$$D(\nabla u_1, \nabla v_b) + D(\nabla u_b, \nabla v_b) + (\nabla.(qu_1), v_b) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_b)_K = (f, v_b)$$

Integrating by parts

$$D(\nabla u_1, \nabla v_b) + (\nabla.(qu_1), v_b) + \sum_K (\nabla.(qu_b) - D\triangle u_b, v_b)_K = (f, v_b)$$

And the stabilized variational form of the advective-diffusive equation reads,

$$\begin{aligned}
& D(\nabla u_1, \nabla v_1) + D(\nabla u_1, \nabla v_b) + D(\nabla u_b, \nabla v_1) + D(\nabla u_b, \nabla v_b) \\
+ \quad & (\nabla.(qu_1), v_1) + (\nabla.(qu_1), v_b) + \sum_K ((\nabla.(qu_b), v_1) + (\nabla.(qu_b), v_b))_K \\
= \quad & (f, v_1) + (f, v_b)
\end{aligned} \tag{10.52}$$

5.10.4 The bubble function

The bubble function for the tetrahedra is defined as,

$$\boldsymbol{b}_T = 4^4 \phi_1\phi_2\phi_3\phi_4 \;\; with$$

$$\phi_i = \frac{\begin{vmatrix} 1 & x & y & z \\ 1 & x_i & y_i & z_i \\ 1 & x_{i+1} & y_{i+1} & z_{i+1} \\ 1 & x_{i+2} & y_{i+2} & z_{i+2} \end{vmatrix}}{\begin{vmatrix} 1 & x_i & y_i & z_i \\ 1 & x_{i+1} & y_{i+1} & z_{i+1} \\ 1 & x_{i+2} & y_{i+2} & z_{i+2} \\ 1 & x_{i+3} & y_{i+3} & z_{i+3} \end{vmatrix}} \;\; and \;\; \phi_4 = 1 - \phi_1 - \phi_2 - \phi_3$$

where $p_1 = (x_1, y_1, z_1)$ $p_2 = (x_2, y_2, z_2)$ $p_3 = (x_3, y_3, z_3)$ $p_4 = (x_4, y_4, z_4)$ are the vertices of the tetrahedra.

We recall the following formulae for barycentric coordinates,

$$\int_T \phi_1^{\alpha_1}\phi_2^{\alpha_2}\phi_3^{\alpha_3} = d!|T|\frac{\alpha_1!\alpha_2!\alpha_3!}{(d+\alpha_1+\alpha_2+\alpha_3)!}$$

$$\frac{|T|}{d!} = \begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix}$$

is the measure of the tetrahedra.

We have to determinate the coefficient for each tetrahedra :

$$\frac{1}{|T|}\frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\nu\int_T |\nabla \boldsymbol{b}_T|^2 d\tau}$$

We verify that $\int_T \boldsymbol{b}_T d\tau = \frac{32}{105}|T|$ and $\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = \frac{4096S}{8505|T|}$ where S is the sum of tetrahedra's faces to the power two . So $\frac{1}{|T|}\frac{\left(\int_T \boldsymbol{b}_T d\tau\right)^2}{\nu\int_T |\nabla \boldsymbol{b}_T|^2 d\tau} = \frac{27|T|^2}{\nu 140 S}$.

Calculus

$$\begin{aligned}\int_T \boldsymbol{b}_T\, d\tau &= \int_T 256\ \phi_1\phi_2\phi_3\phi_4\ d\tau = 256\int_T \phi_1\phi_2\phi_3(1-\phi_1-\phi_2-\phi_3)\ d\tau \\ &= 256\left(\int_T \phi_1\phi_2\phi_3\ d\tau - \int_T \phi_1^2\phi_2\phi_3\ d\tau - \int_T \phi_1\phi_2^2\phi_3\ d\tau - \int_T \phi_1\phi_2\phi_3^2\ d\tau\right) \\ &= 256\left(3!|T|\frac{1}{6!} - 3\times 3!|T|\frac{2}{7!}\right) = \frac{32}{105}|T|\end{aligned}$$

$$\nabla \boldsymbol{b}_T = \nabla(256\ \phi_1\phi_2\phi_3\phi_4) = 256\sum_{i=1}^{4}\nabla\phi_i \prod_{\substack{j\neq i\\ j=1}}^{4}\phi_j$$

$$\begin{aligned}|\nabla \boldsymbol{b}_T|^2 &= \left(256\sum_{i=1}^{4}\nabla\phi_i\prod_{\substack{j\neq i\\ j=1}}^{4}\phi_j\right)^2 \\ &= 256^2\left(\sum_{i=1}^{4}|\nabla\phi_i|^2\prod_{j\neq i}^{4}\phi_j^2 + 2\sum_{\substack{i<j\\ i,j=1}}^{4}\nabla\phi_i.\nabla\phi_j\phi_i\phi_j\prod_{\substack{k\neq i,j\\ k=1}}^{4}\phi_k^2\right)\end{aligned}$$

$$\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = 256^2\left(\sum_{i=1}^{4}|\nabla\phi_i|^2\int_T\prod_{j\neq i}^{4}\phi_j^2\ d\tau + 2\sum_{\substack{i<j\\ i,j=1}}^{j=4}\nabla\phi_i.\nabla\phi_j\int_T\phi_i\phi_j\prod_{\substack{k\neq i,j\\ k=1}}^{4}\phi_k^2\ d\tau\right)$$

$$\int_T \phi_i^2\phi_j^2\phi_k^2 \, d\tau = 3!|T|\frac{8}{9!} = \frac{1}{7560}|T|$$

$$\int_T \phi_i\phi_j^2\phi_k^2\phi_l \, d\tau \ \ avec \ \ \phi_i = 1-\phi_j-\phi_k-\phi_l$$

gives

$$\int_T \phi_j^2\phi_k^2\phi_l \, d\tau - \int_T \phi_j^3\phi_k^2\phi_l \, d\tau - \int_T \phi_j^2\phi_k^3\phi_l \, d\tau - \int_T \phi_j^2\phi_k^2\phi_l^2 \, d\tau$$
$$= 3!|T|\left(\frac{4}{8!}-\frac{12}{9!}-\frac{12}{9!}-\frac{8}{9!}\right) = \frac{1}{15120}|T|$$

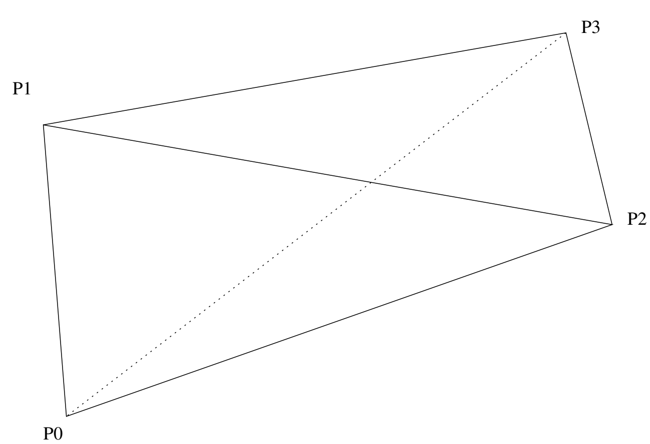
so,

$$\int_T |\nabla \boldsymbol{b}_T|^2 d\tau = 256^2\left(\frac{1}{7560}|T|\sum_{i=1}^{4}|\nabla\phi_i|^2 + 2\frac{1}{15120}|T|\sum_{\substack{i<j\\ i,j=1}}^{j=4}\nabla\phi_i.\nabla\phi_j\right)$$
$$= \frac{8192}{945}|T|\left(\sum_{i=1}^{4}|\nabla\phi_i|^2 + \sum_{\substack{i<j\\ i,j=1}}^{j=4}\nabla\phi_i.\nabla\phi_j\right)$$

It remains to show,

$$\sum_{i=1}^{4}|\nabla\phi_i|^2 + \sum_{\substack{i<j\\ i,j=1}}^{j=4}\nabla\phi_i.\nabla\phi_j = \frac{S}{18|T|^2}$$

where S is the sum of tetrahedra's faces to the power two. Notice that $p_0 = (x_0, y_0, z_0)$ $p_1 = (x_1, y_1, z_1)$ $p_2 = (x_2, y_2, z_2)$ $p_3 = (x_3, y_3, z_3)$ are the vertices of the tetrahedra.



We get the area of the square as follow,

$$\begin{cases} \frac{||\overrightarrow{p_0p_1}\wedge\overrightarrow{p_0p_2}||^2}{4} = S_1^2 \\ \frac{||\overrightarrow{p_0p_3}\wedge\overrightarrow{p_0p_2}||^2}{4} = S_2^2 \\ \frac{||\overrightarrow{p_0p_1}\wedge\overrightarrow{p_0p_3}||^2}{4} = S_3^2 \\ \frac{||\overrightarrow{p_1p_3}\wedge\overrightarrow{p_1p_2}||^2}{4} = S_4^2 \end{cases} \tag{10.53}$$

S_1 is the area of triangle $\triangle p_0p_1p_2$, S_2 the area of triangle $\triangle p_0p_2p_3$, S_3 the area of triangle $\triangle p_0p_1p_3$, S_4 the area of triangle $\triangle p_1p_2p_3$.

We note $S = S_1^2 + S_2^2 + S_3^2 + S_4^2$.

We obtain,

$$\sum_{i=1}^{4} |\nabla\phi_i|^2 + \sum_{\substack{i<j \\ i,j=1}}^{j=4} \nabla\phi_i.\nabla\phi_j = \frac{2\ S}{36|T|^2} = \frac{S}{18|T|^2}.$$

5.10.5 Algebraïc form

We want to rewrite (10.46) in terms of tensor or linear operator.

We start with the problem :

$$\begin{cases} -\nu\Delta\boldsymbol{u} + \nabla p = f \ \ on \ \ \Omega \\ div\ \boldsymbol{u} = 0 \ \ on \ \ \Omega \\ \boldsymbol{u} = 0 \ \ on \ \ \partial\Omega \end{cases} \tag{10.54}$$

$$B(\boldsymbol{u}_L, p_h; \boldsymbol{v}_L, q_h) = \nu \int_\Omega \nabla\boldsymbol{u}_L : \nabla\boldsymbol{v}_L\ d\tau - \int_\Omega div\ \boldsymbol{v}_L\ p_h\ d\tau - \int_\Omega div\ \boldsymbol{u}_L\ q_h\ d\tau = 0$$

But $\boldsymbol{u}_h$, $\boldsymbol{v}_h \in V_h^L$, so, $\boldsymbol{u}_h = \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} \sum_i u_i^1\phi_i \\ \sum_i u_i^2\phi_i \\ \sum_i u_i^3\phi_i \end{pmatrix}$ and $\boldsymbol{v}_h = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} \sum_j v_j^1\phi_j \\ \sum_j v_j^2\phi_j \\ \sum_j v_j^3\phi_j \end{pmatrix}$

The term $\int_\Omega \nabla u^s : \nabla v^s\ d\tau$ for $s = 1..3$ becomes

$$\sum_{k=1}^{3}\sum_{i,j} u_i^s v_j^s \int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k} = \sum_{i,j}\sum_{k=1}^{3} u_i^s v_j^s \int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k}$$

In terms of operator it is equivalent to determinate for $s = 1..3$

$$A_s = \left(\sum_{k=1}^{3}\int_\Omega \frac{\partial\phi_i}{\partial x_k}\frac{\partial\phi_j}{\partial x_k} d\tau\right)_{i,j}$$

With $p_h = \sum_j p^j \phi_j$ and $q_h = \sum_i p^i \phi_i$, the terms $\int_\Omega div\ \boldsymbol{u}\ q_h\ d\tau$ for $s = 1..3$ become

$$\sum_{k=1}^{3} \sum_{i,j} q^j u_i^k \int_\Omega \frac{\partial \phi_i}{\partial x_k} \phi_j d\tau$$

And it is equivalent to determinate for $k = 1..3$:

$$D_{x_k}(p) = \left(\int_\Omega \frac{\partial \phi_i}{\partial x_k} \phi_j d\tau \right)_{i,j}$$

The term $\int_\Omega div\ \boldsymbol{v}\ p_h\ d\tau$ becomes :

$$\sum_{k=1}^{3} \sum_{i,j} p^i v_j^k \int_\Omega \frac{\partial \phi_j}{\partial x_k} \phi_i d\tau$$

This is equivalent to determinate :

$$D_{x_k}(p) = \left(\int_\Omega \frac{\partial \phi_j}{\partial x_k} \phi_i d\tau \right)_{i,j}$$

Finally, the term for $s = 1, 2, 3$,

$$\sum_{T \in \mathcal{T}_h} \alpha_T \int_T \frac{\partial p_h}{\partial x_s} \frac{\partial q_h}{\partial x_s} d\tau$$

becomes

$$\sum_{T \in \mathcal{T}_h} \alpha_T \sum_{i,j} \int_T p^i q^j \frac{\partial \phi_i}{\partial x_s} \frac{\partial \phi_j}{\partial x_s} d\tau$$

This is equivalent to determinate :

$$Kl_T = \left(\sum_{T \in \mathcal{T}_h} \alpha_T \int_T \frac{\partial \phi_i}{\partial x_s} \frac{\partial \phi_j}{\partial x_s} d\tau \right)_{i,j}$$

And the second member with $\boldsymbol{f} = \begin{pmatrix} f^1 \phi_j \\ f^2 \phi_j \\ f^3 \phi_j \end{pmatrix}$ for $k = 1, 2, 3$,

$$\sum_j \sum_i f^k q^i \frac{\partial \phi_i}{\partial x_k} \int_T \phi_j d\tau$$

This is equivalent to determinate,

$$CD_k = \left(\sum_i \alpha_T \frac{\partial \phi_i}{\partial x_k} \int_T \phi_j d\tau \right)_j .f^k$$

So the system may be written in its algebraïc form,

$$\left(\begin{array}{c|c|c|c} A_1 & 0 & 0 & -D_x(p) \\ \hline 0 & A_2 & 0 & -D_y(p) \\ \hline 0 & 0 & A_3 & -D_z(p) \\ \hline -D_x(p) & -D_y(p) & -D_z(p) & -Kl_T \end{array}\right) \left(\begin{array}{c} u^1 \\ \hline u^2 \\ \hline u^3 \\ \hline p_h \end{array}\right) = \left(\begin{array}{c} b_1 - CD_1 \\ b_2 - CD_2 \\ b_3 - CD_3 \\ 0 \end{array}\right)$$

5.11 Results

After giving results in order to illustrate qualitative known test cases for the Stokes problem approximated by $\mathbb{P}_1$-bubble $\mathbb{P}_1$ finite element we borrow a quantitative test case proposed in [9] but fitted to the three dimensions' situation. We compare error given in the suitable norms $e_2(u) = |I_h(u_{ex}) - u_h|_{2,\Omega}$, $e_1(u) = |I_h(u_{ex}) - u_h|_{1,\Omega}$ and $e_{max} = \max_u |I_h(u_{ex}) - u_h|$ for the velocity and the pressure by comparisons to an exact solution and the computed one. Let $u = (u_1, u_2, u_3)$ and p the exact solution given by

$$\begin{aligned} u_i &= 100\, x_i^2\, (1-x_i)^2\, x_j\, (1-x_j)(1-2x_j) \quad i \neq j \\ u_j &= -100\, x_j^2\, (1-x_j)^2\, x_i\, (1-x_i)(1-2x_i) \\ u_k &= 0 \quad k \neq i, j \quad i,j,k \in \{1,2,3\} \\ p &= x_i^3 + x_j^3 + cste \end{aligned}$$

We solve the Stokes problem in a cubic domain $\Omega \subset \mathbb{R}^3$, with the following tangential Dirichlet boundary conditions. We denote by $\Gamma_W, \Gamma_E, \Gamma_N, \Gamma_S, \Gamma_U, \Gamma_D$ the West-East-North-South-Upper-Downer Face of this cube. For the corresponding two opposite faces,

$$\begin{aligned} u_i &= 100\, x_i^2 (1-x_i)\, x_j\, (1-x_j)(1-2x_j) \\ u_j &= 100\, x_j (1-x_j)^2\, x_i\, (1-x_i)(1-2x_i) \\ u_k &= 0 \end{aligned}$$

We summarize results in a table and in figure 5.1 for convergence.

N denotes the number of nodes per edge of the cube. s denotes the slop of the error curves given by $log(e_i)$, $i = 1, 2$ or max as function of $log(h) = log(1/N)$ obtained by linear regression, i.e. an estimate of the numerical order of convergence.

