



HAL
open science

Shape optimization and applications to hydraulic structures: mathematical analysis and numerical approximation

Mostafa Kadiri

► **To cite this version:**

Mostafa Kadiri. Shape optimization and applications to hydraulic structures: mathematical analysis and numerical approximation. Numerical Analysis [math.NA]. Normandie Université, 2019. English. NNT: 2019NORMC214 . tel-02435595

HAL Id: tel-02435595

<https://theses.hal.science/tel-02435595v1>

Submitted on 11 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité MATHÉMATIQUES

Préparée au sein de l'Université de Caen Normandie

**Shape optimization and applications to hydraulic structures
Mathematical analysis and numerical approximation**

**Présentée et soutenue par
Mostafa KADIRI**

**Thèse soutenue publiquement le 10/07/2019
devant le jury composé de**

M. PHILIPPE HELLUY	Professeur des universités, Université de Strasbourg	Rapporteur du jury
M. FRANCOIS JOUVE	Professeur des universités, Université Paris 7 Paris Diderot	Rapporteur du jury
Mme NINA AGUILLON	Maître de conférences, Université Paris 5 Descartes	Membre du jury
Mme NAÏMA DEBIT	Maître de conférences, Université Lyon 1 Claude Bernard	Membre du jury
M. JALAL FADILI	Professeur des universités, ENSICAEN	Membre du jury
M. NICOLAS FORCADEL	Professeur des universités, INSA de Rouen Normandie	Président du jury
M. HOUARI MECHKOUR	Enseignant chercheur, ECE Paris	Membre du jury
M. MOHAMMED LOUAKED	Maître de conférences HDR, Université Caen Normandie	Directeur de thèse

Thèse dirigée par MOHAMMED LOUAKED, Laboratoire de Mathématiques 'Nicolas Oresme' (Caen)



UNIVERSITÉ
CAEN
NORMANDIE





Normandie Université

THESE

Pour obtenir le diplôme de doctorat

Spécialité : Mathématiques appliquées.

Préparée au sein de l'Université de Caen Normandie

Shape optimization and applications to hydraulic structures
mathematical analysis and numerical approximation

**Présentée et soutenue par
Mostafa KADIRI**

**Thèse soutenue publiquement le 10/07/2019
devant le jury composé de**

M. PHILIPPE HELLUY	Professeur des universités, Université de Strasbourg	Rapporteur du jury
M. FRANCOIS JOUVE	Professeur des universités, Université Paris 7 Paris Diderot	Rapporteur du jury
Mme. NINA AGUILLON	Maître de conférences, Université de Paris 5 Descartes	Membre de Jury
Mme. NAÏMA DEBIT	Maître de conférences, Université Lyon 1 Claude Bernard	Membre de Jury
M. JALAL FADILI	Professeur des universités, ENSICAEN	Membre de Jury
M. NICOLAS FORCADEL	Professeur des universités, INSA de Rouen Normandie	Membre de Jury
M. HOUARI MECHKOUR	Enseignant chercheur, ECE PARIS	Membre de Jury
M. MOHAMMED LOUAKED	Maître de conférences HDR, Université de Caen Normandie	Directeur de thèse

Thèse dirigée par Mohammed LOUAKED, Laboratoire de Mathématiques Nicolas Oresme (Caen)



UNIVERSITÉ
CAEN
NORMANDIE



À l'âme de ma Mère

Remerciements

D'abord, je voudrais remercier mon directeur de thèse, Mr Mohammed LOUAKED, pour m'avoir accueilli et m'avoir donné l'occasion d'intégrer l'équipe modélisation et applications. Je lui suis également reconnaissant pour le temps conséquent qu'il m'a accordé, ses qualités pédagogiques et scientifiques, sa franchise et sa sympathie. J'ai beaucoup appris à ses côtés, en particulier sur le plan théorique et numérique et je lui adresse ma gratitude pour le temps consacré avec ses conseils, son implication et les corrections apportées à cette thèse.

J'adresse mes remerciements les plus chaleureux à mon co-encadrant de thèse, Mr Houari MECHKOUR, pour son attention à mes travaux, et pour ses conseils avisés qui ont été prépondérants pour la bonne réussite de cette thèse. J'ai pris un très grand plaisir à travailler avec lui.

Je tiens à remercier les rapporteurs de cette thèse Mr François JOUVE et Mr Philippe HELLUY, pour l'intérêt qu'ils ont apporté à mon travail. Je remercie également Mme Aguillon Nina, Mme Debit Naima, Mr Fadili Jalal, Mr Forcadel Nicolas d'avoir accepté de faire partie de mon jury.

Je remercie les responsables de l'École Centrale d'Électronique de Paris (ECE) pour le financement de cette thèse.

Je désire grandement remercier le directeur de L'ECE Mr Christophe Baujault et l'ancienne directrice Mme Lamia Rouai. Je souhaiterais remercier l'ancienne directrice de recherche de l'ECE Mme Valeria Nuzzo et la nouvelle directrice de recherche Mme Assia SOUKANE. J'associe à ces remerciements Mme Fabienne Coudray.

Mes sincères remerciements au Directeur de Laboratoire de recherche LMNO à Caen Mr Eric Ricard.

Un grand merci à tous les collègues de l'équipe recherche de l'ECE : François, Waleed, Sebti, Rafik, Quentin, Aakash, Manolo, Filippo, Thomas, Jae Yun, Yves, Ben et Serena pour leur sympathie. Tous les moments que j'ai passé à vos côtés étaient inoubliables.

Je désire en outre remercier les membres de l'équipe de recherche LMNO et les membres de l'équipe Modélisation et applications. Toutes mes reconnaissances à Mr Christian DOGBE qui m'a donné la chance d'intégrer la promotion de licence de mathématiques et qui était le point de départ de mon parcours jusqu'à la thèse.

Je voudrais aussi remercier l'équipe administrative Axelle, Anita, Marie et Sonia pour leurs disponibilités et leurs sympathies.

Un merci à mes collègues doctorants et postdoctorants de l'université de Caen Normandie : Nacer, Arnaud, Comba, Etienne, Tuyen, Tin, Angelot, Julien, Vlère, Leonard, Hong, Stavroula, Dalal, Guillaume, Ruben, Thuen et Gilberto. Les moments que j'ai passé à vos côtés m'ont appris beaucoup de choses.

Je tiens à remercier tous les membres de ma famille pour leur soutien quotidien durant toute la période de mes études : mon père Mohammed, mes frères Lahcen et Aziz et mes sœurs Mounia et Amal. J'associe à ces remerciements mes cousines et mes cousins Ali, Youssef, Khalid, Toufik, Soufiane, Said, Mounir, Alae, et Houcine pour leurs conseils et leur aide morale pendant mon séjour en France. Un grand merci à mes oncles paternels Ahmed et Moha ben Aissa et je n'oublie pas mon oncle maternel Mohammed. Une petite pensée à mes neveux et mes nièces et en particulier le nouveau né Imran.

Je voudrais remercier Mr Mohammed Idirssi pour ses conseils et sa sympathie, ses invitations quotidiennes à la maison durant mon séjour à Paris. Je remercie Rabeh, Noui, Ayman, Mazza, et la collocation avec Mahmoud était agréable pendant les années passées à Paris.

Introduction

Un fluide est un milieu matériel parfaitement déformable sous l'action d'une force extérieure. En général on a deux types de fluides, fluides compressibles tels que les gaz et plasma et des fluides incompressibles tels que les liquides. La trajectoire d'une particule décrit le mouvement selon l'écoulement du fluide laminaire ou turbulent. Le premier type d'écoulement est caractérisé par des lignes de courant parallèles qui se traduisent par un faible nombre de Reynolds contrairement au deuxième type d'écoulement où le nombre de Reynolds est élevé et qui présente en tout point un caractère tourbillonnaire. Dans le cadre de cette thèse, nous nous intéressons à l'étude théorique et numérique de ces deux types de modèles d'écoulement, en particulier les écoulements dans des milieux poreux.

Pour faire face à chaque type d'écoulement et sa mise en œuvre, nous nous intéressons à l'optimisation des structures hydrauliques telles que les passes à poissons. Ces structures jouent un rôle d'un pont de passage entre le milieu marin et les rivières favorisant et facilitant la migration quotidienne ou annuelle des espèces aquatiques des eaux salées vers les eaux douces pour la reproduction et pour accomplir leurs cycles biologiques. A titre d'exemple largement connu, le saumon qui est capable de nager des centaines de kilomètres en amont. Quand une barrière artificielle est construite dans un ruisseau, des règlements juridiques européens obligent l'installation d'une passe migratoire afin de permettre aux espèces de surmonter la barrière.

Les passes à poissons sont des structures hydrauliques placées à côté des barrages permettant aux poissons de traverser vers l'amont du fleuve. La plupart de ces structures sont construites en pente divisées en plusieurs bassins. Dans la littérature, de nombreux types de passes à poissons existent, celles de Denil [2], ou bien à fentes verticales VSF (Vertical Slot Fishway) [9]. Ce dernier type fait l'objet de notre étude. Du point de vue physique, la vitesse de l'écoulement du fluide dans ces structures est relativement faible permettant des pauses de repos. De plus ces structures permettent une bonne dissipation d'énergie. Ces caractéristiques nous ont conduit à privilégier et optimiser le choix de ce type de passe à poissons.

Un problème d'optimisation de forme consiste à trouver une forme dite optimale dans un ensemble de domaines admissibles qui minimise une fonction objective reflétant les conditions de confort pour les poissons. La recherche liée à l'optimisation de forme fait partie du domaine de contrôle optimal gouverné par des équations aux dérivées partielles. Le problème de contrôle intervient dans plusieurs domaines d'applications tels que l'ingénierie, la construction navale, aéronautique, génie civil et autres. Au fil de ces dernières années, cet horizon de recherche attire de plus en plus l'attention la communauté mathématiques

appliquées et ingénierie [8, 49, 59, 64, 60, 58].

Pour répondre à notre problématique, nous explorons différents types de modèles d'écoulements. Dans un premier lieu, nous introduisons un système hyperbolique des eaux peu profondes. Ce modèle dérivant des équations de Navier-Stokes incompressible en utilisant la méthode d'approximation de Boussinesq où les variations de densité sont négligeables sauf l'effet gravitationnel. Ce modèle est largement utilisé dans le cas de la modélisation des phénomènes géophysiques à surface libre pour simuler de nombreux phénomènes naturels en raison de son efficacité.

Le but de l'étude est de trouver une structure optimale de la structure hydraulique permettant aux poissons de franchir l'obstacle dans des conditions convenables. Nous donnons une formulation mathématique du canal composé de dix bassins avec des fentes verticales afin d'obtenir un débit moyen adapté à une grande variété d'espèces migratrices. Les équations bidimensionnelles en eaux peu profondes avec terme source s'écrivent

$$\begin{aligned} \frac{\partial H}{\partial t} + \vec{\nabla} \cdot \vec{Q} &= 0 & \text{dans } \Omega \times (0, T) \\ \frac{\partial \vec{Q}}{\partial t} + \vec{\nabla} \cdot \left(\vec{Q} \otimes \frac{\vec{Q}}{H} \right) + gH\vec{\nabla}(H - \eta) &= \vec{f} & \text{dans } \Omega \times (0, T) \end{aligned} \quad (0.0.1)$$

avec H est la hauteur de l'eau; $\vec{u} = (u_1, u_2)$ est le vecteur de vitesse; u_1 et u_2 sont les composantes x et y de la vitesse d'écoulement, respectivement; $Q = (u_1 H, u_2 H)$ est le débit d'eau; η est la géométrie du fond; g est l'accélération de gravité et \vec{f} représente tous les effets de frottements du fond et les forces de la pression atmosphérique. Nous associons au système de Saint-Venant bidimensionnel les conditions initiales et aux limites suivantes

$$\begin{aligned} H(0) &= H_0, \quad \vec{Q}(0) = \vec{Q}_0 & \text{dans } \Omega \\ \vec{Q} \cdot \vec{n} &= 0, \quad \text{curl} \left(\frac{\vec{Q}}{H} \right) = 0 & \text{sur } \gamma_0 \times (0, T) \\ \vec{Q} &= Q_1 \vec{n} & \text{sur } \gamma_1 \times (0, T) \\ H &= H_2 & \text{sur } \gamma_2 \times (0, T) \end{aligned} \quad (0.0.2)$$

avec γ_0 est le bord latéral, γ_1 et γ_2 sont les bord d'entrée et de sortie respectivement.

L'optimisation de forme est liée à une fonction objective qui dépend des capacités de nage du poisson.

Tout d'abord, nous calculons sa dérivée de forme en fonction des variables d'état et des variables adjointes en dérivant un système adjoint. Les résultats d'existence et d'unicité du système adjoint sont prouvés à l'aide des techniques pseudo-différentielles. Une méthode de volume fini est adoptée pour résoudre le problème d'état (0.0.1) combinée avec deux algorithmes de minimisation (algorithme de simplexe et gradient projeté) pour la résolution du problème d'optimisation de forme.

Le deuxième type de modèle basé sur un écoulement peu profond est un modèle multicouches. Le modèle multicouches est introduit pour éviter les calculs coûteux des équations de Navier-Stokes lorsque le débit est encore dans le régime peu profond.

Commençons par les équations de Navier-Stokes non stationnaires

$$\begin{aligned} \partial_t \rho + \nabla \cdot (\rho \vec{\mathbf{u}}) &= 0 & \text{dans } \Omega \times [0, T] \\ \partial_t (\rho \vec{\mathbf{u}}) + \nabla \cdot (\rho \vec{\mathbf{u}} \otimes \vec{\mathbf{u}}) &= \nabla \cdot \Sigma_T + \rho \mathbf{g} & \text{dans } \Omega \times [0, T] \end{aligned} \quad (0.0.3)$$

avec Σ_T est le tenseur donné par

$$\Sigma_T = -p \mathbf{Id} + \mu (\nabla \vec{\mathbf{u}} + (\nabla \vec{\mathbf{u}})')$$

et $\mathbf{g} = (\mathbf{0}, -g)' \in \mathbb{R}^d$ est l'accélération de gravité, et μ est une viscosité donnée.

En suivant la méthode initiée par Audusse et al. [14]. La hauteur de l'eau est divisée en un certain nombre de couches où l'hypothèse de Saint-Venant peut être appliquée à chaque couche. En considérant une vitesse verticale non constante, le profil vertical de la vitesse horizontale est décrit par le système multicouches suivant

$$\begin{aligned} \partial_t h_\alpha + \nabla_x \cdot (h_\alpha \vec{\mathbf{u}}_{H,\alpha}) &= G_{\alpha+\frac{1}{2}} - G_{\alpha-\frac{1}{2}}. \\ \rho \partial_t (h_\alpha \vec{\mathbf{u}}_{H,\alpha}) + \rho \nabla_x \cdot (h_\alpha \vec{\mathbf{u}}_{H,\alpha} \otimes \vec{\mathbf{u}}_{H,\alpha}) - \nabla_x \cdot (\mu h_\alpha (\nabla_x \vec{\mathbf{u}}_{H,\alpha} + (\nabla_x \vec{\mathbf{u}}_{H,\alpha})')) & \\ = \frac{1}{2} \rho G_{\alpha+\frac{1}{2}} (\vec{\mathbf{u}}_{H,\alpha+1} + \vec{\mathbf{u}}_{H,\alpha}) - \frac{1}{2} \rho G_{\alpha-\frac{1}{2}} (\vec{\mathbf{u}}_{H,\alpha} + \vec{\mathbf{u}}_{H,\alpha-1}) & \\ - \mu (\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}}) - \rho g h_\alpha \nabla_x (z_B + h) - h_\alpha \nabla_x p_S. & \end{aligned} \quad (0.0.4)$$

avec $G_{\alpha+1/2}$ est le flux normal de masse à l'interface entre les couches $\alpha - 1$ et α .

Nous montrons l'existence et l'unicité d'une solution forte du système multicouche bidimensionnel. Cela se fait en plusieurs étapes: l'analyse du problème linéarisé, l'obtention d'une estimation a priori et enfin l'utilisation d'un schéma de Picard pour le traitement du cas non linéaire. Afin d'étudier le problème d'optimisation de forme, nous procédons à la généralisation de la technique multicouches aux équations hydrostatiques de Navier-Stokes tridimensionnelles et dérivons un système multicouches 3-D. La description de l'écoulement multicouches de Saint-Venant permet de déduire une structure optimale de la passe à poissons en minimisant une fonction coût donnée modélisant la facilité de passage du poisson à travers la structure vers les zones de reproduction, de croissance ou d'alimentation. Nous dérivons un système adjoint lié au modèle multicouche 3-D obtenu et exprimons la dérivée de domaine de la fonction objective en termes des variables d'état et des variables adjointes. Nous présentons deux procédures de minimisation: la première est un algorithme de simplexe et la deuxième est un algorithme de type gradient (gradient projeté).

Le troisième type d'écoulement d'eau est basé sur le modèle de milieux poreux stationnaire. L'écoulement des fluides dans les milieux poreux est généralement décrit par les équations de Brinkman Forchheimer [29]:

$$\begin{aligned}
-\nu\Delta u + (u.\nabla)u + a |u|^\alpha u + \nabla p &= f \text{ in } \Omega; & \nabla.u &= 0 \text{ dans } \Omega \\
u &= 0 & & \text{sur } \Gamma_0 \\
u &= g & & \text{sur } \Gamma_1 \\
\nu \frac{\partial u}{\partial n} - p\mathbf{n} &= 0 & & \text{sur } \Gamma_2
\end{aligned} \tag{0.0.5}$$

avec u et p représentent respectivement la vitesse et la pression. ν est le coefficient de Brinkman, a est le coefficient de Forchheimer. f désigne les forces extérieures et $\alpha \in [1, 2]$ est un nombre réel choisi de manière appropriée tout au long de ce travail. Γ_0 est le bord latéral, Γ_1 est le bord d'entrée, du haut et du fond, et Γ_2 est le bord de sortie.

Dans un premier temps, nous avons examiné si le problème est bien posé et nous avons montré l'existence et l'unicité d'une solution faible du problème (0.0.5).

Parmi les difficultés liées à l'approximation numérique de la solution du problème (0.0.5), la contrainte d'incompressibilité " $\text{div } u = 0$ " puisque le champs de vitesse u et la pression son dissociés.

De multiple méthodes sont utilisées pour la relaxation de cette contrainte. Nous pouvons citer:

1. La méthode de pénalisation. Nous pouvons citer comme référence; J. Shen [19], et R. Temam [36]

$$\nabla.u^\varepsilon + \varepsilon p^\varepsilon = 0 \text{ in } \Omega \times [0, T]$$

2. Compressibilité artificielle, nous rappelons les travaux de Chorin [20] et de R. Temam [37]

$$\nabla.u^\varepsilon + \varepsilon p_i^\varepsilon = 0 \text{ in } \Omega \times [0, T]$$

3. La pression stabilisée

$$\nabla.u^\varepsilon - \varepsilon \Delta p^\varepsilon = 0 \text{ in } \Omega \times [0, T], \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \text{ in } \partial\Omega \times [0, T]$$

introduit par Temam [36] et Chorin [20, 21].

4. La méthode de pseudo-compressibilité

$$\nabla \cdot u^\varepsilon - \varepsilon \Delta p_t^\varepsilon = 0 \quad \text{in } \Omega \times [0, T], \quad \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \quad \text{in } \partial\Omega \times [0, T]$$

Durant ce travail, une méthode de pénalisation est utilisée pour relaxer la contrainte d'incompressibilité de la vitesse. Le problème perturbé associé au problème (0.0.5) s'écrit

$$\begin{aligned} -\nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f & \text{dans } \Omega \\ \nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 & \text{dans } \Omega \\ u^\varepsilon &= 0 & \text{sur } \Gamma_0 \\ u^\varepsilon &= g & \text{sur } \Gamma_1 \\ \nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 & \text{sur } \Gamma_2 \end{aligned} \tag{0.0.6}$$

Afin d'évaluer l'utilité de l'approche de pénalisation, nous dérivons des estimations à priori utiles pour entamer l'analyse de l'erreur. Nous utilisons une méthode de Faédo-Galerkin pour prouver l'existence d'une solution faible du problème perturbé (1.0.6). Le problème d'optimisation consiste à minimiser une fonction coût. Nous exprimons son gradient de forme en fonction de la vitesse u comme variable d'état, des variables adjointes, et le vecteur unité normal au bord du domaine.

Ces dernières années, la méthode d'éléments finis est utilisée dans plusieurs domaines d'ingénierie et d'analyses numériques [62, 57, 63]. Nous adoptons une méthode d'éléments finis discrète pour approximer la solution du problème pénalisé et établir des estimations à priori afin de prouver la convergence de la solution approchée vers la solution de l'équation de Navier Stokes Forchheimer. Le problème d'optimisation est implémenté en utilisant la méthode adjointe continue et la méthode d'éléments finis.

Le dernier type de modèle d'écoulement est décrit par les équations de Navier-Stokes Forchheimer non stationnaires

$$\begin{aligned} u_t - \nu \Delta u + (u \cdot \nabla) u + a |u|^\alpha u + \nabla p &= f & \text{dans } \Omega \\ \nabla \cdot u &= 0 & \text{dans } \Omega \\ u &= 0 & \text{sur } \Gamma_0 \\ u &= g & \text{sur } \Gamma_1 \\ \nu \frac{\partial u}{\partial n} - p \mathbf{n} &= 0 & \text{sur } \Gamma_2 \\ u|_{t=0} &= u_0 \end{aligned} \tag{0.0.7}$$

Dans un premier lieu, nous montrons l'existence et l'unicité d'une solution faible du problème (0.0.7). Les estimations d'erreur de la vitesse et de la pression sont établies à l'aide d'une méthode d'énergie.

Ensuite, nous entamons l'étude du problème perturbé en deux parties; la première est

consacrée au traitement du problème linéarisé. La deuxième partie traite le problème non linéaire.

Le problème perturbé prend la forme

$$\begin{aligned}
u_t^\varepsilon - \nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f && \text{dans } \Omega \\
\nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 && \text{dans } \Omega \\
u^\varepsilon &= 0 && \text{sur } \Gamma_0 \\
u^\varepsilon &= g && \text{sur } \Gamma_1 \\
\nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 && \text{sur } \Gamma_2 \\
u^\varepsilon|_{t=0} &= u_0 &&
\end{aligned} \tag{0.0.8}$$

Une discrétisation en temps est mise en place en utilisant un schéma d'Euler. Ce dernier est combiné à une discrétisation totale d'éléments finis discrète pour approcher le problème perturbé et dériver une estimation d'erreur de la vitesse et de la pression afin de montrer la convergence de la solution approchée vers la solution du problème initial (0.0.7).

Le problème d'optimisation de forme consiste à trouver une forme optimale de la structure en minimisant une fonction coût. Nous établissons un système adjoint associé au problème perturbé puis nous calculons le gradient de forme de la fonction à minimiser en fonction des variables d'état et adjointes. Le problème d'optimisation est implémenté en utilisant la méthode adjointe continue et la méthode d'éléments finis.

Dans ce qui suit, nous détaillons le lien entre les différents types de modèles d'écoulements mentionnés et les chapitres de cette thèse.

Dans le premier chapitre, nous considérons les équations des eaux peu profondes bidimensionnelles (0.0.1) avec les conditions initiales et aux bords (0.0.2).

Nous traitons ces équations en cinq sections. La première section est consacrée à la formulation mathématique de la structure du VSF. Nous donne la définition de la fonction coût à minimiser comme suit

$$J = \frac{1}{2} \int_0^T \int_\Omega \| \vec{u} - \vec{v} \|^2 dxdt + \frac{\alpha}{2} \int_0^T \int_\Omega |rot(\vec{u})|^2 dxdt \tag{0.0.9}$$

avec $\vec{u} = \frac{\vec{Q}}{H}$ est la vitesse de l'eau et (H, \vec{Q}) sont solutions du système (0.0.1)-(0.0.2). La vitesse $\frac{\vec{Q}}{H}$ est une vitesse cible liée à la capacité de nage des poissons. Elle s'écrit

$$\vec{v}(x_1, x_2) = \begin{cases} (c, 0) & \text{si } x_2 \leq \frac{1}{3} 0.97 \\ (0, 0) & \text{sinon} \end{cases} \tag{0.0.10}$$

La deuxième section est concernée au calcul du système adjoint associé aux équations (0.0.1) et la fonction coût (0.0.9). Le système adjoint obtenu prend la forme

$$\begin{aligned} -\frac{\partial p}{\partial t} + \frac{1}{H^2}(\vec{Q} \cdot \vec{\nabla})\vec{r} \cdot \vec{Q} - gH(\vec{\nabla} \cdot \vec{r}) - g\vec{\nabla}\eta \cdot \vec{r} &= -\left(\frac{\vec{Q}}{H} - v\right)\frac{\vec{Q}}{H^2} - \alpha \vec{curl}(\vec{curl}\left(\frac{\vec{Q}}{H}\right)) \cdot \frac{\vec{Q}}{H^2} \\ -\frac{\partial \vec{r}}{\partial t} - \vec{\nabla}p - \frac{1}{H}(\vec{Q} \cdot \vec{\nabla})\vec{r} - \frac{1}{H}(\vec{\nabla}\vec{r})^t \vec{Q} &= \frac{1}{H}\left(\frac{\vec{Q}}{H} - v\right) + \frac{\alpha}{H}\vec{curl}(\vec{curl}\left(\frac{\vec{Q}}{H}\right)) \end{aligned} \quad (0.0.11)$$

avec les conditions aux limites et finale suivantes

$$\begin{aligned} p(T) = 0, \quad \vec{r}(T) = 0 & \quad \text{dans } \Omega \\ \vec{r} \cdot \vec{n} = 0 & \quad \text{sur } \gamma_0 \times (0, T) \\ (gH - \frac{Q_1^2}{H^2})\vec{r} \cdot \vec{n} = 0 & \quad \text{sur } \gamma_1 \times (0, T) \\ \{p + \frac{1}{H_2}(\vec{Q} \cdot \vec{r})\}\vec{n} + \frac{1}{H_2}(\vec{Q} \cdot \vec{n})\vec{r} - \frac{\alpha}{H_2}\vec{curl}\left(\frac{\vec{Q}}{H_2}\right)\vec{r} = 0 & \quad \text{sur } \gamma_2 \times (0, T) \end{aligned} \quad (0.0.12)$$

avec $(H, \vec{Q}) \in L^2(0, T; \mathbb{R}) \times H^2(0, T; \Omega)$ est solution du système (0.0.1), et $(p, \vec{r}) \in (L^\infty([0, T]; H^2(\Omega)))^2$.

la troisième section est dédiée à la dérivation de domaine de la fonction coût (0.0.9) en fonction des variables d'état (H, \vec{Q}) les variables adjoints (p, \vec{r}) .

Théorème 0.0.1. *Soit $\vec{V} \in Lip(\Omega; \mathbb{R}^2)$. Si les fonctions A , J et L sont différentiables par rapport au domaine Ω dans un ensemble admissible X_0 , alors la dérivée du domaine de la fonction j à $\Omega \in X_0$ prend la forme*

$$\begin{aligned} \frac{\partial}{\partial \Omega} j(\Omega) \cdot \vec{V} &= \frac{\partial}{\partial \Omega} J(\Omega; H, \vec{Q}) \cdot \vec{V} - \frac{\partial}{\partial \Omega} A(\Omega; H, \vec{Q}; p, \vec{r}) \cdot \vec{V} \\ &+ \frac{\partial}{\partial \Omega} L(\Omega; p, \vec{r}) \cdot \vec{V} \end{aligned} \quad (0.0.13)$$

avec

$$\begin{aligned} \frac{\partial}{\partial \Omega} L(\Omega; p, \vec{r}) \cdot \vec{V} &= \int_0^T \int_{\Omega} \vec{f} \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}), \quad \text{for } \vec{V} \in Lip(\Omega; \mathbb{R}^2) \\ \frac{\partial}{\partial \Omega} J(\Omega; H, \vec{Q}) \cdot \vec{V} &= \frac{1}{2} \int_0^T \int_{\Omega} \left\| \frac{\vec{Q}}{H} - v \right\|^2 (\vec{\nabla} \cdot \vec{V}) + \frac{1}{2} \int_0^T \int_{\Omega} |\vec{curl}\left(\frac{\vec{Q}}{H}\right)|^2 (\vec{\nabla} \cdot \vec{V}) \\ \frac{\partial}{\partial \Omega} A(\Omega; H, \vec{Q}; p, \vec{r}) \cdot \vec{V} &= \int_0^T \int_{\Omega} \frac{\partial H}{\partial t} p(\vec{\nabla} \cdot \vec{V}) + \int_0^T \int_{\Omega} (\vec{\nabla} \cdot \vec{Q}) p(\vec{\nabla} \cdot \vec{V}) - \int_0^T \int_{\Omega} (\vec{\nabla} \vec{V})^t : \vec{\nabla} \vec{Q} p \\ &+ \int_0^T \int_{\Omega} \frac{\partial \vec{Q}}{\partial t} \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}) + \int_0^T \int_{\Omega} (\vec{\nabla} \cdot \vec{Q}) \frac{\vec{Q}}{H} \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}) - \int_0^T \int_{\Omega} (\vec{\nabla} \vec{V})^t : \vec{\nabla} \vec{Q} \frac{\vec{Q}}{H} \cdot \vec{r} \\ &+ \int_0^T \int_{\Omega} (\vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}) - \int_0^T \int_{\Omega} (\vec{\nabla} \vec{V} \vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r} + \int_0^T \int_{\Omega} gH \vec{\nabla} H \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}) \\ &\int_0^T \int_{\Omega} gH (\vec{\nabla} \vec{V})^t \vec{\nabla} H \cdot \vec{r} - \int_0^T \int_{\Omega} gH \vec{\nabla} \eta \cdot \vec{r}(\vec{\nabla} \cdot \vec{V}), \quad \text{for } \vec{V} \in Lip(\Omega; \mathbb{R}^2) \end{aligned}$$

pour $\vec{V} \in Lip(\Omega; \mathbb{R}^2)$, et (H, \vec{Q}) est solution du système (0.0.1) avec les condition initiales et aux bords (0.0.2) et (p, \vec{r}) est solution du système adjoint (0.0.11) avec les conditions finales et aux bords (0.0.12).

La quatrième section regroupe des résultats concernant l'existence et l'unicité du système adjoint (0.0.11). Nous commençons par l'écrire sous une forme conservative

$$\frac{\partial \vec{W}}{\partial t} + M(H, \vec{Q}) \frac{\partial \vec{W}}{\partial x} + N(H, \vec{Q}) \frac{\partial \vec{W}}{\partial y} = S(H, \vec{Q}, \vec{W}) \quad (0.0.14)$$

avec $\vec{W} = (p, r_1, r_2)^t$, $c = \sqrt{gH}$

$$M(H, \vec{Q}) = \begin{bmatrix} 0 & c^2 - u_1^2 & -u_1 u_2 \\ 1 & 2u_1 & u_2 \\ 0 & 0 & u_1 \end{bmatrix}, \quad S(H, \vec{Q}, \vec{W}) = \begin{bmatrix} F - g \vec{\nabla} \eta \cdot \vec{r} \\ -G_1 \\ -G_2 \end{bmatrix}$$

et

$$N(H, \vec{Q}) = \begin{bmatrix} 0 & -u_1 u_2 & c^2 - u_2^2 \\ 0 & u_2 & 0 \\ 1 & u_1 & 2u_2 \end{bmatrix}$$

Ensuite, nous montrons l'existence et l'unicité de la solution du système (0.0.14) en utilisant les opérateurs pseudo-différentiels

Théorème 0.0.2. *Si $\vec{W}^0 \in H^s(\mathbb{R}^2)$ et $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, alors il existe une constante $C(t)$ indépendante de $\varepsilon \in]0, 1]$ tel que*

$$\| \vec{W}_\varepsilon \|_{H^s}^2 \leq C(t) \{ \| \vec{W}^0 \|_{H^s}^2 + \| S \|_{L^\infty([0,t], H^s)}^2 \} \quad (0.0.15)$$

□

Corollaire 0.0.1. *Si $\vec{W}^0 \in H^s(\mathbb{R}^2)$ et $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, alors le système (0.0.14) admet une solution $\vec{W} \in H^s(\mathbb{R}^2)$ dans l'intervalle de temps $[0, T]$.* □

Théorème 0.0.3. *Si $\vec{W}^0 \in H^s(\mathbb{R}^2)$ et $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, alors le système (0.0.14) admet une solution unique dans $H^s(\mathbb{R}^2)$.* □

Dans la dernière section, nous adoptons une méthode de volumes finis pour résoudre les équations de Saint Venant (0.0.1) combinée à un algorithme de simplexe ou un algorithme de gradient projeté pour la résolution du problème d'optimisation de forme et minimiser la fonction objective (0.0.9).

Dans le deuxième chapitre, nous nous intéressons aux équations de Navier-Stokes non

stationnaires (0.0.3).

Le but de cette étude et d'adopter une approche multicouches qui consiste à subdiviser la hauteur de l'eau en plusieurs couches α . Nous notons la hauteur de l'eau dans chaque couche par h_α et par $u_{H,\alpha}$ la vitesse horizontale.

Dans la première et la deuxième section, nous donnons une formulation mathématique du problème puis nous dérivons le système multicouche bidimensionnel (0.0.4).

La troisième section est concernée par l'étude de l'existence et l'unicité du système multicouches 2-D. Afin d'étudier l'existence et l'unicité de la solution du système (0.0.4), nous le transformons sous forme d'un système de deux équations l'une est hyperbolique et l'autre est parabolique. Nous considérons $U = (u_1, \dots, u_N)$, le système à étudier s'écrit

$$\left\{ \begin{array}{l} \partial_t h + \partial_x \left(h \sum_{\beta=1}^N l_\beta u_\beta \right) = -G_{1/2} \\ \partial_t U - \nu \partial_{xx} U = -g \partial_x h (1, \dots, 1)^T - g \partial_x z_B (1, \dots, 1)^T - \frac{1}{\rho} \partial_x p_S (1, \dots, 1)^T \\ + G_{1/2} \frac{U}{h} + \left(\sum_{\beta=1}^N l_\beta \partial_x u_\beta \right) U + \left(\sum_{\beta=1}^N l_\beta u_\beta \right) \frac{U}{h} \partial_x h \\ - \partial_x (U^2) - \frac{U^2}{h} \partial_x h + (\mathbf{T}U) \partial_x U + (\mathbf{T}U) U \frac{\partial_x h}{h} - \nu \frac{U}{h} \partial_{xx} h + \frac{\nu}{2} (\mathbf{V}U) \frac{\partial_{xx} z_B}{h} \\ + \frac{\nu}{2} (\mathbf{W}U) \frac{\partial_{xx} h}{h} - \frac{\nu}{2} (\mathbf{V} \partial_x U) \frac{\partial_x z_B}{h} - \frac{\nu}{2} (\mathbf{W} \partial_x U) \frac{\partial_x h}{h} + 2\nu \frac{(\mathbf{X}U)}{h^2} + 2\nu \frac{(\mathbf{Y}U)}{h^2} \\ - G_{1/2} \frac{(\mathbf{Z}U)}{h} \end{array} \right. \quad (0.0.16)$$

où $\mathbf{T}, \mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}$, et \mathbf{Z} sont des matrices données et elles dépendent de h, U, z_B .

Dans cette partie, nous montrons l'existence et l'unicité de la solution du système (0.0.16).

Pour cela, nous montrons ce théorème

Notation 0.0.1. 1. Soient $U \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ et $h \in C(0, T, \mathbf{H}^1(\mathbb{R}))$, on définit la norme

$$\| (U, h) \|_{1-1}^2 = \| U \|_1^2 + \| h \|_1^2$$

2. Soient $U \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ et $h \in C(0, T, \mathbf{H}^2(\mathbb{R}))$, on définit la norme

$$\| (U, h) \|_{1-2}^2 = \| U \|_1^2 + \| h \|_2^2$$

3. Soient $U \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ et $h \in C(0, T, \mathbf{H}^2(\mathbb{R}))$, on définit la norme

$$\| (U, h) \|_{2-2}^2 = \| U \|_2^2 + \| h \|_2^2$$

4. Soient $U \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ et $h \in C(0, T, \mathbf{H}^3(\mathbb{R}))$, on définit la norme

$$\| (U, h) \|_{2_3}^2 = \| U \|_2^2 + \| h \|_3^2$$

Théorème 0.0.4. *On considère le système (0.0.16) avec la condition initiale*

$$(U, h)(0, x) = (U^0(x), h^0(x)) \in H^2(\mathbb{R}) \times H^3(\mathbb{R})$$

On suppose $z_b, p_S \in H^2(\mathbb{R})$, et on note

$$E = 2 \left(\| (U^0, h^0) \|_{2_3} \right)$$

On suppose qu'il existe une constante positive η_0 telle que $\inf_{x \in \mathbb{R}} h^0(x) \geq \eta_0 > 0$. Alors il existe $T > 0$ telle que le système (0.0.16) admet une solution forte (U, h) vérifiant

1. $U \in C(0, T, H^2(\mathbb{R})) \cap C^1(0, T, L^2(\mathbb{R})) \cap L^2(0, T, H^3(\mathbb{R}))$
2. $h \in C(0, T, H^2(\mathbb{R})) \cap C^1(0, T, H^1(\mathbb{R}))$
3. $\forall t \in [0, T]$

$$\forall x \in \mathbb{R}, h(t, x) \geq \inf_{x \in \mathbb{R}} h^0(x)$$

$$\max \left(\| (U, h)(t, \cdot) \|_{2_2}, \left(\int_0^T \| U(\tau, \cdot) \|_3^2 d\tau \right)^{\frac{1}{2}} \right) \leq E$$

La preuve de ce théorème est divisé en trois sous sections; l'estimation des termes sources, l'étude du système linéarisé, et la construction d'une suite convergente vers la solution du système (0.0.16).

Nous écrivons le système (0.0.16) sous la forme compacte

$$\begin{aligned} \partial_t U - \nu \partial_{xx} U &= S = S_b + S_h + S_{p_S} + S_{nl} \\ \partial_t h + \bar{u} \partial_x h &= F = -G_{1/2} - \partial_x(\bar{u})h \end{aligned} \tag{0.0.17}$$

avec

$$\begin{cases} \bar{u} = \sum_{\beta=1}^N l_\beta u_\beta, S_b = -g \partial_x z_B(1, \dots, 1), S_h = -g \partial_x h(1, \dots, 1), S_{p_S} = -g \partial_x p_S(1, \dots, 1) \\ S_{nl} = S_{nl}^{(1)} + S_{nl}^{(2)} + S_{nl}^{(3)} + S_{nl}^{(4)} + S_{nl}^{(5)} + S_{nl}^{(6)} + S_{nl}^{(7)} + S_{nl}^{(8)} + S_{nl}^{(9)} + S_{nl}^{(10)} \end{cases}$$

où

$$S_{nl}^{(1)} = G_{1/2}(\mathbf{I}_N - \mathbf{Z}) \frac{U}{h}; \quad S_{nl}^{(2)} = \left(\sum_{\beta=1}^N l_{\beta} \partial_x u_{\beta} \right) U; \quad S_{nl}^{(3)} = \left(\sum_{\beta=1}^N l_{\beta} u_{\beta} \right) \frac{\partial_x h}{h} U;$$

$$S_{nl}^{(4)} = ((\mathbf{T} - 2\mathbf{I}_N)U) \partial_x U; \quad S_{nl}^{(5)} = ((\mathbf{T} - \mathbf{I}_N)U) U \frac{\partial_x h}{h}; \quad S_{nl}^{(6)} = \frac{\nu}{2} ((\mathbf{W} - 2\mathbf{I}_N)U) \frac{\partial_{xx} h}{h};$$

$$S_{nl}^{(7)} = \frac{\nu}{2} \mathbf{V} U \frac{\partial_{xx} z_B}{h}; \quad S_{nl}^{(8)} = -\frac{\nu}{2} \mathbf{V} U \frac{\partial_x z_B}{h}; \quad S_{nl}^{(9)} = -\frac{\nu}{2} \mathbf{W} \partial_x U \frac{\partial_x h}{h}; \quad S_{nl}^{(10)} = 2\nu(\mathbf{X} + \mathbf{Y}) \frac{U}{h^2}$$

Dans ce qui suit, nous supposons que le fond est non pénétrable: l'échange de masse au fond est nulle ($G_{1/2} = 0$).

Proposition 0.0.1. *On suppose $U(t, \cdot), h(t, \cdot) \in H^2(\mathbb{R})$ tel que $h \geq \eta_0$, avec η_0 est une constante donnée strictement positive. alors on a*

- Soient S, F définis dans (0.0.17), alors $S, F \in H^1(\mathbb{R})$, et nous avons les estimations suivantes

$$\begin{aligned} \| S \|_1 \leq C(1 + \| (U, h) \|_{2_3} + \| (U, h) \|_{2_3}^2 + \| (U, h) \|_{2_3}^3) \| (U, h) \|_{2_3} \\ + g \| z_B \|_2 + \frac{1}{\rho} \| p_S \|_2 \end{aligned} \quad (0.0.18)$$

$$\| F \|_1 \leq C \| (U, h) \|_{2_3}^2 \quad (0.0.19)$$

- Soient $(U, h), (U', h') \in H^2(\mathbb{R}) \times H^3(\mathbb{R})$ tels que

$$\| (U, h) \|_{2_3} \leq E, \quad \| (U', h') \|_{2_3} \leq E, \quad \text{et } h, h' \geq \eta_0 > 0 \quad (0.0.20)$$

où E , et η_0 sont des constantes, alors

$$\| S(U, h) - S(U', h') \|_1 \leq C(1 + E + E^2 + E^3) \| (U - U', h - h') \|_{2_3} \quad (0.0.21)$$

$$\| F(U, h) - F(U', h') \|_1 \leq C_{ME} \| (U - U', h - h') \|_{2_3} \quad (0.0.22)$$

Après avoir établi les estimations des sources termes, nous entamons l'étude du système linéarisé

$$\partial_t U - \nu \partial_{xx} U = S(\tilde{U}, \tilde{h}, \partial_x \tilde{U}, \partial_x \tilde{h}, \partial_{xx} \tilde{h}) := \tilde{S} \quad (0.0.23)$$

$$\partial_t h + \tilde{u} \partial_x h = F(U, \tilde{h}) := \tilde{F} := -\partial_x(\tilde{u})\tilde{h}$$

avec \tilde{U} , et \tilde{h} sont connues. Nous divisons l'étude du système (0.0.23) en deux systèmes; l'un parabolique

$$\begin{aligned} \partial_t U - \nu \partial_{xx} U &:= \tilde{S} \\ U(0, \cdot) &= U^0(\cdot) \in H^2(\mathbb{R}) \end{aligned} \quad (\text{Par})$$

et l'autre hyperbolique

$$\begin{aligned} \partial_t h + \bar{u} \partial_x h &= \tilde{F} \\ h(0, \cdot) &= h^0 \in H^2(\mathbb{R}) \end{aligned} \tag{Hyp}$$

Nous montrons les résultats suivants

Proposition 0.0.2. *Si $\tilde{S} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$ pour $T > 0$, alors*

1. *le problème (Par) admet une solution forte U vérifiant*

$$U \in C(0, T; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R})) \tag{0.0.24}$$

2. *il existe deux constantes positives C_1 et C_2 , qui peuvent dépendre de la viscosité, tel que pour tout $t \in [0, T]$:*

$$\|U(t, \cdot)\|_2^2 + C_1 \int_0^t \|U(\tau, \cdot)\|_3^2 d\tau \leq e^t \left(\|U^0\|_2^2 + C_2 \int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right) \tag{0.0.25}$$

□

Lemme 0.0.1. *Soit $T > 0$ tel que $\tilde{F} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$, et $\bar{u} \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$, $h^0 \in H^2(\mathbb{R})$*

Alors

1. *Le problème (Hyp) admet une solution forte h telle que:*

$$h \in C(0, t; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \tag{0.0.26}$$

2. *Pour tout $t \in [0, T]$, on a l'inégalité d'énergie:*

$$\|h(t, \cdot)\|_1 \leq e^{C_1 E_{\bar{u}} t} \left(\|h^0\|_1 + \int_0^t e^{-C_1 E_{\bar{u}} \tau} \|\tilde{F}(\tau, \cdot)\|_1 d\tau \right) \tag{0.0.27}$$

avec

$$C_1 > 0, \text{ est une constante donnée et } E_{\bar{u}} := \sup_{0 \leq t \leq T} \|\bar{u}\|_2$$

Proposition 0.0.3. *Soit $T > 0$ tel que $\tilde{F} \in C(0, T; \mathbf{H}^2(\mathbb{R}))$, $\tilde{u} \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$ et $h^0 \in H^2(\mathbb{R})$*

Alors

1. *Le problème (Hyp) admet une solution forte h telle que:*

$$h \in C(0, t; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{H}^1(\mathbb{R})) \quad (0.0.28)$$

2. *Pour tout $t \in [0, T]$ on a l'inégalité d'énergie:*

$$\|h(t, \cdot)\|_2 \leq e^{C_2 E_{\tilde{u}} t} \left(\|h^0\|_2 + \int_0^t e^{-C_2 E_{\tilde{u}} \tau} \|\tilde{F}(\tau, \cdot)\|_2 d\tau \right) \quad (0.0.29)$$

avec

$$C_2 > 0, \text{ est une constante donnée et } E_{\tilde{u}} := \sup_{0 \leq t \leq T} \|\tilde{u}\|_2$$

À la fin de la preuve du théorème 0.0.4, nous construisons une suite récursive $(U^{(j)}, h^{(j)}) = (u_1^{(j)}, \dots, u_N^{(j)}, h^{(j)})_{j \in \mathbb{N}}$ comme suit

- $\forall (t, x) \in [0, T] \times \mathbb{R}$, $(U^{(0)}, h^{(0)})(t, x) = (U^0, h^0)(x) \in \mathbf{H}^2(\mathbb{R})$
- $\forall j \in \mathbb{N}$, $(U^{(j+1)}, h^{(j+1)})$ est solution de

$$\begin{cases} \partial_t U^{(j+1)} - \nu \partial_{xx} U^{(j+1)} = S(j) & \text{dans } [0, T] \times \mathbb{R} \\ \partial_t h^{(j+1)} - \bar{u}^{(j)} \partial_x h^{(j+1)} = F^{(j, j+1)} & \text{dans } [0, T] \times \mathbb{R} \\ (U^{(j)}, h^{(j)})(0, \cdot) = (U^{(0)}, h^{(0)})(\cdot) & \text{dans } \mathbb{R} \end{cases} \quad (0.0.30)$$

où pour tout $j \in \mathbb{N}$:

$$\begin{cases} S^{(j)} = S(U^{(j)}, h^{(j)}, \partial_x U^{(j)}, \partial_x h^{(j)}, \partial_{xx} h^{(j)}), \\ F^{(j, j+1)} = -G_{1/2} - \partial_x(\bar{u}^{(j+1)})h^{(j)} \end{cases}$$

nous définissons les constantes

$$E_0 = 2 \| (U^0, h^0) \|_{2,3} \text{ et } \eta_0 = \frac{1}{2} \inf_{x \in \mathbb{R}} h^0(x)$$

Proposition 0.0.4. *Soit $(h^0, U^0) \in \mathbf{H}^3(\mathbb{R}) \times \mathbf{H}^2(\mathbb{R})$. Alors il existe $T > 0$ petit tel que*

- *La suite $(U^{(j)}, h^{(j)})_{j \in \mathbb{N}}$ est bien définie et satisfait, pour tout $t \in [0, T]$ et tout $j \in \mathbb{N}$:*

$$\begin{aligned} U^{(j)} &\in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, t; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R})) \\ h^{(j)} &\in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{H}^1(\mathbb{R})) \end{aligned} \quad (0.0.31)$$

- $\forall (t, x) \in [0, T] \times \mathbb{R}$ et $\forall j \in \mathbb{N}$, on a

$$\begin{aligned} & \| (U^{(j)}, h^{(j)})(t, \cdot) \|_{2,2} \leq E_0 \\ & \left(\int_0^t \| U^{(j)}(\tau, \cdot) \|_3^2 d\tau \right)^{1/2} \leq E_0 \end{aligned} \quad (0.0.32)$$

$$h^{(j)}(t, x) \geq \eta_0 > 0 \quad (0.0.33)$$

Après l'étude du système multicouches 2-D avec l'existence et l'unicité de sa solution, Nous généralisons l'approche mutlicouches aux équations de Navier Stokes hydrostatiques non stationnaires 3-D dans la quatrième section

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \\ \frac{\partial \mathbf{U}}{\partial t} + \nabla_x \cdot (\mathbf{U} \otimes \mathbf{U}) + \frac{\partial (\mathbf{U}w)}{\partial z} + \nabla_x p &= \mu \frac{\partial^2 \mathbf{U}}{\partial z^2} \\ \frac{\partial p}{\partial z} &= -g \end{aligned} \quad (0.0.34)$$

avec les conditions aux bords

$$\begin{aligned} w(t, x, z_b) &= 0 \\ \mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_b) &= k \mathbf{U}(t, x, y, z_b) \\ \frac{\partial \mathbf{U}}{\partial z}(t, x, h) &= 0 \\ p(t, x, h) &= 0 \end{aligned} \quad (0.0.35)$$

et la condition cinématique

$$\frac{\partial h}{\partial t} + \mathbf{U}(t, x, h) \nabla_x h - w(t, x, h) = 0 \quad (0.0.36)$$

où $\mathbf{u}(t, x, z) = (u, v, w)$ est le vecteur vitesse; et $\mathbf{U}(t, x, z) = (u, v)$ est la vitesse horizontale. Nous dérivons un système multicouches à partir du système (0.0.34). Ce système est utilisé pour la modélisation des structures de passe à poissons. Pour cela, nous considérons le système multicouches suivant:

$$\begin{aligned} \frac{\partial h_\alpha}{\partial t} + \nabla_x \cdot \mathbf{q}_\alpha &= 0 && \text{dans } \Omega_\alpha \times [0, T] \\ \frac{\partial \mathbf{q}_\alpha}{\partial t} + \nabla_x \cdot \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \otimes \mathbf{q}_\alpha \right) + \frac{g}{2} \nabla_x (h_\alpha h) &= \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) - g h_\alpha \nabla_x z_b - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \\ + 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} - 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} & \text{for } \alpha = 1, \dots, N && \text{dans } \Omega_\alpha \times [0, T] \end{aligned} \quad (0.0.37)$$

avec les conditions aux bords et initiales suivantes

$$\begin{aligned}
h_\alpha(0) &= h_0, & q_\alpha(0) &= q_0 & \text{dans } \Omega_\alpha \\
q_\alpha \cdot \mathbf{n} &= 0, & \text{curl} \left(\frac{q_\alpha}{h_\alpha} \right) &= 0 & \text{sur } \gamma_0 \times [0, T] \\
q_\alpha &= q_1 \mathbf{n} & & & \text{sur } \gamma_\alpha^1 \times [0, T] \\
h_\alpha &= h_2 & & & \text{sur } \gamma_\alpha^2 \times [0, T]
\end{aligned} \tag{0.0.38}$$

Ensuite, nous dérivons un système adjoint à partir du système (0.0.37),

$$\begin{aligned}
-\frac{\partial p}{\partial t} &+ \frac{1}{h_\alpha^2} (\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} \cdot \mathbf{q}_\alpha + g \nabla h \cdot \mathbf{r} + g \nabla z_b \cdot \mathbf{r} - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha^2} \cdot \mathbf{r} \\
&- 2\nu_\alpha \frac{h_{\alpha+1}^3 \mathbf{q}_\alpha - h_\alpha^2 h_{\alpha+1} \mathbf{q}_{\alpha+1} + 2h_\alpha h_{\alpha+1}^2 \mathbf{q}_\alpha}{(h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1})^2} \cdot \mathbf{r} \\
&+ 2\nu_{\alpha-1} \frac{2h_\alpha^2 h_{\alpha-1} \mathbf{q}_{\alpha-1} - 2h_\alpha h_{\alpha-1}^2 \mathbf{q}_\alpha - h_\alpha h_{\alpha-1}^2 \mathbf{q}_{\alpha-1} - h_{\alpha-1}^3 \mathbf{q}_\alpha}{(h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha)^2} \cdot \mathbf{r} \\
&= - \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) \frac{\mathbf{q}_\alpha}{h_\alpha^2} - \sigma \mathbf{curl}(\text{curl}(\frac{\mathbf{q}_\alpha}{h_\alpha})) \cdot \frac{\mathbf{q}_\alpha}{h_\alpha^2} \quad \text{dans } \Omega_\alpha \times [0, T] \\
-\frac{\partial \mathbf{r}}{\partial t} &- \nabla p - \frac{1}{h_\alpha} (\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} - \frac{1}{h_\alpha} (\nabla \mathbf{r})^t \mathbf{q}_\alpha + k_\alpha \frac{\mathbf{r}}{h_\alpha} + 2\nu_\alpha \frac{h_{\alpha+1} \mathbf{r}}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \\
&+ 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{r}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \\
&= \frac{1}{h_\alpha} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) + \sigma \frac{1}{h_\alpha} \mathbf{curl}(\text{curl}(\frac{\mathbf{q}_\alpha}{h_\alpha})) \quad \text{dans } \Omega_\alpha \times [0, T]
\end{aligned} \tag{0.0.39}$$

avec les conditions finales et aux bords suivantes

$$\begin{aligned}
p(T) &= 0, & \mathbf{r}(T) &= 0 & \text{dans } \Omega_\alpha \\
\mathbf{r} \cdot \mathbf{n} &= 0 & & & \text{sur } \gamma_0 \times [0, T] \\
-\frac{q_1}{h_\alpha^2} \cdot \mathbf{r} \cdot \mathbf{n} &= 0 & & & \text{sur } \gamma_\alpha^1 \times [0, T] \\
(p + \frac{1}{h_2} (\mathbf{q}_\alpha \cdot \mathbf{r})) \mathbf{n} &+ \frac{1}{h_2} (\mathbf{q}_\alpha \cdot \mathbf{n}) \cdot \mathbf{r} - \frac{\sigma}{h_2} \text{curl}(\frac{\mathbf{q}_\alpha}{h_2}) \boldsymbol{\tau} &= 0 & \text{sur } \gamma_\alpha^2 \times [0, T]
\end{aligned} \tag{0.0.40}$$

et nous calculons la dérivée du domaine de la fonction objective à chaque couche α :

$$j_\alpha = \frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left\| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right\|^2 dt dx + \frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} \left| \text{curl}(\frac{\mathbf{q}_\alpha}{h_\alpha}) \right|^2 dt dx \tag{0.0.41}$$

Pour le calcul de la dérivée du domaine de la fonction coût (0.0.41), nous montrons ce résultat

Théorème 0.0.5. Soit $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$, alors la dérivée du domaine de la fonction j_α par rapport à $\Omega_\alpha \in X_\alpha^0$ est donnée par

$$\frac{\partial}{\partial \Omega_\alpha} j_\alpha(\Omega_\alpha) \cdot \mathbf{V} = \frac{\partial}{\partial \Omega_\alpha} J_\alpha(\Omega_\alpha; h_\alpha, \mathbf{q}_\alpha) \cdot \mathbf{V} - \frac{\partial}{\partial \Omega_\alpha} A_\alpha(\Omega_\alpha; h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} \quad (0.0.42)$$

avec

$$J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) = \underbrace{\frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left\| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right\|^2}_{J_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha)} + \underbrace{\frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} \left| \text{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right) \right|^2}_{J_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha)}$$

et

$$\begin{aligned} A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) &= \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{\partial h_\alpha}{\partial t} p}_{A_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} (\nabla_x \cdot \mathbf{q}_\alpha) p}_{A_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{\partial \mathbf{q}_\alpha}{\partial t} \cdot \mathbf{r}}_{A_\alpha^3(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} \nabla_x \cdot \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \otimes \mathbf{q}_\alpha \right) \cdot \mathbf{r}}_{A_\alpha^4(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{g}{2} \nabla_x (h_\alpha h) \cdot \mathbf{r}}_{A_\alpha^5(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} - \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) \cdot \mathbf{r}}_{A_\alpha^6(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} g h_\alpha \nabla_x z_b \cdot \mathbf{r}}_{A_\alpha^7(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r}}_{A_\alpha^8(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &- \underbrace{\int_0^T \int_{\Omega_\alpha} 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \cdot \mathbf{r}}_{A_\alpha^9(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \cdot \mathbf{r}}_{A_\alpha^{10}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \end{aligned}$$

où $(h_\alpha, \mathbf{q}_\alpha)$ est solution du système d'état (0.0.37) avec les conditions initiales et aux bords (0.0.38), et (p, \mathbf{r}) est solution du système adjoint (0.0.39) avec les conditions finales et aux bords (0.0.40). \square

La dernière sections est dédiée aux méthodes numériques utilisées. Nous comparons deux techniques de résolution pour le problème d'optimisation de forme; la première est une méthode qui ne s'appuie pas sur l'utilisation du gradient de la fonction objective dans

ses itérations. la deuxième technique est de type gradient, Elle fait un appel du gradient de la fonction coût à chaque itération. Enfin, nous donnons quelques exemples de tests numériques.

Quant au troisième chapitre, nous nous intéressons aux équations de Brinkman Forchheimer stationnaires (0.0.5).

L'étude dans ce chapitre est faite en plusieurs sections.

Tout d'abord, nous donnons une définition des espaces à utiliser pour la suite. Pour $d=2, 3$ nous définissons

$$X = (H_0^1(\Omega))^d, \quad Y = L^2(\Omega), \quad \hat{Y} = \{q \in H^1(\Omega)\}, \quad M = \left\{ q \in L^2(\Omega); \int_{\Omega} q dx = 0 \right\}$$

et X' est l'espace dual de X .

Dans la première section, nous montrons l'existence et l'unicité d'une solution du problème (0.0.5), puis nous introduisons le problème perturbé (1.0.6). L'existence d'une solution faible du problème pénalisé est prouvée en utilisant une méthode de Féado-Galerkin.

Théorème 0.0.6. *Soit Ω un domaine borné dans \mathbb{R}^n ($n \geq 0$), $f \in X'$ est une fonction donnée. alors le problème perturbé (1.0.6) admet une solution qui satisfait*

$$\frac{\nu}{2} \|\nabla u^\varepsilon\|^2 + a \|u^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p^\varepsilon\|^2 \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (0.0.43)$$

À la fin de cette section, nous établissons les estimations d'erreur de la vitesse et de la pression en norme H^1 et L^2 respectivement.

Théorème 0.0.7. *Soit $(u^\varepsilon, p^\varepsilon)$ solution du problème (1.0.6) et (u, p) est la solution dans $X \times Y/\mathbb{R}$ du système (0.0.5). Alors*

$$(p^\varepsilon) \in Y \quad \text{est borne uni formment en } \varepsilon \quad (0.0.44)$$

et quand $\varepsilon \rightarrow 0$.

$$u^\varepsilon \rightarrow u \quad \text{dans } X \quad (0.0.45)$$

et

$$p^\varepsilon \rightarrow \tilde{p} \quad \text{dans } Y \quad (0.0.46)$$

où \tilde{p} est la projection orthogonale dans Y de p sur $(\ker B^*)^\perp$. et nous avons des estimations suivantes

$$\|u^\varepsilon - u\|_{H^1} + \|p^\varepsilon - \tilde{p}\| \leq C\varepsilon \quad (0.0.47)$$

avec C est une constante indépendante de ε .

où

$$\ker B^* = \{q \in Y \mid (q, \nabla \cdot v) = 0 \text{ for all } v \in X\}$$

Dans la deuxième section, nous entamons l'analyse du problème du contrôle optimal. Le problème d'optimisation de forme consiste à minimiser la fonction suivante

$$J(\Omega) = \frac{1}{2} \int_{\Omega} |u^\varepsilon - u_d|^2 dx + \frac{\sigma}{2} \int_{\Omega} |\operatorname{curl}(u^\varepsilon)|^2 dx \quad (0.0.48)$$

où u^ε est solution du problème (0.0.5) et u_d est une vitesse cible dépendante des capacités de nages des poissons et de leurs natures. σ est un paramètre de vorticit . Le syst me adjoint associ  au probl me (1.0.6) et la fonction co t (0.0.48) s' crivent

$$\begin{aligned} -\nu \Delta v + (\nabla u^\varepsilon)^T \cdot v - (u^\varepsilon \cdot \nabla)v - \frac{1}{2} \nabla(u^\varepsilon \cdot v) + \frac{1}{2} (\nabla \cdot u^\varepsilon)v + a|u^\varepsilon|^{\alpha-2}(|u^\varepsilon|^2 v + \alpha(u^\varepsilon \cdot v)u^\varepsilon) \\ + \nabla q = (u^\varepsilon - u_d) - \sigma \vec{\operatorname{curl}}(\operatorname{curl}(u^\varepsilon)) \quad \text{dans } \Omega \\ \nabla \cdot v + \varepsilon q = 0 \quad \text{dans } \Omega \\ v = 0 \quad \text{sur } \Gamma_0 \cup \Gamma_1 \\ \nu \frac{\partial v}{\partial n} + 2(u^\varepsilon \cdot n)v - nq = \sigma \operatorname{curl}(u^\varepsilon) \cdot \tau \quad \text{sur } \Gamma_2 \end{aligned} \quad (0.0.49)$$

La deuxi me partie de cette section est concern e au calcul du gradient de forme de la fonction objective (0.0.48) en fonction des variables d' tat $u^\varepsilon, p^\varepsilon$ et les variables adjointes v , and q

$$\nabla J = \left[\frac{1}{2} |u^\varepsilon - u_d|^2 + \frac{\sigma}{2} |\operatorname{curl}(u^\varepsilon)|^2 + \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial v}{\partial \mathbf{n}} - \sigma \operatorname{curl}(u^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{n} \quad (0.0.50)$$

La quatri me partie de cet axe se focalise sur l' tude d'une m thode p nalis e d' l ments finis. Nous discr tisons le probl me (0.0.5), et nous associons au syst me obtenu le probl me discret et perturb  suivant

Trouver $u_h^\varepsilon \in X^h$ tel que

$$\begin{aligned} \nu(\nabla u_h^\varepsilon, \nabla v) + ((u_h^\varepsilon \cdot \nabla)u_h^\varepsilon, v_h) + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h) + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon, v_h) \\ + \frac{1}{\varepsilon} I((\nabla \cdot u_h^\varepsilon)(\nabla \cdot v_h)) = (f, v_h) \text{ pour tout } v_h \in X^h \end{aligned} \quad (0.0.51)$$

De mani re analogue que pour le probl me continu, nous d montrons l'existence de la solution du probl me discret et perturb  (0.0.51). Ensuite, nous  tablissons une estimation d'erreur en terme de la norme H^1 et L^2 pour la vitesse et la pression respectivement.

Théorème 0.0.8. Soit $(u,p) \in X \times Y$ solution du problème (0.0.5) et $(u_h^\varepsilon, p_h^\varepsilon)$ définis par (0.0.51) alors

$$\begin{aligned} \|u - u_h^\varepsilon\|_1 \leq & \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_1 \right) \left(1 + \frac{C}{\beta_h} \right) \inf_{v_h \in X^h} \|u - v_h\|_1 \\ & + \frac{C_2}{\nu r} \inf_{q_h \in Y^h} \|p - q_h\| + \frac{C_3}{\beta_h} \varepsilon \|p_h\| \end{aligned} \quad (0.0.52)$$

et

$$\begin{aligned} \|\hat{p} - \hat{p}_h^\varepsilon\|_{Z^h} \leq & \frac{\nu(3-2r) + aC_1}{\beta_h} \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_2 \right) \left(1 + \frac{C_3}{\beta_h} \right) \inf_{v_h \in X^h} \|u - v_h\|_1 \\ & + \left(1 + \frac{C_4}{\beta_h} \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_1 \right) \right) \inf_{q_h \in Y^h} \|p - q_h\| + \frac{C''}{\beta_h^2} \varepsilon \|p_h\| \end{aligned} \quad (0.0.53)$$

□

où l'espace d'éléments finis Z^h est défini par

$$Z^h = \{ \bar{q}_h \in Y^h \mid \bar{q}_h - q_h \in \ker B_h^* \} \quad (0.0.54)$$

avec

$$\ker B_h^* = \{ q_h \in Y^h \mid I(q_h \nabla \cdot v_h) = 0 \quad \forall v_h \in X^h \} \quad (0.0.55)$$

et X^h, Y^h sont les espace d'éléments finis associés à (X, Y) .

Dans la dernière section de ce chapitre, nous adoptons une technique de type gradient pour résoudre notre problème d'optimisation et minimiser la fonction coût (0.0.48) combinée à une méthode d'éléments finis discrète pour approcher le problème (0.0.5).

Le quatrième chapitre se rapporte à l'étude des équations de Navier-Stokes Forchheimer non stationnaires (0.0.7). Il est divisé en plusieurs sections.

Dans la première section, nous nous préoccupons de certaines résultats utilisées pour la suite du chapitre. Nous donnons la formulation mathématique du problème puis nous mettons en place quelques propriétés de régularité sur la fonction f et la pression p .

La deuxième section est divisée en deux parties. Dans la première, nous étudions le cas linéaire des équations de Navier Stokes Forchheimer non stationnaires et nous dérivons une estimations d'erreur qui sera utilisée dans la deuxième partie pour le cas non linéaire.

Théorème 0.0.9. Supposons $u_0 \in (H_0^1(\Omega))^d$ ($d = 2, 3$), et $f \in L^2(\Omega)$, alors

$$\|e(t)\| + \left(\int_0^t \|e(s)\|_1^2 ds \right)^{\frac{1}{2}} \leq C\sqrt{\varepsilon} \quad \forall t \in [0, T_0]$$

Supposons $u_0 \in (H_0^1(\Omega))^d$ ($d = 2, 3$), $f \in L^2(\Omega)$ et $tf_t \in L^2(0, T; H^1(\Omega))$, alors

$$\sqrt{t}\|e(t)\| + t\|e(t)\|_1 + \left(\int_0^t s^2 \|q\|^2 ds \right)^{\frac{1}{2}} \leq C\varepsilon \quad \forall t \in [0, T_0]$$

Ce résultat obtenu sera utilisé pour déduire une estimation générale pour l'analyse d'erreur du système non linéaire

Théorème 0.0.10. *Supposons $u_0 \in X$, $f \in L^2(\Omega)$ et $tf_t \in L^2(0, T; H^1(\Omega))$, alors nous avons les estimations d'erreurs suivantes*

$$\sqrt{t}\|u - u^\varepsilon\| + \sqrt{\nu t}\|u - u^\varepsilon\|_1 + \left(\int_0^t s^2 \|p - p^\varepsilon\|^2 ds \right)^{\frac{1}{2}} \leq C\varepsilon \quad \forall t \in]0, T_0]$$

La troisième section est consacrée au contrôle optimal; nous introduisons un problème d'optimisation de forme, nous calculons une équations adjoint associée au problème perturbé (0.0.8) et la fonction coût suivante

$$J(\Omega) = \frac{1}{2} \int_0^T \int_\Omega |u^\varepsilon - u_d|^2 dxdt + \frac{\sigma}{2} \int_0^T \int_\Omega |\operatorname{curl}(u^\varepsilon)|^2 dxdt \quad (0.0.56)$$

Le système adjoint s'écrit

$$\left\{ \begin{array}{ll} -\mathbf{v}_t - \nu \Delta \mathbf{v} + (\nabla \mathbf{u}^\varepsilon)^T \cdot \mathbf{v} - (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} - \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} + a |\mathbf{u}^\varepsilon|^{\alpha-2} (|\mathbf{u}^\varepsilon|^2 \mathbf{v} + \alpha (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon) \\ \quad + \nabla q = (\mathbf{u}^\varepsilon - \mathbf{u}_d) - \sigma \vec{\operatorname{curl}}(\operatorname{curl}(\mathbf{u}^\varepsilon)) & \text{dans } \Omega \times [0, T] \\ \nabla \cdot \mathbf{v} + \varepsilon q = 0 & \text{dans } \Omega \times [0, T] \\ \mathbf{v} = 0 & \text{sur } (\Gamma_0 \cup \Gamma_1) \times [0, T] \\ \nu \frac{\partial \mathbf{v}}{\partial n} + 2(\mathbf{u}^\varepsilon \cdot n) \mathbf{v} - nq = \sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \cdot \boldsymbol{\tau} & \text{sur } \Gamma_2 \times [0, T] \\ \mathbf{v}(T) = 0 & \end{array} \right. \quad (0.0.57)$$

Ensuite, nous calculons le gradient de forme de la fonction objective (0.0.56) en fonction des variables d'états et adjointes

$$\nabla J = \left[\frac{1}{2} |u^\varepsilon - u_d|^2 + \frac{\sigma}{2} |\operatorname{curl}(u^\varepsilon)|^2 + \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} - \sigma \operatorname{curl}(u^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{n} \quad (0.0.58)$$

Pour ce qui suit, nous définissons les formes suivantes

$$a(\mathbf{u}, \mathbf{v}) = \nu(A^{1/2} \mathbf{u}, A^{1/2} \mathbf{v}), \quad a_\varepsilon(\mathbf{u}, \mathbf{v}) = \nu(A_\varepsilon^{1/2} \mathbf{u}, A_\varepsilon^{1/2} \mathbf{v}), \quad c_\alpha(\mathbf{u}, \mathbf{v}) = (a |\mathbf{u}|^\alpha \mathbf{u}, \mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v} \in X$$

avec $A = -P_{\mathbf{H}}\Delta = \tilde{\Delta}$ est l'opérateur de Stokes et $P_{\mathbf{H}}$ est la projection orthogonale dans $\mathbf{L}^2(\Omega)$ sur \mathbf{H} où $\mathbf{H} = \{\mathbf{v} \in \mathbf{L}^2(\Omega), : \nabla \cdot \mathbf{v} = 0, \mathbf{v} \cdot \mathbf{n}|_{\Gamma} = 0\}$, avec $\Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2$

Dans la quatrième section, nous adoptons une méthode d'éléments finis pour discrétiser le problème continu en espace et un schéma d'Euler pour une discrétisation en temps. La formulation variationnelle de l'équation (0.0.7) et de l'équations pénalisée (0.0.8) s'écrit respectivement comme suit

$$\begin{aligned} & \text{Trouver } (u, p) \in L^\infty(0, T; Y) \cap L^2(0, T; X) \times L^2(0, T; M) \text{ tel que} \\ & (u_t, v) + a(u, v) - d(v, p) + d(u, q) + b(u, u, v) + c_\alpha(u, v) = (f, v), \quad \forall (v, q) \in X \times M \end{aligned} \quad (0.0.59)$$

et

$$\begin{aligned} & \text{Trouver } (u_\varepsilon, p_\varepsilon) \in L^\infty(0, T; Y) \cap L^2(0, T; X) \times L^2(0, T; M) \text{ tel que} \\ & (u_{\varepsilon t}, v) + a(u_\varepsilon, v) - d(v, p_\varepsilon) + d(u_\varepsilon, q) + \varepsilon(p_\varepsilon, q) + \tilde{b}(u_\varepsilon, u_\varepsilon, v) + c_\alpha(u_\varepsilon, v) = (f, v) \\ & \forall (v, q) \in X \times M \end{aligned} \quad (0.0.60)$$

avec les données initiales suivantes; $u(0) = u_0$ et $u_\varepsilon(0) = u_0$.

Nous définissons la discrétisation en temps de la formulation faible perturbée (0.0.60) avec le schéma d'Euler

$$(d_t u_\varepsilon^n, v) + a(u_\varepsilon^n, v) - d(v, p_\varepsilon^n) + d(u_\varepsilon^n, q) + \varepsilon(p_\varepsilon^n, q) + \tilde{b}(u_\varepsilon^n, u_\varepsilon^n, v) + c_\alpha(u_\varepsilon^n, v) = (f(t_n), v) \quad (0.0.61)$$

pour tout $(v, q) \in X \times M$, et $1 \leq n \leq N$, où $0 < \Delta t < 1$ est le pas de temps, $t_n = n\Delta$, $t_N = T$, $(u_\varepsilon^0, p_\varepsilon^0) = (u_0, 0)$ et $d_t u_\varepsilon^n = \frac{u_\varepsilon^n - u_\varepsilon^{n-1}}{\Delta t}$ pour $1 \leq n \leq N$, et $d_t u_\varepsilon^0$ est défini comme suit

$$(d_t u_\varepsilon^0, v) = a(u_0, v) + ((u_0 \cdot \nabla) u_0, v) + c_\alpha(u_0, v) = (f(0), v), \quad \forall v \in X \text{ with } \operatorname{div} v = 0$$

Nous supposons que le couple de données (u_0, f) vérifie

$$\|Au_0\| + \sup_{t \in [0, T]} \{\|f(t)\| + \|f_t(t)\|\} \leq C \quad (A0)$$

Nous obtenons les résultats suivants

Théorème 0.0.11. *Supposons que (A₀) et $c\varepsilon \leq 1$ sont vérifiées, il existe une constante $k_0 > 0$ tel que si $k_0\Delta t \leq 1$, alors*

$$\|A_\varepsilon^{1/2} u_\varepsilon^m\|^2 + \Delta t \sum_{n=1}^m (\|d_t u_\varepsilon^n\|^2 + \|A_\varepsilon u_\varepsilon^n\|^2 + \|p_\varepsilon^n\|^2 + a\|u_\varepsilon^n\|_{L^{\alpha+2}}^{\alpha+2}) \leq C \quad (0.0.62)$$

$$\|d_t u_\varepsilon^m\|^2 + \|A_\varepsilon u_\varepsilon^m\|^2 + \|p_\varepsilon^m\|_1^2 + \Delta t \sum_{n=1}^m \|A_\varepsilon^{1/2} d_t u_\varepsilon^n\|^2 \leq C \quad (0.0.63)$$

pour tout $1 \leq m \leq N$, avec $\|\cdot\|_1$ est la norme de l'espace de Sobolev $H^1(\Omega)$.

Nous supposons pour le couple (X_h, M_h) des espace d'éléments finis associés à X and M, que

$$\text{Il existe un opérateur } r_h : D(A) \cap X \rightarrow X_h \text{ tel que} \quad (A1)$$

$$(\text{div}(u - r_h u), q_h) = 0, \quad \forall q \in M, \quad q_h \in M_h \quad (0.0.64)$$

$$\|A^{1/2}(u - r_h u)\| \leq ch \|Au\|, \|p - \pi_h p\| \leq ch \|p\|_1, \quad \forall p \in H^1(\Omega) \cap M \quad (0.0.65)$$

De plus, nous supposons l'inégalité inverse

$$\|A^{1/2} v_h\| \leq ch^{-1} \|\mathbf{v}_h\|, \quad \forall \mathbf{v}_h \in X_h \quad (0.0.66)$$

et la condition inf-sup

$$\|q\| \leq c \sup_{\mathbf{v}_h \in X_h} \frac{(\text{div}(\mathbf{v}_h), q_h)}{\|A^{1/2} \mathbf{v}_h\|}, \quad \forall q_h \in M_h \quad (0.0.67)$$

Théorème 0.0.12. *Supposons que (A0) et (A1) et $\varepsilon c_0 \leq 1$, $\Delta t k_0 \leq 1$ sont vérifiées. Alors on a*

$$\|\mathbf{u}_{\varepsilon h}^n\|^2 + \Delta t \sum_{n=1}^m \|A^{1/2} u_{\varepsilon h}^n\|^2 + a \Delta t \sum_{n=1}^n \|\mathbf{u}_{\varepsilon h}^n\|_{L^{\alpha+2}}^{\alpha+2} \leq C, \quad (0.0.68)$$

$$\begin{aligned} \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 + \Delta t \sum_{n=1}^m \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 &\leq C + Ch^{-2} \Delta t \sum_{n=1}^m \|A^{1/2} (\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n)\|^2 \|\mathbf{u}_{\varepsilon h}^n\|^2, \\ &+ Ch^{-2} \Delta t \sum_{n=1}^m \|\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2, \end{aligned} \quad (0.0.69)$$

pour tout $1 \leq m \leq N$.

avec $(u_{\varepsilon h}^n, p_{\varepsilon h}^n)$ est solution du problème approximé suivant

$$\begin{aligned} (d_t u_{\varepsilon h}^n, v_h) + a(u_{\varepsilon h}^n, v_h) - d(v_h, p_{\varepsilon h}^n) + d(u_{\varepsilon h}^n, q_h) + \varepsilon(p_{\varepsilon h}^n, q_h) + \tilde{b}(u_{\varepsilon h}^n, u_{\varepsilon h}^n, v_h) \\ + c_\alpha(u_{\varepsilon h}^n, v_h) = (f(t_n), v_h), \quad \forall (v_h, q_h) \in X_h \times M_h \end{aligned} \quad (0.0.70)$$

avec $u_{\varepsilon h}^0 = r_h u_0$, $p_{\varepsilon h}^0 = 0$. et (X_h, M_h) sont les espaces d'éléments finis associés à X et M. Dans la dernière section, nous implémentons une méthode d'élément finis sous Freefem++ pour résoudre le problème perturbé (0.0.8), combiné à une méthode de gradient projeté pour la résolution du problème d'optimisation de forme afin d'en déduire une forme optimale de la structure de type VSF.