$\mathbb{P}_1$-bubble $\mathbb{P}_1$	e_{max}		e_2			
N	p	v	p	v_x	v_y	v_z
10	0.3966957	0.3529597	0.0533052	0.0812666	0.0796666	0.0122089
20	0.2048136	0.0836469	0.0151948	0.0189809	0.0189060	0.0030173
30	0.1380296	0.0370121	0.0076317	0.0082081	0.0081998	0.0013435
40	0.1041125	0.0208432	0.0047935	0.0045508	0.0045510	0.0007567
	e_1					
N	p	v_x	v_y	v_z		
10	0.0690151	0.1397889	0.1478699	0.0216202		
20	0.0209554	0.0334138	0.0359207	0.0053798		
30	0.0114437	0.0145506	0.0156653	0.0024133		
40	0.0076799	0.0080925	0.0087163	0.0013635		

$\mathbb{P}_1$-bubble $\mathbb{P}_1$						
	$e_{max}(p)$	$e_{max}(v)$	$e_2(p)$	$e_2(v_x)$	$e_2(v_y)$	$e_2(v_z)$
s	0.96	2.04	1.74	2.08	2.07	2.01
	$e_1(p)$	$e_1(v_x)$	$e_1(v_y)$	$e_1(v_z)$		
s	1.59	2.02	2.04	1.99		

Figure 5.2, 5.3, 5.4, 5.5 illustrate the velocity and pressure fields. In the following samples, different boundary conditions for basic to complex Stokes flows in different domains have been tested and pictures summarize the results. Great care was paid to verify that the velocity is equal to zero on averaged on Γ, to respect the continuity equation. The driven cavity was obtained with the following boundary condition on the upper bound for each domain,

$$\left\{ \begin{array}{rcl} u_1(x) & = & 4x(1-x) \quad \text{on } \Gamma_U \\ u_2 & = & 0 \end{array} \right.$$

and an homogeneous Dirichlet boundary condition for the other bounds.

For the Poiseuille flow, a parabolic profile at the inlet and at the outlet of the square,

$$\left\{ \begin{array}{rcl} u_2(x) & = & 4x(1-x) \quad \text{on } \Gamma_U \\ u_1 & = & 0 \end{array} \right. ,$$

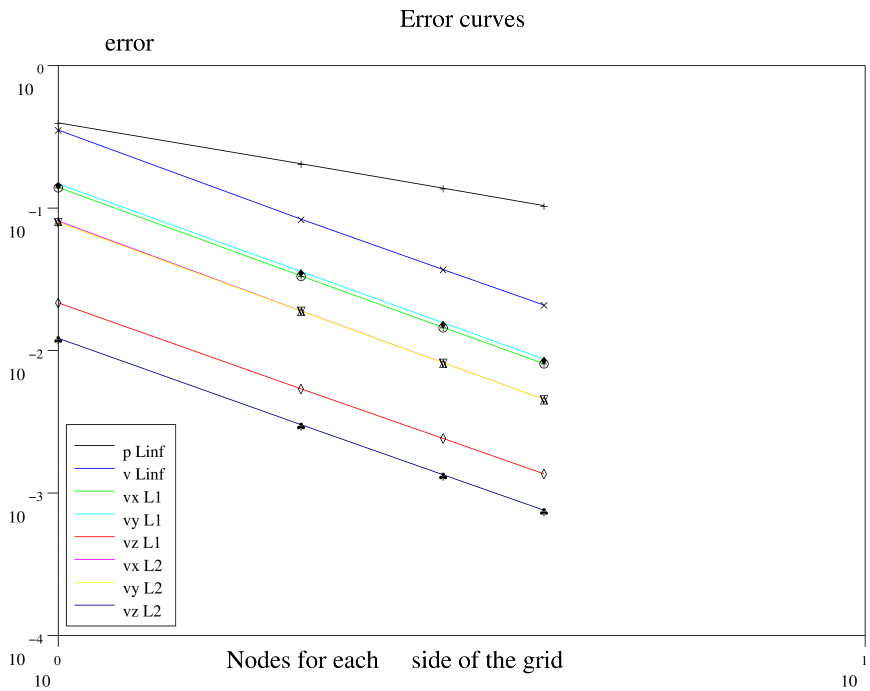


FIG. 5.6 – Error curves for the 3D-Stokes problem

and the cube,

$$\begin{cases} u_1 & = & 16y(1-y)z(1-z) \\ u_2 & = & 0 \\ u_3 & = & 0 \end{cases}.$$

For the complex flow, a constant velocity profile was chosen at the inlet and at the outlet whereas a slip boundary condition was imposed on the upper and downer bound in addition with an homogeneous Dirichlet boundary condition imposed on the circular or spheric bound. Without an exact solution of the problem, the only thing we can say is that the results are in good agreement with the physical sense.

5.12 Conclusion

The Stokes equation describes the motion of an incompressible Newtonian fluid characterized with a low Reynolds number. This is the case when the flow is laminar. A more complete description would be if the inertial terms aren't neglectible anymore. This is the case when the Reynolds number is big, and described by the Navier-Stokes equation. Solving the Navier-Stokes equation lead to the solving of several iterated Stokes problems. The problem is the transition from laminar to turbulent flow and turbulent flow itself that encounter more delays in realistic applications in aeronautic or forecast research fields examples. We explained the $\mathbb{P}_1$+bubble-$\mathbb{P}_1$ Finite Element approximation of the Stokes problem with technical details. It gives an approximation of order h^2 for the velocity and h for the pressure easy to handle, with a low number of degrees of freedom, interesting for 3 dimensions calculus, efficient and accurate. Another advantage is the application of the method to advection-diffusion equation that stabilizes the numerical scheme. Even if the method and the problem are not brand new, the study of more complex flow as in the next section would have been a pipe dream without a good knowledge of this scolar problem.

Bibliographie

[1] *R.B. Kellog and E. Osborn, A regularity result for the Stokes problem in a convex polygon, journal of functional analysis 21, p 397- 431, (1976).*

[2] P.A. Raviart, J.M. Thomas, Primal Hybrid Finite Element Methods for second Order Elliptic Equations , Mathematics of computation, vol. 31, n° 138, (1977), p 391-413.

[3] R. Temam Navier-Stokes Equations and non-linear functional analysis, SIAM, (1983).

[4] F. Brezzi, M. Fortin, Mixed and hybrid Finite Element Methods, Springer-Verlag New-York inc., (1991).

[5] A. Quarteroni, A. Valli, Numerical Approximation of Partial Differential Equations, Springer-Verlag, (1994).

[6] V. Girault, P.A. Raviart, Finite Element methods for Navier-Stokes equations- Theory and algorithms, Springer-Verlag, (1986).

[7] A.Russo, Bubble stabilization of finite element methods for the linearized incompressible Navier-Stokes equations , Comput. Methods appl. Mech. Eng. 132, p 335-343, (1996).

[8] F.Brezzi, L.P. Franca, A.Russo,, Further considerations on residual-free bubbles for advective-diffusive equations, Computer Methods in Applied Mechanics and Engineering, Volume 166, Issues 1-2, 13 November 1998, Pages 25-33.

[9] P.Knobloch, On the application of P_1^{mod} *element to incompressible flow problems , Computing and Visualization in Science 6, Springer-Verlag, p185-195, (2004).*

Chapitre 6

Numerical analysis of flow, transport precipitation and dissolution in a porous medium

Sommaire

6.1 Introduction **138**

6.2 The time discrete numerical scheme **142**

6.2.1 Stability in L^∞ 143

6.2.2 A priori estimates 146

6.2.3 Convergence 149

6.3 A fixed point iteration for the time discrete problems **155**

6.4 Numerical results **159**

6.5 Conclusion **167**

Cette section a fait l'objet des proceedings suivants :

V. Devigne, T. Clopeau "Flow, transport and crystal dissolution in a porous medium" proceeding for the conference Fourth Conference on Applied Mathematics and Scientific Computing, June 19-24, 2005 Brijuni island, Croatia.

V. Devigne, I.S. Pop "A numerical scheme for the micro scale dissolution and precipitation in porous media" proceeding for the conference ENUMATH 2005, July 18-22 2005, Santiago de Compostela (Spain).

6.1 Introduction

In this paper we consider a pore scale model for crystal dissolution and precipitation processes in porous media. This model is studied in [2] and represents the pore–scale analogue of the macroscopic (core scale) model proposed in [10]. The particularity of the model is in the description of the dissolution and precipitation processes taking place on the surface of the grains Γ_G, involving a multi–valued dissolution rate function. Models of similar type are analyzed in a homogenization context in [7], [8], and [15].

Without going into details, we briefly recall the background of the model. A fluid in which cations and anions are dissolved occupies the pores of a porous medium. The boundary of the void space consists of two disjoint parts : the surface of the porous skeleton (the grains), named from now one the internal boundary, and the external part, which is the outer boundary of the domain. Under certain conditions, the ions transported by the fluid can precipitate and form a crystalline solid, which is attached to the internal boundary and thus is immobile. The reverse reaction of dissolution is also possible.

The model proposed in [2] consists of several components : the Stokes flow in the pores, the transport of dissolved ions by convection and diffusion, and dissolution/precipitation reactions on the surface of the porous skeleton (grains). Here we assume that the flow geometry, as well as the fluid properties are not affected by the chemical processes. Therefore the flow component can be completely decoupled from the remaining part of the model. Further, the total electric charge can be considered in order to eliminate the anion from the system. The reason for doing so is in the resulting simplification of the model. The electric charge is satisfying a convection diffusion equation with homogeneous Neumann conditions on the grain surface. In this way the total charge can be decoupled from the chemistry once the fluid flow is determined (see [1] for details).

Our main interest is focused on the chemistry, this being the challenging part of the model. We continue here the analysis in [2] by investigating an appropriate numerical scheme for the dissolution and precipitation component. The present work is closely related to [6]. The numerical scheme proposed there involves also a special treatment of the diffusion on the grain boundary. The coupled system is solved by iterating between the equations in the pores, respectively on the solid matrix. Numerical methods for upscaled models for dissolution and precipitation in porous media are considered in [3], [13], [14], and [4]. The time stepping is performed by first order implicit approaches, and finite elements or finite volumes are employed for the spatial discretization. For the upscaled version of the present model, a numerical algorithm for computing traveling wave solutions is proposed in [10].

We denote the flow domain by $\Omega \subset \mathbb{R}^d$ $(d > 1)$, which is assumed open, connected and bounded. Its boundary $\partial\Omega$ is Lipschitz continuous. By (Γ_G) we mean the surface of the grains (the internal boundary), and by Γ_D the outer boundary of the domain. Let $\vec{\nu}$ denote the outer normal to $\partial\Omega$ and $T > 0$ be a fixed but arbitrarily chosen value of time. For X being Ω, Γ_G, or Γ_D, and any $0 < t \leq T$, we define

$$X^t = (0, t] \times X.$$

To simplify the presentation, we assume that the initial data are compatible (see [1], or [10]). Moreover, the total charge is assumed constant in time and space. Then the model can be reduced to

$$\begin{cases} \partial_t u + \nabla \cdot (\vec{q}u - D\nabla u) & = & 0, & \text{in } \Omega^T, \\ -D\vec{\nu} \cdot \nabla u & = & \varepsilon \tilde{n} \partial_t v, & \text{on } \Gamma_G^T, \\ u & = & 0, & \text{on } \Gamma_D^T, \\ u & = & u_I, & \text{in } \Omega, \text{ for } t = 0, \end{cases} \tag{1.1}$$

for the ion transport, and

$$\begin{cases} \partial_t v & = & D_a(r(u) - w), & \text{on } \Gamma_G^T, \\ w & \in & H(v), & \text{on } \Gamma_G^T, \\ v & = & v_I, & \text{on } \Gamma_G, \text{ for } t = 0, \end{cases} \tag{1.2}$$

for the precipitation and dissolution. H stands for the Heaviside graph,

$$H(u) = \begin{cases} 0, & \text{if } u < 0, \\ [0,1], & \text{if } u = 0, \\ 1, & \text{if } u > 0, \end{cases}$$

while w is the actual value of the dissolution rate. Here v denotes the concentration of the precipitate, which is defined only on the interior boundary Γ_G, while u stands for the cation concentration. The dissolution rate is constant (and scaled to 1) in the presence of crystal, i.e. for $v > 0$ somewhere on Γ_G. In the absence of crystals, the overall rate is either zero, if the fluid is not containing sufficient dissolved ions, or positive.

By $\vec{q}$ we denote the divergence free fluid velocity, and no slip is assumed along the internal grain boundary Γ_G :

$$\nabla \cdot \vec{q} = 0 \text{ in } \Omega, \qquad \vec{q} = \vec{0} \text{ on } \Gamma_G. \tag{1.3}$$

To avoid unnecessary technical complications we also assume that the fluid velocity $\vec{q}$ is bounded in Ω,

$$M_q = \|\vec{q}\|_{\infty,\Omega}. \tag{1.4}$$

For the Stokes model with homogeneous Dirichlet boundary conditions, this holds if, for example, the domain is polygonal (see [9] or [12]).

The present setting is a simplification of the dissolution/precipitation model in [2]. However, the main difficulties that are associated with that model are still present here : the system (1.1)–(1.2) is consisting of a parabolic equation that is coupled through the boundary conditions to a differential inclusion defined on a lower dimensional manifold. For the ease of presentation we have considered here only the case of homogeneous Dirichlet boundary conditions on the external boundary Γ_D, but the results can be extended to more general cases.

All the quantities and variables in the above are assumed dimensionless. D denotes the diffusion coefficient and $\tilde{n}$ the anion valence. D_a represents the ratio of the characteristic precipitation/dissolution time scale and the characteristic transport time scale - the Damköhler number, which is assumed to be of moderate order. By ε we mean the ratio of the characteristic pore scale and the reference (macroscopic) length scale. In a proper scaling (see Remark 1.2 of [2]), this gives

$$\varepsilon \text{ meas}(\Gamma_G) \approx \text{meas}(\Omega). \tag{1.5}$$

This balance is natural for a porous medium, where $\text{meas}(\Gamma_G)$ denotes the total surface of the porous skeleton and $\text{meas}(\Omega)$ the total void volume. Throughout this paper we keep the value of ε fixed, but this can be taken arbitrarily small.

Assuming mass action kinetics, with $[\cdot]_+$ denoting the non-negative part, the precipitation rate is defined by

$$r(u) = [u]_+^{\tilde{m}}[(\tilde{m}u - c)/\tilde{n}]_+^{\tilde{n}}, \tag{1.6}$$

where $\tilde{m}$ is the cation concentration and c the total negative charge, which is assumed here constant in time and space. The analysis here is not restricted to the typical example above, but assumes that r is an increasing, positive, locally Lipschitz continuous function. Further, since c is fixed in (1.6), there exists a unique $u_* \geq 0$ such that $r(u) = 0$ for all $u \leq u_*$, and r is strictly increasing for $u > u_*$. With respect to the reaction rate function $r(u)$ in (1.2) we recall the assumptions proposed in [2] :

(A_r) (i) $r : \mathbb{R} \to [0, \infty)$ is locally Lipschitz in $\mathbb{R}$;

(ii) $r(u) = 0$ for all $u \leq 0$;

(iii) there exists a unique $u_* \geq 0$, such that

$$r(u) = \begin{cases} 0, & \text{for} \quad u \leq u_*, \\ \text{strictly increasing for } u > u_* \text{ with } r(\infty) = \infty; \end{cases}$$

Remark 15 *In the setting above, a unique u^* exists for which $r(u^*) = 1$. If $u = u^*$ for all t and x, then the system is in equilibrium : no precipitation or dissolution occurs, since the precipitation rate is balanced by the dissolution rate regardless of the presence of absence of crystals.*

On the initial data we make the following assumption

(A_D) The initial data are bounded and non-negative ; further we assume $u_I \in H^1_{0,\Gamma_D}(\Omega)$ and $v_I \in L^2(\Gamma_G)$.

Due to the occurrence of the multi–valued dissolution rate, classical solutions do not exists, except for some particular cases. For defining a weak solution we consider the following sets

$$\begin{aligned} \mathcal{U} &:= \{u \in L^2(0,T; H^1_{0,\Gamma_D}(\Omega)) : \partial_t u \in L^2(0,T; H^{-1}(\Omega))\}, \\ \mathcal{V} &:= \{v \in H^1(0,T; L^2(\Gamma_G))\}, \\ \mathcal{W} &:= \{w \in L^\infty(\Gamma_G^T), : 0 \leq w \leq 1\}. \end{aligned}$$

Here we have used standard notations in the functional analysis.