Contents

1	General introduction	35
2	Optimum Hydrodynamic Design using Shallow Water Equations	41
2.1	Introduction	41
2.2	Mathematical model	42
2.3	Adjoint system	45
2.4	Shape derivative	47
2.5	Analysis of the adjoint problem	52
2.5.1	Pseudo-differential tools	52
2.5.2	Main result	54
2.5.3	Existence and Uniqueness using the pseudo-differential calculus	55
2.6	Numerical methods	62
2.6.1	Finite volume method	62
2.6.2	A gradient free algorithm	63
2.6.3	Spectral projected gradient	66
2.6.4	Design examples	69
2.7	Conclusion	73
3	Shape optimization of fishways in multilayer flow based on optimal control theory	75
3.1	Introduction	75
3.2	Multilayer approach	76
3.2.1	Preliminaries	76
3.2.2	Weak solution with discontinuities	78
3.2.3	Vertical velocity	80
3.3	Hydrostatic pressure multilayer model	81
3.4	Analysis of the 2D multilayer system	82
3.4.1	Mass conservation	83
3.4.2	Momentum conservation	83
3.4.3	Existence and uniqueness of solution model	85

3.5	3-D Multilayer system and fishway shape optimization	105
3.5.1	3-D Multilayer system	105
3.5.2	Optimal problem	107
3.5.3	Adjoint system	109
3.5.4	Domain derivative	112
3.6	Numerical resolution	117
3.6.1	Finite volume method	118
3.6.2	A gradient free algorithm	119
3.6.3	Spectral projected gradient	121
3.6.4	Numerical results	123
3.7	Conclusion	130
4	Analysis and finite element approximation of optimal control problem based on porous media model	131
4.1	Introduction	131
4.2	Mathematical model	133
4.2.1	Preliminaries	133
4.2.2	Perturbed equation	140
4.2.3	Error estimates	142
4.3	Control problem	147
4.3.1	Adjoint system	148
4.3.2	Shape gradient	150
4.4	Finite element method	153
4.5	Numerical resolution	163
4.5.1	Numerical implementation of optimization problem	163
4.5.2	Numerical examples	165
4.6	Conclusion	170
5	Hydrodynamic design optimization using non stationary porous media model	173
5.1	Introduction	173
5.2	Mathematical model	175
5.2.1	Preliminaries	175
5.3	Error Analysis	183
5.3.1	Linearised problem	183
5.3.2	Error estimates for the nonlinear perturbed problem	187
5.4	Control problem	191
5.4.1	Adjoint equation	191
5.4.2	Shape gradient	193

5.5	Discrete finite element method	196
5.5.1	Preliminaries	196
5.5.2	Finite element penalty method	201
5.6	Numerical resolution	205
5.6.1	Numerical methods	205
5.6.2	Numerical results	207
5.7	Conclusion	211

Chapter 1

General introduction

A fluid is a deformable material medium in the case of the action of an external force. It consists of free particles which move relative to each other. For example, liquids are not very compressible fluids; they maintain their volume whatever their forms and present a clean surface. Contrariwise, the gases are compressible fluids. A fluid in a liquid or gaseous state is characterised by a density and a viscosity. The density represents the number of molecules per volume unit. The viscosity characterises the velocity of the fluid movement. There are many types of flow. In particular, we cite the laminar and turbulent flow. The first type represents parallel straight current lines. The second one is a disorganised flow with the presence of turbulences.

Among the applications of these water flow models are the problems associated to the optimization of fishways structure. In fact, some types of fish accomplish their daily or annual migrations and for long or short distances to complete their life cycles. We are interested in diadromous fishes which migrate between sea and fresh water. These migratory fish species are threatened by urbanization, agriculture, or dams.

Dams are built through centuries by men for agriculture, navigation, electricity or economic reasons. They represent impassable barriers for migratory species for their reproduction or feeding areas during the rise of the rivers. To meet this challenge, man thought to build fishways for permitting fish to pass with comfortable conditions. Fishways (fish ladder, or fish passage) are hydraulic structures placed near from dams or weirs to allow migratory fishes to cross to upstream river. Most of these channels are built with a slope to the ground and they are divided into some pools. The speed in the channel pools should not be very high to deplete the fish or prevent them from continuing their journeys to the breeding areas. Several types of fish passage are known nowadays. The best references are Clay [1] for the pool and weir type, Katopodis et al. [2] for the Denil fishways and for the vertical slot type Rajaratnam et al. [9].

In this thesis, we are involved in the third type of fishways: vertical slot fishway (VSF). VSF has an area of low velocity for fish to rest during upstream passage, a good energy dissipation. This makes them suitable to sites where fish passage is needed at different times of the year and for a maximum number of small species. Naturally, the aim is to find optimal forms of VSF in order to facilitate the crossing of migratory species through dams in convenient conditions.

The goal of a shape design problem is to find an optimal shape domain in a set of admissible domains when a cost functional achieves a minimum or maximum on this domain. The research of shape design is part of optimal control field governed by partial differential equations and it has variety of applications in engineering. In particular, the design of bridges and dams. Recently, Many researcher and mathematicians have been interested in this field of research [8, 49, 59, 64, 60, 58].

Our purpose is to investigate these structures with different flow models. The first one is based on the shallow water equations. The shallow water systems are a hyperbolic systems, introduced at the end of the nineteenth century by A.J.C Barré de Saint-Venant since 1871 [55]. These equations are derived from the incompressible Navier-Stokes equations using the Boussinesq approximation, in which the density variations are neglected everywhere except in the gravity term. They are used to model free-surface geophysical fluids in shallow flows. Due to their numerical efficiency, They are now widely used for the simulation of many current phenomena: environmental pollution, natural disasters, and climate change.

The aim here is to assess the possibility of using a two-dimensional shallow water model to compute the flow pattern in vertical slot fish ladder and deduce an optimal structure allows fish to cross the obstacle in a convenient conditions. We begin by giving a mathematical formulation of channel consisting of ten pools with vertical slots for obtaining a mean flow suitable for large variety of migratory species. The 2-D shallow-water equations with source terms may be written as

$$\begin{aligned} \frac{\partial H}{\partial t} + \vec{\nabla} \cdot \vec{Q} &= 0 & \text{in } \Omega \times (0, T) \\ \frac{\partial \vec{Q}}{\partial t} + \vec{\nabla} \cdot \left(\vec{Q} \otimes \frac{\vec{Q}}{H} \right) + gH \vec{\nabla} (H - \eta) &= \vec{f} & \text{in } \Omega \times (0, T) \end{aligned} \quad (1.0.1)$$

Where H is the water depth; $\vec{u} = (u_1, u_2)$ is the velocity vector; u_1 and u_2 are the x and y components of flow velocity, respectively; $Q = (u_1 H, u_2 H)$ is the unit-width discharge; η is the bottom geometry; g is the gravitational acceleration and \vec{f} represents all effects of bottom friction and atmospheric pressure. We complete the 2-D Saint-Venant system

with the following initial and boundary conditions

$$\begin{aligned}
H(0) &= H_0, \quad \vec{Q}(0) = \vec{Q}_0 && \text{in } \Omega \\
\vec{Q} \cdot \vec{n} &= 0, \quad \text{curl} \left(\frac{\vec{Q}}{H} \right) = 0 && \text{on } \gamma_0 \times (0, T) \\
\vec{Q} &= Q_1 \vec{n} && \text{on } \gamma_1 \times (0, T) \\
H &= H_2 && \text{sur } \gamma_2 \times (0, T)
\end{aligned} \tag{1.0.2}$$

where γ_0 is lateral boundary, γ_1 et γ_2 are the inflow and outflow boundary respectively. The optimum shape design involve an objective function linked to fish's swimming aptitudes.

Firstly, we compute its shape derivative in term of state variables and adjoint variables according to the derivation of an adjoint system. Existence and uniqueness results for the adjoint system are stated using the Lax symbolic symmetrizer for hyperbolic systems and pseudo-differential techniques. A finite volume method is set up to solve the state problem (1.0.1) combined with two minimizing algorithms (Free gradient algorithm and Spectral projected gradient) for the resolution of the shape optimization problem.

The second sort of shallow flow based model is a multilayer model. The multilayer model avoids the expensive Navier-Stokes equations when the flow is still within the shallow regime.

Starting from the non stationary Navier-Stokes equations

$$\begin{aligned}
\partial_t \rho + \nabla \cdot (\rho \vec{u}) &= 0 && \text{in } \Omega \times [0, T] \\
\partial_t (\rho \vec{u}) + \nabla \cdot (\rho \vec{u} \otimes \vec{u}) &= \nabla \cdot \Sigma_T + \rho \mathbf{g} && \text{in } \Omega \times [0, T]
\end{aligned} \tag{1.0.3}$$

with the stress tensor symbol Σ_T is given by

$$\Sigma_T = -p \mathbf{Id} + \mu (\nabla \vec{u} + (\nabla \vec{u})')$$

and where $\mathbf{g} = (\mathbf{0}, -g)' \in \mathbb{R}^d$ is the gravity acceleration, and μ is a given viscosity.

Following the methodology initiated by Audusse and al. [14]. Then the water height is divide into a number of layers where the hypothesis of Saint-Venant can be applied in each layer. By considering a non constant vertical velocities, the vertical profile of the horizontal velocity is described by the following multilayer system

$$\begin{aligned}
\partial_t h_\alpha + \nabla_{x \cdot} (h_\alpha \vec{u}_{H,\alpha}) &= G_{\alpha+\frac{1}{2}} - G_{\alpha-\frac{1}{2}}. \\
\rho \partial_t (h_\alpha \vec{u}_{H,\alpha}) + \rho \nabla_{x \cdot} (h_\alpha \vec{u}_{H,\alpha} \otimes \vec{u}_{H,\alpha}) &- \nabla_{x \cdot} (\mu h_\alpha (\nabla_x \vec{u}_{H,\alpha} + (\nabla_x \vec{u}_{H,\alpha})')) \\
= \frac{1}{2} \rho G_{\alpha+\frac{1}{2}} (\vec{u}_{H,\alpha+1} + \vec{u}_{H,\alpha}) &- \frac{1}{2} \rho G_{\alpha-\frac{1}{2}} (\vec{u}_{H,\alpha} + \vec{u}_{H,\alpha-1}) \\
-\mu (\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}}) - \rho g h_\alpha \nabla_x (z_B + h) &- h_\alpha \nabla_x p_S.
\end{aligned} \tag{1.0.4}$$

where $G_{\alpha+1/2}$ is the normal mass flux at the interface between layers $\alpha - 1$ and α .

We prove the existence and uniqueness of a strong solution for the 2-D multilayer system. The proof is done in several stages: the analysis of the linearized problem, the obtaining of an a priori estimates and at the end a Picard scheme for the transition to the nonlinear case. In order to deal with the optimum shape design problem, we proceed with the generalization of the multilayer procedure to the 3-D hydrostatic Navier-Stokes equations and extract a 3-D multilayer system. The multilayer Saint-Venant description of flow allows to derive an optimal fishway structures with respect to the shape of the variable domain for some given cost functionals modeling the ease of fish passage over the structure to their breeding, growing or feeding areas. We derive an adjoint system related to the obtained 3-D multilayer model and express the domain derivative of the cost function in terms of the state and adjoint variables. We present two minimizing procedures: the first one is a free gradient algorithm and the second is a gradient type algorithm namely spectral projected gradient.

The third water flow is based on the stationary porous media model. Fluid flow in porous media are usually described by Darcy Brinkman Forchheimer equations [29]:

$$\begin{aligned}
 -\nu\Delta u + (u.\nabla)u + a |u|^\alpha u + \nabla p &= f \text{ in } \Omega; & \nabla.u &= 0 \text{ in } \Omega \\
 u &= 0 & & \text{on } \Gamma_0 \\
 u &= g & & \text{on } \Gamma_1 \\
 \nu\frac{\partial u}{\partial n} - p\mathbf{n} &= 0 & & \text{on } \Gamma_2
 \end{aligned} \tag{1.0.5}$$

where u and p present respectively the velocity and the pressure. ν is Brinkman coefficient, a defines Forchheimer coefficient. f denotes the exterior forces and $\alpha \in [1, 2]$ is a real number chosen appropriately throughout this work. Γ_0 is the lateral boundary, Γ_1 is inflow, top and bottom flow boundaries, and Γ_2 is the outflow boundary.

Firstly, we investigated the well posedness aspect of the problem. We show the existence and uniqueness of a weak solution of the boundary problem (1.0.5).

The importance of the treatment of the incompressibility constraint has been recognized since a long time. Among the method adopted to relax the constraint $\nabla.u = 0$, we cite

1. The penalty method. We refer here to J. Shen [19], R. Temam [36]

$$\nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon = 0 \quad \text{in } \Omega \times [0, T]$$

2. Artificial compressibility, we mention here the works of Chorin [20] and of R. Temam [37]

$$\nabla \cdot u^\varepsilon + \varepsilon p_t^\varepsilon = 0 \quad \text{in } \Omega \times [0, T]$$

3. The projection method

$$\nabla \cdot u^\varepsilon - \varepsilon \Delta p^\varepsilon = 0 \quad \text{in } \Omega \times [0, T], \quad \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \quad \text{in } \partial\Omega \times [0, T]$$

introduced by Temam [36] and Chorin [20, 21] and has a high computational cost.

4. The pseudocompressibility method

$$\nabla \cdot u^\varepsilon - \varepsilon \Delta p_t^\varepsilon = 0 \quad \text{in } \Omega \times [0, T], \quad \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \quad \text{in } \partial\Omega \times [0, T]$$

In this work, a penalty method is used to relax the constraint of incompressibility for the velocity. The penalized problem related to the system (1.0.5) writes

$$\begin{aligned} -\nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f & \text{in } \Omega \\ \nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 & \text{in } \Omega \\ u^\varepsilon &= 0 & \text{on } \Gamma_0 \\ u^\varepsilon &= g & \text{on } \Gamma_1 \\ \nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 & \text{on } \Gamma_2 \end{aligned} \tag{1.0.6}$$

In order to assess the usefulness of the approach, we establish some a priori estimates helpful to build the error analysis. The existence of a weak solution for the penalized problem (1.0.6) is proved using a Faedo-Galerkin method. The design objective is to minimize a cost function. We express the shape gradient of the cost function in terms of the velocity value as a state variable, the adjoint variables and the unit normal vector to the boundary of the domain.

Nowadays, several works are interested in the finite element method employed in almost domain of engineering and numerical analysis [62, 57, 63]. A discrete finite element method is introduced to approximate the solution of the penalized problem. We establish a priori estimate which proves the convergence of the approximate penalized solution to the solution of the Navier Stokes Forchheimer equation. The optimization procedure is implemented using the continuous adjoint method and the finite element method.

The last type of flow model is described by the non stationary Navier-Stokes Forchheimer equations

$$\begin{aligned}
u_t - \nu \Delta u + (u \cdot \nabla)u + a |u|^\alpha u + \nabla p &= f && \text{in } \Omega \\
\nabla \cdot u &= 0 && \text{in } \Omega \\
u &= 0 && \text{on } \Gamma_0 \\
u &= g && \text{on } \Gamma_1 \\
\nu \frac{\partial u}{\partial n} - p \mathbf{n} &= 0 && \text{on } \Gamma_2 \\
u|_{t=0} &= u_0 &&
\end{aligned} \tag{1.0.7}$$

We begin by showing the existence and uniqueness of solution for the initial problem (1.0.7). Error estimates for the velocity and the pressure are established via the energy method.

Next we treat the penalized problem along two parts; the first one is the study of the linearized problem and the second is the treatment of the nonlinear problem.

The perturbed problem is written as

$$\begin{aligned}
u_t^\varepsilon - \nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla)u^\varepsilon + \frac{1}{2}(\nabla \cdot u^\varepsilon)u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f && \text{in } \Omega \\
\nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 && \text{in } \Omega \\
u^\varepsilon &= 0 && \text{on } \Gamma_0 \\
u^\varepsilon &= g && \text{on } \Gamma_1 \\
\nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 && \text{on } \Gamma_2 \\
u^\varepsilon|_{t=0} &= u_0 &&
\end{aligned} \tag{1.0.8}$$

We introduce a time discretization by the use of a backward Euler scheme combined with fully discrete finite element method to approximate the penalized problem and establish an error estimate for the velocity and the pressure which will be used to show the convergence of the approximate solution to the solution of the initial problem.

The shape optimization problem is to find the shape which is optimal in that it minimizes a comfort cost functional. We derive the adjoint system associated to the penalized problem. We compute the gradient in terms of state and adjoint variables. The optimization procedure is implemented using the continuous adjoint method and the finite element method.

Chapter 2

Optimum Hydrodynamic Design using Shallow Water Equations

2.1 Introduction

Many species of salmon, shad, giant catfishes, dorado, sturgeons and eel migrate between the sea and the rivers to complete their life cycle. Free migration routes for fish are crucial to their survival. We take interest in diadromous fish species which immigrate between sea and fresh water.

We distinguish some types of diadromous fish: Anadromous fish (as salmon, smelt, American shad, hickory shad, striped bass, lamprey, gulf sturgeon,...) which live and grow in the salt water, and migrate to freshwater rivers and lakes to reproduce. The Anadromous fish eventually return to freshwater to spawn. About half of all diadromous fish in the world are Anadromous. Adult Catadromous (American eel, European eel, inanga, shortfin eel, longfin eel) live in fresh water, then migrate to the sea for breeding. After hatching, they migrate back to freshwater where they stay until growing into adults. Catadromous fish undertake a great migration from freshwater to spawn in the marine, and they die there due to the effort made for migration. About one quarter of all diadromous fish in the world are catadromous. Amphidromous species (bigmouth sleeper, mountain mullet, sirajo goby, river goby, torrentfish, Dolly Varden) migrate between estuaries and coastal rivers and streams (in both directions). Amphidromous fish live in freshwater for breeding and they leave to the marine for feeding and growing.

The presence of dams without fish passes appears to be a major contributing factor in the decline of migratory species. Fishways have been designed to provide safe passage for migratory species inhabiting the river to get pass towards their breeding or feeding areas. The utility of such systems has been demonstrated around the world. The best general reference here is Clay [1] for the pool and weir type, Katopodis et al. [2] for the Denil

fishways and for the vertical slot type Rajaratnam et al. [9].

Vertical slot ladders are quite common and use a large narrow slot to control water flow and depths in the pools between slots. This allows fish to swim upstream without leaping over an obstacle. This design reasonably handles the seasonal fluctuation in water levels and is not sensitive to impoundment or upstream water surface elevation changes.

The chapter is devoted to the study of vertical slots fishway. The aim of this work is to assess the possibility of using a two-dimensional shallow water model to compute the flow pattern in vertical slot fish ladder and deduce an optimal structure allows fish to cross the obstacle in a convenient conditions.

The reminder of the chapter is organized as follows. The next section is dedicated to the mathematical formulation and the introduction of the objective function related to the optimal shape design. In sections 3 and 4, we use a method of transport to describe a variational domain in the optimization process. Then the adjoint equations are derived by employing the differentiability of an saddle point problem which includes a Lagrange multiplier function. Section 5 is assigned to the well-posedness of the adjoint system using an approach based on pseudo-differential calculus. The last section provides numerical methods (finite volume scheme combined to minimizing algorithms) with some design examples to prove that our approach could be very useful for the practical purpose.

2.2 Mathematical model

A vertical-slot fishway, shown in Figure 1, is a rectangular channel $\Omega \subset \mathbb{R}^2$ with a sloping floor that is divided into 10 pools by baffles. The pools have a double function: they ensure a proper dissipation of the energy of water flowing through the fish pass, and provide resting areas for the fish. It is worth pointing out that the geometric features of each pool are with a width of $0.97m$, a length of $1.213m$, also two transition pools, one at the beginning and other at the end of the channel with the same width and a length of $1.5m$. Inside each pool, two baffles are built. They have a width of $2r = 0.061m$ and are vertical to the lateral fishway boundary. The channel is constructed with a slope relative to the ground.

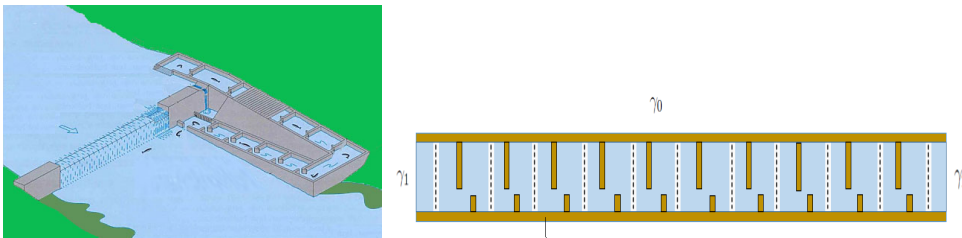


Figure 2.1: Fishway and ground plan Ω : each pool is designed by dashed lines

The shallow water equations are used to simulate a variety of problems related to environment and coastal engineering. These equations can be obtained by integrating the incompressible Navier-Stokes equations in depth and taking into account the kinematic and kinetic boundary conditions. The 2-D shallow-water equations with source terms may be written as

$$\begin{aligned} \frac{\partial H}{\partial t} + \vec{\nabla} \cdot \vec{Q} &= 0 & \text{in } \Omega \times (0, T) \\ \frac{\partial \vec{Q}}{\partial t} + \vec{\nabla} \cdot \left(\vec{Q} \otimes \frac{\vec{Q}}{H} \right) + gH \vec{\nabla} (H - \eta) &= \vec{f} & \text{in } \Omega \times (0, T) \end{aligned} \quad (2.2.1)$$

Where H is the water depth; $\vec{u} = (u_1, u_2)$ is the velocity vector; u_1 and u_2 are the x and y components of flow velocity, respectively; $Q = (u_1 H, u_2 H)$ is the unit-width discharge; η is the bottom geometry; g is the gravitational acceleration and \vec{f} represents all effects of bottom friction and atmospheric pressure.

We introduce three parts of the boundary of Ω : the lateral boundary denoted by γ_0 , the inflow boundary denoted by γ_1 , and the outflow boundary denoted by γ_2 . We take for \vec{n} the unit outer normal vector to boundary. To obtain a well-posed problem, we add to this system an initial and boundary conditions defined by:

$$\begin{aligned} H(0) = H_0, \quad \vec{Q}(0) = \vec{Q}_0 \quad \text{in } \Omega, \quad \vec{Q} \cdot \vec{n} = 0 \quad \text{curl} \left(\frac{\vec{Q}}{H} \right) = 0 \quad \text{on } \gamma_0 \times (0, T), \\ \vec{Q} = Q_1 \vec{n} \quad \text{on } \gamma_1 \times (0, T), \quad H = H_2 \quad \text{on } \gamma_2 \times (0, T) \end{aligned} \quad (2.2.2)$$

The geometry of the vertical slot based on the use of guide elements to lead smooth hydraulic flow into the next slot. The positioning of the guide elements was carried out at two different locations, a and b , which configure the shape of the fish ladder Ω (Figure 2.2).

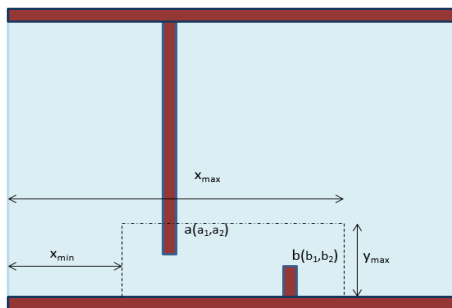


Figure 2.2: Prototype geometry: details of slot and pool

The design variables a and b are subject to constraints in order to ensure a positive influence on the flow in the individual pools. These constraints are formulated as

$$\begin{aligned} \frac{1}{4} 1.213 &\leq y_1, y_3 \leq \frac{3}{4} 1.213 \\ 0 &\leq y_2, y_4 \leq \frac{1}{4} 0.97 \end{aligned} \quad (2.2.3)$$

In order to provide a comfort conditions during the fish passage and permit to a maximum number of fishes to pass to the river upstream, the following constraints are introduced

$$y_3 - y_1 \geq d_1 = 0.1 \quad (2.2.4)$$

$$y_2 - y_4 \geq d_2 = 0.05$$

A shape optimization problem consists in the minimization of a functional $J \in \mathbf{R}$, also called cost function depending on the design variables a and b defining the shape within the admissible constants defining the admissible set X . We have a direct calculation loop for the functional: from a parametrization (a, b) we define a domain $\Omega(a, b)$ on which we compute the state equation solution $W = (H, Q)$ and the cost function $J(\Omega(a, b))$:

$$J : x = (a, b) \in X \longrightarrow \Omega(x) \longrightarrow W(\Omega(x)) \longrightarrow J(x, \Omega(x), W(\Omega(x)))$$

We consider that the shape of the structure is efficient if the associated energy dissipation leads to a velocity of water close to a target velocity \vec{v} related to fishes species and minimizing the flow turbulence in the channel.

The target velocity is given by

$$\vec{v}(x_1, x_2) = \begin{cases} (c, 0) & \text{if } x_2 \leq \frac{1}{3} 0.97 \\ (0, 0) & \text{otherwise} \end{cases} \quad (2.2.5)$$

Thus, we want to minimize the following cost function

$$J = \frac{1}{2} \int_0^T \int_{\Omega} \| \vec{u} - \vec{v} \|^2 dxdt + \frac{\alpha}{2} \int_0^T \int_{\Omega} |rot(\vec{u})|^2 dxdt \quad (2.2.6)$$

with $\alpha \geq 0$ is the vorticity parameter, and $\vec{u} = \frac{\vec{Q}}{H}$ where (H, \vec{Q}) is solution of the shallow water system (2.2.1) with the initial and boundary conditions (2.2.2).

2.3 Adjoint system

We consider (p, \vec{r}) two test functions space in $L^\infty(0, T; H^2(\omega))^2$. The variational formulation of the state system (2.2.1) is given by

Find $(H, \vec{Q}) \in L^\infty(0, T; H^2(\omega))^2$ such that

$$A(\omega, H, \vec{Q}; p, \vec{r}) = L(\omega; p, \vec{r}) \quad \forall (p, \vec{r}) \in L^\infty(0, T; H^2(\omega))^2$$

where

$$A(\omega, H, \vec{Q}; p, \vec{r}) = \int_\omega \int_0^T \left\{ \frac{\partial H}{\partial t} + \vec{\nabla} \cdot \vec{Q} \right\} \cdot p + \int_\omega \int_0^T \left\{ \frac{\partial \vec{Q}}{\partial t} + \vec{\nabla} \cdot \left(\frac{\vec{Q}}{H} \otimes \vec{Q} \right) + gH\vec{\nabla}(H - \eta) \right\} \vec{r}$$

$$L(\omega; p, \vec{r}) = \int_\omega \int_0^T \vec{f} \cdot \vec{r}$$

We reformulate the objective function as $J(\omega, H, \vec{Q}) = j(\omega)$, such that

$$J(\omega, H, \vec{Q}) = \frac{1}{2} \int_\omega \int_0^T \left\| \frac{\vec{Q}}{H} - \vec{v} \right\|^2 + \frac{\alpha}{2} \int_\omega \int_0^T \left| \text{curl} \left(\frac{\vec{Q}}{H} \right) \right|^2$$

Theorem 2.3.1. *let $(H, \vec{Q}) \in L^2(0, T; \mathbb{R}) \times H^2(0, T; \omega)$ solution of the system (2.2.1), and $(p, \vec{r}) \in L^2(0, T; \mathbb{R}) \times H^2(0, T; \omega)$ two test functions, then the adjoint system associated to the state system (2.2.1) take the form*

$$\begin{cases} -\frac{\partial p}{\partial t} + \frac{1}{H^2} (\vec{Q} \cdot \vec{\nabla}) \vec{r} \cdot \vec{Q} - gH(\vec{\nabla} \cdot \vec{r}) - g\vec{\nabla} \eta \cdot \vec{r} = -\left(\frac{\vec{Q}}{H} - \vec{v} \right) \frac{\vec{Q}}{H^2} - \alpha \vec{\text{curl}}(\text{curl}(\frac{\vec{Q}}{H})) \cdot \frac{\vec{Q}}{H^2} \\ -\frac{\partial \vec{r}}{\partial t} - \vec{\nabla} p - \frac{1}{H} (\vec{Q} \cdot \vec{\nabla}) \vec{r} - \frac{1}{H} (\vec{\nabla} \vec{r})^t \vec{Q} = \frac{1}{H} \left(\frac{\vec{Q}}{H} - \vec{v} \right) + \frac{\alpha}{H} \vec{\text{curl}}(\text{curl}(\frac{\vec{Q}}{H})) \end{cases} \quad (2.3.1)$$

with final and boundary conditions

$$\begin{cases} p(T) = 0 & \text{in } \omega \\ \vec{r}(T) = 0 & \text{in } \omega \\ \vec{r} \cdot \vec{n} = 0 & \text{in } \gamma_0 \times (0, T) \\ (gH - \frac{Q_1^2}{H^2}) \vec{r} \cdot \vec{n} = 0 & \text{in } \gamma_1 \times (0, T) \\ \left\{ p + \frac{1}{H_2} (\vec{Q} \cdot \vec{r}) \right\} \vec{n} + \frac{1}{H_2} (\vec{Q} \cdot \vec{n}) \vec{r} - \frac{\alpha}{H_2} \text{curl} \left(\frac{\vec{Q}}{H_2} \right) \vec{r} = 0 & \text{in } \gamma_2 \times (0, T) \end{cases} \quad (2.3.2)$$

Proof. We consider the Lagrangian function (for details see [4],[7] given by

$$L_{ag}(\omega, H, \vec{Q}; p, \vec{r}) = J(\omega, H, \vec{Q}) - A(\omega, H, \vec{Q}; p, \vec{r}) + L(\omega, H, \vec{Q}; p, \vec{r})$$

By integration by parts of the function $A(\omega, H, \vec{Q}; p, r)$, deriving in H the function A and J and using the adjoint state ($\partial_H L_{ag}(\omega, H, \vec{Q}; p, \vec{r}) = 0$), we obtain the following equation: in $\omega \times (0, T)$

$$-\frac{\partial p}{\partial t} + \frac{1}{H^2}(\vec{Q} \cdot \vec{\nabla})\vec{r} \cdot \vec{Q} - gH(\vec{\nabla} \cdot \vec{r}) - g\vec{\nabla} \eta \cdot \vec{r} = -\left(\frac{\vec{Q}}{H} - v\right)\frac{\vec{Q}}{H^2} - \alpha \vec{curl}(\vec{curl}\left(\frac{\vec{Q}}{H}\right)) \cdot \frac{\vec{Q}}{H^2}$$

with the following final and boundary conditions:

$$\begin{cases} p(T) = 0 & \text{in } \omega \\ \vec{r} \cdot \vec{n} = 0 & \text{in } \gamma_0 \times (0, T) \\ (gH - \frac{Q_1^2}{H^2})\vec{r} \cdot \vec{n} = 0 & \text{in } \gamma_1 \times (0, T) \end{cases}$$

By integration by parts of the function $A(\omega, H, \vec{Q}; p, r)$, deriving in \vec{Q} the function A and J and using the adjoint state ($\partial_{\vec{Q}} L_{ag}(\omega, H, \vec{Q}; p, \vec{r}) = 0$), we obtain the following equation: in $\omega \times (0, T)$

$$-\frac{\partial \vec{r}}{\partial t} - \vec{\nabla} p - \frac{1}{H}(\vec{Q} \cdot \vec{\nabla})\vec{r} - \frac{1}{H}(\vec{\nabla} \vec{r})^t \vec{Q} = \frac{1}{H}\left(\frac{\vec{Q}}{H} - \vec{v}\right) + \frac{\alpha}{H} \vec{curl}(\vec{curl}\left(\frac{\vec{Q}}{H}\right))$$

with the following final and boundary conditions:

$$\begin{cases} \vec{r}(T) = 0 & \text{in } \omega \\ \vec{r} \cdot \vec{n} = 0 & \text{in } \gamma_0 \times (0, T) \\ \left\{p + \frac{1}{H_2}(\vec{Q} \cdot \vec{r})\right\}\vec{n} + \frac{1}{H_2}(\vec{Q} \cdot \vec{n})\vec{r} - \frac{\alpha}{H_2} \vec{curl}\left(\frac{\vec{Q}}{H_2}\right)\vec{r} = 0 & \text{in } \gamma_2 \times (0, T) \end{cases}$$

Finally, the adjoint system is given by:

$$\begin{cases} -\frac{\partial p}{\partial t} + \frac{1}{H^2}(\vec{Q} \cdot \vec{\nabla})\vec{r} \cdot \vec{Q} - gH(\vec{\nabla} \cdot \vec{r}) - g\vec{\nabla} \eta \cdot \vec{r} = -\left(\frac{\vec{Q}}{H} - v\right)\frac{\vec{Q}}{H^2} - \alpha \vec{curl}(curl(\frac{\vec{Q}}{H})).\frac{\vec{Q}}{H^2} \\ -\frac{\partial \vec{r}}{\partial t} - \vec{\nabla} p - \frac{1}{H}(\vec{Q} \cdot \vec{\nabla})\vec{r} - \frac{1}{H}(\vec{\nabla} \vec{r})^t \vec{Q} = \frac{1}{H}\left(\frac{\vec{Q}}{H} - \vec{v}\right) + \frac{\alpha}{H} \vec{curl}(curl(\frac{\vec{Q}}{H})) \end{cases} \quad (2.3.3)$$

with final and boundary conditions:

$$\begin{cases} p(T) = 0 & in \ \omega \\ \vec{r}(T) = 0 & in \ \omega \\ \vec{r} \cdot \vec{n} = 0 & in \ \gamma_0 \times (0, T) \\ (gH - \frac{Q_1^2}{H^2})\vec{r} \cdot \vec{n} = 0 & in \ \gamma_1 \times (0, T) \\ \{p + \frac{1}{H_2}(\vec{Q} \cdot \vec{r})\}\vec{n} + \frac{1}{H_2}(\vec{Q} \cdot \vec{n})\vec{r} - \frac{\alpha}{H_2} curl(\frac{\vec{Q}}{H_2})\vec{r} = 0 & in \ \gamma_2 \times (0, T) \end{cases} \quad (2.3.4)$$

□

2.4 Shape derivative

In this section we give a description of the domain derivative of the objective function j . Let ω_0 be a fixed domain, and we suppose that it is a bounded open subset of \mathbb{R}^2 with Lipschitz boundary. We denote by $Lip(\omega_0; \mathbb{R}^2)$ the Lipschitz applications domain, which is defined by

$$\begin{aligned} Lip(\omega_0; \mathbb{R}^2) &= \{\phi : \omega_0 \rightarrow \mathbb{R}^2 \ / \ \exists k \geq 0 \text{ such that} \\ &\quad \|\phi(x) - \phi(y)\| \leq k \|x - y\|, \ \forall (x, y) \in \omega_0 \times \omega_0\} \end{aligned}$$

We denote by O_0 the subset of bi-Lipschitz homeomorphisms of ω_0 . It is an open subset of $Lip(\omega_0; \mathbb{R}^2)$ and it is defined by

$$O_0 = \{\psi : \omega_0 \rightarrow \psi(\omega_0), \text{ bijective} \ / \ \psi \in Lip(\omega_0; \mathbb{R}^2), \psi^{-1} \in Lip(\psi(\omega_0); \mathbb{R}^2)\}$$

The admissible domain denoted by X_0 is a set of domains ω which are bound by bijective Lipschitz functions to the fixed domain ω_0 . It is given as follow

$$X_0 = \{\omega = \psi(\omega_0) \ / \ \psi \in O_0\} \quad (2.4.1)$$

Theorem 2.4.1. *Let $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$, and if the functions A , J and L are differentiable with respect to the domain ω in the admissible set X_0 , then the domain derivative of j at $\omega \in X_0$ takes the form*

$$\begin{aligned} \frac{\partial}{\partial \omega} j(\omega) \cdot \vec{\mathbf{V}} &= \frac{\partial}{\partial \omega} J(\omega; H, \vec{Q}) \cdot \vec{\mathbf{V}} - \frac{\partial}{\partial \omega} A(\omega; H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} \\ &+ \frac{\partial}{\partial \omega} L(\omega; p, \vec{r}) \cdot \vec{\mathbf{V}} \end{aligned} \quad (2.4.2)$$

where

$$\begin{aligned} \frac{\partial}{\partial \omega} A(\omega; H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} &= \int_0^T \int_{\omega} \frac{\partial H}{\partial t} p(\vec{\nabla} \cdot \vec{\mathbf{V}}) + \int_0^T \int_{\omega} (\vec{\nabla} \cdot \vec{Q}) p(\vec{\nabla} \cdot \vec{\mathbf{V}}) \\ &- \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}})^t : \vec{\nabla} \vec{Q} p + \int_0^T \int_{\omega} \frac{\partial \vec{Q}}{\partial t} \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}) \\ &+ \int_0^T \int_{\omega} (\vec{\nabla} \cdot \vec{Q}) \frac{\vec{Q}}{H} \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}) - \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}})^t : \vec{\nabla} \vec{Q} \frac{\vec{Q}}{H} \cdot \vec{r} \\ &+ \int_0^T \int_{\omega} (\vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}) - \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}} \vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r} \\ &+ \int_0^T \int_{\omega} g H \vec{\nabla} H \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}) - \int_0^T \int_{\omega} g H (\vec{\nabla} \vec{\mathbf{V}})^t \vec{\nabla} H \cdot \vec{r} \\ &- \int_0^T \int_{\omega} g H \vec{\nabla} \eta \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}), \quad \text{for } \vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2) \end{aligned}$$

$$\frac{\partial}{\partial \omega} L(\omega; p, \vec{r}) \cdot \vec{\mathbf{V}} = \int_0^T \int_{\omega} \vec{f} \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}), \quad \text{for } \vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$$

$$\frac{\partial}{\partial \omega} J(\omega; H, \vec{Q}) \cdot \vec{\mathbf{V}} = \frac{1}{2} \int_0^T \int_{\omega} \left\| \frac{\vec{Q}}{H} - \vec{v} \right\|^2 (\vec{\nabla} \cdot \vec{\mathbf{V}}) + \frac{1}{2} \int_0^T \int_{\omega} |\text{curl}(\frac{\vec{Q}}{H})|^2 (\vec{\nabla} \cdot \vec{\mathbf{V}})$$

for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$, and with (H, \vec{Q}) is solution of the state system (2.1) with initial and boundary conditions (2.2), and (p, \vec{r}) is solution of the adjoint system (3.1) with final and boundary conditions (3.2).

Proof. Let $\omega \in X_0$, such that $\omega = \psi(\omega_0)$ with $\psi \in O_0$. We consider $\mathbf{F} \in Lip(\omega; \mathbb{R}^2)$ a bi-Lipschitz homeomorphism such that $\omega = \mathbf{F}(\omega_0) \in X_0$. Let $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$ such that $\vec{\mathbf{V}} = \mathbf{F} - I$.

We define the function d by $d : \omega \in X_0 \rightarrow d(\omega) \in \mathbb{R}$, and we introduce a transported function called \underline{d}

$$\underline{d} : \mathbf{F} \in Lip(\omega; \mathbb{R}^2) \rightarrow \underline{d}(\mathbf{F}) = d(\mathbf{F}(\omega)) = d(\omega) \in \mathbb{R}.$$

Thus, the domain derivative of d at a given $\omega \in X_0$ writes

$$\frac{\partial}{\partial \omega} d(\omega) \cdot \vec{V} = \frac{\partial}{\partial \mathbf{F}} d(I) \cdot \vec{V}, \quad \text{for } \vec{V} \in Lip(\omega; \mathbb{R}^2) \quad (2.4.3)$$

The calculation of the domain derivative of j is done by introducing the transported functions of A , L , and J . In order to realize this computation, we define the function A as a sum of a number of functions as follow

$$A(\omega, H, \vec{Q}; p, \vec{r}) = \sum_{k=1}^6 A^k(\omega, H, \vec{Q}; p, \vec{r})$$

Explicitly

$$\begin{aligned} A(\omega, H, \vec{Q}; p, \vec{r}) &= \underbrace{\int_0^T \int_{\omega} \frac{\partial H}{\partial t} p}_{A^1(\omega, H, \vec{Q}; p, \vec{r})} + \underbrace{\int_0^T \int_{\omega} (\vec{\nabla} \cdot \vec{Q}) p}_{A^2(\omega, H, \vec{Q}; p, \vec{r})} + \underbrace{\int_0^T \int_{\omega} \frac{\partial \vec{Q}}{\partial t} \cdot \vec{r}}_{A^3(\omega, H, \vec{Q}; p, \vec{r})} \\ &+ \underbrace{\int_0^T \int_{\omega} \vec{\nabla} \cdot \left(\frac{\vec{Q}}{H} \otimes \vec{Q} \right) \cdot \vec{r}}_{A^4(\omega, H, \vec{Q}; p, \vec{r})} + \underbrace{\int_0^T \int_{\omega} g H \vec{\nabla}(H) \cdot \vec{r}}_{A^5(\omega, H, \vec{Q}; p, \vec{r})} + \underbrace{\int_0^T \int_{\omega} -g H \vec{\nabla} \eta \cdot \vec{r}}_{A^6(\omega, H, \vec{Q}; p, \vec{r})} \end{aligned}$$

With a similar manner we define J as a sum of

$$J(\omega, H, \vec{Q}) = \sum_{k=1}^2 J^k(\omega, H, \vec{Q})$$

Since,

$$J(\omega, H, \vec{Q}) = \underbrace{\frac{1}{2} \int_0^T \int_{\omega} \left\| \frac{\vec{Q}}{H} - \vec{V} \right\|^2}_{J^1(\omega, H, \vec{Q})} + \underbrace{\frac{\alpha}{2} \int_0^T \int_{\omega} \left| \text{curl} \left(\frac{\vec{Q}}{H} \right) \right|^2}_{J^2(\omega, H, \vec{Q})}$$

and

$$L(\omega; p, \vec{r}) = \int_0^T \int_{\omega} \vec{f} \cdot \vec{r}$$

Next, for computing the domain derivative of each function above, we introduce its transported function. In particular for the first function

$$A^1(\omega, H, \vec{Q}; p, \vec{r}) = \int_0^T \int_{\omega} \frac{\partial H}{\partial t} p,$$

we introduce the transported function

$$\begin{aligned} \underline{A}^1(\mathbf{F}, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) &= A^1(\mathbf{F}(\omega), \underline{H} \circ \mathbf{F}^{-1}, \underline{Q} \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{r} \circ \mathbf{F}^{-1}) \\ &= A^1(\omega, H, \vec{Q}; p, \vec{r}) = \int_0^T \int_{\omega} \frac{\partial H}{\partial t} \underline{p} | \det(\vec{\nabla} \mathbf{F}) | \end{aligned}$$

As it is known, $\mathbf{F} \rightarrow | \det(\vec{\nabla} \mathbf{F}) |$ is a differentiable application. Its derivative writes

$$\frac{\partial}{\partial \mathbf{F}} | \det(\vec{\nabla} \mathbf{F}) | (I) \cdot \vec{\mathbf{V}} = \vec{\nabla} \cdot \vec{\mathbf{V}}, \quad \text{for } \vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$$

and we have for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\frac{\partial}{\partial \omega} A^1(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} = \frac{\partial}{\partial \mathbf{F}} \underline{A}^1(I, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) \cdot \vec{\mathbf{V}} = \int_0^T \int_{\omega} \frac{\partial H}{\partial t} p (\vec{\nabla} \cdot \vec{\mathbf{V}})$$

For the second function $A^2(\omega, H, \vec{Q}; p, \vec{r})$, we consider the following transported function

$$\begin{aligned} \underline{A}^2(\mathbf{F}, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) &= A^2(\mathbf{F}(\omega), \underline{H} \circ \mathbf{F}^{-1}, \underline{Q} \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{r} \circ \mathbf{F}^{-1}) \\ &= A^2(\omega, H, \vec{Q}; p, \vec{r}) = \int_0^T \int_{\omega} (\vec{\nabla} \mathbf{F}^{-1} \circ \mathbf{F})^t : \vec{\nabla}(\underline{Q}) \underline{p} | \det(\vec{\nabla} \mathbf{F}) | \end{aligned}$$

where the inner product "X:Y" is expressed as $X : Y = \sum_{k,l=1}^2 X_{kl} Y_{kl}$.

Taking into account that the application $\mathbf{F} \rightarrow \vec{\nabla} \mathbf{F}^{-1} \circ \mathbf{F}$ is differentiable. Its derivative takes the form

$$\frac{\partial}{\partial \mathbf{F}} (\vec{\nabla} \mathbf{F}^{-1} \circ \mathbf{F})(I) \cdot \vec{\mathbf{V}} = -\vec{\nabla} \cdot \vec{\mathbf{V}}, \quad \text{for } \vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$$

Then, for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\begin{aligned} \frac{\partial}{\partial \omega} A^2(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} &= \frac{\partial}{\partial \mathbf{F}} \underline{A}^2(I, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) \cdot \vec{\mathbf{V}} \\ &= \int_0^T \int_{\omega} (\vec{\nabla} \cdot \vec{Q}) p (\vec{\nabla} \cdot \vec{\mathbf{V}}) - \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}})^t : (\vec{\nabla} \vec{Q}) p \end{aligned}$$

For the third function $A^3(\omega, H, \vec{Q}; p, \vec{r})$, we consider the following transported function

$$\begin{aligned}
\underline{A}^3(\mathbf{F}, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) &= A^3(\mathbf{F}(\omega), \underline{H} \circ \mathbf{F}^{-1}, \underline{Q} \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{r} \circ \mathbf{F}^{-1}) \\
&= A^3(\omega, H, \vec{Q}; p, \vec{r}) = \int_0^T \int_{\omega} \frac{\partial \vec{Q}}{\partial t} \cdot \vec{r} | \det(\vec{\nabla} \mathbf{F}) |
\end{aligned}$$

Consequently, for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\frac{\partial}{\partial \omega} A^3(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} = \frac{\partial}{\partial \mathbf{F}} \underline{A}^3(I, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) \cdot \vec{\mathbf{V}} = \int_0^T \int_{\omega} \frac{\partial \vec{Q}}{\partial t} \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}})$$

For the function $A^4(\omega, H, \vec{Q}; p, \vec{r})$, we consider the following transported function

$$\begin{aligned}
\underline{A}^4(\mathbf{F}, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) &= A^4(\mathbf{F}(\omega), \underline{H} \circ \mathbf{F}^{-1}, \underline{Q} \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{r} \circ \mathbf{F}^{-1}) = A^4(\omega, H, \vec{Q}; p, \vec{r}) \\
&= \int_0^T \int_{\omega} (\vec{\nabla} \mathbf{F}^{-1} \circ \mathbf{F})^t : \vec{\nabla} \left(\frac{\vec{Q}}{\underline{H}} \otimes \underline{Q} \right) \cdot \vec{r} | \det(\vec{\nabla} \mathbf{F}) |
\end{aligned}$$

With a similar manner used for the function $A^2(\omega, H, \vec{Q}; p, \vec{r})$ we obtain for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\begin{aligned}
\frac{\partial}{\partial \omega} A^4(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} &= \frac{\partial}{\partial \mathbf{F}} \underline{A}^4(I, \underline{H}, \underline{Q}; \underline{p}, \underline{r}) \cdot \vec{\mathbf{V}} \\
&= \int_0^T \int_{\omega} (\vec{\nabla} \cdot \vec{Q}) \frac{\vec{Q}}{H} \cdot \vec{r} (\vec{\nabla} \vec{\mathbf{V}}) - \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}})^t : (\vec{\nabla} \vec{Q}) \frac{\vec{Q}}{H} \cdot \vec{r} \\
&\quad + \int_0^T \int_{\omega} (\vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r} (\vec{\nabla} \vec{\mathbf{V}}) - \int_0^T \int_{\omega} (\vec{\nabla} \vec{\mathbf{V}} \vec{Q} \cdot \vec{\nabla}) \frac{\vec{Q}}{H} \cdot \vec{r}
\end{aligned}$$

Finally, with same calculations, we get the domain derivative of the other functions for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\frac{\partial}{\partial \omega} A^5(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} = \int_0^T \int_{\omega} g H \vec{\nabla} H \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}}) - \int_0^T \int_{\omega} g H (\vec{\nabla} \vec{\mathbf{V}})^t \vec{\nabla} H \cdot \vec{r}$$

$$\frac{\partial}{\partial \omega} A^6(\omega, H, \vec{Q}; p, \vec{r}) \cdot \vec{\mathbf{V}} = - \int_0^T \int_{\omega} g H \vec{\nabla} \eta \cdot \vec{r} (\vec{\nabla} \cdot \vec{\mathbf{V}})$$

For $J(\omega, H, \vec{Q}; p, \vec{r})$ and $L(\omega, H, \vec{Q}; p, \vec{r})$, we follow the same steps computation above. Then we get for $\vec{\mathbf{V}} \in Lip(\omega; \mathbb{R}^2)$

$$\begin{aligned} \frac{\partial}{\partial \omega} J^1(\omega, H, \vec{Q}) \cdot \vec{V} &= \frac{1}{2} \int_0^T \int_{\omega} \left| \frac{\vec{Q}}{H} - \vec{V} \right|^2 (\vec{\nabla} \cdot \vec{V}) \\ \frac{\partial}{\partial \omega} J^2(\omega, H, \vec{Q}) \cdot \vec{V} &= \frac{\alpha}{2} \int_0^T \int_{\omega} \left| \text{curl} \left(\frac{\vec{Q}}{H} \right) \right|^2 (\vec{\nabla} \cdot \vec{V}) \\ \frac{\partial}{\partial \omega} L(\omega; p, \vec{r}) \cdot \vec{V} &= \int_0^T \int_{\omega} \vec{f} \cdot \vec{r} (\vec{\nabla} \cdot \vec{V}) \end{aligned}$$

Finally, By injecting all previous results in (2.4.2), we obtain the formula for the domain derivative of the objective function j at $\omega \in X_0$ in terms of (H, \vec{Q}) solution of (2.2.1)-(2.2.2), and (p, \vec{r}) solution of (2.3.1)-(2.3.2). \square

2.5 Analysis of the adjoint problem

2.5.1 Pseudo-differential tools

In this section, we present some properties on pseudo-differential operators. We restrict ourselves to the main tools which will be used in the following. We consider ω is a set of \mathbb{R}^2 . We define S as the space C^∞ of complex-valued functions defined in \mathbb{R}^2 , which decrease faster than any polynomial function of $|x|$ as $|x| \rightarrow \infty$ (their derivative are so).

The spatial Fourier transform $\hat{p}(\xi)$ of $p(x) \in S$ writes

$$\hat{p}(\xi) = \int_{\mathbb{R}^2} p(x) e^{-ix\xi} dx$$

conversely, the inverse formula takes the form

$$p(x) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{p}(\xi) e^{-ix\xi} d\xi$$

Consider a sequence of scalar $\alpha = (\alpha_j)_{j=1,2}$, the α derivative is given by

$$D^\alpha p(x) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \xi^\alpha \hat{p}(\xi) e^{ix\xi} d\xi$$

where $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$, and $D_j = \frac{1}{i} \frac{\partial}{\partial x_j}$. We note the integral above by $p(D) = D^\alpha$. $p(D)$ is called the differential operator with symbol $p(\xi) = \xi^\alpha$. The differential operator depends the space variable too in general. More generally, the differential operator of order k can be written as

$$p(x, D) = \sum_{|\alpha| \leq k} a_k(x) D^\alpha$$

for some smooth coefficients $a_\alpha(x)$. The symbol of this differential operator writes

$$p(x, \xi) = \sum_{|\alpha| \leq k} a_\alpha \xi^\alpha$$

For instance, for $u \in S$, we have the following formula

$$p(x, D)u(x) = \mathcal{F}_{\xi \rightarrow x}^{-1}(p(x, \xi)\hat{u}(\xi)) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} p(x, \xi)\hat{u}(\xi)e^{ix\xi} d\xi$$

There are some restrictions about the differential operators. In particular they are associated only to polynomial symbols. More generally, we can treat a general class of symbols by introducing the pseudodifferential operators. For instance, if $p(x, \xi)$ is a functions in the researched admissible class, it verifies

$$|\partial_\xi^\alpha p(x, \xi)| \leq c_\alpha \Lambda(\xi)^{m-|\alpha|}$$

and

$$|\partial_x^\alpha p(x, \xi)| \leq c_\alpha \Lambda(\xi)^m$$

We define the Sobolev space $H^s(\mathbb{R}^2)$, $s \in \mathbb{R}$ as

$$H^s(\mathbb{R}^2) = \{W \in S' : (1 + |\xi|^2)^{\frac{s}{2}} \hat{W}(\xi) \in L^2(\mathbb{R}^2)\}$$

$\|\cdot\|_{H^s}$ is the usual Sobolev norm based on L^2 ,

$$\|W\|_{H^s} = \|\Lambda^s W\|_{L^2(\mathbb{R})}, \text{ with } \Lambda^s = (1 - \Delta)^{\frac{s}{2}}$$

where Δ is the Laplace operator.

Let $p(x, \xi)$ be a C^∞ function. $p(x, \xi) \in S_{1,0}^m$, $m \in \mathbb{R}$ means that

$$|D_x^\beta D_\xi^\alpha p(x, \xi)| \leq C_{\alpha,\beta} (1 + |\xi|)^{m-|\alpha|}, \forall \alpha, \beta \in \mathbb{R}$$

And $p(x, \xi) \in S^m$ means that:

$$p(x, \xi) \sim \sum_{j \leq m} p_j(x, \xi)$$

where $p_j(x, \xi)$ is such that: for $|\xi| \geq 1$,

$$p_j(x, r\xi) = r^j p_j(x, \xi), \quad |\xi| \geq 1, \quad r \geq 1$$

Pseudo-differential operator P defined by:

$$P(x, D)\psi = \int p(x, \xi) \hat{\psi}(\xi) e^{ix\xi} d\xi \quad (2.5.1)$$

is called of order m if $p(x, \xi) \in S^m$. $OP(\Sigma)$ denotes the set of operators with symbols in Σ .

2.5.2 Main result

In this section, we establish a conservative form of the adjoint system obtained in the previous section. We present a standard method to construct a symmetrizer of the found system. In order to obtain a conservative form of our adjoint system (2.3.1), we set $\vec{Q} = (Hu_1, Hu_2)^t$, $\vec{r} = (r_1, r_2)^t$ and $\vec{v} = (v_1, v_2)^t$. We consider the following scalar quantities; $F = (\frac{\vec{Q}}{H} - \vec{v}) \cdot \frac{\vec{Q}}{H^2} + \alpha \vec{curl}(curl(\frac{\vec{Q}}{H})) \cdot \frac{\vec{Q}}{H^2}$, $G_1 = \frac{1}{H}(\frac{u_1}{H} - v_1) + \alpha \frac{1}{H}(\partial_{yx}u_2 - \partial_{yy}u_1)$ and $G_2 = \frac{1}{H}(\frac{u_2}{H} - v_2) + \alpha \frac{1}{H}(\partial_{xx}u_2 - \partial_{xy}u_1)$.

After heavy calculus the system (2.3.1) takes its conservative form

$$\frac{\partial \vec{W}}{\partial t} + M(H, \vec{Q}) \frac{\partial \vec{W}}{\partial x} + N(H, \vec{Q}) \frac{\partial \vec{W}}{\partial y} = S(H, \vec{Q}, \vec{W}) \quad (2.5.2)$$

Where $\vec{W} = (p, r_1, r_2)^t$, $c = \sqrt{gH}$

$$M(H, \vec{Q}) = \begin{bmatrix} 0 & c^2 - u_1^2 & -u_1 u_2 \\ 1 & 2u_1 & u_2 \\ 0 & 0 & u_1 \end{bmatrix}, \quad S(H, \vec{Q}, \vec{W}) = \begin{bmatrix} F - g\vec{\nabla}\eta \cdot \vec{r} \\ -G_1 \\ -G_2 \end{bmatrix}$$

and

$$N(H, \vec{Q}) = \begin{bmatrix} 0 & -u_1 u_2 & c^2 - u_2^2 \\ 0 & u_2 & 0 \\ 1 & u_1 & 2u_2 \end{bmatrix}$$

Definition 2.5.1. *The system (2.5.2) is hyperbolic if the matrix $E = M(H, \vec{Q}).n_1 + N(H, \vec{Q}).n_2$ is diagonalizable.*

Proposition 2.5.1. *If $u_1, u_2, c \in \mathbb{R}$, then the system (2.5.2) is hyperbolic.*

Proof: The matrix $E = M(H, \vec{Q}).n_1 + N(H, \vec{Q}).n_2$ writes

$$E = \begin{bmatrix} 0 & (c^2 - u_1^2)n_1 - u_1u_2n_2 & -u_1u_2n_1 + (c^2 - u_2^2)n_2 \\ n_1 & 2u_1n_1 + u_2n_2 & u_2n_1 \\ n_2 & u_1n_2 & u_1n_1 + 2u_2n_2 \end{bmatrix}$$

which is diagonalizable, and has the following three eigenvalues

$$\lambda_1 = u_1n_1 + u_2n_2 - c, \quad \lambda_2 = u_1n_1 + u_2n_2, \quad \lambda_3 = u_1n_1 + u_2n_2 + c$$

Moreover, The eigenvectors of the matrix E writes

$$X = \frac{1}{2c} \begin{bmatrix} u_1n_1 + u_2n_2 + c & 2(u_1n_2 - u_2n_1) & -(u_1n_1 + u_2n_2) + c \\ -n_1 & -2n_2 & n_1 \\ -n_2 & 2n_1 & n_2 \end{bmatrix}$$

and we have

$$X^{-1} = \begin{bmatrix} 1 & u_1 - c.n_1 & u_2 - c.n_2 \\ 0 & -c.n_2 & c.n_1 \\ 1 & u_1 + c.n_1 & u_2 + c.n_2 \end{bmatrix}$$

2.5.3 Existence and Uniqueness using the pseudo-differential calculus

2.5.3.1 Construction of symmetrizer

The objective of this subsection is to construct a symmetrizer of the system (2.5.2) . In order to achieve this, we introduce a standard definition of the symmetrizer notion.

Definition 2.5.2. *Let $R(t, x, \vec{W}, \xi)$ be a smooth function on $\mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^3 \times \mathbb{R}_*^2$, homogeneous of degree 0 in ξ . $R(t, x, \vec{W}, \xi)$ is called a symmetrizer of the system (2.5.2) if*

$$R(t, x, \vec{W}, \xi) \text{ is a positive definite matrix,}$$

and

$$R(t, x, \vec{W}, \xi) \cdot \{M(H, \vec{Q})\xi_1 + N(H, \vec{Q})\xi_2\} \text{ is self-adjoint for each } (t, x, \vec{W}, \xi_1, \xi_2).$$

□

We define the following operator P

$$P(t, x, \vec{W}, \xi) = M(H, \vec{Q})\xi_1 + N(H, \vec{Q})\xi_2 \quad (2.5.3)$$

We will diagonalize the matrix $P(t, x, \vec{W}, \xi)$ and we will use the eigenvectors of its adjoint matrix to define the symmetrizer of the system (2.5.2). In general, it becomes difficult in some case to find a symmetrizer which verifies the properties in the definition above. In the following, we present a standard method to construct a symmetrizer for the system (2.5.2).

Proposition 2.5.2. *If $u_1, u_2, \xi_1, \xi_2 \in \mathbb{R}$, then the matrix $P(t, x, \vec{W}, \xi)$ is diagonalizable.*

Proof. The matrix $P(t, x, \vec{W}, \xi)$ has three distinct eigenvalues

$$\alpha_1 = u_1\xi_1 + u_2\xi_2 - c, \quad \alpha_2 = u_1\xi_1 + u_2\xi_2, \quad \alpha_3 = u_1\xi_1 + u_2\xi_2 + c$$

The conjugate operator $P(t, x, \vec{W}, \xi)^*$ of P is diagonalizable, and its eigenvectors write

$$P_1 = \begin{bmatrix} 1 \\ u_1 - c.\xi_1 \\ u_2 - c.\xi_2 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 \\ -c.\xi_2 \\ c.\xi_1 \end{bmatrix}, \quad P_3 = \begin{bmatrix} 1 \\ u_1 + c.\xi_1 \\ u_2 + c.\xi_2 \end{bmatrix}$$

The symmetrizer of the system (2.5.2) writes

$$R(t, x, \vec{W}, \xi) = P_1P_1^* + P_2P_2^* + P_3P_3^* \quad (2.5.4)$$

Explicitly

$$R(t, x, \vec{W}, \xi) = 2 \begin{bmatrix} 1 & u_1 & u_2 \\ u_1 & u_1^2 + \frac{c^2}{2}(\xi_1^2 + 1) & u_1u_2 + \frac{c^2}{2}\xi_1\xi_2 \\ u_2 & u_1u_2 + \frac{c^2}{2}\xi_1\xi_2 & u_2^2 + \frac{c^2}{2}(\xi_2^2 + 1) \end{bmatrix} \quad (2.5.5)$$

Which is self-adjoint, and we note that

$$RP = P^*R = 2 \begin{bmatrix} \alpha_2 & u_1\alpha_2 + c^2\xi_1 & u_2\alpha_2 + c^2\xi_2 \\ u_1\alpha_2 + c^2\xi_1 & a & b \\ u_2\alpha_2 + c^2\xi_2 & b & c \end{bmatrix}$$

where $a = \frac{1}{2}\{(2u_1^2 + c^2(n_1^2 + 1))\alpha_2 + 4u_1c^2n_1\}$, $b = \frac{1}{2}\{(2u_1u_2 + c^2\xi_1\xi_2)\alpha_2 + c^2(u_1\xi_2 + u_2\xi_1)\}$, and $c = \frac{1}{2}\{(2u_2^2 + c^2(n_2^2 + 1))\alpha_2 + 4u_2c^2n_2\}$.

Finally, in order to derive an energy estimate we define the operator Q using the symmetrizer R as:

$$Q = \frac{1}{2}(R + R^*) + K\Lambda^{-1} \quad (2.5.6)$$

where K is selected to make Q a positive definite operator in L^2 .

Lemma 2.5.1. *Let the operator P be defined by the expression in (2.5.3). If $Q \in C^1(S^0)$ then*

$$|\langle QPv, v \rangle| \leq C(\|v\|_{C^1} \|v\|_{L^2}^2) \quad (2.5.7)$$

Lemma 2.5.2. *(Kato Ponce commutator [10]) Let the operator P be defined by the expression in (2.5.3), then*

$$\|[\Lambda^s, P]v\|_{L^2} \leq C\{\|\nabla P\|_{\infty} \|v\|_{H^{s-1}} + \|P\|_{H^s} \|v\|_{\infty}\}$$

2.5.3.2 Existence and uniqueness

The objective of this section is to prove the existence and uniqueness of the adjoint system solution. In order to assume that, we introduce the Friedrich's mollifier $\{J_\varepsilon\}_{\varepsilon>0}$, and we define the regularized system obtained from (2.5.2) by:

$$\begin{aligned} \partial_t \vec{W}_\varepsilon + J_\varepsilon M(H, Q) \partial_x (J_\varepsilon \vec{W}_\varepsilon) + J_\varepsilon N(H, Q) \partial_y (J_\varepsilon \vec{W}_\varepsilon) &= J_\varepsilon \cdot S(\vec{W}_\varepsilon) \\ \vec{W}_\varepsilon(0, x) &= \vec{W}(0, x) \end{aligned} \quad (2.5.8)$$

We define the following operators

$$P_\varepsilon(\cdot) = J_\varepsilon M(H, Q) \partial_x(\cdot) + J_\varepsilon N(H, Q) \partial_y(\cdot) \quad (2.5.9)$$

and

$$g_\varepsilon(\vec{W}_\varepsilon) = J_\varepsilon \cdot S(\vec{W}_\varepsilon) \quad (2.5.10)$$

The definition of the symmetrizer R and the regularized operator Q_ε using R_ε , will help us to derive an energy estimate by applying Λ^s and symmetrizer (2.5.4) to the system (2.5.8) to estimate $\langle \Lambda^s \vec{W}_\varepsilon, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle$

$$\frac{d}{dt} \langle \Lambda^s \vec{W}_\varepsilon, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle = 2 \langle \Lambda^s \frac{\partial \vec{W}_\varepsilon}{\partial t}, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle + \langle \Lambda^s \vec{W}_\varepsilon, Q'_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle \quad (2.5.11)$$

The second term of the right hand side can be raised as the following: from definition 2.5.2 the construction of the symmetrizer R_ε gives

$$|\langle \Lambda^s \vec{W}_\varepsilon, R'_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle| \leq C(\|\vec{W}_\varepsilon\|_{C^1}) \|\vec{W}_\varepsilon\|_{H^s}^2$$

Using the description of the regularized operator Q_ε in (2.5.6), we get

$$|\langle \Lambda^s \vec{W}_\varepsilon, Q'_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle| \leq C(\|\vec{W}_\varepsilon\|_{C^1}) \|\vec{W}_\varepsilon\|_{H^s}^2$$

It remains to deal with the first term of (2.5.11)

$$\langle \Lambda^s \frac{\partial \vec{W}_\varepsilon}{\partial t}, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle = \langle Q_\varepsilon \Lambda^s J_\varepsilon P_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle + \langle Q_\varepsilon \Lambda^s g_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle$$

The second term in right hand side can be raised as

$$|\langle Q_\varepsilon \Lambda^s g_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle| \leq C(\|\vec{W}_\varepsilon\|_{H^s}) \|\vec{W}_\varepsilon\|_{H^s}$$

Next, we will try to raise the first term. It can be decomposed as

$$\begin{aligned} \langle Q_\varepsilon \Lambda^s J_\varepsilon P_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle &= \langle Q_\varepsilon P_\varepsilon \Lambda^s J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle + \langle Q_\varepsilon [\Lambda^s, P_\varepsilon] J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle \\ &\quad + \langle [Q_\varepsilon \Lambda^s, J_\varepsilon] P_\varepsilon J_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle \end{aligned} \quad (2.5.12)$$

The operator $[Q_\varepsilon \Lambda^s, J_\varepsilon]$ is bounded in the H^{s-1} norm. the term $\langle [Q_\varepsilon \Lambda^s, J_\varepsilon] P_\varepsilon J_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle$ can be raised by

$$\langle [Q_\varepsilon \Lambda^s, J_\varepsilon] P_\varepsilon J_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle \leq C(\|\vec{W}_\varepsilon\|_{C^1}) \|P_\varepsilon J_\varepsilon \vec{W}_\varepsilon\|_{H^{s-1}} \|\vec{W}_\varepsilon\|_{H^s}$$

Using Moser's inequality (see [5]) for the term $\|P_\varepsilon J_\varepsilon \vec{W}_\varepsilon\|_{H^{s-1}}$ yields

$$\langle [Q_\varepsilon \Lambda^s, J_\varepsilon] P_\varepsilon J_\varepsilon \vec{W}_\varepsilon, \Lambda^s \vec{W}_\varepsilon \rangle \leq C(\|\vec{W}_\varepsilon\|_{H^s}) \|\vec{W}_\varepsilon\|_{H^s} \quad (2.5.13)$$

The operator $[\Lambda^s, P_\varepsilon]$ has order s , and we use the results of lemma 2.5.2 to raise

$$\begin{aligned} |\langle Q_\varepsilon [\Lambda^s, P_\varepsilon] J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle| &\leq C \|\Lambda^s \vec{W}_\varepsilon\|_{L^2} \{ \|\nabla P_\varepsilon\|_\infty \|\vec{W}_\varepsilon\|_{H^{s-1}} \\ &\quad + \|P_\varepsilon\|_{H^s} \|\vec{W}_\varepsilon\|_\infty \} \end{aligned}$$

Since

$$\begin{aligned} |\langle Q_\varepsilon [\Lambda^s, P_\varepsilon] J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle| &\leq C \|\vec{W}_\varepsilon\|_{H^s} \{ \|\nabla P_\varepsilon\|_\infty \|\vec{W}_\varepsilon\|_{H^{s-1}} \\ &\quad + \|P_\varepsilon\|_{H^s} \|\vec{W}_\varepsilon\|_\infty \} \end{aligned}$$

From the previous expression and using the Sobolev embedding we deduce:

$$|\langle Q_\varepsilon [\Lambda^s, P_\varepsilon] J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle| \leq C(\|\vec{W}_\varepsilon\|_{H^s}) \|\vec{W}_\varepsilon\|_{H^s} \quad (2.5.14)$$

We use the results of Lemma.5.1 for the first term of the equation (5.4.13)

$$|\langle Q_\varepsilon P_\varepsilon \Lambda^s J_\varepsilon \vec{W}_\varepsilon, \Lambda^s J_\varepsilon \vec{W}_\varepsilon \rangle| \leq C(\|\vec{W}_\varepsilon\|_{H^s}) \|\vec{W}_\varepsilon\|_{H^s} \quad (2.5.15)$$

The three inequalities (5.4.14)-(2.5.15) and the Cauchy inequality yields

$$\frac{d}{dt} \langle \Lambda^s \vec{W}_\varepsilon, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle \leq C(t) \{ \|\vec{W}\|_{H^s}^2 + \|S\|_{L^\infty([0,t], H^s)}^2 \} \quad (2.5.16)$$

Using the previous results, we prove the following theorem

Theorem 2.5.1. *Let $\vec{W}^0 \in H^s(\mathbb{R}^2)$ and $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, then there exists a constant $C(t)$ independent of $\varepsilon \in]0, 1]$ such that*

$$\|\vec{W}_\varepsilon\|_{H^s}^2 \leq C(t) \{ \|\vec{W}^0\|_{H^s}^2 + \|S\|_{L^\infty([0,t], H^s)}^2 \} \quad (2.5.17)$$

Proof. We observe that the product $\langle \cdot, Q_\varepsilon \cdot \rangle$ defines an equivalent norm in H^s and use Gronwall's inequality to get

$$\|\vec{W}_\varepsilon\|_{H^s}^2 \leq C(t) \{ \|\vec{W}^0\|_{H^s}^2 + \|S\|_{L^\infty([0,t],H^s)}^2 \} \quad (2.5.18)$$

□

Corollary 2.5.1. *Let $\vec{W}^0 \in H^s(\mathbb{R}^2)$ and $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, then the system (2.5.2) have a solution $\vec{W} \in H^s(\mathbb{R}^2)$ in the time interval $[0, T]$.*

Proof. Let $\varepsilon \in]0, 1]$, Using theorem.5.1 in a Banach space, we have the energy inequality in terms of \vec{W}_ε

$$\frac{d}{dt} \langle \Lambda^s \vec{W}_\varepsilon, Q_\varepsilon \Lambda^s \vec{W}_\varepsilon \rangle \leq C([0, T_\varepsilon]) \{ \|\vec{W}_\varepsilon\|_{H^s}^2 + \|S\|_{L^\infty([0,T],H^s)} \} \quad (2.5.19)$$

The existence of a function $S(t)$ independent of $\varepsilon \in]0, 1]$ and an interval time $[0, T]$ yields from Gronwall's inequality:

$$\|\vec{W}_\varepsilon\|_{H^s} \leq S(t) \quad \forall t \in [0, T]$$

then $\{\vec{W}_\varepsilon, \varepsilon \in]0, 1]\}$ converges to a limit \vec{W} solution of the system (2.5.2). The sequence $\{\vec{W}_\varepsilon, \varepsilon \in]0, 1]\}$ is bounded in $C([0, T], H^s(B(0, r)))$ and $\{\frac{\partial \vec{W}_\varepsilon}{\partial t}, \varepsilon \in]0, 1]\}$ is bounded in $C([0, T], H^{s-1}(B(0, r)))$ by definition of (2.5.8) with $B(0, r)$ is a ball. And we have the compact inclusion $H^s(B(0, r)) \subset H^{s-1}(B(0, r))$.

Ascoli's theorem with interpolation inequalities arguments gives the existence of a subsequence \vec{W}_{ε_j} such that

$$\vec{W}_{\varepsilon_j} \longrightarrow \vec{W} \text{ in } C([0, T], C^1(B(0, r)))$$

the limit \vec{W} verifies the system (2.5.2) in the distribution sense. □

Theorem 2.5.2. *Let $\vec{W}^0 \in H^s(\mathbb{R}^2)$ and $S \in L_{loc}^\infty(\mathbb{R}, H^s(\mathbb{R}^2))$, $s > \frac{5}{2}$, the system (2.5.2) has an unique solution in $H^s(\mathbb{R}^2)$.*

Proof: We suppose that the problem (2.5.2) has two solutions \vec{W}_1 et \vec{W}_2 such that:

$$\begin{aligned}\frac{\partial \vec{W}_1}{\partial t} &= -[M(H, \vec{Q})\partial_x(\vec{W}_1) + N(H, \vec{Q})\partial_y(\vec{W}_1)] + S(H, \vec{Q}, \vec{W}_1) \\ \frac{\partial \vec{W}_2}{\partial t} &= -[M(H, \vec{Q})\partial_x(\vec{W}_2) + N(H, \vec{Q})\partial_y(\vec{W}_2)] + S(H, \vec{Q}, \vec{W}_2)\end{aligned}\tag{2.5.20}$$

We define $\vec{W} = \vec{W}_1 - \vec{W}_2$, then \vec{W} satisfies the following equation:

$$\begin{aligned}\frac{\partial \vec{W}}{\partial t} &= -[M(H, \vec{Q})\partial_x(\vec{W}_1) + N(H, \vec{Q})\partial_y(\vec{W}_1)]\vec{W} - [M(H, \vec{Q})\partial_x(\vec{W}_1) + N(H, \vec{Q})\partial_y(\vec{W}_1)]\vec{W}_2 \\ &\quad + [M(H, \vec{Q})\partial_x(\vec{W}_2) + N(H, \vec{Q})\partial_y(\vec{W}_2)]\vec{W}_2 + S(H, \vec{Q}, \vec{W}_1) - S(H, \vec{Q}, \vec{W}_2)\end{aligned}\tag{2.5.21}$$

Other side, we have

$$S(H, \vec{Q}, \vec{W}_1) - S(H, \vec{Q}, \vec{W}_2) = s(H, \vec{Q}, \vec{W}_1, \vec{W}_2)(\vec{W}_1 - \vec{W}_2)$$

where $s(H, \vec{Q}, \vec{W}_1, \vec{W}_2) = \int_0^1 g'(\tau\vec{W}_1 + (1-\tau)\vec{W}_2)d\tau$, with $g(\cdot) = S(H, \vec{Q}, \cdot)$ and

$$[M(H, \vec{Q})\partial_x(\vec{W}_2 - \vec{W}_1) + N(H, \vec{Q})\partial_y(\vec{W}_2 - \vec{W}_1)] = (\vec{W}_1 - \vec{W}_2).a(H, \vec{Q}, \vec{W}_1, \vec{W}_2)$$

Consequently

$$\frac{\partial \vec{W}}{\partial t} = [M(H, \vec{Q})\partial_x(\vec{W}_1) + N(H, \vec{Q})\partial_y(\vec{W}_1)]\vec{W} + \vec{W}.a(\vec{W}_1, \vec{W}_2, D)\vec{W}_2$$

the operator $[M(H, \vec{Q})\partial_x(\cdot) + N(H, \vec{Q})\partial_y(\cdot)]$ is strictly hyperbolic and using the calculus of the proof of theorem.5.1 we infer

$$\frac{d}{dt} \|\vec{W}\|_{L^2}^2 \leq C(t) \|\vec{W}\|_{L^2}^2\tag{2.5.22}$$

Using Gronwall's inequality we deduce $\vec{W} \equiv 0$.

2.6 Numerical methods

The shallow water equations are a set of nonlinear hyperbolic equations. The nonlinear character combined with the hyperbolic type of the equations can lead to discontinuous solutions in finite time. In order to formulate simple and robust numerical procedures, the two-dimensional shallow water equations (2.2.1) are cast in conservation form with source terms

$$\partial_t U + \partial_x F(U) + \partial_y G(U) = S(U) \quad (2.6.1)$$

where

$$U = \begin{pmatrix} H \\ Hu_1 \\ Hu_2 \end{pmatrix}, \quad F(U) = \begin{pmatrix} Hu_1 \\ Hu_1^2 + \frac{1}{2}gH^2 \\ Hu_1u_2 \end{pmatrix}, \quad G(U) = \begin{pmatrix} Hu_2 \\ Hu_1u_2 \\ Hu_2^2 + \frac{1}{2}gH^2 \end{pmatrix}$$

and

$$S(U) = \begin{pmatrix} 0 \\ f_1 + gH\partial_x\eta \\ f_2 + gH\partial_y\eta \end{pmatrix}$$

2.6.1 Finite volume method

Finite volume schemes for the shallow water systems consist in using an upwinding of the fluxes. The problem domain is first discretized into a set of triangular cells T_i forming an unstructured computational mesh. Let Δt be the constant time step and define $t_n = n\Delta t$ for $n = 0, \dots, N$. At each discrete time t_n , we note U_i^n the approximated solution value.

Denote by $E(U) = (F(U), G(U))$ the physical fluxes. By integrating the equation (2.6.1) on a triangle T_i , we obtain

$$\int_{T_i} U_t + \int_{T_i} \nabla \cdot E(U) = \int_{T_i} S(U) \quad (2.6.2)$$

We note \vec{n}_i the normal on the edges of triangle T_i . Using the divergence formula $\int_{T_i} \nabla \cdot E(U) = \int_{\partial T_i} E(U) \cdot n_i d\Gamma$. The equation (2.6.2) takes the form

$$\int_{T_i} U_t + \int_{\partial T_i} E(U) \cdot n_i d\Gamma = \int_{T_i} S(U) \quad (2.6.3)$$

The term $\int_{\partial T_i} E(U) \cdot n_i d\Gamma$ can be calculated as

$$\int_{\partial T_i} E(U) \cdot n_i d\Gamma = \sum_{j=1}^3 E_{ij} \cdot n_{ij} \cdot dl_{ij}$$

The equation (2.6.3) becomes

$$|T_i| U_t + \sum_{j=1}^3 E_{ij} \cdot n_{ij} \cdot dl_{ij} = |T_i| S \quad (2.6.4)$$

where n_{ij} is the normal on the edge T_i/T_j , E_{ij} are the discrete fluxes on the interface T_i/T_j and dl_{ij} is the length of the interface T_i/T_j .

Thus, we now have an equation for each cell i of the form

$$U_t = -\frac{1}{|T_i|} \sum_{j=1}^3 E_{ij} \cdot n_{ij} \cdot dl_{ij} + S \quad (2.6.5)$$

We make a finite difference approximation to the time derivative to obtain the scheme

$$U_i^{n+1} = U_i^n - \frac{dt}{|T_i|} \cdot \sum_{j=1}^3 E_{ij} \cdot n_{ij} \cdot dl_{ij} + dt \cdot S \quad (2.6.6)$$

Finding the value of the fluxes at the interface is of primary importance. A variety of approximation techniques have been developed to allow efficient calculation of the solution to the Riemann problem. The Roe solver is used to evaluate the term $\sum_{j=1}^3 E_{ij} \cdot n_{ij} \cdot dl_{ij}$.

2.6.2 A gradient free algorithm

The design variables related to the shape Ω depend on the two positions of the slot $y = (a, b) = (y_1, y_2, y_3, y_4)$ (Figure 2.2). We redefine the objective function (2.2.6) in the following way $\bar{\Phi}_1 : \mathbb{R}^4 \rightarrow \mathbb{R}$ where $\bar{\Phi}_1(y) = J(\Omega(y))$. The finite volume scheme (2.6.6) yields, for each time t_n , an approximated velocity $\vec{u}_i^n = \frac{\vec{Q}_i^n}{H_i^n}$ which induces an approximate objective function

$$\bar{\Phi}_1(y) = \frac{\Delta t}{2} \sum_{n=1}^N \sum_{e \in T_i} \left[\int_e \|\vec{u}_i^n - \vec{v}\|^2 + \alpha \int_e |\text{curl}(\vec{u}_i^n)|^2 \right] \quad (2.6.7)$$

Finally, we collect all linear constraints (2.2.3) and (2.2.4) in a function $\vec{\phi}_2 : \mathbb{R}^4 \rightarrow \mathbb{R}^{10}$

$$\vec{\phi}_2(y_1, y_2, y_3, y_4) = \begin{pmatrix} \frac{1}{4}1.213 - y_1, & \frac{1}{4}1.213 - y_3, & y_1 - \frac{3}{4}1.21, & y_3 - \frac{3}{4}1.213, & -y_2, \\ -y_4, & y_2 - \frac{1}{2}0.97, & y_4 - \frac{1}{2}0.97, & 0.1 - y_3 + y_1, & 0.05 - y_2 + y_4 \end{pmatrix} \quad (2.6.8)$$

These constraints are satisfied if and only if $\vec{\phi}_2(y) \leq 0$. The functional $\bar{\Phi}_1$ can be penalized to include geometric and state constraints

$$\Phi(y) = \bar{\Phi}_1(y) + \beta \sum_{j=1}^{10} \max\{(\vec{\phi}_2(y))_j, 0\} \quad (2.6.9)$$

with β is a penalty parameter.

Due to the essentially geometric nature of the problem, we propose a direct search technique for solving the discretized control problem. The Nelder Mead "simplex" algorithm is one of the most widely used methods for nonlinear optimization. The method attempts to minimize a scalar-valued nonlinear function using only function values, without any derivative information. The method constructs a sequence of simplices as approximations to an optimal point. To describe Nelder-Mead iterations, we begin with an arbitrary simplex of 5 vertices y_1, y_2, \dots, y_5 . We evaluate and order our function on these vertices $\Phi(y_1) \leq \Phi(y_2) \leq \dots \leq \Phi(y_5)$. The vertex associated to the maximal value is replaced with a new point $y(\nu) = (1 + \nu)y^* - \nu y_5$, where y^* is the centroid of the convex hull $\{y_1, \dots, y_4\}$. The value of ν is chosen from this set of values: $\nu_\delta = -0.5$, $\nu_\gamma = 0.5$, $\nu_\alpha = 1$, $\nu_\beta = 2$. The choice of these values is determined according to the following algorithm

Calculate and sort $\Phi(y_1), \Phi(y_2), \dots, \Phi(y_5)$

While $|\Phi(y_5) - \Phi(y_1)|$ is not sufficiently small, calculate $y(\nu_\beta)$ and $\Phi_\beta = \Phi(y(\nu_\beta))$ then

- a) If $\Phi_\beta \leq \Phi(y_1)$ then calculate $\Phi_\alpha = \Phi(y(\nu_\alpha))$. If $\Phi_\beta \leq \Phi_\alpha$, replace y_5 with $y(\nu_\alpha)$; otherwise replace y_5 with $y(\nu_\beta)$. Go to (f)
- b) If $\Phi(y_1) \leq \Phi_\beta \leq \Phi(y_4)$ then replace y_5 with y_β and go to (f)
- c) If $\Phi(y_4) \leq \Phi_\beta \leq \Phi(y_5)$, then calculate $\phi_\gamma = \Phi(y(\nu_\gamma))$. If $\Phi_\gamma \leq \Phi_\beta$ replace y_5 with $x(\nu_\gamma)$ and go to (f). Otherwise go to (e)
- d) if $\Phi(y_5) \leq \Phi_\beta$ then calculate $\phi_\delta = \Phi(y(\nu_\delta))$. If $\Phi_\delta \leq \Phi_{y_5}$, replace y_5 with $y(\nu_\delta)$ and go to (f). Otherwise go to (e)
- e) For $j = 2, \dots, 5$, set $y_j = y_1 + \frac{1}{2}(y_j - y_1)$

f) Resort values of Φ at each resulting vertex

To prevent stagnation at non-optimal point, a modification proposed by Kelley [3] is used. This technique consists to replace the current simplex by a smaller one.

Theorem 2.6.1. *Let f a strictly convex function. Then if Nelder-Mead method is applied to f , then no shrink steps are performed.*

Proof. Shrink steps occur when the out and inside contraction are tried and fail. Firstly, we consider that the out contraction is taken. Out contraction occurs when $\Phi(x_n) \leq \Phi_\beta \leq \Phi(x_{n+1})$. $x(\nu_\gamma) = \frac{1}{2}x^* + \frac{1}{2}x(\nu_\beta)$ and if we use the convexity of Φ we have:

$$\Phi_\gamma = \Phi(x(\nu_\gamma)) = \Phi\left(\frac{1}{2}x^* + \frac{1}{2}x(\nu_\beta)\right) < \frac{1}{2}\Phi(x^*) + \frac{1}{2}\Phi(x(\nu_\beta)) \leq \max\{\Phi(x^*), \Phi_\beta\}$$

$\Phi(x^*) = \Phi\left(\frac{1}{n} \sum_{i=1}^n x_i\right)$. Using the convexity of Φ we have $\Phi(x^*) \leq \Phi(x^n)$, then $\max\{\Phi(x^*), \Phi_\beta\} = \Phi_\beta$. Finally $\Phi_\gamma \leq \Phi_\beta$ and the out contraction is taken (the shrink step is not applied).

Secondly, the inside contraction is performed only if $\Phi(x_{n+1}) \leq \Phi_\beta$. $x(\nu_\delta)$ can be written as convex combination $x(\nu_\delta) = \frac{1}{2}x^* + \frac{1}{2}x^{n+1}$. If we use the convexity of Φ we obtain:

$$\Phi_\delta = \Phi(x(\nu_\delta)) = \Phi\left(\frac{1}{2}x^* + \frac{1}{2}x^{n+1}\right) < \frac{1}{2}\Phi(x^*) + \frac{1}{2}\Phi(x^{n+1}) \leq \max\{\Phi(x^*), \Phi(x^{n+1})\}$$

$\max\{\Phi(x^*), \Phi(x^{n+1})\} = \Phi(x^{n+1})$, then the inside contraction is taken.

For the rest, let Φ_k^l be the value of Φ on the vertex l ($l = 1, \dots, n+1$) and iteration k .

Theorem 2.6.2. *Let Φ a bounded functiona from below on \mathbb{R}^n . Then if the Nelder-Med algorithm is applied to Φ then*

1. *The sequence $\{\Phi_k^1\}$ is convergent.*
2. *If only a finite number of shrinks occur, then all the $n+1$ sequences $\{\Phi_k^l\}$, $l=1, \dots, n+1$, converge and their limits satisfy $\Phi_*^1 \leq \Phi_*^2 \leq \dots \leq \Phi_*^{n+1}$ (where $\Phi_*^l = \lim_{k \rightarrow +\infty} \Phi_k^l$).*
3. *If only a finite number of non-shrinks occur, then all the simplex vertices converge to a single point.*

Proof All the values of the sequence $\{\Phi_k^1\}$ are bounded by below by zero (Φ is bounded from below by zero). $\{\Phi_k^1\}$ is a decreasing sequence, then it is convergent.

For $l = 1, \dots, n$, $\Phi_k^l \leq \Phi_k^{l+1}$, then if we pass to limit

$$\Phi_*^l \leq \Phi_*^{l+1} \quad \forall l \in [1, n]$$

2.6.3 Spectral projected gradient

The second adopted approach for resolving our shape optimization problem is consisting in use of a gradient technique called Spectral projected gradient (**SPG**). This method use the projection into a closed and convex Ω_0 subset of \mathbb{R}^4 of all the points $y \in \mathbb{R}^4$ satisfying (2.2.3)-(2.2.4).

We denote $\sigma_1 = \sigma_3 = \frac{1}{4}1.213$, $\delta_1 = \delta_3 = \frac{3}{4}1.213$, $\sigma_2 = \sigma_4 = 0$, $\delta_2 = \delta_4 = \frac{1}{4}0.97$. Then the admissible set Ω_0 is defined as

$$\Omega_0 = \{y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : \sigma_i \leq y_i \leq \delta_i, \quad i = 1, \dots, 4, \\ y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2\} \quad (2.6.10)$$

The optimization problem can be reformulated as

$$\min_{y \in \Omega_0} j(y). \quad (2.6.11)$$

Through the iterations of the appointed algorithm above the gradient of the cost function is called and used before calling the projection function as follow:

- (a). (Initialization). Let $\bar{y} \in \Omega_0$, and let $\varepsilon > 0$ be a positive tolerance.
- (b). (Search direction computation) Let $d = P_{\Omega_0}(\bar{y} - \eta \nabla j(\bar{y})) - \bar{y}$, with η is a positive constant given by
 - First iteration $\eta = 1$.
 - Other iterations: Let \bar{y} be the current point and \tilde{y} teh previous point. Calculate $x = \bar{y} - \tilde{y}$ and $y = \nabla j(\bar{y}) - \nabla j(\tilde{y})$. Then, if $x^T y > 0$, take $\eta = \frac{x^T x}{x^T y}$; elsewhere, take η as a fixed positive value.
- (c). (Termination) If $d = 0$ (in practice, $\|d\|_2 < \varepsilon$), then stop: \bar{y} is a stationary point of j on Ω_0 .
- (d). (Step-size) Calculate a value $\xi \in (0, 1]$ such that $j(\bar{y} + \xi d) \leq j(\bar{y}) + \xi \delta \nabla j(\bar{y})^T d$, with $\delta > 0$ (usually, $\delta \in [10^{-4}, 10^{-1}]$). For the determination of a suitable step-size ξ by an iterative way, we choose ζ as a positive constant (usually, $\zeta = 2$), take $\theta_1 = j(\bar{y})$ and $\theta_2 = \delta \nabla j(\bar{y})^T d$, and then for $p = 0, 1, 2, \dots$ define $\xi = \frac{1}{\zeta^p}$, and stop when $j(\bar{y} + \xi d) \leq \theta_1 + \xi \theta_2$.

(e) (Update) Define $\tilde{y} = \bar{y} + \xi d$, and go to (a) with $\bar{y} = \tilde{y}$.

The value of j in this algorithm is computed from the expression (2.6.7). And $y = P_{\Omega_0}(z)$ is the projection of $z \in \mathbb{R}^4$ onto Ω_0 . It is calculated by minimizing a quadratic function of the distance of z to Ω_0 as follow

$$\min_{y \in \Omega_0} \frac{1}{2} \|y - z\|_2^2 = \min_{y \in \Omega_0} \frac{1}{2} z^T z - z^T y + \frac{1}{2} y^T y \quad (2.6.12)$$

It leads to minimize

$$\left\{ \begin{array}{l} \min_{(y_1, y_2, y_3, y_4)} \sum_{i=1}^4 \left\{ \frac{1}{2} z_i^2 - z_i y_i + \frac{1}{2} y_i^2 \right\} \\ \text{subject to } \sigma_i \leq y_i \leq \delta_i, i = 1, \dots, 4 \\ y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{array} \right.$$

\Leftrightarrow

$$\left\{ \begin{array}{l} \min_{(y_1, y_3)} \frac{1}{2} (z_1^2 + z_3^2) - (z_1 y_1 + z_3 y_3) + \frac{1}{2} (y_1^2 + y_3^2) \\ \text{subject to } \sigma_i \leq y_i \leq \delta_i, i = 1, 3 \\ y_3 - y_1 \geq d_1, \end{array} \right.$$

and

$$\left\{ \begin{array}{l} \min_{(y_2, y_4)} \frac{1}{2} (z_2^2 + z_4^2) - (z_2 y_2 + z_4 y_4) + \frac{1}{2} (y_2^2 + y_4^2) \\ \text{subject to } \sigma_i \leq y_i \leq \delta_i, i = 2, 4 \\ y_2 - y_4 \geq d_2, \end{array} \right.$$

In order to resolve the two quadratic problems, we use the Karush-Kuhn-Tucker (KKT) technique, and the way of resolution of the two above problems is analogous. To release that, we begin by the first quadratic problem and we choose $\sigma = \sigma_1 = \sigma_3$, and $\delta = \delta_1 = \delta_3$, then from the five constraints, only three are kept. Finally the optimization problem is reformulated as

$$\begin{array}{ll} \min_{(x_1, x_2)} & c - (l_1 x_1 + l_2 x_2) + \frac{1}{2} (x_1^2 + x_2^2) \\ \text{subject to} & x_2 \leq \delta, \\ & x_1 \geq \sigma, \\ & x_2 - x_1 \geq d_1, \end{array}$$

where σ , δ , d_1 , c , and l_i , $i=1,2$, are real numbers.

We consider

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -\delta \\ \sigma \\ d_1 \end{pmatrix},$$

Then the linear constraints write

$$Ax - v = b, \quad v \geq 0,$$

where $v = (v_1, v_2, v_3)^T$ is the slack variables. The cost function is strictly convex and quadratic in \mathbb{R}^2 . Then the optimisation problem has a unique solution which satisfies the KKT conditions

$$\begin{aligned} v &= Ax - b \\ -l + x &= A^T w, \\ v \geq 0, \quad w &\geq 0, \\ v^T w &= 0, \end{aligned}$$

where $l = (l_1, l_2, l_3)$, and $w = (w_1, w_2, w_3)$ is the vector of the Lagrange multipliers associated to the three constraints above.

The unique optimal solution writes

$$x = l + A^T w$$

which gives

$$\begin{cases} x_1 = l_1 + w_2 - w_3 \\ x_2 = l_2 - w_1 + w_3 \end{cases}$$

For computing the Lagrange multipliers w_i , we resolve the following LCP problem

$$\begin{aligned} v &= (-b + Al) + AA^T w \\ v &\geq 0, \quad w \geq 0 \\ v^T w &= 0 \end{aligned}$$

Finally, the gradient of the cost function $\nabla j(\bar{y}) = \left(\frac{\partial j}{\partial \bar{y}_1}(\bar{y}), \frac{\partial j}{\partial \bar{y}_2}(\bar{y}), \frac{\partial j}{\partial \bar{y}_3}(\bar{y}), \frac{\partial j}{\partial \bar{y}_4}(\bar{y}) \right)$ called along the algorithm iteration above can be approximated for a fixed $\bar{y} \in \Omega_0$ and $t > 0$ small enough, by

$$\frac{\partial j}{\partial y_i}(\bar{y}) \approx \frac{j(\bar{y} + te_i) - j(\bar{y})}{t}, \quad \text{for } i = 1, \dots, 4.$$

2.6.4 Design examples

The characteristics of the flow in vertical slot fishways depended mainly on the specific pool design, such as the geometry of the pool. We conduct three different designs of vertical slot fishways with I or L shaped baffle.

The two different types of vertical slot fishways (VSF): a standard VSF and a multi-slot VSF were tested with the same physical and numerical conditions. All initial and boundary conditions are taken constants. $H_0 = H(x, 0) = 0.5m$, $Q_0 = Q(x, 0) = (0; 0)m^2s^{-1}$, $Q_1 = -0.065/0.97 m^2s^{-1}$, and $H_2 = 0.5 m$. the target velocity is $c = 0.8 m.s^{-1}$. The vorticity parameter is $\alpha = 0$. The penalty parameter is $\beta = 500$. In these experiments, we consider that only the bottom friction stress is applied associated to Chezy coefficient of 57.36. For fish passage with comfortable conditions, we put $d_1 = 0.1$ and $d_2 = 0.05$ (see comfort constraints (2.2.4)).

2.6.4.1 Vertical Slot Fishway ("I" shaped baffle-rectangular slots)

The fishway model consisted of a 12.13 m long, 0.97 m wide and 1.213 m deep flume, which contains 10 pools. There are two rectangular vertical slots in each pool with the geometric dimensions that are detailed in figure 3.5 (see [6]).

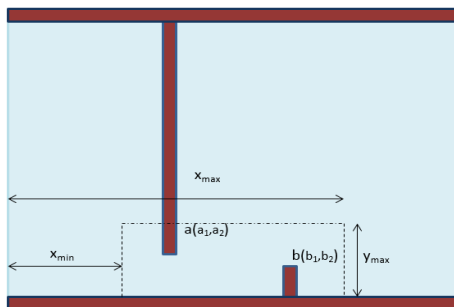


Figure 2.3: Geometric characteristics of a pool with "I" shaped baffle

In the case of the initial random shape (Figure 2.4), we observe that the main flow followed a curved path as it crosses the slot, and directly hit the opposite wall. Two low velocity recirculation regions appeared in the two sides of the jet. In the optimal shape configuration (Figure 5.1), the curved trajectory of velocity is removed, and the velocity is close to the uniform target velocity v for the optimal points $a = (0.5721, 0.1487)$ and $b = (0.8786, 0.0520)$.

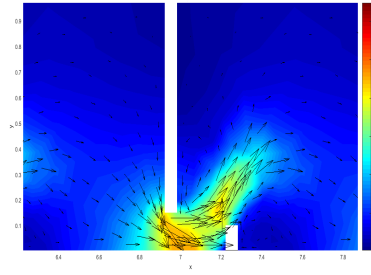


Figure 2.4: Initial random velocity field in the central pool

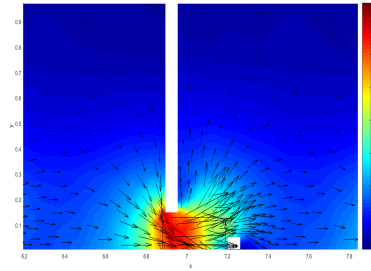


Figure 2.5: Optimal velocity field in the central pool based on gradient free algorithm

2.6.4.2 Vertical Slot Fishway ("I" shaped baffle-oblique slots)

We considered the fishway under study, whose scheme is shown in Figure 3.8. We maintain the same hydrodynamic and numerical conditions as the previous test. The time interval for the simulation was $T = 300s$.

The results show a recirculation regions flowing in opposite directions for the random

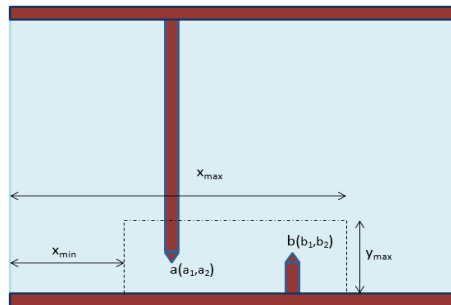


Figure 2.6: Geometric characteristics of a pool with "I" shaped baffle

initial shape (Figure 3.9). In the case of the optimal shape the two large recirculation regions at both sides of the slot are highly reduced (Figure 3.12). the obtained optimal points are $a = (0.6170, 0.1477)$ and $b = (0.8792, 0.0554)$.

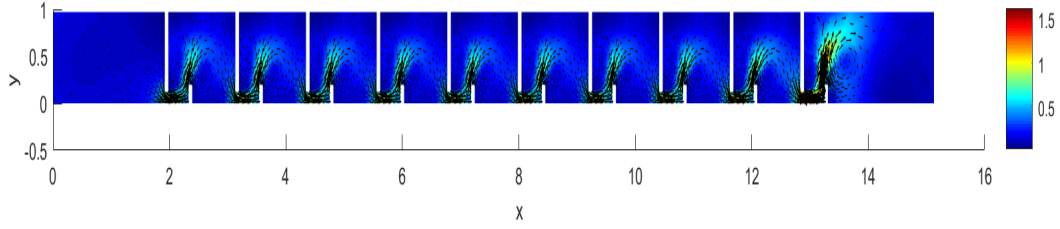


Figure 2.7: Non Optimal Shape and corresponding velocity

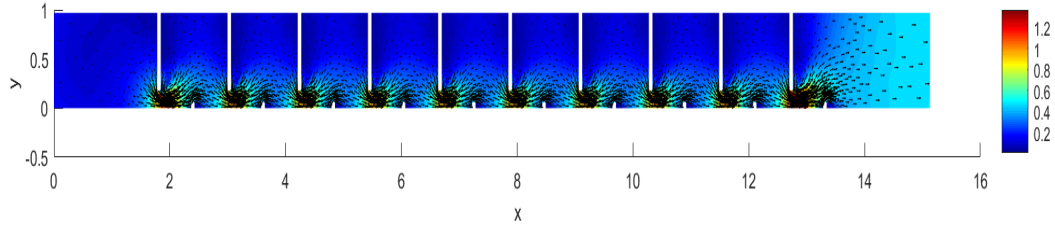


Figure 2.8: Optimal Shape based on spectral projected gradient algorithm and corresponding velocity

2.6.4.3 Multi-Slot Fishway ("L" shaped baffle)

The third test consists to consider the same rectangular form of fishway shown in Figure 2.2, but the two vertical baffles are replaced by three vertical baffles (Figure 3.13). The third slot is made to use it as a break waves.

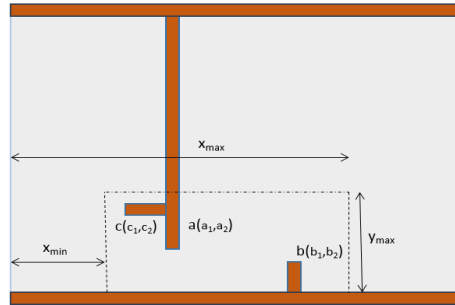


Figure 2.9: Geometric characteristics of a pool with "L" shaped baffle

We consider that the form of fishway depends only on the three points $a(a_1, a_2)$, $b(b_1, b_2)$ and $c(c_1, c_2)$. The constraints (2.2.3) are rewritten as

$$\begin{cases} \frac{1}{4} 1.213 \leq a_1, b_1, c_1 \leq \frac{3}{4} 1.213 \\ 0 \leq a_2, b_2, c_2 \leq \frac{1}{2} 0.97 \end{cases} \quad (2.6.13)$$

We replace the constraints (2.2.4) by the following constraints

$$\begin{cases} b_1 - a_1 \geq d_1 = 0.1 \\ a_2 - b_2 \geq d_2 = 0.05 \\ a_1 - c_1 \geq d_3 = \frac{1}{2}0.0305 \\ c_2 - a_2 \geq d_4 = \frac{1}{2}0.0305 \end{cases} \quad (2.6.14)$$

Finally, the objective function conserves the same writing. We collect all six ten constraints defined in (2.6.13)- (2.6.14) in a function $\vec{\Phi}_2 : \mathbb{R}^6 \rightarrow \mathbb{R}^{16}$ defined by:

$$\begin{aligned} \vec{\Phi}_2(a_1, a_2, b_1, b_2, c_1, c_2) = & \left(\frac{1}{4}1.213 - a_1, \frac{1}{4}1.213 - b_1, \frac{1}{4}1.213 - c_1, a_1 - \frac{3}{4}1.21, \right. \\ & b_1 - \frac{3}{4}1.213, c_1 - \frac{3}{4}1.213, -a_2, -b_2, -c_2, a_2 - \frac{1}{2}0.97, \\ & b_2 - \frac{1}{2}0.97, c_2 - \frac{1}{2}0.97, 0.1 - b_1 + a_1, 0.05 - a_2 + b_2, \\ & \left. \frac{1}{2}0.0305 - a_1 + c_1, \frac{1}{2}0.0305 + a_2 - c_2 \right) \end{aligned} \quad (2.6.15)$$

We use the free gradient algorithm defined previously (Nelder-Mead) to minimise the penalty function written as:

$$\Phi(y) = \bar{\Phi}_1(y) + \beta \sum_{j=1}^m \max\{(\vec{\Phi}_2(y))_j, \vec{0}\} \quad (2.6.16)$$

The numerical modeling was carried out using the same physical and numerical requirements as the previous tests. The time interval for the simulation was $T = 300s$. The obtained optimal points are $a = (0.5521, 0.1581)$, $b = (0.7770, 0.0818)$, $c = (0.3500, 0.2431)$.

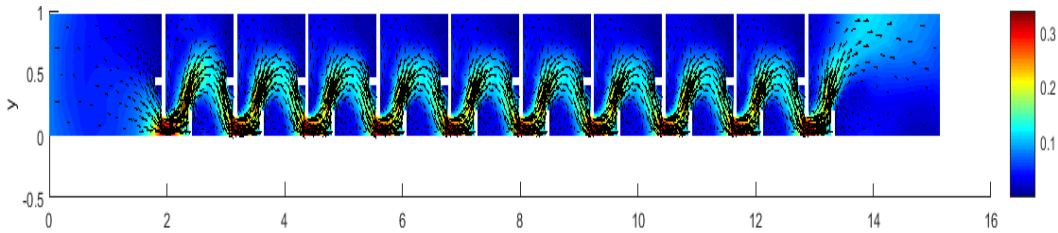


Figure 2.10: Non Optimal Shape and corresponding velocity

Figure 3.14 shows the water velocity at final time of the simulation corresponding to the initial random configuration. Three different regions are distinguished: a main flow region where the maximum velocities in the pool occur and two low velocity recirculation regions. The main flow first passed the slot, then flowed back after blocking by the "L" shaped baffle.

In the controlled case (Figure 3.16), the circulation areas near slots are strongly reduced. The results pointed out that design is better in terms of velocity and flow pattern.

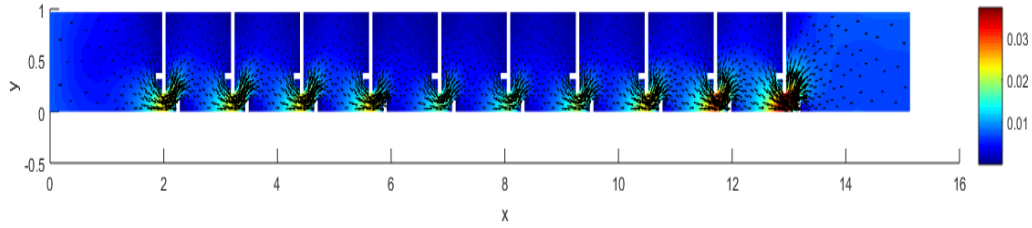


Figure 2.11: Optimal Shape based on gradient free algorithm and corresponding velocity

2.7 Conclusion

The optimal shape design techniques combined with a robust total variation diminishing scheme for solving the state system can be considered as useful tools for practical fishway design purpose.

Residence in an area of large scale turbulence will cause the fish to become confused, lose equilibrium and have a reduced swimming aptitude. Simulations of variant configurations provide a detailed flow structures in vertical slot fish ladder and allow to identify hydraulic issues and propose an appropriate type of construction. Of the three designs studied, design 3 is recommended for practical use, in terms of velocity and flow pattern.

Chapter 3

Shape optimization of fishways in multilayer flow based on optimal control theory

3.1 Introduction

Through centuries, dams are built for electricity, for navigation, or for agricultural reasons. These constructions are considered as barriers for some fish species, especially when crossing to the river upstream. For this, we need to ensure building a specific hydraulic structure to facilitate fish immigration between saline and fresh water called fishways.

Fishways are designed to enable migratory species to cross the river going to their feeding or growing areas. Those structures are useful to control the water flow and to assure a safe passage of small species. Different designs of fish ways exist; the pool and weir type [1], the Denil fishways [2], and the vertical slot type [9].

In this chapter we are interested in the vertical slot fishways, which are used frequently nowadays in a large range of hydraulic and biologic environments, as the structure is built with some slots vertical to the boundary of the channel, allowing a good control of the flow and the elevation of the water in the areas near from slots. This way of structure design allows migratory fish to cross dams more easily.

Considering flows with large friction coefficients or with important wind effects, the horizontal velocity can hardly be approximated by a vertical constant velocity as in the Saint-Venant system. A multilayer approach can be a good alternative to the computational issues associated with the free surface Navier-Stokes system. This procedure consists to

divide the water height to a number of layers where the hypothesis of Saint-Venant can be applied in each layer. The method leads to a precise description of the vertical profile of the horizontal velocity.

The multilayer Saint-Venant description of flow allows to derive an optimal fishway structures with respect to the shape of the variable domain for some given cost functionals modeling the ease of fish passage over the structure to their breeding, growing or feeding areas.

The chapter is organized as follows. The section 2 is devoted to the deduction of a general multilayer Saint-Venant system from the incompressible Navier-Stokes equations. In section 3 will look more closely to a particular model under the hypothesis of hydrostatic pressure. Section 4 deals with the analysis of the local in time existence and uniqueness of solution for the 2D multilayer system. A 3D version of the multilayer model is used to examine an optimal control problem applied to fishway structures, and derive the adjoint system of the multilayer system obtained from the hydrostatic hypothesis in section 5. Finally in the last section, we compare a free gradient and a gradient type algorithm with some numerical examples with structures already used for fishway modeling [6],[11]. The experiments prove a best control of the water flow in the fishway channel.

3.2 Multilayer approach

3.2.1 Preliminaries

In this section we are interested in the multilayer model which to be an approximation of the Navier-Stokes equations. This procedure consists in a discretization of the fluid domain into a number of layers and establishing a velocity vector and scalar pressure which are independent of the vertical variable. We suppose that the vertical velocity is linear with respect to the vertical variable z .

Let T be a given positive constant. We denote by Ω the fluid domain. The non-stationary Navier-Stokes system in a d -dimensional space ($d = 2,3$) takes the form:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \vec{u}) = 0 & \text{on } [0, T] \times \Omega \\ \partial_t (\rho \vec{u}) + \nabla \cdot (\rho \vec{u} \otimes \vec{u}) = \nabla \cdot \Sigma_T + \rho \mathbf{g} & \text{on } [0, T] \times \Omega \end{cases} \quad (3.2.1)$$

with the stress tensor symbol Σ_T is given by

$$\Sigma_T = -p \mathbf{Id} + \mu (\nabla \vec{u} + (\nabla \vec{u})')$$

and where $\mathbf{g} = (\mathbf{0}, -g)' \in \mathbb{R}^d$ is the gravity acceleration, and μ is a given viscosity.

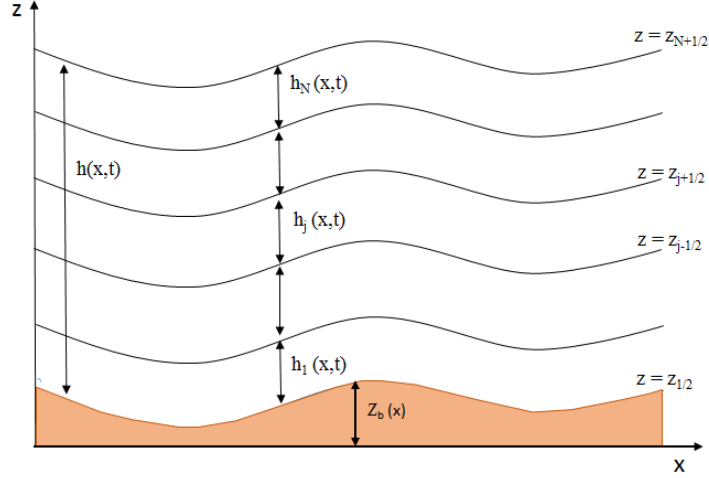


Figure 3.1: description of the multilayer division of the fluid domain.

For obtaining a multilayer system, we introduce a vertical discretization along the z direction of the water depth into $N \in \mathbb{N}^*$ layers of thickness $h_\alpha(t, x)$ with $N+1$ interfaces $\Gamma_{\alpha+1/2}(t)$ of equation $z = z_{\alpha+1/2}(t, x)$ for $\alpha = 0, 1, \dots, N$ (see figure 1). We denote by $z_b = z_{1/2}$ the equation of the bottom interface $\Gamma_b(t)$ and by $z_s = z_{N+1/2}$ the equation of the free surface interfaces $\Gamma_s(t)$. the thickness of each layer α is $h_\alpha = z_{\alpha+1/2} - z_{\alpha-1/2}$ where $z_{\alpha+1/2} = z_b + \sum_{\beta=1}^{\alpha} h_\beta$ for $\alpha = 1, \dots, N$. Finally the height of fluid is defined as

$$h = z_s - z_b = \sum_{\alpha=1}^N h_\alpha.$$

The boundary of the fluid domain is divided into $\Gamma_b(t)$ the bottom boundary, $\Gamma_s(t)$ the free surface boundary and $\Theta(t)$ the vertical inflow/outflow boundary. We set $\partial\Omega(t) = \Gamma_b(t) \cup \Gamma_s(t) \cup \Theta(t)$. The domain flow $\Omega(t)$ is consisting of a set of subdomain $\Omega_\alpha(t)$ in each layer. Let $\overline{\Omega}(t) = \cup_{\alpha=1}^N \overline{\Omega_\alpha}(t)$ where:

$$\begin{aligned} \Omega_\alpha(t) &= \{(x, z); x \in I_H(t) \text{ and } z_{\alpha-1/2} < z < z_{\alpha+1/2}\} \\ \partial\Omega_\alpha(t) &= \Gamma_{\alpha-1/2}(t) \cup \Gamma_{\alpha+1/2}(t) \cup \Theta_\alpha(t), \text{ with} \\ \Theta_\alpha(t) &= \{(x, z); x \in \partial I_H(t) \text{ and } z_{\alpha-1/2} < z < z_{\alpha+1/2}\} \end{aligned} \quad (3.2.2)$$

where $I_H(t)$ is the projection of the fluid domain onto the horizontal plane. The inflow/outflow boundary can be considered as a set of boundary domains in each layer α . It is given by the expression $\overline{\Theta}(t) = \cup_{\alpha=1}^N \overline{\Theta_\alpha}(t)$.

We introduce some notations:

- (i) the differential operator is defined as $\nabla = (\partial_{x_1}, \dots, \partial_{x_{d-1}}, \partial_z)$, then we set $\bar{\nabla} := (\partial_t; \nabla) = (\partial_t, \partial_{x_1}, \dots, \partial_{x_{d-1}}, \partial_z)$, and $\nabla = (\partial_{x_1}, \dots, \partial_{x_{d-1}})$

- (ii) we denote by $\vec{\eta}_{T,\alpha+1/2}$ the time-space unit normal vector to the interface $\Gamma_\alpha(t)$ and is given by

$$\vec{\eta}_{T,\alpha+1/2} = \frac{(\partial_t z_{\alpha+1/2}, \nabla_x z_{\alpha+1/2}, -1)'}{\sqrt{1 + |\partial_t z_{\alpha+1/2}|^2 + |\nabla_x z_{\alpha+1/2}|^2}}$$

- (iii) we denote by $\vec{\eta}_{\alpha+1/2}$ the space unit normal vector to the interface $\Gamma_\alpha(t)$, and is defined as

$$\vec{\eta}_{\alpha+1/2} = \frac{(\nabla_x z_{\alpha+1/2}, -1)'}{\sqrt{1 + |\nabla_x z_{\alpha+1/2}|^2}}$$

- (iv) if \mathbf{a} and \mathbf{b} are two tensors of sizes (n,m) and (n,p) respectively, $(\mathbf{a};\mathbf{b})$ is the tensor concatenation of \mathbf{a} and \mathbf{b} of size $(n,m+p)$

The following section gives some results about the jump condition of the mass and momentum conservation law.

3.2.2 Weak solution with discontinuities

Definition 3.2.1. *We consider that the velocity vector $\vec{\mathbf{u}}$, the pressure p and the density ρ are possibly discontinuous across the interfaces $\Gamma_{\alpha+1/2}(t)$ for $\alpha = 1, \dots, N-1$. Then $(\vec{\mathbf{u}}, p, \rho)$ is a weak solution of (2.1) if we have the following proprieties:*

(i) *In each layer $\Omega_\alpha(t)$ the triplet $(\vec{\mathbf{u}}, p, \rho)$ is a standard weak solution of (2.1).*

(ii) *$(\vec{\mathbf{u}}, p, \rho)$ verifies the following conditions at $\Gamma_{\alpha+1/2}(t)$, for $\alpha = 0, \dots, N$:*

(a) *For the mass conservation law,*

$$[(\rho; \rho \vec{\mathbf{u}})]_{|\Gamma_{\alpha+1/2}(t)} \cdot \vec{\eta}_{T,\alpha+1/2} = 0 \quad (3.2.3)$$

where $[(a; b)]_{|\Gamma_{\alpha+1/2}(t)}$ denotes the jump of the pair $(a; b)$ across $\Gamma_{\alpha+1/2}(t)$,

$$[(a; b)]_{|\Gamma_{\alpha+\frac{1}{2}}(t)} = ((a; b)_{|\Omega_{\alpha+1}} - (a; b)_{|\Omega_\alpha})_{|\Gamma_{\alpha+\frac{1}{2}}(t)}$$

(b) *For the momentum conservation law,*

$$[(\rho \vec{\mathbf{u}}; \rho \vec{\mathbf{u}} \otimes \vec{\mathbf{u}} - \Sigma_T)]_{|\Gamma_{\alpha+1/2}(t)} \cdot \vec{\eta}_{T,\alpha+1/2} = 0 \quad (3.2.4)$$

□

In the following we assume that the velocity vector and the pressure are independent of the vertical variable and the vertical velocity is linear with respect to z . And we write:

$$\vec{\mathbf{u}}|_{\Omega_\alpha(t)} := \vec{\mathbf{u}}_\alpha := (\vec{\mathbf{u}}_{H,\alpha}, w_\alpha)', \quad p_\alpha = p|_{\Omega_\alpha(t)}$$

with

$$\partial_z \vec{\mathbf{u}}_{H,\alpha} = 0, \quad \partial_z p_\alpha = 0, \quad \partial_z w_\alpha = d_\alpha(t, x) \quad (3.2.5)$$

for a smooth function $d_\alpha(t, x)$. Where $\vec{\mathbf{u}}_{H,\alpha}$ and w_α , respectively represent the horizontal and vertical velocities on layer α .

In order to satisfy the mass and momentum conservation two jump conditions at the interfaces must be verified

(i) Mass conservation jump conditions

The vector $\vec{\mathbf{u}}_{H,\alpha}$ verifies

$$\vec{\mathbf{u}}_{H,\alpha-1/2}^+(t, x) = \vec{\mathbf{u}}_{H,\alpha+1/2}^-(t, x) = \vec{\mathbf{u}}_{H,\alpha}(t, x) \quad (3.2.6)$$

where

$$u_{H,\alpha+1/2}^- := (u_{H,\alpha}|_{\Omega_\alpha(t)})|_{\Gamma_{\alpha+1/2}(t)} \text{ and } u_{H,\alpha+1/2}^+ := (u_{H,\alpha}|_{\Omega_{\alpha+1}(t)})|_{\Gamma_{\alpha+1/2}(t)}.$$

Then the mass conservation jump conditions are verified by $\vec{\mathbf{u}}$ if

$$G_{\alpha+1/2} := G_{\alpha+1/2}^- = G_{\alpha+1/2}^+ \quad (3.2.7)$$

where

$$\begin{cases} G_{\alpha+1/2}^+ = \partial_t z_{\alpha+1/2} + \vec{\mathbf{u}}_{H,\alpha+1} \cdot \nabla_x z_{\alpha+1/2} - w_{\alpha+1/2}^+ \\ G_{\alpha+1/2}^- = \partial_t z_{\alpha+1/2} + \vec{\mathbf{u}}_{H,\alpha} \cdot \nabla_x z_{\alpha+1/2} - w_{\alpha+1/2}^- \end{cases}$$

and $G_{\alpha+1/2}^+$ is the normal mass flux at the interface $\Gamma_{\alpha+1/2}(t)$

(ii) Momentum conservation jump conditions

We suppose that the tensor $D(\vec{\mathbf{u}}) = \nabla \vec{\mathbf{u}} + (\nabla \vec{\mathbf{u}})'$ has an approximation $D_{\alpha+1/2}^\pm$ at the interface $\Gamma_{\alpha+1/2}$. It verifies the following conditions:

$$\mu D_{\alpha+1/2}^\pm \cdot \vec{\mathbf{n}}_{\alpha+1/2} = \mu \tilde{D}_{\alpha+1/2} \cdot \vec{\mathbf{n}}_{\alpha+1/2} \pm \frac{1}{2} \frac{G_{\alpha+1/2}}{\sqrt{1 + |\nabla_x z_{\alpha+1/2}|^2}} [\vec{\mathbf{u}}]|_{\Gamma_{\alpha+1/2}(t)} \quad (3.2.8)$$

where

$$\tilde{D}_{\alpha+1/2} = \begin{pmatrix} D_H \left(\frac{\vec{\mathbf{u}}_{H,\alpha+1/2}^+ + \vec{\mathbf{u}}_{H,\alpha+1/2}^-}{2} \right) & \left(\nabla_x \left(\frac{w_{H,\alpha+1/2}^+ + w_{H,\alpha+1/2}^-}{2} \right) \right)' + \mathbf{Q}_{H,\alpha+1/2} \\ \left(\nabla_x \left(\frac{w_{H,\alpha+1/2}^+ + w_{H,\alpha+1/2}^-}{2} \right) \right) + \mathbf{Q}'_{H,\alpha+1/2} & 2Q_{v,\alpha+1/2} \end{pmatrix} \quad (3.2.9)$$

The unknown \vec{Q} is a new variable introduced for approximating the second order derivatives in z of \vec{u} . The following equation must be satisfied

$$\mathbf{Q} - \partial_z \vec{u} = 0, \quad \text{with } \mathbf{Q} = (\mathbf{Q}_H, Q_v) \quad (3.2.10)$$

In the following, we assume that the fluid is incompressible in all the fluid domain (i.e the density is a positive constant)

3.2.3 Vertical velocity

The horizontal velocity is determined and is independent of the vertical variable z . We use it to define the vertical velocity profile. We consider $z \in]z_{\alpha-1/2}, z_{\alpha+1/2}[$, the integration with respect to z of the first equation of system (3.2.1) gives

$$w_\alpha(t, x, z) = w_{\alpha-1/2}^+(t, x) - (z - z_{\alpha-1/2}) \nabla_x \cdot \vec{u}_{H,\alpha}(t, x), \quad \text{for } \alpha = 1, \dots, N$$

Using the conditions (3.2.7) at the interfaces, we obtain

$$w_{\alpha+\frac{1}{2}}^+ = (\vec{u}_{H,\alpha+1} - \vec{u}_{H,\alpha}) \cdot \nabla_x z_{\alpha+1/2} + w_{\alpha+1/2}^- \quad (3.2.11)$$

The algorithm in follow computes the vertical velocity using the horizontal velocities:

1. Let $G_{1/2}$ be known and using (3.2.7) at the bottom, The quantity $w_{1/2}^+$ writes

$$w_{\frac{1}{2}}^+ = \vec{u}_{H,1} \cdot \nabla_x z_B + \partial_t z_B - G_{1/2}$$

2. Then, for $\alpha = 1, \dots, N$ and $z \in]z_{\alpha-1/2}, z_{\alpha+1/2}[$, we have

$$\begin{cases} w_\alpha(t, x, z) = w_{\alpha-1/2}^+(t, x) - (z - z_{\alpha-1/2}) \nabla_x \cdot \vec{u}_{H,\alpha}(t, x) \\ w_{\alpha+1/2}^+ = (\vec{u}_{H,\alpha+1} - \vec{u}_{H,\alpha}) \cdot \nabla_x z_{\alpha+1/2} + w_{\alpha+1/2}^- \end{cases} \quad (3.2.12)$$

where

$$w_{\alpha+1/2}^- = w_{\alpha|\Gamma_{\alpha+1/2}(t)} = w_{\alpha-1/2}^+ - h_\alpha \nabla_x \cdot \vec{u}_{H,\alpha}.$$

Then the velocity vector \vec{u} writes $\vec{u}(t, x, z)|_{\Omega_\alpha(t)} = \vec{u}_\alpha(t, x, z)$ for $\alpha = 1, \dots, N$ with

$$\vec{u}_\alpha(t, x, z) = \left(\vec{u}_{H,\alpha}(t, x), w_{\alpha-1/2}^+(t, x) - (z - z_{\alpha-1/2}) \nabla_x \cdot \vec{u}_{H,\alpha}(t, x) \right)', \quad (3.2.13)$$

where $w_{\alpha-1/2}^+(t, x)$ comes from (3.2.12).

3.3 Hydrostatic pressure multilayer model

Now, we consider the hypothesis of hydrostatic pressure to construct our model. That is expressed as:

$$p_\alpha(t, x, z) = p_{\alpha+1/2}(t, x) + \rho g(z_{\alpha+1/2} - z) \quad (3.3.1)$$

with

$$p_{\alpha+1/2}(t, x) = p_S(t, x) + \rho g \sum_{\beta=\alpha+1}^N h_\beta(t, x) \quad (3.3.2)$$

where $p_{\alpha+1/2}$ is the kinematic pressure at the interface $\Gamma_{\alpha+1/2}(t)$, and p_S represents the free surface pressure.

The pressure variable is written in terms of layer depth in each layer α .

Remark 3.3.1. *The unknowns of the model are the layer depth h_α , and the velocities $\vec{\mathbf{u}}_\alpha$.*

Definition 3.3.1. *The weak formulation of (3.2.1) in $\Omega_\alpha(t)$ writes for all $\varphi \in L^2(\Omega_\alpha(t))$ and for all $\mathbf{v} \in H^1(\Omega_\alpha(t))^3$*

$$\left\{ \begin{array}{l} 0 \\ \int_{\Omega_\alpha(t)} \rho \mathbf{g} \cdot \mathbf{v} d\Omega \end{array} \right. = \left\{ \begin{array}{l} \int_{\Omega_\alpha(t)} (\nabla \cdot \vec{\mathbf{u}}_\alpha) \varphi d\Omega \\ \int_{\Omega_\alpha(t)} \rho \partial_t \vec{\mathbf{u}}_\alpha \cdot \mathbf{v} d\Omega + \int_{\Omega_\alpha(t)} \rho (\vec{\mathbf{u}}_\alpha \cdot \nabla \vec{\mathbf{u}}_\alpha) \cdot \mathbf{v} d\Omega \\ + \int_{\Omega_\alpha(t)} \mu (\nabla \vec{\mathbf{u}}_\alpha + (\nabla \vec{\mathbf{u}}_\alpha)') : \nabla \mathbf{v} d\Omega - \int_{\Omega_\alpha(t)} p_\alpha \nabla \cdot \mathbf{v} d\Omega \\ + \int_{\Gamma_{\alpha+1/2}(t)} \left(\left(\boldsymbol{\Sigma}_{T,\alpha+1/2}^- + p_{\alpha+1/2} \mathbf{Id} \right) \cdot \vec{\mathbf{n}}_{\alpha+1/2} \right) \cdot \mathbf{v} d\Gamma \\ - \int_{\Gamma_{\alpha-1/2}(t)} \left(\left(\boldsymbol{\Sigma}_{T,\alpha-1/2}^- + p_{\alpha-1/2} \mathbf{Id} \right) \cdot \vec{\mathbf{n}}_{\alpha-1/2} \right) \cdot \mathbf{v} d\Gamma, \end{array} \right. \quad (3.3.3)$$

We consider that THE velocity-pressure pairs verify (3.2.5), this leads to choose the test functions such that $\partial_z \varphi = 0$ and

$$\mathbf{v} = (\mathbf{v}_H(t, x), (z - z_B)V(t, x))' \quad (3.3.4)$$

where $\mathbf{v}_H(t, x)$ and $V(t, x)$ do not depend on z .

- Mass conservation

We consider a scalar test function $\varphi = \phi(t, x) \in L^2(\Omega_\alpha(t))$ independent of z . by integrating, projecting on the horizontal axis and taking into account (3.2.7), we obtain the mass conservation laws

$$\partial_t h_\alpha + \nabla_x \cdot (h_\alpha \vec{\mathbf{u}}_{H,\alpha}) = G_{\alpha+1/2} - G_{\alpha-1/2}, \quad \alpha = 1, \dots, N. \quad (3.3.5)$$

Where $G_{N+\frac{1}{2}}$ represents the mass exchange at the free surface, and the quantity $G_{1/2}$ represents the mass exchange at the bottom.

- Momentum conservation

We choose test functions $\vec{v} \in H^1(\Omega_\alpha)$ verifying (3.3.4). By respecting this structure of test function and projecting on the horizontal axis, then taking into account (3.2.8), we obtain the horizontal momentum conservation laws, for $\alpha = 1, \dots, N$,

$$\begin{aligned} & \rho \partial_t (h_\alpha \vec{u}_{H,\alpha}) + \rho \nabla_{x \cdot} (h_\alpha \vec{u}_{H,\alpha} \otimes \vec{u}_{H,\alpha}) - \nabla_{x \cdot} (\mu h_\alpha (\nabla_x \vec{u}_{H,\alpha} + (\nabla_x \vec{u}_{H,\alpha})')) \\ &= \frac{1}{2} \rho G_{\alpha+\frac{1}{2}} (\vec{u}_{H,\alpha+1} + \vec{u}_{H,\alpha}) - \frac{1}{2} \rho G_{\alpha-\frac{1}{2}} (\vec{u}_{H,\alpha} + \vec{u}_{H,\alpha-1}) \\ & - \mu (\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}}) - \rho g h_\alpha \nabla_x (z_B + h) - h_\alpha \nabla_x p_S \end{aligned} \quad (3.3.6)$$

where the term $\vec{K}_{\alpha+\frac{1}{2}}$ comes from the expression (3.2.8) and we have

$$\vec{K}_{\alpha+\frac{1}{2}} = \left(\sqrt{1 + |\nabla_x z_{\alpha+\frac{1}{2}}|^2} \right) [\tilde{D}_{\alpha+\frac{1}{2}} \cdot \vec{\eta}_{\alpha+\frac{1}{2}}]_H,$$

with $[\cdot]_H$ is the horizontal components of the vector. Then

$$\vec{K}_{\alpha+\frac{1}{2}} = D_H \left(\frac{\vec{u}_{H,\alpha+\frac{1}{2}}^+ + \vec{u}_{H,\alpha+\frac{1}{2}}^-}{2} \right) \cdot \nabla_x z_{\alpha+\frac{1}{2}} - \left(\nabla_x \left(\frac{w_{H,\alpha+\frac{1}{2}}^+ + w_{H,\alpha+\frac{1}{2}}^-}{2} \right) \right)' - \vec{Q}_{H,\alpha+\frac{1}{2}} \quad (3.3.7)$$

Finally, We introduce the kinematic viscosities $\nu = \frac{\mu}{\rho}$. With the hydrostatic hypothesis the multilayer system in d-dimensional case, for $\alpha = 1, \dots, N$ takes the form,

$$\left\{ \begin{array}{l} \partial_t h_\alpha + \nabla_{x \cdot} (h_\alpha \vec{u}_{H,\alpha}) = G_{\alpha+\frac{1}{2}} - G_{\alpha-\frac{1}{2}} \cdot \\ \rho \partial_t (h_\alpha \vec{u}_{H,\alpha}) + \rho \nabla_{x \cdot} (h_\alpha \vec{u}_{H,\alpha} \otimes \vec{u}_{H,\alpha}) - \nabla_{x \cdot} (\mu h_\alpha (\nabla_x \vec{u}_{H,\alpha} + (\nabla_x \vec{u}_{H,\alpha})')) \\ = \frac{1}{2} \rho G_{\alpha+\frac{1}{2}} (\vec{u}_{H,\alpha+1} + \vec{u}_{H,\alpha}) - \frac{1}{2} \rho G_{\alpha-\frac{1}{2}} (\vec{u}_{H,\alpha} + \vec{u}_{H,\alpha-1}) \\ - \mu (\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}}) - \rho g h_\alpha \nabla_x (z_B + h) - h_\alpha \nabla_x p_S. \end{array} \right. \quad (3.3.8)$$

with $\vec{K}_{\alpha+\frac{1}{2}}$ comes from (3.3.7).

After the obtaining of the horizontal velocity proprieties and the description of the vertical velocity profile in the general case, we are interested on the study of the multilayer system with a linear axis of the horizontal variable $x \in \mathbb{R}$. Some proprieties of existence and uniqueness of solution will be revealed in the following part.

3.4 Analysis of the 2D multilayer system

Let us now consider the 2D multilayer system which is defined as a particular case of the d-dimensional multilayer model (3.3.8). We study the existence and uniqueness of solution of this system in three parts; the first one is the estimate of the source terms, Next we treat a linearised system, and finally we use a Picard scheme to construct a convergent sequence to the solution of our model.

3.4.1 Mass conservation

Let us denote by h_α the water height in each layer and by h the height of water. And let for $\alpha = 1, \dots, N$, $h_\alpha = l_\alpha h$ with l_α a positive constant which represents the fraction of the water height,

$$\sum_{\alpha=1}^N l_\alpha = 1 \quad (3.4.1)$$

In this case, we can denote the horizontal velocities $\vec{u}_{H,\alpha}$ by u_α . If we use (3.4.1), and sum the equations (3.3.5) with respect to $\alpha = 1, \dots, N$, we obtain

$$G_{\alpha+\frac{1}{2}} - G_{1/2} = \sum_{\beta=1}^{\alpha} (\partial_t h_\beta + \partial_x (h_\beta u_\beta)) \quad (3.4.2)$$

We consider $G_{N+\frac{1}{2}} = 0$, for $\alpha = N$ the continuity equation writes

$$\partial_t h + \partial_x \left(h \sum_{\beta=1}^N l_\beta u_\beta \right) = -G_{1/2} \quad (3.4.3)$$

We will use the obtained results in the previous subsection to extract the momentum equation.

3.4.2 Momentum conservation

We use the equation (3.4.3) and rewriting (3.4.2) with a simple manner to get

$$\begin{aligned} G_{\alpha+\frac{1}{2}} &= G_{1/2} + \sum_{\beta=1}^{\alpha} l_\beta (\partial_t h + \partial_x (h u_\beta)) \\ &= G_{1/2} + \sum_{\beta=1}^{\alpha} l_\beta \left(\partial_x (h u_\beta) - \sum_{\gamma=1}^N \partial_x (l_\gamma h u_\gamma) - G_{1/2} \right) \end{aligned}$$

Consequently

$$G_{\alpha+\frac{1}{2}} = (1 - L_\alpha) G_{1/2} + \sum_{\gamma=1}^N \xi_{\alpha,\gamma} \partial_x (h u_\gamma), \quad \alpha = 1, \dots, N \quad (3.4.4)$$

where $L_\alpha := l_1 + \dots + l_\alpha$, and for $\alpha, \gamma \in \{1, \dots, N\}$:

$$\xi_{\alpha,\gamma} := \sum_{\beta=1}^{\alpha} (\delta_{\beta\gamma} - l_\beta) l_\gamma = \begin{cases} (1 - L_\alpha) l_\gamma & \text{if } \gamma \leq \alpha, \\ -L_\alpha l_\gamma & \text{otherwise} \end{cases}$$

where $\delta_{\beta\gamma}$ is the standard Kronecker symbol.

Now using (3.4.1) and the expression (3.4.4), the equation (3.3.6) takes the form

$$\begin{aligned}
& l_\alpha \partial_t (hu_\alpha) + l_\alpha \partial_x (hu_\alpha^2) - l_\alpha \partial_x (2\nu h \partial_x u_\alpha) + l_\alpha \partial_x \left[h \left(\frac{1}{\rho} p_S + g \frac{h}{2} \right) \right] - l_\alpha \frac{1}{\rho} p_S \partial_x h \\
& + \sum_{\gamma=1}^N \frac{1}{2} [(u_\alpha + u_{\alpha-1}) \xi_{\alpha-1,\gamma} - (u_{\alpha+1} + u_\alpha) \xi_{\alpha,\gamma}] \partial_x (hu_\gamma) \\
& = - l_\alpha g h \partial z_B - \nu \left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) \\
& - \frac{1}{2} [((u_\alpha + u_{\alpha-1})(1 - L_{\alpha-1}) - ((u_{\alpha+1} + u_\alpha)(1 - L_\alpha))] G_{1/2},
\end{aligned} \tag{3.4.5}$$

for $\alpha = 1, \dots, N$, with $L_0 = 0$ and

$$\begin{cases} \vec{K}_{\alpha+\frac{1}{2}} = 2\partial_x u_{\alpha+\frac{1}{2}} z_{\alpha+\frac{1}{2}} - \partial_z w_{\alpha+\frac{1}{2}} - \vec{Q}_{H,\alpha+\frac{1}{2}} \\ u_{\alpha+\frac{1}{2}} := \frac{u_{\alpha+\frac{1}{2}}^+ + u_{\alpha+\frac{1}{2}}^-}{2} \quad \text{and} \quad w_{\alpha+\frac{1}{2}} := \frac{w_{\alpha+\frac{1}{2}}^+ + w_{\alpha+\frac{1}{2}}^-}{2} \end{cases}$$

We define q_α variable of the horizontal discharges in the layer $\Omega_\alpha(t)$, for $\alpha = 1, \dots, N$. It is written as $q_\alpha = hu_\alpha$. Using (3.4.3) and (3.4.5) the model to be analysed is the following

$$\left\{ \begin{aligned} & \partial_t h + \partial_x \left(h \sum_{\beta=1}^N l_\beta u_\beta \right) = - G_{1/2} \\ & \partial_t (q_\alpha) + \partial_x \left(\frac{q_\alpha^2}{2} + g \frac{h^2}{2} + \frac{1}{\rho} p_S h - 2\nu (\partial_x q_\alpha - \frac{q_\alpha}{h} \partial_x h) \right) - \frac{1}{\rho} p_S \partial_x h \\ & + \sum_{\gamma=1}^N \frac{1}{2hl_\alpha} [(q_\alpha + q_{\alpha-1}) \xi_{\alpha-1,\gamma} - (q_{\alpha+1} + q_\alpha) \xi_{\alpha,\gamma}] \partial_x (q_\gamma) = g h \partial z_B \\ & - \frac{\nu}{l_\alpha} \left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) - \frac{1}{2hl_\alpha} [(q_\alpha + q_{\alpha-1})(1 - L_{\alpha-1}) - ((q_{\alpha+1} + q_\alpha)(1 - L_\alpha))] G_{1/2}, \end{aligned} \right. \tag{3.4.6}$$

The flow equation is based on the evaluation of the source term $\left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right)$. Therefore, the expression of the quantity $\left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right)$ can be expressed using the variables h_α , u_α , z_B and $Q_{H,\alpha}$ as follow

$$\begin{aligned}
& \left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) = \\
& 2\partial_x u_{\alpha+\frac{1}{2}} z_{\alpha+\frac{1}{2}} - 2\partial_x u_{\alpha-\frac{1}{2}} z_{\alpha-\frac{1}{2}} - \partial_x (w_{\alpha+\frac{1}{2}} - w_{\alpha-\frac{1}{2}}) - \left(Q_{H,\alpha+\frac{1}{2}} - Q_{H,\alpha-\frac{1}{2}} \right).
\end{aligned}$$

From the expression of the vertical velocity,

$$\begin{aligned}
w_{\alpha+\frac{1}{2}} - w_{\alpha-\frac{1}{2}} &= \frac{1}{2}(w_{\alpha+\frac{1}{2}}^+ + w_{\alpha+\frac{1}{2}}^- - w_{\alpha-\frac{1}{2}}^+ - w_{\alpha-\frac{1}{2}}^-) \\
&= \frac{1}{2} \left(2w_{\alpha+\frac{1}{2}}^- + (u_{\alpha+1} - u_{\alpha})\partial_x z_{\alpha+\frac{1}{2}} - 2w_{\alpha-\frac{1}{2}}^+ + (u_{\alpha} - u_{\alpha-1})\partial_x z_{\alpha-\frac{1}{2}} \right) \\
&= \frac{1}{2} \left((u_{\alpha+1} - u_{\alpha})\partial_x z_{\alpha+\frac{1}{2}} + (u_{\alpha} - u_{\alpha-1})\partial_x z_{\alpha-\frac{1}{2}} - 2h_{\alpha}\partial_x u_{\alpha} \right)
\end{aligned}$$

Observing that $z_{\alpha+\frac{1}{2}} = z_B + L_{\alpha}h$ we infer that

$$\begin{aligned}
\left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) &= \partial_x(u_{\alpha+1} + u_{\alpha})\partial_x(z_B + L_{\alpha}h) - \partial_x(u_{\alpha} + u_{\alpha-1})\partial_x(z_B + L_{\alpha-1}h) \\
&\quad + \partial_x(h_{\alpha}\partial_x u_{\alpha}) - \frac{1}{2}\partial_x((u_{\alpha+1} - u_{\alpha})\partial_x(z_B + L_{\alpha}h)) \\
&\quad - \frac{1}{2}\partial_x((u_{\alpha} - u_{\alpha-1})\partial_x(z_B + L_{\alpha-1}h)) - \left(Q_{H,\alpha+\frac{1}{2}} - Q_{H,\alpha-\frac{1}{2}} \right)
\end{aligned} \tag{3.4.7}$$

Finally, the expression (3.4.7) is expressed as

$$\begin{aligned}
\left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) &= \partial_x^2(h_{\alpha}u_{\alpha}) - \frac{1}{2}(u_{\alpha+1} - u_{\alpha-1})\partial_x^2 z_B \\
&\quad - \frac{1}{2}(L_{\alpha}u_{\alpha+1} + l_{\alpha}u_{\alpha} - L_{\alpha-1}u_{\alpha-1})\partial_x^2 h \\
&\quad + \frac{1}{2}\partial_x(u_{\alpha+1} - u_{\alpha-1})\partial_x z_B \\
&\quad + \frac{1}{2}\partial_x(L_{\alpha}u_{\alpha+1} + l_{\alpha}u_{\alpha} - L_{\alpha-1}u_{\alpha-1})\partial_x h \\
&\quad - \left(Q_{H,\alpha+\frac{1}{2}} - Q_{H,\alpha-\frac{1}{2}} \right)
\end{aligned} \tag{3.4.8}$$

where,

$$Q_{H,\alpha+\frac{1}{2}} = 2 \frac{u_{\alpha+1} - u_{\alpha}}{h_{\alpha}l_{\alpha} + h_{\alpha+1}l_{\alpha+1}}$$

The following section gives some results for the existence and uniqueness of the 2-D multilayer system. We begin by casting the system (3.4.6) into a system composed of two equations; the first one is a hyperbolic equations associated to the height of water, and an parabolic equation associated to the water flow.

3.4.3 Existence and uniqueness of solution model

The definition of the 2D multilayer system permits to study more of its characteristics. In particular, the existence and uniqueness of its solution. In order to achieve this, we

consider the following system:

$$\left\{ \begin{array}{l} \partial_t h + \partial_x \left(h \sum_{\beta=1}^N l_\beta u_\beta \right) = -G_{1/2} \\ \partial_t (h_\alpha u_\alpha) + \partial_x (h_\alpha u_\alpha^2) - \partial_x (2\nu h_\alpha \partial_x u_\alpha) \\ + \sum_{\gamma=1}^N \frac{1}{2} [(u_\alpha + u_{\alpha-1})\xi_{\alpha-1,\gamma} - (u_\alpha + u_{\alpha+1})\xi_{\alpha,\gamma}] \partial_x (h u_\gamma) \\ = -gh_\alpha \partial_x h - gh_\alpha \partial_x z_B - \frac{h_\alpha}{\rho} \partial_x p_S - \nu (\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}}) \\ - \frac{1}{2} [(u_\alpha + u_{\alpha-1})(1 - L_{\alpha-1}) - (u_{\alpha+1} + u_\alpha)(1 - L_\alpha)] G_{1/2} \end{array} \right. \quad (3.4.9)$$

In order to cast the system as parabolic type system, firstly we notice that

$$\begin{aligned} \left(\vec{K}_{\alpha+\frac{1}{2}} - \vec{K}_{\alpha-\frac{1}{2}} \right) &= h_\alpha \partial_{xx} u_\alpha + 2\partial_x h_\alpha \partial_x u_\alpha + u_\alpha \partial_{xx} h_\alpha - \frac{1}{2} (u_{\alpha+1} - u_{\alpha-1}) \partial_x^2 z_B \\ &\quad - \frac{1}{2} (L_\alpha u_{\alpha+1} + l_\alpha u_\alpha - L_{\alpha-1} u_{\alpha-1}) \partial_x^2 h + \frac{1}{2} \partial_x (u_{\alpha+1} - u_{\alpha-1}) \partial_x z_B \\ &\quad + \frac{1}{2} \partial_x (u_{\alpha+1} - u_{\alpha-1}) \partial_x z_B \\ &\quad + \frac{1}{2} \partial_x (L_\alpha u_{\alpha+1} + l_\alpha u_\alpha - L_{\alpha-1} u_{\alpha-1}) \partial_x h - \left(Q_{H,\alpha+\frac{1}{2}} - Q_{H,\alpha-\frac{1}{2}} \right) \\ &\quad - \left(Q_{H,\alpha+\frac{1}{2}} - Q_{H,\alpha-\frac{1}{2}} \right) \end{aligned} \quad (3.4.10)$$

Let $U = (u_1, u_2, \dots, u_N)$, the unknown couple (h, U) satisfy the following system

$$\left\{ \begin{array}{l} \partial_t h + \partial_x \left(h \sum_{\beta=1}^N l_\beta u_\beta \right) = -G_{1/2} \\ \partial_t U - \nu \partial_{xx} U = -g \partial_x h (1, \dots, 1)^T - g \partial_x z_B (1, \dots, 1)^T - \frac{1}{\rho} \partial_x p_S (1, \dots, 1)^T \\ + G_{1/2} \frac{U}{h} + \left(\sum_{\beta=1}^N l_\beta \partial_x u_\beta \right) U + \left(\sum_{\beta=1}^N l_\beta u_\beta \right) \frac{U}{h} \partial_x h \\ - \partial_x (U^2) - \frac{U^2}{h} \partial_x h + (\mathbf{T}U) \partial_x U + (\mathbf{T}U) U \frac{\partial_x h}{h} - \nu \frac{U}{h} \partial_{xx} h + \frac{\nu}{2} (\mathbf{V}U) \frac{\partial_{xx} z_B}{h} \\ + \frac{\nu}{2} (\mathbf{W}U) \frac{\partial_{xx} h}{h} - \frac{\nu}{2} (\mathbf{V} \partial_x U) \frac{\partial_x z_B}{h} - \frac{\nu}{2} (\mathbf{W} \partial_x U) \frac{\partial_x h}{h} + 2\nu \frac{(\mathbf{X}U)}{h^2} + 2\nu \frac{(\mathbf{Y}U)}{h^2} \\ - G_{1/2} \frac{(\mathbf{Z}U)}{h} \end{array} \right. \quad (3.4.11)$$

Where $\mathbf{T}, \mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}$, and \mathbf{Z} are matrices defined as:

$$\mathbf{T} = (T_{\alpha,\beta}) \text{ with } \begin{cases} T_{\alpha,\beta} = \frac{1}{2l_\alpha}(\xi_{\alpha,\beta} - \xi_{\alpha-1,\beta}), & \text{if } \alpha = \beta \\ T_{\alpha,\beta} = -\frac{1}{2l_\alpha}\xi_{\alpha-1,\beta}, & \text{if } \alpha = \beta + 1 \\ T_{\alpha,\beta} = \frac{1}{2l_\alpha}\xi_{\alpha,\beta}, & \text{if } \alpha = \beta - 1 \\ T_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases},$$

$$\mathbf{V} = (V_{\alpha,\beta}) \text{ with } \begin{cases} V_{\alpha,\beta} = \frac{1}{l_\alpha}, & \text{if } \alpha = \beta + 1 \\ V_{\alpha,\beta} = -\frac{1}{l_\alpha}, & \text{if } \alpha = \beta - 1 \\ V_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases}$$

$$\mathbf{W} = (W_{\alpha,\beta}) \text{ with } \begin{cases} W_{\alpha,\beta} = 1, & \text{if } \alpha = \beta \\ W_{\alpha,\beta} = -\frac{L_{\alpha-1}}{l_\alpha}, & \text{if } \alpha = \beta + 1 \\ W_{\alpha,\beta} = \frac{L_\alpha}{l_\alpha}, & \text{if } \alpha = \beta - 1 \\ W_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases}$$

$$\mathbf{X} = (X_{\alpha,\beta}) \text{ with } \begin{cases} X_{\alpha,\beta} = \frac{1}{l_\alpha(l_\alpha + l_{\alpha+1})}, & \text{if } \alpha = \beta \text{ for } \alpha = 1, \dots, N-1 \\ X_{\alpha,\beta} = -\frac{1}{l_\alpha(l_\alpha + l_{\alpha+1})}, & \text{if } \alpha = \beta - 1 \\ X_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases}$$

$$\mathbf{Y} = (Y_{\alpha,\beta}) \text{ with } \begin{cases} Y_{\alpha,\beta} = \frac{1}{l_\alpha(l_\alpha + l_{\alpha-1})}, & \text{if } \alpha = \beta \text{ for } \alpha = 2, \dots, N \\ Y_{\alpha,\beta} = -\frac{1}{l_\alpha(l_\alpha + l_{\alpha-1})}, & \text{if } \alpha = \beta + 1 \\ Y_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases}$$

and

$$\mathbf{Z} = (Z_{\alpha,\beta}) \text{ with } \begin{cases} Z_{\alpha,\beta} = l_\alpha, & \text{if } \alpha = \beta \text{ for } \alpha = 2, \dots, N \\ Z_{\alpha,\beta} = -(1 - L_\alpha), & \text{if } \alpha = \beta - 1 \\ Z_{\alpha,\beta} = 1 - L_{\alpha-1}, & \text{if } \alpha = \beta + 1 \\ Z_{\alpha,\beta} = 0, & \text{otherwise,} \end{cases}$$

Definition 3.4.1. 1. Let $U \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ and $h \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ we define the norm

$$\| (U, h) \|_{1_1}^2 = \| U \|_1^2 + \| h \|_1^2$$

2. Let $U \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ and $h \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ we define the norm

$$\| (U, h) \|_{1_2}^2 = \| U \|_1^2 + \| h \|_2^2$$

3. Let $U \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ and $h \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ we define the norm

$$\| (U, h) \|_{2_2}^2 = \| U \|_2^2 + \| h \|_2^2$$

4. Let $U \in C(0, T, \mathbf{H}^2(\mathbb{R}))$ and $h \in C(0, T, \mathbf{H}^3(\mathbb{R}))$ we define the norm

$$\| (U, h) \|_{2_3}^2 = \| U \|_2^2 + \| h \|_3^2$$

Theorem 3.4.1. We consider the system (3.4.11) with initial condition

$$(U, h)(0, x) = (U^0(x), h^0(x)) \in H^2(\mathbb{R}) \times H^3(\mathbb{R})$$

We suppose $z_b, p_S \in H^2(\mathbb{R})$, and we denote

$$E = 2 \left(\| (U^0, h^0) \|_{2_3} \right)$$

We suppose that there exists a positive constant η_0 such that $\inf_{x \in \mathbb{R}} h^0(x) \geq \eta_0 > 0$. Then there exists $T > 0$ such that the system (3.4.11) admits an unique strong solution (U, h) satisfying

$$1. U \in C(0, T, H^2(\mathbb{R})) \cap C^1(0, T, L^2(\mathbb{R})) \cap L^2(0, T, H^3(\mathbb{R}))$$

$$2. h \in C(0, T, H^2(\mathbb{R})) \cap C^1(0, T, H^1(\mathbb{R}))$$

$$3. \forall t \in [0, T]$$

$$\forall x \in \mathbb{R}, h(t, x) \geq \inf_{x \in \mathbb{R}} h^0(x)$$

$$\max \left(\| (U, h)(t, \cdot) \|_{2_2}, \left(\int_0^T \| U(\tau, \cdot) \|_3^2 d\tau \right)^{\frac{1}{2}} \right) \leq E$$

The proof of this theorem will be detailed in the following subsections. We begin by the estimation of source terms. Next, we treat a linearised multilayer system, and finally, we construct a convergent subsequence which converges to the system solution.