Definition 6.1.1 *A triple $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times \mathcal{W}$ is called a weak solution of (1.1) and (1.2) if $(u(0), v(0)) = (u_I, v_I)$ and if*

$$(\partial_t u, \varphi)_{\Omega^T} + D(\nabla u, \nabla\varphi)_{\Omega^T} - (\vec{q}u, \nabla\varphi)_{\Omega^T} = -\varepsilon\tilde{n}(\partial_t v, \varphi)_{\Gamma_G^T}, \tag{1.7}$$

$$\begin{aligned} (\partial_t v, \theta)_{\Gamma_G^T} &= D_a(r(u) - w, \theta)_{\Gamma_G^T}, \\ w &\in H(v) \quad \textit{a.e. in } \Gamma_G^T, \end{aligned} \tag{1.8}$$

for all $(\varphi, \theta) \in L^2(0,T; H^1_{0,\Gamma_D}(\Omega)) \times L^2(\Gamma_G^T)$.

The existence of a weak solution is proven in [2], Theorem 2.21. Moreover, with

$$M_u := \max\{\|u_I\|_{\infty,\Omega}, u^*\}, \tag{1.9}$$

$$M_v := \max\{\|v_I\|_{\infty,\Omega}, 1\}, \quad C_v := \frac{r(M_u)D_a}{M_v}, \tag{1.10}$$

a weak solution satisfies

$$
\begin{aligned}
0 \le u \le M_u, &\quad \text{a.e. in } \Omega^T \\
0 \le v(t,\cdot) \le M_v e^{C_v t}, &\quad \text{for all } t \in [0,T] \text{ and a.e. on } \Gamma_G, \\
0 \le w \le 1, &\quad \text{a.e. on } \Gamma_G^T,
\end{aligned}
\tag{1.11}
$$

respectively

$$
\begin{aligned}
&\|u(t)\|_\Omega^2 + \|\nabla u\|_{\Omega^T}^2 + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))}^2 \\
&\qquad\qquad +\varepsilon \|v(t)\|_{\Gamma_G}^2 + \varepsilon \|\partial_t v\|_{\Gamma_G^T}^2 \le C,
\end{aligned}
\tag{1.12}
$$

for all $0 \le t \le T$. Here $C > 0$ is a generic constant. The proof is based on regularization arguments and provides a solution for which, in addition, we have

$$
w = r(u) \qquad \text{a.e. in } \{v = 0\} \cap \Gamma_G^T.
$$

6.2 The time discrete numerical scheme

In this section we analyze a semi–implicit numerical scheme for the system (1.1)–(1.2). To overcome the difficulties that are due to the multi–valued dissolution rate we approximate the Heaviside graph by

$$
H_\delta(v) := \begin{cases} 0, & \text{if } v \le 0, \\ v/\delta, & \text{if } v \in (0,\delta), \\ 1, & \text{if } v \ge \delta, \end{cases}
\tag{2.1}
$$

where $\delta > 0$ is a small regularization parameter.

Next we consider a time stepping that is implicit in u and explicit in v. Though possible here as well, an implicit discretization of v would involve an additional nonlinearity in v without bringing any significant improvement of the results.

With $N \in \mathbb{N}$, $\tau = T/N$, and $t_n = n\tau$, the approximation (u^n, v^n) of $(u(t_n), v(t_n))$ is the solution of the following problem :

Problem $\mathbf{P}_\delta^n$: *Given u^{n-1}, v^{n-1}, compute $u^n \in H^1_{0,\Gamma_D}(\Omega)$, and $v^n \in L^2(\Gamma_G)$ such that*

$$
\begin{aligned}
(u^n - u^{n-1}, \phi)_\Omega + \tau D(\nabla u^n, \nabla\phi)_\Omega - \tau(\vec{q}u^n, \nabla\phi)_\Omega& \\
+\epsilon\tilde{n}(v^n - v^{n-1}, \phi)_{\Gamma_G} &= 0,
\end{aligned}
\tag{2.2}
$$

$$
(v^n, \theta)_{\Gamma_G} = (v^{n-1}, \theta)_{\Gamma_G} + \tau D_a(r(u^n) - H_\delta(v^{n-1}), \theta)_{\Gamma_G},
\tag{2.3}
$$

for all $\phi \in H^1_{0,\Gamma_D}(\Omega)$ and $\theta \in L^2(\Gamma_G)$.

Here $n = 1, \ldots, N$, while $u^0 = u_I$ and $v^0 = v_I$. For the consistency with the original setting we approximate the dissolution rate $w(t_n)$ by

$$w^n := H_\delta(v^n). \tag{2.4}$$

To simplify the notations, we have given up the subscript δ for the solution triple (u^n, v^n, w^n).

Remark 16 *As we will see later, for guaranteeing the stability of the scheme the regularization parameter δ should be chosen such that $\delta > \tau D_a$. This is the only restriction that is related to the explicit discretization of v. Further, the convergence result is obtained assuming $\delta = O(\tau^\alpha)$ for any $\alpha \in (0,1)$, which is consistent with the previous restriction.*

Due to the explicit discretization of v, the ion transport equation in Problem P_δ^n can be decoupled from the time discrete dissolution and precipitation equation. This can be done by replacing the Γ_G scalar product in (2.2) by the last term in (2.3). For obtaining a solution for Problem P_δ^n, one has to solve first an elliptic problem with a nonlinear boundary condition. Once u^n is obtained, v^n can be determined straightforwardly from (2.3). We will use this writing later in Section 6.5, where a linear iteration is discussed. As a complementary result, the iteration proposed there can also be employed for proving the existence and uniqueness for Problem P_δ^n.

6.2.1 Stability in L^∞

All the estimates in this section should be interpreted in the almost everywhere (a.e.) sense. As follows from (1.11), the concentrations u and v, as well as the dissolution rate w are non-negative and bounded. Here we prove similar results for the time discrete concentrations u^n and v^n. The bounds for w^n are following straightforwardly from (2.4).

Lemma 6.2.1 *Assume $\delta \geq \tau D_a$ and that u^{n-1} and v^{n-1} are non-negative. Then u^n and v^n are non-negative as well.*

Proof. We start with the estimate in v^n. With $[\cdot]_-$ denoting the negative cut, we test (2.3) with $\theta := [v^n]_-$ and obtain

$$\|[v^n]_-\|^2_{\Gamma_G} = \tau D_a \left(r(u^n), [v^n]_-\right)_{\Gamma_G} + \left(v^{n-1} - \tau D_a H_\delta(v^{n-1}), [v^n]_-\right)_{\Gamma_G}$$

In view of (A_r), the first term on the right is non-positive. Further, since $v^{n-1} \geq 0$ and $\delta \geq \tau D_a$, by the construction of H_δ we have

$$v^{n-1} - \tau D_a H_\delta(v^{n-1}) \geq v^{n-1}(1 - \tau D_a/\delta) \geq 0,$$

almost everywhere on Γ_G. Hence the second term on the right is negative as well. This yields

$$\|[v^n]_-\|^2_{\Gamma_G} \leq 0,$$

implying the assertion for v^n.

For proving that u^n is non-negative we proceed in a similar manner. Testing (2.2) with $\phi := [u^n]_-$ gives

$$\begin{aligned} &\|[u^n]_-\|^2_\Omega + \tau D\,\|\nabla [u^n]_-\|^2_\Omega - \tau\,(\vec{q}u^n, \nabla [u^n]_-)_\Omega \\ &\qquad + \epsilon\tilde{n}\left(v^n - v^{n-1}, [u^n]_-\right)_{\Gamma_G} = (u^{n-1}, [u^n]_-)_\Omega. \end{aligned} \tag{2.5}$$

The first two terms in the above are non-negative, whereas the third one is vanishing. This follows from

$$\begin{aligned} (\vec{q}u^n, \nabla [u^n]_-)_\Omega &= \tfrac{1}{2}(\vec{q}, \nabla [u^n]_-^2)_\Omega \\ &= \tfrac{1}{2}(\vec{\nu}\cdot\vec{q}, [u^n]_-^2)_{\Gamma_D\cup\Gamma_G} - \tfrac{1}{2}(\nabla\cdot\vec{q}, [u^n]_-^2)_\Omega, \end{aligned}$$

and the boundary conditions on $\partial\Omega$, since $\nabla\cdot\vec{q} = 0$ in Ω.

Further, since $[u^n]_- \leq 0$ a.e. and it belongs to $H^1_{0,\Gamma_D}(\Omega)$, its trace $[u^n]_-|_{\Gamma_G}$ is a non-positive $L^2(\Gamma_G)$ function. Testing (2.3) with $[u^n]_-|_{\Gamma_G}$ gives

$$\begin{aligned} (v^n - v^{n-1}, [u^n]_-)_{\Gamma_G} &= \tau D_a(r(u^n) - H_\delta(v^{n-1}), [u^n]_-)_{\Gamma_G} \\ &= -\tau D_a(H_\delta(v^{n-1}), [u^n]_-)_{\Gamma_G} \geq 0, \end{aligned}$$

where we have used (A$_r$) and the positivity of H_δ.

Finally, the term on the right is non-positive, since $u^{n-1} \geq 0$. Using these observations into (2.5) shows that $[u^n]_- = 0$ almost everywhere in Ω, implying the result. ■

Remark 17 *Since we assume that the initial data are non-negative, Lemma 6.2.1 show that for all $n = 0, \ldots, N$ the time-discrete concentrations u^n and v^n are positive. As stated in Remark 16, the restriction $\delta \geq \tau D_a$ is due to the explicit treatment of v in (2.3). A fully implicit scheme would make it unnecessary.*

Now we turn our attention to the upper bounds for u^n and v^n. First, with M_u defined in (1.9) we have

Lemma 6.2.2 *If $u^{n-1} \leq M_u$, then the same holds for u^n.*

Proof. We test (2.2) with $\phi := [u^n - M_u]_+$, the non-negative part of $u^n - M_u$. This gives

$$\begin{aligned} &\|[u^n - M_u]_+\|^2_\Omega + \tau D\|\nabla[u^n - M_u]_+\|^2_\Omega - \tau(\vec{q}u^n, \nabla[u^n - M_u]_+)_\Omega \\ &\quad = (u^{n-1} - M_u, [u^n - M_u]_+)_\Omega - \epsilon\tilde{n}(v^n - v^{n-1}, [u^n - M_u]_+)_{\Gamma_G}. \end{aligned}$$

Arguing as in the proof of Lemma 6.2.1, we first observe that the convection term vanishes. Further, since $u^{n-1} \leq M_u$, the first term on the right is negative. Finally, for the last term we have

$$(v^n - v^{n-1}, [u^n - M_u]_+)_{\Gamma_G} = \tau D_a(r(u^n) - H_\delta(v^{n-1}), [u^n - M_u]_+)_{\Gamma_G}.$$

By the definition of M_u, whenever $u^n \geq M_u$ we have $r(u^n) \geq 1 \geq H_\delta(v^{n-1})$. This implies the positivity of the above scalar product. We are therefore left with

$$\|[u^n - M_u]_+\|_\Omega^2 + \tau D\|\nabla[u^n - M_u]_+\|_\Omega^2 \leq 0,$$

implying that $u^n \leq M_u$. ■

Remark 18 *Since $u_I \leq M_u$, Lemma 6.2.2 implies $u^n \leq M_u$ for all the time discrete approximations of u.*

The upper estimates for v are similar.

Lemma 6.2.3 *With M_v and C_v defined in* (1.10)*, assume that $v^{n-1} \leq M_v e^{C_v(n-1)\tau}$. Then $v^n \leq M_v e^{C_v n\tau}$.*

Proof. Testing (2.3) with $\theta := [v^n - M_v e^{C_v n\tau}]_+$ gives

$$\begin{aligned}\|[v^n - M_v e^{C_v n\tau}]_+\|_{\Gamma_G}^2 &= (v^{n-1} - M_v e^{C_v(n-1)\tau}, [v^n - M_v e^{C_v n\tau}]_+)_{\Gamma_G}\\ &+M_v(e^{C_v(n-1)\tau} - e^{C_v n\tau}, [v^n - M_v e^{C_v n\tau}]_+)_{\Gamma_G} \\ &+\tau D_a(r(u^n) - H_\delta(v^{n-1}), [v^n - M_v e^{C_v n\tau}]_+)_{\Gamma_G}.\end{aligned} \tag{2.6}$$

We denote the terms on the right by I_1, I_2 and I_3. We first notice that the assumption on v^{n-1} imply $I_1 \leq 0$. Further, since $0 \leq u^n \leq M_u$, $H_\delta(v^{n-1}) \geq 0$, and due to the monotonicity of r we obtain

$$I_4 \leq \tau D_a(r(M_u), [v^n - M_v e^{C_v n\tau}]_+)_{\Gamma_G}.$$

This gives

$$I_3 + I_4 \leq M_v(\tau C_v + e^{C_v(n-1)\tau}(1 - e^{\tau C_v}), [v^n - M_v e^{Cn\tau}]_+)_{\Gamma_G} \leq 0.$$

Here we have used the elementary inequality $e^x \geq 1 + x$, as well as $e^{C_v(n-1)\tau} \geq 1$. In this way (2.6) becomes

$$\|[v^n - M_v e^{C_v n\tau}]_+\|_{\Gamma_G}^2 \leq 0,$$

implying the upper bounds for v^n. ■

Remark 19 *As before, since $v_I \leq M_v$, it follows that $v^n \leq M_v e^{Cn\tau}$ for all $n = 1, \ldots, N$.*

Remark 20 *The essential bounds provided by Lemma 6.2.1, 6.2.2 and 6.2.3 are uniform in δ, assuming that $\delta \geq \tau D_a$. Moreover, τ only appears in the upper bounds for v^n. Nevertheless, given a $t \in [0, T]$, for any $\tau > 0$ and n such that $n\tau \leq t$ we have $v^n \leq M_v e^{C_v t}$. Therefore we can say that the estimates are uniform in τ as well.*

6.2.2 A priori estimates

We continue the analysis of the numerical scheme (2.2)–(2.3) by giving some energy estimates for the sequence of time discrete concentrations $\{(u^n, v^n),\ n = 0, \ldots, N\}$. We start with the estimates in v, which are depending on meas(Γ_G).

Lemma 6.2.4 *For any $n \geq 1$ we have :*

$$\begin{aligned} \left\|v^n - v^{n-1}\right\|_{\Gamma_G} &\leq \tau D_a\, r(M_u)\, \textit{meas}(\Gamma_G)^{1/2}, \qquad \textit{and} &(2.7)\\ \|v^n\|_{\Gamma_G} &\leq \|v_I\|_{\Gamma_G} + n\tau D_a\, r(M_u)\, \textit{meas}(\Gamma_G)^{1/2}. &(2.8)\end{aligned}$$

Proof. Testing (2.3) with $\theta := [v^n - v^{n-1}] \in L^2(\Gamma_G)$ and applying Cauchy's inequality gives,

$$\|v^n - v^{n-1}\|^2_{\Gamma_G} \leq \tau D_a \|r(u^n) - H_\delta(v^{n-1})\|_{\Gamma_G} \|v^n - v^{n-1}\|_{\Gamma_G} \tag{2.9}$$

The essential bounds on u^n and v^n, together with the assumptions on r and H_δ imply $-1 \leq r(u^n) - H_\delta(v^{n-1}) \leq r(M_u)$. By (1.9) we have $r(M_u) \geq 1$, implying the first estimates.

In a similar manner, by we taking $\theta := v^n \in L^2(\Gamma_G)$ in (2.3) and applying the Cauchy inequality we obtain

$$\|v^n\|^2_{\Gamma_G} \leq \|v^{n-1}\|_{\Gamma_G} \|v^n\|_{\Gamma_G} + \tau D_a\, r(M_u)\, \text{meas}(\Gamma_G)^{1/2} \|v^n\|_{\Gamma_G}. \tag{2.10}$$

Simplifying with $\|v^n\|_{\Gamma_G}$ and applying the inequality backward we immediately obtain the estimate (2.8). ∎

Remark 21 *Notice that the estimates in Lemma 6.2.4 are δ–independent. Next for $n \leq N$, the term on the right in (2.8) is bounded uniformly in τ as well. Finally, due to* (1.5) *we can replace meas(Γ_G) by C/ϵ, where C does not depend on δ, τ, or ϵ.*

Now we proceed by the estimates for u.