3.4.3.1 Estimates on the source terms

The system (3.4.11) cast in the form

$$\begin{cases} \partial_t U - \nu \partial_{xx} U = S \\ \partial_t h + \bar{u} \partial_x h = F \end{cases} \quad (3.4.12)$$

with

$$S = S_b + S_h + S_{p_S} + S_{nl}$$

$$F = -G_{1/2} - \partial_x(\bar{u})h$$

where

$$\begin{cases} \bar{u} = \sum_{\beta=1}^N l_\beta u_\beta \\ S_b = -g \partial_x z_B(1, \dots, 1) \\ S_h = -g \partial_x h(1, \dots, 1) \\ S_{p_S} = -g \partial_x p_S(1, \dots, 1) \\ S_{nl} = S_{nl}^{(1)} + S_{nl}^{(2)} + S_{nl}^{(3)} + S_{nl}^{(4)} + S_{nl}^{(5)} + S_{nl}^{(6)} + S_{nl}^{(7)} + S_{nl}^{(8)} + S_{nl}^{(9)} + S_{nl}^{(10)} \end{cases}$$

with

$$S_{nl}^{(1)} = G_{1/2}(\mathbf{I}_N - \mathbf{Z}) \frac{U}{h}; \quad S_{nl}^{(2)} = \left(\sum_{\beta=1}^N l_\beta \partial_x u_\beta \right) U; \quad S_{nl}^{(3)} = \left(\sum_{\beta=1}^N l_\beta u_\beta \right) \frac{\partial_x h}{h} U;$$

$$S_{nl}^{(4)} = ((\mathbf{T} - 2\mathbf{I}_N)U) \partial_x U; \quad S_{nl}^{(5)} = ((\mathbf{T} - \mathbf{I}_N)U) U \frac{\partial_x h}{h}; \quad S_{nl}^{(6)} = \frac{\nu}{2} ((\mathbf{W} - 2\mathbf{I}_N)U) \frac{\partial_{xx} h}{h};$$

$$S_{nl}^{(7)} = \frac{\nu}{2} \mathbf{V} U \frac{\partial_{xx} z_B}{h}; \quad S_{nl}^{(8)} = -\frac{\nu}{2} \mathbf{V} U \frac{\partial_x z_B}{h}; \quad S_{nl}^{(9)} = -\frac{\nu}{2} \mathbf{W} \partial_x U \frac{\partial_x h}{h};$$

$$S_{nl}^{(10)} = 2\nu(\mathbf{X} + \mathbf{Y}) \frac{U}{h^2}$$

Lemma 3.4.1. (*Moser inequality [5]*). *Let $f, g \in H^s(\mathbb{R})$ for $s \geq 1$. Then $fg \in H^s(\mathbb{R})$ and there exists $C_M > 0$ independent of f and g such that:*

$$\|fg\|_s \leq C_M \|f\|_s \|g\|_s \quad (3.4.13)$$

In what follows, we suppose that the bottom is non penetrable. Then the mass exchange at the bottom is $G_{1/2} = 0$. We also give some results about the estimates for the source terms

Proposition 3.4.1. *We suppose that $U(t, \cdot), h(t, \cdot) \in H^2(\mathbb{R})$ such that $h \geq \eta_0$, with η_0 a given constant. Then we have*

- *Let S, F defined as in (3.4.12), then $S, F \in H^1(\mathbb{R})$, and we obtain the following estimates*

$$\begin{aligned} \|S\|_1 \leq C(1 + \|(U, h)\|_{2_3} + \|(U, h)\|_{2_3}^2 + \|(U, h)\|_{2_3}^3) \|(U, h)\|_{2_3} \\ + g \|z_B\|_2 + \frac{1}{\rho} \|p_S\|_2 \end{aligned} \quad (3.4.14)$$

$$\|F\|_1 \leq C \|(U, h)\|_{2_3}^2 \quad (3.4.15)$$

- *Let $(U, h), (U', h') \in H^2(\mathbb{R}) \times H^3(\mathbb{R})$ such that*

$$\|(U, h)\|_{2_3} \leq E, \quad \|(U', h')\|_{2_3} \leq E, \quad \text{and } h, h' \geq \eta_0 > 0 \quad (3.4.16)$$

Where E , and η_0 are constants, then

$$\|S(U, h) - S(U', h')\|_1 \leq C(1 + E + E^2 + E^3) \|(U - U', h - h')\|_{2_3} \quad (3.4.17)$$

$$\|F(U, h) - F(U', h')\|_1 \leq C_{ME} \|(U - U', h - h')\|_{2_3} \quad (3.4.18)$$

Proof. Firstly, for more simplicity we denote $\|\cdot\|_{H^s(\mathbf{R})}$ by $\|\cdot\|_s$, for $s \geq 1$. And $\|\cdot\|_{L^2(\mathbf{R})}$ by $\|\cdot\|$

Estimates on F:

$$\begin{aligned} \|F\|_1 &= \|\partial_x(\bar{u})h\|_1 \\ &\leq C \|\partial_x(\bar{u})h\|_1 \\ &\leq C \|(U, h)\|_2^2 \quad \text{using (3.4.13)} \end{aligned}$$

For estimating the variation of $F(\cdot, \cdot)$, we write

$$\begin{aligned} \partial_x(\bar{u})h - \partial_x(\bar{u}')h' &= \partial_x(\bar{u} - \bar{u}')h + \partial_x(\bar{u}')h - \partial_x(\bar{u}')h' \\ &= \partial_x(\bar{u} - \bar{u}')h + \partial_x(\bar{u}')(h - h') \end{aligned}$$

which gives

$$\|F(U, h) - F(U', h')\|_1 \leq C_M \|U - U'\|_2 + C_M \|h - h'\|_2$$

Estimates on S: we begin by estimating S_h , S_b , and S_{ps} :

$$\| S_b \|_1 \leq g \| z_b \|_2$$

$$\| S_h \|_1 \leq g \| h \|_2$$

$$\| S_{ps} \|_1 \leq \frac{1}{\rho} \| p_S \|_2$$

For S_{nl} , we estimate all its components

$$\left\| \frac{U}{h} \right\|_1^2 = \left\| \frac{U}{h} \right\|^2 + \left\| \partial_x \left(\frac{U}{h} \right) \right\|^2$$

and

$$\partial_x \left(\frac{U}{h} \right) = \frac{h \cdot \partial_x U - \partial_x h \cdot U}{h^2} = \frac{\partial_x U}{h} - \frac{\partial_x h}{h^2} U$$

Then

$$\left\| \partial_x \left(\frac{U}{h} \right) \right\|^2 \leq 2 \left(\frac{\| \partial_x U \|^2}{\eta_0^2} + \frac{1}{\eta_0^4} \| U \|_\infty^2 \| \partial_x h \|^2 \right)$$

The Sobolev embedding of $L^\infty \hookrightarrow H^1$, yields to

$$\left\| \partial_x \left(\frac{U}{h} \right) \right\|^2 \leq 2 \frac{\| \partial_x U \|^2}{\eta_0^2} + \frac{1}{\eta_0^4} \| U \|_1^2 \| \partial_x h \|^2$$

Since

$$\left\| \frac{U}{h} \right\|_1 \leq \| U \|_1 (1 + \| h \|_1)$$

With a same approach, we obtain

$$\| S_{nl}^{(2)} \|_1 \leq C \| U \|_2^2$$

$$\| S_{nl}^{(3)} \|_1 \leq C(\eta_0) \| U \|_1^2 \| h \|_2 (1 + \| h \|_2)$$

$$\| S_{nl}^{(4)} \|_1 \leq C(\mathbf{T}) \| U \|_2^2$$

$$\| S_{nl}^{(5)} \|_1 \leq C(\mathbf{T}) \| U \|_1^2 \| h \|_2 (1 + \| h \|_2)$$

$$\| S_{nl}^{(6)} \|_1 \leq C(\nu, \mathbf{W}) \| U \|_1 \| \partial_x h \|_2 (1 + \| h \|_1)$$

$$\| S_{nl}^{(7)} \|_1 \leq C(\nu, \mathbf{V}, \eta_0, z_B) \| U \|_1 (1 + \| h \|_1)$$

$$\| S_{nl}^{(8)} \|_1 \leq C(\nu, \mathbf{V}, \eta_0, z_B) \| U \|_1 (1 + \| h \|_1)$$

$$\| S_{nl}^{(9)} \|_1 \leq C(\eta_0, \mathbf{W}, \nu) \| U \|_2 \| h \|_2 (1 + \| h \|_2)$$

$$\| S_{nl}^{(10)} \|_1 \leq C(\eta_0, \nu, \mathbf{X}, \mathbf{Y}) \| U \|_1 (1 + \| h \|_1)$$

Now let $S = S(U, h)$ and $S' = S(U', h')$, We compute the variation of $S(\dots)$

$$\begin{aligned}
S_h^i - S_h^{\prime i} &= -g\partial_x(h - h') \quad \forall i = 1, \dots, N \\
S_{nl}^{(2)} - S_{nl}^{\prime(2)} &= \left(\sum_{\beta=1}^N l_\beta \partial_x u_\beta \right) (U - U') + \left(\sum_{\beta=1}^N l_\beta (\partial_x u_\beta - \partial_x u'_\beta) \right) U' \\
S_{nl}^{(3)} - S_{nl}^{\prime(3)} &= \left(\sum_{\beta=1}^N l_\beta (u_\beta - u'_\beta) \right) \frac{U \partial_x h}{h} + \left(\sum_{\beta=1}^N l_\beta u'_\beta \right) \frac{\partial_x h}{h} (U - U') \\
&\quad + \left(\sum_{\beta=1}^N l_\beta u'_\beta \right) \frac{U'}{h} \partial_x (h - h') + \left(\sum_{\beta=1}^N l_\beta u'_\beta \right) \frac{U' \partial_x h'}{hh'} (h' - h) \\
S_{nl}^{(4)} - S_{nl}^{\prime(4)} &= ((\mathbf{T} - 2\mathbf{I}_N)U) \partial_x (U - U') + ((\mathbf{T} - 2\mathbf{I}_N)(U - U')) \partial_x U' \\
S_{nl}^{(5)} - S_{nl}^{\prime(5)} &= ((\mathbf{T} - \mathbf{I}_N)U) \frac{U}{h} \partial_x h + ((\mathbf{T} - \mathbf{I}_N)U') \frac{U - U'}{h} \partial_x h \\
&\quad + ((\mathbf{T} - \mathbf{I}_N)U') \frac{U'}{h} \partial_x (h - h') + ((\mathbf{T} - \mathbf{I}_N)U') \frac{U' (h' - h)}{hh'} \partial_x h' \\
S_{nl}^{(6)} - S_{nl}^{\prime(6)} &= \frac{\nu}{2} \left[((\mathbf{W} - 2\mathbf{I}_N)(U - U')) \frac{\partial_{xx} h}{h} + ((\mathbf{W} - 2\mathbf{I}_N)U') \frac{1}{h} \partial_{xx} (h - h') \right. \\
&\quad \left. + ((\mathbf{W} - 2\mathbf{I}_N)U') \partial_{xx} h' \frac{h' - h}{hh'} \right] \\
S_{nl}^{(7)} - S_{nl}^{\prime(7)} &= \frac{\nu}{2} \left((\mathbf{V}(U - U')) \frac{1}{h} + (\mathbf{V}U') \frac{h' - h}{hh'} \right) \partial_{xx} z_B \\
S_{nl}^{(8)} - S_{nl}^{\prime(8)} &= -\frac{\nu}{2} \left((\mathbf{V}(U - U')) \frac{1}{h} + (\mathbf{V}U') \frac{h' - h}{hh'} \right) \partial_x z_B \\
S_{nl}^{(9)} - S_{nl}^{\prime(9)} &= -\frac{\nu}{2} \left((\mathbf{W} \partial_x (U - U')) \frac{\partial_x h}{h} + (\mathbf{W} (\partial_x U')) \frac{1}{h} \partial_x (h - h') \right. \\
&\quad \left. + (\mathbf{W} (\partial_x U')) \partial_x h' \frac{h' - h}{hh'} \right) \\
S_{nl}^{(10)} - S_{nl}^{\prime(10)} &= 2\nu \left(((\mathbf{X} + \mathbf{Y})(U - U')) \frac{1}{h^2} + ((\mathbf{X} + \mathbf{Y})U') \frac{(h' - h)(h' + h)}{h^2 h'^2} \right)
\end{aligned}$$

Then using **Lemma 3.4.1**, and (3.4.16) yields

$$\begin{aligned}
\| S_h - S_h' \|_1 &\leq C \| h - h' \|_2 \\
\| S_{nl}^{(2)} - S_{nl}^{\prime(2)} \|_1 &\leq E \| U - U' \|_2 \\
\| S_{nl}^{(3)} - S_{nl}^{\prime(3)} \|_1 &\leq \frac{C}{\eta_0} E^2 \| U - U' \|_1 + \frac{C}{\eta_0} E^2 \| h - h' \|_2 + \frac{C}{\eta_0^2} E^3 \| h - h' \|_1 \\
\| S_{nl}^{(4)} - S_{nl}^{\prime(4)} \|_1 &\leq 2CE \| U - U' \|_2 \\
\| S_{nl}^{(5)} - S_{nl}^{\prime(5)} \|_1 &\leq 2\frac{C}{\eta_0} E^2 \| U - U' \|_1 + \frac{C}{\eta_0} E^2 \| h - h' \|_2 + \frac{C}{\eta_0^2} E^3 \| h - h' \|_1 \\
\| S_{nl}^{(6)} - S_{nl}^{\prime(6)} \|_1 &\leq \frac{C}{\eta_0} E \| U - U' \|_1 + \frac{C}{\eta_0} E \| h - h' \|_3 + \frac{C}{\eta_0^2} E^2 \| h - h' \|_1 \\
\| S_{nl}^{(7)} - S_{nl}^{\prime(7)} \|_1 &\leq \frac{C}{\eta_0} \| U - U' \|_1 + \frac{C}{\eta_0^2} E \| h - h' \|_1
\end{aligned}$$

$$\begin{aligned}
\| S_{nl}^{(8)} - S_{nl}'^{(8)} \|_1 &\leq \frac{C}{\eta_0} \| U - U' \|_1 + \frac{C}{\eta_0^2} E \| h - h' \|_1 \\
\| S_{nl}^{(9)} - S_{nl}'^{(9)} \|_1 &\leq \frac{C}{\eta_0} E \| U - U' \|_2 + \frac{C}{\eta_0} E \| h - h' \|_2 + \frac{C}{\eta_0^2} E^2 \| h - h' \|_1 \\
\| S_{nl}^{(10)} - S_{nl}'^{(10)} \|_1 &\leq \frac{C}{\eta_0^2} \| U - U' \|_1 + \frac{C}{\eta_0^4} E^2 \| h - h' \|_1
\end{aligned}$$

Adding these inequalities, we get (3.4.17). \square

3.4.3.2 Study of the linearised problem

Now, We consider the following linearised problem of the system (3.4.11):

$$\begin{cases} \partial_t U - \nu \partial_{xx} U = S(\tilde{U}, \tilde{h}, \partial_x \tilde{U}, \partial_x \tilde{h}, \partial_{xx} \tilde{h}) := \tilde{S} \\ \partial_t h + \tilde{u} \partial_x h = F(U, \tilde{h}) := \tilde{F} := -G_{1/2} - \partial_x(\tilde{u}) \tilde{h} \end{cases} \quad (3.4.19)$$

where \tilde{U} , and \tilde{h} are known. We separate the problem (3.4.19) into two Cauchy problems

$$\begin{cases} \partial_t U - \nu \partial_{xx} U := \tilde{S} \\ U(0, \cdot) = U^0(\cdot) \in H^2(\mathbb{R}) \end{cases} \quad (\text{Par})$$

and

$$\begin{cases} \partial_t h + \tilde{u} \partial_x h = \tilde{F} \\ h(0, \cdot) = h^0 \in H^2(\mathbb{R}) \end{cases} \quad (\text{Hyp})$$

Proposition 3.4.2. *Let $\tilde{S} \in C(0, T, \mathbf{H}^1(\mathbb{R}))$ for $T > 0$.*

1. *The initial value problem (Par) has an unique strong solution \mathbf{U} which satisfies:*

$$\mathbf{U} \in C(0, T; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R})) \quad (3.4.20)$$

2. *There exist two positive constants C_1 and C_2 , depending only on the viscosity such that for any $t \in [0, T]$:*

$$\| U(t, \cdot) \|_2^2 + C_1 \int_0^t \| U(\tau, \cdot) \|_3^2 d\tau \leq e^t \left(\| U^0 \|_2^2 + C_2 \int_0^t \| \tilde{S}(\tau, \cdot) \|_1^2 d\tau \right) \quad (3.4.21)$$

Proof. Existence: We note K^t the Green Kernel of the operator $\partial_t - \nu \partial_{xx}$. Then, the solution of problem (Par) is given by

$$\forall (t, x) \in [0, T] \times \mathbb{R}, \quad U(t, x) = [K^t * U^0](x) + \int_0^t K^{t-s} * \tilde{S}(s, \cdot) ds$$

Then $\mathbf{U} \in C(0, T; \mathbf{H}^2(\mathbb{R}))$. Moreover, we note that

$$\partial_t \mathbf{U} = \nu \partial_{xx} \mathbf{U} + \tilde{S} \in C(0, T; \mathbf{L}^2(\mathbb{R}))$$

Consequently

$$\mathbf{U} \in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R}))$$

Uniqueness. We multiply the first equation of system (Par) by \mathbf{U} and integrate with respect to x , we obtain, for any $\lambda > 0$:

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{U}\|^2 + \nu \|\partial_x \mathbf{U}\|^2 \leq \frac{\lambda}{2} \|\mathbf{U}\|^2 + \frac{1}{2\lambda} \|\tilde{S}\|^2$$

Derive (Par) with respect to x , multiplying by $\partial_x \mathbf{U}$ and integrating in space, yields, for any $\lambda > 0$:

$$\frac{1}{2} \frac{d}{dt} \|\partial_x \mathbf{U}\|^2 + \nu \|\partial_{xx} \mathbf{U}\|^2 \leq \frac{\lambda}{2} \|\partial_x \mathbf{U}\|^2 + \frac{1}{2\lambda} \|\partial_x \tilde{S}\|^2.$$

Finally, we derive (Par) two times with respect to x , multiply by $\partial_{xx} \mathbf{U}$ and integrate in space. We obtain an approximation in terms of derivative of \tilde{S} . After the integration of the source term, we have for any $\lambda > 0$,

$$\left| \int_{\mathbb{R}} \partial_{xx} \tilde{S} \partial_{xx} \mathbf{U} dx \right| = \left| \int_{\mathbb{R}} \partial_x \tilde{S} \partial_{xxx} \mathbf{U} dx \right| \leq \frac{\lambda}{2} \|\partial_{xxx} \mathbf{U}\|^2 + \frac{1}{2\lambda} \|\partial_x \tilde{S}\|^2$$

Let $C_1 := 2\nu - \lambda > 0$. By summing the previous inequalities we obtain:

$$\frac{d}{dt} \|\mathbf{U}\|_2^2 + C_1 \|\mathbf{U}\|_3^2 \leq \lambda \|\mathbf{U}\|_2^2 + \frac{1}{\lambda} \|\tilde{S}\|_1^2.$$

The Gronwall Lemma gives the estimates (3.4.21). The uniqueness of solution is given by the priori estimate (3.4.21). \square

Lemma 3.4.2. *We suppose $u \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$ for $T > 0$. Then there exists $C > 0$ such that*

$$\|\partial_{xx}(u\partial_x h) - u\partial_{xxx} h\| \leq C \sup_{0 \leq \tau \leq T} \|u(\tau, \cdot)\|_2 \|h\|_2, \quad \forall h \in L^\infty(0, T, \mathbf{H}^2(\mathbb{R}))$$

Proof. We compute the difference

$$\partial_{xx}(u\partial_x h) - u\partial_{xxx} h = 2\partial_x(u)\partial_{xx} h + \partial_{xx}(u)\partial_x h$$

Using the embedding of $\mathbf{L}^\infty \hookrightarrow \mathbf{H}^1$ we obtain

$$\|(\partial_{xx}(u\partial_x h) - u\partial_{xxx} h)(t, \cdot)\| \leq C \sup_{0 \leq \tau \leq T} \|u(\tau, \cdot)\|_2 \|h(t, \cdot)\|_2$$

\square

Lemma 3.4.3. *Let $T > 0$ such that $\tilde{F} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$, and $\tilde{u} \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$, $h^0 \in H^2(\mathbb{R})$*

Then

1. *The problem (Hyp) has an unique strong solution h such that:*

$$h \in C(0, t; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \quad (3.4.22)$$

2. *For all $t \in [0, T]$ we have the energy inequality for $k = 1, 2$:*

$$\|h(t, \cdot)\|_1 \leq e^{C_1 E_{\tilde{u}} t} \left(\|h^0\|_1 + \int_0^t e^{-C_1 E_{\tilde{u}} \tau} \|\tilde{F}(\tau, \cdot)\|_1 d\tau \right) \quad (3.4.23)$$

with

$$C_1 > 0, \text{ is a given constant and } E_{\tilde{u}} := \sup_{0 \leq t \leq T} \|\tilde{u}\|_2$$

Proof. Uniqueness. We multiply (Hyp) by h and integrate with respect to x we obtain

$$\frac{1}{2} \frac{d}{dt} \|h\|^2 = - \int_{\mathbb{R}} \tilde{u} \partial_x \left(\frac{h^2}{2} \right) + \int_{\mathbb{R}} \tilde{F} h$$

Using the Holder inequality applied to the right hand side and integrating by parts its first term, we get

$$\frac{1}{2} \frac{d}{dt} \|h\|^2 \leq \frac{1}{2} E_{\tilde{u}} \|h\|^2 + \|\tilde{F}\| \|h\|$$

We divide the previous inequality by $\|h\|$ we get

$$\frac{d}{dt} \|h\| \leq C E_{\tilde{u}} \|h\| + \|\tilde{F}\| \quad (3.4.24)$$

for some constant $C > 0$.

Deriving one time (Hyp) with respect to x , multiplying by $\partial_x h$ and thank to

$$\int_{\mathbb{R}} \partial_x (\tilde{u} \partial_x h) \partial_x h dx = - \frac{1}{2} \int_{\mathbb{R}} \tilde{u} (\partial_x h)^2 dx$$

We obtain the estimate

$$\frac{d}{dt} \|\partial_x h\| \leq C E_{\tilde{u}} \|\partial_x h\| + \|\partial_x \tilde{F}\| \quad (3.4.25)$$

By combining the inequalities (3.4.24) and (3.4.25) we have for some constant $C > 0$:

$$\frac{d}{dt} \|h\|_1 \leq C E_{\tilde{u}} \|h\|_1 + \|\tilde{F}\|_1 \quad (3.4.26)$$

The Gronwall lemma gives (3.4.23).

Existence. We use the characteristic method: we define

$$\begin{cases} \frac{d}{ds}\chi(s, x, t) = u(s, \chi(s, x, t)) \\ \chi(t_0, x, t) = x_0 \in \mathbb{R} \end{cases}$$

where $\chi(s, x, t)$ is a characteristic curve. Then the solution of (Hyp) takes the form:

$$\forall t \in [0, T], \forall y \in \mathbb{R}, \quad h(t, y) = h^0(\chi(0, y, t)) + \int_0^t \tilde{F}(\tau, y, t) d\tau$$

which gives the smoothness of h in $C(0, T; \mathbf{H}^1(\mathbb{R})) \cap \mathbf{C}^1(0, T; \mathbf{L}^2(\mathbb{R}))$. \square

Proposition 3.4.3. *Let $T > 0$ such that $\tilde{F} \in C(0, T; \mathbf{H}^2(\mathbb{R}))$, $\tilde{u} \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$ and $h^0 \in H^2(\mathbb{R})$*

Then

1. *The problem (Hyp) has an unique strong solution h such that:*

$$h \in C(0, t; \mathbf{H}^2(\mathbb{R})) \cap \mathbf{C}^1(0, T; \mathbf{H}^1(\mathbb{R})) \quad (3.4.27)$$

2. *For all $t \in [0, T]$ we have the energy inequality*

$$\|h(t, \cdot)\|_2 \leq e^{C_2 E_{\tilde{u}} t} \left(\|h^0\|_2 + \int_0^t e^{-C_2 E_{\tilde{u}} \tau} \|\tilde{F}(\tau, \cdot)\|_2 d\tau \right) \quad (3.4.28)$$

with

$$C_2 > 0, \text{ is a given constant and } E_{\tilde{u}} := \sup_{0 \leq t \leq T} \|\tilde{u}\|_2$$

Proof. Uniqueness. We differentiate (Hyp) two times with respect to x , we multiply by $\partial_{xx}h$ and we integrate in space:

$$\frac{1}{2} \frac{d}{dt} \|\partial_{xx}h\|^2 = - \int_{\mathbb{R}} \partial_{xx}(\tilde{u}\partial_x h) \partial_{xx}h + \int_{\mathbb{R}} \partial_{xx}\tilde{F} \partial_{xx}h$$

by integration by parts we have:

$$\int_{\mathbb{R}} u \partial_{xxx}h \partial_{xx}h = - \frac{1}{2} \int_{\mathbb{R}} u \partial_x((\partial_{xx}h)^2) = + \frac{1}{2} \int_{\mathbb{R}} \partial_x(u)(\partial_{xx}h)^2$$

Since

$$\frac{d}{dt} \|\partial_{xx}h\| \leq C E_{\tilde{u}} \|h\|_2 + \|\partial_{xx}\tilde{F}\| \quad (3.4.29)$$

Adding (3.4.24)-(3.4.26)-(3.4.29), the Gronwall lemma gives (3.4.28).

Existence. We differentiate (Hyp) with respect to x . We define $\phi := \partial_x h$, then it verifies

$$\begin{cases} \partial_t \phi + \tilde{u} \partial_x \phi = \partial_x \tilde{F} - \partial_x(\tilde{u})\phi \\ \phi(0, x) = \partial_x h^0(x) \in \mathbf{H}^1(\mathbb{R}) \end{cases}$$

We follow an iterative technique to solve the previous problem with the initial condition

$$\phi^0(t, x) = \partial_x h^0(x), \forall (t, x) \in [0, T] \times \mathbb{R}$$

Define, for $j \geq 1$, $\phi^{(j)}$ with

$$\begin{cases} \partial_t \phi^{(j)} + \bar{u} \partial_x \phi^{(j)} = \partial_x \tilde{F} - \partial_x(\bar{u}) \phi^{(j-1)} \\ \phi^{(j)}(0, x) = \partial_x h^0(x) \quad \forall x \in \mathbb{R} \end{cases}$$

Since

$$\| \partial_x \tilde{F} - \partial_x(\bar{u}) \phi^{(j-1)} \|_1 \leq \| \tilde{F} \|_2 + C E_{\bar{u}} \| \phi^{(j-1)} \|_1$$

thus, we have $\phi^{(j)} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$ according to the lemma 4.3. To prove the convergence of $(\phi^{(j)})_j$ to $\partial_x h$, we write

$$\partial_t (\phi^{(j+1)} - \phi^{(j)}) = \partial_x(\bar{u}) (\phi^{(j)} - \phi^{(j-1)})$$

if we use j times the energy estimate (3.4.28), we obtain

$$\begin{aligned} \| \phi^{(j+1)} - \phi^{(j)} \|_1 &\leq e^{C_2 E_{\bar{u}} t} \int_0^t e^{-C_2 E_{\bar{u}} \tau} C_2 E_{\bar{u}} \| \phi^{(j)} - \phi^{(j-1)} \|_1 d\tau \\ &\leq \dots \leq e^{C_2 E_{\bar{u}} t} \frac{(C_2 E_{\bar{u}} t)^j}{j!} \left(2 \| \partial_x h^0 \|_1 + \int_0^t e^{C_2 E_{\bar{u}} \tau} \| \partial_x \tilde{F}(\tau) \|_1 d\tau \right) \end{aligned}$$

The right term tends to zero as j goes to $+\infty$. This gives the convergence of $(\phi^{(j)})_j$ to $\partial_x h$ in \mathbf{H}^1 , and then the \mathbf{H}^2 -regularity of h . \square

Proposition 3.4.4. *Let $\tilde{S} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$, $\tilde{h} \in C(0, T; \mathbf{H}^3(\mathbb{R}))$, and $\bar{u} \in L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$, then*

1. *the system*

$$\begin{cases} \partial_t U - \nu \partial_{xx} U = \tilde{S} \\ \partial_t h + \bar{u} \partial_x h = \tilde{F} \end{cases}$$

has an unique strong solution (U, h) . which belong to

$$\begin{aligned} U &\in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R})) \\ h &\in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{H}^1(\mathbb{R})) \end{aligned} \tag{3.4.30}$$

2. If we denote $E_{\tilde{u}, \tilde{h}} = \max \left\{ \sup_{0 \leq t \leq T} \|\tilde{u}\|_2, \sup_{0 \leq t \leq T} \|\tilde{h}\|_2 \right\}$, there exists $C, K > 0$ such that: for any $t \in [0, T]$

$$\|(U, h)(t, \cdot)\|_{2_2} \leq K e^{C(1+E_{\tilde{u}, \tilde{h}})^2 t} \left[\|(U^0, h^0)\|_{2_2} + \left(\int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right)^{1/2} \right]$$

and

$$\left(\int_0^t \|U(\tau, \cdot)\|_3^2 d\tau \right)^{1/2} \leq K e^{C(1+E_{\tilde{u}, \tilde{h}})^2 t} \left[\|(U^0, h^0)\|_{2_2} + \left(\int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.31)$$

Proof. Uniqueness. According to the assumptions, the source term \tilde{F} of the hyperbolic equation belongs to

$$C(0, T; \mathbf{H}^1(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^2(\mathbb{R}))$$

and we observe that

$$\|\tilde{F}(\tau, \cdot)\|_2 \leq C_B(1 + E_{\tilde{u}, \tilde{h}}) \|U((\tau, \cdot))\|_3 \quad (3.4.32)$$

for some constant C_B depending of z_B .

The Cauchy Schwartz inequality gives:

$$\int_0^t e^{-C(k)E_{\tilde{u}, \tilde{h}}\tau} \|\tilde{F}(\tau, \cdot)\|_k d\tau \leq e^{C_4(1+E_{\tilde{u}, \tilde{h}})^2 t} \left(\int_0^t \|U(\tau, \cdot)\|_3^2 d\tau \right)^{1/2} \quad (3.4.33)$$

with C_4 is a constant depending on C_B and $C(k)$.

The inequality (3.4.21) gives the existence of two constant depending only on the viscosity C'_1, C'_2 , such that

$$\|U(t, \cdot)\|_2 \leq C'_1 e^{C'_2 t} \left[\|U^0\|_2 + \left(\int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.34)$$

and

$$\left(\int_0^t \|U(\tau, \cdot)\|_3^2 d\tau \right)^{1/2} \leq C'_1 e^{C'_2 t} \left[\|U^0\|_2 + \left(\int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.35)$$

Combining inequalities (3.4.34), (3.4.35) and (3.4.33) yields to

$$\int_0^t e^{-C(k)E_{\tilde{u}, \tilde{h}}\tau} \|\tilde{F}(\tau, \cdot)\|_k d\tau \leq C'_1 e^{C'_5(1+E_{\tilde{u}, \tilde{h}})^2 t} \left[\|U^0\|_2 + \left(\int_0^t \|\tilde{S}(\tau, \cdot)\|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.36)$$

with $C_5 = C'_2 + C_4$.

The estimations (3.4.31) result from inequalities (3.4.28) and (3.4.36). The uniqueness of the solution is proved by the priori estimate (3.4.31).

Existence. By the proposition 3.4.1 $\tilde{S} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$ and $\tilde{F} \in \mathbf{H}^1(\mathbb{R})$. The proposition 3.4.2 gives the existence and uniqueness of U such that

$$U \in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R}))$$

Since

$$\tilde{F} \in C(0, T; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^2(\mathbb{R}))$$

and by the lemma 3.4.3 and the proposition 3.4.3, we obtain the existence and the uniqueness of h which satisfies:

$$h \in C(0, T; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R}))$$

To get more regularity in space, we derive the problem with respect to x , and

$$\begin{cases} \partial_t(\partial_x U) - \nu \partial_{xx}(\partial_x U) = \partial_x \tilde{S} \\ \partial_t(\partial_x h) + \tilde{u} \partial_x(\partial_x h) = \partial_x(\tilde{F}) - \partial_x(\tilde{u}) \partial_x(h) \\ (\partial_x U^0, \partial_x h^0) \in \mathbf{H}^1(\mathbb{R}) \end{cases}$$

We can follow the same process as in the latest part of Proposition 3.4.3 to complete the proof. \square

Lemma 3.4.4. *Let $\tilde{S} \in C(0, T; \mathbf{L}^2(\mathbb{R}))$, $\tilde{h} \in C(0, T; \mathbf{H}^1(\mathbb{R}))$, and $\tilde{u} \in L^\infty(0, T; \mathbf{H}^1(\mathbb{R}))$, then*

1. *The system*

$$\begin{cases} \partial_t U - \nu \partial_{xx} U = \tilde{S} \\ \partial_t h + \tilde{u} \partial_x h = \tilde{F} \end{cases}$$

has an unique strong solution (U, h) . which belong to

$$\begin{aligned} U &\in C(0, T; \mathbf{H}^1(\mathbb{R})) \cap C^1(0, T; \mathbf{L}^2(\mathbb{R})) \\ h &\in C(0, T; \mathbf{H}^1(\mathbb{R})) \end{aligned} \tag{3.4.37}$$

2. *There exists $C, K > 0$ such that: for any $t \in [0, T]$*

$$\| (U, h)(t, \cdot) \|_{1_{-1}} \leq K e^{C(1+E_{\tilde{u}, \tilde{h}})^2 t} \left[\| (U^0, h^0) \|_{1_{-1}} + \left(\int_0^t \| \tilde{S}(\tau, \cdot) \|^2 d\tau \right)^{1/2} \right] \tag{3.4.38}$$

where

$$E_{\bar{u}, \bar{h}} = \max \left\{ \sup_{0 \leq t \leq T} \|\bar{u}\|_1, \sup_{0 \leq t \leq T} \|\bar{h}\|_1 \right\}$$

Proof. The proof of this lemma is similar to the one of the proposition 3.4.4. \square

For proving the existence of solution of the non-linear system (3.4.12), we construct a sequence which is solution of a linear problem and converges to the strong solution of the non linear problem. This complete the proof of the theorem (3.4.1).

3.4.3.3 Picard's iterative scheme

We define a recursive sequence $(U^{(j)}, h^{(j)}) = (u_1^{(j)}, \dots, u_N^{(j)}, h^{(j)})_{j \in \mathbb{N}}$ as follows:

- $\forall (t, x) \in [0, T] \times \mathbb{R}$, $(U^{(0)}, h^{(0)})(t, x) = (U^0, h^0)(x) \in \mathbf{H}^2(\mathbb{R}) \times \mathbf{H}^3(\mathbb{R})$
- $\forall j \in \mathbb{N}$, $(U^{(j+1)}, h^{(j+1)})$ is a solution of:

$$\begin{cases} \partial_t U^{(j+1)} - \nu \partial_{xx} U^{(j+1)} = S(j) & \text{on } [0, T] \times \mathbb{R} \\ \partial h^{(j+1)} - \bar{u}^{(j)} \partial_x h^{(j+1)} = F^{(j,j+1)} & \text{on } [0, T] \times \mathbb{R} \\ (U^{(j)}, h^{(j)})(0, \cdot) = (U^{(0)}, h^{(0)})(\cdot) & \text{on } \mathbb{R} \end{cases} \quad (3.4.39)$$

where for any $j \in \mathbb{N}$:

$$\begin{cases} S^{(j)} = S(U^{(j)}, h^{(j)}, \partial_x U^{(j)}, \partial_x h^{(j)}, \partial_{xx} h^{(j)}), \\ F^{(j,j+1)} = -G_{1/2} - \partial_x(\bar{u}^{(j+1)})h^{(j)} \end{cases}$$

We also define

$$\begin{cases} E_0 = 2 \| (U^0, h^0) \|_{2,3} \\ \eta_0 = \frac{1}{2} \inf_{x \in \mathbb{R}} h^0(x) \end{cases}$$

Proposition 3.4.5. *Let $(h^0, U^0) \in \mathbf{H}^3(\mathbb{R}) \times \mathbf{H}^2(\mathbb{R})$. Then there exists a small $T > 0$ such that*

- *The sequence $(U^{(j)}, h^{(j)})_{j \in \mathbb{N}}$ is well defined and satisfies, for any $t \in [0, T]$ and any $j \in \mathbb{N}$:*

$$\begin{aligned} U^{(j)} &\in C(0, T, \mathbf{H}^2(\mathbb{R})) \cap C^1(0, t; \mathbf{L}^2(\mathbb{R})) \cap L^2(0, T; \mathbf{H}^3(\mathbb{R})) \\ h^{(j)} &\in C(0, T; \mathbf{H}^2(\mathbb{R})) \cap C^1(0, T; \mathbf{H}^1(\mathbb{R})) \end{aligned} \quad (3.4.40)$$

- $\forall (t, x) \in [0, T] \times \mathbb{R}$ and $\forall j \in \mathbb{N}$, we have:

$$\begin{aligned} & \| (U^{(j)}, h^{(j)})(t, \cdot) \|_{2_2} \leq E_0 \\ & \left(\int_0^t \| U^{(j)}(\tau, \cdot) \|_3^2 d\tau \right)^{1/2} \leq E_0 \end{aligned} \quad (3.4.41)$$

$$h^{(j)}(t, x) \geq \eta_0 > 0 \quad (3.4.42)$$

For the proof, we introduce the following regularization result (for more details see [12])

Lemma 3.4.5. *Let $(U, h) \in C(0, T; \mathbf{H}^1(\mathbb{R}))$ then There exist two positive constants C and C' such that*

$$\| C_0^\varepsilon \|_1 \leq C \| U \|_2 \quad (3.4.43)$$

$$\| C_1^\varepsilon \|_1 \leq C' \| (h, U) \|_2 \quad (3.4.44)$$

Proof. (of proposition 3.4.5) Applying Proposition 3.4.4 on $(U^{(0)}, h^{(0)})$, gives the existence $(U^{(1)}, h^{(1)}) \in C(0, t; \mathbf{H}^2(\mathbb{R}))$ for any $t > 0$. The characteristic formula applied to $h^{(1)}$, yields, $\forall t > 0$:

$$\begin{aligned} h^{(1)}(t, y) &= h^0(\chi(0, y, t)) + \int_0^t F^{(0,1)}(s, \chi(s, y, t)) ds \\ &\geq 2\eta_0 + C(E_0)t \\ &\geq \eta_0 \end{aligned}$$

Which proves (3.4.42) for $j = 1$.

We apply the inequality (3.4.31) in Proposition 3.4.4 to $(U^{(1)}, h^{(1)})$, we get for any $t \leq T_1$:

$$\| (U^{(1)}, h^{(1)})(t, \cdot) \|_{2_2} \leq K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(\int_0^t \| S^{(0)}(\tau, \cdot) \|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.45)$$

and

$$\left(\int_0^t \| U^{(1)}(\tau, \cdot) \|_3^2 d\tau \right)^{1/2} \leq K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(\int_0^t \| S^{(0)}(\tau, \cdot) \|_1^2 d\tau \right)^{1/2} \right] \quad (3.4.46)$$

We use the Proposition 3.4.1 with $S^{(0)}$, we obtain

$$\max \left(\| (U^{(1)}, h^{(1)})(t, \cdot) \|_{2_2}, \left(\int_0^t \| U^{(1)}(\tau, \cdot) \|_3^2 d\tau \right)^{1/2} \right) \leq K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(C(E_0)(1 + E_0 + E_0^2 + E_0^3) + g \| z_b \|_2 + \frac{1}{\rho} \| p_S \|_2 \right) \sqrt{t} \right]$$

We define for $t \in \mathbb{R}^+$

$$\varsigma(t) = K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(C(E_0)(1 + E_0 + E_0^2 + E_0^3) + g \| z_b \|_2 + \frac{1}{\rho} \| p_S \|_2 \right) \sqrt{t} \right]$$

The continuity of ς regarding to time t allows to find $0 < T_1 = T_1(\eta_0, E_0, z_b, p_S)$ such that (3.4.41) is satisfied for any $t \leq T_1$. We choose $T := T_1$

Next, we pass from j to $j+1$. If for any $j \in \mathbb{N}$, $(U^{(j)}, h^{(j)})$ satisfies (3.4.40), (3.4.41) and (3.4.42) for any $t \leq T_1$, Proposition 4.4 shows the existence of $(U^{(j+1)}, h^{(j+1)})$.

We apply the inequality (3.4.31) to $(U^{(j+1)}, h^{(j+1)})$, we get with similar manner, for any $t \leq T_1$:

$$\| (U^{(j+1)}(t, \cdot), h^{(j+1)}(t, \cdot)) \|_{2_2} \leq K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(\int_0^t \| S^{(j)}(\tau, \cdot) \|_1^2 d\tau \right)^{1/2} \right]$$

Finally, we use the inequality revealed in Proposition 3.4.1 with $S^{(j)}$, we obtain

$$K e^{C(1+E_{\bar{u}, \bar{h}})^2 t} \left[\frac{E_0}{2} + \left(C(E_0)(1 + E_0 + E_0^2 + E_0^3) + g \| z_b \|_2 + \frac{1}{\rho} \| p_S \|_2 \right) \sqrt{t} \right] \| (U^{(j+1)}, h^{(j+1)})(t, \cdot) \|_{2_2} \leq$$

As the previous way, for $T = T_1$, The estimation (3.4.41) is satisfied for any $t \leq T_1$. We choose $T := T_1$.

The inequality (3.4.28) for rank $j+1$ can be shown as the same manner of rank 1, this ends the proof with $T := T_1$

Now, it remains to prove the convergence of the sequence built above. In this order, we will proceed as the following way, for $j > 1$ the couple $(U^{(j+1)} - U^{(j)}, h^{(j+1)} - h^{(j)})$ satisfies

$$\begin{cases} (\partial_t - \nu \partial_{xx}) (U^{(j+1)} - U^{(j)}) = S^{(j)} - S^{(j)} \\ \partial_t (h^{(j+1)} - h^{(j)}) - \bar{u}^{(j)} \partial_x (h^{(j+1)} - h^{(j)}) = F^{(j,j+1)} - F^{(j-1,j)} - \partial_x h^{(j)} (\bar{u}^{(j)} - \bar{u}^{(j-1)}) \\ (U^{(j+1)} - U^{(j)}, h^{(j+1)} - h^{(j)}) (0, \cdot) = 0 \end{cases} \quad (3.4.47)$$

The source term of the second equation can be written as

$$\begin{aligned}\tilde{F} &= F^{(j,j+1)} - F^{(j-1,j)} - \partial_x h^{(j)}(\bar{u}^{(j)} - \bar{u}^{(j-1)}) \\ &= -h^{(j)}(\bar{u}^{(j+1)} - \bar{u}^{(j)}) - \partial_x \bar{u}^{(j)}(h^{(j)} - h^{(j-1)}) - \partial_x h^{(j)}(\bar{u}^{(j)} - \bar{u}^{(j-1)})\end{aligned}$$

Using Lemma 3.4.1 yields to

$$\|\tilde{F}\|_{1\leq} \leq CE_0 \|(U^{(j)} - U^{(j-1)}, h^{(j)} - h^{(j-1)})\|_{1_{-1}} + C_B(1 + E_0) \|U^{(j+1)} - U^{(j)}\|_2 \quad (3.4.48)$$

and

$$\|\tilde{F}\|_{2\leq} \leq CE_0 \|(U^{(j)} - U^{(j-1)}, h^{(j)} - h^{(j-1)})\|_{2_{-2}} + C_B(1 + E_0) \|U^{(j+1)} - U^{(j)}\|_3 \quad (3.4.49)$$

With a similar manner as (3.4.17) we obtain

$$\|S^{(j)} - S^{(j-1)}\| \leq C(1 + E_0 + E_0^2 + E_0^3) \|(U^{(j)} - U^{(j-1)}, h^{(j)} - h^{(j-1)})\|_{1_{-1}}$$

Using (3.4.38) proved in the lemma 4.4, for any $t \leq T$ we have the following estimate

$$\begin{aligned}&\|(U^{(j+1)} - U^{(j)}, h^{(j+1)} - h^{(j)})(t, \cdot)\|_{1_{-1}} \\ &\leq C(E_0)e^{C(1+E_0)^2 t} \left(\int_0^t \|(U^{(j)} - U^{(j-1)}, h^{(j)} - h^{(j-1)})(\tau, \cdot)\|_{1_{-1}}^2 \right)^{1/2}\end{aligned}$$

We apply the previous inequality j times to get

$$\begin{aligned}&\|(U^{(j+1)} - U^{(j)}, h^{(j+1)} - h^{(j)})(t, \cdot)\|_{1_{-1}} \\ &\leq C(E_0)^j e^{Kt} \left(\sup_{0 \leq t \leq T} \|(U^{(1)} - U^{(0)}, h^{(1)} - h^{(0)})\|_{1_{-1}}^2 \right)^{1/2} \left(\frac{1}{K^j j!} \right)\end{aligned}$$

with $K = C(1 + E_0)^2$.

Which gives the existence of the subsequence $(U^{(j)}, h^{(j)})_j$ such as

$$(U^{(j)}, h^{(j)}) \xrightarrow{j \rightarrow \infty} (U, h) \text{ strongly in } C(0, t; \mathbf{H}^1(\mathbb{R}))$$

Moreover, Proposition 4.5 gives, up to subsequence, the convergence:

$$U^{(j)} \xrightarrow{j \rightarrow \infty} U \text{ weakly in } L^2(0, T; \mathbf{H}^3(\mathbb{R})),$$

while for every fixed $t \leq T$:

$$(U^{(j)}, h^{(j)}) \xrightarrow{j \rightarrow \infty} (U, h) \text{ weakly in } \mathbb{H}^2(\mathbb{R})$$

The solution (U, h) for the system (3.4.12) belongs to $C(0, T; \mathbf{H}^1(\mathbb{R})) \cap L^\infty(0, T; \mathbf{H}^2(\mathbb{R}))$, satisfying $\forall t, x \in [0, T] \times \mathbb{R}$

$$h(t, x) \geq \eta_0 > 0,$$

and

$$\max \left(\| (U, h)(t, \cdot) \|_{2,2} \left(\int_0^t \| U(\tau, \cdot) \|_3^2 d\tau \right)^{1/2} \right) \leq E_0$$

Finally, it remains to prove $(U, h) \in C(0, T; \mathbf{H}^2(\mathbb{R}))$. By regularizing: we consider $(U^\varepsilon, h^\varepsilon) = (\rho_\varepsilon * U, \rho_\varepsilon * h)$, where $\rho_\varepsilon *$ is the Friedrich's mollifier with respect to the space variable. If we apply $\rho_\varepsilon *$ to system (3.4.12) we get:

$$\begin{cases} \partial_t U^\varepsilon - \nu \partial_{xx} U^\varepsilon = S^\varepsilon + C_0^\varepsilon \\ \partial_t h^\varepsilon - \bar{u}^\varepsilon \partial_x h^\varepsilon = F^\varepsilon + C_1^\varepsilon \\ (U^\varepsilon, h^\varepsilon)(0, \cdot) = (\rho_\varepsilon * U^0, \rho_\varepsilon * h^0) \in C^\infty \end{cases} \quad (3.4.50)$$

where $S^\varepsilon = \rho_\varepsilon * S$ $F^\varepsilon = \rho_\varepsilon * F$ and

$$\begin{cases} C_0^\varepsilon = (\partial_t - \nu \partial_{xx})(U^\varepsilon) - \rho_\varepsilon * (\partial_t U - \nu \partial_{xx} U) \\ C_1^\varepsilon = (\partial_t h^\varepsilon - \partial_x(h^\varepsilon \bar{u}^\varepsilon)) - \rho_\varepsilon * (\partial_t h - \partial_x(h \bar{u})) \end{cases}$$

The lemma 3.4.5 gives as ε goes to zero:

$$\begin{cases} C_0^\varepsilon, C_1^\varepsilon \longrightarrow 0, \\ (U^\varepsilon, h^\varepsilon) \longrightarrow (U, h). \end{cases}$$

Therefore, at the uniform limit we have $(U, h) \in C(0, T; \mathbf{H}^2(\mathbb{R}))$. The energy estimate gives the uniqueness and this end the proof of theorem 3.4.1. \square

The study of the incompressible Navier-stokes equations inspires an approach which we have to use to establish a multilayer system, and show the existence and uniqueness of its solution. In the following section, we are interested in design optimisation of vertical slots fishway. For this aim, we will use the hydrostatic Navier-Stokes equations, and deduce the associated multilayer system.

At the first step, we consider the height of water as a subdivision of layers α , and find the height and the horizontal velocity of water in each layer α . The next step consists to find an optimal shape of fishway corresponding to a velocity close to a given target velocity. Finally, it remains to derive the adjoint system from the multilayer state system in each layer, associated to the hydrostatic Navier-Stokes equations.

3.5 3-D Multilayer system and fishway shape optimization

In order to treat an optimal problem requiring a 2-D horizontal velocity, we are oriented towards a generalization of the previous multilayer method to the 3-D hydrostatic Navier-Stokes equations.

We suppose that $\mathbf{u}(t, x, z) = (u, v, w)$ is the velocity vector and let $\mathbf{U}(t, x, z) = (u, v)$ be the horizontal velocity. The hydrostatic model of Navier-Stokes system writes

$$\begin{cases} \nabla \cdot \mathbf{u} = 0 \\ \frac{\partial \mathbf{U}}{\partial t} + \nabla_x \cdot (\mathbf{U} \otimes \mathbf{U}) + \frac{\partial(\mathbf{U}w)}{\partial z} + \nabla_x p = \mu \frac{\partial^2 \mathbf{U}}{\partial z^2} \\ \frac{\partial p}{\partial z} = -g \end{cases} \quad (3.5.1)$$

with the boundary conditions

$$\begin{aligned} w(t, x, z_b) &= 0 \\ \mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_b) &= k\mathbf{U}(t, x, y, z_b) \\ \frac{\partial \mathbf{U}}{\partial z}(t, x, h) &= 0 \\ p(t, x, h) &= 0 \end{aligned} \quad (3.5.2)$$

and the kinematic boundary condition

$$\frac{\partial h}{\partial t} + \mathbf{U}(t, x, h) \nabla_x h - w(t, x, h) = 0 \quad (3.5.3)$$

The combination of the third equation of the system (3.5.1) and the fourth boundary condition of (3.5.2) gives the expression of the pressure $p(t, x, z)$

$$p(t, x, z) = g(h - z) \quad (3.5.4)$$

The following subsection describes the technique for extracting the 3-D multilayer system from the system (3.5.1).

3.5.1 3-D Multilayer system

The water height is divided in the vertical direction z into a set of layers. In each layer α , for $\alpha = 1, \dots, N$, let us denote by h_α the water height, where

$$h_\alpha = z_{\alpha+1/2} - z_{\alpha-1/2}, \quad \text{with } z_{\alpha+1/2} = z_b + \sum_{\beta=1}^{\alpha} h_\beta$$

with z_b is the bathymetry of the bottom.

and we define the averaged horizontal velocity \mathbf{u}_α as

$$\mathbf{u}_\alpha(t, x) = \frac{1}{h_\alpha} \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} \mathbf{U}(t, x, z) dz$$

We introduce an indicator function $\theta_\alpha(t, x, z)$ in each layer

$$\theta_\alpha(t, x, z) = \begin{cases} 1 & \text{if } z_{\alpha-1/2} \leq z \leq z_{\alpha+1/2} \\ 0 & \text{otherwise} \end{cases}, \text{ for } \alpha = 1, \dots, N$$

which verifies

$$\frac{\partial \theta_\alpha}{\partial t} + \nabla_x \cdot (\theta_\alpha \mathbf{U}) + \frac{\partial \theta_\alpha w}{\partial z} = 0 \quad (3.5.5)$$

Integrating the equation (3.5.5) in the vertical direction z , taking into account the condition (3.5.3) in each layer α , and using of the definition of $\mathbf{u}_\alpha(t, x)$ yields

$$\frac{\partial h_\alpha}{\partial t} + \nabla_x \cdot (h_\alpha \mathbf{u}_\alpha) = 0$$

Now, we integrate the equation (3.5.4) in each layer α , the form of hydrostatic pressure corresponds to

$$\int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} \nabla_x p dz = \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} g \nabla_x h dz = g \nabla_x h \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} dz = gh_\alpha \nabla_x h \quad (3.5.6)$$

If we integrate the second equation of (3.5.1) in the vertical direction z , and take into account the kinematic boundary condition (3.5.3), we get for $\alpha = 1, \dots, N$

$$\begin{aligned} \frac{\partial}{\partial t} \int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} \mathbf{U} dz + \nabla_x \cdot \left(\int_{z_{\alpha-1/2}}^{z_{\alpha+1/2}} \mathbf{U} \otimes \mathbf{U} dz \right) + gh_\alpha \nabla_x h = -gh_\alpha \nabla_x z_b \\ + \mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_{\alpha+1/2}) - \mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_{\alpha-1/2}) \end{aligned} \quad (3.5.7)$$

We apply a finite difference method in the vertical direction to the term

$$\mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_{\alpha+1/2}) - \mu \frac{\partial \mathbf{U}}{\partial z}(t, x, z_{\alpha-1/2}),$$

we obtain for a horizontal velocity \mathbf{u}_α and water height h_α in each layer α , the following multilayer model of the system (3.5.1)

$$\left\{ \begin{array}{l} \frac{\partial h_\alpha}{\partial t} + \nabla_x \cdot (h_\alpha \mathbf{u}_\alpha) = 0 \\ \frac{\partial h_\alpha \mathbf{u}_\alpha}{\partial t} + \nabla_x \cdot (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + gh_\alpha \nabla_x h = -gh_\alpha \nabla_x z_b - k_\alpha \mathbf{u}_\alpha \\ + 2\nu_\alpha \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - 2\nu_{\alpha-1} \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \text{ for } \alpha = 1, \dots, N \end{array} \right. \quad (3.5.8)$$

with

$$k_\alpha = \begin{cases} k & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha \neq 1 \end{cases}, \nu_\alpha = \begin{cases} 0 & \text{if } \alpha = 0 \\ \mu & \text{if } \alpha = 1, \dots, N \\ 0 & \text{if } \alpha = N \end{cases}$$

We can write the previous system in terms of conservative components in the left hand side as follow:

$$\left\{ \begin{array}{l} \frac{\partial h_\alpha}{\partial t} + \nabla_{x \cdot} (h_\alpha \mathbf{u}_\alpha) = 0 \\ \frac{\partial h_\alpha \mathbf{u}_\alpha}{\partial t} + \nabla_{x \cdot} (h_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha) + \frac{g}{2} \nabla_x (h_\alpha h) = \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) - gh_\alpha \nabla_x z_b - k_\alpha \mathbf{u}_\alpha \\ + 2\nu_\alpha \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - 2\nu_{\alpha-1} \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \quad \text{for } \alpha = 1, \dots, N \end{array} \right. \quad (3.5.9)$$

In the following, we give a definition for the shape optimization problem. We also define the geometry of the fishway to be optimized and give the system which governs the channel.

3.5.2 Optimal problem

In this section we are interesting in fishway structures. They are composed of succession of pools. There are also two transition pools, one in the beginning and the second at the end of structure. Two baffles are built in each pool and are vertical to the lateral channel boundary. The fishway is built with a slop relative to the ground (Figure 3.6)

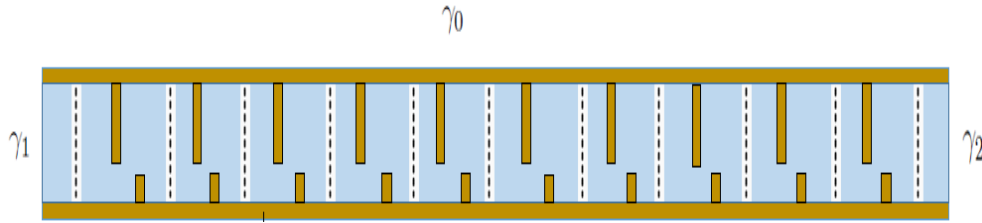


Figure 3.2: Prototype geometry: details of ten pools

We divide the water height into N layers, and we denote by h_α the height of water in each layer $\alpha = 1, \dots, N$. And by u_α the horizontal velocity in these layers.

If we denote by $\Omega \in \mathbb{R}^2$ the fishway channel, and by Ω_α the domain of layer α defined in (3.2.2). If we define $q_\alpha = h_\alpha u_\alpha$, then water height and horizontal flow in each layer α

along an interval time $[0, T]$ are governed by the following multilayer system

$$\left\{ \begin{array}{l} \frac{\partial h_\alpha}{\partial t} + \nabla_x \cdot \mathbf{q}_\alpha = 0 \quad \text{in } \Omega_\alpha \times [0, T] \\ \frac{\partial \mathbf{q}_\alpha}{\partial t} + \nabla_x \cdot \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \otimes \mathbf{q}_\alpha \right) + \frac{g}{2} \nabla_x (h_\alpha h) = \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) - g h_\alpha \nabla_x z_b - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \\ + 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} - 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \quad \text{in } \Omega_\alpha \times [0, T] \\ \text{for } \alpha = 1, \dots, N \end{array} \right. \quad (3.5.10)$$

We denote by γ_0 the lateral channel boundary, and by γ_α^1 and γ_α^2 the inflow/outflow boundary in each layer α respectively. Then the initial and boundary conditions are written as:

$$\left\{ \begin{array}{l} h_\alpha(0) = h_0, \quad q_\alpha(0) = q_0 \quad \text{in } \Omega_\alpha \\ q_\alpha \cdot \mathbf{n} = 0, \quad \text{curl} \left(\frac{q_\alpha}{h_\alpha} \right) = 0 \quad \text{in } \gamma_0 \times [0, T] \\ q_\alpha = q_1 \mathbf{n} \quad \text{in } \gamma_\alpha^1 \times [0, T] \\ h_\alpha = h_2 \quad \text{in } \gamma_\alpha^2 \times [0, T] \end{array} \right. \quad (3.5.11)$$

We take into account that the global geometry of each pool depends only of the two points a and b at the end of baffles (see figure 4.2). The position of these points change along flow for controlling the water velocity and permitting to maximum of fishes species to across the dam construction. In this order, two constraints are written:

$$\begin{aligned} \frac{1}{4} 1.213 \leq y_1, y_3 \leq \frac{3}{4} 1.213 \\ 0 \leq y_2, y_4 \leq \frac{1}{4} 0.97 \end{aligned} \quad (3.5.12)$$

Other constraints are adding for assuming a comfortable passage of a maximum number of fishes. These constraints are written as:

$$\begin{aligned} y_3 - y_1 \geq d_1 = 0.1 \\ y_2 - y_4 \geq d_2 = 0.05 \end{aligned} \quad (3.5.13)$$

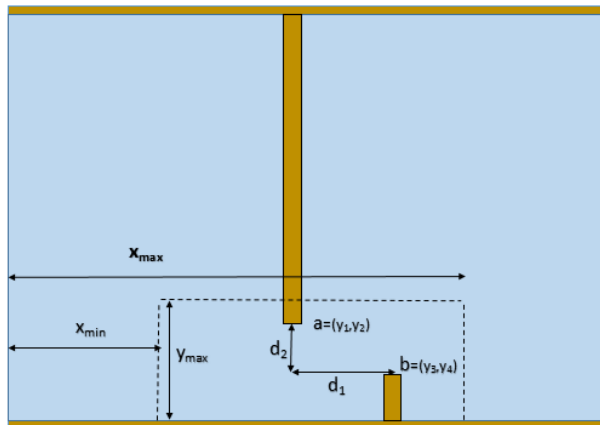


Figure 3.3: Prototype geometry: details of slot and pool

The optimization problem consists in the minimization of a functional $J(\cdot) \in \mathbb{R}$, also called objective function depending on the variables a and b (Figure 4.2). The purpose of our work is to establish an optimal fishway structure and obtaining a velocity close to a target velocity named \mathbf{v} in each layer α depended of fishes swimming ability. The velocity \mathbf{v} is given by:

$$\mathbf{v}(x_1, x_2) = \begin{cases} (c, 0) & \text{if } x_2 \leq \frac{1}{3} 0.97 \\ (0, 0) & \text{otherwise} \end{cases} \quad (3.5.14)$$

and the objective function in each layer is defined as:

$$j_\alpha = \frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left\| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right\|^2 dxdt + \frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} |curl(\frac{\mathbf{q}_\alpha}{h_\alpha})|^2 dxdt \quad (3.5.15)$$

with $\sigma \geq 0$ is the vorticity parameter, and $\mathbf{q}_\alpha = h_\alpha \mathbf{u}_\alpha$ where $(h_\alpha, \mathbf{q}_\alpha)$ is solution of the multilayer system (3.5.10) with the initial and boundary conditions (3.5.11).

The global objective function to be minimized takes the form

$$J = \sum_{\alpha=1}^N j_\alpha \quad (3.5.16)$$

In the following subsection, we derive an adjoint system related to the state system (3.5.10) and the objective function (3.5.15).

3.5.3 Adjoint system

We consider (p, \vec{r}) two test functions space in $L^\infty(0, T; H^2(\Omega))^2$. The variational formulation of the state system (3.5.10) in each layer α for $\alpha = 1, \dots, N$ is given by

Find $(h_\alpha, \mathbf{q}_\alpha) \in L^\infty(0, T; H^2(\Omega))^2$ such that

$$A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) = L_\alpha(\Omega_\alpha; p, \mathbf{r}) \quad \forall (p, \vec{r}) \in L^\infty(0, T; H^2(\Omega))^2$$

where

$$\begin{aligned} A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) &= \int_0^T \int_{\Omega_\alpha} \left\{ \frac{\partial h_\alpha}{\partial t} + \nabla \cdot \mathbf{q}_\alpha \right\} \cdot p \\ &+ \int_0^T \int_{\Omega_\alpha} \left\{ \frac{\partial \mathbf{q}_\alpha}{\partial t} + \nabla \cdot \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \otimes \mathbf{q}_\alpha \right) + \frac{g}{2} \nabla (h_\alpha h) \right\} \cdot \mathbf{r} \\ &- \int_0^T \int_{\Omega_\alpha} \left\{ \frac{g}{2} h^2 \nabla \left(\frac{h_\alpha}{h} \right) - g h_\alpha \nabla z_b - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \right\} \cdot \mathbf{r} \\ &- \int_0^T \int_{\Omega_\alpha} \left\{ 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} - 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \right\} \cdot \mathbf{r} \end{aligned}$$

We rewrite the cost function in each layer as $J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) = j_\alpha(\Omega)$, such that

$$J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) = \frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left\| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right\|^2 dt dx + \frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} \left| \operatorname{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right) \right|^2 dt dx \quad (3.5.17)$$

Theorem 3.5.1. *Let $(p, \vec{r}) \in L^\infty(0, T; H^2(\Omega))^2$ be two tests functions. The adjoint system of the equations (3.5.10) with initial and boundary conditions (3.5.11) writes*

$$\left\{ \begin{aligned} & -\frac{\partial p}{\partial t} + \frac{1}{h_\alpha^2} (\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} \cdot \mathbf{q}_\alpha + g \nabla h \cdot \mathbf{r} + g \nabla z_b \cdot \mathbf{r} - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha^2} \cdot \mathbf{r} \\ & - 2\nu_\alpha \frac{h_{\alpha+1}^3 \mathbf{q}_\alpha - h_\alpha^2 h_{\alpha+1} \mathbf{q}_{\alpha+1} + 2h_\alpha h_{\alpha+1}^2 \mathbf{q}_\alpha}{(h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1})^2} \cdot \mathbf{r} \\ & + 2\nu_{\alpha-1} \frac{2h_\alpha^2 h_{\alpha-1} \mathbf{q}_{\alpha-1} - 2h_\alpha h_{\alpha-1}^2 \mathbf{q}_\alpha - h_\alpha h_{\alpha-1}^2 \mathbf{q}_{\alpha-1} - h_{\alpha-1}^3 \mathbf{q}_\alpha}{(h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha)^2} \cdot \mathbf{r} \\ & = - \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) \frac{\mathbf{q}_\alpha}{h_\alpha^2} - \sigma \operatorname{curl}(\operatorname{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right)) \cdot \frac{\mathbf{q}_\alpha}{h_\alpha^2} \quad \text{in } \Omega_\alpha \times [0, T] \\ & -\frac{\partial \mathbf{r}}{\partial t} - \nabla p - \frac{1}{h_\alpha} (\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} - \frac{1}{h_\alpha} (\nabla \mathbf{r})^t \mathbf{q}_\alpha + k_\alpha \frac{\mathbf{r}}{h_\alpha} + 2\nu_\alpha \frac{h_{\alpha+1} \mathbf{r}}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \\ & + 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{r}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \\ & = \frac{1}{h_\alpha} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) + \sigma \frac{1}{h_\alpha} \operatorname{curl}(\operatorname{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right)) \quad \text{in } \Omega_\alpha \times [0, T] \end{aligned} \right. \quad (3.5.18)$$

with the final and boundary conditions

$$\begin{aligned}
p(T) &= 0, \quad \mathbf{r}(T) = 0 && \text{in } \Omega_\alpha \\
\mathbf{r} \cdot \mathbf{n} &= 0 && \text{in } \gamma_0 \times [0, T] \\
-\frac{q_1}{h_\alpha^2} \cdot \mathbf{r} \cdot \mathbf{n} &= 0 && \text{in } \gamma_1 \times [0, T] \\
\left(p + \frac{1}{h_2}(\mathbf{q}_\alpha \cdot \mathbf{r})\right) \mathbf{n} + \frac{1}{h_2}(\mathbf{q}_\alpha \cdot \mathbf{n}) \cdot \mathbf{r} - \frac{\sigma}{h_2} \operatorname{curl}\left(\frac{\mathbf{q}_\alpha}{h_2}\right) \boldsymbol{\tau} &= 0 && \text{in } \gamma_\alpha^2 \times [0, T]
\end{aligned} \tag{3.5.19}$$

where $(h_\alpha, \mathbf{q}_\alpha)$ is solution of (3.5.10)-(3.5.11)

Proof. We define a Lagrangian function as the following way:

$$L_{ag}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) = J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) - A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) + L_\alpha(\Omega_\alpha; p, \mathbf{r})$$

By applying an integration by parts on the function $A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$, deriving with respect to h_α the function A_α and J_α , and using the adjoint state of L_{ag} (see [4])

$$\frac{\partial}{\partial h_\alpha} L_{ag}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \tilde{h}_\alpha = 0$$

for any arbitrary direction \tilde{h}_α , to obtain the following system: in $\Omega_\alpha \times [0, T]$

$$\begin{aligned}
& -\frac{\partial p}{\partial t} + \frac{1}{h_\alpha^2}(\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} \cdot \mathbf{q}_\alpha + g \nabla h \cdot \mathbf{r} + g \nabla z_b \cdot \mathbf{r} - k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha^2} \cdot \mathbf{r} \\
& - 2\nu_\alpha \frac{h_{\alpha+1}^3 \mathbf{q}_\alpha - h_\alpha^2 h_{\alpha+1} \mathbf{q}_{\alpha+1} + 2h_\alpha h_{\alpha+1}^2 \mathbf{q}_\alpha}{(h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1})^2} \cdot \mathbf{r} \\
& + 2\nu_{\alpha-1} \frac{2h_\alpha^2 h_{\alpha-1} \mathbf{q}_{\alpha-1} - 2h_\alpha h_{\alpha-1}^2 \mathbf{q}_\alpha - h_\alpha h_{\alpha-1}^2 \mathbf{q}_{\alpha-1} - h_{\alpha-1}^3 \mathbf{q}_\alpha}{(h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha)^2} \cdot \mathbf{r} \\
& = - \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) \frac{\mathbf{q}_\alpha}{h_\alpha^2} - \sigma \operatorname{curl}\left(\operatorname{curl}\left(\frac{\mathbf{q}_\alpha}{h_\alpha}\right)\right) \cdot \frac{\mathbf{q}_\alpha}{h_\alpha^2}
\end{aligned}$$

with final and boundary conditions

$$\begin{aligned}
p(T) &= 0 && \text{in } \Omega_\alpha \\
\mathbf{r} \cdot \mathbf{n} &= 0 && \text{in } \gamma_0 \times [0, T] \\
-\frac{q_1}{h_\alpha^2} \cdot \mathbf{r} \cdot \mathbf{n} &= 0 && \text{in } \gamma_1 \times [0, T]
\end{aligned}$$

Next, we derive with respect to \mathbf{u}_α the functions A_α and J_α , then we use the adjoint state

$$\frac{\partial}{\partial \mathbf{q}_\alpha} L_{ag}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \tilde{\mathbf{q}}_\alpha = 0$$

for any arbitrary direction $\tilde{\mathbf{q}}_\alpha$, to get: in $\Omega_\alpha \times [0, T]$

$$\begin{aligned}
-\frac{\partial \mathbf{r}}{\partial t} - \nabla p - \frac{1}{h_\alpha}(\mathbf{q}_\alpha \cdot \nabla) \mathbf{r} - \frac{1}{h_\alpha}(\nabla \mathbf{r})^t \mathbf{q}_\alpha + k_\alpha \frac{\mathbf{r}}{h_\alpha} + 2\nu_\alpha \frac{h_{\alpha+1} \mathbf{r}}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \\
+ 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{r}}{h_{\alpha-1}^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} = \frac{1}{h_\alpha} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right) + \sigma \frac{1}{h_\alpha} \mathbf{curl}(\mathbf{curl}(\frac{\mathbf{q}_\alpha}{h_\alpha}))
\end{aligned}$$

with final and boundary conditions:

$$\begin{aligned}
\mathbf{r}(T) &= 0 && \text{in } \Omega_\alpha \\
\mathbf{r} \cdot \mathbf{n} &= 0 && \text{in } \gamma_0 \times [0, T] \\
(p + \frac{1}{h_2}(\mathbf{q}_\alpha \cdot \mathbf{r})) \mathbf{n} + \frac{1}{h_2}(\mathbf{q}_\alpha \cdot \mathbf{n}) \cdot \mathbf{r} - \frac{\sigma}{h_2} \mathbf{curl}(\frac{\mathbf{q}_\alpha}{h_2}) \boldsymbol{\tau} &= 0 && \text{in } \gamma_\alpha^2 \times [0, T]
\end{aligned}$$

□

The following subsection gives results about the domain derivative of the objective function (3.5.15) in each layer α in terms of adjoint and state variables.

3.5.4 Domain derivative

The study of the cost function will be done in each layer α . Now, let Ω_α^0 be a fixed domain, and we suppose that it is a bounded open subset of \mathbb{R}^2 with Lipschitz boundary. We denote by $Lip(\Omega_\alpha^0; \mathbb{R}^2)$ the Lipschitz applications domain, which is defined by

$$\begin{aligned}
Lip(\Omega_\alpha^0; \mathbb{R}^2) &= \{ \phi : \Omega_\alpha^0 \rightarrow \mathbb{R}^2 / \exists k \geq 0 \text{ such that} \\
&\| \phi(x) - \phi(y) \| \leq k \| x - y \|, \forall (x, y) \in \Omega_\alpha^0 \times \Omega_\alpha^0 \}
\end{aligned}$$

We denote by O_α^0 the subset of bi-Lipschitz homeomorphisms of Ω_α^0 . It is an open subset of $Lip(\Omega_\alpha^0; \mathbb{R}^2)$ and is given by

$$O_\alpha^0 = \{ \psi : \Omega_\alpha^0 \rightarrow \psi(\Omega_\alpha^0), \text{ bijective} / \psi \in Lip(\Omega_\alpha^0; \mathbb{R}^2), \psi^{-1} \in Lip(\psi(\Omega_\alpha^0); \mathbb{R}^2) \}$$

Then, the admissible domain denoted by X_α^0 is a set of domains Ω_α which are bound by bijective Lipschitz functions to the fixed domain Ω_α^0 . It is given as the follow way

$$X_\alpha^0 = \{ \Omega_\alpha = \psi(\Omega_\alpha^0) / \psi \in O_\alpha^0 \} \quad (3.5.20)$$

Theorem 3.5.2. *Let $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$, then the domain derivative of j_α at $\Omega_\alpha \in X_\alpha^0$ takes the form*

$$\begin{aligned}
\frac{\partial}{\partial \Omega_\alpha} j_\alpha(\Omega_\alpha) \cdot \mathbf{V} &= \frac{\partial}{\partial \Omega_\alpha} J_\alpha(\Omega_\alpha; h_\alpha, \mathbf{q}_\alpha) \cdot \mathbf{V} - \frac{\partial}{\partial \Omega_\alpha} A_\alpha(\Omega_\alpha; h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} \\
&+ \frac{\partial}{\partial \Omega_\alpha} L_\alpha(\Omega_\alpha; p, \mathbf{r}) \cdot \mathbf{V}
\end{aligned} \quad (3.5.21)$$

where $(h_\alpha, \mathbf{q}_\alpha)$ is solution of the state system (3.5.10) with initial and boundary conditions (3.5.11), and (p, \mathbf{r}) is solution of the adjoint system (3.5.18) with final and boundary conditions (3.5.19).

In the following, we give the computation details for each derivative term cited in the previous theorem.

Let $\Omega_\alpha \in X_\alpha^0$, such that $\Omega_\alpha = \psi(\Omega_\alpha^0)$ with $\psi \in O_\alpha^0$. We consider $\mathbf{F} \in Lip(\Omega_\alpha; \mathbb{R}^2)$ a bi-Lipschitz homeomorphism such that $\omega_\alpha = \mathbf{F}(\Omega_\alpha) \in X_\alpha^0$. Let $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$ such that $\mathbf{V} = \mathbf{F} - I$.

We define the function d by $d : \Omega_\alpha \in X_\alpha^0 \rightarrow d(\omega_\alpha) \in \mathbb{R}$, and we introduce a transported function called \underline{d} as the follow way

$$\underline{d} : \mathbf{F} \in Lip(\Omega_\alpha; \mathbb{R}^2) \rightarrow \underline{d}(\mathbf{F}) = d(\mathbf{F}(\Omega_\alpha)) = d(\omega_\alpha) \in \mathbb{R}.$$

Thus, the domain derivative of d at a given $\Omega_\alpha \in X_\alpha^0$ is defined by

$$\frac{\partial}{\partial \Omega_\alpha} d(\Omega_\alpha) \cdot \mathbf{V} = \frac{\partial}{\partial \mathbf{F}} \underline{d}(I) \cdot \mathbf{V}, \quad \text{for } \mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2) \quad (3.5.22)$$

The computation of the domain derivative of j_α is done by introducing the transported functions of A_α , L_α , and J_α . For this aim, we define the function A_α as a sum of a number of functions as follow

$$\begin{aligned} A_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) &= \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{\partial h_\alpha}{\partial t} p}_{A_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} (\nabla_x \cdot \mathbf{q}_\alpha) p}_{A_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{\partial \mathbf{q}_\alpha}{\partial t} \cdot \mathbf{r}}_{A_\alpha^3(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} \nabla_x \cdot \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \otimes \mathbf{q}_\alpha \right) \cdot \mathbf{r}}_{A_\alpha^4(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{g}{2} \nabla_x (h_\alpha h) \cdot \mathbf{r}}_{A_\alpha^5(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &- \underbrace{\int_0^T \int_{\Omega_\alpha} \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) \cdot \mathbf{r}}_{A_\alpha^6(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} + \underbrace{\int_0^T \int_{\Omega_\alpha} g h_\alpha \nabla_x z_b \cdot \mathbf{r}}_{A_\alpha^7(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r}}_{A_\alpha^8(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} - \underbrace{\int_0^T \int_{\Omega_\alpha} 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \cdot \mathbf{r}}_{A_\alpha^9(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \\ &+ \underbrace{\int_0^T \int_{\Omega_\alpha} 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \cdot \mathbf{r}}_{A_\alpha^{10}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})} \end{aligned}$$

We kept J_α as

$$J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) = \underbrace{\frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left\| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right\|^2}_{J_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha)} + \underbrace{\frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} \left| \operatorname{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right) \right|^2}_{J_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha)}$$

Next, for computing the domain derivative of each function above, we introduce its transported function. In particular for the first function

$$A_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) = \int_0^T \int_{\Omega_\alpha} \frac{\partial h_\alpha}{\partial t} p,$$

we introduce the transported function

$$\begin{aligned} \underline{A}_\alpha^1(\mathbf{F}, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) &= A_\alpha^1(\mathbf{F}(\Omega_\alpha), \underline{h}_\alpha \circ \mathbf{F}^{-1}, \underline{\mathbf{q}}_\alpha \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{\mathbf{r}} \circ \mathbf{F}^{-1}) \\ &= A_\alpha^1(\omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) = \int_0^T \int_{\Omega_\alpha} \frac{\partial \underline{h}_\alpha}{\partial t} \underline{p} \, | \det(\nabla_x \mathbf{F}) | \end{aligned}$$

As it is known, $\mathbf{F} \rightarrow | \det(\nabla_x \mathbf{F}) |$ is a differentiable application. Its derivative writes

$$\frac{\partial}{\partial \mathbf{F}} | \det(\nabla_x \mathbf{F}) | (I) \cdot \mathbf{V} = \nabla_x \cdot \mathbf{V}, \quad \text{for } \mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$$

Hence, for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned} \frac{\partial}{\partial \Omega_\alpha} A_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \frac{\partial}{\partial \mathbf{F}} \underline{A}_\alpha^1(I, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) \cdot \mathbf{V} \\ &= \int_0^T \int_{\Omega_\alpha} \frac{\partial h_\alpha}{\partial t} p (\nabla_x \cdot \mathbf{V}) \end{aligned}$$

For the second function $A_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$, we consider the following transported function

$$\begin{aligned} \underline{A}_\alpha^2(\mathbf{F}, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) &= A_\alpha^2(\mathbf{F}(\Omega_\alpha), \underline{h}_\alpha \circ \mathbf{F}^{-1}, \underline{\mathbf{q}}_\alpha \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{\mathbf{r}} \circ \mathbf{F}^{-1}) \\ &= A_\alpha^2(\omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \\ &= \int_0^T \int_{\Omega_\alpha} (\nabla_x \mathbf{F}^{-1} \circ \mathbf{F})^t : \nabla_x(\underline{\mathbf{q}}_\alpha) \underline{p} \, | \det(\nabla_x \mathbf{F}) | \end{aligned}$$

where the inner product "X:Y" is expressed as $X : Y = \sum_{k,l=1}^2 X_{kl} Y_{kl}$.

Taking into account that the application $\mathbf{F} \rightarrow \nabla_x \mathbf{F}^{-1} \circ \mathbf{F}$ is differentiable and the derivative is written as

$$\frac{\partial}{\partial \mathbf{F}} (\nabla_x \mathbf{F}^{-1} \circ \mathbf{F})(I) \cdot \mathbf{V} = -\nabla_x \cdot \mathbf{V}, \quad \text{for } \mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$$

Then, for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned} \frac{\partial}{\partial \Omega_\alpha} A_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \frac{\partial}{\partial \mathbf{F}} A_\alpha^2(I, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) \cdot \mathbf{V} \\ &= \int_0^T \int_{\Omega_\alpha} (\nabla_x \cdot \mathbf{q}_\alpha) p (\nabla_x \cdot \mathbf{V}) \\ &\quad - \int_0^T \int_{\Omega_\alpha} (\nabla_x \mathbf{V})^t : (\nabla_x \mathbf{q}_\alpha) p \end{aligned}$$

For the third function $A_\alpha^3(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$, we define the transported function

$$\begin{aligned} \underline{A}_\alpha^3(\mathbf{F}, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) &= A_\alpha^3(\mathbf{F}(\Omega_\alpha), \underline{h}_\alpha \circ \mathbf{F}^{-1}, \underline{\mathbf{q}}_\alpha \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{\mathbf{r}} \circ \mathbf{F}^{-1}) \\ &= A_\alpha^3(\omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) = \int_0^T \int_{\Omega_\alpha} \frac{\partial \underline{\mathbf{q}}_\alpha}{\partial t} \cdot \underline{\mathbf{r}} \mid \det(\nabla_x \mathbf{F}) \mid \end{aligned}$$

Consequently, for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned} \frac{\partial}{\partial \Omega_\alpha} A_\alpha^3(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \frac{\partial}{\partial \mathbf{F}} \underline{A}_\alpha^3(I, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) \cdot \mathbf{V} \\ &= \int_0^T \int_{\Omega_\alpha} \frac{\partial \underline{\mathbf{q}}_\alpha}{\partial t} \cdot \underline{\mathbf{r}} (\nabla_x \cdot \mathbf{V}) \end{aligned}$$

For the function $A_\alpha^4(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$, the associated transported function

$$\begin{aligned} \underline{A}_\alpha^4(\mathbf{F}, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) &= A_\alpha^4(\mathbf{F}(\Omega_\alpha), \underline{h}_\alpha \circ \mathbf{F}^{-1}, \underline{\mathbf{q}}_\alpha \circ \mathbf{F}^{-1}; \underline{p} \circ \mathbf{F}^{-1}, \underline{\mathbf{r}} \circ \mathbf{F}^{-1}) \\ &= A_\alpha^4(\omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \\ &= \int_0^T \int_{\Omega_\alpha} (\nabla_x \mathbf{F}^{-1} \circ \mathbf{F})^t : \nabla_x \left(\frac{\underline{\mathbf{q}}_\alpha}{\underline{h}_\alpha} \otimes \underline{\mathbf{q}}_\alpha \right) \cdot \underline{\mathbf{r}} \mid \det(\nabla_x \mathbf{F}) \mid \end{aligned}$$

As the same manner used for the function $A_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$ we obtain for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned}
\frac{\partial}{\partial \Omega_\alpha} A_\alpha^4(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \frac{\partial}{\partial \mathbf{F}} A_\alpha^4(I, \underline{h}_\alpha, \underline{\mathbf{q}}_\alpha; \underline{p}, \underline{\mathbf{r}}) \cdot \mathbf{V} \\
&= \int_0^T \int_{\Omega_\alpha} (\nabla_x \cdot \mathbf{q}_\alpha) \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r} (\nabla_x \mathbf{V}) \\
&\quad - \int_0^T \int_{\Omega_\alpha} (\nabla_x \mathbf{V})^t : (\nabla_x \mathbf{q}_\alpha) \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r} \\
&\quad + \int_0^T \int_{\Omega_\alpha} (\mathbf{q}_\alpha \cdot \nabla_x) \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r} (\nabla_x \mathbf{V}) \\
&\quad - \int_0^T \int_{\Omega_\alpha} (\nabla_x \mathbf{V} \mathbf{q}_\alpha \cdot \nabla_x) \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r}
\end{aligned}$$

Finally, we get the domain derivative of the other functions for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned}
\frac{\partial}{\partial \Omega_\alpha} A_\alpha^5(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \int_0^T \int_{\Omega_\alpha} \frac{g}{2} \nabla_x (h_\alpha h) \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V}) \\
&\quad - \int_0^T \int_{\Omega_\alpha} \frac{g}{2} (\nabla_x \mathbf{V})^t \nabla_x (h_\alpha h) \cdot \mathbf{r}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \Omega_\alpha} A_\alpha^6(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= \int_0^T \int_{\Omega_\alpha} \frac{g}{2} h^2 \nabla_x \left(\frac{h_\alpha}{h} \right) \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V}) \\
&\quad - \int_0^T \int_{\Omega_\alpha} \frac{g}{2} h^2 (\nabla_x \mathbf{V})^t \nabla_x \left(\frac{h_\alpha}{h} \right) \cdot \mathbf{r}
\end{aligned}$$

$$\frac{\partial}{\partial \Omega_\alpha} A_\alpha^7(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} = \int_0^T \int_{\Omega_\alpha} g h_\alpha \nabla_x z_b \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V})$$

$$\frac{\partial}{\partial \Omega_\alpha} A_\alpha^8(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} = \int_0^T \int_{\Omega_\alpha} k_\alpha \frac{\mathbf{q}_\alpha}{h_\alpha} \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V})$$

$$\frac{\partial}{\partial \Omega_\alpha} A_\alpha^9(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} = \int_0^T \int_{\Omega_\alpha} 2\nu_\alpha \frac{h_\alpha \mathbf{q}_{\alpha+1} - h_{\alpha+1} \mathbf{q}_\alpha}{h_{\alpha+1}^2 h_\alpha + h_\alpha^2 h_{\alpha+1}} \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V})$$

and

$$\frac{\partial}{\partial \Omega_\alpha} A_\alpha^{10}(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r}) \cdot \mathbf{V} = \int_0^T \int_{\Omega_\alpha} 2\nu_{\alpha-1} \frac{h_{\alpha-1} \mathbf{q}_\alpha - h_\alpha \mathbf{q}_{\alpha-1}}{h_\alpha^2 h_{\alpha-1} + h_{\alpha-1}^2 h_\alpha} \cdot \mathbf{r} (\nabla_x \cdot \mathbf{V})$$

For $J_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$ and $L_\alpha(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha; p, \mathbf{r})$, we follow the same approach and we get for $\mathbf{V} \in Lip(\Omega_\alpha; \mathbb{R}^2)$

$$\begin{aligned}\frac{\partial}{\partial \Omega_\alpha} J_\alpha^1(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) \cdot \mathbf{V} &= \frac{1}{2} \int_0^T \int_{\Omega_\alpha} \left| \frac{\mathbf{q}_\alpha}{h_\alpha} - \mathbf{v} \right|^2 (\nabla_x \cdot \mathbf{V}) \\ \frac{\partial}{\partial \Omega_\alpha} J_\alpha^2(\Omega_\alpha, h_\alpha, \mathbf{q}_\alpha) \cdot \mathbf{V} &= \frac{\sigma}{2} \int_0^T \int_{\Omega_\alpha} \left| \text{curl} \left(\frac{\mathbf{q}_\alpha}{h_\alpha} \right) \right|^2 (\nabla_x \cdot \mathbf{V}) \\ \frac{\partial}{\partial \Omega_\alpha} L_\alpha(\Omega_\alpha; p, \mathbf{r}) \cdot \mathbf{V} &= 0\end{aligned}$$

By using all the previous results in (3.5.21), we obtain the formula for the domain derivative of the objective function j_α at $\Omega_\alpha \in X_\alpha^0$ in each layer α .

3.6 Numerical resolution

In this section, we formulate the multilayer system (3.5.10) into a conservative system with source term. let us denote $\mathbf{X}_\alpha = (h_\alpha, \mathbf{q}_\alpha)^T$ with $\mathbf{q}_\alpha = (q_{\alpha,x}, q_{\alpha,y})$, then the system (3.5.10) can be written as

$$\frac{\partial \mathbf{X}_\alpha}{\partial t} + \nabla \cdot \mathbf{F}(h, \mathbf{X}_\alpha) = \mathbf{S}_p(h, \mathbf{X}_\alpha) + \mathbf{S}_z(\mathbf{X}_\alpha) + \mathbf{S}_\nu(\mathbf{X}_{\alpha-1}, \mathbf{X}_\alpha, \mathbf{X}_{\alpha+1}) \quad \text{for } \alpha = 1, \dots, N \quad (3.6.1)$$

with

$$\mathbf{F}(h, \mathbf{X}_\alpha) = \begin{pmatrix} q_{\alpha,x} & q_{\alpha,y} \\ \frac{q_{\alpha,x}^2}{h_\alpha} + \frac{g}{2} h_\alpha h & \frac{q_{\alpha,x} q_{\alpha,y}}{h_\alpha} \\ \frac{q_{\alpha,x} q_{\alpha,y}}{h_\alpha} & \frac{q_{\alpha,y}^2}{h_\alpha} + \frac{g}{2} h_\alpha h \end{pmatrix} \quad (3.6.2)$$

$$\mathbf{S}(h, \mathbf{X}_\alpha) = \begin{pmatrix} 0 \\ \frac{g}{2} h^2 \partial_x \left(\frac{h_\alpha}{h} \right) \\ \frac{g}{2} h^2 \partial_y \left(\frac{h_\alpha}{h} \right) \end{pmatrix}, \quad \mathbf{S}_p(\mathbf{X}_\alpha) = \begin{pmatrix} 0 \\ -gh_\alpha \partial_x z_b \\ -gh_\alpha \partial_y z_b \end{pmatrix} \quad (3.6.3)$$

and

$$\mathbf{S}_\nu(\mathbf{X}_{\alpha-1}, \mathbf{X}_\alpha, \mathbf{X}_{\alpha+1}) = \begin{pmatrix} 0 \\ -k_\alpha \mathbf{u}_\alpha + 2\nu_\alpha \frac{\mathbf{u}_{\alpha+1} - \mathbf{u}_\alpha}{h_{\alpha+1} + h_\alpha} - 2\nu_{\alpha-1} \frac{\mathbf{u}_\alpha - \mathbf{u}_{\alpha-1}}{h_\alpha + h_{\alpha-1}} \end{pmatrix} \quad (3.6.4)$$

The approximation of the system of conservation laws (3.6.1) is based on a finite volume type discretization.

3.6.1 Finite volume method

Finite volume schemes for the shallow water systems consist in using an upwinding of the fluxes. The problem domain is first discretized into a set of triangular cells T_i forming an unstructured computational mesh. Let Δt be the constant time step and define $t_n = n\Delta t$ for $n = 0, \dots, N$. At each discrete time t_n , we note $X_\alpha^{i,n}$ the approximated solution value. By integrating the equation (3.6.1) on a triangle T_i , we obtain

$$\int_{T_i} \partial_t X_\alpha + \int_{T_i} \nabla \cdot \mathbf{F}(h, \mathbf{X}_\alpha) = \int_{T_i} S(h, X_{\alpha+1}, X_\alpha, X_{\alpha-1}) \quad (3.6.5)$$

where $S(h, X_{\alpha+1}, X_\alpha, X_{\alpha-1}) = \mathbf{S}_p(h, \mathbf{X}_\alpha) + \mathbf{S}_z(\mathbf{X}_\alpha) + \mathbf{S}_\nu(\mathbf{X}_{\alpha-1}, \mathbf{X}_\alpha, \mathbf{X}_{\alpha+1})$, and \vec{n}_i is the normal on the edges of triangle T_i .

By applying the divergence formula, the equation (3.6.5) writes as

$$\int_{T_i} \partial_t X_\alpha + \int_{\partial T_i} \mathbf{F}(h, \mathbf{X}_\alpha) \cdot \vec{n}_i d\Gamma = \int_{T_i} S(h, X_{\alpha+1}, X_\alpha, X_{\alpha-1}) \quad (3.6.6)$$

Using a gauss formula for the second term of the left hand side to obtain

$$\int_{\partial T_i} \mathbf{F}(h, \mathbf{X}_\alpha) \cdot \vec{n}_i d\Gamma = \sum_{j=1}^3 F_{ij} \cdot n_{ij} \cdot dl_{ij}$$

The equation (3.6.6) takes the form

$$|T_i| \partial_t X_\alpha + \sum_{j=1}^3 F_{ij} \cdot n_{ij} \cdot dl_{ij} = |T_i| S \quad (3.6.7)$$

where n_{ij} is the normal on the edge T_i/T_j , F_{ij} are the discrete fluxes on the interface T_i/T_j and dl_{ij} is the length of the interface T_i/T_j .

Therefore, we now have an equation for each cell i of the form

$$\partial_t X_\alpha = -\frac{1}{|T_i|} \sum_{j=1}^3 F_{ij} \cdot n_{ij} \cdot dl_{ij} + S \quad (3.6.8)$$

A finite difference approximation is adopted to obtain the scheme

$$X_\alpha^{i,n+1} = X_\alpha^{i,n} - \frac{\Delta t}{|T_i|} \cdot \sum_{j=1}^3 F_{ij} \cdot n_{ij} \cdot dl_{ij} + \Delta t \cdot S \quad (3.6.9)$$

The Roe solver is used to evaluate the term $\sum_{j=1}^3 F_{ij} \cdot n_{ij} \cdot dl_{ij}$.

In what follow, we will present two approaches. The first one does not need the gradient of the cost function through its algorithm iterations. The second technique is called spectral projected gradient. It calls the gradient of the functional we want to minimize along the iterations. Finally we compare the results of each used technique.

3.6.2 A gradient free algorithm

The geometry of the vertical slot based on the use of guide elements to lead smooth hydraulic flow into the next slot.

The positioning of the guide elements was carried out at two different locations, $y = (a, b) = (y_1, y_2, y_3, y_4)$ (Figure 4.2), which configure the shape of the fish ladder. We formulate the shape optimization problem as a constrained problem depending on all linear constraints (3.5.12) and (3.5.13). We redefine the objective function (3.5.15) in the following way $\Phi_\alpha^1 : \mathbb{R}^4 \rightarrow \mathbb{R}$ where $\Phi_\alpha^1(y) = J_\alpha(\Omega_\alpha(y))$. The discrete approximation of the objective function in each layer α can be written as

$$\bar{\Phi}_\alpha^1(y) = \frac{\Delta t}{2} \sum_{n=1}^N \sum_{e \in T_i} \left[\int_e \|\mathbf{u}_\alpha^{i,n} - \vec{v}\|^2 + \alpha \int_e |\text{curl}(\mathbf{u}_\alpha^{i,n})|^2 \right] \quad (3.6.10)$$

Here $\mathbf{u}_\alpha^{i,n}$ is the discrete horizontal velocity obtained by the finite volume method (3.6.9). Next, we collect all linear constraints (3.5.12) and (3.5.13) in a function

$$\begin{aligned} \phi_\alpha^2 : \mathbb{R}^4 &\rightarrow \mathbb{R}^{10} \\ \phi_\alpha^2(y_1, y_2, y_3, y_4) &= \left(\frac{1}{4}1.213 - y_1, \frac{1}{4}1.213 - y_3, y_1 - \frac{3}{4}1.21, y_3 - \frac{3}{4}1.213, -y_2, \right. \\ &\quad \left. -y_4, y_2 - \frac{1}{2}0.97, y_4 - \frac{1}{2}0.97, 0.1 - y_3 + y_1, 0.05 - y_2 + y_4 \right) \end{aligned} \quad (3.6.11)$$

These constraints are satisfied if and only if $\phi_\alpha^2(y) \leq 0$. Lastly, we penalize the functional $\bar{\Phi}_\alpha^1$ as follow

$$\Phi_\alpha(y) = \bar{\Phi}_\alpha^1(y) + \beta \sum_{j=1}^{10} \max\{(\phi_\alpha^2(y))_j, 0\} \quad (3.6.12)$$

with β is a penalty parameter.

The discrete objective function to minimize writes

$$\Phi(y) = \sum_{\alpha=1}^N \Phi_{\alpha}(y) \quad (3.6.13)$$

For solving the non linear optimization problem, firstly we adopt a gradient free method named "Nelder-Mead" simplex technique. The Nelder-Mead technique, begins with an arbitrary simplex of 5 vertices y_1, y_2, \dots, y_5 . it evaluates and order the cost function on these vertices $\Phi(y_1) \leq \Phi(y_2) \leq \dots \leq \Phi(y_5)$. The vertex associated to the maximal value is replaced with a new point $y(\nu) = (1 + \nu)y^* - \nu y_5$, where y^* is the centroid of the convex hull $\{y_1, \dots, y_4\}$. The value of ν is chosen from this set of values: $\nu_{\delta} = -0.5$, $\nu_{\gamma} = 0.5$, $\nu_{\alpha} = 1$, $\nu_{\beta} = 2$. The choice of these values is determined according to the following algorithm

Calculate and sort $\Phi(y_1), \Phi(y_2), \dots, \Phi(y_5)$

While $|\Phi(y_5) - \Phi(y_1)|$ is not sufficiently small, calculate $y(\nu_{\beta})$ and $\Phi_{\beta} = \Phi(y(\nu_{\beta}))$ then

- a) If $\Phi_{\beta} \leq \Phi(y_1)$ then calculate $\Phi_{\alpha} = \Phi(y(\nu_{\alpha}))$. If $\Phi_{\beta} \leq \Phi_{\alpha}$, replace y_5 with $y(\nu_{\alpha})$; otherwise replace y_5 with $y(\nu_{\beta})$. Go to (f)
- b) If $\Phi(y_1) \leq \Phi_{\beta} \leq \Phi(y_4)$ then replace y_5 with y_{β} and go to (f)
- c) If $\Phi(y_4) \leq \Phi_{\beta} \leq \Phi(y_5)$, then calculate $\phi_{\gamma} = \Phi(y(\nu_{\gamma}))$. If $\Phi_{\gamma} \leq \Phi_{\beta}$ replace y_5 with $x(\nu_{\gamma})$ and go to (f). Otherwise go to (e)
- d) if $\Phi(y_5) \leq \Phi_{\beta}$ then calculate $\phi_{\delta} = \Phi(y(\nu_{\delta}))$. If $\Phi_{\delta} \leq \Phi_{y_5}$, replace y_5 with $y(\nu_{\delta})$ and go to (f). Otherwise go to (e)
- e) For $j = 2, \dots, 5$, set $y_j = y_1 + \frac{1}{2}(y_j - y_1)$
- f) Resort values of Φ at each resulting vertex

To prevent stagnation at non-optimal point, a modification proposed by Kelley [3] is used. This technique consists to replace the current simplex by a smaller one.

3.6.3 Spectral projected gradient

The second adopted approach to solve the shape optimization problem is consisting in the use of a gradient technique called Spectral projected gradient. This method use the projection into a closed and convex ω_α subset of \mathbb{R}^4 of all the points $y \in \mathbb{R}^4$ satisfying (5.12)-(5.13) in each layer α .

We denote $a_1 = a_3 = \frac{1}{4}1.213$, $b_1 = b_3 = \frac{3}{4}1.213$, $a_2 = a_4 = 0$, $b_2 = b_4 = \frac{1}{4}0.97$. Then the admissible set ω_α is defined as

$$\omega_\alpha = \{y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : \begin{aligned} & a_i \leq y_i \leq b_i, \quad i = 1, \dots, 4, \\ & y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{aligned}\} \quad (3.6.14)$$

The optimization problem can be reformulated as

$$\min_{y \in \omega_\alpha} J_\alpha(y). \quad (3.6.15)$$

Through the iterations of the appointed algorithm above, the gradient of the cost function is called and used before calling the projection function as follow

- a) Initialization: Let $\bar{y} \in \omega_\alpha$, and let $\varepsilon > 0$ be a positive tolerance
- b) Search direction computation: Let $d = P_{\omega_\alpha}(\bar{y} - \eta \nabla J_\alpha(\bar{y})) - \bar{y}$, where ∇J_α is the gradient of the cost function and η is a positive constant given by
 - First iteration $\eta = 1$
 - Other iterations: Let \bar{y} be the current point and \tilde{y} the previous point. Calculate $x = \bar{y} - \tilde{y}$ and $y = \nabla J_\alpha(\bar{y}) - \nabla J_\alpha(\tilde{y})$. Then, if $x^T y > 0$, take $\eta = \frac{x^T x}{x^T y}$; elsewhere, take η as a fixed positive value
- c) Termination: If $d = 0$ (in practice, $\|d\|_2 < \varepsilon$), then stop: \bar{y} is a stationary point of J_α on ω_α
- d) Step-size: Calculate a value $\xi \in (0, 1]$ such that $J_\alpha(\bar{y} + \xi d) \leq J_\alpha(\bar{y}) + \xi \delta \nabla J_\alpha(\bar{y})^T d$, with $\delta > 0$ (usually, $\delta \in [10^{-4}, 10^{-1}]$). For the determination of a suitable step-size ξ by an iterative way, we choose ζ as a positive constant (usually, $\zeta = 2$), take $\theta_1 = J_\alpha(\bar{y})$ and $\theta_2 = \delta \nabla J_\alpha(\bar{y})^T d$, and then for $p = 0, 1, 2, \dots$ define $\xi = \frac{1}{\zeta^p}$, and stop when $J_\alpha(\bar{y} + \xi d) \leq \theta_1 + \xi \theta_2$
- e) Update: Define $\tilde{y} = \bar{y} + \xi d$, and go to (a) with $\bar{y} = \tilde{y}$

The value of J_α is computed using the expression (3.6.10). Let $y = P_{\omega_\alpha}(z)$ is the projection of $z \in \mathbb{R}^4$ onto ω_α . $P_{\omega_\alpha}(z)$ is deduced by minimizing a quadratic function of the distance

of z to ω_α as follow

$$\min_{y \in \omega_\alpha} \frac{1}{2} \|y - z\|_2^2 = \min_{y \in \omega_\alpha} \frac{1}{2} z^T z - z^T y + \frac{1}{2} y^T y \quad (3.6.16)$$

which is equivalent to

$$\Leftrightarrow \left\{ \begin{array}{l} \min_{(y_1, y_2, y_3, y_4)} \sum_{i=1}^4 \left\{ \frac{1}{2} z_i^2 - z_i y_i + \frac{1}{2} y_i^2 \right\} \\ \text{subject to } a_i \leq y_i \leq b_i, i = 1, \dots, 4 \\ y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{l} \min_{(y_1, y_3)} \frac{1}{2} (z_1^2 + z_3^2) - (z_1 y_1 + z_3 y_3) + \frac{1}{2} (y_1^2 + y_3^2) \\ \text{subject to } a_i \leq y_i \leq b_i, i = 1, 3 \\ y_3 - y_1 \geq d_1, \end{array} \right. \text{and}$$

$$\left\{ \begin{array}{l} \min_{(y_2, y_4)} \frac{1}{2} (z_2^2 + z_4^2) - (z_2 y_2 + z_4 y_4) + \frac{1}{2} (y_2^2 + y_4^2) \\ \text{subject to } a_i \leq y_i \leq b_i, i = 2, 4 \\ y_2 - y_4 \geq d_2, \end{array} \right.$$

In order to solve the two quadratic problems, we use the Karush-Kuhn-Tucker (KKT) technique: Begin by the first quadratic problem and we choose $a = a_1 = a_3$, and $b = b_1 = b_3$, then from the five constraints, only three are kept. Finally the optimization problem is reformulated as

$$\min_{(x_1, x_2)} c - (l_1 x_1 + l_2 x_2) + \frac{1}{2} (x_1^2 + x_2^2)$$

$$\text{subject to } x_2 \leq b,$$

$$x_1 \geq a,$$

$$x_2 - x_1 \geq d_1,$$

where a, b, d_1, c , and $l_i, i=1,2$, are real numbers.

Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -b \\ a \\ d_1 \end{pmatrix},$$

Then the linear constraints write as

$$Ax - v = b, \quad v \geq 0,$$

where $v = (v_1, v_2, v_3)^T$ is the slack variables. The cost function is strictly convex and quadratic in \mathbb{R}^2 . Then the optimisation problem has an unique solution which satisfies the KKT conditions

$$\begin{aligned} v &= Ax - b \\ -l + x &= A^T w, \\ v &\geq 0, \quad w \geq 0, \\ v^T w &= 0, \end{aligned}$$

with $l = (l_1, l_2, l_3)$, and $w = (w_1, w_2, w_3)$ is the vector of the Lagrange multipliers associated to the three constraints above.

The unique optimal solution satisfies

$$x = l + A^T w$$

For computing the Lagrange multipliers w_i , we solve the following LCP problem

$$\begin{aligned} v &= (-b + Al) + AA^T w \\ v &\geq 0, \quad w \geq 0 \\ v^T w &= 0 \end{aligned}$$

3.6.4 Numerical results

3.6.4.1 Transcritical flow over a bump

To validate the numerical method based on the use of the multilayer system treated in this work, we introduce the test transcritical flow over a bump which is proposed by Castro et al. [13].

The domain is a channel of 21 m of length and 2 m of width with a bump of 5.75 m of length and it has a height of 0.2 m. For the boundary condition, we take a discharges of 2 m^3/s at the inflow boundary, and a water height of 0.6 m in the outflow boundary. The vertical viscosity is 0.001 m^2/s and the Strickler coefficient is 25.

Figure 4.3 represents the 3D mesh for the flow over a bump. The 3D horizontal velocity module for six layers is represented in figure 3.5. The used method may to capture the shock as shown in figure 3.5.

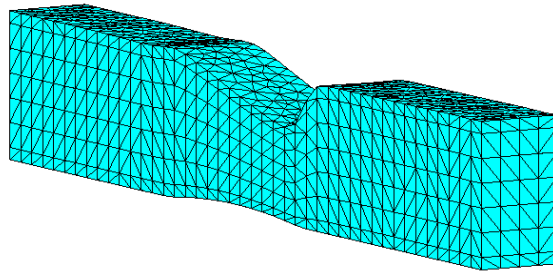


Figure 3.4: 3D mesh of the flow

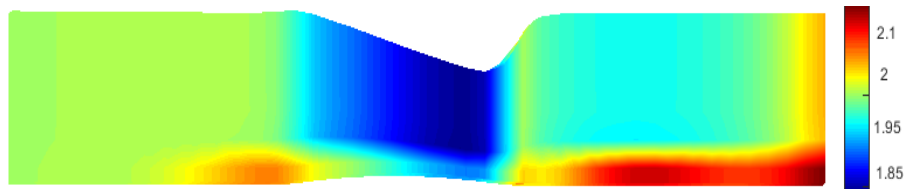


Figure 3.5: velocity field of 6 layers

3.6.4.2 Fishway Tests

The nature of the water flow in vertical slot fishways is influenced by the geometry of pools. We carry out three different tests of designs of vertical slot fishways with I or L shaped baffle. The fishway structures are composed of succession of pools and built with a slop relative to the ground

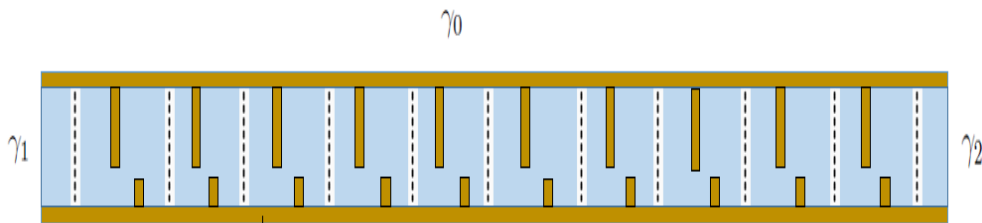


Figure 3.6: Ten pools fishway configuration used in the study

The numerical simulation for the three tests is done with the same physical and numerical conditions. All initial and boundary conditions are taken constants. $h_0 = h_\alpha(x, 0) = 0.083m$, $q_0 = q_\alpha(x, 0) = (0; 0)m^2s^{-1}$, $q_1 = -0.065/0.97 m^2s^{-1}$, and $h_2 = 0.5 m$. the target velocity is $c = 0.8 m.s^{-1}$. The vorticity parameter is $\alpha = 0$. The penalty parameter is $\beta = 500$. In these experiments, we consider that only the bottom friction stress is applied associated to Chezy coefficient of 57.36. For fish passage with comfortable conditions, we put $d_1 = 0.1$ and $d_2 = 0.05$ (see comfort constraints (3.5.13)).

3.6.4.2.1 Slot Fishway ("I" shaped baffle-rectangular slots)

Numerical simulations are focused to study the rectangular baffles as shown in figure 4.2 (see [6]). For initial random shape (figure 3.7 and 3.9), we observe circulation region near form baffles compared to the optimal shape the velocity is almost uniform and is close to the target velocity \mathbf{v} . The turbulence flow is removed for optimal points positions $a_{NM} = (0.4640; 0.1595)$ and $b_{NM} = (0.8920; 0.0378)$ for the Nelder Mead algorithm (NM) and $a_{SPG} = (0.4930; 0.1646)$ and $b_{SPG} = (0.7102; 0.0685)$ for the spectral projected gradient (SPG).

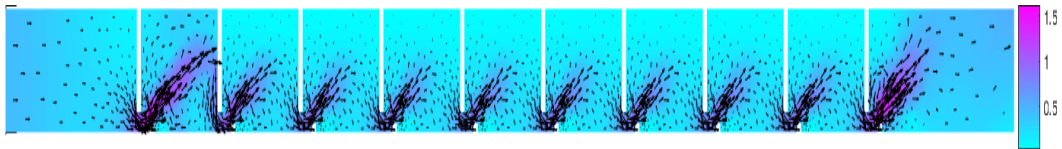


Figure 3.7: SPG: Non Optimal Shape and corresponding velocity in the ten pools at $T = 300$

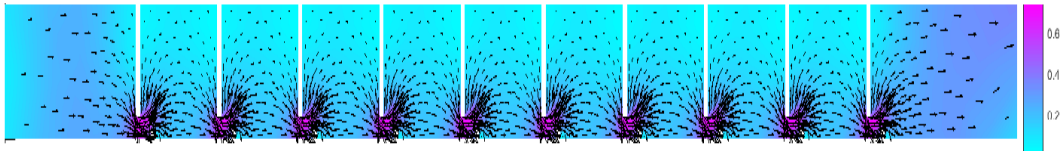


Figure 3.8: SPG: Optimal Shape and corresponding velocity in the ten pools at $T = 300$

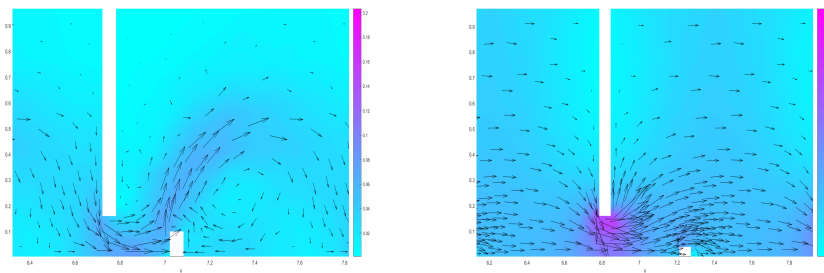
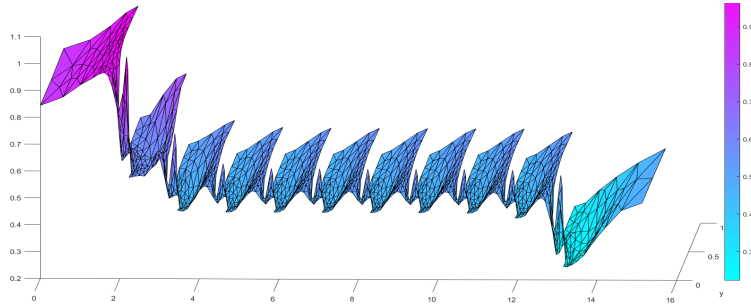
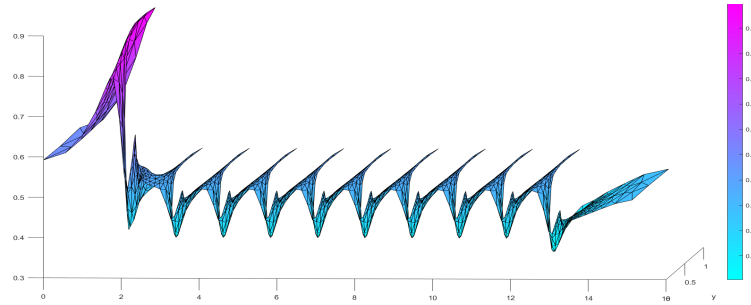


Figure 3.9: NM: Initial (left) and optimal (right) velocities for central pool at $T = 300$

Figure 3.10: NM: Initial water elevation at $T = 300$ Figure 3.11: NM: Optimal water elevation at $T = 300$

3.6.4.2.2 Vertical Slot Fishway ("I" shaped baffle-oblique slots)

We consider that the baffles under study are oblique (Figure 3.12), instead of rectangular baffles as done in the first experiment.

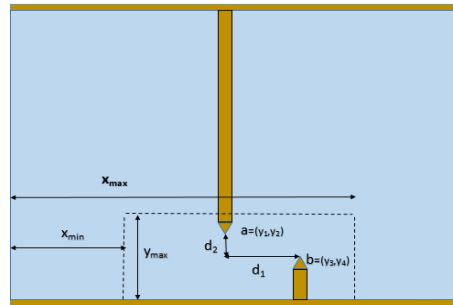


Figure 3.12: Scheme of the first pool

We observe that the circulation area are removed (Figure 3.13 and 3.14 right) compared to the velocity in the initial structure (Figure 3.13 and 3.14 left) for the optimal points $\mathbf{a}_{NM} = (0.4645, 0.1478)$ and $\mathbf{b}_{NM} = (0.7625, 0.0769)$ for Nelder Mead technique, and $\mathbf{a}_{SPG} = (0.4075, 0.1050)$ and $\mathbf{b}_{SPG} = (0.7990, 0.0350)$ for the spectral projected gradient.

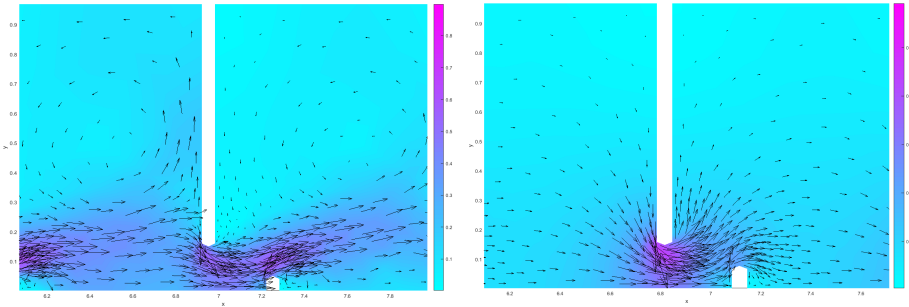


Figure 3.13: NM: Initial (left) and optimal (right) velocities for central pool at $T=300$

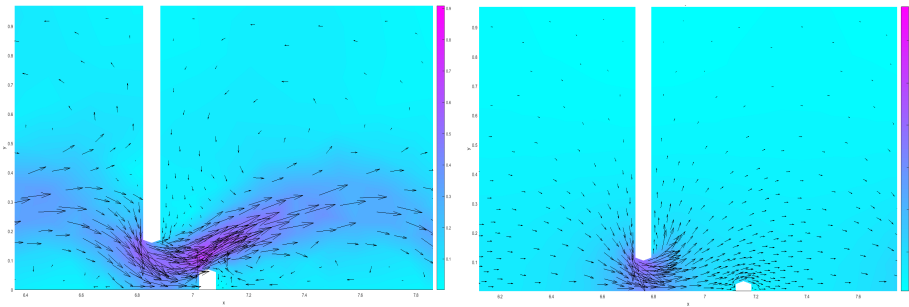


Figure 3.14: SPG: Initial (left) and optimal (right) velocities for central pool at $T=300$

3.6.4.2.3 Multi-Slot Fishway ("L" shaped baffle)

The third test consists to consider the same structure of fishway (Figure 4.2), but the two vertical baffles are replaced by three vertical baffles Figure 3.15.

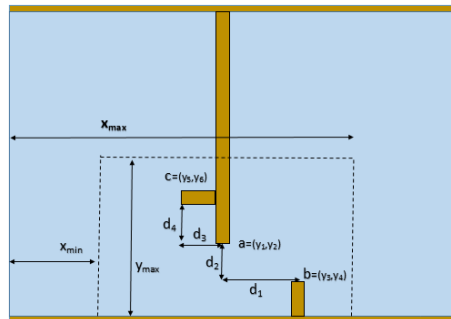


Figure 3.15: Scheme of the first pool

We consider the form of fishway depends now on the three points $a(y_1, y_2)$, $b(y_3, y_4)$ and $c(y_5, y_6)$. The comfort constraints are rewritten as

$$\begin{aligned} \frac{1}{4} 1.213 \leq y_1, y_3, y_5 \leq \frac{3}{4} 1.213 \\ 0 \leq y_2, y_4, y_6 \leq \frac{1}{2} 0.97 \end{aligned} \quad (3.6.17)$$

The stability constraints are written as

$$\begin{aligned} y_3 - y_1 &\geq d_1 = 0.1 \\ y_2 - y_4 &\geq d_2 = 0.05 \\ y_1 - y_5 &\geq d_3 = \frac{1}{2} 0.0305 \\ y_6 - y_2 &\geq d_4 = \frac{1}{2} 0.0305 \end{aligned} \quad (3.6.18)$$

Finally, in each layer α , the objective function conserves the same writing. We collect all sixteen constraints defined in (3.6.17) and (3.6.18) in a function

$$\begin{aligned} \Phi_\alpha^2 : \mathbb{R}^6 &\longrightarrow \mathbb{R}^{16} \\ \Phi_\alpha^2(y_1, y_2, y_3, y_4, y_5, y_6) &= \left(\frac{1}{4} 1.213 - y_1, \frac{1}{4} 1.213 - y_3, \frac{1}{4} 1.213 - y_5, y_1 - \frac{3}{4} 1.213, \right. \\ & y_3 - \frac{3}{4} 1.213, y_5 - \frac{3}{4} 1.213, -y_2, -y_4, -y_6, y_2 - \frac{1}{2} 0.97, \\ & y_4 - \frac{1}{2} 0.97, y_6 - \frac{1}{2} 0.97, 0.1 - y_3 + y_1, 0.05 - y_2 + y_4, \\ & \left. \frac{1}{2} 0.0305 - y_1 + y_5, \frac{1}{2} 0.0305 + y_2 - y_6 \right) \end{aligned} \quad (3.6.19)$$

The associated penalty function takes the form

$$\Phi_\alpha(y) = \bar{\Phi}_\alpha^1(y) + \beta \sum_{j=1}^{16} \max\{(\Phi_\alpha^2(y))_j, 0\} \quad (3.6.20)$$

For the spectral projected gradient, the matrix A and b are replaced by

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \quad b = \left(-\frac{3}{4} 1.213, \frac{1}{4} 1.213, \frac{1}{2} 0.0305, 0.1\right)^T$$

for the horizontal projection, and

$$A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \quad b = \left(-\frac{1}{2} 0.97, 0, \frac{1}{2} 0.0305, 0.05\right)^T$$

for the vertical projection.

For numerical simulations, the geometric parameters are $d_1 = 0.1$, $d_2 = 0.05$, $d_3 = \frac{1}{2}0.0305$, $d_4 = \frac{1}{2}0.0305$. The obtained optimal position points are $a_{NM} = (0.5160, 0.1663)$, $b_{NM} = (0.7964, 0.0437)$ and $c_{NM} = (0.3446, 0.2927)$ for Nelder mead algorithm, and $a_{SPG} = (0.4786; 0.1548)$, $b_{SPG} = (0.6660; 0.0892)$, and $c_{SPG} = (0.3448; 0.3735)$ for the spectral projected gradient.

We observe the disappearance of the circulation area in the optimal structure (Figure 3.17 and 3.18) compared to the initial random design (Figure 3.16 and 3.19).

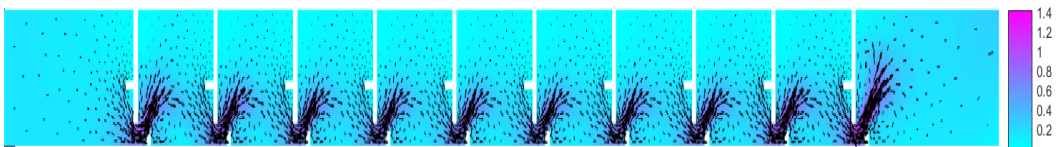


Figure 3.16: NM: Non Optimal Shape and corresponding velocity at T=300

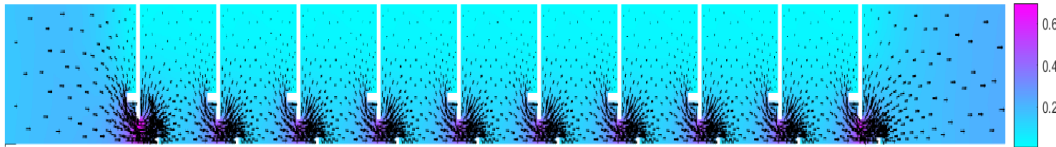


Figure 3.17: NM: Optimal Shape and corresponding velocity at T=300

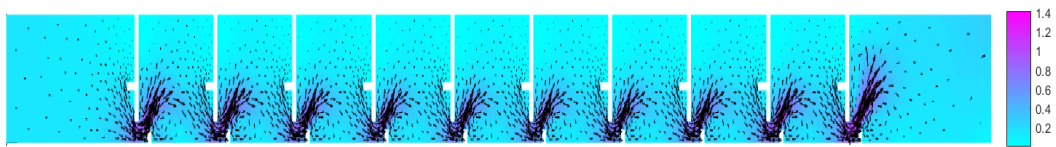


Figure 3.18: SPG: Non Optimal Shape and corresponding velocity at T=300

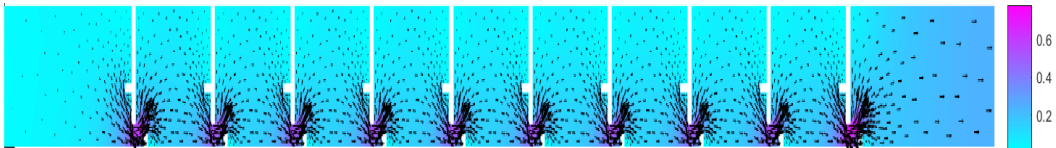


Figure 3.19: SPG: Optimal Shape and corresponding velocity at T=300

3.7 Conclusion

The multilayer model is adopted to avoid the expensive Navier-Stokes equations and is used in order to compute the free surface flow in vertical slot fishways. Design of effective fishways is becoming increasingly important. Numerical tests prove that the 2D multilayer Saint-Venant combined with an optimal shape design techniques is an alternative to 3D hydrostatic Navier-Stokes system. The numerical simulations evaluated the performance of various regular pool geometries. The study underlines the important influence of slot layout on flow characteristics. This study also assessed that the multi-slot design is suitable in terms of velocity and flow pattern.

Chapter 4

Analysis and finite element approximation of optimal control problem based on porous media model

4.1 Introduction

The shape optimisation problem consists to find an optimal shape in an admissible set. It is a branch of the control domain by the partial differential equations.

The purpose of this chapter is the study of different proprieties for the porous media equations based on the Navier Stokes equations with damping. We begin by showing the existence and uniqueness of the solution for the following boundary problem

$$\begin{aligned} -\nu\Delta u + (u.\nabla)u + a |u|^\alpha u + \nabla p &= f & \text{in } \Omega \\ \nabla.u &= 0 & \text{in } \Omega \\ u &= 0 & \text{on } \Gamma_0 \\ u &= g & \text{on } \Gamma_1 \\ \nu \frac{\partial u}{\partial \mathbf{n}} - p\mathbf{n} &= 0 & \text{on } \Gamma_2 \end{aligned} \tag{4.1.1}$$

where $\Omega \in \mathbb{R}^d$ ($d=2, 3$) is the fluid domain; Γ_0 is the lateral boundary, Γ_1 is inflow, top and bottom flow boundaries, and Γ_2 is the outflow boundary; u is the velocity; p is the pressure; ν is the viscosity; \mathbf{n} is the unit normal vector to the boundary $\partial\Omega$. $\alpha \in [0, 2]$ and $a > 0$ are real numbers. We introduce a perturbed problem related to the state system (4.1.1) as

$$\begin{aligned}
-\nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f && \text{in } \Omega \\
\nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 && \text{in } \Omega \\
u^\varepsilon &= 0 && \text{on } \Gamma_0 \\
u^\varepsilon &= g && \text{on } \Gamma_1 \\
\nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 && \text{on } \Gamma_2
\end{aligned} \tag{4.1.2}$$

Along this chapter, we restrict the treatment to the case $g = 0$ for the system (4.1.1). We study the existence and uniqueness of a weak solution by using a Faedo-Galerkin method. A penalized problem is introduced to derive an error priori estimates and show that the perturbed problem has a convergent solution as $O(\varepsilon)$ where ε is the penalty parameter. We study a penalized finite element method approximation to establish existence and uniqueness and to derive an error estimates for the approximated solution.

This kind of water flow is used to study a shape optimization problem of fishways. We look for finding an optimal shape of fishways by minimizing a cost function. Such optimal structure allow to a maximum number of fishes to cross to the river upstream.

The construction of dams is a very old activity; the first known works date back to 5,000 years and are located in the Middle East. It has continued to grow ever since. These structure are used usually for agriculture, electricity, or for navigation, But they constitute impassable obstacles to the free movement of fish and may contribute to the extinction of fish populations as Salomon, Trout, and Eels. In order to resolve achieve this challenge, men think to built the hydraulic structure which permit to migratory species to cross to the river upstream. Fishways represent a comfortable way for migratory species to across through dams. Our objective in this chapter is to simply the fishes passage to their production or feeding areas and assure to provide a rest area between pools. In our study we are interested by the vertical slots fishways (see [9]).

Vertical slots fishways are used frequently in the world. The vertical slots provide a good energy dissipation and create a quiet areas between pool which allows to fishes to have some rest before crossing one pool to another. Such structure is designed with a slope to the ground. It permits to small species to pass the circulation zones in the pool with less efforts which is not the case for other types of fishways; as the pool and weir type (Clay [1]) and the Denil type (Katopodis et al. [2]).

The chapter is organized as follows. The next section deals with the analysis of the

perturbed problem and the derivation of an error estimates to prove the convergence of the penalized solution. the section 3 is devoted to optimal control through which we will minimize a cost functional to obtain an optimal shape design of the fishways structures. A penalty finite element method approximation is introduced. Then we derive an error estimates and approximate the solution of the problem by discrete penalized solution. Finally in the last section, we adopt a gradient type algorithm with some numerical examples for 2D and 3D cases with structures already used for fishway modelling [6],[11].

4.2 Mathematical model

4.2.1 Preliminaries

We restrict the study of the system (4.1.1) to the case $g = 0$ along this section. The generalized case is treated in the introduction section and along study of the control problem based on the use of the penalized problem.

For the rest of the chapter, we denote by $\|\cdot\|$ the L^2 -norm and by $\|\cdot\|_k$ the norm of $H^k(\Omega)$. And we define the following Sobolev spaces

$$X = (H_0^1(\Omega))^d, \quad Y = L^2(\Omega), \quad \hat{Y} = \{q \in H^1(\Omega)\}, \quad M = \left\{ q \in L^2(\Omega); \int_{\Omega} q dx = 0 \right\}$$

$$\hat{V}_0(\Omega) = \{u \in (H^1(\Omega))^d | u = 0 \text{ on } \Gamma_0 \cup \Gamma_1\}$$

$$V_0(\Omega) = \{u \in (H^2(\Omega))^d | u = 0 \text{ on } \Gamma_0 \cup \Gamma_1\}$$

$$\hat{V}_g(\Omega) = \{u \in (H^1(\Omega))^d | u = 0 \text{ on } \Gamma_0, u = g \text{ on } \Gamma_1\}$$

$$V_g(\Omega) = \{u \in (H^2(\Omega))^d | u = 0 \text{ on } \Gamma_0, u = g \text{ on } \Gamma_1\}$$

$$\mathbb{X} = \{v \in (H_0^1(\Omega))^d | (q, \nabla \cdot v) = 0, \forall q \in L^2(\Omega)\}$$

We give the property of monotonicity for any mapping F given by $F : \mathbf{x} \mapsto |\mathbf{x}|^{\alpha} \mathbf{x}$

$$(|\mathbf{u}|^{\alpha} \mathbf{u} - |\mathbf{v}|^{\alpha} \mathbf{v}, \mathbf{u} - \mathbf{v}) \geq 0 \quad (4.2.1)$$

We also denote by $C_{0,div}^{\infty}$ the set of all C^{∞} real vector functions $u = (u_1, \dots, u_d)$ ($d=2, 3$) with compact support in Ω such that $\nabla \cdot u = 0$.

We begin by showing here the existence and uniqueness results which will be used in the next sections. We prove that the velocity is bounded in term of the L^2 norm. For the existence of the solution, we use a Faedo-Galerkin method.

4.2.1.1 Existence of weak solution

In this subsection, we study the system

$$\begin{aligned} -\nu\Delta u + (u.\nabla)u + a|u|^\alpha u + \nabla p &= f && \text{in } \Omega \\ \nabla.u &= 0 && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega \end{aligned} \tag{4.2.2}$$

The weak formulation of the system (4.2.2) takes the form

Find $(u, p) \in C_{0,div}^\infty(\Omega) \times Y$ such that

$$\begin{aligned} \nu(\nabla u, \nabla v) + ((u.\nabla)u, v) + a(|u|^\alpha u, v) - (p, \nabla.v) &= (f, v) \quad \forall v \in C_{0,div}^\infty(\Omega) \\ (\nabla.u, q) &= 0 \quad \forall q \in M \end{aligned} \tag{4.2.3}$$

The next result proves the existence of the velocity and show that it is bounded in norm L^2 .

Theorem 4.2.1. *Let Ω be a bounded domain in \mathbb{R}^d ($d = 2, 3$), $f \in X'$ is a given exterior function. There exist at least a weak solution of (4.2.2) satisfying*

$$\frac{\nu}{2} \|\nabla u\|^2 + a \|u\|_{L^{\alpha+2}}^{\alpha+2} \leq \frac{1}{2\nu} \|f\|_{X'}^2, \quad \text{for } 0 \leq \alpha \leq 10 \tag{4.2.4}$$

and

$$\|p\| \leq C(\|f\|_{X'} + \|\nabla u\| + \|\nabla u\|^2 + \|u\|_{L^{\alpha+2}}) \tag{4.2.5}$$

for some positive constant C .

For the proof, we use the coming lemmas

Lemma 4.2.1. *Suppose that Ω is a bounded locally Lipschitzian domain of \mathbb{R}^d ($d=2, 3$), $\Omega_0 \subseteq \Omega$, $\Omega_0 \neq \emptyset$. Let $F \in X'$ and $u \in X$ such that*

$$(F, u) = 0$$

then there exists a unique $P \in L^2(\Omega)$ satisfying

$$(F, u) = \int_{\Omega} P \nabla.u dx \quad \text{and} \quad \int_{\Omega} P dx = 0$$

Lemma 4.2.2. *Assume that Ω is an arbitrary domain in \mathbb{R}^d ($d=2, 3$), and let $f \in X'$. Then*

1. *The vector $V \in H_{loc}^1(\Omega)$ satisfies (4.2.3) for all $v \in C_{0,div}^\infty(\Omega)$ if and only if there exists $P \in L_{loc}^2(\Omega)$ satisfying the identity*

$$\nu(\nabla V, \nabla w) + ((V \cdot \nabla)V, w) + (a|V|^\alpha V, w) = (P, \nabla \cdot w) + (f, w) \quad (4.2.6)$$

for all $w \in C_{0,div}^\infty(\Omega)$.

2. *If, Moreover, Ω is bounded and Lipschitzian and $f \in X'$, $V \in (H^1(\Omega))^d$ ($d=2, 3$), then*

$$P \in Y, \text{ with } \int_{\Omega} P dx = 0$$

and (4.2.6) holds for all $w \in X$.

Proof. It is clear that (4.2.6) implies (4.2.3). If Ω is locally Lipschitzian, the functional

$$F(w) := \nu(\nabla V, \nabla w) + ((V \cdot \nabla)V, w) + (a|V|^\alpha V, w) - (f, w) \quad (4.2.7)$$

is linear and bounded in $w \in X$ and it becomes zero when $w \in X$ and $\nabla \cdot w = 0$. Using the Corollary 3.5.1 in [22] and lemma 4.2.1, there exists $P \in Y$ such that

$$F(w) = (P, \nabla \cdot w) \quad (4.2.8)$$

for all $w \in X$, thus satisfying (4.2.6). If Ω is an arbitrary domain, we use Corollary 3.5.2 in [22] to deduce the existence of $P \in L_{loc}^2(\Omega)$ satisfying (4.2.8) $\forall w \in C_{0,div}^\infty(\Omega)$. \square

Lemma 4.2.3. *Consider $F: \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a continuous function such that for some $K > 0$*

$$F(\xi) \cdot \xi > 0$$

for all $\xi \in \mathbb{R}^m$ with $|\xi| = K$. Then there exists $\xi_0 \in \mathbb{R}^m$ with $|\xi_0| \leq K$ such that $F(\xi_0) = 0$.

Proof. (of theorem (4.2.1)) Existence of the velocity field: We consider an orthonormal basis (w_k) in X and define the sequence $(u_m) \in X$ satisfying

$$\nu(\nabla u_m, \nabla w_k) + ((u_m \cdot \nabla)u_m, w_k) + (a|u_m|^\alpha u_m, w_k) = (f, w_k) \quad (4.2.9)$$

where for each $m \in \mathbb{N}$

$$u_m = \sum_{k=1}^m \xi_{km} w_k \quad (4.2.10)$$

We multiply (4.2.9) by ξ_{km} and sum with respect to k to obtain

$$\nu \|\nabla u_m\|^2 + a \|u_m\|_{L^{\alpha+2}}^{\alpha+2} = (f, u_m) \leq \|f\|_{X'} \|\nabla u_m\| \quad (4.2.11)$$

By Hölder inequality we get

$$\frac{\nu}{2} \|\nabla u_m\|^2 + a \|u_m\|_{L^{\alpha+2}}^{\alpha+2} \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.2.12)$$

Next, we introduce the following function

$$(F(\xi), \xi) := \nu(\nabla\xi, \nabla\xi) + a(|\xi|^\alpha \xi, \xi) - (f, \xi)$$

We have

$$\begin{aligned} (F(\xi), \xi) &\geq \nu \|\xi\|^2 + a \|\xi\|_{L^{\alpha+2}}^{\alpha+2} - \|f\|_{X'} \|\xi\| \\ &= (\nu \|\xi\| - \|f\|_{X'}) \|\xi\| + a \|\xi\|_{L^{\alpha+2}}^{\alpha+2} \end{aligned}$$

when $\|\xi\| = K$, $K > \frac{1}{\nu} \|f\|_{X'}$. The lemma 4.2.3 gives the existence of u_m solution of (4.2.9) for each m . Moreover, the sequence (u_m) is bounded by (4.2.12), then there exists a subsequence denoted (u_m) which

$$u_m \rightharpoonup u \text{ in } X \cap L^{\alpha+2}(\Omega) \quad (4.2.13)$$

From the Sobolev embedding theorem we get

$$u_m \longrightarrow u \text{ in } L^{6-\varepsilon} \quad (4.2.14)$$

with ε is some positive constant. For the convergence of three quantities in (4.2.9), we proceed as follow; as $m \rightarrow \infty$ we have

$$(\nabla u_m, \nabla w_k) \longrightarrow (\nabla u, \nabla w_k) \quad (4.2.15)$$

For the two remaining terms, we develop

$$\begin{aligned} &|((u_m \cdot \nabla)u_m, w_k) - ((u \cdot \nabla)u, w_k)| \\ &\leq |((u_m - u) \cdot \nabla)u_m, w_k| + |((u \cdot \nabla)(u_m - u), w_k)| \\ &= T^{(1)} + T^{(2)} \end{aligned}$$

The first quantity can be raised by

$$\begin{aligned} T^{(1)} &\leq \|w_k\|_{L^6} \|\nabla u_m\| \|u_m - u\|_{L^3} \\ &\leq c \|w_k\|_1 \|\nabla u_m\| \|u_m - u\|_{L^3} \end{aligned}$$

By (4.2.14) we obtain

$$\lim_{m \rightarrow \infty} T^{(1)} = 0$$

It remains to raise the term $T^{(2)}$

$$\begin{aligned} T^{(2)} &\leq |((u \cdot \nabla)w_k, (u_m - u))| \\ &\leq \|u\|_{L^6} \|\nabla w_k\| \|u_m - u\|_{L^3} \\ &\leq C \|w_k\| \|u_m - u\|_{L^3} \end{aligned}$$

Therefore,

$$\lim_{m \rightarrow \infty} T^{(2)} = 0$$

Finally, we obtain the ensuing convergence result

$$|((u_m \cdot \nabla)u_m, w_k) - ((u \cdot \nabla)u, w_k)| \rightarrow 0, \quad m \rightarrow \infty \quad (4.2.16)$$

For the last term, we use

$$\begin{aligned} |(|u_m|^\alpha u_m, w_k) - (|u|^\alpha u, w_k)| &\leq |(|u_m|^\alpha (u_m - u), w_k)| + |((|u_m|^\alpha - |u|^\alpha)u, w_k)| \\ &= T^{(3)} + T^{(4)} \end{aligned}$$

For the first term $T^{(3)}$, thank to the embedding theorem

$$T^{(3)} \leq C \sup |w_k| \cdot \|u_m\|_{L^{\alpha+2}}^\alpha \|u_m - u\|_{L^{6-\varepsilon}}$$

and according to $\alpha = 10 - 2\varepsilon$ we obtain

$$\lim_{m \rightarrow \infty} T^{(3)} = 0$$

The last term $T^{(4)}$ can be raised with the same approach

$$T^{(4)} \leq C \|u_m - u\|_{L^{6-\varepsilon}} \cdot \|u_m\|_{L^{\alpha+2}}^\alpha \|w_k\|_{L^p}$$

For $p = \infty$ and $\alpha = 10 - 2\varepsilon$, we have

$$\lim_{m \rightarrow \infty} T^{(4)} = 0$$

Therefore

$$\lim_{m \rightarrow \infty} (|u_m|^\alpha u_m, w_k) = (|u|^\alpha u, w_k) \quad (4.2.17)$$

From (4.2.16)-(4.2.17), the velocity $u \in X$ satisfies the equation

$$\nu(\nabla u, \nabla w_k) + ((u \cdot \nabla)u, w_k) + (a |u|^\alpha u, w_k) = (f, w_k) \quad (4.2.18)$$

Finally, we write any $\varphi \in X$ as a limit of a linear combination of w_k . Then the velocity $u \in X$ satisfies

$$\nu(\nabla u, \nabla \varphi) + ((u \cdot \nabla)u, \varphi) + (a |u|^\alpha u, \varphi) = (f, \varphi) \quad \text{for all } \varphi \in X \quad (4.2.19)$$

Using (4.2.12) one gets

$$\frac{\nu}{2} \|\nabla u\|^2 + a \|u\|_{L^{\alpha+2}}^{\alpha+2} \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.2.20)$$

Existence of the pressure field: Supposing that Ω is locally Lipschitzian, thanks to lemma (4.2.2): there exists $p \in Y$ satisfying (4.2.3) and

$$\int_{\Omega} p dx = 0 \quad (4.2.21)$$

Consider the problem

$$\begin{aligned} \nabla \cdot W &= p \\ W &\in X \end{aligned} \quad (4.2.22)$$

$$\|W\|_1 \leq \|p\|$$

As $p \in Y$ and satisfies (4.2.21), the problem (4.2.22) has a solution thanks to the theorem 3.3.1 in [22]. Setting $w = W$ in (4.2.6) to obtain

$$\|p\|^2 \leq C (\|f\|_{X'} + \|u\|_1 + \|u\|_{L^{\alpha+2}} + \|u\|_1^2) \|p\| \quad (4.2.23)$$

□

Lemma 4.2.4. *Consider $u, v \in X$ such that (4.2.20) holds for u and v , then*

$$\| |u|^\alpha u - |v|^\alpha v \| \leq C \| \nabla(u - v) \| \quad (4.2.24)$$

After the study of the existence of the solution for the boundary problem, we proceed to treat the uniqueness of the solution.

4.2.1.2 Uniqueness of weak solution

This theorem states the uniqueness of the weak solution for the boundary problem (4.2.2).

Theorem 4.2.2. *Let Ω be a bounded domain in \mathbb{R}^d ($d=2, 3$), locally Lipschitz, and $f \in X'$. If ν is "big enough" and $\frac{2}{3} \leq \alpha \leq 2$, then the weak solution of (4.2.2) is unique.*

Proof. Let us suppose that the problem (4.2.2) has two weak solutions u_1 and u_2 .

Define $\varphi = u_1 - u_2$ and write

$$\nu \| \nabla(u_1 - u_2) \|^2 + (((u_1 - u_2) \cdot \nabla) u_1, u_1 - u_2) + (a|u_1|^\alpha u_1 - a|u_2|^\alpha u_2, u_1 - u_2) = 0$$

which leads to

$$\begin{aligned} &\nu \| \nabla(u_1 - u_2) \|^2 + (((u_1 - u_2) \cdot \nabla) u_1, u_1 - u_2) \\ &+ (a|u_1|^\alpha (u_1 - u_2), (u_1 - u_2)) + a(|u_1|^\alpha - |u_2|^\alpha) u_2, u_1 - u_2) = 0 \end{aligned}$$

This gives

$$\begin{aligned} \nu \|\nabla(u_1 - u_2)\|^2 &= -(((u_1 - u_2) \cdot \nabla)u_1, u_1 - u_2) \\ &\quad - (a|u_1|^\alpha(u_1 - u_2), (u_1 - u_2)) \\ &\quad - a(|u_1|^\alpha - |u_2|^\alpha)u_2, u_1 - u_2 \end{aligned} \quad (4.2.25)$$

The first term on the right hand side in (4.2.25) implies

$$\begin{aligned} | -(((u_1 - u_2) \cdot \nabla)u_1, u_1 - u_2) | &\leq \|\nabla u_1\| \cdot \|u_1 - u_2\|_{L^4}^2 \\ &\leq \frac{c_1}{\nu} \|f\|_{X'} \|u_1 - u_2\|_{L^6}^2 \end{aligned} \quad (4.2.26)$$

Using the embedding of $L^6 \hookrightarrow H^1$

$$| -(((u_1 - u_2) \cdot \nabla)u_1, u_1 - u_2) | \leq \frac{c_1}{\nu} \|f\|_{X'} \|u_1 - u_2\|_1^2 \quad (4.2.27)$$

We now turn to the other terms using $1 \leq \frac{3\alpha}{2} \leq \alpha + 2$, i.e., $\frac{2}{3} \leq \alpha \leq 4$, we obtain

$$\begin{aligned} | - (a|u_1|^\alpha(u_1 - u_2), (u_1 - u_2)) | &= |a|u_1|^\alpha|u_1 - u_2|^2| \\ &\leq a \| |u_1|^\alpha \|_{\frac{3}{2}} \| |u_1 - u_2|^2 \|_{L^3} \\ &\leq a \| |u_1| \|_{\frac{3\alpha}{2}} \|u_1 - u_2\|_{L^6}^2 \\ &\leq ca \|u_1\|_{L^{\alpha+2}}^\alpha \|\nabla(u_1 - u_2)\|^2 \\ &\leq c_2 a \|f\|_{X'} \|\nabla(u_1 - u_2)\|^2 \end{aligned} \quad (4.2.28)$$

The last term on the right of (4.2.25), and according to $1 \leq \frac{3\alpha}{2} \leq \alpha + 2$, i.e., $\frac{2}{3} \leq \alpha \leq 4$ we obtain

$$| - a(|u_1|^\alpha - |u_2|^\alpha)u_2, (u_1 - u_2) | \leq c_3 a \alpha \|f\|_{X'} \|\nabla(u_1 - u_2)\|^2 \quad (4.2.29)$$

Replacing (4.2.26), (4.2.28) and (4.2.29) in (4.2.25) yields

$$\nu \|\nabla(u_1 - u_2)\|^2 - \left(\frac{c_1}{\nu} \|f\|_{X'} + c_2 a \|f\|_{X'} + c_3 a \alpha \|f\|_{X'} \right) \|\nabla(u_1 - u_2)\|^2 \leq 0$$

We conclude, for ν big enough, that $u_1 = u_2$. \square

The established inequalities prove that the velocity u is bounded. In the next subsection we treat the penalized problem related to the Navier Stokes equations with damping (4.2.2). We use a penalty method to overcome the difficulties associated with the constraint " $\nabla \cdot u = 0$ ".

Among the method adopted to relax the constraint $\nabla \cdot u = 0$ we cite

1. The penalty method. We refer here to J. Shen [19], R. Temam [36]

$$\nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon = 0 \quad \text{in } \Omega \times [0, T]$$

2. Artificial compressibility, we mention here the works of Chorin [20] and of R. Temam [37]

$$\nabla \cdot u^\varepsilon + \varepsilon p_t^\varepsilon = 0 \quad \text{in } \Omega \times [0, T]$$

3. The projection method

$$\nabla \cdot u^\varepsilon - \varepsilon \Delta p^\varepsilon = 0 \quad \text{in } \Omega \times [0, T], \quad \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \quad \text{in } \partial\Omega \times [0, T]$$

introduced by Temam [36] and Chorin [20, 21] and has a high computational cost.

4. The pseudocompressibility method

$$\nabla \cdot u^\varepsilon - \varepsilon \Delta p_t^\varepsilon = 0 \quad \text{in } \Omega \times [0, T], \quad \frac{\partial p^\varepsilon}{\partial \mathbf{n}} = 0 \quad \text{in } \partial\Omega \times [0, T]$$

4.2.2 Perturbed equation

Let us now give the penalized formulation associated to the Navier Stokes equations with damping

$$\begin{aligned} \nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f \quad \text{in } \Omega \\ \varepsilon p^\varepsilon + \nabla \cdot u^\varepsilon &= 0 \quad \text{in } \Omega \\ u^\varepsilon &= 0 \quad \text{on } \partial\Omega \end{aligned} \quad (4.2.30)$$

The penalized weak formulation of the equation (4.2.30) writes for $\varepsilon > 0$

Find $(u^\varepsilon, p^\varepsilon) \in X \times Y$ such that

$$\begin{aligned} \nu (\nabla u^\varepsilon, \nabla v) + ((u^\varepsilon \cdot \nabla) u^\varepsilon, v) + \frac{1}{2} ((\nabla \cdot u^\varepsilon), u^\varepsilon \cdot v) + a (|u^\varepsilon|^\alpha u^\varepsilon, v) + \frac{1}{\varepsilon} (\nabla \cdot u^\varepsilon, \nabla \cdot v) &= (f, v) \\ \varepsilon (p^\varepsilon, q) + (\nabla \cdot u^\varepsilon, q) &= 0 \end{aligned} \quad (4.2.31)$$

for all $v \in X$ and $q \in Y$.

Here are some elementary proprieties of the functional defined by $T(w, u, v) = ((w \cdot \nabla) u, v) + \frac{1}{2} ((\nabla \cdot w), u \cdot v)$ for $(w, u, v) \in X^3$. T is trilinear and there exists a positive constant $C > 0$ such that

$$T(w, u, v) \leq C \|w\|_1 \|u\|_1 \|v\|_1 \quad \forall (w, u, v) \in X^3 \quad (4.2.32)$$

Moreover, T satisfies

$$((w \cdot \nabla)v, v) + \frac{1}{2}((\nabla \cdot w), v \cdot v) = 0 \text{ for all } w, v \in X \quad (4.2.33)$$

and

$$((w \cdot \nabla)u, v) + \frac{1}{2}((\nabla \cdot w), u \cdot v) = -((w \cdot \nabla)v, u) - \frac{1}{2}((\nabla \cdot w), v \cdot u) \text{ for all } (w, u, v) \in X^3 \quad (4.2.34)$$

We begin with an existence result.

4.2.2.1 Existence of weak solution of the perturbed problem

The following theorem asserts that the solution for the perturbed problem (4.2.30) exists and it is bounded.

Theorem 4.2.3. *Let Ω a bounded domain in \mathbb{R}^d ($d = 2, 3$), $f \in X'$ is a given exterior function. Then the perturbed problem (4.2.30) has a weak solution satisfying*

$$\frac{\nu}{2} \|\nabla u^\varepsilon\|^2 + a \|u^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p^\varepsilon\|^2 \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.2.35)$$

Proof. We consider (w_k) an orthonormal basis of X and (r_j) an orthonormal basis of Y . Let $X_m = \langle w_1, \dots, w_m \rangle$ and $Y_m = \langle r_1, \dots, r_m \rangle$ for some positive integer m . We define the sequence (u_m^ε) and (p_m^ε)

$$u_m^\varepsilon = \sum_{k=1}^m \xi_{km} w_k, \quad p_m^\varepsilon = \sum_{k=1}^m \psi_{km} r_j \quad (4.2.36)$$

the approximate solution of u^ε and p^ε such that

$$\begin{aligned} \nu(\nabla u_m^\varepsilon, \nabla w_k) + (u_m^\varepsilon \cdot \nabla)u_m^\varepsilon, w_k) + \frac{1}{2}(\nabla \cdot u_m^\varepsilon, u_m^\varepsilon \cdot w_k) + a(|u_m^\varepsilon|^\alpha u_m^\varepsilon, w_k) \\ -(p_m^\varepsilon, \nabla \cdot w_k) = (f, w_k) \quad \text{for } k = 1, \dots, m \end{aligned} \quad (4.2.37)$$

$$\varepsilon(p_m^\varepsilon, r_j) + (\nabla \cdot u_m^\varepsilon, r_j) = 0, \quad \text{for } j = 1, \dots, m \quad (4.2.38)$$

We multiply by ξ_{km} the equation (4.2.37) and by ψ_{jm} the equation (4.2.37) and sum with respect to k and j to get

$$\nu \|\nabla u_m^\varepsilon\|^2 + a \|u_m^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p_m^\varepsilon\|^2 = (f, u_m^\varepsilon) \leq \|f\|_{X'} \|u_m^\varepsilon\| \quad (4.2.39)$$

Using the Hölder inequality to have

$$\frac{\nu}{2} \|\nabla u_m^\varepsilon\|^2 + a \|u_m^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p_m^\varepsilon\|^2 \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.2.40)$$

For the convergence of the all terms on the left hand side in (4.2.37), we follow a similar process as in the previous section. u^ε satisfies

$$\begin{aligned} \nu(\nabla u^\varepsilon, \nabla w_k) + (u^\varepsilon \cdot \nabla) u^\varepsilon, w_k) + \frac{1}{2}(\nabla \cdot u^\varepsilon, u^\varepsilon \cdot w_k) + a(|u^\varepsilon|^\alpha u^\varepsilon, w_k) \\ - (p^\varepsilon, \nabla \cdot w_k) = (f, w_k) \quad \text{for } k = 1, \dots, m \end{aligned} \quad (4.2.41)$$

For all $\varphi \in H^1(\Omega)$ written as a linear combination of w_k , u^ε and p^ε satisfy

$$\begin{aligned} \nu(\nabla u^\varepsilon, \nabla \varphi) + (u^\varepsilon \cdot \nabla) u^\varepsilon, \varphi) + \frac{1}{2}(\nabla \cdot u^\varepsilon, u^\varepsilon \cdot \varphi) + a(|u^\varepsilon|^\alpha u^\varepsilon, \varphi) \\ - (p^\varepsilon, \nabla \cdot \varphi) = (f, \varphi) \quad \forall \varphi \in H^1(\Omega) \end{aligned} \quad (4.2.42)$$

Using (4.2.40) to infer that

$$\frac{\nu}{2} \|\nabla u^\varepsilon\|^2 + a \|u^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p^\varepsilon\|^2 \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.2.43)$$

□

In the following section, we firstly use the fact that u^ε is bounded to show that p^ε is bounded. Next, We extract a subsequence convergent to the boundary problem solution. Finally, we establish the error estimates.