Lemma 6.2.5 *Assume $\tau \leq \tau_0$, with $\tau_0 > 0$ a fixed value that will be given below. For the time discrete solute concentrations we have*

$$\begin{aligned} \tau \sum_{n=1}^{N} \|\nabla u^n\|_{\Omega}^2 &\leq C && (2.11)\\ \sum_{n=1}^{N} \|u^n - u^{n-1}\|_{\Omega}^2 &\leq C\sqrt{\tau} && (2.12)\\ \tau \sum_{n=1}^{N} \|\nabla(u^n - u^{n-1})\|_{\Omega}^2 &\leq C\sqrt{\tau} && (2.13)\\ \sum_{n=1}^{N} \|u^n - u^{n-1}\|_{\Gamma_G}^2 &\leq C. && (2.14) \end{aligned}$$

Here C is a generic constant that does not depend on δ and τ.

Proof. We start by testing (2.2) with $\phi = u^n$. This gives

$$\begin{aligned} &(u^n - u^{n-1}, u^n)_{\Omega} + \tau D \|\nabla u^n\|_{\Omega}^2 \\ &\quad + \tau(\vec{q} u^n, \nabla u^n)_{\Omega} + \varepsilon \tilde{n}(v^n - v^{n-1}, u^n)_{\Gamma_G} = 0. \end{aligned} \tag{2.15}$$

We denote the terms on the right by $T_1, \ldots, T_4$. We have

$$T_1 = \frac{1}{2}(\|u^n\|_{\Omega}^2 - \|u^{n-1}\|_{\Omega}^2 + \|u^n - u^{n-1}\|_{\Omega}^2).$$

Further, T_2 is non-negative, whereas T_3 vanishes as argued in the proof of Lemma 6.2.1.

Before estimating the last term we notice the existence of positive constants C_1 and C_2 such that

$$\|\varphi\|_{\Gamma_G}^2 \leq C_1 \|\varphi\|_{\Omega}^2 + C_2 \|\varphi\|_{\Omega} \|\nabla \varphi\|_{\Omega} \tag{2.16}$$

for all $\varphi \in H^1_{0,\Gamma_D}(\Omega)$. This can be obtained by following the proof of the trace theorem (e.g. see [5], Theorem 1.5.1.10). By the inequality of means

$$ab \leq \frac{\delta}{2}a^2 + \frac{1}{2\delta}b^2, \qquad \text{for all } a, b, \text{ and } \delta > 0, \tag{2.17}$$

we get

$$\|\varphi\|_{\Gamma_G}^2 \leq \left(C_1 + \frac{C_2^2}{4\rho}\right) \|\varphi\|_{\Omega}^2 + \rho \|\nabla \varphi\|_{\Omega}^2. \tag{2.18}$$

Here $\rho > 0$ can be taken arbitrarily small.

With $M := \varepsilon \tilde{n} D_a\, r(M_u)\ \mathrm{meas}(\Gamma_G)^{1/2}$, we use (2.7) to estimate T_4 :

$$|T_4| \leq \tau M \|u^n\|_{\Gamma_G} \leq \frac{\tau M^2}{4} + \tau \|u^n\|_{\Gamma_G}^2.$$

Now we use (2.18) with $\delta = D/2$ and obtain

$$|T_4| \leq \frac{\tau M^2}{4} + \tau \left(C_1 + \frac{C_2^2}{2D}\right) \|u^n\|_{\Omega}^2 + \frac{\tau D}{2} \|\nabla u^n\|_{\Omega}^2.$$

Combining these estimates into (2.15), summing up for $n = 1, \dots, N$ and multiplying the result by 2 yields

$$\|u^N\|_\Omega^2 + \sum_{n=1}^N \|u^n - u^{n-1}\|_\Omega^2 + \tau D \sum_{n=1}^N \|\nabla u^n\|_\Omega^2 \le \|u_I\|_\Omega^2 + \frac{TM^2}{2} + C,$$

where $C = T\big(C_1 + C_2^2/(2D)\big) M_u^2 \mathrm{meas}(\Omega)$. This estimate follows by the essential bounds on u^n. An alternative proof can be given based on the discrete Gronwall lemma. In this way we have proven (2.11). Notice that we have also obtained

$$\sum_{n=1}^N \|u^n - u^{n-1}\|_\Omega^2 \le C,$$

which is not so good as (2.12).

To proceed with (2.12) and (2.13) we notice that, since $u_I \in H^1_{0,\Gamma_D}$, $u^n - u^{n-1}$ is a H^1_{0,Γ_D} function for all $n \ge 1$. Testing (2.2) with $u^n - u^{n-1}$ gives

$$\begin{aligned} &\|u^n - u^{n-1}\|_\Omega^2 + \tau D(\nabla u^n, \nabla(u^n - u^{n-1}))_\Omega \\ &-\tau(\vec{q} u^n, \nabla(u^n - u^{n-1}))_\Omega + \epsilon \tilde{n} \left(v^n - v^{n-1}, u^n - u^{n-1}\right)_\Omega = 0. \end{aligned} \tag{2.19}$$

Denoting the terms in the above by $I_1, \dots, I_4$, we have

$$I_2 = \frac{\tau D}{2}(\|\nabla u^n\|_\Omega^2 - \|\nabla u^{n-1}\|_\Omega^2 + \|\nabla(u^n - u^{n-1})\|_\Omega^2).$$

Recalling (1.4), the inequality of means gives

$$\begin{aligned} |I_3| &= \tau(\nabla \cdot (\vec{q} u^n), u^n - u^{n-1})_\Omega \le \tau M_q \|\nabla u^n\|_\Omega \|u^n - u^{n-1}\|_\Omega \\ &\le \tfrac{1}{2}\|u^n - u^{n-1}\|_\Omega^2 + \tfrac{\tau^2 M_q^2}{2}\|\nabla u^n\|_\Omega^2. \end{aligned}$$

With M defined above, for I_4 we use the estimate (2.7) and obtain

$$|I_4| \le \varepsilon \tilde{n} \|v^n - v^{n-1}\|_{\Gamma_G} \|u^n - u^{n-1}\|_{\Gamma_G} \le \frac{(\tau M)^2}{4\mu_1} + \mu_1 \|u^n - u^{n-1}\|_{\Gamma_G}^2,$$

where $\mu_1 > 0$ will be chosen below. Now we can take $\rho = \tau D/(4\mu_1)$ into (2.18) to obtain

$$|I_4| \le \frac{(\tau M)^2}{4\mu_1} + \mu_1 \left(C_1 + \frac{\mu_1 C_2^2}{\tau D}\right) \|u^n - u^{n-1}\|_\Omega^2 + \frac{\tau D}{4}\|\nabla(u^n - u^{n-1})\|_\Omega^2.$$

Using the above estimates into (2.19), taking $\mu_1 = \sqrt{\tau D}/C_2$, multiplying the result by 2 and summing up for $n = 1, \dots, N$, (2.11) gives

$$\begin{aligned} &\left(\tfrac{1}{2} - \tfrac{\sqrt{\tau D}}{C_2}\right) \textstyle\sum_{n=1}^N \|u^n - u^{n-1}\|_\Omega^2 + \tfrac{\tau D}{2} \textstyle\sum_{n=1}^N \|\nabla(u^n - u^{n-1})\|_\Omega^2 \\ &\le \tau D \|\nabla u_I\|_\Omega^2 + \tfrac{TM^2 C_2}{\sqrt{D}} \tau^{1/2} + \tau C. \end{aligned}$$

For $\tau_0 := DC_2^2/(16C_1^2)$ and $\tau \le \tau_0$, since $u_I \in H^1_{0,\Gamma_D}$ the inequality above immediately implies (2.12) and (2.13).

Finally, using (2.18) with $\rho = \tau^{1/2}$, (2.14) is a direct consequence of (2.12) and (2.13). ■

Remark 22 *As in Remark 21, if the medium is ϵ–periodic, the generic constant C in Lemma 6.2.5 does not depend on ϵ. To see this we recall* (1.5)*, and Lemma 3 of [7], saying that there exists a constant $C > 0$, independent of ε, such that*

$$\varepsilon\|\varphi\|_{\Gamma_G}^2 \leq C\left(\|\varphi\|_\Omega^2 + \varepsilon^2\|\nabla\varphi\|_\Omega^2\right)$$

for all $\varphi \in H^1(\Omega)$. Then the boundary term in (2.19) yields a constant C which does not depend on ε as well.

6.2.3 Convergence

In this part we proceed by proving the convergence of the numerical scheme defined in (2.2)–(2.3) to a weak solution of the system (1.1)–(1.2), as given in Definition 6.1.1. The multivalued dissolution rate hinders us in obtaining useful error estimates. Therefore convergence will be achieved by compactness arguments. In doing so we mainly follow the ideas in [2].

Considering the sequence of time discrete triples $\{(u^n, v^n, w^n), n = 1, \ldots, N\}$ solving the problems P_δ^n, where w^n is defined in (2.4), we construct an approximation of the solution of (1.1) and (1.2) for all times $t \in [0, T]$. Specifically, define for any $n = 1, \ldots, N$ and $t \in (t_{n-1}, t_n]$

$$Z^\tau(t) := z^n \frac{(t - t_{n-1})}{\tau} + z^{n-1}\frac{(t_n - t)}{\tau}, \tag{2.20}$$

where (z, Z) stands for (u, U) or (v, V), and define

$$W^\tau(t) := H_\delta(V^\tau(t)). \tag{2.21}$$

Notice that U_τ, V_τ, and W_τ do not only depend on τ, but also on the regularization parameter δ.

For the time continuous triple $\{U^\tau, V^\tau, W^\tau)\}$ we can use the uniform L^∞ bounds, as well as the a priori estimates in Lemmata 6.2.4 and 6.2.5, to derive δ–independent estimates that are similar to those for the solution defined in Definition 6.1.1 (see [2]).

Lemma 6.2.6 *Assume $\delta \geq \tau D_a$. Then for (U^τ, V^τ, W^τ) we have :*

$$0 \leq U^\tau \leq \quad M_u, \qquad \textit{a.e. in } \Omega^T, \tag{2.22}$$

$$0 \leq V^\tau \leq M_v e^{CT},\ 0 \leq W^\tau \leq \quad 1, \qquad \textit{a.e. in } \Gamma_G^T, \tag{2.23}$$

$$\|U^\tau(t)\|_\Omega + \|V^\tau(t)\|_{\Gamma_G}^2 \leq \quad C, \qquad \textit{for all } 0 \leq t \leq T, \tag{2.24}$$

$$\|\partial_t U^\tau\|_{L^2(0,T;H^{-1}(\Omega))}^2 + \|\nabla U^\tau\|_{\Omega^T}^2 + \|\partial_t V^\tau\|_{\Gamma_G^T}^2 \leq C. \tag{2.25}$$

Here $C > 0$ is a generic constant that does not depend on τ or δ.

Proof. The essential bounds in (2.22) and (2.23) are a direct consequence of Lemmata 6.2.1, 6.2.2, 6.2.3, and of (2.4). The same holds for (2.24), for which we employ the stability estimates in Lemmata 6.2.4 and 6.2.5.

To proceed with the gradient estimates in (2.25) we notice that

$$\begin{aligned}\int_0^T \|\nabla U^\tau(t)\|_\Omega^2 dt &\leq 2\sum_{n=1}^{N} \tau \|\nabla u^{n-1}\|_\Omega^2 \\ &\quad + \int_{t_{n-1}}^{t_n} 2\frac{(t-t_{n-1})^2}{\tau^2}\|\nabla(u^n - u^{n-1})\|_\Omega^2 dt \\ &\leq 2\tau\sum_{n=1}^{N}\|\nabla u^{n-1}\|_\Omega^2 + \frac{2\tau}{3}\sum_{n=1}^{N}\|\nabla(u^n - u^{n-1})\|_\Omega^2 \leq C.\end{aligned}$$

Here we have used the estimate in (2.11) and (2.13).

The estimate on $\partial_t V^\tau$ follows by (2.7),

$$\int_0^T \|\partial_t V^\tau(t)\|_{\Gamma_G}^2 dt = \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\|\frac{v^n - v^{n-1}}{\tau}\right\|_{\Gamma_G}^2 dt \leq CN\tau = CT.$$

Finally, for estimating $\partial_t U^\tau$ we notice that

$$(\partial_t U^\tau(t), \phi) = \left(\frac{u^n - u^{n-1}}{\tau}, \phi\right),$$

for all $\phi \in H^1_{0,\Gamma_D}(\Omega)$ and all $t \in (t_{n-1}, t_n]$. By (2.2) this gives

$$\begin{aligned}|(\partial_t U^\tau, \phi)| &\leq D\,\|\nabla u^n\|_\Omega\,\|\nabla\phi\|_\Omega + M_q\,\|\nabla u^n\|_\Omega\,\|\phi\|_\Omega \\ &\quad + \frac{\epsilon\tilde{n}}{\tau}\left\|v^n - v^{n-1}\right\|_{\Gamma_G}\,\|\phi\|_{\Gamma_G}.\end{aligned}$$

Using the trace theorem we obtain

$$\begin{aligned}|(\partial_t U^\tau, \phi)| &\leq (D + M_q)\|\nabla u^n\|_\Omega\|\phi\|_{H^1(\Omega)} \\ &\quad + C(\Omega)\frac{\epsilon\tilde{n}}{\tau}\|v^n - v^{n-1}\|_{\Gamma_G}\|\phi\|_{H^1(\Omega)},\end{aligned} \tag{2.26}$$

for any $\phi \in H^1_{0,\Gamma_D}(\Omega)$. This implies that

$$\|\partial_t U^\tau(t)\|_{H^{-1}(\Omega)} \leq C\left(\|\nabla u^n\|_\Omega + \frac{\epsilon\tilde{n}}{\tau}\|v^n - v^{n-1}\|_{\Gamma_G}\right), \tag{2.27}$$

for all $t \in (t_{n-1}, t_n]$, and the remaining part is a direct consequence of the L^∞ and stability estimates. ■

Remark 23 *For an ϵ–periodic medium, the estimates in Lemma 6.2.6 can be made ϵ–independent. This follows from the remarks 21 and 22. In this case the estimates* (2.24) *and*

(2.25) *become*

$$\|U^\tau(t)\|_\Omega + \|\partial_t U^\tau\|^2_{L^2(0,T;H^{-1}(\Omega))} + \|\nabla U^\tau\|^2_{\Omega^T} + \epsilon\left(\|V^\tau(t)\|^2_{\Gamma_G} + \|\partial_t V^\tau\|^2_{\Gamma_G^T}\right) \leq C,$$

for all $0 \leq t \leq T$*, where* C *does not depend on* τ*,* δ*, or* ε*.*

Having the τ and δ uniform estimates in Lemma 6.2.6, we can proceed by sending τ and δ to 0. For any $\tau > 0$ and $\delta \geq \tau D_a$ we have $(U^\tau, V^\tau, W^\tau) \in \mathcal{U} \times \mathcal{V} \times L^\infty(\Gamma_G^T)$. As stated in Remark 16, δ should also satisfy $\delta = O(\tau^\alpha)$ with $\alpha \in (0,1)$, the reasons for this being made clear later in this paragraph. From now one we assume that both restrictions are satisfied. Then obviously $\tau \searrow 0$ implies the same for δ. Compactness arguments give the existence of a triple $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times L^\infty(\Gamma_G^T)$ and a subsequence $\tau \searrow 0$, such that

a)	$U^\tau \to u$	weakly in	$L^2((0,T); H^1_{0,\Gamma_D}(\Omega))$,
b)	$\partial_t U^\tau \to \partial_t u$	weakly in	$L^2((0,T); H^{-1}(\Omega))$,
c)	$V^\tau \to v$	weakly in	$L^2((0,T); L^2(\Gamma_G))$,
d)	$\partial_t V^\tau \to \partial_t v$	weakly in	$L^2(\Gamma_G^T)$,
e)	$W^\tau \to w$	weakly–star in	$L^\infty(\Gamma_G^T)$.