4.2.3 Error estimates

We consider that $(u^\varepsilon, p^\varepsilon)$ are solutions of the problem (4.2.30) and let s be the norm

$$s = \sup_{u, v, w \in X} \frac{|((w \cdot \nabla) u \cdot v) + \frac{1}{2}(\nabla \cdot w, u \cdot v)|}{\|w\|_1 \|u\|_1 \|v\|_1} \quad (4.2.44)$$

Then, the data ν and f are assumed to satisfy

$$\frac{s}{\nu^2} \|f\|_{X'} < 1 \quad (4.2.45)$$

The equation (4.2.45) can be written differently; there is a fixed r with $0 < r < 1$ such that

$$\frac{s}{\nu^2} \|f\|_{X'} \leq 1 - r \quad (4.2.46)$$

This lemma gives the convergence of non linear terms which will be used in this section.

Lemma 4.2.5. *If $u_k \rightarrow u \in X$. Then for all $v \in X$*

$$\begin{aligned} \lim_{k \rightarrow \infty} ((u_k \cdot \nabla) u_k, v) &= ((u \cdot \nabla) u, v) \\ \lim_{k \rightarrow \infty} (\nabla \cdot u_k, u_k \cdot v) &= (\nabla \cdot u, u \cdot v) \\ \lim_{k \rightarrow \infty} (\nabla \cdot u_k, \nabla \cdot v) &= (\nabla \cdot u, \nabla \cdot v) \\ \lim_{k \rightarrow \infty} (|u_k|^\alpha u_k, v) &= (|u|^\alpha u, v) \end{aligned} \quad (4.2.47)$$

Proof. For the last term, we use Lemma 4.2.4

$$\| (|u_k|^\alpha u_k - |u|^\alpha u) \| \leq \| \nabla(u_k - u) \|$$

□

We introduce the inf-sup condition: there exists a constant $\beta > 0$ such that

$$\sup_{v \in X(v \neq 0)} \frac{(q, \nabla \cdot v)}{\|v\|_1} \geq \beta \| \hat{q} \|_Z \quad \text{for all } q \in Y \quad (4.2.48)$$

where

$$\hat{q} = \{ \bar{q} \in Y | \bar{q} - q = \text{constant} \}, \quad Z = \{ \hat{q} | q \in Y \} \quad \text{and} \quad \| \hat{q} \|_Z = \inf_{x \in \mathbb{R}} \| q + x \| \quad (4.2.49)$$

We also define the operators $B \in \mathcal{L}(X, Y')$ and $B^* \in \mathcal{L}(Y, X')$ by

$$(q, \nabla \cdot v) = (q, Bv)_{Y \times Y'} = (B^*q, v)_{X' \times X} \quad (4.2.50)$$

We have

$$Bv = \nabla \cdot v \quad \text{and} \quad B^*q = \nabla q \quad (4.2.51)$$

and

$$\text{Ker } B^* = \{ q \in Y | (q, \nabla \cdot v) = 0 \text{ for all } v \in X \} = \{ \text{constants} \}$$

The ensuing theorem shows that the perturbed pressure is bounded. The convergence of the couple $(u^\varepsilon, p^\varepsilon)$ is asserted by the extraction of a subsequence convergent to the solution for the boundary problem (4.2.2). We also give a priori estimate for the velocity and the pressure.

Theorem 4.2.4. *Let $(u^\varepsilon, p^\varepsilon)$ be solution of the problem (4.2.30) and (u, p) the solution in $X \times Y/\mathbb{R}$ of the problem (4.2.2). Then we have*

$$(p^\varepsilon) \in Y \quad \text{is uniformly bounded in } \varepsilon \quad (4.2.52)$$

and as $\varepsilon \rightarrow 0$.

$$u^\varepsilon \rightarrow u \quad \text{in } X \quad (4.2.53)$$

and

$$p^\varepsilon \rightarrow \tilde{p} \quad \text{in } Y \quad (4.2.54)$$

where \tilde{p} is the orthogonal projection in Y of p onto $(\text{Ker } B^*)^\perp$. And we have the following estimates

$$\| u^\varepsilon - u \|_1 + \| p^\varepsilon - \tilde{p} \| \leq C\varepsilon \quad (4.2.55)$$

with C is a positive constant non depending of ε .

Proof. We use the second equation in (4.2.30) into the first equation in (4.2.31) we obtain

$$-(p^\varepsilon, \nabla \cdot v) = (f, v) - \nu(\nabla u^\varepsilon, \nabla v) - ((u^\varepsilon \cdot \nabla)u^\varepsilon, v) - \frac{1}{2}(\nabla u^\varepsilon, u^\varepsilon \cdot v) - (a|u^\varepsilon|^\alpha u^\varepsilon, v) \text{ for all } v \in X$$

Thanks to (4.2.32) and Lemma 4.2.4

$$|(p^\varepsilon, \nabla \cdot v)| \leq (\|f\|_{X'} + \nu \|u^\varepsilon\|_1 + C_1 \|u^\varepsilon\|_1^2 + C_2 a \|u^\varepsilon\|_1) \|v\|_1 \text{ for all } v \in X$$

Applying this in the inf-sup condition with $q = p^\varepsilon \in Y$, and using the fact that ∇u^ε is uniformly bounded, we get

$$\|p^\varepsilon\|_Z \leq \frac{C}{\beta} \quad (4.2.56)$$

with C is a constant no depending of ε . Using the second equation of (4.2.30) and the definition of $\ker B^*$

$$(p^\varepsilon, q) = 0 \text{ for all } q \in \ker B^*$$

Then

$$p^\varepsilon \in (\ker B^*)^\perp \text{ in } Y$$

Since

$$\|p^\varepsilon\|_Z = \|p^\varepsilon\|$$

Therefore

$$\|p^\varepsilon\| \leq \frac{C}{\beta} \quad (4.2.57)$$

(p^ε) is uniformly bounded in ε .

The couple $(u^\varepsilon, p^\varepsilon)$ is uniformly bounded in $X \times Y$. Then there exists a subsequence without lose of generality denoted $(u^\varepsilon, p^\varepsilon)$ convergent to $(u, \tilde{p}) \in X \times Y$. We know already that $(\ker B^*)^\perp$ is closed and convex that is gives $(\ker B^*)^\perp$ is weakly closed and $p^\varepsilon \in (\ker B^*)^\perp \rightarrow \tilde{p} \in (\ker B^*)^\perp$.

Let us now show that (u, \tilde{p}) is solution of the problem (4.2.2). We use the second equation of (4.2.30), we have

$$(q, \nabla \cdot u^\varepsilon) = -\varepsilon \left(q, -\frac{1}{\varepsilon} \nabla \cdot u^\varepsilon \right) = -\varepsilon (q, p^\varepsilon)$$

Therefore

$$|(q, \nabla \cdot u^\varepsilon)| \leq \varepsilon \|q\| \|p^\varepsilon\| \quad (4.2.58)$$

And by lemma (4.2.5), $\nabla \cdot u^\varepsilon \rightarrow \nabla \cdot u$ in $L^2(\Omega)$. we get

$$(q, \nabla \cdot u^\varepsilon) \rightarrow (q, \nabla \cdot u) \text{ as } \varepsilon \rightarrow 0$$

Passing to limits in (4.2.58) and using (4.2.57) yields

$$(q, \nabla \cdot u) = 0 \text{ for all } q \in Y$$

Therefore the second equation of (4.2.2) is verified.

Using lemma 4.2.5 and the previous result gives

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} (\nabla u^\varepsilon, \nabla v) &= (\nabla u, \nabla v) && \text{for all } v \in X \\ \lim_{\varepsilon \rightarrow 0} ((u^\varepsilon \cdot \nabla) u^\varepsilon, \nabla v) &= ((u \cdot \nabla) u, \nabla v) && \text{for all } v \in X \\ \lim_{\varepsilon \rightarrow 0} (\nabla \cdot u^\varepsilon, u^\varepsilon \cdot v) &= 0 && \text{for all } v \in X \\ \lim_{\varepsilon \rightarrow 0} (|u^\varepsilon|^\alpha u^\varepsilon, \nabla v) &= (|u|^\alpha u, \nabla v) && \text{for all } v \in X \\ \lim_{\varepsilon \rightarrow 0} (p^\varepsilon, \nabla \cdot v) &= (p, \nabla \cdot v) && \text{for all } v \in X \end{aligned} \quad (4.2.59)$$

We replace these limits in the first equation (4.2.30) implies that the limit of $(u^\varepsilon, p^\varepsilon)$ is solution of the problem (4.2.2).

Finally, it remains to establish the estimate (4.2.55). Subtracting the first equation (4.2.2) from the first equation of (4.2.30) and using the second equation of (4.2.30) gives

$$\begin{aligned} (p^\varepsilon - \tilde{p}, \nabla \cdot v) &= \nu (\nabla (u^\varepsilon - u), \nabla v) + ((u^\varepsilon \cdot \nabla) u^\varepsilon, v) - ((u \cdot \nabla) u, v) + \frac{1}{2} (\nabla \cdot u^\varepsilon, u^\varepsilon \cdot v) \\ &\quad + (a |u^\varepsilon|^\alpha u^\varepsilon, v) - (a |u|^\alpha u, v) \quad \forall v \in X \end{aligned} \quad (4.2.60)$$

Taking $v = u^\varepsilon - u$ yields

$$\begin{aligned} (p^\varepsilon - p, \nabla \cdot (u^\varepsilon - u)) &= \nu \|u^\varepsilon - u\|_1^2 + ((u^\varepsilon \cdot \nabla) u^\varepsilon, u^\varepsilon - u) + \frac{1}{2} (\nabla \cdot u^\varepsilon, u^\varepsilon \cdot (u^\varepsilon - u)) \\ &\quad - ((u \cdot \nabla) u, (u^\varepsilon - u)) + a(|u^\varepsilon|^\alpha u^\varepsilon, (u^\varepsilon - u)) - a(|u|^\alpha u, (u^\varepsilon - u)) \end{aligned} \quad (4.2.61)$$

Rewriting the fourth term on the right hand side in (4.2.61),

$$\begin{aligned} -((u \cdot \nabla) u, u^\varepsilon - u) &= (((u^\varepsilon - u - u^\varepsilon) \cdot \nabla)(u^\varepsilon - u + u^\varepsilon), (u^\varepsilon - u)) \\ &= (((u^\varepsilon - u) \cdot \nabla) u^\varepsilon, (u^\varepsilon - u)) + (((u^\varepsilon - u) \cdot \nabla)(u - u^\varepsilon), (u^\varepsilon - u)) \\ &\quad - ((u^\varepsilon \cdot \nabla) u^\varepsilon, (u^\varepsilon - u)) - ((u^\varepsilon \cdot \nabla)(u - u^\varepsilon), (u^\varepsilon - u)) \end{aligned}$$

We note that

$$\begin{aligned} -((u \cdot \nabla) u, u^\varepsilon - u) &= (((u^\varepsilon - u) \cdot \nabla) u^\varepsilon, (u^\varepsilon - u)) - ((u^\varepsilon \cdot \nabla) u^\varepsilon, (u^\varepsilon - u)) \\ &\quad - \frac{1}{2} (\nabla \cdot (u^\varepsilon - u), u^\varepsilon \cdot (u^\varepsilon - u)) - \frac{1}{2} (\nabla \cdot u^\varepsilon, u^\varepsilon \cdot (u^\varepsilon - u)) \end{aligned} \quad (4.2.62)$$

Replacing the previous result in (4.2.61) yields to

$$\begin{aligned}
(p^\varepsilon - \tilde{p}, \nabla \cdot (u^\varepsilon - u)) &= \nu \|u^\varepsilon - u\|_1^2 + ((u^\varepsilon - u) \cdot \nabla) u^\varepsilon, (u^\varepsilon - u) \\
&\quad + \frac{1}{2} (\nabla \cdot (u^\varepsilon - u), u^\varepsilon \cdot (u^\varepsilon - u)) + a(|u^\varepsilon|^\alpha u^\varepsilon - |u|^\alpha u, (u^\varepsilon - u)) \\
&\geq \nu \|u^\varepsilon - u\|_1^2 - s \|u^\varepsilon - u\|_1^2 \|u^\varepsilon\|_1
\end{aligned} \tag{4.2.63}$$

We use the fact that (u^ε) is bounded and by the inequality (4.2.46) we obtain

$$\|u^\varepsilon\|_1 \leq \frac{\nu}{s}(1-r)$$

Consequently

$$((p^\varepsilon - \tilde{p}), \nabla \cdot (u^\varepsilon - u)) \geq \nu r \|u^\varepsilon - u\|_1^2$$

By the second equation in (4.2.30) we have

$$\|u^\varepsilon - u\|_1^2 \leq \frac{\varepsilon}{\nu r} (\tilde{p} - p^\varepsilon, p^\varepsilon) = \frac{\varepsilon}{\nu r} (\tilde{p} - p^\varepsilon, \tilde{p}) - \frac{\varepsilon}{\nu r} \|\tilde{p} - p^\varepsilon\|^2 \tag{4.2.64}$$

Therefore

$$\|u^\varepsilon - u\|_1^2 \leq \frac{\varepsilon}{\nu r} \|\tilde{p} - p^\varepsilon\| \|\tilde{p}\|. \tag{4.2.65}$$

We now observe that

$$((u^\varepsilon \cdot \nabla) u^\varepsilon, v) = (((u^\varepsilon \cdot \nabla) u^\varepsilon, v) + ((u \cdot \nabla) u^\varepsilon, v) \tag{4.2.66}$$

$$-((u \cdot \nabla) u, v) = ((u \cdot \nabla)(u^\varepsilon - u), v) - ((u \cdot \nabla) u^\varepsilon, v) \tag{4.2.67}$$

$$\frac{1}{2} \nabla \cdot u^\varepsilon, u^\varepsilon \cdot v = \frac{1}{2} (\nabla \cdot (u^\varepsilon - u), u^\varepsilon \cdot v) \tag{4.2.68}$$

Injecting (4.2.66)-(4.2.68) in (4.2.60) to get

$$\begin{aligned}
(p^\varepsilon - \tilde{p}, \nabla \cdot v) &= \nu (\nabla \cdot (u^\varepsilon - u), \nabla v) + (((u^\varepsilon - u) \cdot \nabla) u^\varepsilon, v) + ((u \cdot \nabla)(u^\varepsilon - u), v) \\
&\quad + \frac{1}{2} (\nabla \cdot (u^\varepsilon - u), u^\varepsilon \cdot v) + a(|u^\varepsilon|^\alpha u^\varepsilon - |u|^\alpha u, v)
\end{aligned} \tag{4.2.69}$$

Using (4.2.69) in the condition inf-sup (4.2.48) gives with $q = p - \tilde{p}$

$$\|p^\varepsilon - \tilde{p}\|_Z \leq \frac{1}{\beta} (\nu \|u^\varepsilon - u\|_1 + s(\|u^\varepsilon\|_1 + \|u\|_1) \|u^\varepsilon - u\|_1 + C_1 a \|u^\varepsilon - u\|_1)$$

Note that $p^\varepsilon, \tilde{p} \in (\ker B^*)^\perp$, and from (4.2.46) we obtain

$$\|p^\varepsilon - \tilde{p}\| \leq \frac{1}{\beta} (\nu + 2(1-r)\nu + C_1 a) \|u^\varepsilon - u\|_1$$

Therefore

$$\| p^\varepsilon - \tilde{p} \| \leq \frac{A}{\beta} \| u^\varepsilon - u \|_1 \quad (4.2.70)$$

where $A = (3 - 2r)\nu + C_1 a$.

Putting (4.2.70) in (4.2.65) to have

$$\| u^\varepsilon - u \|_1 \leq \frac{A}{\nu r \beta} \| \tilde{p} \| \varepsilon. \quad (4.2.71)$$

Finally, we substitute (4.2.71) in (4.2.70) to infer that

$$\| p^\varepsilon - \tilde{p} \| \leq \frac{A^2}{\nu r \beta^2} \| \tilde{p} \| \varepsilon. \quad (4.2.72)$$

□

After the study of different proprieties about the boundary and penalized problem, we pass to the shape optimization. We introduce in the ensuig section the shape optimization problem. We derive an adjoint equation related to the perturbed equation. Finally, we compute the gradient of the cost function in terms of state and adjoint variables.

4.3 Control problem

During this section, we introduce a shape optimization problem to minimize an objective function. We derive an adjoint system related to the coming penalized problem

$$\begin{aligned} -\nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla) u^\varepsilon + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon + a |u^\varepsilon|^\alpha u^\varepsilon + \nabla p^\varepsilon &= f \quad \text{in } \Omega \\ \nabla \cdot u^\varepsilon + \varepsilon p^\varepsilon &= 0 \quad \text{in } \Omega \\ u^\varepsilon &= 0 \quad \text{on } \Gamma_0 \\ u^\varepsilon &= g \quad \text{on } \Gamma_1 \\ \nu \frac{\partial u^\varepsilon}{\partial n} - p^\varepsilon \mathbf{n} &= 0 \quad \text{on } \Gamma_2 \end{aligned} \quad (4.3.1)$$

and compute the shape gradient of the cost function in terms of state and adjoint variables. For simplicity we assume that $f = 0$. We first define the objective function

$$J(\Omega) = J_1(\Omega) + J_2(\Omega) \quad (4.3.2)$$

where

$$J_1(\Omega) = \frac{1}{2} \int_{\Omega} |u^\varepsilon - u_d|^2 dx, \text{ and } J_2(\Omega) = \frac{\sigma}{2} \int_{\Omega} |\text{curl}(u^\varepsilon)|^2 dx$$

with u_d is the target velocity. It depends on the ability of fish swimming.

We introduce also the following functions

$$F(\Omega, u, p, v, q) = \int_{\Omega} \left(\nu \nabla u^\varepsilon : \nabla v + (u^\varepsilon \cdot \nabla) u^\varepsilon \cdot v + \frac{1}{2} (\nabla \cdot u^\varepsilon) u^\varepsilon \cdot v + a |u^\varepsilon|^\alpha u^\varepsilon \cdot v - p^\varepsilon \nabla \cdot v \right) dx \\ - \int_{\Omega} q \nabla \cdot u^\varepsilon dx - \varepsilon \int_{\Omega} q p^\varepsilon dx$$

Then the optimization problem can be expressed as

$$\min_{\Omega \in \Omega_{ad}} J(\Omega) = \frac{1}{2} \int_{\Omega} |u^\varepsilon - u_d|^2 dx + \frac{\sigma}{2} \int_{\Omega} |\text{curl}(u^\varepsilon)|^2 dx \quad (4.3.3) \\ \text{such that } (u^\varepsilon, p^\varepsilon) \text{ satisfies the equation (4.3.1)}$$

4.3.1 Adjoint system

We begin by deriving an adjoint system associated with the penalized system (4.3.1). This theorem describes the adjoint system related to the general penalized problem (4.3.1).

Theorem 4.3.1. *Let $(u^\varepsilon, p^\varepsilon) \in \hat{V}_g(\Omega) \times Y$. The adjoint system associated to the equation (4.3.1) takes the form*

$$-\nu \Delta v + (\nabla u^\varepsilon)^T \cdot v - (u^\varepsilon \cdot \nabla) v - \frac{1}{2} \nabla (u^\varepsilon \cdot v) + \frac{1}{2} (\nabla \cdot u^\varepsilon) v + a |u^\varepsilon|^{\alpha-2} (|u^\varepsilon|^2 v + \alpha (u^\varepsilon \cdot v) u^\varepsilon) \\ + \nabla q = (u^\varepsilon - u_d) - \sigma \vec{\text{curl}}(\text{curl}(u^\varepsilon)) \quad \text{on } \Omega \\ \nabla \cdot v + \varepsilon q = 0 \quad \text{on } \Omega \\ v = 0 \quad \text{on } \Gamma_0 \cup \Gamma_1 \\ \nu \frac{\partial v}{\partial n} + 2(u^\varepsilon \cdot n) v - n q = \sigma \text{curl}(u^\varepsilon) \cdot \tau \quad \text{on } \Gamma_2 \quad (4.3.4)$$

Proof. Let $L(\Omega, u^\varepsilon, p^\varepsilon, v, q)$ be a Lagrangian functional defined by

$$L(\Omega, u^\varepsilon, p^\varepsilon, v, q) = J(\Omega) - F(\Omega, u^\varepsilon, p^\varepsilon, v, q) \quad (4.3.5)$$

We first derive L with respect to the state variable p in any direction $\tilde{p} \in M$ to obtain

$$\frac{\partial L}{\partial p^\varepsilon}(\Omega, u^\varepsilon, p^\varepsilon, v, q) \cdot \tilde{p} = \int_{\Omega} \tilde{p} \nabla \cdot v dx + \int_{\Omega} \varepsilon \tilde{p} q dx$$

The variation \tilde{p} is arbitrary, we get

$$\nabla \cdot v + \varepsilon q = 0 \quad \text{on } \Omega \quad (4.3.6)$$

Next, we derive L with respect of the state variable u^ε in the arbitrary direction $\tilde{u} \in V_0(\Omega)$ to have

$$\begin{aligned}
0 &= \frac{\partial L}{\partial u^\varepsilon}(\Omega, u^\varepsilon, p^\varepsilon, v, q) \cdot \tilde{u} \\
&= \int_{\Omega} (u^\varepsilon - u_d) \cdot \tilde{u} dx + \sigma \int_{\Omega} \text{curl}(u^\varepsilon) \cdot \text{curl}(\tilde{u}) dx \\
&\quad - \int_{\Omega} \left[\nu \nabla \tilde{u} : \nabla v + (\tilde{u} \cdot \nabla) u^\varepsilon \cdot v + (u^\varepsilon \cdot \nabla) \tilde{u} \cdot v + \frac{1}{2} (\nabla \cdot \tilde{u}) u^\varepsilon \cdot v + (\nabla \cdot u^\varepsilon) \tilde{u} \cdot v \right] dx \\
&\quad + \int_{\Omega} q \nabla \cdot \tilde{u} dx - \int_{\Omega} [a |u^\varepsilon|^\alpha \tilde{u} + a \alpha |u^\varepsilon|^{\alpha-2} (u^\varepsilon \cdot \tilde{u}) u^\varepsilon] \cdot v dx \\
&= \int_{\Omega} [(u^\varepsilon - u_d) - \sigma \vec{\text{curl}}(\text{curl}(u^\varepsilon)) + \nu \Delta u^\varepsilon - (\nabla u^\varepsilon)^T \cdot v + (u^\varepsilon \cdot \nabla) v \\
&\quad + \frac{1}{2} \nabla (u^\varepsilon \cdot v) - \frac{1}{2} (\nabla \cdot u^\varepsilon) v - (a |u^\varepsilon|^\alpha v + a \alpha |u^\varepsilon|^{\alpha-2} (u^\varepsilon \cdot v) u^\varepsilon - \nabla q)] \cdot \tilde{u} dx \\
&\quad + \int_{\Gamma_2} \left(\sigma \text{curl}(u^\varepsilon) \tau - \left(\nu \frac{\partial v}{\partial n} + 2(u^\varepsilon \cdot n) v - n q \right) \right) \cdot \tilde{u} ds
\end{aligned}$$

Considering an arbitrary direction \tilde{u} which vanishes in neighbourhood of the boundary Γ_2 , we write

$$\begin{aligned}
-\nu \Delta v + (\nabla u^\varepsilon)^T \cdot v - (u^\varepsilon \cdot \nabla) v - \frac{1}{2} \nabla (u^\varepsilon \cdot v) + \frac{1}{2} (\nabla \cdot u^\varepsilon) v + a |u^\varepsilon|^{\alpha-2} (|u^\varepsilon|^2 v + \alpha (u^\varepsilon \cdot v) u^\varepsilon) \\
+ \nabla q = (u^\varepsilon - u_d) - \sigma \vec{\text{curl}}(\text{curl}(u^\varepsilon))
\end{aligned}$$

An arbitrary \tilde{u} in Γ_2 gives

$$\nu \frac{\partial v}{\partial n} + 2(u^\varepsilon \cdot n) v - n q = \sigma \text{curl}(u^\varepsilon) \cdot \tau \quad \text{on } \Gamma_2$$

Finally we obtain the adjoint system

$$\begin{aligned}
-\nu \Delta v + (\nabla u^\varepsilon)^T \cdot v - (u^\varepsilon \cdot \nabla) v - \frac{1}{2} \nabla (u^\varepsilon \cdot v) + \frac{1}{2} (\nabla \cdot u^\varepsilon) v + a |u^\varepsilon|^{\alpha-2} (|u^\varepsilon|^2 v + \alpha (u^\varepsilon \cdot v) u^\varepsilon) \\
+ \nabla q = (u^\varepsilon - u_d) - \sigma \vec{\text{curl}}(\text{curl}(u^\varepsilon)) \quad \text{on } \Omega
\end{aligned}$$

$$\nabla \cdot v + \varepsilon q = 0 \quad \text{on } \Omega$$

$$v = 0 \quad \text{on } \Gamma_0 \cup \Gamma_1$$

$$\nu \frac{\partial v}{\partial n} + 2(u^\varepsilon \cdot n) v - n q = \sigma \text{curl}(u^\varepsilon) \cdot \tau \quad \text{on } \Gamma_2$$

(4.3.7)

□

4.3.2 Shape gradient

In this subsection, we express the shape gradient using the state and adjoint systems.

Consider Ω be a reference domain in \mathbb{R}^2 , the perturbation of Ω by the velocity method is described as the flow defined by the initial value problem

$$\begin{aligned} \frac{d\chi}{dt}(t, \mathbf{X}) &= \mathbf{V}(t, \chi(t)), \\ \chi(0, \mathbf{X}) &= \mathbf{X}, \end{aligned} \tag{4.3.8}$$

Let $T_t(\mathbf{X}) = \chi(t, \mathbf{X})$, $\forall \mathbf{X} \in \Omega$.

Let Ω_t be a perturbation domain of Ω and $J(\Omega)$ be a functional associated to Ω_t . The shape derivative of the functional $J(\Omega_t)$ at Ω in the direction of the deformation field \mathbf{V} is written as

$$dJ(\Omega; \mathbf{V}) = \lim_{t \rightarrow 0} \frac{1}{t} (J(\Omega_t) - J(\Omega)).$$

If $dJ(\Omega; \mathbf{V})$ exists for all $\mathbf{V} \in \mathcal{C}([0, T]; (\mathcal{D}^k(\mathbb{R}^2))^2)$, for small positive constant T , the functional J is called shape differentiable at Ω and its shape gradient satisfies

$$dJ(\Omega, \mathbf{V}) = (\nabla J, \mathbf{V})_{((\mathcal{D}^k(\mathbb{R}^2))^2)' \times (\mathcal{D}^k(\mathbb{R}^2))^2}$$

We assume that the boundaries Γ_1 and Γ_2 are fixed in the admissible set of domains. And we define the velocity field admissible domain as follow

$$V_{ad} = \{ \mathbf{V} \in \mathcal{C}^0(0, \tau; (\mathcal{C}^2(\mathbb{R}^2))^2) \mid \mathbf{V} = 0 \text{ on } \Gamma_1 \cup \Gamma_2 \}$$

The use of the velocity \mathbf{V} for $t \geq 0$, implies the transformation of the domain Ω into $\Omega_t = T_t(\Omega)$ by the velocity method with formulation (4.3.8).

Let us find an expression of the derivative of the saddle point problem $j(t)$ with respect to t , where

$$j(t) = \min_{(u_t^\varepsilon, p_t^\varepsilon) \in V_g(\Omega_t) \times \hat{Y}(\Omega_t)} \max_{(v_t, q_t) \in V_0(\Omega_t) \times \hat{Y}(\Omega_t)} L(\Omega_t, u_t^\varepsilon, p_t^\varepsilon, v_t, q_t)$$

and (u_t, p_t) and (v_t, q_t) are solutions of (4.3.1) and (4.3.4) in the perturbed domain Ω_t , respectively.

We consider the Hilbert spaces which depend on the parameter t defined by

$$\begin{aligned} V_0(\Omega_t) &= \{ u^\varepsilon \circ T_t^{-1} : u^\varepsilon \in V_0(\Omega) \} \\ V_g(\Omega_t) &= \{ u^\varepsilon \circ T_t^{-1} : u^\varepsilon \in V_g(\Omega) \} \\ \hat{Y}(\Omega_t) &= \{ p^\varepsilon \circ T_t^{-1} : p^\varepsilon \in \hat{Y}(\Omega) \} \end{aligned}$$

Since T_t and T_t^{-1} are diffeomorphisms, the parametrisation do not influence $j(t)$, and we have

$$j(t) = \min_{(u^\varepsilon, p^\varepsilon) \in V_g(\Omega) \times \hat{Y}(\Omega)} \max_{(v, q) \in V_0(\Omega) \times \hat{Y}(\Omega)} L(\Omega_t, u^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, v \circ T_t^{-1}, q \circ T_t^{-1})$$

We introduce the following functions which depend on the parameter t

$$\begin{aligned} l_1(t) &= \frac{1}{2} \int_{\Omega_t} |u^\varepsilon \circ T_t^{-1} - u_d|^2 dx + \frac{\sigma}{2} \int_{\Omega} |\operatorname{curl}(u^\varepsilon \circ T_t^{-1})|^2 dx, \\ l_2(t) &= \int_{\Omega_t} [\nu \nabla(u^\varepsilon \circ T_t^{-1}) : \nabla(v \circ T_t^{-1}) + ((u^\varepsilon \circ T_t^{-1}) \cdot \nabla)(u^\varepsilon \circ T_t^{-1}) \cdot (v \circ T_t^{-1}) \\ &\quad + a |u^\varepsilon \circ T_t^{-1}|^\alpha (u^\varepsilon \circ T_t^{-1}) \cdot (v \circ T_t^{-1}) - (p^\varepsilon \circ T_t^{-1}) \nabla \cdot (v \circ T_t^{-1})] dx \\ &\quad - \int_{\Omega_t} (q \circ T_t^{-1}) \nabla \cdot (u^\varepsilon \circ T_t^{-1}) dx - \varepsilon \int_{\Omega_t} (q \circ T_t^{-1}) (p^\varepsilon \circ T_t^{-1}) dx \end{aligned}$$

The Lagrangian functional writes

$$L(\Omega_t, u^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, v \circ T_t^{-1}, q \circ T_t^{-1}) = l_1(t) - l_2(t)$$

If $\Phi : [0, \tau] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is sufficiently smooth, we have the following Hadamard formula

$$\frac{d}{dt} \int_{\Omega_t} \Phi(t, x) dx \Big|_{t=0} = \int_{\Omega} \frac{\partial \Phi}{\partial t}(0, x) dx + \int_{\partial \Omega} \Phi(0, x) V(0, X) \cdot n ds \quad (4.3.9)$$

Let $\mathbf{V}(0, X) \in V_{ad}$, and observe that $\mathbf{V}(0, X) = \mathbf{V}$. Therefore we can derive the shape gradient using the formula (4.3.9):

$$\frac{d}{dt} L(\Omega_t, u^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, v \circ T_t^{-1}, q \circ T_t^{-1}) \Big|_{t=0} = l'_1(0) - l'_2(0) \quad (4.3.10)$$

where

$$\begin{aligned} l'_1(0) &= \int_{\Omega} (u^\varepsilon - u_d) \cdot (-\nabla u^\varepsilon \cdot \mathbf{V}) dx + \sigma \int_{\Omega} \operatorname{curl}(u^\varepsilon) \operatorname{curl}(-\nabla u^\varepsilon \cdot \mathbf{V}) dx \\ &\quad + \frac{1}{2} \int_{\Gamma_0} (|u^\varepsilon - u_d|^2 \mathbf{V} \cdot n) ds + \frac{\sigma}{2} \int_{\Gamma_0} (|\operatorname{curl}(u^\varepsilon)|^2 \mathbf{V} \cdot n) ds \end{aligned} \quad (4.3.11)$$

$$\begin{aligned} l'_2(0) &= \int_{\Omega} [\nu \nabla(-\nabla u^\varepsilon \cdot \mathbf{V}) : \nabla v + \nu \nabla u^\varepsilon : \nabla(-\nabla v \cdot \mathbf{V}) \\ &\quad + ((-\nabla u^\varepsilon \cdot \mathbf{V}) \cdot \nabla u^\varepsilon) \cdot v + (u^\varepsilon \cdot \nabla(-\nabla u^\varepsilon \cdot \mathbf{V})) \cdot v + (u^\varepsilon \cdot \nabla u^\varepsilon) \cdot (-\nabla v \cdot \mathbf{V}) \\ &\quad + \frac{1}{2} \nabla \cdot (-\nabla u^\varepsilon \cdot \mathbf{V}) u^\varepsilon \cdot v + \frac{1}{2} \nabla \cdot u^\varepsilon (-\nabla u^\varepsilon \cdot \mathbf{V}) \cdot v + \frac{1}{2} \nabla \cdot u^\varepsilon u^\varepsilon \cdot (-\nabla v \cdot \mathbf{V}) \\ &\quad + a \alpha |u^\varepsilon|^{\alpha-2} (-\nabla u^\varepsilon \cdot \mathbf{V} \cdot u^\varepsilon) (u^\varepsilon \cdot v) + a |u^\varepsilon|^\alpha (-\nabla u^\varepsilon \cdot \mathbf{V}) \cdot v \\ &\quad + a |u^\varepsilon|^\alpha u^\varepsilon \cdot (-\nabla v \cdot \mathbf{V}) - (\nabla p \cdot \mathbf{V}) \nabla \cdot v \\ &\quad - p \nabla \cdot (-\nabla v \cdot \mathbf{V}) - q \nabla \cdot (-\nabla u^\varepsilon \cdot \mathbf{V}) - (\nabla \cdot u^\varepsilon) (-\nabla q \cdot \mathbf{V})] dx \end{aligned}$$

Using Green's formula and the condition $u^\varepsilon = 0$ on Γ_0 , we obtain

$$\begin{aligned}
l'_2(0) = & - \int_{\Omega} [(-\nu \Delta u^\varepsilon + (u^\varepsilon \cdot \nabla)u^\varepsilon + \frac{1}{2} \nabla \cdot u^\varepsilon u^\varepsilon + a|u^\varepsilon|^\alpha u^\varepsilon + \nabla p) \cdot (\nabla v \cdot \mathbf{V})] dx \\
& + \int_{\Omega} (\nabla \cdot u^\varepsilon)(\nabla q \cdot \mathbf{V}) dx + \int_{\Omega} (\nabla \cdot v)(\nabla p \cdot \mathbf{V}) dx + \int_{\Omega} [-\nu \Delta v + (\nabla u^\varepsilon)^T \cdot v - (u^\varepsilon \cdot \nabla)v \\
& - \frac{1}{2} \nabla(u^\varepsilon \cdot v) + \frac{1}{2} (\nabla \cdot u^\varepsilon)v + a|u^\varepsilon|^{\alpha-2}(|u^\varepsilon|^2 v + \alpha(u^\varepsilon \cdot v)u^\varepsilon) - \nabla q] \cdot (\nabla u^\varepsilon \cdot \mathbf{V}) dx \\
& - \int_{\Gamma_0} [\nu \frac{\partial v}{\partial \mathbf{n}} + 2(u^\varepsilon \cdot \mathbf{n})v - nq] \cdot (\nabla u^\varepsilon \cdot \mathbf{V}) ds - \int_{\Gamma_0} [\frac{\partial u^\varepsilon}{\partial \mathbf{n}} - nq] \cdot (\nabla v \cdot \mathbf{V}) ds \\
& + \int_{\Gamma_0} [\nu \nabla u^\varepsilon : \nabla v + (u^\varepsilon \cdot \nabla u^\varepsilon) \cdot v + a|u^\varepsilon|^\alpha u^\varepsilon \cdot v - p \nabla \cdot v - q \nabla \cdot u^\varepsilon] \mathbf{V} \cdot \mathbf{n} ds
\end{aligned} \tag{4.3.12}$$

Replacing (4.3.11) and (4.3.12) in (4.3.10) and using the fact that $(u^\varepsilon, p^\varepsilon)$ is solution of (4.3.1) and (v, q) is solution of (4.3.4) respectively yields to

$$\begin{aligned}
dJ(\Omega; \mathbf{V}) &= \frac{d}{dt} L|_{t=0} \\
&= \frac{1}{2} \int_{\Gamma_0} (|u^\varepsilon - u_d|^2 + \sigma |curl(u^\varepsilon)|^2) \mathbf{V} \cdot \mathbf{n} ds + \int_{\Gamma_0} [\nu \frac{\partial v}{\partial \mathbf{n}} + nq] \cdot (\nabla u^\varepsilon \cdot \mathbf{V}) ds \\
&\quad + \int_{\Gamma_0} [\nu \frac{\partial u^\varepsilon}{\partial \mathbf{n}} - n p] \cdot (\nabla v \cdot \mathbf{V}) ds - \int_{\Gamma_0} [(\nu \nabla u^\varepsilon : \nabla v) \mathbf{V} \cdot \mathbf{n} ds \\
&\quad - \sigma \int_{\Gamma_0} (\frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot (curl(u^\varepsilon) \wedge \mathbf{n})) \mathbf{V} \cdot \mathbf{n} ds
\end{aligned} \tag{4.3.13}$$

Note that $u^\varepsilon = 0, v=0$ in Γ_0 we have

$$\mathbf{n} \cdot (\nabla u^\varepsilon \cdot \mathbf{V}) = \nabla u^\varepsilon \cdot (\mathbf{n} \otimes \mathbf{n}) \cdot \mathbf{V} \cdot \mathbf{n} = \nabla u^\varepsilon \cdot \mathbf{n} \cdot \mathbf{n} (\mathbf{V} \cdot \mathbf{n}) = (\nabla \cdot u^\varepsilon)(\mathbf{V} \cdot \mathbf{n}) = 0, \quad \forall x \in \Gamma_0 \tag{4.3.14}$$

$$\frac{\partial v}{\partial \mathbf{n}} \cdot (\nabla u^\varepsilon \cdot \mathbf{V}) = \nabla u^\varepsilon \cdot (\mathbf{n} \otimes \mathbf{n}) \cdot \mathbf{V} \cdot \frac{\partial v}{\partial \mathbf{n}} = \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \frac{\partial v}{\partial \mathbf{n}} (\mathbf{V} \cdot \mathbf{n}) = (\nabla u^\varepsilon : \nabla v) \mathbf{V} \cdot \mathbf{n} \tag{4.3.15}$$

With a similar manner we obtain

$$\mathbf{n} \cdot (\nabla v \cdot \mathbf{V}) = 0, \quad \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot (\nabla v \cdot \mathbf{V}) = \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \frac{\partial v}{\partial \mathbf{n}} (\mathbf{V} \cdot \mathbf{n}) = (\nabla u^\varepsilon : \nabla v) \mathbf{V} \cdot \mathbf{n} \tag{4.3.16}$$

Substituting (4.3.14)-(4.3.16) in (4.3.13) the shape derivative takes the form

$$dJ(\Omega; \mathbf{V}) = \int_{\Gamma_0} \left[\frac{1}{2} |u^\varepsilon - u_d|^2 + \frac{\sigma}{2} |curl(u^\varepsilon)|^2 + \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial v}{\partial \mathbf{n}} - \sigma curl(u^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{V} \cdot \mathbf{n} ds \tag{4.3.17}$$

Consequently, the shape gradient writes

$$\nabla J = \left[\frac{1}{2} |u^\varepsilon - u_d|^2 + \frac{\sigma}{2} |curl(u^\varepsilon)|^2 + \frac{\partial u^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial v}{\partial \mathbf{n}} - \sigma curl(u^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{n} \tag{4.3.18}$$

After the computation of the gradient of the objective function and the derivation of the adjoint system, we adopt a discrete finite element method to discretize the penalized problem. We establish an error estimates to prove the convergence of the discrete approximate solution to the boundary problem solution.

4.4 Finite element method

We consider a finite element space Ω_h of Ω . Let $\bar{\Omega}_h = \cup_k \bar{T}_k$, with $\{T_k\}$ is a partition of Ω into non overlapping elements. Let us introduce the finite element dimensional space X^h of $X = (H_0^1(\Omega))^d$, $d=2$ or 3 as follow

$$X^h = \{v \in (C(\bar{\Omega}))^d \mid v|_{T_k} \in Q^k \text{ for all } T_k \subset \Omega_h\} \quad (4.4.1)$$

where Q^k is the space of polynomials functions of degree k given by

$$Q^k = \text{span} \left\{ \prod_{i=1}^d x_i^{b_i} \mid 0 \leq b_i \leq k \right\} \quad (4.4.2)$$

We define the integration rule $I(\cdot)$ by

$$\int_{\Omega} g dx \simeq I(g) = \sum_{k=1}^E \sum_{i=1}^{G_k} g(x_i^k) w_i^k \quad (4.4.3)$$

where $x_i^k \in \bar{\Omega}_k$, are the integration point coordinates, w_i^k , $1 \leq i \leq G_k$, $1 \leq k \leq E$, are the integration weights of the Gauss rule $I(\cdot)$ on any element Ω_k and E is the number of elements in the mesh.

The pressure space Y^h is piecewise discontinuous defined with the points in the integration above. We have for any $q_h \in Y^h$

$$I(q_h \nabla \cdot v_h) = (q_h, \nabla \cdot v_h) \quad \text{for all } v_h \in X^h \quad (4.4.4)$$

For $f \in (X^h)'$, and for $\varepsilon > 0$ fixed we introduce the approximate perturbed problem reads as

Find $u_h^\varepsilon \in X^h$ such that

$$\begin{aligned} \nu(\nabla u_h^\varepsilon, \nabla v) + ((u_h^\varepsilon \cdot \nabla) u_h^\varepsilon, v_h) + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h) + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon, v_h) \\ + \frac{1}{\varepsilon} I((\nabla \cdot u_h^\varepsilon)(\nabla \cdot v_h)) = (f, v_h) \quad \text{for all } v_h \in X^h \end{aligned} \quad (4.4.5)$$

Moreover, the approximate pressure verifies

$$p_h^\varepsilon(x_i^e) = -\frac{1}{\varepsilon} \nabla \cdot u_h^\varepsilon(x_i^e), \quad 1 \leq i \leq G, \quad 1 \leq e \leq E \quad (4.4.6)$$

We assume again that the condition (4.2.46) is satisfied by f

$$\frac{s_h}{\nu} \|f\|_{(X^h)'} \leq \frac{s}{\nu^2} \|f\|_{X'} \leq 1 - r \quad (4.4.7)$$

The next result focuses on the existence of a weak solution for the approximate perturbed problem (4.4.5).

Theorem 4.4.1. *Assume that the regularity of the data (4.4.7) holds for $h, \varepsilon > 0$. We have the following proprieties*

1. *There exists a solution to the approximate perturbed problem (4.4.5).*
2. *The sequence $(u_h^\varepsilon) \in X^h$ is uniformly bounded in ε .*
3. *The following estimates*

$$\frac{\nu}{2} \|\nabla u_h^\varepsilon\|^2 + a \|u_h^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p_h^\varepsilon\|^2 \leq \frac{1}{2\nu} \|f\|_{X'}^2 \quad (4.4.8)$$

hold.

Proof. We follow similar steps as for the continuous problem in theorem 4.2.1 to obtain

$$\|u_h^\varepsilon\|_{X^h} \leq \frac{1}{\nu} \|f\|_{(X^h)'} \quad (4.4.9)$$

□

We set the inf-sup stability condition for the approximate problem. We suppose, there exists a constant $\beta_h > 0$, such that

$$\sup_{v_h \in X^h} \frac{I(q_h \nabla \cdot v_h)}{\|v_h\|_V} \geq \beta_h \|\hat{q}_h\|_{Z^h} \quad \text{for all } q_h \in Y^h \quad (4.4.10)$$

We introduce the finite element space Z^h with

$$\hat{q}_h = \{\bar{q}_h \in Y^h \mid \bar{q}_h - q_h \in \text{Ker} B_h^*\} \quad (4.4.11)$$

We also define the operators $B_h \in \mathcal{L}(X^h, (Y^h)')$, $B_h^* \in \mathcal{L}(Y^h, (X^h)')$ by

$$I(q_h \nabla \cdot v_h) = (q_h, B_h v_h)_{(Y^h)' \times Y^h} = (B_h^* q_h, v_h)_{(X^h)' \times X^h}, \quad \text{for all } (v_h, q_h) \in X^h \times Y^h \quad (4.4.12)$$

The kernel of the operator B_h^* is defined by

$$\text{Ker} B_h^* = \{q_h \in Y^h \mid I(q_h \nabla \cdot v_h) = 0 \quad \forall v_h \in X^h\} \quad (4.4.13)$$

The norm on Z^h in (4.4.10) writes

$$\|\hat{q}_h\|_{Z^h} = \inf_{q_h^0 \in \text{Ker} B_h^*} \|q_h^0 + q_h\| \quad (4.4.14)$$

The main idea of these results is that the pressure is bounded in ε and that the solution for the approximate penalized problem (4.4.5) converges to the solution of a boundary approximate problem.

Theorem 4.4.2. Consider that u_h^ε is solution of the problem (4.4.5) in X^h and the pressure p_h^ε defined by (4.4.6) in Y^h . Then

- (i) The approximate discrete pressure p_h^ε is uniformly bounded in ε .
- (ii) The couple $(u_h^\varepsilon, p_h^\varepsilon)$ converges to the solution (u_h, p_h) of the problem

Find $(u_h, p_h) \in X^h \times Y^h$ such that

$$\begin{aligned} \nu(\nabla u_h, \nabla v_h) + ((u_h \cdot \nabla)u_h, v_h) + \frac{1}{2}((\nabla \cdot u_h), u_h \cdot v_h) + a(|u_h|^\alpha u_h, v_h) \\ - (p_h, \nabla \cdot v_h) = (f, v_h) \text{ for all } v_h \in X^h \\ (q_h, \nabla \cdot u_h) = 0 \text{ for all } q_h \in Y^h \end{aligned} \quad (4.4.15)$$

Proof. The theorem 4.4.1 ensures that $(u_h^\varepsilon) \in X^h$ is uniformly bounded in ε . Replacing (4.4.6) in (4.4.5)

$$\begin{aligned} I(p_h^\varepsilon \nabla \cdot v_h) = - (f, v_h) + \nu(\nabla u_h^\varepsilon, \nabla v_h) + ((u_h^\varepsilon \cdot \nabla)u_h^\varepsilon, v_h) \\ + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h) + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon, v_h) \end{aligned}$$

Substituting this results in the discrete inf-sup condition (4.4.10),

$$\| \hat{p}_h^\varepsilon \|_{Z^h} \leq \frac{1}{\beta_h} \left\{ \| f \|_{(X^h)'} + \nu \| u_h^\varepsilon \|_1 + c_1 \| u_h^\varepsilon \|_1^2 + c_2 \| u_h^\varepsilon \|_1 \right\} \leq \frac{C}{\beta_h}. \quad (4.4.16)$$

Or $p_h^\varepsilon \in (Ker B_h^*)^\perp$, we infer that

$$\| p_h^\varepsilon \|_{Z^h} = \| p_h^\varepsilon \|$$

Therefore, (4.4.16) gives

$$\| p_h^\varepsilon \| \leq \frac{C}{\beta_h} \quad (4.4.17)$$

Since C is independent of ε .

The sequence $(u_h^\varepsilon, p_h^\varepsilon)$ is bounded, then there exists a subsequence denoted $(u_h^\varepsilon, p_h^\varepsilon)$ without lose of generality which converge to $(u_h, p_h) \in X^h \times Y^h$. And we have

$$\lim_{\varepsilon \rightarrow 0} I(q_h \nabla \cdot u_h^\varepsilon) = I(q_h \nabla u_h), \text{ for all } q_h \in Y^h \quad (4.4.18)$$

Using (4.4.6),

$$I(q_h \nabla \cdot u_h^\varepsilon) = -\varepsilon I(q_h p_h^\varepsilon) \leq \varepsilon \| q_h \| \| p_h^\varepsilon \|$$

We use the fact that (p_h^ε) is bounded, and we pass to limits as $\varepsilon \rightarrow 0$ to obtain

$$I(q_h \nabla \cdot v_h) = 0$$

Therefore the second equation of (4.4.15) is established. For the first equation of (4.4.15) we use

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \nu(\nabla u_h^\varepsilon, \nabla v_h) &= \nu(\nabla u_h, \nabla v_h) && \text{for all } v_h \in X^h \\ \lim_{\varepsilon \rightarrow 0} ((u_h^\varepsilon \cdot \nabla) u_h^\varepsilon, v_h) &= ((u_h \cdot \nabla) u_h, v_h) && \text{for all } v_h \in X^h \\ \lim_{\varepsilon \rightarrow 0} \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h) &= \frac{1}{2}(\nabla \cdot u_h, u_h \cdot v_h) && \text{for all } v_h \in X^h \\ \lim_{\varepsilon \rightarrow 0} (|u_h^\varepsilon|^\alpha u_h^\varepsilon, v_h) &= (|u_h|^\alpha u_h, v_h) && \text{for all } v_h \in X^h \\ \lim_{\varepsilon \rightarrow 0} (p_h^\varepsilon, \nabla \cdot v_h) &= (p_h, \nabla \cdot v_h) && \text{for all } v_h \in X^h \end{aligned}$$

□

We are now in a position to give a priori estimates of the error for the solution of the discrete penalized problem (4.4.5) and the solution of the approximate state problem (4.4.15).

Lemma 4.4.1. *Consider $(u_h^\varepsilon, p_h^\varepsilon) \in X^h \times Y^h$ defined by (4.4.5) and $(u_h, p_h) \in X^h \times Y^h$ solution of the problem (4.4.15). Then*

$$\|u_h^\varepsilon - u_h\|_1 \leq \varepsilon \frac{C_1}{\beta_h} \|p_h\| \quad (4.4.19)$$

and

$$\|p_h^\varepsilon - p_h\| \leq \varepsilon \frac{C_2}{\beta_h^2} \|p_h\| \quad (4.4.20)$$

where C_1 and C_2 are two positive constant non depending in ε .

Proof. We use the definition of the discrete pressure in (4.4.6) in the equation (4.4.5) and we subtract (4.4.15) to have

$$\begin{aligned} \nu(\nabla(u_h^\varepsilon - u_h), \nabla v_h) + ((u_h^\varepsilon \cdot \nabla) u_h^\varepsilon, v_h) + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h) + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon - |u_h|^\alpha u_h, v_h) \\ - ((u_h \cdot \nabla) u_h, v_h) - \frac{1}{2}(\nabla \cdot u_h, u_h \cdot v_h) + (p_h - p_h^\varepsilon, \nabla \cdot v_h) = 0 \quad \text{for all } v_h \in X^h \end{aligned} \quad (4.4.21)$$

With some calculus, for $v_h \in X^h$ we obtain

$$((u_h^\varepsilon \cdot \nabla) u_h^\varepsilon, v_h) = (((u_h^\varepsilon - u_h) \cdot \nabla) u_h^\varepsilon, v_h) + ((u_h \cdot \nabla) u_h^\varepsilon, v_h), \quad (4.4.22)$$

$$-((u_h \cdot \nabla) u_h, v_h) = -((u_h \cdot \nabla)(u_h - u_h^\varepsilon), v_h) - ((u_h \cdot \nabla) u_h^\varepsilon, v_h) \quad (4.4.23)$$

$$\frac{1}{2}((\nabla \cdot u_h^\varepsilon), u_h^\varepsilon \cdot v_h) = \frac{1}{2}(\nabla \cdot (u_h^\varepsilon - u_h), u_h^\varepsilon \cdot v_h) + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot v_h), \quad (4.4.24)$$

$$-\frac{1}{2}((\nabla \cdot u_h), u_h \cdot v_h) = \frac{1}{2}((\nabla \cdot u_h), (u_h^\varepsilon - u_h) \cdot v_h) - \frac{1}{2}(\nabla \cdot u_h, u_h^\varepsilon \cdot v_h) \quad (4.4.25)$$

Replacing (4.4.22)-(4.4.25) in (4.4.21), yields to

$$\begin{aligned} (p_h^\varepsilon - p_h, \nabla \cdot v_h) &= \nu(\nabla(u_h^\varepsilon - u_h), \nabla v_h) + (((u_h^\varepsilon - u_h) \cdot \nabla)u_h^\varepsilon, v_h) + ((u_h \cdot \nabla)(u_h^\varepsilon - u_h), v_h) \\ &\quad + \frac{1}{2}(\nabla \cdot (u_h^\varepsilon - u_h), u_h^\varepsilon \cdot v_h) + \frac{1}{2}(\nabla \cdot u_h, (u_h^\varepsilon - u_h) \cdot v_h) + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon - |u|^\alpha u_h, v_h) \end{aligned} \quad (4.4.26)$$

Substituting in (4.4.10) and using (4.4.7) and (4.4.9)

$$\| \hat{p}_h^\varepsilon - \hat{p}_h \|_{Z^h} \leq \frac{\nu + 2\nu(1-r) + aC_3}{\beta_h} \| u_h^\varepsilon - u_h \|_1 \quad (4.4.27)$$

Or, p_h^ε and p_h are in $(\text{Ker} B_h^*)^\perp$, then

$$\| \hat{p}_h^\varepsilon - \hat{p}_h \|_{Z^h} = \| \hat{p}_h^\varepsilon - \hat{p}_h \|$$

Consequently, (4.4.27) gives

$$\| p_h^\varepsilon - p_h \| \leq \frac{\nu(3-2r) + aC_3}{\beta_h} \| u_h^\varepsilon - u_h \|_1 \quad (4.4.28)$$

Consider $v_h = u_h^\varepsilon - u_h$ in the equation (4.4.21) to write

$$\begin{aligned} (p_h^\varepsilon - p_h, \nabla \cdot (u_h^\varepsilon - u_h)) &= \nu \| u_h^\varepsilon - u_h \|_1^2 + ((u_h^\varepsilon \cdot \nabla)u_h^\varepsilon, u_h^\varepsilon - u_h) + \frac{1}{2}(\nabla \cdot u_h^\varepsilon, u_h^\varepsilon \cdot (u_h^\varepsilon - u_h)) \\ &\quad - ((u_h \cdot \nabla)u_h, u_h^\varepsilon - u_h) - \frac{1}{2}((\nabla \cdot u_h, u_h \cdot (u_h^\varepsilon - u_h)) \\ &\quad + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon, u_h^\varepsilon - u_h) - a(|u_h|^\alpha u_h, u_h^\varepsilon - u_h) \end{aligned} \quad (4.4.29)$$

Using $u_h = -(u_h^\varepsilon - u_h - u_h^\varepsilon)$ and thanks to (5.2.13), and the inequality (4.4.27), (4.4.29) takes the form

$$\begin{aligned} (p_h^\varepsilon - p_h, \nabla \cdot (u_h^\varepsilon - u_h)) &= \nu \| u_h^\varepsilon - u_h \|_1^2 + (((u_h^\varepsilon - u_h) \cdot \nabla)u_h^\varepsilon, (u_h^\varepsilon - u_h)) \\ &\quad + \frac{1}{2}((\nabla \cdot (u_h^\varepsilon - u_h), u_h^\varepsilon \cdot (u_h^\varepsilon - u_h)) \\ &\quad + a(|u_h^\varepsilon|^\alpha u_h^\varepsilon, u_h^\varepsilon - u_h) - a(|u_h|^\alpha u_h, u_h^\varepsilon - u_h) \\ &\geq \nu \| u_h^\varepsilon - u_h \|_1^2 - s \| u_h^\varepsilon - u_h \|_1 \| u_h^\varepsilon \|_1 . \end{aligned}$$

By (4.4.7) and (4.4.9), we infer

$$(p_h^\varepsilon - p_h, \nabla \cdot (u_h^\varepsilon - u_h)) \geq \nu r \| u_h^\varepsilon - u_h \|_1^2$$

Using (4.4.15)

$$\| u_h^\varepsilon - u_h \|_1^2 \leq \frac{1}{\nu r} (p_h^\varepsilon - p_h, \nabla \cdot u_h^\varepsilon)$$

The use of the property (4.4.6) gives

$$\begin{aligned} \|u_h^\varepsilon - u_h\|_1^2 &\leq \frac{\varepsilon}{\nu r} (p_h^\varepsilon - p_h, p_h^\varepsilon) \\ &= \frac{\varepsilon}{\nu r} (p_h^\varepsilon - p_h, p_h) - \frac{\varepsilon}{\nu r} \|p_h^\varepsilon - p_h\|^2 \end{aligned}$$

Therefore

$$\|u_h^\varepsilon - u_h\|_1^2 \leq \frac{\varepsilon}{\nu r} \|p_h^\varepsilon - p_h\| \|p_h\| \quad (4.4.30)$$

Substituting (4.4.30) in (4.4.28) gives

$$\|u_h^\varepsilon - u_h\|_1^2 \leq \frac{(3-2r)\nu + aK_1}{\nu r \beta_h} \varepsilon \|p_h\| \quad (4.4.31)$$

Replacing this results in (4.4.28) to get

$$\|p_h^\varepsilon - p_h\|_1^2 \leq \frac{(\nu(3-2r) + aK_1)^2}{\nu r \beta_h^2} \varepsilon \|p_h\| \quad (4.4.32)$$

Finally the desired estimates (4.4.19) and (4.4.20) are obtained with $C_1 = \frac{(3-2r)\nu + aK_1}{\nu r}$ and $C_2 = \frac{(\nu(3-2r) + aK_1)^2}{\nu r}$. \square

Theorem 4.4.3. *Assume that (u, p) and (u_h, p_h) are solutions of the problem (4.2.2) and (4.4.15) respectively. Then*

$$\|u_h - u\|_1 \leq \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_1\right) \left(1 + \frac{C}{\beta_h}\right) \|u - v_h\|_1 + \frac{C_2}{\nu r} \|p - q_h\| \quad (4.4.33)$$

for all $v_h \in X^h$ and all $q_h \in Y^h$, and

$$\|\hat{p} - \hat{p}_h\|_{Z^h} \leq \frac{\nu(3-2r) + aC}{\beta_h} \|u - u_h\|_1 + \left(1 + \frac{C'}{\beta_h}\right) \|p - q_h\| \quad \text{for all } q_h \in Y^h \quad (4.4.34)$$

Proof. We introduce the finite space $\mathbb{X}^h = \{v_h \in X^h \mid (q_h, \nabla \cdot v_h) = 0 \text{ for all } q_h \in Y^h\}$. Since $u_h \in \mathbb{X}^h$ and $v_h \in \mathbb{X}^h$, let $v_h = u_h - v_h$ in (4.4.15) to get

$$\begin{aligned} \nu(\nabla u_h, \nabla(u_h - v_h)) + ((u_h \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\ + a(|u_h|^\alpha u_h, u_h - v_h) = (f, u_h - v_h) \quad \text{for all } v_h \in \mathbb{X}^h \end{aligned} \quad (4.4.35)$$

Subtracting the quantity

$$\nu(\nabla v_h, \nabla(u_h - v_h)) + ((v_h \cdot \nabla)v_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot v_h, v_h \cdot (u_h - v_h)) + a(|v_h|^\alpha v_h, u_h - v_h)$$

from the both sides of (4.4.35), multiplying (4.2.2) by $v = u_h - v_h$ and substituting here in the right hand side

$$\begin{aligned}
& \nu(\nabla(u_h - v_h), \nabla(u_h - v_h)) + ((u_h \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\
& + a(|u_h|^\alpha u_h, u_h - v_h) - ((v_h \cdot \nabla)v_h, u_h - v_h) - \frac{1}{2}(\nabla \cdot v_h, v_h \cdot (u_h - v_h)) \\
& - a(|v_h|^\alpha v_h, u_h - v_h) \\
& = \nu(\nabla u, \nabla(u_h - v_h)) + ((u \cdot \nabla)u, u_h - v_h) + a(|u|^\alpha u, u_h - v_h) \\
& - (p, \nabla \cdot (u_h - v_h)) - \nu(\nabla v_h, \nabla(u_h - v_h)) - ((v_h \cdot \nabla)v_h, u_h - v_h) \\
& - \frac{1}{2}(\nabla \cdot v_h, v_h \cdot (u_h - v_h)) - a(|v_h|^\alpha v_h, u_h - v_h) \text{ for all } v_h \in \mathbb{X}^h
\end{aligned} \tag{4.4.36}$$

Or, $v_h \in \mathbb{X}^h$, $(p, \nabla \cdot (u_h - v_h)) = (p - q_h, \nabla \cdot (u_h - v_h))$. Adding and subtracting $((u \cdot \nabla)v_h, u_h - v_h)$ on the right hand side in the previous result

$$\begin{aligned}
& \nu(\nabla(u_h - v_h), \nabla(u_h - v_h)) + ((u_h \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\
& + a(|u_h|^\alpha u_h - |v_h|^\alpha v_h, u_h - v_h) - ((v_h \cdot \nabla)v_h, u_h - v_h) - \frac{1}{2}(\nabla \cdot v_h, v_h \cdot (u_h - v_h)) \\
& = \nu(\nabla(u - v_h), \nabla(u_h - v_h)) + ((u \cdot \nabla)(u - v_h), u_h - v_h) + a(|u|^\alpha u - |v_h|^\alpha v_h, u_h - v_h) \\
& - (p - q_h, \nabla \cdot (u_h - v_h)) + ((u - v_h \cdot \nabla)v_h, u_h - v_h) - \frac{1}{2}(\nabla \cdot (u - v_h), v_h \cdot (u_h - v_h))
\end{aligned} \tag{4.4.37}$$

for all $v_h \in \mathbb{X}^h$.

The first part of the equality above can be simplified as

$$\begin{aligned}
& \nu(\nabla(u_h - v_h), \nabla(u_h - v_h)) + ((u_h \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\
& + a(|u_h|^\alpha u_h - |v_h|^\alpha v_h, u_h - v_h) - ((v_h \cdot \nabla)v_h, u_h - v_h) - \frac{1}{2}(\nabla \cdot v_h, v_h \cdot (u_h - v_h)) \\
& = \nu(\nabla(u_h - v_h), \nabla(u_h - v_h)) + ((u_h \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\
& - (((v_h - u_h) \cdot \nabla)(v_h - u_h), u_h - v_h) - ((u_h \cdot \nabla)(v_h - u_h), u_h - v_h) \\
& - (((v_h - u_h) \cdot \nabla)u_h, u_h - v_h) - ((u_h \cdot \nabla)u_h, u_h - v_h) \\
& - \frac{1}{2}(\nabla \cdot (v_h - u_h), (v_h - u_h) \cdot (u_h - v_h)) - \frac{1}{2}((\nabla \cdot u_h), (v_h - u_h) \cdot (u_h - v_h)) \\
& - \frac{1}{2}(\nabla \cdot (v_h - u_h), u_h \cdot (u_h - v_h)) - \frac{1}{2}(\nabla \cdot u_h, u_h \cdot (u_h - v_h)) \\
& + a(|u_h|^\alpha u_h - |v_h|^\alpha v_h, u_h - v_h) \\
& = \nu(\nabla(u_h - v_h), \nabla(u_h - v_h)) + (((u_h - v_h) \cdot \nabla)u_h, u_h - v_h) + \frac{1}{2}(\nabla \cdot (u_h - v_h), u_h \cdot (u_h - v_h)) \\
& + a(|u_h|^\alpha u_h - |v_h|^\alpha v_h, u_h - v_h)
\end{aligned} \tag{4.4.38}$$

$$\geq \nu \|u_h - v_h\|_1^2 - s \|u_h - v_h\|_1^2 \|u_h\|_1 \geq \nu r \|u_h - v_h\|_1^2 \tag{4.4.39}$$

Substituting (4.4.39) in (4.4.37) and thanks to (4.4.7), the fact that (u_h) is bounded and the Cauchy Schwartz inequality, we have

$$\begin{aligned} \|u_h - v_h\|_1 \leq & \frac{1}{r} \left(1 + \frac{s}{\nu} (\|u\|_1 + \|v_h\|_1) + \frac{a}{\nu} C_1 \right) \|u - v_h\|_1 \\ & + \frac{C_2}{\nu r} \|p - q_h\| \end{aligned} \quad (4.4.40)$$

for all $v_h \in \mathbb{X}^h$ and $q_h \in Y^h$.

Simplifying (4.4.40) gives

$$\begin{aligned} \|u - u_h\|_1 \leq & \left[1 + \frac{1}{r} \left(1 + \frac{s}{\nu} (\|u\|_1 + \|v_h\|_1) + \frac{a}{\nu} C_1 \right) \right] \|u - v_h\|_1 \\ & + \frac{C_2}{\nu r} \|p - q_h\| \end{aligned} \quad (4.4.41)$$

We define $v_h = \prod_h u$ as the orthogonal projection of u on \mathbb{X}^h . We have

$$\|u - \prod_h u\|_1 = \inf_{v_h \in \mathbb{X}^h} \|u - v_h\|_1$$

and

$$\|\prod_h u\|_1 \leq \|u\|_1$$

Using these properties in (4.4.41) and the fact that (u) is bounded yields

$$\|u - u_h\|_1 \leq \left[\frac{3}{r} - 1 + \frac{a}{\nu r} C_1 \right] \inf_{v_h \in \mathbb{X}^h} \|u - v_h\|_1 + \frac{C_2}{\nu r} \|p - q_h\| \quad (4.4.42)$$

It remains to estimate the term $\inf_{v_h \in \mathbb{X}^h} \|u - v_h\|_1$. For this end, we solve an auxiliary approximate penalized variational Stokes problem:

let $v_h \in X^h$, for all $\varepsilon > 0$, Find $z_h^\varepsilon \in X^h$ such that

$$(\nabla z_h^\varepsilon, \nabla w_h) - \frac{1}{\varepsilon} I((\nabla \cdot z_h^\varepsilon)(\nabla \cdot w_h)) = (\nabla v_h, \nabla w_h) \quad \text{for all } w_h \in X^h \quad (4.4.43)$$

There exists a unique solution $z_h^\varepsilon \in X^h$ of the problem (4.4.43). We define the associated pressure $\pi_h^\varepsilon \in Y^h$ as

$$\pi_h^\varepsilon(x_i^e) = -\frac{1}{\varepsilon} \nabla \cdot z_h^\varepsilon(x_i^e), \quad 1 \leq i \leq G, \quad 1 \leq e \leq E \quad (4.4.44)$$

The sequence $(z_h^\varepsilon, \pi_h^\varepsilon)$ converges in $\mathbb{X}^h \times Y^h$ to the solution $(z_h, \pi_h) \in X^h \times Y^h$ of the problem

Find $(z_h, \pi_h) \in X^h \times Y^h$ such that

$$\begin{aligned} (\nabla z_h, \nabla w_h) - (\pi_h, \nabla w_h) &= (\nabla v_h, \nabla w_h) & \text{for all } w_h \in X^h \\ (q_h, \nabla \cdot z_h) &= 0 & \text{for all } q_h \in Y^h \end{aligned} \quad (4.4.45)$$

So,

$$\|v_h - z_h\|_1 = \sup_{w_h \in X^h} \frac{(\nabla(v_h - z_h), \nabla w_h)}{\|w_h\|_1}$$

Using (4.4.45) to infer that

$$\|v_h - z_h\|_1 = \sup_{w_h \in X^h} \frac{(\pi_h, \nabla \cdot w_h)}{\|w_h\|_1}$$

The discrete inf-sup condition gives

$$\|v_h - z_h\|_1 \geq \beta_h \|\hat{z}_h\|_{Z^h} \quad (4.4.46)$$

Consider $w_h = u_h - z_h$ in (4.4.45). Since $(\pi_h, \nabla \cdot z_h) = 0$ and combine (4.4.45) and (4.2.2) to have $(\pi_h, \nabla \cdot u) = 0$ and since $(q_h, \nabla \cdot (u - v_h)) = 0$ for all $q_h \in \text{Ker} B_h^*$, we obtain

$$\|z_h - v_h\|_1^2 \leq C \|\pi_h + q_h^0\| \|u - v_h\|_1$$

The infimum over all $q_h \in \text{Ker} B_h^*$ implies

$$\|z_h - v_h\|_1^2 \leq C \|\hat{\pi}_h\|_{Z^h} \|u - v_h\|_1,$$

By the inequality (4.4.46),

$$\|z_h - v_h\|_1 \leq \frac{C}{\beta_h} \|u - v_h\|_1$$

which gives

$$\|u - z_h\|_1 \leq \left(1 + \frac{C}{\beta_h}\right) \|u - v_h\|_1$$

$z_h \in \mathbb{X}^h$ according to the problem (4.4.45). The infimum over all $z_h \in \mathbb{X}^h$ gives

$$\inf_{z_h \in \mathbb{X}^h} \|u - z_h\|_1 \leq \left(1 + \frac{C}{\beta_h}\right) \|u - v_h\|_1 \quad \text{for all } v_h \in X^h \quad (4.4.47)$$

Substituting the (4.4.47) in (4.4.42) gives the estimate (4.4.33).

Now, let us derive the pressure estimate. From the discrete inf-sup condition (4.4.10) applied to $q_h = q_h - p_h$ we have

$$\|\hat{q}_h - \hat{p}_h\|_{Z^h} \leq \frac{1}{\beta_h} \sup_{v_h \in X^h} \frac{I((q_h - p_h)\nabla \cdot v_h)}{\|v_h\|_1}. \quad (4.4.48)$$

Using (4.4.4),

$$I((q_h - p_h)\nabla \cdot v_h) = (q_h - p, \nabla \cdot v_h) + (p - p_h, \nabla \cdot v_h) \quad (4.4.49)$$

Multiplying (4.2.2) by v_h and subtracting (4.4.15)

$$\begin{aligned}
(p - p_h, \nabla \cdot v_h) &= \nu(\nabla(u - u_h), \nabla v_h) + ((u \cdot \nabla)u, v_h) - ((u_h \cdot \nabla)u_h, v_h) \\
&\quad - \frac{1}{2}((\nabla \cdot u_h), u_h \cdot v_h) + a|u|^\alpha u, v_h) - a|u_h|^\alpha u_h, v_h) \\
&\text{for all } v_h \in X^h
\end{aligned}$$

Using similar calculus as in (4.2.66)- (4.2.71), the triangle inequality and (4.4.7) to get

$$(p - p_h, \nabla \cdot v_h) \leq [\nu \|u - u_h\|_1 + 2\nu(1-r)\|u - u_h\|_1 + aC\|u - u_h\|_1] \|v_h\|_1$$

Injecting this result in (4.4.49) and using (4.4.46) yields to

$$\|\hat{q}_h - \hat{p}_h\|_{Z^h} \leq \frac{\nu(3-2r) + aC}{\beta_h} \|u - u_h\|_1 + \frac{C'}{\beta_h} \|p - q_h\|$$

Consequently

$$\begin{aligned}
\|\hat{p} - \hat{p}_h\|_{Z^h} &\leq \|\hat{p} - \hat{q}_h\|_{Z^h} + \frac{\nu(3-2r) + aC}{\beta_h} \|u - u_h\|_1 + \frac{C'}{\beta_h} \|p - q_h\| \\
&\leq \frac{\nu(3-2r) + aC}{\beta_h} \|u - u_h\|_1 + \left(1 + \frac{C'}{\beta_h}\right) \|p - q_h\|
\end{aligned}$$

which finish the proof. \square

Theorem 4.4.4. *Let $(u, p) \in X \times Y$ be the solution of (4.2.2) and $(u_h^\varepsilon, p_h^\varepsilon)$ be defined by (4.4.5) then*

$$\begin{aligned}
\|u - u_h^\varepsilon\|_1 &\leq \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_1\right) \left(1 + \frac{C}{\beta_h}\right) \inf_{v_h \in X^h} \|u - v_h\|_1 \\
&\quad + \frac{C_2}{\nu r} \inf_{q_h \in Y^h} \|p - q_h\| + \frac{C_3}{\beta_h} \varepsilon \|p_h\|
\end{aligned} \tag{4.4.50}$$

and

$$\begin{aligned}
\|\hat{p} - \hat{p}_h^\varepsilon\|_{Z^h} &\leq \frac{\nu(3-2r) + aC_1}{\beta_h} \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_2\right) \left(1 + \frac{C_3}{\beta_h}\right) \inf_{v_h \in X^h} \|u - v_h\|_1 \\
&\quad + \left(1 + \frac{C_4}{\beta_h} \left(\frac{3}{r} - 1 + \frac{a}{\nu r} C_1\right)\right) \inf_{q_h \in Y^h} \|p - q_h\| + \frac{C''}{\beta_h^2} \varepsilon \|p_h\|
\end{aligned} \tag{4.4.51}$$

Proof. For the first inequality, we use the triangle inequality

$$\|u - u_h^\varepsilon\|_1 \leq \|u - u_h\|_1 + \|u_h - u_h^\varepsilon\|_1$$

Using the inequalities (4.4.33) and (4.4.19). The infimum over $v_h \in X^h$, $q_h \in Y^h$ implies the estimate (4.4.50).

With a similar manner

$$\begin{aligned} \|\hat{p} - \hat{p}_h^\varepsilon\|_{Z^h} &\leq \|\hat{p} - \hat{p}_h\|_{Z^h} + \|\hat{p}_h - \hat{p}_h^\varepsilon\|_{Z^h} \\ &= \|\hat{p} - \hat{p}_h\|_{Z^h} + \|\hat{p}_h - \hat{p}_h^\varepsilon\| \end{aligned}$$

□

Remark 4.4.1. For ε a small positive constant, the inequality (4.4.20) show that $\|p_h\|$ is bounded.

4.5 Numerical resolution

4.5.1 Numerical implementation of optimization problem

In this subsection we study the approach called spectral projected gradient to solve the shape optimization problem which consists to minimize the discrete cost function defined by

$$\bar{\Phi}(y) = \frac{\Delta t}{2} \sum_{n=1}^N \sum_{e \in T_i} \left[\int_e \|\mathbf{u}^{n,i} - \vec{v}\|^2 + \alpha \int_e |\operatorname{curl}(\mathbf{u}^{n,i})|^2 \right]. \quad (4.5.1)$$

The used method is based on the successive projections into a closed and convex Ω_0 subset of \mathbb{R}^4 of all the points $y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$ satisfying

$$\begin{aligned} \frac{1}{4} 1.213 &\leq y_1, y_3 \leq \frac{3}{4} 1.213 \\ 0 &\leq y_2, y_4 \leq \frac{1}{4} 0.97 \end{aligned} \quad (4.5.2)$$

and for a comfortable fish passage other constraints are set up:

$$\begin{aligned} y_3 - y_1 &\geq d_1 = 0.1 \\ y_2 - y_4 &\geq d_2 = 0.05 \end{aligned} \quad (4.5.3)$$

We denote $a_1 = a_3 = \frac{1}{4} 1.213$, $b_1 = b_3 = \frac{3}{4} 1.213$, $a_2 = a_4 = 0$, $b_2 = b_4 = \frac{1}{4} 0.97$. Then the admissible set Ω_0 is defined as

$$\Omega = \{y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : \begin{aligned} &a_i \leq y_i \leq b_i, \quad i = 1, \dots, 4, \\ &y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{aligned}\} \quad (4.5.4)$$

The optimization problem writes

$$\min_{y \in \Omega_0} J(y). \quad (4.5.5)$$

The optimal position $y \in \Omega_0$ is found through the following algorithm using the computation of the cost function gradient and the projection functions.