It remains to show that the limit triple (u, v, w) solves (1.1)–(1.2) weakly. This is done in the following

Theorem 6.2.7 *The limit triple* (u, v, w) *is a weak solution of* (1.1)–(1.2) *in the sense of Definition 6.1.1. Moreover, for the dissolution rate we have*

$$w = r(u) \qquad \text{a.e. in } \{v = 0\} \cap \Gamma_G^T.$$

Proof. By the weak convergence, all the estimates stated in Lemma 6.2.6 hold for the limit triple (u, v, w). Furthermore, for any $t \in (t_{n-1}, t_n]$, by (2.2) we have

$$\begin{aligned} &(\partial_t U^\tau(t), \phi)_\Omega + D(\nabla U^\tau(t), \nabla\phi)_\Omega \\ &\qquad + (\nabla \cdot (\vec{q} U^\tau(t)), \phi)_\Omega + \epsilon\tilde{n}(\partial_t V^\tau(t), \phi)_{\Gamma_G} \\ &= D(\nabla(U^\tau(t) - u^n), \nabla\phi)_\Omega + (\nabla.(\vec{q}(U^\tau(t) - u^n)), \phi)_\Omega, \end{aligned} \tag{2.28}$$

for all $\phi \in H^1_{0,\Gamma_D}(\Omega)$. Denoting the terms on the right by $I_1(t)$ and $I_2(t)$, taking $\phi \in L^2(0,T; H^1_{0,\Gamma_D}(\Omega))$, and integrating (2.28) in time gives

$$\begin{aligned} &(\partial_t U^\tau, \phi)_{\Omega^T} + D(\nabla U^\tau, \nabla\phi)_{\Omega^T} \\ &\qquad + (\nabla \cdot (\vec{q}\, U^\tau), \phi)_{\Omega^T} + \epsilon\tilde{n}(\partial_t V^\tau, \phi)_{\Gamma_G^T} \\ &= \textstyle\sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_1(t) + I_2(t) dt. \end{aligned} \tag{2.29}$$

The definition (2.20) of U^τ implies for almost all $0 \le t \le T$

$$|I_1(t)| \le D\frac{t_n - t}{\tau}\|\nabla(u^n - u^{n-1})\|_\Omega \|\nabla\phi\|_\Omega, \tag{2.30}$$

respectively

$$|I_2(t)| \le M_q\frac{t_n - t}{\tau}\|u^n - u^{n-1}\|_\Omega \|\nabla\phi\|_\Omega. \tag{2.31}$$

In (2.31) we have integrated by parts and used the essential bounds (1.4) on $\vec{q}$. Now we can proceed by estimating the terms on the right in (2.29). For I_1 we get

$$\begin{aligned}
\left|\sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_1(t)dt\right| &\le \sum_{n=1}^N \int_{t_{n-1}}^{t^n} \tfrac{t_n - t}{\tau} D\|\nabla(u^n - u^{n-1})\|_\Omega \|\nabla\phi(t)\|_\Omega dt \\
&\le \sum_{n=1}^N \left(\tfrac{\tau}{3}D^2\|\nabla(u^n - u^{n-1})\|_\Omega^2\right)^{\frac12} \left(\int_{t_{n-1}}^{t^n} \|\nabla\phi(t)\|_\Omega^2 dt\right)^{\frac12} \\
&\le \tfrac{1}{2\sqrt{3}}\tau^{\frac14}\int_0^T \|\nabla\phi(t)\|_\Omega^2 dt + \tfrac{\tau^{\frac34}}{2\sqrt{3}}\sum_{n=1}^N D^2\|\nabla(u^n - u^{n-1})\|_\Omega^2 \\
&\le \tfrac{1}{2\sqrt{3}}\left(\int_0^T \|\nabla\phi(t)\|_\Omega^2 dt + C\right)\tau^{\frac14}.
\end{aligned}$$

In the above we have use the inequality of means (2.17) with $\delta = \tau^{1/4}$, as well as the estimates (2.12) and (2.13). In a similar manner, for I_2 we get

$$\left|\sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_1(t)dt\right| \le \left(\tfrac{1}{2\sqrt{3}}\int_0^T \|\nabla\phi(t)\|_\Omega^2 dt + C\tau^{\frac12}\right)\tau^{\frac14}.$$

Letting $\tau \searrow 0$, the weak convergence of U^τ and V^τ, as well as the estimates above imply that u and v satisfy (1.7).

For the dissolution–precipitation equation (1.8) we start by analyzing the behavior of U^τ on Γ_G^T. The a–priori estimates, together with Lemma 9 and Corollary 4 of [19], imply

$$U^\tau \to u \quad \text{strongly in} \quad C([0,T]; H^{-s}(\Omega)) \cap L^2(0,T; H^s(\Omega))$$

for any $s \in (0,1)$. Then the trace theorem, see for example Satz 8.7 of [23], gives

$$U^\tau \to u \quad \text{strongly in} \quad L^2(\Gamma_G^T).$$

Taking into account the Lipschitz continuity of r, this yields

$$r(U^\tau) \to r(u) \quad \text{strongly in} \quad L^2(\Gamma_G^T).$$

Further we proceed as for (1.7). For any $t_{n-1} < t \le t_n$ and $\theta \in L^2(\Gamma_G)$ we rewrite (2.3) as

$$\begin{aligned}
(\partial_t V^\tau(t), \theta)_{\Gamma_G} &= D_a(r(U^\tau(t)) - W^\tau(t), \theta)_{\Gamma_G} \\
&\quad + D_a(r(u^n) - r(U^\tau(t)), \theta)_{\Gamma_G} \qquad (2.32)\\
&\quad + D_a(W^\tau(t) - H_\delta(v^{n-1}), \theta)_{\Gamma_G}.
\end{aligned}$$

Denoting the last two terms on the right by $I_3(t)$ and $I_4(t)$, with $\theta \in L^2(\Gamma_G^T)$, we integrate (2.32) in time and obtain

$$(\partial_t V^\tau, \theta)_{\Gamma_G^T} = D_a(r(U^\tau) - W^\tau, \theta)_{\Gamma_G^T} + \sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_3(t) + I_4(t) dt. \tag{2.33}$$

For almost all $0 \le t \le T$, since r and H_δ are Lipschitz, we use the definition of V^τ and W^τ in (2.20) and (2.21) to obtain

$$|I_3(t)| \le D_a L_r \frac{(t_n - t)}{\tau} \|u^n - u^{n-1}\|_{\Gamma_G} \|\theta\|_{\Gamma_G}, \tag{2.34}$$

respectively

$$|I_4(t)| \le \frac{D_a}{\delta} \frac{(t_n - t)}{\tau} \|v^n - v^{n-1}\|_{\Gamma_G} \|\theta\|_{\Gamma_G}. \tag{2.35}$$

Using the estimate in (2.14), the first sum in (2.33) is bounded by

$$\begin{aligned}
\left|\sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_3(t) dt\right| &\le D_a L_r \sum_{n=1}^N \int_{t_{n-1}}^{t^n} \tfrac{t_n - t}{\tau} \|u^n - u^{n-1}\|_{\Gamma_G} \|\theta(t)\|_{\Gamma_G} dt \\
&\le D_a L_r \sum_{n=1}^N \left(\tfrac{\tau}{3}\|u^n - u^{n-1}\|_{\Gamma_G}^2\right)^{\frac{1}{2}} \left(\int_{t_{n-1}}^{t^n} \|\theta(t)\|_{\Gamma_G}^2 dt\right)^{\frac{1}{2}} \\
&\le \tfrac{D_a L_r}{2\sqrt{3}} \left(\|\theta\|_{\Gamma_G^T}^2 + \sum_{n=1}^N \|u^n - u^{n-1}\|_{\Gamma_G}^2\right) \tau^{\frac{1}{2}} \\
&\le \tfrac{D_a L_r}{2\sqrt{3}} \left(\|\theta\|_{\Gamma_G^T}^2 + C\right) \tau^{\frac{1}{2}}.
\end{aligned}$$

In the above we have use the inequality of means (2.17) with $\delta = \tau^{1/2}$. Similarly, for the last sum in (2.33), by (2.7) we get

$$\begin{aligned}
\left|\sum_{n=1}^N \int_{t_{n-1}}^{t^n} I_4(t) dt\right| &\le \tfrac{D_a}{\delta} \sum_{n=1}^N \int_{t_{n-1}}^{t^n} \tfrac{t_n - t}{\tau} \|v^n - v^{n-1}\|_{\Gamma_G} \|\theta(t)\|_{\Gamma_G} dt \\
&\le \tfrac{D_a}{\delta} \sum_{n=1}^N \left(\tfrac{\tau}{3}\|v^n - v^{n-1}\|_{\Gamma_G}^2\right)^{\frac{1}{2}} \left(\int_{t_{n-1}}^{t^n} \|\theta(t)\|_{\Gamma_G}^2 dt\right)^{\frac{1}{2}} \\
&\le \tfrac{1}{2\sqrt{3}} \left(\tau {D_a}^2 \|\theta\|_{\Gamma_G^T}^2 + \sum_{n=1}^N \|v^n - v^{n-1})\|_{\Gamma_G}^2\right) \frac{1}{\delta} \\
&\le \tfrac{{D_a}^2}{2\sqrt{3}} \left(\|\theta\|_{\Gamma_G^T}^2 + r(M_u)^2 \operatorname{meas}(\Gamma_G)\right) \frac{\tau}{\delta}.
\end{aligned}$$

Letting now $\tau \searrow 0$, with $\delta = O(\tau^\alpha)$ for an $\alpha \in (0, 1)$, the above estimates together with the weak convergence of $\partial_t V^\tau$, the strong convergence of $r(U^\tau)$, as well as the weak–star convergence of W^τ implies that u, v and w satisfy (1.8_1).

It only remains to show that the dissolution rate satisfies (1.8_2). The major impediment is the lacking pointwise V^τ convergence. To resolve this we follow the procedure in Theorem 2.21 of [2]. We define

$$\underline{V}(t, x) := \liminf_{\tau \searrow 0} V^\tau(t, x) \ge 0 \qquad \text{a.e. in } \Gamma_G^T,$$

and decompose Γ_G^T in two subsets defined as

$$S_1 = \{\underline{V} > 0\} \qquad \text{and} \qquad S_2 = \{\underline{V} = 0\}.$$

The equalities should be understood in the almost everywhere sense. Obviously we have $\Gamma_G^T = S_1 \cup S_2$. As in [2] we show that $v > 0$ and $w = 1$ in S_1, while $v = 0$ and $w = r(u) \in [0,1]$ in S_2.

For proving the first assertion we follow exactly the steps in [2], with (V^τ, W^τ) replacing the pair (v_δ, w_δ) there. To show the assertion in S_2 we use the reaction equation (2.33) and rule out the possibility that $v > 0$ in S_2. The weakly–star convergence of W^τ implies that $\int_0^t W^\tau \to \int_0^t w$ weakly–star in $L^\infty(\Gamma_G^T)$, and therefore $\liminf_{\tau\searrow 0} \int_0^t W^\tau \leq \int_0^t w$ a.e. in Γ_G^T. Further, both (1.8_1) and (2.33) hold in $L^2(\Gamma_G^T)$ and thus a.e. in Γ_G^T. For any $n = 1, \ldots, N$ and a.e. $x \in \Gamma_G$ and $t \in (t_{n-1}, t_n]$ we integrate the two reaction equations in time and obtain

$$\begin{aligned} V^\tau &= v_0 + D_a\left(\int_0^t (r(U^\tau) - W^\tau) + \Sigma_u - \Sigma_w\right), \\ &= v + D_a\left(\int_0^t (r(U^\tau) - r(u))\right. \\ &\qquad\qquad \left. - \int_0^t (W^\tau - w) + \Sigma_u - \Sigma_w\right), \end{aligned} \tag{2.36}$$

with

$$\begin{aligned} \Sigma_u &= \sum_{p=1}^{n-1} \int_{t_{p-1}}^{t_p} (r(u^p) - r(U^\tau)) + \int_{t_{n-1}}^t (r(u^p) - r(U^\tau)), \\ \Sigma_w &= \sum_{p=1}^{n-1} \int_{t_{p-1}}^{t_p} (H_\delta(v^{p-1}) - W^\tau)) + \int_{t_{n-1}}^t (r(u^p) - r(U^\tau)). \end{aligned}$$

We now analyze the Σ terms in the above. For a.e $(t,x) \in \Gamma_G^T$ we have

$$\begin{aligned} |\Sigma_u| &\leq L_r \sum_{p=1}^n \int_{t_{p-1}}^{t_p} \tfrac{t_p - t}{\tau} |u^p - u^{p-1}| \leq \tfrac{L_r}{2}\tau\left(n \sum_{p=1}^n |u^p - u^{p-1}|^2\right)^{\frac{1}{2}}. \\ &\leq \tfrac{L_r}{2}\left(\tau T \sum_{p=1}^N |u^p - u^{p-1}|^2\right)^{\frac{1}{2}}. \end{aligned}$$

By the estimates (2.14), the $L^2(\Gamma_G)$ norm of the term on the right vanishes as $\tau \searrow 0$. This also implies the convergence to 0 a.e.

For Σ_w we obtain analogously

$$\begin{aligned} |\Sigma_w| &\leq \frac{\tau}{2\delta}\left(N \sum_{p=1}^N |v^p - v^{p-1}|^2\right)^{\frac{1}{2}} \\ &\leq \frac{\tau^{(1+\alpha)/2}}{2\delta}\left(\tau^{-\alpha}\, T \sum_{p=1}^N |v^p - v^{p-1}|^2\right)^{\frac{1}{2}}. \end{aligned}$$

Since $\delta = O(\tau^\alpha)$, the estimates (2.7) imply the pointwise convergence of the right hand side a.e as $\tau \searrow 0$.

With $(t,x) \in S_2$, we take the $\liminf_{\tau \searrow 0}$ in (2.36). The Σ–terms vanish, as seen above. Further, since U^τ converges strongly on Γ_G^T and thus pointwisely, we find

$$0 = v - D_a \liminf_{\tau \searrow 0} \int_0^t (W^\tau - w) \geq v \qquad \text{a.e. in } S_2.$$

Therefore $v = 0$ in S_2. Moreover, since $\partial_t v \in L^2(\Gamma_G^T)$, it follows that $\partial_t v = 0$ a.e. in S_2, and therefore $w = r(u)$ with $0 \leq w \leq 1$. ■

Remark 24 *Besides the convergence of the numerical scheme, Theorem 6.2.7 also provides an alternative existence proof for a weak solution of* (1.1)–(1.2). *In [2] this is obtained by fixed point arguments, whereas here the solution is the limit of a time discrete approximating sequence.*

6.3 A fixed point iteration for the time discrete problems

In this section we analyze a linear iteration scheme for solving the non-linear time discrete problems P_δ^n. The nonlinearity appears in the boundary condition (2.3). Numerical experiments based on a Newton type iteration led to instabilities in the form of negative concentrations, or to a precipitation in the undersaturated regime ($u \leq u^*$). Moreover, there is no guarantee of convergence, unless the time stepping is not small enough. Here we discuss an alternative fixed point approach that provides stable results. Moreover, the scheme converges linearly in H^1, regardless of the initial iteration or of the parameters ε, τ and δ.

Assume u^{n-1} and v^{n-1} given and satisfying the bounds in Lemmata 6.2.1, 6.2.2, and 6.2.3. To construct the iteration scheme we proceed as discussed in the beginning of Section 6.2 and decouple the ion transport equation from the dissolution/precipitation equation on the boundary. Using (2.3), (2.2) can rewritten as

$$\begin{aligned} (u^n - u^{n-1}, \phi)_\Omega \quad &+ \quad \tau D(\nabla u^n, \nabla \phi)_\Omega + \tau(\nabla.(\vec{q} u^n), \phi)_\Omega \\ &+ \quad \tau \epsilon \tilde{n} D_a (r(u^n) - H_\delta(v^{n-1}), \theta)_{\Gamma_G} = 0, \end{aligned} \tag{3.1}$$

for all $\phi \in H^1_{0,\Gamma_D}(\Omega)$. This is a scalar elliptic equation with nonlinear boundary conditions on Γ_G. We first construct a sequence $\{u^{n,i}, i \geq 0\}$ approximating the solution u^n of (3.1). Once this is computed, we use (2.3) for directly determining v^n.

Let L_r be the Lipschitz constant of the precipitation rate r on the interval $[0, M_u]$. With

$$\mathcal{K} := \{u \in H^1_{0,\Gamma_D}(\Omega) / \ 0 \leq u \leq M_u \text{ a. e. in } \Omega\}, \tag{3.2}$$

and for a given $u^{n,i-1} \in \mathcal{K}$, we define $u^{n,i}$ as the solution of the linear elliptic equation

$$
\begin{aligned}
&(u^{n,i} - u^{n-1}, \phi)_\Omega + \tau D(\nabla u^{n,i}, \nabla \phi)_\Omega + \tau(\nabla(\vec{q} u^{n,i}), \phi)_\Omega \\
&\quad = \tau \epsilon \tilde{n} D_a L_r (u^{n,i-1} - u^{n,i}, \phi)_{\Gamma_G} \qquad (3.3) \\
&\qquad - \tau \epsilon \tilde{n} D_a (r(u^{n,i-1}) - H_\delta(v^{n-1}), \theta)_{\Gamma_G},
\end{aligned}
$$

for all $\phi \in H^1_{0,\Gamma_D}(\Omega)$. The starting point of the iteration can be chosen arbitrarily in $\mathcal{K}$. A good initial guess is $u^{n,0} = u^{n-1}$.

Comparing the above to (3.1), disregarding the superscripts $i-1$ and i, the only difference is in the appearance of the first term on the right in (3.3). In the case of convergence, this term vanishes, so $u^{n,i}$ approaches u^n. Before making this sentence more precise we mention that the above construction is common in the analysis of nonlinear elliptic problems, in particular when sub- or supersolutions are sought (see, e. g., [22], pp. 96). In [17], this approach is placed in a fixed point context, for approximating the solution of an elliptic problem with a nonlinear and possibly unbounded source term (see also [24]). The same ideas are followed in [20] and [16], where similar schemes are considered for the implicit discretization of a degenerate (fast diffusion) problem in both conformal and mixed formulation. We also mention here [21] for a related work on nonlinear elliptic equations.

Since (3.3) defines a fixed point iteration, only a linear convergence rate is to be expected. This drawback is compensated by the stability of the approximation sequence, as well as its guaranteed convergence. These statements are made precise below.

Lemma 6.3.1 *Assume $0 \leq u^{n,i-1} \leq M_u$ almost everywhere on Ω. Then $u^{n,i}$ solving (3.3) satisfies the same bounds.*

Proof. This can be shown by following the ideas used in proving Lemmata 6.2.1 and 1.9. We omit the details here. ■ Starting the iteration with $u^{n,0} \in \mathcal{K}$, a straightforward mathematical induction argument shows the stability of the entire sequence $\{u^{n,i}, i \geq 0\}$. To prove the convergence of the scheme we let

$$e^{n,i} := u^n - u^{n,i} \qquad (3.4)$$

denote the error at iteration i and define the H^1–equivalent norm

$$|||f|||^2 := \|f\|^2_\Omega + \tau D \|\nabla f\|^2_\Omega + \zeta \|f\|^2_{\Gamma_G}. \qquad (3.5)$$

Here f is any function in $H^1_{0,\Gamma_D}(\Omega)$, and the constant $\zeta > 0$ is defined as

$$\zeta := \frac{\tau}{2} \epsilon \tilde{n} D_a L_r. \qquad (3.6)$$

Notice that if τ is reasonably small, $\zeta < 1$.

The lemma below shows that the iteration error defined in (3.4) is a contraction in the norm (3.5).

Lemma 6.3.2 *For $\tau < 2/(\epsilon \tilde{n} D_a L_r)$, an i–independent constant $0 < \gamma < 1$ exists such that*

$$|||e^{n,i}|||^2 \leq \gamma |||e^{n,i-1}|||^2,$$

provided $u^{n,i-1}$ satisfies the bounds in Lemma 6.3.1.

Proof. With ζ given above, we start by adding $\zeta(u^n, \phi)_{\Gamma_G}$ on both side of (2.2). Subtracting (3.3) from the resulting equation gives

$$\begin{aligned}\left(e^{n,i}, \phi\right)_\Omega + \tau D \left(\nabla e^{n,i}, \nabla \phi\right)_\Omega - \tau \left(q e^{n,i}, \nabla \phi\right)_\Omega + 2\zeta \left(e^{n,i}, \phi\right)_{\Gamma_G} \\ = 2\zeta \left(e^{n,i-1}, \phi\right)_{\Gamma_G} + \tau \epsilon \tilde{n} D_a \left(r(u^n) - r(u^{(i-1)}), \phi\right)_{\Gamma_G}.\end{aligned}$$

Since $\nabla \cdot q = 0$ and $e^{n,i}$ has a vanishing trace on Γ_D, taking $\phi = e^{n,i}$ into above yields

$$\begin{aligned}&\|e^{n,i}\|_\Omega^2 + \tau D \|\nabla e^{n,i}\|_\Omega^2 + 2\zeta \|e^{n,i}\|_{\Gamma_G}^2 \\ &\quad \leq \tau \epsilon \tilde{n} D_a \|L_r e^{n,i-1} - \left(r(u^n) - r(u^{(i-1)})\right) \|_{\Gamma_G} \|e^{n,i}\|_{\Gamma_G} \\ &\quad \leq 2\zeta \|e^{n,i-1}\|_{\Gamma_G} \|e^{n,i}\|_{\Gamma_G}.\end{aligned}$$

We then immediately get

$$|||e^{n,i}|||^2 \leq \zeta(1 - \alpha + \alpha) \|e^{n,i-1}\|_{\Gamma_G}^2, \tag{3.7}$$

where $||| \cdot |||$ is introduced in (3.5) and α is an arbitrary constant in $(0, 1)$ to be chosen below.

Using the trace estimate (2.16) and the inequality of means, for any $\beta > 0$ we have

$$\|e^{n,i-1}\|_{\Gamma_G}^2 \leq \left(C_1 + \frac{C_2^2}{4\beta}\right) \|e^{n,i-1}\|_\Omega^2 + \beta \|\nabla e^{n,i-1}\|_\Omega^2.$$

Using this into (3.7) and taking $\alpha \in (0, 1)$, $\beta > 0$, and $\gamma > 0$ satisfying the constraints

$$\begin{aligned}\zeta\alpha\left(C_1 + \tfrac{C_2^2}{4\beta}\right) &\leq \gamma, \\ \zeta\alpha\beta &\leq \tau D \gamma, \\ \zeta(1-\alpha) &\leq \gamma,\end{aligned} \tag{3.8}$$

we obtain

$$\begin{aligned}|||e^{n,i}|||^2 \leq\ & (1-\alpha)\zeta \|e^{n,i-1}\|_{\Gamma_G}^2 + \alpha\beta\zeta \|\nabla e^{n,i-1}\|_\Omega^2 \\ & + \alpha\zeta \left(C_1 + \tfrac{C_2^2}{4\beta}\right) \|e^{n,i-1}\|_\Omega^2 \leq \gamma |||e^{n,i-1}|||^2.\end{aligned} \tag{3.9}$$

With

$$\beta = \frac{\tau D}{2}\left(C_1 + \sqrt{C_1^2 + \frac{C_2^2}{\tau D}}\right), \quad \text{and} \quad \alpha = \left(1 + \frac{C_1}{2} + \frac{1}{2}\sqrt{C_1^2 + \frac{C_2^2}{\tau D}}\right)^{-1},$$

the restrictions on α and β are fulfilled. Further, this choice also gives identical lower bounds for γ in (3.8), for which we can now take

$$\gamma = \zeta\left(C_1 + \sqrt{C_1^2 + \frac{C_2^2}{\tau D}}\right)\left(2 + C_1 + \sqrt{C_1^2 + \frac{C_2^2}{\tau D}}\right)^{-1}. \tag{3.10}$$

By the assumptions on τ we have $\gamma < 1$, hence the iteration error is contractive. ■

Remark 25 *The iteration* (3.3) *defines an operator $\mathcal{T} : \mathcal{K} \to \mathcal{K}$. Following the steps in the proof of Lemma 6.3.2, we can show that $\mathcal{T}$ is a contraction with respect to the norm defined in* (3.5)*. Therefore $\mathcal{T}$ has a unique fixed point, yielding the existence and uniqueness of a solution for the nonlinear equation* (3.1)*. This immediately implies the existence and uniqueness of a solution for Problem P_δ^n.*

Lemma 6.3.2 implies the linear convergence in H^1 of the iteration sequence $\{u^{n,i}, i \geq 0\}$. Moreover, its limit is the solution u^n of (3.1).

Theorem 6.3.3 *With $u^{n,0} \in H^1_{0,\Gamma_D}(\Omega)$ bounded essentially by 0 and M_u, if $\tau < 2/(\epsilon\tilde{n}D_a L_r)$, the iteration (3.3) is convergent. Specifically, for all $i > 0$ we have*

$$|||e^{n,i}|||^2 \leq \gamma^i |||e^{n,0}|||^2.$$

Remark 26 *The contraction constant γ in the above is bounded from above by $\zeta = \tau\epsilon\tilde{n}D_a L_r/2$. Notice that the first term in the H^1–equivalent norm in* (3.5) *does not depend on τ. As $\tau \searrow 0$, γ approaches 0, implying a fast convergence of the iteration at least in L^2 sense. In the numerical computations presented in the following section, 3 to 4 iterations were enough for obtaining a good numerical approximation of the time discrete solution.*

As stated in Theorem 6.3.3, convergence is achieved as $i \to \infty$. In practice we stop this procedure after a finite (small) number of iterations. This means that at each time step we are adding an iteration error to the time discrete approximation. This error depends on the number of iterations performed per time step, as well as on the initial iteration error. As mentioned before, the solution at the previous time step can be used for initiating the iteration. In the remaining part of this section we show that by this choice the total error is vanishing as $\tau \searrow 0$.

To make this statement rigorous we assuming that i iterations are performed at each time step n. The computed solution $\tilde{u}^n = u^{n,i}$ will only be an approximation of u^n, the solution of (3.1). Let now $\tilde{e}^n$ denote the error at the time step n,

$$\tilde{e}^n := u^n - \tilde{u}^n.$$

Based on the stability estimates in Lemma 6.2.5, we can estimate the total error that is accumulated in the numerical approximation of the time discrete sequence $\{u^n, n = 1, \dots, N\}$.

Lemma 6.3.4 *Assume that, for each $n = 1, \dots, N$, i iterations* (3.3) *are performed by starting with $u^{n,0} = \tilde{u}^{n-1} = u^{n-1,i}$. We have*

$$\sum_{n=1}^{N} |||\tilde{e}^n||| \le C \frac{\gamma^{i/2}}{\tau^{1/4}}.$$

Remark 27 *Since $\gamma = O(\tau)$, the total error vanishes as $\tau \searrow 0$.*

Proof. With $u^{n,0} = \tilde{u}^{n-1}$, the initial error at the time step n is given by

$$e^{n,0} = u^n - \tilde{u}^{n-1} = u^n - u^{n-1} + \tilde{e}^{n-1}.$$

By Theorem 6.3.3 the error at the time step n can be estimated as

$$|||\tilde{e}^n||| \le \gamma^{i/2} |||e^{n,0}||| \le \gamma^{i/2} \left(|||u^n - u^{n-1}||| + |||\tilde{e}^{n-1}||| \right).$$

Repeating this estimates inductively for $n = 1, \dots, N$, since $\tilde{e}^0 = 0$ we obtain

$$|||\tilde{e}^n||| \le \sum_{k=1}^{n} \gamma^{(n+1-k)i/2} |||u^k - u^{k-1}|||. \tag{3.11}$$

Adding the above for $n = 1$ up to N gives

$$\sum_{n=1}^{N} |||\tilde{e}^n||| \le \frac{\gamma^{i/2}}{1-\gamma} \sum_{k=1}^{N} \left(1 - \gamma^{(N+1-n)i/2} \right) |||u^n - u^{n-1}|||.$$

The proof can be completed straightforwardly by applying the stability estimates in Lemma 6.2.5. ■

6.4 Numerical results

In this section we present some numerical simulations obtained for the undersaturated regime. Extensive numerical results for both dissolution and precipitation, and for high or low Damköhler numbers, will be presented in a forthcoming paper. The present computations are carried out in a reference cell Ω, where the d–dimensional hypercube $(-L, L)^d$ ($d = 2$, or 3, and $L = 1$) is including a grain represented as a d–dimensional ball of radius $R = 0.5$ centered in the origin. Then Γ_G represents the surface of this ball. Similar computations are presented in [2]. In the simplified geometry considered there, a strip, the occurrence of a dissolution front can be proven rigorously. After a waiting time t^*, this front moves in the flow direction. The present computations are revealing similar features.

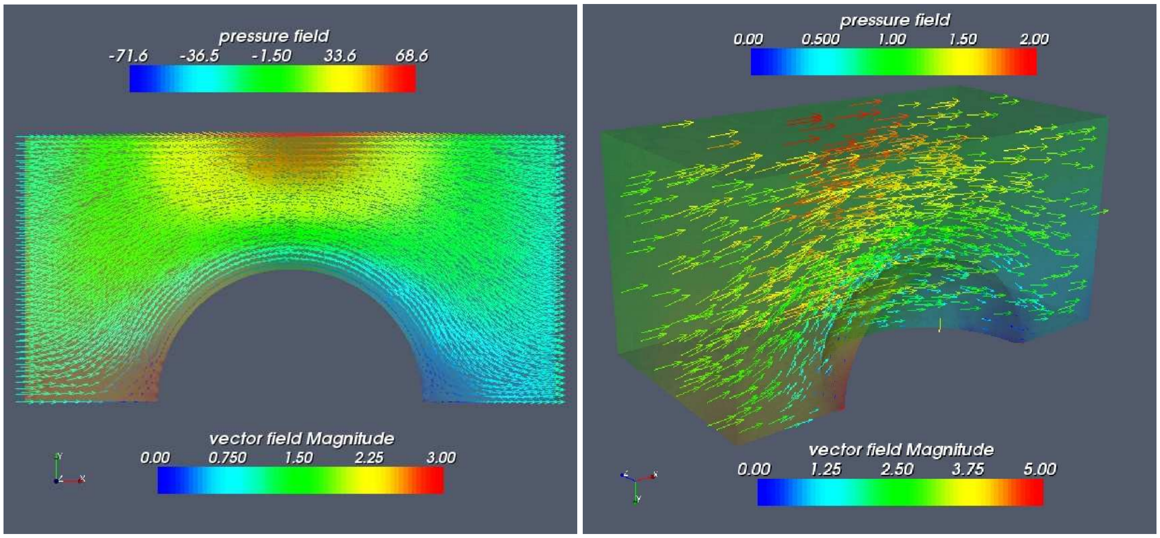


FIG. 6.1 – Velocity and pressure field

We have used the following parameters and rate function :

$$D = 0.25,\ 0.5 \text{ (d=2 or 3 resp.)}$$

$$D_a = 1, 5, 10\ \varepsilon = 1,\ \tilde{m} = \tilde{n} = 1.0,\ \text{and } r(u,c) = \frac{10}{9}[u]_+[u-0.1]_+.$$

The initial and (external) boundary conditions are

$$\begin{aligned} v_I &= v_0 > 0, && \text{on } \Gamma_G, \\ u_I &= u^*, && \text{in } \Omega, \\ u &= u_*, && \text{if } x = -1, t > 0, \\ \partial_\nu u &= 0, && \text{on } \partial\Omega \backslash \{\Gamma_G \cup \{x = -1\}\}. \end{aligned}$$

In the above we have taken $u_* = 0.1$, and $u^* = 1.0$. For the two dimensional case we took $v_0 = 0.01$, while $v_0 = 0.1$ in three spatial dimensions. By the maximum principle, only dissolution is possible. This follows by the L^∞ estimates.

The fluid velocity $\vec{q}$ is solving the Stokes model in Ω,

$$\begin{cases} \mu \triangle \vec{q} &= \nabla p, & \text{in } \Omega, \\ \nabla \cdot \vec{q} &= 0, & \text{in } \Omega, \\ \vec{q} &= \vec{0}, & \text{on } \Gamma_G, \\ \vec{q} &= \vec{q}_D, & \text{on } \Gamma_D, \end{cases}$$

where $\mu = 0.01$ and $\vec{q_D} = (2,0),\ (3,0)$ for $d = 2$ or 3 respectively.

The numerical approximation of $\vec{q}$ is obtained by employing the bubble stabilized finite element method proposed in [18] (see also [11]). Figure 6.1 displays the numerical approximation of $\vec{q}$ in two and three dimensions.

For symmetry reasons, if $d = 2$ the computations are restricted to the upper half of the domain. Similarly, we consider only the upper left part of Ω if $d = 3$. In doing so we impose homogeneous Neumann conditions on the newly occurring symmetry boundaries. The computations are up to $T = 0.3, 0.5$, for $d = 2$ or 3 respectively when the entire crystal has been dissolved. In the time discrete scheme (2.2)–(2.3) we took a fixed time step $\tau = 10^{-4}$. Here $\delta = 10^{-4}$. As mentioned in the appendix of [1] (see also Theorem 6.2.7), we set $w = \min\{r(u), 1\}$ whenever $v = 0$.

The nonlinear problems are solved by the linearization given in (3.3), with $u^{n,0} = u^{n-1}$. The procedure is stopped once the maximal difference between two successive iterations is reduced below 10^{-3}, and 3 to 4 iterates were sufficient. We use piecewise linear finite elements for the spatial discretization of the time discrete problems. All the computations are implemented in the research software *SciFEM* (Scilab Finite Element Method)[1].