Let $\eta = 1$, $\bar{y} \in \Omega$ and $\varepsilon > 0$ be a positive tolerance.

1. Compute $d = P_{\Omega_0}(\bar{y} - \eta \nabla J(\bar{y})) - \bar{y}$, with ∇J is the gradient of the cost function J and η is a positive constant given by
 - Consider \bar{y} the current point and \tilde{y} the previous point. Compute $x = \bar{y} - \tilde{y}$ and $y = \nabla J(\bar{y}) - \nabla J(\tilde{y})$. If $x^T y > 0$, take $\eta = \frac{x^T x}{x^T y}$; elsewhere, take η as a fixed positive value.
2. Stop when $d = 0$ (in practice, $\|d\| < \varepsilon$) and \bar{y} is a stationary point of J on Ω_0 .
3. Set $\tilde{y} = \bar{y} + \xi d$, with ξ is defined as the step size. And go to (1) with $\bar{y} = \tilde{y}$.

The value $y = P_{\Omega}(z)$ is the projection of $z \in \mathbb{R}^4$ onto Ω . It is computed by minimizing a quadratic function of the distance of z to Ω as follow

$$\min_{y \in \Omega} \frac{1}{2} \|y - z\|^2 = \min_{y \in \Omega} \frac{1}{2} z^T z - z^T y + \frac{1}{2} y^T y \quad (4.5.6)$$

which is equivalent to

$$\begin{aligned} \min_{(y_1, y_2, y_3, y_4)} \quad & \sum_{i=1}^4 \left\{ \frac{1}{2} z_i^2 - z_i y_i + \frac{1}{2} y_i^2 \right\} \\ \text{subject to} \quad & a_i \leq y_i \leq b_i, i = 1, \dots, 4 \\ & y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{aligned}$$

In order to solve the two quadratic problems, we use the Karush-Kuhn-Tucker (KKT) technique: Choose $a = a_1 = a_3$, and $b = b_1 = b_3$. We need three constraints with the optimization problem

$$\begin{aligned} \min_{(x_1, x_2)} \quad & c - (l_1 x_1 + l_2 x_2) + \frac{1}{2} (x_1^2 + x_2^2) \\ \text{subject to} \quad & x_2 \leq b, \\ & x_1 \geq a, \\ & x_2 - x_1 \geq d_1, \end{aligned}$$

where a, b, d_1, c , and $l_i, i=1,2$, are real numbers.

Consider the following matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -b \\ a \\ d_1 \end{pmatrix},$$

Then the linear constraints write

$$Ax - v = b, \quad v \geq 0,$$

where $v = (v_1, v_2, v_3)^T$ is the slack variables. The optimisation problem is solvable with unique solution due to the convexity of the cost function in \mathbb{R}^2 . The obtained solution satisfies the KKT conditions

$$\begin{aligned} v &= Ax - b \\ -l + x &= A^T \lambda, \\ v &\geq 0, \quad \lambda \geq 0, \\ v^T \lambda &= 0, \end{aligned}$$

where $l = (l_1, l_2, l_3)$, and $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is the vector of the Lagrange multipliers associated to the three constraints above. The unique optimal solution writes

$$x = l + A^T \lambda$$

which gives

$$\begin{cases} x_1 = l_1 + \lambda_2 - \lambda_3 \\ x_2 = l_2 - \lambda_1 + \lambda_3 \end{cases}$$

For computing the Lagrange multipliers λ_i , we solve a LCP problem which takes the form

$$\begin{aligned} v &= (-b + Al) + AA^T \lambda \\ v &\geq 0, \quad \lambda \geq 0 \\ v^T \lambda &= 0 \end{aligned}$$

Remark 4.5.1. *The gradient of the cost function is computed using the formula established in section 3.*

4.5.2 Numerical examples

The techniques for the optimal design, developed in the previous sections, are investigated. We verify the approach for the calculation of mean flow in a vertical slot fishways (VSF). There exist various approaches to design a VSF, whose objective is to dissipate the energy of the inlet jet and to create flow conditions in the pool that enable the fish to ascend in comfort circumstances. We realize three different tests of designs of VSF with I and L shaped baffle that are the most adopted for upstream passage of fish in river obstructions.

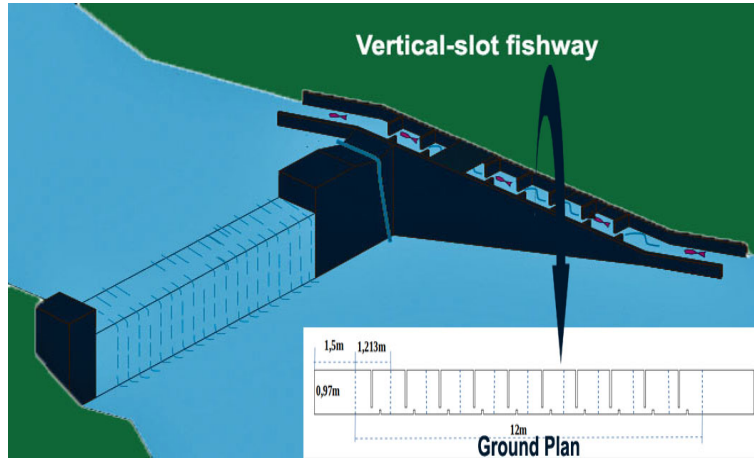


Figure 4.1: Conceptual layout of a vertical slot fishway

The structure of a fishway is consisting of a succession of ten pools (Figure 4.4). Each pool has a length of $1.213m$ and a width of $0.97m$. We also have two transition pools, one at the beginning and the other at the end of the channel with a length of $1.5m$ and a same width as the other pools.

The numerical simulation for the three tests is done with the same physical and numerical conditions. All initial and boundary conditions are taken constants. the inflow velocity is $q_1 = 0.1 m^2s^{-1}$, the target velocity is $c = 0.8 m.s^{-1}$. The vorticity parameter is $\sigma = 0$. The viscosity is 0.01 . The exterior function is $f = 0$. For fish passage with comfortable conditions, we put $d_1 = 0.1$ and $d_2 = 0.05$.

4.5.2.1 Vertical Slot Fishway (2-D "L" shaped baffle-oblique slots)

The set of Equations 4.2.31 is solved by the finite element method 4.4.5 combined with projected gradient for the resolution of the shape optimization problem 4.5.1-4.5.5. The implementation uses the FreeFem++ software.

FreeFem++ is a partial differential equation solver that has an advanced automatic mesh generator, it uses fast algorithms such as the multi-frontal method UMFPACK with several triangular finite elements. The central pool of the computational domain that is shown in Figure 4.2.

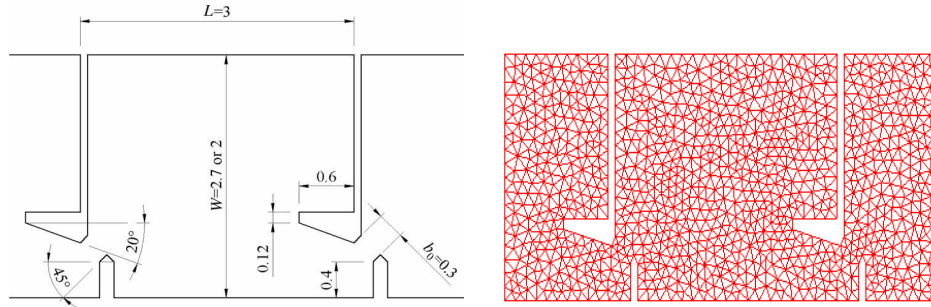


Figure 4.2: Fish pass geometry and mesh in the central pool

In Figure 4.3 calculated flow contour plots of the approximated velocity components both for the random and the optimal shape.

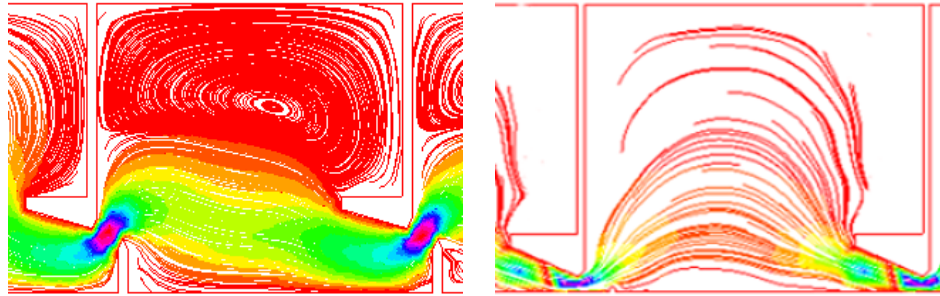


Figure 4.3: Initial (left) and optimal (right) velocities for central pool

We observe that the circulation areas near of slots are removed for the optimal shape see Figure 4.3 (right) compared to the initial random structure see Figure 4.3 (left).

4.5.2.2 Slot Fishway (3-D "I" shaped baffle-rectangular slots)

This numerical example performs a 3-D computational fluid dynamics model for VSF. We consider a 3-D channel of ten pools. The geometry of slots in each pool is rectangular as shown in the Figure 4.4.

The existence of the slots in the fishway provides a good dissipation of energy and quit areas between pools allowing to fishes to have a rest in the pool before the begin to across to an other pool (Figure 4.4). The ensuing results present the velocity for ten pools and compare the water flow in the random and the optimal shape.

As for the initial random shape (Figure 4.5), there are a circulation areas near from slots compared to the obtained optimal shape (Figure 4.6). The results are clearly observed for the central pool. Moreover, the present 3-D calculations showed that the common assumptions in VSF that the flow is 2-D is not in general valid.

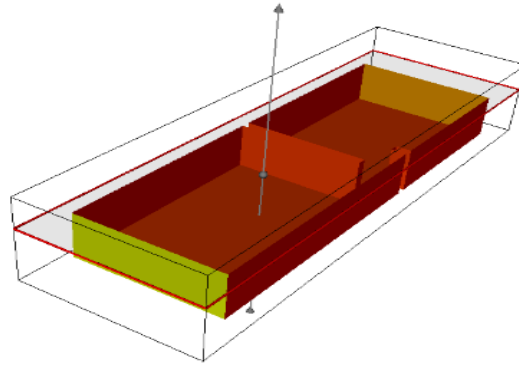


Figure 4.4: Configuration of central pool

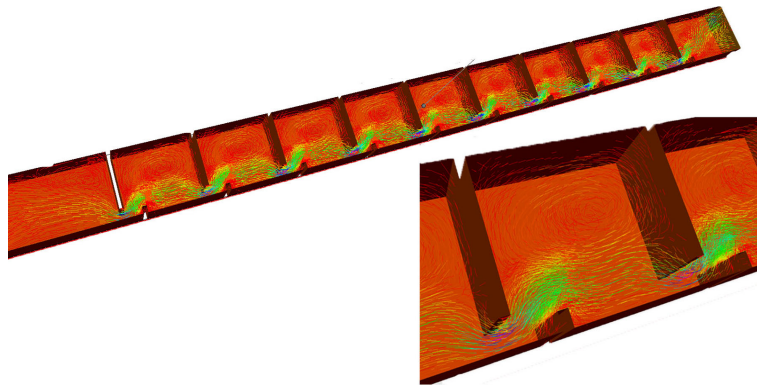


Figure 4.5: Velocity fields for non optimal ten pool configuration

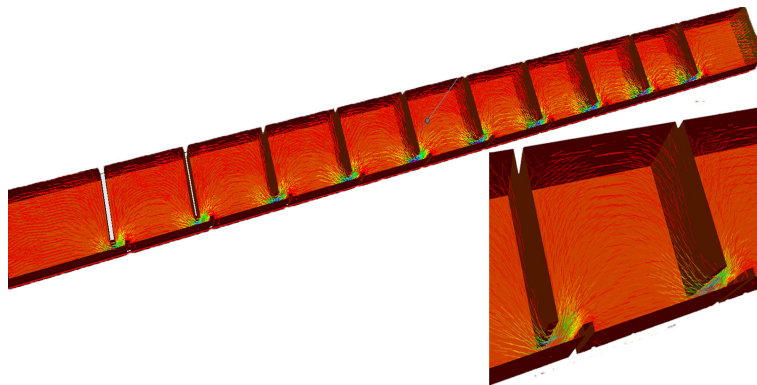


Figure 4.6: Velocity fields for optimal ten pool configuration

The corresponding cost function reduction is given in figure 4.7. We observe a decreasing of the cost function J in terms of the nature of the fish ladder geometry. As soon as the guidelines slots become optimal the objective function reached a minimum value.

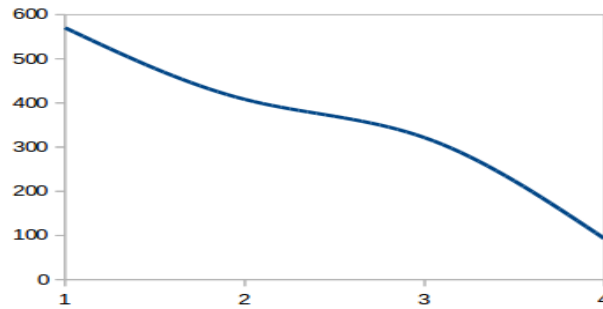


Figure 4.7: Objective function evolution

4.5.2.3 Slot Fishway (3-D "L" shaped baffle-oblique slots)

We present results for 3-D existed fishway design in nature with L shaped oblique baffle. The different dimensions for the initial structure used are (Figure 4.2(left) and Figure 4.8 (left)) slot width $b_0 = 0.3$ m, basin length $L = 3$ m, and widths $W = 2$ m. Length and width proportions are $\frac{L}{b_0} = 10$ and $\frac{W}{b_0} = 6.6$. These standard values are issued from current VSF designs [1, 9]. A triangular discretization in space is used to discretize the fluid domain Ω . The density of mesh is large near from slots compared to the other zones. Numerical simulation are done with 19000 triangles.

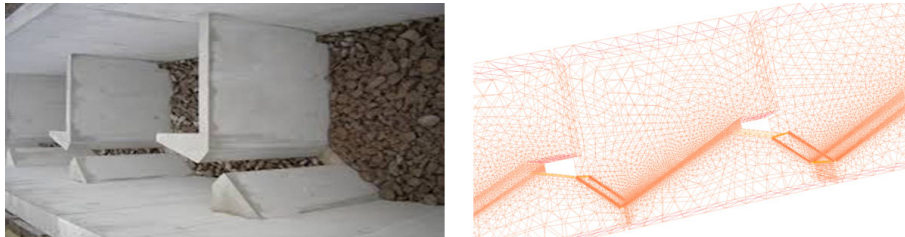


Figure 4.8: 3-D Example of Fiswhay with oblique slots and mesh in the central pool

The following results present the velocity for ten pools and compare the water flow in the initial and the optimal shape. As for the initial structure (Figure 4.9), there are a circulation areas near from slots compared to the obtained optimal shape (Figure 4.10). We observe clearly the recirculation regions flowing in opposite directions for the initial shape compared to results in the optimal shape where the recirculation areas near from slots are reduced.

The figure 4.11 illustrates the decreasing of the cost function J in terms of the nature of the fish pass structure.

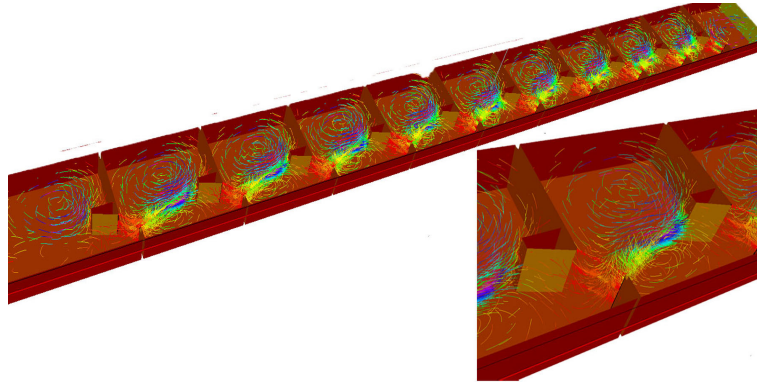


Figure 4.9: Non optimal velocity for ten pool

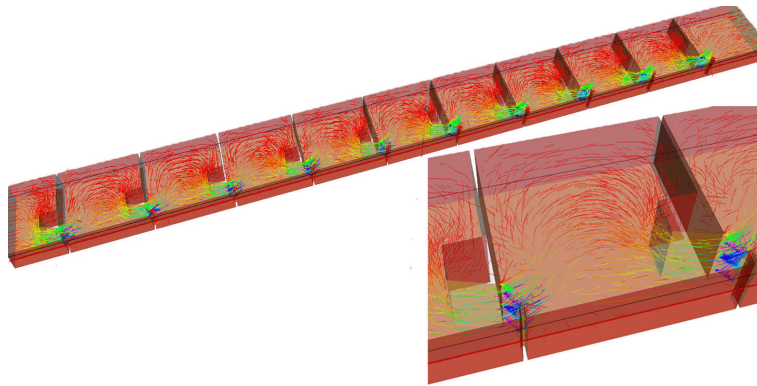


Figure 4.10: Optimal velocity for ten pool

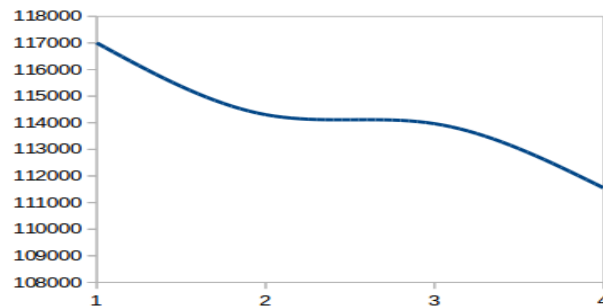


Figure 4.11: Objective function evolution

4.6 Conclusion

The design optimization is extended to the penalty porous media model and has been successfully used in hydraulic shape optimization. Based on the Navier Stokes Forchheimer model, the penalized system has been analyzed and a finite element method are considered.

Error estimates of the velocity and the pressure are derived in term of H^1 and L^2 norm respectively. 2-D and 3-D examples have been presented, and the corresponding results have demonstrated the usefulness and robustness of the approach.

Chapter 5

Hydrodynamic design optimization using non stationary porous media model

5.1 Introduction

In this chapter, we are interested in studying the non-stationary Navier Stokes Forchheimer equations

$$\left\{ \begin{array}{ll} \mathbf{u}_t - \nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + a |\mathbf{u}|^\alpha \mathbf{u} + \nabla p = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = 0 & \text{on } \Gamma_0, \\ \mathbf{u} = g & \text{on } \Gamma_1, \\ \nu \frac{\partial \mathbf{u}}{\partial \mathbf{n}} - p \mathbf{n} = 0 & \text{on } \Gamma_2, \\ \mathbf{u}|_{t=0} = \mathbf{u}_0 & \text{on } \Omega, \end{array} \right. \quad (5.1.1)$$

where Ω is an open bounded set in \mathbb{R}^d ($d = 2$ or $d = 3$) with a sufficiently smooth boundary $\partial\Omega$. The unknown functions here are $\mathbf{u}(x, t) = (u_1(x, t), u_2(x, t), u_3(x, t))$ and $p = p(x, t)$, which stand for the velocity field and the pressure of the flow, respectively. In damping term, $\alpha > 1$ and $a > 0$ are two constants. The given function $\mathbf{u}_0 = \mathbf{u}_0(x)$ is the initial velocity and the constant $\nu > 0$ represents the viscosity coefficient of the flow. The vector \mathbf{n} is the unit normal vector to the boundary $\partial\Omega = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2$, where Γ_0 is the lateral boundary, Γ_1 is inflow, top and bottom flow boundaries and Γ_2 is the outflow boundary.

The resolution of the problem (5.1.1) could be difficult numerically, caused by the incompressibility constraint $\nabla \cdot \mathbf{u} = 0$. A popular strategy to overcome this difficulty is to

relax the incompressibility term with the penalty approach. Our purpose is to use a penalty method which was introduced by Courant [23] and used by Temam [36] as an application for studying the Navier Stokes equations, to treat the Navier Stokes Forchheimer equations. The penalty method is to establish an approximation of the couple (\mathbf{u}, p) solution of (5.1.1) by $(\mathbf{u}^\varepsilon, p^\varepsilon)$ solution of the following penalized problem

$$\left\{ \begin{array}{ll} \mathbf{u}_t^\varepsilon - \nu \Delta \mathbf{u}^\varepsilon + (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{u}^\varepsilon + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{u}^\varepsilon + a |\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon + \nabla p^\varepsilon = \mathbf{f} & \text{in } \Omega, \\ \nabla \cdot \mathbf{u}^\varepsilon + \varepsilon p^\varepsilon = 0 & \text{in } \Omega, \\ \mathbf{u}^\varepsilon = 0 & \text{on } \Gamma_0, \\ \mathbf{u}^\varepsilon = g & \text{on } \Gamma_1, \\ \nu \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} - p^\varepsilon \mathbf{n} = 0 & \text{on } \Gamma_2, \\ \mathbf{u}^\varepsilon|_{t=0} = \mathbf{u}_0 & \text{in } \Omega. \end{array} \right. \quad (5.1.2)$$

In this chapter, we take into consideration the case of Dirichlet boundary conditions $u|_{\partial\Omega} = 0$. We prove the existence and uniqueness of the initial problem solution. We separate the penalized problem into two parts; the linearised problem and the nonlinear one. We derive an adjoint equation from the penalized problem and compute the shape gradient of the objective function. Finally we set up a time discretization and we employ it with a discrete finite element method to discretize the penalized problem and approximate the discrete perturbed equation solution.

The penalized problem model is considered to study an optimization problem related to fish pass structure, which consist to find an optimal shape of fishway structures by minimizing a cost function. The procedure 1 is to adopt a finite element method to solve the non-stationary Navier Stokes equations with damping and combine it with a gradient type algorithm called spectral projected gradient to solve the optimization problem in order to find a velocity close to a target velocity associated with fish swimming capabilities.

The chapter is organized as follow; The second section is devoted to the preliminaries useful for the analysis. A linearised problem is treated to show some error estimates. In the last part of this section we study a non linear problem to establish a general error priori estimates. The section 3 deals with the shape design. We define a shape optimization problem. We derive the adjoint system and compute the shape gradient of the cost function. The section 4 treat a time discretization by the use of a backward Euler scheme combined with a finite element method to approximate the penalized problem and derive an error estimate for the velocity and the pressure. Finally, in the last section, we

use a spectral projected gradient method to solve the shape optimization problem and we show some numerical results to prove the efficiency of this technique.

5.2 Mathematical model

5.2.1 Preliminaries

For the upcoming section we will study the following unsteady incompressible Navier Stokes Forchheimer equations

$$\mathbf{u}_t - \nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + a|\mathbf{u}|^\alpha \mathbf{u} + \nabla p = \mathbf{f} \text{ in } \Omega \times [0, T] \quad (5.2.1)$$

$$\nabla \cdot \mathbf{u} = 0 \text{ in } \Omega \times [0, T], \quad \mathbf{u}|_{\partial\Omega} = 0, \quad \mathbf{u}|_{t=0} = \mathbf{u}_0 \quad (5.2.2)$$

We consider the following notations¹

$$\mathbf{H} = \{ \mathbf{v} \in \mathbf{L}^2(\Omega), : \nabla \cdot \mathbf{v} = 0, \mathbf{v} \cdot \mathbf{n}|_\Gamma = 0 \}, \text{ where } \Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2$$

$$\mathbf{V} = \{ \mathbf{v} \in H_0^1(\Omega) : \nabla \cdot \mathbf{v} = 0 \}$$

$$X = (H_0^1(\Omega))^d \ (d = 2, 3), \ Y = L^2(\Omega), \ \hat{Y} = \{ q \in H^1(\Omega) \}, \ M = \left\{ q \in L^2(\Omega); \int_\Omega q dx = 0 \right\}$$

$$\hat{V}_0(\Omega) = \{ u \in (H^1(\Omega))^d | u = 0 \text{ on } \Gamma_0 \cup \Gamma_1 \}$$

$$\hat{V}_g(\Omega) = \{ u \in (H^1(\Omega))^d | u = 0 \text{ on } \Gamma_0, u = g \text{ on } \Gamma_1 \}$$

$$V_0(\Omega) = \{ u \in (H^2(\Omega))^d | u = 0 \text{ on } \Gamma_0 \cup \Gamma_1 \}$$

$$V_g(\Omega) = \{ u \in (H^2(\Omega))^d | u = 0 \text{ on } \Gamma_0, u = g \text{ on } \Gamma_1 \}$$

$$\mathbb{X} = \{ v \in (H_0^1(\Omega))^d \mid (q, \nabla \cdot \mathbf{v}) = 0, \forall q \in L^2(\Omega) \}$$

$$H^{-2}(\Omega): \text{ the dual space of } H^2(\Omega) \cap H_0^1(\Omega)$$

$$P_{\mathbf{H}}: \text{ the orthogonal projector in } \mathbf{L}^2(\Omega) \text{ onto } \mathbf{H}$$

$A = -P_{\mathbf{H}}\Delta = \tilde{\Delta}$: the Stokes operator, which is an unbounded positive self-adjoint closed operator in \mathbf{H} with domain $D(A) = \mathbf{H}^2(\Omega) \cap \mathbf{V}$. We denote also by $\|\cdot\|_k$ the norm of $\mathbf{H}^k(\Omega)$, In particular, we will use $\|\cdot\|$ to denote the norm in $\mathbf{L}^2(\Omega)$ and (\cdot, \cdot) to denote the scalar product in $\mathbf{L}^2(\Omega)$.

Now, we introduce some operators associated with the Navier Stokes equations and the perturbed problem

$$B(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \cdot \nabla) \mathbf{v}, \quad \tilde{B}(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \cdot \nabla) \mathbf{v} + \frac{1}{2}(\nabla \cdot \mathbf{u}) \mathbf{v},$$

$$b(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (B(\mathbf{u}, \mathbf{v}), \mathbf{w}), \quad \tilde{b}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (\tilde{B}(\mathbf{u}, \mathbf{v}), \mathbf{w}).$$

We observe that

$$b(\mathbf{u}, \mathbf{v}, \mathbf{v}) = 0 \quad \forall \mathbf{u} \in \mathbf{H}, \mathbf{v} \in \mathbf{H}_0^1(\Omega). \quad (5.2.3)$$

We can also verify by integration by parts that

¹ The vector functions and vector spaces will be indicated by boldface type.

$$\tilde{b}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \frac{1}{2} [b(\mathbf{u}, \mathbf{v}, \mathbf{w}) - b(\mathbf{u}, \mathbf{w}, \mathbf{v})] \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega).$$

It is easy to verify that \tilde{b} verifies

$$b(\mathbf{u}, \mathbf{v}, \mathbf{w}) = -b(\mathbf{u}, \mathbf{w}, \mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega). \quad (5.2.4)$$

Consequently, we have

$$\tilde{b}(\mathbf{u}, \mathbf{v}, \mathbf{v}) = 0 \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1(\Omega). \quad (5.2.5)$$

We have also the two following inequality which will be used in the upcoming sections

$$\tilde{b}(\mathbf{u}, \mathbf{v}, \mathbf{w}) \leq \|\mathbf{u}\|_1 \|\mathbf{v}\|_1^{\frac{1}{2}} \|\mathbf{v}\|_2^{\frac{1}{2}} \|\mathbf{w}\| \quad \forall \mathbf{v} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega), \quad \forall \mathbf{u}, \mathbf{w} \in \mathbf{H}_0^1(\Omega). \quad (5.2.6)$$

$$\tilde{b}(\mathbf{u}, \mathbf{v}, \mathbf{w}) \leq \begin{cases} \|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \|\mathbf{w}\|_1 & \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ \|\mathbf{u}\|_2 \|\mathbf{v}\| \|\mathbf{w}\|_1 & \forall \mathbf{u} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega), \quad \forall \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ \|\mathbf{u}\|_2 \|\mathbf{v}\|_1 \|\mathbf{w}\| & \forall \mathbf{u} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega), \quad \forall \mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ \|\mathbf{u}\|_1 \|\mathbf{v}\|_2 \|\mathbf{w}\| & \forall \mathbf{v} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega), \quad \forall \mathbf{u}, \mathbf{w} \in \mathbf{H}_0^1(\Omega). \end{cases} \quad (5.2.7)$$

We introduce $A_\varepsilon u = -\nu \Delta u - \frac{1}{\varepsilon} \nabla(\nabla \cdot \mathbf{u})$, which is the operator related to the penalty method. It is a positive self-adjoint operator from $\mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$ onto $\mathbf{L}^2(\Omega)$ and the powers A^β of A ($\beta \in \mathbb{R}$) are well defined and we have

$$(A^{1/2} \mathbf{u}, A^{1/2} \mathbf{v}) = (\nabla \mathbf{u}, \nabla \mathbf{v}), \quad (A_\varepsilon^{1/2} \mathbf{u}, A_\varepsilon^{1/2} \mathbf{v}) = (A^{1/2} \mathbf{u}, A^{1/2} \mathbf{v}) + \frac{1}{\varepsilon} (\operatorname{div} \mathbf{u}, \operatorname{div} \mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in X$.

For ε sufficiently small, we have

$$\begin{aligned} \|\Delta \mathbf{u}\| &\leq C \|A_\varepsilon \mathbf{u}\| \quad \forall \mathbf{u} \in \mathbf{H}^2(\Omega) \cap X, \\ \|\nabla \mathbf{u}\| &\leq C \|A_\varepsilon^{\frac{1}{2}} \mathbf{u}\| \quad \forall \mathbf{u} \in X, \\ \|A_\varepsilon^{-1} \mathbf{u}\| &\leq C \|\mathbf{u}\|_{-2} \quad \forall \mathbf{u} \in \mathbf{H}^{-2}(\Omega). \end{aligned} \quad (5.2.8)$$

for C a positive constant.

Lemma 5.2.1. *We have the following Gagliardo-Nirenberg inequalities*

$$\|\mathbf{v}\|_{L^4} \leq c \|\mathbf{v}\|^{1/2} \|A^{1/2} \mathbf{v}\|^{1/2}, \quad \|\mathbf{v}\| \leq c \|A^{1/2} \mathbf{v}\|, \quad \forall v \in X \quad (5.2.9)$$

$$\|\nabla \mathbf{v}\|_{L^4} \leq c \|A^{1/2} \mathbf{v}\|^{1/2} \|A \mathbf{v}\|^{1/2}, \quad \|A^{1/2} \mathbf{v}\| \leq c \|A \mathbf{v}\|, \quad \forall v \in D(A) \quad (5.2.10)$$

$$\|\mathbf{v}\|_{L^\infty} \leq c \|\mathbf{v}\|^{1/2} \|A \mathbf{v}\|^{1/2}, \quad \forall v \in D(A) \quad (5.2.11)$$

where c is a general positive constant depending only on Ω .

Lemma 5.2.2. *There exists a constant $c_0 > 0$ depending on Ω and such that if $c\varepsilon \leq 1$,*

$$\|A\mathbf{v}\| \leq c_0\|A_\varepsilon\mathbf{v}\|, \quad \|A^{1/2}\mathbf{v}\| \leq c_0\|A_\varepsilon^{1/2}\mathbf{v}\| \quad (5.2.12)$$

We write some lemma of Gronwall type which will be used in the following.

Lemma 5.2.3. *(Gronwall lemma) Let $y(t)$, $h(t)$, $g(t)$ and $f(t)$ be non negative functions verifying*

$$y(t) + \int_0^t h(s)ds \leq y(0) + \int_0^t (g(s)y(s) + f(s))ds \quad \forall 0 \leq t \leq T, \quad \text{with} \quad \int_0^T g(t)dt \leq M,$$

then

$$y(t) + \int_0^t h(s)ds \leq \exp(M)(y(0) + \int_0^t f(s)ds) \quad \forall 0 \leq t \leq T.$$

Lemma 5.2.4. *(discrete Gronwall lemma). Let y^n , h^n , g^n , f^n be nonnegative series verifying*

$$y^m + k \sum_{n=0}^m h^n \leq k \sum_{n=0}^m (g^n y^n + f^n), \quad \text{with} \quad k \sum_{n=0}^{T/k} g^n \leq M \quad \forall 0 \leq m \leq T/k.$$

Suppose $kg^n < 1 \quad \forall n$ and let $\sigma = \max_{0 \leq n \leq T/R} (1 - kg^n)^{-1}$, then

$$y^m + k \sum_{n=1}^m h^n \leq \exp(\sigma M)(B + k \sum_{n=0}^m f^n), \quad \forall m \leq T/k$$

We give the property of monotonicity for any mapping F given by $F : \mathbf{x} \mapsto |\mathbf{x}|^\alpha \mathbf{x}$

$$(|\mathbf{u}|^\alpha \mathbf{u} - |\mathbf{v}|^\alpha \mathbf{v}, \mathbf{u} - \mathbf{v}) \geq 0 \quad (5.2.13)$$

The following Sobolev inequality is useful to deal with the nonlinear form F .

$$\| \cdot \|_p \leq \| \nabla \mathbf{v} \|, \quad 1 \leq p \leq 6 \quad (5.2.14)$$

It is Known that if $\partial\Omega$ is smooth enough, we have for any $v \in \mathbf{V} \cap \mathbf{H}^2(\Omega)$

$$\|\mathbf{v}\|_2 \leq \|\tilde{\Delta}\mathbf{v}\| \quad (5.2.15)$$

We define the following functional $G(\mathbf{u}) = PF(\mathbf{u})$. In the coming result, we derive an estimate for the velocity's second derivative, the velocity's derivative with respect to time and the pressure.

Lemma 5.2.5. *We assume that the given data \mathbf{u}_0 and \mathbf{f} satisfy the following regularity*

$$\mathbf{u}_0 \in \mathbf{V} \cap \mathbf{H}^2(\Omega), \quad \mathbf{f}, \mathbf{f}_t \in \mathbf{L}^2(\Omega) \quad (S1)$$

Then there exists $T_1 \leq T$ such that the solution of the system (5.2.1)-(5.2.2) satisfies:

$$\|\mathbf{u}(t)\|_2 + \|\mathbf{u}_t(t)\| + \|p(t)\|_1 \leq C, \quad \forall t \in [0, T_1]. \quad (5.2.16)$$

Proof. Take the inner product of the equation (5.2.1) with $\tilde{\Delta}\mathbf{u}$ and integrate all to get

$$\frac{1}{2} \frac{d}{dt} \|\nabla \mathbf{u}\|^2 + \gamma \|\tilde{\Delta} \mathbf{u}\|^2 = -b(\mathbf{u}, \mathbf{u}, \tilde{\Delta} \mathbf{u}) - a(|\mathbf{u}|^\alpha \mathbf{u}, \tilde{\Delta} \mathbf{u}) + (\mathbf{f}, \tilde{\Delta} \mathbf{u}). \quad (5.2.17)$$

Combining the inequalities (5.2.8)-(5.2.11) with Young inequality, we have

$$b(\mathbf{u}, \mathbf{u}, \tilde{\Delta} \mathbf{u}) \leq \frac{\gamma}{6} \|\tilde{\Delta} \mathbf{u}\|^2 + C' \|\nabla \mathbf{u}\|^6, \quad (5.2.18)$$

$$a(|\mathbf{u}|^\alpha \mathbf{u}, \tilde{\Delta} \mathbf{u}) \leq \frac{\gamma}{6} \|\tilde{\Delta} \mathbf{u}\|^2 + C \|\nabla \mathbf{u}\|^{2(\alpha+1)}. \quad (5.2.19)$$

Substituting (5.2.18) and (5.2.19) into (5.2.17) and using the fact that

$$(\mathbf{f}, \tilde{\Delta} \mathbf{u}) \leq \frac{\gamma}{6} \|\tilde{\Delta} \mathbf{u}\|^2 + \frac{3}{2\gamma} \|\mathbf{f}\|^2,$$

we obtain

$$\frac{1}{2} \frac{d}{dt} \|\nabla \mathbf{u}\|^2 + \frac{\gamma}{2} \|\tilde{\Delta} \mathbf{u}\|^2 \leq C' \|\nabla \mathbf{u}\|^6 + \frac{3}{2\gamma} \|\mathbf{f}\|^2. \quad (5.2.20)$$

In particular, It is equivalent to solve the following equation (by assuming $C' \geq 1$):

$$\frac{d}{dt} y(t) \leq C' y(t)^3, \quad \text{with } y(t) = \|\nabla \mathbf{u}(t)\|^2 + C_1, \quad y(0) = \|\nabla \mathbf{u}(0)\|^2 + C_1$$

and $C_1 = \frac{3}{2\gamma} \sup_{0 \leq t \leq T} \|\mathbf{f}\|^2$. The equation above has a solution in $[0, T_1]$ where $T_1 \leq \frac{1}{2C'y(0)^2} = \frac{1}{C_0}$ and

$$y(t) \leq \sqrt{2}y(0), \quad t \in [0, T_1].$$

So, for all $t \in [0, T_1]$ with $T_1 = \min\{T_1, \frac{1}{C_0}\}$, we have

$$\|\nabla \mathbf{u}(t)\|^2 \leq \sqrt{2}(\|\nabla \mathbf{u}(0)\|^2 + C_1) := C_3, \quad (5.2.21)$$

Consequently, using (5.2.20) and (5.2.21), to infer that

$$\sup_{0 \leq t \leq T_1} \|\nabla \mathbf{u}(t)\|^2 + \gamma \int_0^{T_1} \|\tilde{\Delta} \mathbf{u}(t)\|^2 dt \leq C, \quad \forall t \in [0, T_1]. \quad (5.2.22)$$

Deriving one time the equation (5.2.1) with respect to t , we obtain

$$\mathbf{u}_{tt} - \gamma \Delta \mathbf{u}_t + aF'(\mathbf{u})\mathbf{u}_t + B(\mathbf{u}_t, \mathbf{u}) + B(\mathbf{u}, \mathbf{u}_t) + \nabla p_t = \mathbf{f}_t \quad (5.2.23)$$

Multiplying the relation (5.2.23) with \mathbf{u}_t , we get

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{u}_t\|^2 + \gamma \|\nabla \mathbf{u}_t\|^2 = -a(F'(\mathbf{u})\mathbf{u}_t, \mathbf{u}_t) - b(\mathbf{u}_t, \mathbf{u}, \mathbf{u}_t) + (\mathbf{f}_t, \mathbf{u}_t).$$

Since $(F'(\mathbf{u})\mathbf{u}_t) \cdot \mathbf{u}_t$ is positive definite, using the inequalities (5.2.8), Hölder's and Sobolev's inequalities yields to

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\mathbf{u}_t\|^2 + \gamma \|\nabla \mathbf{u}_t\|^2 &\leq c \|\mathbf{u}_t\| \|\mathbf{u}\|_1^{\frac{1}{2}} \|\mathbf{u}\|_2^{\frac{1}{2}} \|\nabla \mathbf{u}_t\| + \frac{1}{2} \|\mathbf{f}_t\|^2, \\ &\leq \frac{\gamma}{2} \|\nabla \mathbf{u}_t\|^2 + C(\|\mathbf{u}\|_1 \|\mathbf{u}\|_2) \|\mathbf{u}_t\|^2 + \frac{1}{2} \|\mathbf{f}_t\|^2. \end{aligned}$$

Using Cauchy-Schwarz inequality, estimates (5.2.15) and (5.2.22), to have:

$$\int_0^{T_1} \|\mathbf{u}(s)\|_1 \|\mathbf{u}(s)\|_2 ds \leq \left(\int_0^{T_1} \|\mathbf{u}(s)\|_1^2 ds \right)^{\frac{1}{2}} \left(\int_0^{T_1} \|\mathbf{u}(s)\|_2^2 ds \right)^{\frac{1}{2}} \leq C. \quad (5.2.24)$$

Applying Gronwall Lemma to (5.2.24), we obtain:

$$\|\mathbf{u}_t\|^2 + \gamma \int_0^{T_1} \|\nabla \mathbf{u}_t(s)\|^2 ds \leq C, \quad \forall t \in [0, T_1]. \quad (5.2.25)$$

Multiplying the equations (5.2.1) by $\tilde{\Delta} \mathbf{u}$, we get

$$\begin{aligned} \gamma \|\tilde{\Delta} \mathbf{u}\|^2 &= -(\mathbf{u}_t, \tilde{\Delta} \mathbf{u}) - a(|\mathbf{u}|^\alpha \mathbf{u}, \tilde{\Delta} \mathbf{u}) - b(\mathbf{u}, \mathbf{u}, \tilde{\Delta} \mathbf{u}) - (\mathbf{f}, \tilde{\Delta} \mathbf{u}) \\ &\leq \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}\|^2 + 2\gamma \|\mathbf{u}_t\|^2 + a(|\mathbf{u}|^\alpha \mathbf{u}, \tilde{\Delta} \mathbf{u}) + |b(\mathbf{u}, \mathbf{u}, \tilde{\Delta} \mathbf{u})| + \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}\|^2 + 2\gamma \|f\|^2. \end{aligned}$$

Similarly as in (5.2.19) and (5.2.18) we control the nonlinear terms. Using (5.2.21), (5.2.25) and (5.2.15), we infer that

$$\|\mathbf{u}\|_2 \leq C, \quad \forall t \in [0, T_1]. \quad (5.2.26)$$

For the pressure estimate, we use the equation (5.2.1) to have

$$\|\nabla p\| \leq (\|\mathbf{u}_t\| + \gamma \|\nabla \mathbf{u}\| + a\|\mathbf{u}\|^{\alpha+1} + \|B(\mathbf{u}, \mathbf{u})\|). \quad (5.2.27)$$

Using the Hölder's and the inequalities (5.2.8), the estimates (5.2.21) and (5.2.26), since $\mathbf{H}^1(\Omega) \hookrightarrow \mathbf{L}^4(\Omega)$, we obtain :

$$\begin{aligned} \|B(\mathbf{u}, \mathbf{u})\|^2 &\leq C \|\mathbf{u}\|_{\alpha+2} \|\nabla \mathbf{u}\|_{\frac{2(\alpha+2)}{\alpha}} \\ &\leq C \|\mathbf{u}\|_{\alpha+2} \|\Delta \mathbf{u}\|_2^{\left(\frac{10-\alpha}{\alpha+8}\right)} \|\mathbf{u}\|_{\alpha+2}^{2\left(\frac{\alpha-1}{\alpha+8}\right)} \\ &\leq C \|\Delta \mathbf{u}\|_2^{2\left(\frac{10-\alpha}{\alpha+8}\right)} \|\mathbf{u}\|_{\alpha+2}^{3\left(\frac{\alpha+2}{\alpha+8}\right)} \leq C. \end{aligned}$$

Substituting the inequalities (5.2.21), (5.2.25) and (5.2.26) in (5.2.27), we obtain the pressure estimate. \square

In the case where the initial data $\mathbf{u}(0)$ and $\mathbf{f}(0)$ satisfy some nonlocal compatibility conditions, we take $t_0 > 0$ and we assume that we have an initial data (\mathbf{u}_0, p_0) such that

$$\|\mathbf{u}_0 - \mathbf{u}(t_0)\| \leq Ck^2, \quad \|\nabla(u_0 - \mathbf{u}(t_0))\| + \|\nabla(p_0 - p(t_0))\| \leq Ck.$$

Lemma 5.2.6. *Under the same assumptions of the lemma 5.2.5, we suppose in addition that*

$$\mathbf{f}_t, \mathbf{f}_{tt} \in C([0, T], \underline{L}^2(\Omega)) \quad (\text{S3})$$

Then, the solution of the system (5.2.1) satisfies

$$\|\mathbf{u}_t(t)\|_2^2 + \|\nabla p_t\| + \int_{t_0}^t (\|\mathbf{u}_{tt}(s)\|_2^2 + \|p_{tt}(s)\|_1^2) ds \leq C, \quad \forall t \in [t_0, T_1]. \quad (5.2.28)$$

Proof. Deriving one time the equation (5.2.1) with respect to t , to have

$$\mathbf{u}_{tt} - \gamma \tilde{\Delta} \mathbf{u}_t + aG'(\mathbf{u})\mathbf{u}_t + B(\mathbf{u}_t, \mathbf{u}) + B(\mathbf{u}, \mathbf{u}_t) = \mathbf{f}_t \quad (5.2.29)$$

Multiplying the equation (5.2.29) by $\tilde{\Delta} \mathbf{u}_t$, we obtain:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla \mathbf{u}_t\|^2 + \gamma \|\tilde{\Delta} \mathbf{u}_t\|^2 &= -(\mathbf{f}_t, \tilde{\Delta} \mathbf{u}_t) + b(\mathbf{u}_t, \mathbf{u}, \tilde{\Delta} \mathbf{u}_t) + b(\mathbf{u}, \mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t) \\ &\quad + a(F'(\mathbf{u})\mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t) \\ &\leq \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}_t\|^2 + 2\|\mathbf{f}_t\|^2 + |b(\mathbf{u}_t, \mathbf{u}, \tilde{\Delta} \mathbf{u}_t)| + |b(\mathbf{u}, \mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \\ &\quad + a|(F'(\mathbf{u})\mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \end{aligned}$$

With a similar manner as in lemma 5.2.5 we write

$$|b(\mathbf{u}_t, \mathbf{u}, \tilde{\Delta} \mathbf{u}_t)| \leq \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}_t\|^2 + C\|\mathbf{u}\|_2^2 \|\nabla \mathbf{u}_t\|^2 \quad (5.2.30)$$

$$|b(\mathbf{u}, \mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \leq \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}_t\|^2 + C\|\mathbf{u}\|_2^2 \|\nabla \mathbf{u}_t\|^2. \quad (5.2.31)$$

Next,

$$a|(F'(\mathbf{u})\mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \leq C(\alpha + 1)\|\mathbf{u}\|_\infty^\alpha \|\mathbf{u}_t\| \|\tilde{\Delta} \mathbf{u}_t\| \leq \frac{\gamma}{8} \|\tilde{\Delta} \mathbf{u}_t\|^2 + C\|\mathbf{u}\|_\infty^{2\alpha} \|\nabla \mathbf{u}_t\|^2 \quad (5.2.32)$$

Combining the previous estimates together and using the inequality (5.2.14), (5.2.21) and (5.2.26), to obtain

$$\frac{d}{dt} \|\nabla \mathbf{u}_t\|^2 + \gamma \|\tilde{\Delta} \mathbf{u}_t\|^2 \leq 4\|\mathbf{f}_t\|^2 + C\|\nabla \mathbf{u}_t\|^2. \quad (5.2.33)$$

Thanks to Gronwall's inequality, we infer that

$$\sup_{t_0 \leq t \leq T_1} \|\nabla \mathbf{u}_t(t)\|^2 + \gamma \int_{t_0}^t \|\tilde{\Delta} \mathbf{u}_t(s)\|^2 ds \leq C, \quad \forall t \in [t_0, T_1]. \quad (5.2.34)$$

Now, differentiating the equation (5.2.23) with respect to t , we have

$$\mathbf{u}_{ttt} - \gamma \Delta \mathbf{u}_{tt} + aF''(\mathbf{u})|\mathbf{u}_t|^2 + aF'(\mathbf{u})\mathbf{u}_{tt} + 2B(\mathbf{u}_t, \mathbf{u}_t) + B(\mathbf{u}_{tt}, \mathbf{u}) + B(\mathbf{u}, \mathbf{u}_{tt}) + \nabla p_{tt} = \mathbf{f}_{tt}. \quad (5.2.35)$$

Taking the inner product of the relation (5.2.35) with \mathbf{u}_{tt} , to get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\mathbf{u}_{tt}\|^2 + \gamma \|\nabla \mathbf{u}_{tt}\|^2 &= -a(F''(\mathbf{u})|\mathbf{u}_t|^2, \mathbf{u}_{tt}) - a(F'(\mathbf{u})\mathbf{u}_{tt}, \mathbf{u}_{tt}) \\ &\quad - 2b(\mathbf{u}_t, \mathbf{u}_t, \mathbf{u}_{tt}) - b(\mathbf{u}_{tt}, \mathbf{u}, \mathbf{u}_{tt}) + (\mathbf{f}_{tt}, \mathbf{u}_{tt}). \end{aligned}$$

Since $(F'(\mathbf{u})\mathbf{u}_{tt}) \cdot \mathbf{u}_{tt}$ is positive definite, using Young's inequality, we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\mathbf{u}_{tt}\|^2 + \gamma \|\nabla \mathbf{u}_{tt}\|^2 &\leq \frac{1}{2} \|\mathbf{f}_{tt}\|^2 + |\beta(F''(\mathbf{u})|\mathbf{u}_t|^2, \mathbf{u}_{tt})| + 2|b(\mathbf{u}_t, \mathbf{u}_t, \mathbf{u}_{tt})| \\ &\quad + |b(\mathbf{u}_{tt}, \mathbf{u}, \mathbf{u}_{tt})|. \end{aligned}$$

Thanks to (5.2.14), (5.2.21) and (5.2.34), since $\mathbf{H}^1(\Omega) \hookrightarrow \mathbf{L}^4(\Omega)$,

$$\begin{aligned} |(F''(\mathbf{u})|\mathbf{u}_t|^2, \mathbf{u}_{tt})| &\leq \alpha(\alpha+1)(|\mathbf{u}|^{\alpha-1}|\mathbf{u}_t|^2, |\mathbf{u}_{tt}|) \\ &\leq \alpha(\alpha+1)\|\mathbf{u}\|_{\infty}^{\alpha-1}\|\mathbf{u}_t\|_4^2\|\mathbf{u}_{tt}\| \\ &\leq C\|\mathbf{u}_{tt}\|^2 + C\|\mathbf{u}\|_{\infty}^{2(\alpha-1)}\|\nabla \mathbf{u}_t\|^4 \leq C\|\mathbf{u}_{tt}\|^2 + C. \end{aligned} \quad (5.2.36)$$

Using Young, Hölder inequalities, the inequality (5.2.8)- (5.2.11) and estimate (5.2.16), we have

$$|b(\mathbf{u}_{tt}, \mathbf{u}, \mathbf{u}_{tt})| \leq \|\mathbf{u}_{tt}\|_1 \|\mathbf{u}\|_2 \|\mathbf{u}_{tt}\| \leq \frac{\gamma}{2} \|\nabla \mathbf{u}_{tt}\|^2 + C\|\mathbf{u}_{tt}\|^2. \quad (5.2.37)$$

$$|b(\mathbf{u}_t, \mathbf{u}_t, \mathbf{u}_{tt})| \leq \|\mathbf{u}_t\|_2 \|\mathbf{u}_t\|_1 \|\mathbf{u}_{tt}\| \leq C\|\nabla \mathbf{u}_t\|^2 + \|\mathbf{u}_t\|_2^2 \|\mathbf{u}_{tt}\|^2. \quad (5.2.38)$$

Replacing (5.2.36)-(5.2.38) in (5.2.36), using (5.2.34) and applying Gronwall Lemma yield

$$\|\mathbf{u}_{tt}\|^2 + \gamma \int_{t_0}^{T_1} \|\nabla \mathbf{u}_{tt}(s)\|^2 ds \leq C. \quad (5.2.39)$$

Multiplying the equations (5.2.1) with $\tilde{\Delta} \mathbf{u}_t$, we obtain:

$$\begin{aligned} \gamma \|\tilde{\Delta} \mathbf{u}_t\|^2 &= (\mathbf{u}_{tt}, \tilde{\Delta} \mathbf{u}_t) - (\mathbf{f}_t, \tilde{\Delta} \mathbf{u}_t) - a(F'(\mathbf{u})\mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t) - b(\mathbf{u}_t, \mathbf{u}, \tilde{\Delta} \mathbf{u}_t) \\ &\quad - b(\mathbf{u}, \mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t) \\ &\leq \frac{5\gamma}{2} \|\mathbf{u}_{tt}\|^2 + \frac{\gamma}{5} \|\tilde{\Delta} \mathbf{u}_t\|^2 + \frac{5\gamma}{2} \|\mathbf{f}_t\|^2 + a|(F'(\mathbf{u})\mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \\ &\quad + |b(\mathbf{u}_t, \mathbf{u}, \tilde{\Delta} \mathbf{u}_t)| + |b(\mathbf{u}, \mathbf{u}_t, \tilde{\Delta} \mathbf{u}_t)| \end{aligned}$$

For bounding the nonlinear terms we apply a similar estimations as in (5.2.30)-(5.2.32), thanks to (5.2.15), (5.2.39) and (5.2.34), we deduct

$$\|\mathbf{u}_t\|_2 \leq C, \quad \forall t \in [t_0, T_1]. \quad (5.2.40)$$

We can obtain the remaining estimates in (5.2.28) by following same arguments of Lemma 5.2.5. Notice that (5.2.34) is an analogue of $\int_{t_0}^t \|\tilde{\Delta} \mathbf{u}_{tt}(s)\|^2 ds$ for \mathbf{u}_{tt} instead of \mathbf{u}_t . \square

Moreover, if we assume that

$$t\mathbf{f}_t \in L^2(0, T_1; \mathbf{H}^1(\Omega)), \quad (\text{S2})$$

We obtain the pressure estimate

$$tp_t \in L^2(0, T_1; \mathbf{H}^1(\Omega)) \quad (5.2.41)$$

In this section we focus on the existence, the uniqueness and the regularity for the penalized problem associated to porous media model (5.2.1)-(5.2.2).

The associated penalized problem to the equations (5.2.1)-(5.2.2) writes

$$\mathbf{u}_t^\varepsilon - \nu \Delta \mathbf{u}^\varepsilon + \tilde{B}(\mathbf{u}^\varepsilon, \mathbf{u}^\varepsilon) + a|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon + \nabla p^\varepsilon = \mathbf{f}, \text{ in } \Omega \times [0, T], \quad (5.2.42)$$

$$\nabla \cdot \mathbf{u}^\varepsilon + \varepsilon p^\varepsilon = 0, \text{ in } \Omega \times [0, T], \quad \mathbf{u}^\varepsilon|_{\partial\Omega} = 0, \quad \mathbf{u}^\varepsilon|_{t=0} = \mathbf{u}_0, \quad (5.2.43)$$

where $\tilde{B}(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \cdot \nabla) \mathbf{v} + \frac{1}{2}(\nabla \cdot \mathbf{u}) \mathbf{v}$ is the modified bilinear term.

Using the operator A_ε , we can reformulate the perturbed system (5.2.1)-(5.2.2) as

$$\mathbf{u}_t^\varepsilon + A_\varepsilon \mathbf{u}^\varepsilon + \tilde{B}(\mathbf{u}^\varepsilon, \mathbf{u}^\varepsilon) + a|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon = \mathbf{f}, \quad (5.2.44)$$

$$\mathbf{u}^\varepsilon|_{\partial\Omega} = 0, \quad \mathbf{u}^\varepsilon|_{t=0} = \mathbf{u}_0 \quad (5.2.45)$$

Lemma 5.2.7. *Under the assumption (S1), there exists $T_2 < T$ and a positive constant C independent of ε such that*

$$\|\mathbf{u}^\varepsilon\|^2 + \frac{\nu}{2} \int_0^t \|\mathbf{u}^\varepsilon\|_1^2 ds + a \int_0^t \|\mathbf{u}^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} ds + \varepsilon \int_0^t \|p^\varepsilon\|^2 ds \leq C, \quad \forall t \in [0, T_2]. \quad (5.2.46)$$

Proof. We use the inner product of (5.2.42) with \mathbf{u}^ε and (5.2.43) with p^ε to obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\mathbf{u}^\varepsilon\|^2 + \nu \|\nabla \mathbf{u}^\varepsilon\|^2 + a \|\mathbf{u}^\varepsilon\|_{L^{\alpha+2}}^{\alpha+2} + \varepsilon \|p^\varepsilon\|^2 &= (\mathbf{f}, \mathbf{u}^\varepsilon) \\ &\leq \frac{1}{2\nu} \|\mathbf{f}\|^2 + \frac{\nu}{2} \|\nabla \mathbf{u}^\varepsilon\|^2 \end{aligned}$$

By integrating with respect to t , we get the inequality (5.2.46). \square

Theorem 5.2.1. *Consider $\mathbf{u}, v \in X$ such that (5.2.46) holds for u and v , then*

$$\|F(\mathbf{u}) - F(\mathbf{v})\| \leq C \|\mathbf{u} - \mathbf{v}\| \quad (5.2.47)$$

For the upcoming sections we work in $[0, T_0]$ with $T_0 = \min(T_1, T_2)$. For simplicity we denote by $\|\mathbf{v}\|_{L^p(X)}$ the norm $(\int_0^{T_0} \|\mathbf{v}\|_X^p dt)^{1/p}$ in $L^p(0, T_0; X)$. C is the generic constant than can take different values in different places.

5.3 Error Analysis

The main result of this section is stated in Theorem 5.3.2. The proof is split into two steps: the error of the linear case is given in the first step and the second one is related to the error behavior for the fully nonlinear problem.

5.3.1 Linearised problem

The results of this subsection will be used in the next subsection to complete the analysis of the nonlinear Navier Stokes equations.

We consider the linearised Navier Stokes equations at $\mathbf{u} = 0$:

$$\begin{aligned} \mathbf{u}_t - \nu \Delta \mathbf{u} + \nabla p &= \mathbf{f} \\ \nabla \cdot \mathbf{u} &= 0, \quad \mathbf{u}(0) = \mathbf{u}_0 \end{aligned} \tag{5.3.1}$$

The associated perturbed problem to the system (5.3.1) takes the form

$$\begin{aligned} \mathbf{u}^\varepsilon - \nu \Delta \mathbf{u}^\varepsilon + \nabla p^\varepsilon &= \mathbf{f} \\ \nabla \cdot \mathbf{u}^\varepsilon + \varepsilon p^\varepsilon &= 0, \quad \mathbf{u}^\varepsilon(0) = \mathbf{u}_0 \end{aligned} \tag{5.3.2}$$

Let $\mathbf{e} = \mathbf{u} - \mathbf{u}^\varepsilon$ and $q = p - p^\varepsilon$. Subtracting (5.3.2) from (5.3.1) to get

$$\mathbf{e}_t - \nu \Delta \mathbf{e} + \nabla q = 0 \tag{5.3.3}$$

$$\nabla \cdot \mathbf{e} + \varepsilon q = \varepsilon p, \quad \mathbf{e}(0) = 0 \tag{5.3.4}$$

The following result summarizes the error relating to the linearized problem

Lemma 5.3.1. *Let assume (S1), we have*

$$\|\mathbf{e}\|_{L^\infty(\mathbf{L}^2)} + \sqrt{\nu} \|\mathbf{e}\|_{L^2(\mathbf{H}^1)} + \sqrt{\varepsilon} \|q\|_{L^2(\mathbf{L}^2)} \leq C\sqrt{\varepsilon} \tag{5.3.5}$$

$$\|\mathbf{e}\|_{L^2(\mathbf{L}^2)} \leq C\varepsilon \tag{5.3.6}$$

Proof. Multiplying (5.3.3) by \mathbf{e} and (5.3.4) by q and summing up, we obtain

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{e}\|^2 + \nu \|\nabla \mathbf{e}\|^2 + \varepsilon \|q\|^2 = \varepsilon(p, q) \leq \frac{\varepsilon}{2} \|q\|^2 + \frac{\varepsilon}{2} \|p\|^2$$

We integrate the previous inequality from 0 to $t \leq T_0$, and we use the fact that $\mathbf{e}(0) = 0$ and lemma 5.2.5 to have

$$\|\mathbf{e}\|_{L^\infty(\mathbf{L}^2)}^2 + \nu\|\mathbf{e}\|_{L^2(\mathbf{H}^1)}^2 + \varepsilon\|q\|_{L^2(\mathbf{L}^2)}^2 \leq C\varepsilon$$

which is equivalent to (5.3.5).

To deduce the assertion (5.3.6), we need to introduce the auxiliary problem: for $0 < t \leq T_0$ we define (\mathbf{w}, ϕ)

$$\begin{aligned} \mathbf{w}_s + \nu\Delta\mathbf{w} + \nabla\phi &= \mathbf{e}(s) \quad \forall 0 < s \leq t, \\ \nabla\cdot\mathbf{w} &= 0, \quad \mathbf{w}(t) = 0. \end{aligned} \tag{5.3.7}$$

Firstly, we derive the following inequality

$$\nu\|\mathbf{w}\|_{L^2(\mathbf{H}^2)} + \|\nabla\phi\|_{L^2(\mathbf{L}^2)} \leq C\|\mathbf{e}\|_{L^2(\mathbf{L}^2)} \tag{5.3.8}$$

For this aim, we use the inner product of (5.3.7) with $A\mathbf{w}$, and we integrate from 0 to t , to obtain

$$\nu\|\mathbf{w}\|_{L^2(\mathbf{H}^2)} + \|\nabla\mathbf{w}(0)\|_{L^2(\mathbf{L}^2)} \leq C\|\mathbf{e}\|_{L^2(\mathbf{L}^2)}$$

Using the projection operator $P_{\mathbf{H}}$ on (5.3.7) to get

$$\|\mathbf{w}_s\|_{L^2(\mathbf{L}^2)} \leq C\|\mathbf{e}\|_{L^2(\mathbf{L}^2)}$$

We use the equation (5.3.7) again to get

$$\|\nabla\phi\|_{L^2(\mathbf{L}^2)} \leq C\|\mathbf{e}\|_{L^2(\mathbf{L}^2)}$$

This latter inequality completes the demonstration of (5.3.8).

We now multiply the equation (5.3.7) by $\mathbf{e}(s)$, using (5.3.3) and fact $\nabla\cdot\mathbf{w} = 0$ we obtain

$$\begin{aligned} \|\mathbf{e}\|^2 &= (\mathbf{w}_s, \mathbf{e}) + \nu(\Delta\mathbf{w}, \mathbf{e}) + (\nabla\phi, \mathbf{e}) \\ &= \frac{d}{ds}(\mathbf{w}, \mathbf{e}) - (\mathbf{e}_s, \mathbf{w}) - \nu(\nabla\mathbf{e}, \nabla\mathbf{w}) + (\nabla\phi, \mathbf{e}) \\ &= \frac{d}{ds}(\mathbf{w}, \mathbf{e}) + (\nabla q, \mathbf{w}) - (\phi, \nabla\cdot\mathbf{e}) = \frac{d}{ds}(\mathbf{w}, \mathbf{e}) - \varepsilon(\phi, p^\varepsilon) \end{aligned}$$

We integrate from 0 to t , and the fact $\mathbf{w}(t) = \mathbf{e}(0) = 0$ gives

$$\int_0^t \|\mathbf{e}\|^2 ds = - \int_0^t \varepsilon(\phi, p^\varepsilon) ds \leq \delta\|\phi\|_{L^2(\mathbf{L}^2)}^2 + C_\delta\varepsilon^2\|p^\varepsilon\|_{L^2(\mathbf{L}^2)}^2$$

with C_δ is a positive constant depending only of δ . Applying (5.3.8) and choosing δ small enough to have

$$\int_0^t \|\mathbf{e}\|^2 ds \leq C\varepsilon^2\|p^\varepsilon\|_{L^2(\mathbf{L}^2)} \leq C\varepsilon^2 \quad \forall t \in [0, T_0].$$

□

The following result is necessary for the subsequent investigations.

Lemma 5.3.2. *Provided (S1) and (S2) are satisfied, we have*

$$\int_0^t s^2 \|p_t^\varepsilon\|^2 ds \leq C \quad \forall t \in [0, T_0].$$

Proof. We derive one time with respect to t on (5.3.3) and (5.3.4), we get

$$\mathbf{e}_{tt} - \nu \Delta \mathbf{e}_t + \nabla q_t = 0 \quad (5.3.9)$$

$$\nabla \cdot \mathbf{e}_t + \varepsilon q_t = \varepsilon p_t, \quad \mathbf{e}(0) = 0 \quad (5.3.10)$$

Multiplying (5.3.3) by $t\mathbf{e}_t$ and (5.3.10) by tq to get

$$\begin{aligned} t\|\mathbf{e}\|^2 + \frac{\nu}{2} \frac{d}{dt} t \|\nabla \mathbf{e}\|^2 + \frac{\varepsilon}{2} \frac{d}{dt} t \|q\|^2 &= \frac{\nu}{2} \|\nabla \mathbf{e}\|^2 + \frac{\varepsilon}{2} \|q\|^2 + \varepsilon t(p_t, q) \\ &\leq \frac{\nu}{2} \|\nabla \mathbf{e}\|^2 + \varepsilon \|q\|^2 + \frac{\varepsilon t^2}{2} \|p_t\|^2 \end{aligned}$$

Integrating from 0 to t , using Lemma 5.3.6, (5.2.41) and Gronwall lemma, we derive

$$\int_0^t s \|\mathbf{e}_t\|^2 ds + t \|\mathbf{e}\|_1^2 + \varepsilon t \|q\|^2 \leq C\varepsilon \quad (5.3.11)$$

We Multiply (5.3.9) by $t^2\mathbf{e}_t$ and (5.3.10) by t^2q_t and we sum up we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} t^2 \|\mathbf{e}_t\|^2 + \nu t^2 \|\nabla \mathbf{e}_t\|^2 + \varepsilon t^2 \|q_t\|^2 &= t \|\mathbf{e}_t\|^2 + \varepsilon t^2 (p_t, q_t) \\ &\leq t \|\mathbf{e}_t\|^2 + \frac{\varepsilon t^2}{2} \|q_t\|^2 + \frac{\varepsilon t^2}{2} \|p_t\|^2 \end{aligned}$$

Integrating over $[0, t]$, using (5.3.11) and (5.2.41) we get

$$\varepsilon \int_0^t s^2 \|q_t\|^2 ds \leq C \int_0^t s \|\mathbf{e}_t\|^2 ds + C\varepsilon \int_0^t s^2 \|p_t\|^2 ds \leq C\varepsilon$$

The results follows from the previous inequality and (5.2.41). \square

Lemma 5.3.3. *We suppose that (S1) and (S2) are satisfied then*

$$t \|\mathbf{e}(t)\|^2 + \nu \int_0^t s \|\nabla \mathbf{e}(s)\|^2 ds + \varepsilon \int_0^t s \|q(s)\|^2 ds \leq C\varepsilon^2 \quad \forall t \in [0, T_0], \quad (5.3.12)$$

$$t^2 \|\nabla \mathbf{e}(t)\|^2 + \int_0^t s^2 \|q(s)\|^2 ds \leq C\varepsilon^2 \quad \forall t \in [0, T_0]. \quad (5.3.13)$$

Proof. We consider the decomposition

$$\mathbf{H}_0^1(\Omega) = V \oplus V^\perp, \quad \text{where } V^\perp = \{(-\Delta)^{-1} \nabla q, \quad q \in L^2(\Omega)\}$$

and $\mathbf{v} = (-\Delta)^{-1}\nabla q$ if and only if $-\Delta\mathbf{v} = \nabla q$ and $\mathbf{v}|_{\partial\Omega} = 0$. We note that for $p(t) \in L^2(\Omega)/\mathbb{R}$ there exists a unique $\varphi(t) \in V^\perp$ such that $\nabla \cdot \varphi(t) = p(t)$ with

$$\|\varphi(t)\|_1 \leq C\|p(t)\| \quad \forall t \in [0, T_0]. \quad (5.3.14)$$

Moreover, if $p_t(t) \in L^2(\Omega)/\mathbb{R}$, we then have $\nabla \cdot \varphi_t(t) = p_t(t)$ such that

$$\|\varphi_t(t)\|_1 \leq C\|p_t(t)\| \quad \forall t \in [0, T_0]. \quad (5.3.15)$$

Multiplying (5.3.3) by $t\mathbf{e}$ and (5.3.4) by tq , we sum up and use (5.3.3), we infer

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} t \|\mathbf{e}\|^2 + t\nu \|\nabla \mathbf{e}\|^2 + \varepsilon t \|q\|^2 &= \frac{1}{2} \|\mathbf{e}\|^2 + \varepsilon t (p, q) = \frac{1}{2} \|\mathbf{e}\|^2 + \varepsilon t (\nabla \cdot \varphi, q) \\ &= \frac{1}{2} \|\mathbf{e}\|^2 - \varepsilon t (\nabla q, \varphi) = \frac{1}{2} \|\mathbf{e}\|^2 + \varepsilon t (\mathbf{e}_t, \varphi) + \varepsilon \nu t (\nabla \mathbf{e}, \nabla \varphi) \\ &= \frac{1}{2} \|\mathbf{e}\|^2 + \varepsilon \frac{d}{dt} t (\mathbf{e}, \varphi) - \varepsilon (\mathbf{e}, \varphi) - \varepsilon t (\mathbf{e}, \varphi_t) + \varepsilon \nu t (\nabla \mathbf{e}, \nabla \varphi) \end{aligned} \quad (5.3.16)$$

By the assumption (S1), we have $\sqrt{t}p \in L^\infty(0, T_0; H^1(\Omega)/\mathbb{R})$. Using (5.3.14) we obtain

$$\varepsilon t (\mathbf{e}(t), \varphi(t)) \leq \frac{t}{4} \|\mathbf{e}(t)\|^2 + \varepsilon^2 t \|\varphi\|^2 \leq \frac{t}{4} \|\mathbf{e}(t)\|^2 + C\varepsilon^2$$

By integrating (5.3.16) from 0 to t , using the previous inequality, the Cauchy-Schwartz inequality, Lemma 5.3.6, (5.2.41) and (5.3.15) we deduce

$$\begin{aligned} t \|\mathbf{e}(t)\|^2 + \int_0^t (\nu s \|\nabla \mathbf{e}(s)\|^2 + \varepsilon s \|q(s)\|^2) ds \\ \leq C\varepsilon^2 + C \int_0^t \|\mathbf{e}(s)\|^2 ds + c\varepsilon^2 \int_0^t \|\varphi(s)\|^2 ds + C\varepsilon^2 \int_0^t s^2 \|\varphi_t(s)\|^2 ds \\ \leq C\varepsilon^2 + C\varepsilon^2 \int_0^t (\|p\|^2 + s^2 \|p_t\|^2) ds \leq C\varepsilon^2 \end{aligned}$$

Taking the inner product of (5.3.3) with $t^2\mathbf{e}_t$, (5.3.10) with t^2q we deduct

$$t^2 \|\mathbf{e}_t\|^2 + \frac{1}{2} \frac{d}{dt} t^2 \|\nabla \mathbf{e}\|^2 + \frac{\varepsilon}{2} \frac{d}{dt} t^2 \|q\|^2 = t \|\mathbf{e}\|_1^2 + \varepsilon t \|q\|^2 + \varepsilon t^2 (p_t, q) \quad (5.3.17)$$

Using equation (5.3.3) and the inequality (5.3.15) to obtain

$$\begin{aligned} \varepsilon t^2 (p_t, q) &= \varepsilon t^2 (\nabla \cdot \varphi_t, q) = -\varepsilon t^2 (\varphi_t, \nabla q) = \varepsilon t^2 (\mathbf{e}_t, \varphi_t) + \varepsilon t^2 (\nabla \mathbf{e}, \nabla \varphi_t) \\ &\leq \frac{t^2}{2} \|\mathbf{e}_t\|^2 + c\varepsilon^2 t^2 \|\varphi_t\|^2 + t^2 \|\nabla \mathbf{e}\|^2 + \varepsilon^2 t^2 \|\nabla \varphi_t\|^2 \\ &\leq \frac{t^2}{2} \|\mathbf{e}_t\|^2 + t^2 \|\nabla \mathbf{e}\|^2 + C\varepsilon^2 t^2 \|p_t\|^2 \end{aligned}$$

We integrate (5.3.17) and using the Gronwall lemma to have

$$\int_0^t s^2 \|\mathbf{e}_t(s)\|^2 ds + t^2 \|\nabla \mathbf{e}\|^2 + \varepsilon t^2 \|q(t)\|^2 \leq C\varepsilon^2 \quad \forall t \in [0, T_0]$$

Again from the equation (5.3.3)

$$\|q\|^2 \leq C\|\nabla q\|_{-1}^2 \leq C(\|\Delta e\|_{-1}^2 + \|e_t\|_{-1}^2) \leq C(\|\mathbf{e}\|_1^2 + \|\mathbf{e}_t\|^2)$$

Consequently

$$\int_0^{T_0} s^2 \|q\|^2 ds \leq C \int_0^{T_0} s^2 (\|\mathbf{e}\|_1^2 + \|\mathbf{e}_t\|^2) ds \leq C\varepsilon^2$$

□

The result below describes the behavior of the linear error part.

Theorem 5.3.1. *Assume the assumption (S1). Then, there exists a constant C depending on the given data such that*

$$\|\mathbf{e}(t)\| + \left(\int_0^t \|\mathbf{e}(s)\|_1^2 ds \right)^{\frac{1}{2}} \leq C\sqrt{\varepsilon} \quad \forall t \in [0, T_0].$$

If we assume moreover the hypothesis (S2), then

$$\sqrt{t}\|\mathbf{e}(t)\| + t\|\mathbf{e}(t)\|_1 + \left(\int_0^t s^2 \|q\|^2 ds \right)^{\frac{1}{2}} \leq C\varepsilon \quad \forall t \in [0, T_0].$$

holds.

The following subsection treats the transfer of the results that have been derived for the linear case to the non linear one.

5.3.2 Error estimates for the nonlinear perturbed problem

We consider the following intermediate linear equations

$$\mathbf{v}_t - \nu\Delta\mathbf{v} + \nabla\gamma = f - B(\mathbf{u}, \mathbf{u}) - a|\mathbf{u}|^\alpha\mathbf{u} \quad (5.3.18)$$

$$\nabla \cdot \mathbf{v} + \varepsilon\gamma = 0, \quad \mathbf{v}(0) = \mathbf{u}_0, \quad (5.3.19)$$

with \mathbf{u} is solution of the Navier-Stokes equation with damping (5.1.1).

Taking $\xi = \mathbf{v} - \mathbf{u}$, $\phi = \gamma - p$, and subtracting (5.3.18)-(5.3.19) from (5.1.1), we get

$$\begin{aligned} \xi_t - \nu\Delta\xi + \nabla\phi &= 0 \\ \nabla \cdot \xi + \varepsilon\phi &= -\varepsilon p, \quad \xi(0) = 0. \end{aligned} \quad (5.3.20)$$

The next Lemma describes the control of the errors through the non-linearities.

Lemma 5.3.4. *We suppose that (S1) and (S2) are valid. Then we have*

$$\left(\int_0^t \|\xi(s)\|^2 ds \right)^{\frac{1}{2}} + \sqrt{t} \|\xi(t)\| + t \|\xi(t)\|_1 + \left(\int_0^t s^2 \|\phi(s)\|^2 ds \right)^{\frac{1}{2}} \leq C\varepsilon \quad \forall t \in [0, T_0].$$

Proof. The lemma is a consequence of Lemma.5.3.1 and Theorem.5.3.1 applied to the problem (5.3.20). \square

We now take $\eta = \mathbf{u}^\varepsilon - \mathbf{v}$, $q = p^\varepsilon - \gamma$, and subtracting (5.3.18)-(5.3.19) from (5.2.42)-(5.2.43) to obtain

$$\eta_t - \nu \Delta \eta + \tilde{B}(\mathbf{u}^\varepsilon, \mathbf{u}^\varepsilon) - \tilde{B}(\mathbf{u}, \mathbf{u}) + a(|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha \mathbf{u}) + \nabla q = 0 \quad (5.3.21)$$

$$\nabla \cdot \eta + \varepsilon q = 0, \quad \eta(0) = 0 \quad (5.3.22)$$

Or

$$\begin{aligned} \tilde{B}(\mathbf{u}^\varepsilon, \mathbf{u}^\varepsilon) - \tilde{B}(\mathbf{u}, \mathbf{u}) &= \tilde{B}(\mathbf{u}^\varepsilon, \mathbf{u}^\varepsilon - \mathbf{u}) + \tilde{B}(\mathbf{u}^\varepsilon - \mathbf{u}, \mathbf{u}^\varepsilon) \\ &= \tilde{B}(\mathbf{u}^\varepsilon, \xi + \eta) + \tilde{B}(\xi + \eta, \mathbf{u}^\varepsilon) \end{aligned} \quad (5.3.23)$$

We rewrite the equation (5.3.21) as

$$\eta_t + A_\varepsilon \eta + \tilde{B}(\mathbf{u}^\varepsilon, \xi + \eta) + \tilde{B}(\xi + \eta, \mathbf{u}^\varepsilon) + a(|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha \mathbf{u}) = 0 \quad (5.3.24)$$

The following theorem states a general priori estimate for the errors $\mathbf{e} = \mathbf{u} - \mathbf{u}^\varepsilon$ and $q = p - p^\varepsilon$; with (\mathbf{u}, p) are solutions of the state problem (5.1.1), and $(\mathbf{u}^\varepsilon, p^\varepsilon)$ are solutions of the penalized problem (5.2.42).