In Figure 6.2, 6.3 and 6.4 we show the cation concentration u for $D_a = 1, 5, 10$.

Figure 6.5 and 6.6 display the precipitate concentration v.

[1]Finite Element Library used with scilab, developped by T. Clopeau and V. Devigne for research in Scientific Computing

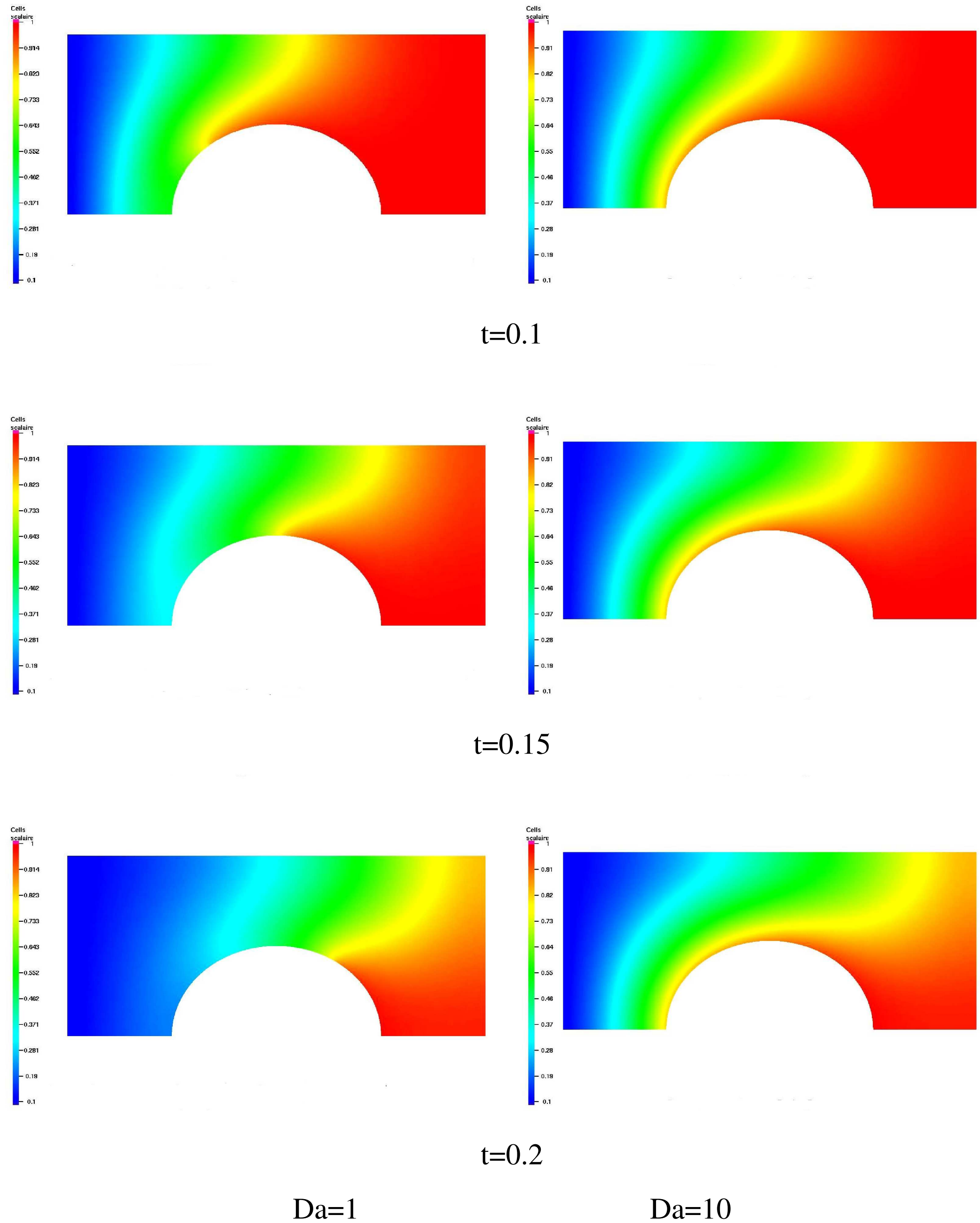


FIG. 6.2 – Concentration of cations

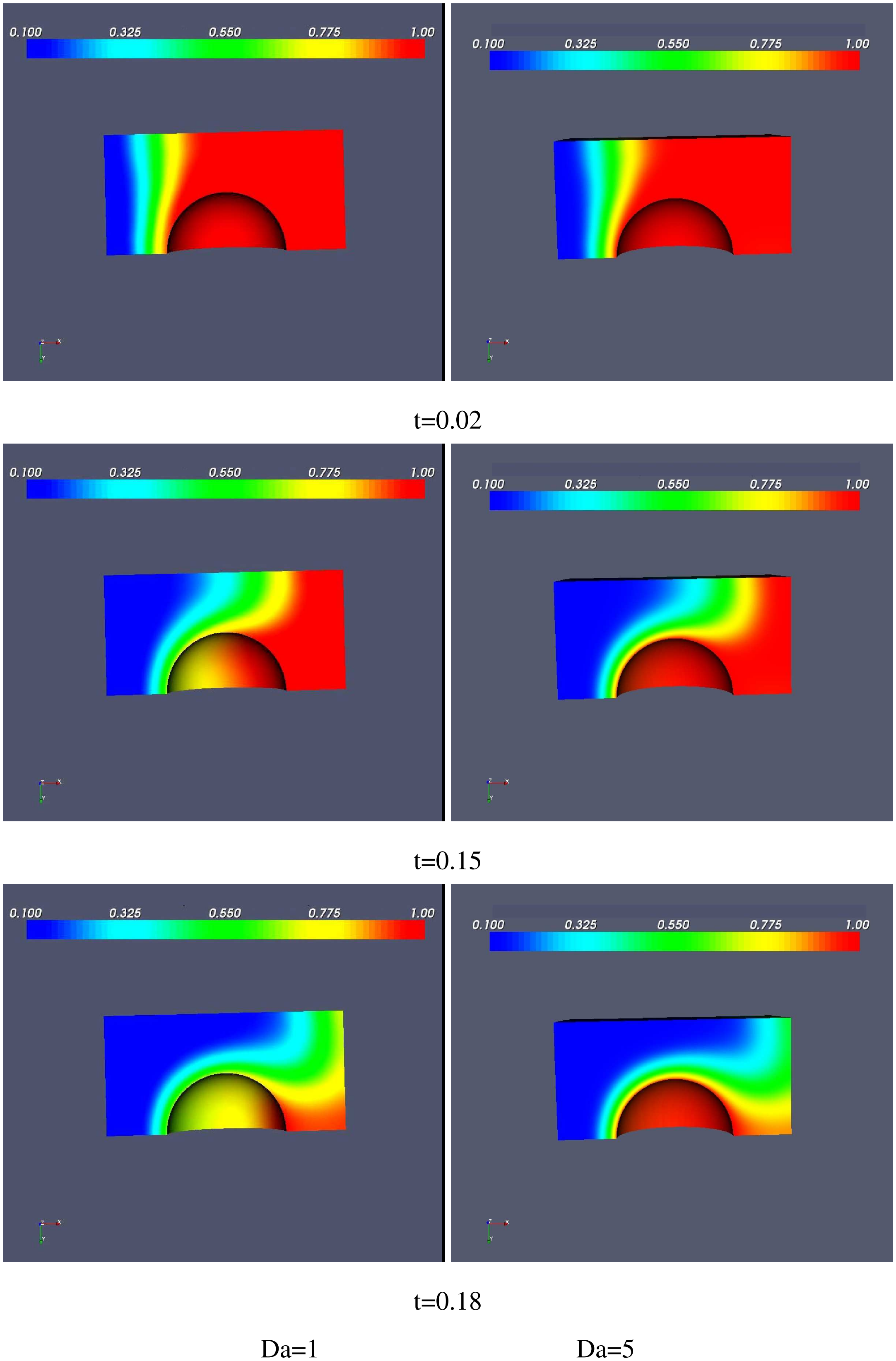
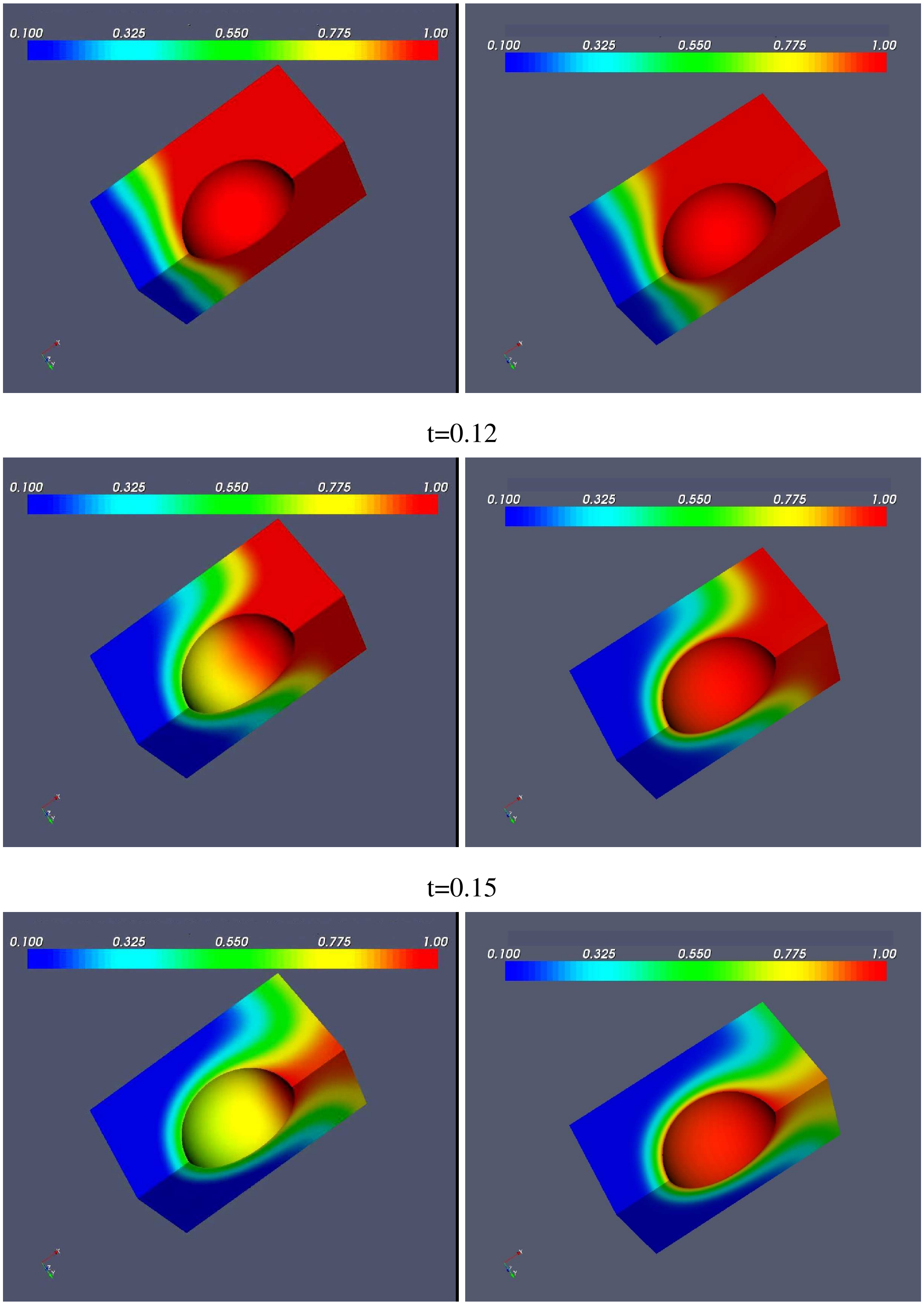


FIG. 6.3 – Concentration of cations, lateral view



t=0.12

t=0.15

t=0.18

Da=1 Da=5

FIG. 6.4 – Concentration of cations, below view

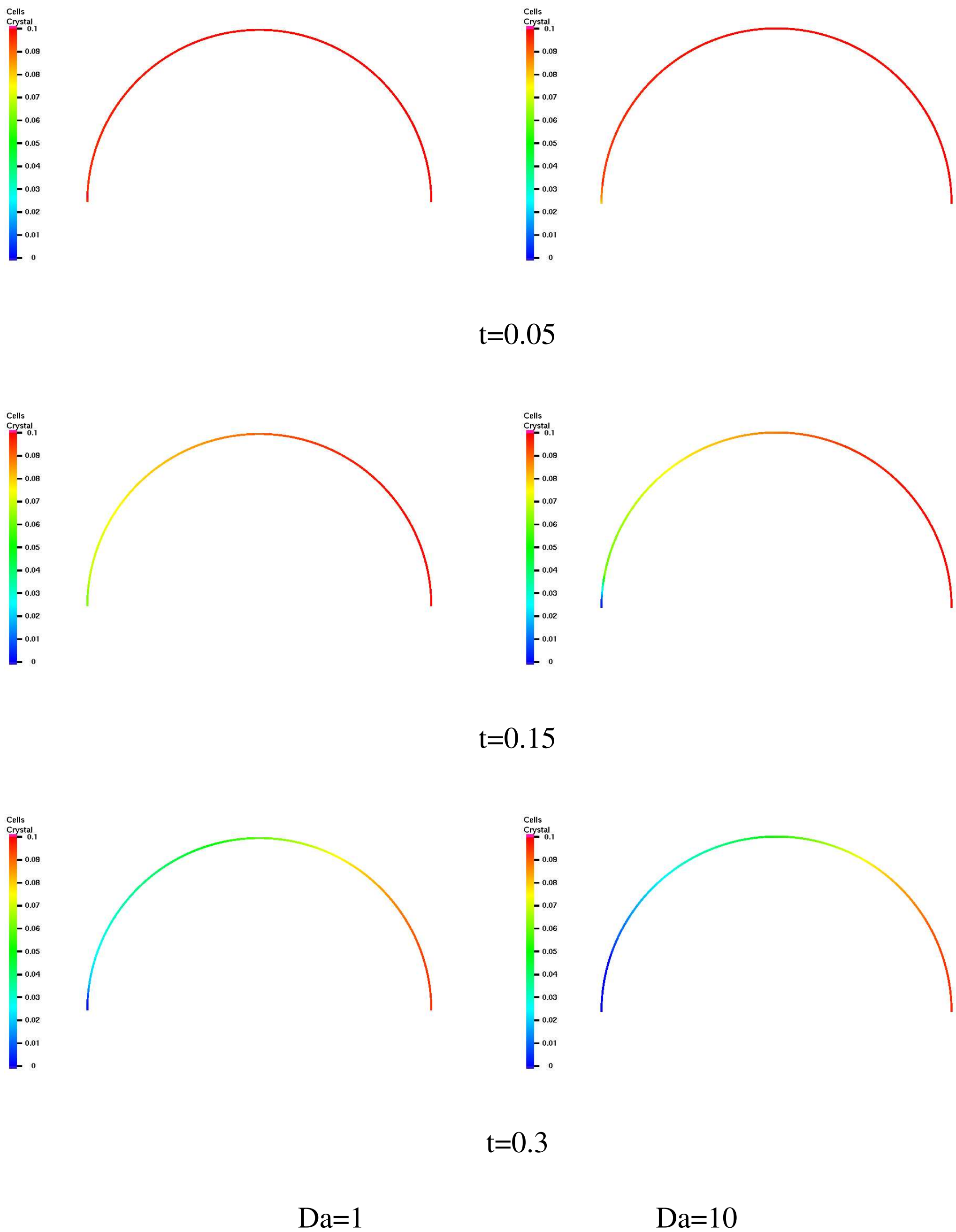


FIG. 6.5 – Concentration of crystal

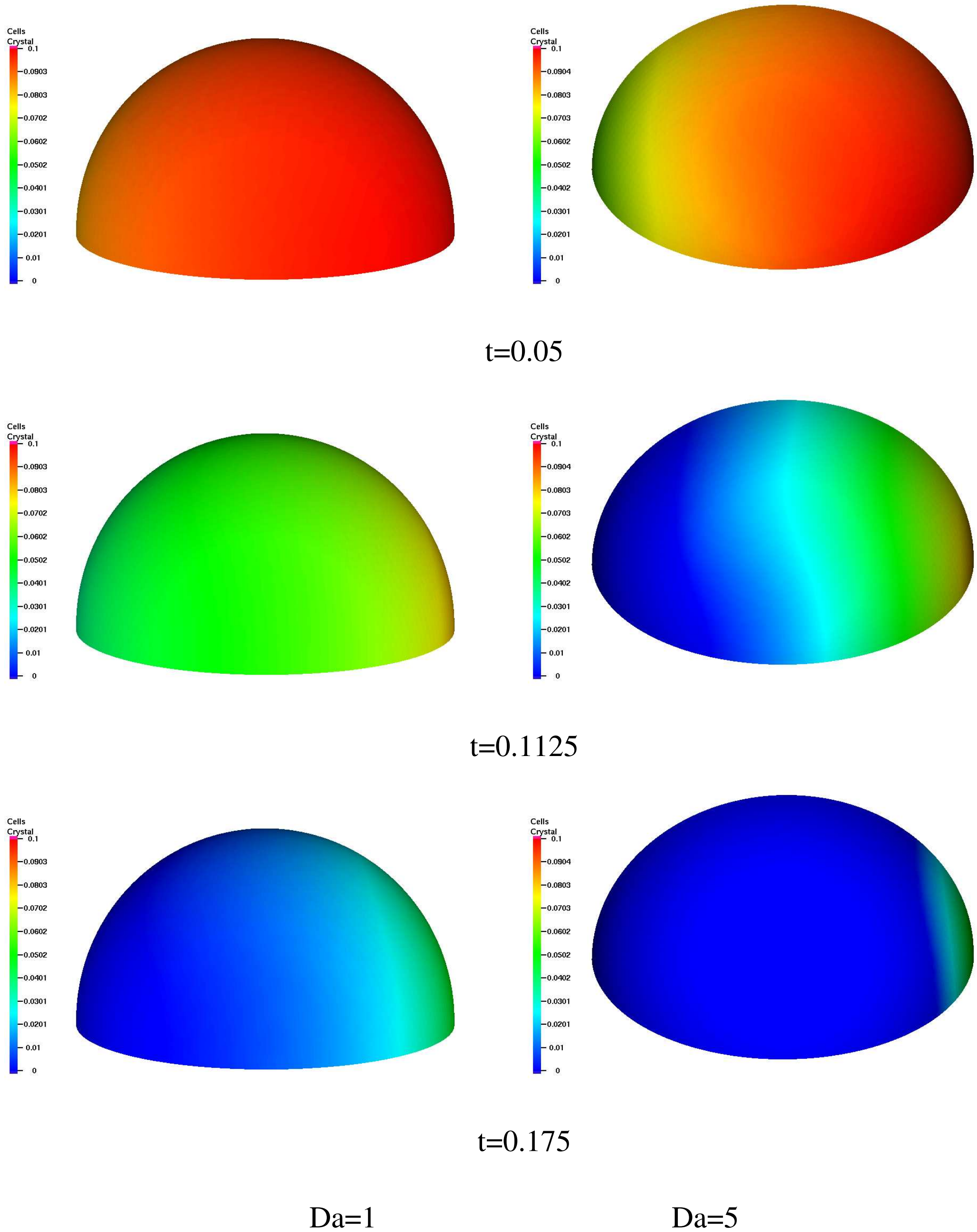


FIG. 6.6 – Concentration of crystal D=0.5

6.5 Conclusion

We have analyzed a numerical scheme for the time discretization of (1.1)–(1.2). This is a simplification of the pore scale model for crystal dissolution and precipitation in porous media. The main difficulties associated with the model, a parabolic problem that is coupled through the boundary to a differential inclusion, are still present in this setting.

The numerical scheme is implicit in u and explicit in v. A regularization step is employed for dealing with the multivalued dissolution rate. We prove the stability of the scheme in both L^∞, as well as energy norms. Further, compactness arguments are used for showing that the scheme is convergent.

A fixed point iteration is proposed for solving the nonlinear time discrete problems. This linearization is stable and converges linearly, regardless of the discretization parameters and the starting point.

Acknowledgment The work of C.J. van Duijn and I.S. Pop was supported by the Dutch government through the national program BSIK : knowledge and research capacity, in the ICT project BRICKS (http ://www.bsik-bricks.nl), theme MSV1. This research was initiated while V. Devigne spent six months at the Technische Universiteit Eindhoven, supported through an European Community Marie Curie Fellowship (contract number HPMT-CT-2001-00422).

Nous venons de présenter l'analyse d'un schéma numérique pour une discrétisation en temps. Il s'agit d'une simplification du modèle à l'échelle du pore pour la dissolution/précipitation en milieu poreux. La plus grosse difficulté associée au modèle, un problème parabolique couplé à la frontière par une inclusion différentielle, a été traitée sans dénaturer le modèle en respectant sa complexité.

Le schéma numérique est implicite en u et explicite en v. Une étape de régularisation est utilisée pour le taux de dissolution multivoque. Nous avons prouvé la stabilité du schéma dans les normes du maximum et de l'énergie. Des arguments de compacité ont montré la convergence. Une itération de type point fixe est proposée pour résoudre le problème discrétisé en temps non linéaire. Cette discrétisation est stable et convergente, au regard des paramètres de discrétisation et du point de départ.

Bibliographie

[1] C. J. van Duijn, P. Knabner, Travelling wave behaviour of crystal dissolution in porous media flow, European J. Appl. Math. **8** (1997), 49–72.

[2] C. J. van Duijn, I. S. Pop, Crystal dissolution and precipitation in porous media : pore scale analysis, J. Reine Angew. Math. **577** (2004), 171–211.

[3] R. Eymard, T. Gallouët, R. Herbin, D. Hilhorst, M. Mainguy, Instantaneous and noninstantaneous dissolution : approximation by the finite volume method, in Actes du 30ème Congrès d'Analyse Numérique : CANum '98 (Arles, 1998), R. Boyer et al. (Eds.), ESAIM Proc. **6**, Soc. Math. Appl. Indust., Paris (1999), 41–55.

[4] B. Faugeras, J. Pousin, F. Fontvieille, An efficient numerical scheme for precise time integration of a diffusion-dissolution/precipitation chemical system, Math. Comp. **75** (2006), 209–222.

[5] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and Studies in Mathematics **24**, Pitman (Advanced Publishing Program), Boston, MA (1985).

[6] U. Hornung, W. Jäger, A model for chemical reactions in porous media, in J. Warnatz, W. Jäger (Eds.), *Complex Chemical Reaction Systems. Mathematical Modeling and Simulation*, Chemical Physics **47**, Springer, Berlin, (1987).

[7] U. Hornung, W. Jäger, Diffusion, convection, adsorption, and reaction of chemicals in porous media, J. Differential Equations **92** (2001), 199–225.

[8] U. Hornung, W. Jäger, A. Mikelić, Reactive transport through an array of cells with semi-permeable membranes, RAIRO Modél. Math. Anal. Numér., **28** (1994), 59–94.

[9] R. B. Kellogg, J. E. Osborn, A regularity result for the Stokes problem in a convex polygon, J. Functional Analysis **21** (1976), 397–431.

[10] P. Knabner, C. J. van Duijn, S. Hengst, An analysis of crystal dissolution fronts in flows through porous media. Part 1 : Compatible boundary conditions, Adv. Water Res. **18** (1995), 171–185.

[11] P. Knobloch, On the application of the $p\,\mathrm{sp}\,\mathrm{mod}_1$ element to incompressible flow problems, Comput. Vis. Sci. **6** (2004), 185–195.

[12] J. R. Kweona, R. B. Kellogg, Compressible Stokes problem on nonconvex polygonal domains, J. Differential Equations **176** (2001), 290–314.

[13] E. Maisse, J. Pousin, Diffusion and dissolution/precipitation in an open porous reactive medium, J. Comput. Appl. Math. **82** (1997), 279–290.

[14] E. Maisse, J. Pousin, Finite element approximation of mass transfer in a porous medium with non equilibrium phase change, J. Numer. Math. **12** (2004), 207–231.

[15] A. Mikelić, V. Devigne, C. J. van Duijn, Rigorous upscaling of the reactive flow through a pore, under dominant Peclet and Damkohler numbers, CASA Report 05-19, Eindhoven University of Technology (2005).

[16] I. S. Pop, F. Radu, P. Knabner, Mixed finite elements for the Richards' equation : linearization procedure, J. Comput. Appl. Math. **168** (2004) 365–373.

[17] I. S. Pop, W. A. Yong, On the existence and uniqueness of a solution for an elliptic problem, Studia Univ. Babeş-Bolyai Math. **45** (2000), 97–107.

[18] A. Russo, Bubble stabilization of finite element methods for the linearized incompressible Navier-Stokes equations, Comput. Methods appl. Mech. Engrg. **132** (1996), 335–343.

[19] J. Simon, Compact sets in the space $L^p(0,T;B)$, Ann. Mat. Pura Appl.(4) **146** (1987), 65 - 96.

[20] M. Slodička, A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media, SIAM J. Sci. Comput. **23** (2002), 1593–1614.

[21] M. Slodička, A robust linearisation scheme for a nonlinear elliptic boundary value problem : error estimates, ANZIAM J. **46** (2005), 449–470.

[22] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer Verlag, New York, 1983.

[23] J. Wloka, *Partielle Differentialgleichungen*, B. G. Teubner, Stuttgart (1982).

[24] W. A. Yong, I. S. Pop, A numerical approach to porous medium equations, Preprint **96-50** (SFB 359), IWR, University of Heidelberg (1996).

Chapitre 7

Questions ouvertes et travaux futurs - Open questions and future works

Nous espérons que le lecteur a pu trouver ici une étude claire et aussi complète que possible des écoulements en milieu poreux par le problème de Darcy, l'analyse numérique du problème de Stokes pour l'écoulement incompressible de fluide visqueux et un début de réponse concernant le problème posé par les frontières rugueuses avec la loi de Navier, les balbutiements dans la compréhension des processus chimiques dans un milieu poreux. Nous avons essayé de garder un équilibre entre les aspects techniques et théoriques dans un souci pédagogique.

La question des approximations des équations de Darcy et Stokes peut être considérée comme close, les éléments finis présentés sont spécifiques à chacun des problèmes. La méthode non-conforme peut être étendue à des quadrangles ou héxaèdres réguliers la rendant accessible à des domaines plus généraux. Mais l'interrogation ici est d'ordre technique et n'implique pas de difficultés conceptuelles ou de modélisation. Plus ouvertes restent les questions posées par les lois de paroi et la chimie.

Dans tous les aspects du problème, la technique d'homogénéisation conduit à des lois sous la forme de conditions aux limites de type mixte faisant intervenir le paramètre caractéristique ϵ (la taille des pores, ou des rugosités) et le tenseur de Beavers et Joseph ou de Navier. Les composantes tangentielle et normale de la vitesse sont liées par une condition aux limites de type Robin ou Fourier. Ces matrices restent inconnues et dépendent des caractéristiques du matériau. Il reste ici, un problème d'ingénierie difficile qui pourrait trouver un écho favorable dans le monde industriel. Nous sommes persuadés que l'homogénéisation apporte par le calcul des alternatives aux outils statistiques.

La chimie des écoulements en milieu poreux a été abordée de deux façons différentes. Dans la première, la chimie et la géométrie sont volontairement simplifiées de façon à obtenir un problème modèle et donner les bases de l'application de la théorie de l'homogénéisation en gardant à l'esprit l'objectif principal : une justification de la dispersion de Taylor dans les tubes capillaires. Dans la seconde, un grand soin a été porté à la description du milieu au niveau microscopique et à la chimie. Cette complexité n'a pas permis de proposer un modèle à plus grande échelle à cause de difficultés techniques. Cependant cette deuxième étude a laissé le champ libre à de nombreuses simulations beaucoup plus faciles à fournir que dans l'approche de Taylor à cause de la raideur du problème physique bidimensionnel. La théorie d'homogénéisation développée pour confirmer la dispersion de Taylor est prometteuse et d'autres travaux futurs sur une description plus complète de la chimie ainsi qu'une géométrie plus complexe pourront être envisagés.

We expect the reader to find here a clear and an "as complete as possible" study about flows in a porous medium, with the Darcy's problem, about incompressible free fluid flow with the numerical analysis of the Stokes problem and a part of the answer concerning the problem at a rough interface with the study of Navier's law, the first-fruits of understanding of chemical processes in a porous medium. We tried to keep an equilibrium between a numerical and an analytical point of view in a pedagogic manner.

Concerning the approximations of Darcy and Stokes equations, the finite element presented are specific to each problem, and the question is closed in that sense. The non-conforming method could be extended to non regular quadrangles or hexaedras for more general domains. But the question here is only technical and doesn't bring conceptual neither modeling difficulty. The questions related to the wall laws and the chemistry are more opened.

In every situation, the results of the homogenization technique lead to a law in the form of a mixed boundary condition involving the characteristic parameter of the interface ϵ (the pore size) and a tensor (Beavers and Joseph, Navier's matrices). The tensor links the normal component of the velocity to the tangential ones. The tensor is unknown and depend on the characteristic of the material. A difficult engineer type problem still remains and is of interest for the industrial world. We are confident that homogenization theory may bring alternatives to statistic tools by the calculus.

The question of chemistry for flows in a porous medium was considered from both angle starting from two different descriptions of the porous medium but governed by two different approaches. In the first one, the chemistry as well as the geometry is consciously oversimplified to provide a toy problem for the first steps of homogenization theory applications keeping in mind a goal : a justification of Taylor's dispersion in a capillary tube. The second one is attached to a complex description of the chemical reaction keeping the balance between the dissolution and the precipitation applied to a closed description of the domain. This is its drawback, technical difficulties don't allow us to give an upscaled problem for complex geometries and chemistry. Nevertheless it gave the free hands to several simulations not so easy to provide for the Taylor's approach due to the rough character of the 2D-problem. The homogenization theory developed for the Taylor's dispersion is promising and further work will focused on a more complex description of the chemistry and a more complex geometry.

résumé : En sciences de l'environnement et plus particulièrement en hydrogéologie les problèmes de nature phénoménologique nombreux conduisent bien souvent à l'étude des Équations aux Dérivées Partielles (EDP's) au travers des non moins nombreux modèles qui en découlent.

Si chaque phénomène physique, mécanique, chimique ou autres pris indépendamment et à une échelle suffisamment fine est aujourd'hui bien compris et relativement aisé à modéliser il n'en est pas de même pour les problèmes multiphysiques, physico-chimique, les écoulements au voisinage de domaines de structures différentes ou même dans l'appréhension de ces phénomènes à des échelles plus grandes méso et macroscopique.

La compréhension des conditions aux limites et leur modélisation reste une étape clef dans l'étude de ces phénomènes naturels.

Nous verrons à travers différents problèmes comment il est possible de résoudre numériquement et en partie ces difficultés par des techniques d'analyse mathématique adéquates.

Des résultats de simulations réalisées au moyen d'un logiciel de résolution d'EDP's baptisé SciFEM (Scilab Finite Element Method) conçu dans le cadre de la thèse illustreront notre démarche.

Titre en anglais : Flows and Particular Boundary Conditions applied in Hydrogeology and Mathematical Theory of Dissolution/Precipitation process in porous media

abstract : In environmental sciences and more specifically in hydrogeology, most of phenomenological problems lead the scientist in front of the study of Partial Differential Equations (PDE's) through the various different models they are coming from.

A natural phenomenon may be studied from different angles. It depends on its main type which could be physical, mechanical, chemical ... Considered independently and under assumption of sufficient fine scale of observation, this phenomenon is reasonably well understood and modeled. This is not the case for multi-physics, coupled physics and chemistry problems, flow problems closed to interfaces of domains with different structure where the phenomenon itself is not clearly handled. What 's happening if the same microscopic behavior is considered at a larger (meso or macrocscopic) scale ?

A good understanding of the boundary conditions is required as well as their modeling which is the "key" in the study of natural phenomenon.

We will see with different problems how it is possible to solve numerically and partially

those difficulties with techniques based on recent mathematical analysis.

Results of simulations realized with a PDE's software called SciFEM (for Scilab Finite Element Method) created during the thesis will emphasize our discussion.

Mots-clefs : homogénéisation, éléments finis non-conforme, écoulements, Darcy, Stokes, fonction bulle, loi de Navier, dissolution/ précipitation, milieu poreux, dispersion

Discipline : Mathématiques et applications des mathématiques

Intitulé et adresse du laboratoire :

Institut Camille Jordan - UMR 5208 -

Université Claude Bernard Lyon1.

43 bd du 11 novembre 1918 - 69622 Villeurbanne.