Theorem 5.3.2. *Assume that (S1) and (S2) hold. Then we have*

$$\sqrt{t} \|\mathbf{u}(t) - \mathbf{u}^\varepsilon(t)\| + \sqrt{\nu t} \|\mathbf{u}(t) - \mathbf{u}^\varepsilon(t)\|_1 + \left(\int_0^t s^2 \|p(t) - p^\varepsilon(t)\|^2 ds \right)^{\frac{1}{2}} \leq C\varepsilon \quad \forall t \in]0, T_0].$$

Proof. Multiplying (5.3.24) by $A_\varepsilon^{-1} \eta$, we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|A_\varepsilon^{\frac{1}{2}} \eta\|^2 + \nu \|\eta\|^2 &= -\tilde{b}(\mathbf{u}^\varepsilon, \xi + \eta, A_\varepsilon^{-1} \eta) - \tilde{b}(\xi + \eta, \mathbf{u}, A_\varepsilon^{-1} \eta) - a(|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha \mathbf{u}, A_\varepsilon^{-1} \eta) \\ &= I_1 + I_2 + I_3 \end{aligned}$$

Using (5.2.7), (5.2.8), and Cauchy Schwartz's inequality we derive that

$$\begin{aligned} I_1 &\leq C \|\mathbf{u}^\varepsilon\|_2 \|\xi + \eta\| \|A_\varepsilon^{-1} \eta\|_1 \leq C \|\mathbf{u}^\varepsilon\|_2 \|\xi + \eta\| \|\nabla A_\varepsilon^{-1} \eta\| \\ &\leq C \|\mathbf{u}^\varepsilon\|_2 (\|\xi\| + \|\eta\|) \|A_\varepsilon^{-\frac{1}{2}} \eta\| \leq \frac{\nu}{4} \|\eta\|^2 + C \|\xi\|^2 + C \|\mathbf{u}^\varepsilon\|_2^2 \|A_\varepsilon^{-\frac{1}{2}} \eta\|^2 \end{aligned}$$

With the same argument, we have

$$I_2 \leq \frac{\nu}{4} \|\eta\|^2 + C \|\xi\|^2 + C \|\mathbf{u}\|_2^2 \|A_\varepsilon^{-\frac{1}{2}} \eta\|^2$$

For the third term we proceed similarly

$$\begin{aligned} I_3 &\leq C \|\mathbf{u}^\varepsilon - \mathbf{u}\|_2 \|A_\varepsilon^{-\frac{1}{2}} \eta\| \\ &\leq C (\|\mathbf{u}^\varepsilon\|_2^2 + \|\mathbf{u}\|_2^2) + C \|A_\varepsilon^{-\frac{1}{2}} \eta\|^2 \end{aligned}$$

Adding the above inequalities we get

$$\frac{d}{dt} \|A_\varepsilon^{-\frac{1}{2}} \eta\|^2 \nu \|\eta\|^2 \leq C \|\xi\|^2 + C(1 + \|\mathbf{u}^\varepsilon\|_2^2 + \|\mathbf{u}\|_2^2) \|A_\varepsilon^{-\frac{1}{2}} \eta\|^2 + C(\|\mathbf{u}^\varepsilon\|_2^2 + \|\mathbf{u}\|_2^2) \quad (5.3.25)$$

Thanks to lemma 5.2.5, $\int_0^{T_0} (\|\mathbf{u}^\varepsilon\|_2^2 + \|\mathbf{u}\|_2^2) dt \leq C$, we can apply the Gronwall lemma to (5.3.25), using lemma 2.8 to obtain

$$\|A_\varepsilon^{-\frac{1}{2}} \eta(t)\|^2 + \nu \int_0^t \|\eta(s)\|^2 ds \leq C \int_0^t \|\xi(s)\|^2 ds \leq C \varepsilon^2 \quad \forall t \in [0, T_0]. \quad (5.3.26)$$

We now multiply the equation (5.3.21) by $t\eta$ and (5.3.22) by tq and we sum them up, then using (5.2.7) and Cauchy Schwartz's inequality one gets

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} t \|\eta\|^2 + \nu t \|\nabla \eta\|^2 + \varepsilon t \|q\|^2 &= \frac{1}{2} \|\eta\|^2 - t \tilde{b}(\mathbf{u}^\varepsilon, \xi + \eta, \eta) - t \tilde{b}(\xi + \eta, \mathbf{u}, \eta) \\ &\quad - at(|\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha \mathbf{u}, \eta) \\ &\leq \frac{1}{2} \|\eta\|^2 + Ct \|\mathbf{u}^\varepsilon\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|\mathbf{u}\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|\mathbf{u}^\varepsilon - \mathbf{u}\|_1 \|\eta\| \\ &\leq \frac{1}{2} \|\eta\|^2 + Ct \|\mathbf{u}^\varepsilon\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|\mathbf{u}\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|(\mathbf{u}^\varepsilon - \mathbf{v}) - (\mathbf{u} - \mathbf{v})\|_1 \|\eta\| \\ &\leq \frac{1}{2} \|\eta\|^2 + Ct \|\mathbf{u}^\varepsilon\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|\mathbf{u}\|_2 \|\xi + \eta\|_1 \|\eta\| + Ct \|\eta - \xi\|_1 \|\eta\| \\ &\leq \frac{1}{2} \|\eta\|^2 + \frac{\nu t}{2} \|\nabla \eta\|^2 + Ct \|\xi\|_1^2 + Ct (\|\mathbf{u}^\varepsilon\|^2 + \|\mathbf{u}\|^2) \|\eta\|^2 \end{aligned}$$

We integrate from 0 to t and use (5.3.26), Lemma 2.8 and the Gronwall lemma we get

$$t \|\eta(t)\|^2 + \nu \int_0^t s \|\nabla \eta(s)\|^2 ds + \varepsilon \int_0^t s \|q(s)\|^2 ds \leq C \varepsilon^2 \quad (5.3.27)$$

We now derive one time with respect to the equation (5.3.22)

$$\nabla \cdot \eta_t + \varepsilon q_t = 0 \quad (5.3.28)$$

Multiplying the equation (5.3.21) by $t^2\eta_t$, and (5.3.28) by t^2q , summing them up to get

$$\begin{aligned} t^2\|\eta_t\|^2 + \frac{\nu}{2}\frac{d}{dt}t^2\|\nabla\eta\|^2 + \frac{\varepsilon}{2}\frac{d}{dt}t^2\|q\|^2 &= \nu t\|\nabla\eta\|^2 + \varepsilon t\|q\|^2 \\ -t^2\tilde{b}(\mathbf{u}^\varepsilon, \xi + \eta, \eta_t) - t^2\tilde{b}(\xi + \eta, \mathbf{u}, \eta_t) - at^2(|\mathbf{u}^\varepsilon|^\alpha\mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha\mathbf{u}, \eta_t) \end{aligned} \quad (5.3.29)$$

For the first nonlinear term in the right hand side, we proceed as follow

$$\begin{aligned} t^2\tilde{b}(\mathbf{u}^\varepsilon, \xi + \eta, \eta_t) &\leq t^2\|\mathbf{u}^\varepsilon\|_2\|\xi + \eta\|_1\|\eta_t\| \\ &\leq \frac{t^2}{6}\|\eta_t\|^2 + Ct^2\|\mathbf{u}^\varepsilon\|_2^2(\|\xi\|_1^2 + \|\eta\|_1^2) \\ &\leq \frac{t^2}{6}\|\eta_t\|^2 + C\varepsilon^2\|\mathbf{u}^\varepsilon\|_2^2 + Ct^2\|\mathbf{u}^\varepsilon\|_2^2\|\nabla\eta\|^2 \end{aligned}$$

With the same manner we obtain

$$t^2\tilde{b}(\xi + \eta, \mathbf{u}, \eta_t) \leq \frac{t^2}{6}\|\eta_t\|^2 + C\varepsilon^2\|\mathbf{u}\|_2^2 + Ct^2\|\mathbf{u}\|_2^2\|\nabla\eta\|^2$$

For the last non linear term we have

$$\begin{aligned} at^2(|\mathbf{u}^\varepsilon|^\alpha\mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha\mathbf{u}, \eta_t) &\leq at^2\|\eta_t\|\|\mathbf{u}^\varepsilon - \mathbf{u}\|_1 \\ &\leq \frac{t^2}{6}\|\eta_t\|^2 + Ct^2\|\eta - \xi\|_1^2 \end{aligned}$$

Summing the above inequalities into (5.3.29) to get

$$\begin{aligned} t^2\|\eta_t\|^2 + \nu\frac{d}{dt}t^2\|\nabla\eta\|^2 + \varepsilon\frac{d}{dt}t^2\|q\|^2 &\leq \nu t\|\nabla\eta\|^2 + \varepsilon t\|q\|^2 \\ + C(\varepsilon^2 + t^2\|\nabla\eta\|^2)(\|\mathbf{u}^\varepsilon\|_2^2 + \|\mathbf{u}\|_2^2) + Ct^2\|\eta - \xi\|_1^2 \end{aligned}$$

We integrate from 0 to t , using (5.3.27) and the Gronwall lemma we obtain

$$\int_0^t s^2\|\eta_t\|^2 ds + \nu t^2\|\nabla\eta(t)\|^2 + \varepsilon t^2\|q(t)\|^2 \leq C\varepsilon^2 \quad (5.3.30)$$

Using (5.2.7), we have

$$\begin{aligned} \|\tilde{B}(\mathbf{u}^\varepsilon, \xi + \eta)\|_{-1} &\leq C(\|\mathbf{u}^\varepsilon\|_1\|\xi + \eta\|_1) \leq C(\|\mathbf{u}^\varepsilon\|_1\|\xi\|_1 + \|\eta\|) \\ \|\tilde{B}(\xi + \eta, \mathbf{u})\|_{-1} &\leq C(\|\mathbf{u}\|_1\|\xi + \eta\|_1) \leq C(\|\mathbf{u}\|_1\|\xi\|_1 + \|\eta\|_1) \\ a(|\mathbf{u}^\varepsilon|^\alpha\mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha\mathbf{u}) &\leq C(\|\xi\|_1 + \|\eta\|_1) \end{aligned}$$

From (5.3.21) and (5.3.23) we obtain

$$\nabla q = -\eta_t + \nu\Delta\eta - \tilde{B}(\mathbf{u}^\varepsilon, \xi + \eta) - \tilde{B}(\xi + \eta, \mathbf{u}) - a(|\mathbf{u}^\varepsilon|^\alpha\mathbf{u}^\varepsilon - |\mathbf{u}|^\alpha\mathbf{u})$$

consequently, by applying the above estimates on the previous equation we get

$$\int_0^{T_0} s^2\|q\|^2 ds \leq \int_0^{T_0} s^2\|\nabla q\|_{-1}^2 ds \leq C\varepsilon^2$$

which end the proof of the theorem 4.2.2. \square

5.4 Control problem

5.4.1 Adjoint equation

In this section we derive an adjoint equation related to the equation (5.1.2). For simplicity we assume that $\mathbf{f} = 0$.

First, we define the objective function

$$J(\Omega) = J_1(\Omega) + J_2(\Omega), \quad (5.4.1)$$

where

$$J_1(\Omega) = \frac{1}{2} \int_0^T \int_{\Omega} |\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 dt dx, \text{ and } J_2(\Omega) = \frac{\sigma}{2} \int_0^T \int_{\Omega} |\mathit{curl}(\mathbf{u}^\varepsilon)|^2 dt dx.$$

Then the optimization problem can be expressed as

$$\begin{aligned} \min_{\Omega \in Q_{ad}} J(\Omega) &= \frac{1}{2} \int_0^T \int_{\Omega} |\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 dt dx + \frac{\sigma}{2} \int_0^T \int_{\Omega} |\mathit{curl}(\mathbf{u}^\varepsilon)|^2 dt dx \\ &\text{such that } (\mathbf{u}^\varepsilon, p^\varepsilon) \text{ is solution of (5.1.2)} \end{aligned} \quad (5.4.2)$$

Theorem 5.4.1. *Let $(\mathbf{u}^\varepsilon, p^\varepsilon) \in H^2(0, T; \Omega) \times Y$ be solution of the penalized problem (5.1.2). Then the adjoint equation associated to the equation (5.1.2) takes the form*

$$\left\{ \begin{array}{ll} -\mathbf{v}_t - \nu \Delta \mathbf{v} + (\nabla \mathbf{u}^\varepsilon)^T \cdot \mathbf{v} - (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} - \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} + a |\mathbf{u}^\varepsilon|^{\alpha-2} (|\mathbf{u}^\varepsilon|^2 \mathbf{v} + \alpha (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon) \\ \quad + \nabla q = (\mathbf{u}^\varepsilon - \mathbf{u}_d) - \sigma \vec{\mathit{curl}}(\mathit{curl}(\mathbf{u}^\varepsilon)) & \text{in } \Omega \times [0, T] \\ \nabla \cdot \mathbf{v} + \varepsilon q = 0 & \text{in } \Omega \times [0, T] \\ \mathbf{v} = 0 & \text{on } (\Gamma_0 \cup \Gamma_1) \times [0, T] \\ \nu \frac{\partial \mathbf{v}}{\partial n} + 2(\mathbf{u}^\varepsilon \cdot n) \mathbf{v} - n q = \sigma \mathit{curl}(\mathbf{u}^\varepsilon) \cdot \boldsymbol{\tau} & \text{on } \Gamma_2 \times [0, T] \\ \mathbf{v}(T) = 0 & \end{array} \right. \quad (5.4.3)$$

Proof. We define also the following function

$$\begin{aligned} F(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q) &= \int_0^T \int_{\Omega} \left(\mathbf{u}_t^\varepsilon + \nu \nabla \mathbf{u}^\varepsilon : \nabla \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{u}^\varepsilon \cdot \mathbf{v} + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{u}^\varepsilon \cdot \mathbf{v} + a |\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon \cdot \mathbf{v} \right) dt dx \\ &\quad - \int_0^T \int_{\Omega} \left(p^\varepsilon \nabla \cdot \mathbf{v} - q \nabla \cdot \mathbf{u}^\varepsilon - \varepsilon q p^\varepsilon \right) dt dx. \end{aligned}$$

Let $L(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q)$ be a Lagrangian functional defined by

$$L(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q) = J(\Omega) - F(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q), \quad (5.4.4)$$

we first derive L with respect to the state variable p in any direction $\tilde{p} \in M(\Omega)$, we obtain

$$\frac{\partial L}{\partial p}(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q) \cdot \tilde{p} = \int_0^T \int_\Omega (\tilde{p} \nabla \cdot \mathbf{v} + \varepsilon \tilde{p} q) dt dx,$$

the variation \tilde{p} is arbitrary, we get

$$\nabla \cdot \mathbf{v} + \varepsilon q = 0 \quad \text{on } \Omega. \quad (5.4.5)$$

Next, we derive L with respect of the state variable \mathbf{u}^ε in the arbitrary direction $\tilde{\mathbf{u}} \in V_0(\Omega)$ we have

$$\begin{aligned} 0 &= \frac{\partial L}{\partial \mathbf{u}^\varepsilon}(\Omega, \mathbf{u}^\varepsilon, p^\varepsilon, \mathbf{v}, q) \cdot \tilde{\mathbf{u}}, \\ &= \int_0^T \int_\Omega \left((\mathbf{u}^\varepsilon - \mathbf{u}_d) \cdot \tilde{\mathbf{u}} \sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \cdot \operatorname{curl}(\tilde{\mathbf{u}}) \right) dt dx \\ &\quad - \int_0^T \int_\Omega \left(-\mathbf{v}_t \tilde{\mathbf{u}} + \nu \nabla \tilde{\mathbf{u}} : \nabla \mathbf{v} + (\tilde{\mathbf{u}} \cdot \nabla) \mathbf{u}^\varepsilon \cdot \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla) \tilde{\mathbf{u}} \cdot \mathbf{v} + \frac{1}{2} (\nabla \cdot \tilde{\mathbf{u}}) \mathbf{u}^\varepsilon \cdot \mathbf{v} + (\nabla \cdot \mathbf{u}^\varepsilon) \tilde{\mathbf{u}} \cdot \mathbf{v} \right) dt dx \\ &\quad + \int_0^T \int_\Omega \left(q \nabla \cdot \tilde{\mathbf{u}} - (a |\mathbf{u}^\varepsilon|^\alpha \tilde{\mathbf{u}} + a \alpha |\mathbf{u}^\varepsilon|^{\alpha-2} (\mathbf{u}^\varepsilon \cdot \tilde{\mathbf{u}}) \mathbf{u}^\varepsilon) \cdot \mathbf{v} \right) dt dx, \\ &= \int_0^T \int_\Omega \left(\mathbf{u}^\varepsilon - \mathbf{u}_d \right) - \sigma \vec{\operatorname{curl}}(\operatorname{curl}(\mathbf{u}^\varepsilon)) + \mathbf{v}_t + \nu \Delta \mathbf{u}^\varepsilon - (\nabla \mathbf{u}^\varepsilon)^T \cdot \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} \\ &\quad + \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) - \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} - (a |\mathbf{u}^\varepsilon|^\alpha \mathbf{v} + a \alpha |\mathbf{u}^\varepsilon|^{\alpha-2} (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon - \nabla q) \cdot \tilde{\mathbf{u}} dt dx \\ &\quad + \int_0^T \int_{\Gamma_2} \left(\sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \tau - \left(\nu \frac{\partial \mathbf{v}}{\partial n} + 2(\mathbf{u}^\varepsilon \cdot n) \mathbf{v} - n q \right) \right) \cdot \tilde{\mathbf{u}} dt dx + \int_\Omega [\mathbf{v} \cdot \tilde{\mathbf{u}}]_0^T dx. \end{aligned}$$

Considering an arbitrary direction $\tilde{\mathbf{u}}$ which vanishes in neighbourhood of the boundary Γ_2 , we obtain

$$-\mathbf{v}_t - \nu \Delta \mathbf{v} + (\nabla \mathbf{u})^T \cdot \mathbf{v} - (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} - \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} + a |\mathbf{u}^\varepsilon|^{\alpha-2} (|\mathbf{u}^\varepsilon|^2 \mathbf{v} + \alpha (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon) + \nabla q = (\mathbf{u}^\varepsilon - \mathbf{u}_d) - \sigma \vec{\operatorname{curl}}$$

and an arbitrary $\tilde{\mathbf{u}}$ in Γ_2 gives

$$\nu \frac{\partial \mathbf{v}}{\partial n} + 2(u \cdot n) \mathbf{v} - n q = \sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \cdot \tau \quad \text{on } \Gamma_2$$

Finally we obtain the adjoint equation

$$\left\{ \begin{array}{l} -\mathbf{v}_t - \nu \Delta \mathbf{v} + (\nabla \mathbf{u}^\varepsilon)^T \cdot \mathbf{v} - (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} - \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} + a |\mathbf{u}^\varepsilon|^{\alpha-2} (|\mathbf{u}^\varepsilon|^2 \mathbf{v} + \alpha (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon) \\ \quad + \nabla q = (\mathbf{u}^\varepsilon - \mathbf{u}_d) - \sigma \vec{curl}(curl(\mathbf{u}^\varepsilon)) \quad \text{in } \Omega \times [0, T] \\ \nabla \cdot \mathbf{v} + \varepsilon q = 0 \quad \text{in } \Omega \times [0, T] \\ \mathbf{v} = 0 \quad \text{on } (\Gamma_0 \cup \Gamma_1) \times [0, T] \\ \nu \frac{\partial \mathbf{v}}{\partial n} + 2(\mathbf{u}^\varepsilon \cdot n) \mathbf{v} - nq = \sigma curl(\mathbf{u}^\varepsilon) \cdot \tau \quad \text{on } \Gamma_2 \times [0, T] \\ \mathbf{v}(T) = 0 \end{array} \right. \quad (5.4.6)$$

□

5.4.2 Shape gradient

In this subsection, we will try to express the shape gradient using the velocity method (see Céa [4]), via the state and adjoint problems.

Let Ω be a reference domain in \mathbb{R}^2 , the perturbation of Ω by the velocity method is described as the velocity field

$$\mathbf{V}(t)(x) = \mathbf{V}(t, x), \forall x \in \Omega, t \geq 0.$$

It can generate transformations

$$T_t(\mathbf{V})\mathbf{X} = x(t, \mathbf{X}), t \geq 0, \mathbf{X} \in \Omega,$$

the flow defined by the initial value problem

$$\left\{ \begin{array}{l} \frac{d\chi}{dt}(t, \mathbf{X}) = \mathbf{V}(t, \chi(t)), \\ \chi(0, \mathbf{X}) = \mathbf{X}, \end{array} \right. \quad (5.4.7)$$

with the initial value X given. Let Ω_t be a perturbation domain of Ω and $J(\Omega)$ be a functional associated to Ω_t . The shape derivative of the functional $J(\Omega_t)$ at Ω in the direction of the deformation field \mathbf{V} is written as

$$dJ(\Omega; \mathbf{V}) = \lim_{t \rightarrow 0} \frac{J(\Omega_t) - J(\Omega)}{t}.$$

If $dJ(\Omega; \mathbf{V})$ exists for all $\mathbf{V} \in \mathcal{C}([0, T]; (\mathcal{D}^k(\mathbb{R}^2))^2)$, for small positive constant T , the functional J is called shape differentiable at Ω and its shape gradient verifies

$$dJ(\Omega; \mathbf{V}) = (\nabla J, \mathbf{V})_{((\mathcal{D}^k(\mathbb{R}^2))^2)' \times (\mathcal{D}^k(\mathbb{R}^2))^2}$$

We assume that the boundaries Γ_1 and Γ_2 are fixed in the admissible set of domains, and we define the velocity field admissible domain as follow

$$V_{ad} = \{ \mathbf{V} \in \mathcal{C}^0(0, \tau; (\mathcal{C}^2(\mathbb{R}^2))^2) \mid \mathbf{V} = 0 \quad \text{on } \Gamma_1 \cup \Gamma_2 \}.$$

The use of the velocity \mathbf{V} for $t \geq 0$, implies the transformation of the domain Ω into $\Omega_t = T_t(\Omega)$ by the velocity method with formulation (5.4.7). Let us find an expression of the derivative of the saddle point problem $j(t)$ with respect to \mathbf{t} where

$$j(t) = \min_{(\mathbf{u}^\varepsilon, p_y) \in V_g(\Omega_t) \times \hat{Y}(\Omega_t)} \max_{(\mathbf{v}_t, q_t) \in V_0(\Omega_t) \times \hat{Y}(\Omega_t)} L(\Omega_t, \mathbf{u}_t^\varepsilon, p_t^\varepsilon, \mathbf{v}_t, q_t)$$

with (\mathbf{u}, p) and (\mathbf{v}, q) are solutions of (5.1.1) and (5.4.3) in the perturbed domain Ω_t , respectively.

We consider the Hilbert spaces which depend on the parameter t defined by

$$V_0(\Omega_t) = \{ \mathbf{u}^\varepsilon \circ T_t^{-1} : \mathbf{u}^\varepsilon \in V_0(\Omega) \}$$

$$V_g(\Omega_t) = \{ \mathbf{u}^\varepsilon \circ T_t^{-1} : \mathbf{u}^\varepsilon \in V_g(\Omega) \}$$

$$\hat{Y}(\Omega_t) = \{ p^\varepsilon \circ T_t^{-1} : p^\varepsilon \in \hat{Y}(\Omega) \}$$

since T_t and T_t^{-1} are diffeomorphisms, the parametrisation do not influence $j(t)$. And we have

$$j(t) = \min_{(\mathbf{u}^\varepsilon, p^\varepsilon) \in V_g(\Omega) \times \hat{Y}(\Omega)} \max_{(\mathbf{v}, q) \in V_0(\Omega) \times \hat{Y}(\Omega)} L(\Omega_t, \mathbf{u}^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, \mathbf{v} \circ T_t^{-1}, q \circ T_t^{-1})$$

we define the following functions which depend on the parameter t

$$\begin{aligned} l_1(t) &= \frac{1}{2} \int_0^T \int_{\Omega_t} |\mathbf{u}^\varepsilon \circ T_t^{-1} - \mathbf{u}_d|^2 dx + \frac{\sigma}{2} \int_0^T \int_{\Omega_t} |\text{curl}(\mathbf{u}^\varepsilon \circ T_t^{-1})|^2 dt dx, \\ l_2(t) &= \int_0^T \int_{\Omega_t} [\nu \nabla(\mathbf{u}^\varepsilon \circ T_t^{-1}) : \nabla(\mathbf{v} \circ T_t^{-1}) + ((\mathbf{u}^\varepsilon \circ T_t^{-1}) \cdot \nabla)(\mathbf{u}^\varepsilon \circ T_t^{-1}) \cdot (\mathbf{v} \circ T_t^{-1}) \\ &\quad + a |\mathbf{u}^\varepsilon \circ T_t^{-1}|^\alpha (\mathbf{u}^\varepsilon \circ T_t^{-1}) \cdot (\mathbf{v} \circ T_t^{-1}) - (p^\varepsilon \circ T_t^{-1}) \nabla \cdot (\mathbf{v} \circ T_t^{-1})] dt dx \\ &\quad - \int_0^T \int_{\Omega_t} (q \circ T_t^{-1}) \nabla \cdot (\mathbf{u}^\varepsilon \circ T_t^{-1}) dt dx - \varepsilon \int_0^T \int_{\Omega_t} (q \circ T_t^{-1}) (p^\varepsilon \circ T_t^{-1}) dt dx \end{aligned}$$

The Lagrangian functional writes

$$L(\Omega_t, \mathbf{u}^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, \mathbf{v} \circ T_t^{-1}, q \circ T_t^{-1}) = l_1(t) - l_2(t)$$

If $\Phi : [0, \tau] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is sufficiently smooth, we have the following Hadamard formula

$$\frac{d}{dt} \int_0^T \int_{\Omega_t} \Phi(t, x) dx \Big|_{t=0} = \int_0^T \int_{\Omega} \frac{\partial \Phi}{\partial t}(0, x) dx + \int_0^T \int_{\partial \Omega} \Phi(0, x) \mathbf{v}(0, X) \cdot \mathbf{n} ds \quad (5.4.8)$$

We consider $\mathbf{V}(0, X) \in V_{ad}$, and observe that $\mathbf{V}(0, X) = \mathbf{V}$. Therefore we can derive the shape gradient using the formula (5.4.7)

$$\frac{d}{dt}L(\Omega_t, \mathbf{u}^\varepsilon \circ T_t^{-1}, p^\varepsilon \circ T_t^{-1}, \mathbf{v} \circ T_t^{-1}, q \circ T_t^{-1})|_{t=0} = l'_1(0) - l'_2(0) \quad (5.4.9)$$

where

$$\begin{aligned} l'_1(0) &= \int_0^T \int_\Omega (\mathbf{u}^\varepsilon - \mathbf{u}_d) \cdot (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) dx + \sigma \int_0^T \int_\Omega \operatorname{curl}(\mathbf{u}^\varepsilon) \operatorname{curl}(-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) dx \\ &\quad + \frac{1}{2} \int_0^T \int_{\Gamma_0} (|\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 \mathbf{V} \cdot \mathbf{n}) ds + \frac{\sigma}{2} \int_0^T \int_{\Gamma_0} (|\operatorname{curl}(\mathbf{u}^\varepsilon)|^2 \mathbf{V} \cdot \mathbf{n}) ds \end{aligned} \quad (5.4.10)$$

$$\begin{aligned} l'_2(0) &= \int_0^T \int_\Omega [(-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V})_t v + (-\nabla \mathbf{v} \cdot \mathbf{V})_t \mathbf{u}^\varepsilon + \nu \nabla(-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) : \nabla \mathbf{v} + \nu \nabla \mathbf{u}^\varepsilon : \nabla(-\nabla \mathbf{v} \cdot \mathbf{V}) \\ &\quad + ((-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) \cdot \nabla \mathbf{u}^\varepsilon) \cdot \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla(-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V})) \cdot \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^\varepsilon) \cdot (-\nabla \mathbf{v} \cdot \mathbf{V}) \\ &\quad + \frac{1}{2} \nabla \cdot (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) \mathbf{u}^\varepsilon \cdot \mathbf{v} + \frac{1}{2} \nabla \cdot \mathbf{u}^\varepsilon (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) \cdot \mathbf{v} + \frac{1}{2} \nabla \cdot \mathbf{u}^\varepsilon \mathbf{u}^\varepsilon \cdot (-\nabla \mathbf{v} \cdot \mathbf{V}) \\ &\quad + a\alpha |\mathbf{u}^\varepsilon|^{\alpha-2} (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V} \cdot \mathbf{u}^\varepsilon) (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + a |\mathbf{u}^\varepsilon|^\alpha (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) \cdot \mathbf{v} \\ &\quad + a |\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon \cdot (-\nabla \mathbf{v} \cdot \mathbf{V}) - (\nabla p \cdot \mathbf{V}) \nabla \cdot \mathbf{v} \\ &\quad - p \nabla \cdot (-\nabla \mathbf{v} \cdot \mathbf{V}) - q \nabla \cdot (-\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) - (\nabla \cdot \mathbf{u}^\varepsilon) (-\nabla q \cdot \mathbf{V})] dx \end{aligned}$$

Using Green's formula and the condition $\mathbf{u}^\varepsilon = 0$ on Γ_0 , we obtain

$$\begin{aligned} l'_2(0) &= - \int_0^T \int_\Omega [(\mathbf{u}_t^\varepsilon - \nu \Delta \mathbf{u}^\varepsilon + (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{u}^\varepsilon + \frac{1}{2} \nabla \cdot \mathbf{u}^\varepsilon \mathbf{u}^\varepsilon + a |\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon + \nabla p) \cdot (\nabla \mathbf{v} \cdot \mathbf{V})] dt dx \\ &\quad + \int_0^T \int_\Omega (\nabla \cdot \mathbf{u}^\varepsilon) (\nabla q \cdot \mathbf{V}) dt dx + \int_0^T \int_\Omega (\nabla \cdot \mathbf{v}) (\nabla p \cdot \mathbf{V}) dt dx + \int_0^T \int_\Omega [-\mathbf{v}_t - \nu \Delta \mathbf{v} + (\nabla \mathbf{u}^\varepsilon)^T \cdot \mathbf{v} \\ &\quad - (\mathbf{u}^\varepsilon \cdot \nabla) \mathbf{v} - \frac{1}{2} \nabla (\mathbf{u}^\varepsilon \cdot \mathbf{v}) + \frac{1}{2} (\nabla \cdot \mathbf{u}^\varepsilon) \mathbf{v} + a |\mathbf{u}^\varepsilon|^{\alpha-2} (|\mathbf{u}^\varepsilon|^2 v + \alpha (\mathbf{u}^\varepsilon \cdot \mathbf{v}) \mathbf{u}^\varepsilon) - \nabla q] \cdot (\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) dt dx \\ &\quad - \int_0^T \int_{\Gamma_0} [\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} + 2(\mathbf{u}^\varepsilon \cdot \mathbf{n}) \mathbf{v} - nq] \cdot (\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) dt ds - \int_0^T \int_{\Gamma_0} [\frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} - nq] \cdot (\nabla \mathbf{v} \cdot \mathbf{V}) dt ds \\ &\quad + \int_0^T \int_{\Gamma_0} [\nu \nabla \mathbf{u}^\varepsilon : \nabla \mathbf{v} + (\mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^\varepsilon) \cdot \mathbf{v} + a |\mathbf{u}^\varepsilon|^\alpha \mathbf{u}^\varepsilon \cdot \mathbf{v} - p \nabla \cdot \mathbf{v} - q \nabla \cdot \mathbf{u}^\varepsilon] \mathbf{V} \cdot \mathbf{n} dt ds \end{aligned} \quad (5.4.11)$$

Replacing (5.4.10) and (5.4.11) in (5.4.9) and using the fact that $(\mathbf{u}^\varepsilon, p^\varepsilon)$ is solution of (5.1.2) and (\mathbf{v}, q) is solution of (5.4.3) respectively yields

$$\begin{aligned}
dJ(\Omega; \mathbf{V}) &= \frac{d}{dt} L|_{t=0} \\
&= \frac{1}{2} \int_0^T \int_{\Gamma_0} (|\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 + \sigma |\operatorname{curl}(\mathbf{u}^\varepsilon)|^2) \mathbf{V} \cdot \mathbf{n} ds + \int_0^T \int_{\Gamma_0} [\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} + \mathbf{n}q] \cdot (\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) dt ds \\
&\quad + \int_0^T \int_{\Gamma_0} [\nu \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} - \mathbf{n}p] \cdot (\nabla \mathbf{v} \cdot \mathbf{V}) dt ds - \int_0^T \int_{\Gamma_0} [(\nu \nabla \mathbf{u}^\varepsilon : \nabla \mathbf{v}) \mathbf{V} \cdot \mathbf{n} dt ds \\
&\quad - \sigma \int_0^T \int_{\Gamma_0} (\frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot (\operatorname{curl}(\mathbf{u}^\varepsilon) \wedge \mathbf{n})) \mathbf{V} \cdot \mathbf{n} dt ds
\end{aligned} \tag{5.4.12}$$

Note that $\mathbf{u}^\varepsilon = 0$ and $\mathbf{v} = 0$ in Γ_0 we have

$$\mathbf{n} \cdot (\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) = \nabla \mathbf{u}^\varepsilon \cdot (\mathbf{n} \otimes \mathbf{n}) \cdot \mathbf{V} \cdot \mathbf{n} = \nabla \mathbf{u}^\varepsilon \cdot \mathbf{n} \cdot \mathbf{n} (\mathbf{V} \cdot \mathbf{n}) = (\nabla \cdot \mathbf{u}^\varepsilon) (\mathbf{V} \cdot \mathbf{n}) = 0, \quad \forall x \in \Gamma_0 \tag{5.4.13}$$

$$\frac{\partial \mathbf{v}}{\partial \mathbf{n}} \cdot (\nabla \mathbf{u}^\varepsilon \cdot \mathbf{V}) = \nabla \mathbf{u}^\varepsilon \cdot (\mathbf{n} \otimes \mathbf{n}) \cdot \mathbf{V} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{n}} = \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{n}} (\mathbf{V} \cdot \mathbf{n}) = (\nabla \mathbf{u}^\varepsilon : \nabla \mathbf{v}) \mathbf{V} \cdot \mathbf{n} \tag{5.4.14}$$

Similarly we obtain

$$\mathbf{n} \cdot (\nabla \mathbf{v} \cdot \mathbf{V}) = 0, \quad \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot (\nabla \mathbf{v} \cdot \mathbf{V}) = \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{n}} (\mathbf{V} \cdot \mathbf{n}) = (\nabla \mathbf{u}^\varepsilon : \nabla \mathbf{v}) \mathbf{V} \cdot \mathbf{n} \tag{5.4.15}$$

Substituting (5.4.13), (5.4.14) and (5.4.15) in (5.4.12) the shape derivative takes the form

$$dJ(\Omega; \mathbf{V}) = \int_0^T \int_{\Gamma_0} \left[\frac{1}{2} |\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 + \frac{\sigma}{2} |\operatorname{curl}(\mathbf{u}^\varepsilon)|^2 + \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} - \sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{V} \cdot \mathbf{n} dt ds \tag{5.4.16}$$

Consequently, the shape gradient writes

$$\nabla J = \left[\frac{1}{2} |\mathbf{u}^\varepsilon - \mathbf{u}_d|^2 + \frac{\sigma}{2} |\operatorname{curl}(\mathbf{u}^\varepsilon)|^2 + \frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} \cdot \left(2\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}} - \sigma \operatorname{curl}(\mathbf{u}^\varepsilon) \wedge \mathbf{n} \right) \right] \mathbf{n} \tag{5.4.17}$$

5.5 Discrete finite element method

5.5.1 Preliminaries

In this section we write some results and notations which will be used in this work.

We define the following forms

$$a(\mathbf{u}, \mathbf{v}) = \nu (A^{1/2} \mathbf{u}, A^{1/2} \mathbf{v}), \quad a_\varepsilon(\mathbf{u}, \mathbf{v}) = \nu (A_\varepsilon^{1/2} \mathbf{u}, A_\varepsilon^{1/2} \mathbf{v}), \quad c_\alpha(\mathbf{u}, \mathbf{v}) = (a |\mathbf{u}|^\alpha \mathbf{u}, \mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v} \in X$$

$$d(\mathbf{v}, q) = (\nabla \cdot \mathbf{v}, q), \quad \forall \mathbf{v} \in X, \forall q \in M$$

We shall make two standing assumption on the datum \mathbf{u}_0 and \mathbf{f}

The initial velocity $\mathbf{u}_0(x) \in D(A)$ with $\operatorname{div} \mathbf{u}_0 = 0$ and the forcing function

$$\mathbf{f}(x, t) \in H^{1,\infty}(0, T; Y) \text{ satisfy} \tag{A1}$$

$$\|\mathbf{A}\mathbf{u}_0\| + \sup_{t \in [0, T]} \{\|\mathbf{f}(t)\| + \|\mathbf{f}_t(t)\|\} \leq C,$$

for some positive constant C .

We recall the two weak formulations concerning the initial problem and its penalized.

Find $(\mathbf{u}, p) \in L^\infty(0, T; Y) \cap L^2(0, T; X) \times L^2(0, T; M)$ such that

$$(\mathbf{u}_t, \mathbf{v}) + a(\mathbf{u}, \mathbf{v}) - d(\mathbf{v}, p) + d(\mathbf{u}, q) + b(\mathbf{u}, \mathbf{u}, \mathbf{v}) + c_\alpha(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}), \forall (\mathbf{v}, q) \in X \times M \tag{5.5.1}$$

and

Find $(\mathbf{u}_\varepsilon, p_\varepsilon) \in L^\infty(0, T; Y) \cap L^2(0, T; X) \times L^2(0, T; M)$ such that

$$(\mathbf{u}_{\varepsilon t}, \mathbf{v}) + a(\mathbf{u}_\varepsilon, \mathbf{v}) - d(\mathbf{v}, p_\varepsilon) + d(\mathbf{u}_\varepsilon, q) + \varepsilon(p_\varepsilon, q) + \tilde{b}(\mathbf{u}_\varepsilon, \mathbf{u}_\varepsilon, \mathbf{v}) + c_\alpha(\mathbf{u}_\varepsilon, \mathbf{v}) = (\mathbf{f}, \mathbf{v})$$

for all $(\mathbf{v}, q) \in X \times M$

$$\tag{5.5.2}$$

with the initial data $\mathbf{u}(0) = \mathbf{u}_0$ and $\mathbf{u}_\varepsilon(0) = \mathbf{u}_0$ respectively.

Now, we define the time discretization of the penalized weak formulation (5.5.2) by the backward Euler scheme

$$(d_t \mathbf{u}_\varepsilon^n, \mathbf{v}) + a(\mathbf{u}_\varepsilon^n, \mathbf{v}) - d(\mathbf{v}, p_\varepsilon^n) + d(\mathbf{u}_\varepsilon^n, q) + \varepsilon(p_\varepsilon^n, q) + \tilde{b}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, \mathbf{v}) + c_\alpha(\mathbf{u}_\varepsilon^n, \mathbf{v}) = (f(t_n), \mathbf{v}) \tag{5.5.3}$$

for all $(\mathbf{v}, q) \in X \times M$, and $1 \leq n \leq N$, where $0 < \Delta t < 1$ is the time step size, $t_n = n\Delta$, $t_N = T$, $(\mathbf{u}_\varepsilon^0, p_\varepsilon^0) = (\mathbf{u}_0, 0)$ and $d_t \mathbf{u}_\varepsilon^n = \frac{\mathbf{u}_\varepsilon^n - \mathbf{u}_\varepsilon^{n-1}}{\Delta t}$ for $1 \leq n \leq N$, and $d_t \mathbf{u}_\varepsilon^0$ satisfies the following equation

$$(d_t \mathbf{u}_\varepsilon^0, \mathbf{v}) = a(\mathbf{u}_0, \mathbf{v}) + ((\mathbf{u}_0 \cdot \nabla) \mathbf{u}_0, \mathbf{v}) + c_\alpha(\mathbf{u}_0, \mathbf{v}) = (\mathbf{f}(0), \mathbf{v}), \forall \mathbf{v} \in X \text{ with } \operatorname{div} \mathbf{v} = 0$$

By using the inequality (5.2.11), we infer that

$$\begin{aligned} \|d_t \mathbf{u}_\varepsilon^0\| &\leq \nu \|\mathbf{A}\mathbf{u}_0\| + \|(\mathbf{u}_0 \cdot \nabla) \mathbf{u}_0\| + \|A^{1/2} \mathbf{u}_0\| + \|\mathbf{f}(0)\| \\ &\leq c_\nu \|\mathbf{A}\mathbf{u}_0\| + c \|\mathbf{u}_0\| \|A^{1/2} \mathbf{u}_0\|^2 + \|\mathbf{f}(0)\| \end{aligned} \tag{5.5.4}$$

We begin with a general regularity result, on the perturbed solution, useful for the establishment of an error bound associated with the finite elements discretization.

Theorem 5.5.1. *Under the assumptions (A1) and $C\varepsilon \leq 1$, there is a constant $k_0 > 0$ such that if $k_0\Delta t \leq 1$, then the following estimates hold*

$$\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^m\|^2 + \Delta t \sum_{n=1}^m (\|d_t\mathbf{u}_\varepsilon^n\|^2 + \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 + \|p_\varepsilon^n\|^2 + a\|\mathbf{u}_\varepsilon^n\|_{L^{\alpha+2}}^{\alpha+2}) \leq C, \quad (5.5.5)$$

$$\|d_t\mathbf{u}_\varepsilon^m\|^2 + \|A_\varepsilon\mathbf{u}_\varepsilon^m\|^2 + \|p_\varepsilon^m\|_1^2 + \Delta t \sum_{n=1}^m \|A_\varepsilon^{1/2}d_t\mathbf{u}_\varepsilon^n\|^2 \leq C, \quad (5.5.6)$$

for all $1 \leq m \leq N$.

Proof. Taking $(\mathbf{v}, q) = 2(\mathbf{u}_\varepsilon^n, p_\varepsilon^n)\Delta t$ in (5.5.3), using (5.5.3) and the relation

$$2(\mathbf{u} - \mathbf{v}, \mathbf{u}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2, \quad \forall \mathbf{u}, \mathbf{v} \in Y \quad (5.5.7)$$

to have

$$\|\mathbf{u}_\varepsilon^n\|^2 - \|\mathbf{u}_\varepsilon^{n-1}\|^2 + 2\nu\|A^{1/2}\mathbf{u}_\varepsilon^n\|^2\Delta t + a\Delta t \sum_{n=1}^m \|\mathbf{u}_\varepsilon^n\|_{L^{\alpha+2}}^{\alpha+2} \leq 2\|\mathbf{f}(t_n)\| \|\mathbf{u}_\varepsilon^n\| \Delta t$$

We sum this inequality from 1 to m and using (5.2.9) and the Young inequality, to obtain

$$\|\mathbf{u}_\varepsilon^m\|^2 + \nu\Delta t \sum_{n=1}^m \|A^{1/2}\mathbf{u}_\varepsilon^n\|^2 + a\Delta t \sum_{n=1}^m \|\mathbf{u}_\varepsilon^n\|_{L^{\alpha+2}}^{\alpha+2} \leq \|\mathbf{u}_0\|^2 + \nu^{-1}c\Delta t \sum_{n=1}^m \|\mathbf{f}(t_n)\|^2 \leq C \quad (5.5.8)$$

We can establish from (5.5.3) that

$$d_t\mathbf{u}_\varepsilon^n + \nu A_\varepsilon\mathbf{u}_\varepsilon^n + B(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n) + a|\mathbf{u}_\varepsilon^n|^\alpha\mathbf{u}_\varepsilon^n = \mathbf{f}(t_n) \quad (5.5.9)$$

Taking the scalar product of (5.5.9) with $(\nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)\Delta t$ in Y and using the relation

$$2(A_\varepsilon^{1/2}(\mathbf{u} - \mathbf{v}), A_\varepsilon^{1/2}\mathbf{u}) = \|A_\varepsilon^{1/2}\mathbf{u}\|^2 - \|A_\varepsilon^{1/2}\mathbf{v}\|^2 + \|A_\varepsilon^{1/2}(\mathbf{u} - \mathbf{v})\|^2, \quad \forall \mathbf{u}, \mathbf{v} \in X \quad (5.5.10)$$

we obtain

$$\begin{aligned} & \|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2 - \|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^{n-1}\|^2 + \nu^{-1}\|d_t\mathbf{u}_\varepsilon^n\|^2\Delta t + \nu\|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2\Delta t \\ & + c_\alpha(\mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)\Delta t + \tilde{b}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)\Delta t \\ & = (\mathbf{f}(t_n), \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)\Delta t \end{aligned} \quad (5.5.11)$$

Using (5.2.11)-(5.2.12), we get

Firstly

$$\begin{aligned}
& |\tilde{b}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)| \\
& \leq c\|A^{1/2}\mathbf{u}_\varepsilon^n\| \|\mathbf{u}_\varepsilon^n\|_{L^\infty} \|\nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n\| \\
& \leq \frac{1}{6\nu} \|d_t\mathbf{u}_\varepsilon^n\|^2 + \frac{\nu}{6} \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 + \nu^{-1}c\|\mathbf{u}_\varepsilon^n\|^2 \|A^{1/2}\mathbf{u}_\varepsilon^n\|^2 \|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2
\end{aligned}$$

Secondly

$$\begin{aligned}
|c_\alpha(\mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)| & \leq \frac{1}{6\nu} \|d_t\mathbf{u}_\varepsilon^n\|^2 + \frac{\nu}{6} \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 \\
& + \nu^{-1}c\|A^{1/2}\mathbf{u}_\varepsilon^n\|^2 \|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2
\end{aligned}$$

Finally

$$|(\mathbf{f}(t_n), \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n)| \leq \frac{1}{6\nu} \|d_t\mathbf{u}_\varepsilon^n\|^2 + \frac{\nu}{6} \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 + \nu^{-1}c\|\mathbf{f}(t_n)\|^2$$

Combining these inequalities with (5.5.11) to obtain

$$\begin{aligned}
\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2 - \|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^{n-1}\|^2 + \frac{1}{2}(\nu^{-1}\|d_t\mathbf{u}_\varepsilon^n\|^2 + \nu\|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2)\Delta t \\
+ c_\alpha(\mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n) \leq d_n\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2\Delta t + \nu^{-1}c\|\mathbf{f}(t_n)\|^2\Delta t
\end{aligned}$$

with $d_n = \nu^{-1}c(1 + \nu^{-2}\|\mathbf{u}_\varepsilon^n\|^2)\|A^{1/2}\mathbf{u}_\varepsilon^n\|^2$. Summing this inequality from 1 to m and observing that $A_\varepsilon^{1/2}\mathbf{u}_0 = A^{1/2}\mathbf{u}_0$, we get

$$\begin{aligned}
\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^m\|^2 + \frac{1}{2}\Delta t \sum_{n=1}^m (\nu^{-1}\|d_t\mathbf{u}_\varepsilon^n\|^2 + \nu\|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2) + \Delta t \sum_{n=1}^m c_\alpha(\mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n) \\
\leq \|A^{1/2}\mathbf{u}_0\|^2 + \Delta t \sum_{n=1}^m d_n\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2 + \nu^{-1}c\Delta t \sum_{n=1}^m \|\mathbf{f}(t_n)\|^2 \\
\leq C + \Delta t \sum_{n=1}^m d_n\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^2
\end{aligned} \tag{5.5.12}$$

If we take Δt such that $d_n\Delta t \leq \frac{1}{2}$, by applying lemma 2.4 to (5.5.12) we infer

$$\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^m\|^2 + \frac{1}{2}\Delta t \sum_{n=1}^m (\nu^{-1}\|d_t\mathbf{u}_\varepsilon^n\|^2 + \nu\|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2) + \Delta t \sum_{n=1}^m c_\alpha(\mathbf{u}_\varepsilon^n, \nu^{-1}d_t\mathbf{u}_\varepsilon^n + A_\varepsilon\mathbf{u}_\varepsilon^n) \leq C \exp(2\Delta t \sum_{n=1}^m d_n) \tag{5.5.13}$$

for all $1 \leq m \leq N$. By using (5.2.12), (5.5.8), and (5.5.12) there exists a constant $C_0 > 0$ such that

$$2\Delta t \sum_{n=1}^m d_n \leq C \text{ and } d_m = \nu^{-1}c(1 + \nu^{-2}\|\mathbf{u}_\varepsilon^m\|^2)\|A^{1/2}\mathbf{u}_\varepsilon^m\|^2 \leq \frac{1}{2}C_0 \tag{5.5.14}$$

for all $1 \leq m \leq N$.

Next, using (5.5.3) to obtain

$$\begin{aligned} & (d_{tt}\mathbf{u}_\varepsilon^n, \mathbf{v}) + a_\varepsilon(d_t\mathbf{u}_\varepsilon^n, \mathbf{v}) + \tilde{b}(d_t\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, \mathbf{v}) + \tilde{b}(\mathbf{u}_\varepsilon^{n-1}, d_t\mathbf{u}_\varepsilon^n, \mathbf{v}) \\ & + a(d_t(|\mathbf{u}_\varepsilon^n|^\alpha \mathbf{u}_\varepsilon^n), \mathbf{v}) = \left(\frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} \mathbf{f}_t(t) dt, \mathbf{v} \right), \quad \forall v \in X \end{aligned} \quad (5.5.15)$$

We take $v = 2d_t\mathbf{u}_\varepsilon^n \Delta t$ in (5.5.15) and we use (5.2.5) and (5.5.7) to obtain

$$\begin{aligned} & \|d_t\mathbf{u}_\varepsilon^n\|^2 - \|d_t\mathbf{u}_\varepsilon^{n-1}\|^2 + 2\nu \|A_\varepsilon^{1/2} d_t\mathbf{u}_\varepsilon^n\|^2 \Delta t + 2\tilde{b}(d_t\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, d_t\mathbf{u}_\varepsilon^n) \Delta t \\ & + 2a(d_t(|\mathbf{u}_\varepsilon^n|^\alpha \mathbf{u}_\varepsilon^n), d_t\mathbf{u}_\varepsilon^n) \Delta t \leq 2 \left(\int_{t_{n-1}}^{t_n} \mathbf{f}_t(t) dt, d_t\mathbf{u}_\varepsilon^n \right) \end{aligned} \quad (5.5.16)$$

Since $2a(d_t(|\mathbf{u}_\varepsilon^n|^\alpha \mathbf{u}_\varepsilon^n), d_t\mathbf{u}_\varepsilon^n) \geq 0$ thanks to (5.2.13). We use (5.2.9) and (5.2.12) to get

Firstly

$$\begin{aligned} 2|\tilde{b}(d_t\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n, d_t\mathbf{u}_\varepsilon^n)| & \leq c \|d_t\mathbf{u}_\varepsilon^n\| \|A_\varepsilon^{1/2} d_t\mathbf{u}_\varepsilon^n\| \|A_\varepsilon^{1/2} \mathbf{u}_\varepsilon^n\| \\ & + c \|d_t\mathbf{u}_\varepsilon^n\|^{1/2} \|A_\varepsilon^{1/2} d_t\mathbf{u}_\varepsilon^n\|^{3/2} \|A_\varepsilon^{1/2} \mathbf{u}_\varepsilon^n\|^{1/2} \|\mathbf{u}_\varepsilon^n\|^{1/2} \\ & \leq \frac{\nu}{3} \|A_\varepsilon d_t\mathbf{u}_\varepsilon^n\|^2 + d_n \|d_t\mathbf{u}_\varepsilon^n\|^2 \end{aligned}$$

Secondly

$$\left| \left(\int_{t_{n-1}}^{t_n} \mathbf{f}_t(t) dt, d_t\mathbf{u}_\varepsilon^n \right) \right| \leq \frac{\nu}{3} \|A_\varepsilon d_t\mathbf{u}_\varepsilon^n\|^2 + \nu^{-1} c \int_{t_{n-1}}^{t_n} \|\mathbf{f}(t)\|^2 dt$$

Adding these inequalities and replacing them in (5.5.16) we have

$$\begin{aligned} & \|d_t\mathbf{u}_\varepsilon^n\|^2 - \|d_t\mathbf{u}_\varepsilon^{n-1}\|^2 + \nu \|A_\varepsilon^{1/2} d_t\mathbf{u}_\varepsilon^n\|^2 \Delta t \\ & \leq (\nu^{-1} c + d_n) \|d_t\mathbf{u}_\varepsilon^n\|^2 \Delta t + \nu^{-1} c \int_{t_{n-1}}^{t_n} \|\mathbf{f}(t)\|^2 dt \end{aligned} \quad (5.5.17)$$

Summing this inequality from 1 to m and using (5.5.4), (5.5.13) and (5.5.14) we obtain

$$\|d_t\mathbf{u}_\varepsilon^m\|^2 + \nu \Delta t \sum_{n=1}^m \|A_\varepsilon^{1/2} d_t\mathbf{u}_\varepsilon^n\|^2 \leq C, \quad \forall 1 \leq m \leq N \quad (5.5.18)$$

Finally, we can establish from (5.5.3), (5.5.15) and the inf sup condition [?] that

$$\nu \|A_\varepsilon \mathbf{u}_\varepsilon^n\| \leq \|d_t\mathbf{u}_\varepsilon^n\| + \|\tilde{B}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n)\| + \| |\mathbf{u}_\varepsilon^n|^\alpha \mathbf{u}_\varepsilon^n \| + \|\mathbf{f}(t_n)\| \quad (5.5.19)$$

$$\|p_\varepsilon^n\|_1 \leq c \|d_t\mathbf{u}_\varepsilon^n\| + \nu c \|A_\varepsilon \mathbf{u}_\varepsilon^n\| + c \|\tilde{B}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n)\| + c \| |\mathbf{u}_\varepsilon^n|^\alpha \mathbf{u}_\varepsilon^n \| + c \|\mathbf{f}(t_n)\| \quad (5.5.20)$$

Applying (5.2.11)-(5.2.12), we get

$$\begin{aligned} \|\tilde{B}(\mathbf{u}_\varepsilon^n, \mathbf{u}_\varepsilon^n)\| &\leq c\|A^{1/2}\mathbf{u}_\varepsilon^n\| \|\mathbf{u}_\varepsilon^n\|_{L^\infty} \leq c\|A^{1/2}\mathbf{u}_\varepsilon^n\| \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^{1/2} \|\mathbf{u}_\varepsilon^n\|^{1/2} \\ &\leq \frac{\nu}{4} \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 + \nu^{-1}c \|\mathbf{u}_\varepsilon^n\| \|A^{1/2}\mathbf{u}_\varepsilon^n\|^2 \end{aligned}$$

and

$$\|\mathbf{u}_\varepsilon^n\|^\alpha \|\mathbf{u}_\varepsilon^n\| \leq c\|A_\varepsilon\mathbf{u}_\varepsilon^n\|$$

Replacing these inequalities in (5.5.19) and (5.5.20) and using (5.2.12), we obtain

$$\nu\|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2 \leq \nu^{-1}\|d_t\mathbf{u}_\varepsilon^n\|^2 + \nu^{-3}c\|\mathbf{u}_\varepsilon^n\|^2\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^4 + \nu^{-1}c\|f(t_n)\|^2 \quad (5.5.21)$$

$$\|p_\varepsilon^n\|_1^2 \leq c(\|d_t\mathbf{u}_\varepsilon^n\|^2 + \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2) + c\nu^{-2}\|\mathbf{u}_\varepsilon^n\|^2\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^4 + \|f(t_n)\|^2 \quad (5.5.22)$$

$$\begin{aligned} \Delta t \sum_{n=1}^m \|p_\varepsilon^n\|_1^2 &\leq c\Delta t \sum_{n=1}^m (\|d_t\mathbf{u}_\varepsilon^n\|^2 + \|A_\varepsilon\mathbf{u}_\varepsilon^n\|^2) \\ &\quad + c\Delta t \sum_{n=1}^m (\nu^{-2}\|\mathbf{u}_\varepsilon^n\|^2\|A_\varepsilon^{1/2}\mathbf{u}_\varepsilon^n\|^4 + \|f(t_n)\|^2) \end{aligned} \quad (5.5.23)$$

Combining the inequalities (5.5.21)-(5.5.23) with (5.5.13)-(5.5.14) and (5.5.18), we end the proof of the theorem 5.5.1 \square

5.5.2 Finite element penalty method

Let us consider a real positive parameter h . We denote by (X_h, M_h) the finite element pair of (X, M) . It is a discretization of these domain into triangles τ .

We introduce the L^2 -orthogonal projections: $\pi_h : M \rightarrow M_h$ defined by

$$(\pi_h q, q_h) = (q, q_h), \quad \forall q \in M, \quad q_h \in M_h$$

We shall make standing the assumption for the finite element pair (X_h, M_h)

$$\text{There exists a mapping } \mathbf{r}_h : D(A) \cap X \rightarrow X_h \text{ such that} \quad (\text{A2})$$

$$(\text{div}(\mathbf{u} - \mathbf{r}_h\mathbf{u}), q_h) = 0, \quad \forall q \in M, \quad q_h \in M_h \quad (5.5.24)$$

$$\|A^{1/2}(\mathbf{u} - \mathbf{r}_h\mathbf{u})\| \leq ch\|Au\|, \|p - \pi_h p\| \leq ch\|p\|_1, \quad \forall p \in H^1(\Omega) \cap M \quad (5.5.25)$$

Moreover, we suppose the inverse inequality

$$\|A^{1/2}v_h\| \leq ch^{-1}\|\mathbf{v}_h\|, \quad \forall \mathbf{v}_h \in X_h \quad (5.5.26)$$

and the inf-sup condition

$$\|q\| \leq c \sup_{\mathbf{v}_h \in X_h} \frac{(div(\mathbf{v}_h), q_h)}{\|A^{1/2}\mathbf{v}_h\|}, \quad \forall q_h \in M_h \quad (5.5.27)$$

Let us introduce the following finite element discretization of the formulation (5.5.3)

$$\begin{aligned} & (d_t \mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) + a(\mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) - d(\mathbf{v}_h, p_{\varepsilon h}^n) + d(\mathbf{u}_{\varepsilon h}^n, q_h) + \varepsilon(p_{\varepsilon h}^n, q_h) + \tilde{b}(\mathbf{u}_{\varepsilon h}^n, \mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) \\ & + c_\alpha(\mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) = (f(t_n), \mathbf{v}_h), \quad \forall (\mathbf{v}_h, q_h) \in X_h \times M_h \end{aligned} \quad (5.5.28)$$

where $u_{\varepsilon h}^0 = r_h \mathbf{u}_0$, $p_{\varepsilon h}^0 = 0$, and $\{\mathbf{u}_{\varepsilon h}^n\}_{n=1}^N \subset X_h$ and $\{p_{\varepsilon h}^n\}_{n=1}^N \subset M_h$ are the finite element approximations of $\{u_\varepsilon^n\}_{n=1}^N \subset X$ and $\{p_\varepsilon^n\}_{n=1}^N \subset M$ respectively.

Theorem 5.5.2. *Suppose (A1) and (A2) and $\varepsilon c_0 \leq 1$, $\Delta t k_0 \leq 1$ are valid. Then the following proprieties holds*

$$\|\mathbf{u}_{\varepsilon h}^m\|^2 + \Delta t \sum_{n=1}^m \|A^{1/2} u_{\varepsilon h}^n\|^2 + a \Delta t \sum_{n=1}^m \|\mathbf{u}_{\varepsilon h}^n\|_{L^{\alpha+2}}^{\alpha+2} \leq C, \quad (5.5.29)$$

$$\begin{aligned} \|A^{1/2} \mathbf{u}_{\varepsilon h}^m\|^2 + \Delta t \sum_{n=1}^m \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 & \leq C + Ch^{-2} \Delta t \sum_{n=1}^m \|A^{1/2}(\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n)\|^2 \|\mathbf{u}_{\varepsilon h}^n\|^2, \\ & + Ch^{-2} \Delta t \sum_{n=1}^m \|\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2, \end{aligned} \quad (5.5.30)$$

for all $1 \leq m \leq N$.

Proof. In a similar manner as the estimation (5.5.8) in the Theorem 5.5.1, we can derive (5.5.29).

Next we can establish from (5.5.28) that

$$\begin{aligned} & (d_t \mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) + a(\mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) - d(\mathbf{v}_h, p_{\varepsilon h}^n) + d(d_t \mathbf{u}_{\varepsilon h}^n, q_h) + \varepsilon(d_t p_{\varepsilon h}^n, q_h) + \tilde{b}(\mathbf{u}_{\varepsilon h}^n, \mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) \\ & + c_\alpha(\mathbf{u}_{\varepsilon h}^n, \mathbf{v}_h) = (f(t_n), \mathbf{v}_h), \quad \forall (\mathbf{v}_h, q_h) \in X_h \times M_h \end{aligned} \quad (5.5.31)$$

Setting $(\mathbf{v}_h, q_h) = 2(d_t \mathbf{u}_{\varepsilon h}^n, p_{\varepsilon h}^n) \Delta t$ in (5.5.31) and using (5.5.10), we get

$$\begin{aligned} & 2\|d_t \mathbf{u}_{\varepsilon h}^n\|^2 \Delta t + \nu \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 - \nu \|A^{1/2} u_{\varepsilon h}^{n-1}\|^2 + \varepsilon (\|p_{\varepsilon h}^n\|^2 - \|p_{\varepsilon h}^{n-1}\|^2) \\ & + 2c_\alpha(\mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n) \Delta t + 2\tilde{b}(\mathbf{u}_{\varepsilon h}^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n) \Delta t = 2(f(t_n), d_t \mathbf{u}_{\varepsilon h}^n) \Delta t \end{aligned} \quad (5.5.32)$$

Applying (5.2.9)-(5.2.11), (5.2.5) and (5.5.26), we obtain

$$2|\tilde{b}(\mathbf{u}_{\varepsilon h}^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)| \leq 2|\tilde{b}(\mathbf{u}_{\varepsilon h}^n - \mathbf{u}_\varepsilon^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)| + 2|\tilde{b}(\mathbf{u}_\varepsilon^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)|$$

$$\begin{aligned}
2|\tilde{b}(\mathbf{u}_\varepsilon^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)| &\leq c \|A \mathbf{u}_\varepsilon^n\| \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\| \|d_t \mathbf{u}_{\varepsilon h}^n\| \\
&\leq \frac{1}{4} \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 + c \|A \mathbf{u}_\varepsilon^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 \\
2|\tilde{b}(\mathbf{u}_{\varepsilon h}^n - \mathbf{u}_\varepsilon^n, \mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)| &\leq c \|\mathbf{u}_{\varepsilon h}^n - \mathbf{u}_\varepsilon^n\|_{L^4} \|\mathbf{u}_{\varepsilon h}^n\|_{L^4} \|A^{1/2} d_t \mathbf{u}_{\varepsilon h}^n\| \\
&\quad + c \|A^{1/2} (\mathbf{u}_{\varepsilon h}^n - \mathbf{u}_\varepsilon^n)\|^2 \|\mathbf{u}_{\varepsilon h}^n\|_{L^4} \|A^{1/2} d_t \mathbf{u}_{\varepsilon h}^n\|_{L^4} \\
&\leq \frac{1}{4} \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 + ch^{-2} \|A^{1/2} (\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n)\|^2 \|\mathbf{u}_{\varepsilon h}^n\|^2 \\
&\quad + ch^{-2} \|\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 \\
2|c_\alpha(\mathbf{u}_{\varepsilon h}^n, d_t \mathbf{u}_{\varepsilon h}^n)| &\leq \frac{1}{4} \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 + c \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 \\
2|(\mathbf{f}(t_n), d_t \mathbf{u}_{\varepsilon h}^n)| &\leq \frac{1}{4} \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 + c \|\mathbf{f}(t_n)\|^2
\end{aligned}$$

Replacing these inequalities in (5.5.32). with some sample calculation we obtain

$$\begin{aligned}
\nu \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 - \nu \|A^{1/2} \mathbf{u}_{\varepsilon h}^{n-1}\|^2 + \|d_t \mathbf{u}_{\varepsilon h}^n\|^2 \Delta t + \varepsilon (\|p_{\varepsilon h}^n\|^2 - \|p_{\varepsilon h}^{n-1}\|^2) \\
\leq c \|A \mathbf{u}_\varepsilon^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 \Delta t + ch^{-2} \|A^{1/2} (\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n)\|^2 \|\mathbf{u}_{\varepsilon h}^n\|^2 \Delta t \\
+ \|\mathbf{u}_\varepsilon^n - \mathbf{u}_{\varepsilon h}^n\|^2 \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\| \Delta t + c \|A^{1/2} \mathbf{u}_{\varepsilon h}^n\|^2 + c \|\mathbf{f}(t_n)\|^2
\end{aligned} \tag{5.5.33}$$

Summing these inequality from 1 to m and using Theorem 5.5.1 and (5.2.12) we get (5.5.30) \square

In order to establish the error estimates of the finite element penalty method, we introduce the following Galerkin projection $R_h: (X, M) \rightarrow X_h$, $Q_h: (X, M) \rightarrow M_h$ defined by

$$\begin{aligned}
a(\mathbf{u} - R_h(\mathbf{u}, p), \mathbf{v}_h) - d(\mathbf{v}_h, p - Q_h(\mathbf{u}, p)) + d(\mathbf{u} - R_h(\mathbf{u}, p), q_h) + \varepsilon(p - Q_h(\mathbf{u}, p), q_h) = 0 \\
\forall (\mathbf{v}_h, q_h) \in X_h \times M_h,
\end{aligned} \tag{5.5.34}$$

for all $(\mathbf{u}, p) \in X \times M$ with $\operatorname{div} \mathbf{u} + \varepsilon p = 0$.

Lemma 5.5.1. *Under the assumptions (A1), (A2), $c_0 \varepsilon \leq 1$ and $k_0 \Delta t \leq 1$, the Galerkin projection (R_h, Q_h) verifies*

$$\begin{aligned}
\bullet \quad \|\mathbf{u} - R_h(\mathbf{u}, p)\| + h \|A^{1/2} (\mathbf{u} - R_h(\mathbf{u}, p))\| + h \|p - Q_h(\mathbf{u}, p)\| \\
\leq c_\nu h (\|A^{1/2} \mathbf{u}\| + \|p\|), \quad \forall (\mathbf{u}, p) \in X \times M
\end{aligned} \tag{5.5.35}$$

$$\begin{aligned}
\bullet \quad \|\mathbf{u} - R_h(\mathbf{u}, p)\| + h \|A^{1/2} (\mathbf{u} - R_h(\mathbf{u}, p))\| + h \|p - Q_h(\mathbf{u}, p)\| \\
\leq C_\nu h^2 (\|A \mathbf{u}\| + \|p\|_1), \quad \forall (\mathbf{u}, p) \in D(A) \times H^1(\Omega) \cap M
\end{aligned} \tag{5.5.36}$$

with $\text{div} \mathbf{u} + \varepsilon p = 0$, where c_ν is a general positive constant depending only on Ω and ν .

Proof. The Galerkin projection has its stability from (5.5.27) and (5.5.34), and we have

$$\nu \|A^{1/2} R_h(\mathbf{u}, p)\|^2 \leq \frac{\nu}{2} \|A^{1/2} R_h(\mathbf{u}, p)\|^2 + c_\nu (\|A^{1/2} \mathbf{u}\|^2 + \|p\|^2) \quad (5.5.37)$$

$$\|Q_h(\mathbf{u}, p)\| \leq c_\nu (\|A^{1/2} R_h(\mathbf{u}, p)\| + \|A^{1/2} \mathbf{u}\| + \|p\|) \quad (5.5.38)$$

Using (5.5.37)-(5.5.38) and the triangular inequality, yields to

$$\|A^{1/2}(\mathbf{u} - R_h(\mathbf{u}, p))\| + \|p - Q_h(\mathbf{u}, p)\| \leq c_\nu (\|A^{1/2} \mathbf{u}\| + \|p\|) \quad (5.5.39)$$

for all $(\mathbf{u}, p) \in X \times M$ with $\text{div} u + \varepsilon p = 0$.

Let consider the dual problem: find $(\Phi, \Psi) \in X \times M$ such that for all $(\mathbf{v}, q) \in X \times M$

$$a(\mathbf{v}, \Phi) + d(\mathbf{v}, \Psi) - d(\Phi, q) + \varepsilon(q, \Psi) = (\mathbf{v}, u - R_h(\mathbf{u}, p)) \quad (5.5.40)$$

By using the lemma 4.2, the problem (5.5.40) has a unique solution (Φ, Ψ) satisfying

$$\|A\mathbf{u}\| + \|\Psi\|_1 \leq c_\nu \|\mathbf{u} - R_h(\mathbf{u}, p)\| \quad (5.5.41)$$

Taking $(\mathbf{v}, q) = (\mathbf{u} - R_h(\mathbf{u}, p), p - Q_h(\mathbf{u}, p))$ in (5.5.40) and $(\mathbf{v}_h, q_h) = (\mathbf{r}_h \Phi, \pi_h \Psi)$ in (5.5.34) and using (5.5.25) and (5.5.41) we infer that

$$\begin{aligned} |u - R_h(\mathbf{u}, p)|^2 &= a(\mathbf{u} - R_h(\mathbf{u}, p), \Phi - \mathbf{r}_h \Phi) + d(\mathbf{u} - R_h(\mathbf{u}, p), \Psi - \pi_h \Psi) \\ &\quad - d(\Phi - \mathbf{r}_h \Phi, p - Q_h(\mathbf{u}, p)) + \varepsilon(p - Q_h(\mathbf{u}, p), \Psi - \pi_h \Psi) \\ &\leq c_\nu h (\|A^{1/2}(\mathbf{u} - R_h(\mathbf{u}, p))\| + \|p - Q_h(\mathbf{u}, p)\|) (\|A\Phi\| + \|\psi\|_1) \\ &\leq c_\nu h (\|A^{1/2}(\mathbf{u} - R_h(\mathbf{u}, p))\| + \|p - Q_h(\mathbf{u}, p)\|) \|\mathbf{u} - R_h(\mathbf{u}, p)\| \end{aligned} \quad (5.5.42)$$

The estimation (5.5.35) is a consequence of (5.5.42) and (5.5.39).

We consider $(\mathbf{u}, p) \in D(A) \times (H^1(\Omega) \cap M)$ with $\text{div} u + \varepsilon p = 0$. based on (5.5.24)-(5.5.25) and (5.5.34), we deduce that

$$\begin{aligned} \|A^{1/2}(\mathbf{r}_h \mathbf{u} - R_h(\mathbf{u}, p))\| + \|\pi_h p - Q_h(\mathbf{u}, p)\| \\ \leq c_\nu (\|A^{1/2}(\mathbf{u} - \mathbf{r}_h \mathbf{u})\| + \|\pi_h p - p\|) \\ \leq c_\nu h (\|A\mathbf{u}\| + \|p\|_1) \end{aligned}$$

Therefore, using the triangles inequality and (5.5.25) yields to

$$\|A^{1/2}(\mathbf{u} - R_h(\mathbf{u}, p))\| + \|p - Q_h(\mathbf{u}, p)\| \leq c_\nu h (\|A\mathbf{u}\| + \|p\|_1) \quad (5.5.43)$$

Combining (5.5.43) and (5.5.42) to obtain

$$\|\mathbf{u} - R_h(\mathbf{u}, p)\| \leq c_\nu h^2 (\|A\mathbf{u}\| + \|p\|_1) \quad (5.5.44)$$

Combining (5.5.43)-(5.5.44), we obtain the estimation (5.5.36) \square

5.6 Numerical resolution

5.6.1 Numerical methods

In this subsection we study the approach of spectral projected gradient for solving the shape optimization problem. This method use the projection into a closed and convex Ω_0 subset of \mathbb{R}^4 of all the points $y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$ satisfying

$$\begin{aligned} \frac{1}{4} 1.213 &\leq y_1, y_3 \leq \frac{3}{4} 1.213 \\ 0 &\leq y_2, y_4 \leq \frac{1}{4} 0.97 \end{aligned} \quad (5.6.1)$$

and for a comfortable fish passage other constraints are added as

$$y_3 - y_1 \geq d_1 = 0.1 \quad (5.6.2)$$

$$y_2 - y_4 \geq d_2 = 0.05$$

We denote $a_1 = a_3 = \frac{1}{4}1.213$, $b_1 = b_3 = \frac{3}{4}1.213$, $a_2 = a_4 = 0$, $b_2 = b_4 = \frac{1}{4}0.97$. Then the admissible set Ω is defined as

$$\Omega_0 = \{y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4 : \begin{aligned} &a_i \leq y_i \leq b_i, \quad i = 1, \dots, 4, \\ &y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{aligned}\} \quad (5.6.3)$$

The optimization problem takes the form

$$\min_{y \in \Omega_0} J(y). \quad (5.6.4)$$

The optimal position $y \in \Omega_0$ is found through the following algorithm using the computation of the cost function gradient and the projection functions.

- a) Begin with $\bar{y} \in \Omega$ and $\varepsilon > 0$ a positive tolerance.
- b) Compute the projection $d = P_{\Omega_0}(\bar{y} - \eta \nabla J(\bar{y})) - \bar{y}$, where ∇J is the gradient of the cost function with η is a positive constant chosen according to particular conditions.
- c) Stop if $d = 0$ (in practice, $\|d\| < \varepsilon$).
- e) Define $\tilde{y} = \bar{y} + \xi d$, with ξ is the descent step. Go to (a) with $\bar{y} = \tilde{y}$.

The value of J in this algorithm is given by the ensuing discrete formula

$$\bar{\Phi}(y) = \frac{\Delta t}{2} \sum_{n=1}^N \sum_{e \in T_i} \left[\int_e \|\mathbf{u}^{n,i} - \vec{v}\|^2 + \alpha \int_e |\text{curl}(\mathbf{u}^{n,i})|^2 \right]. \quad (5.6.5)$$

$y = P_{\Omega_0}(z)$ is the projection of $z \in \mathbb{R}^4$ in Ω onto Ω_0 . It is computed by minimizing a quadratic function of the distance of z to Ω as follow

$$\min_{y \in \Omega} \frac{1}{2} \|y - z\|^2 = \min_{y \in \Omega} \frac{1}{2} z^T z - z^T y + \frac{1}{2} y^T y \quad (5.6.6)$$

which is equivalent to

$$\left\{ \begin{array}{l} \min_{(y_1, y_2, y_3, y_4)} \sum_{i=1}^4 \left\{ \frac{1}{2} z_i^2 - z_i y_i + \frac{1}{2} y_i^2 \right\} \\ \text{subject to } a_i \leq y_i \leq b_i, i = 1, \dots, 4 \\ y_3 - y_1 \geq d_1, \quad y_2 - y_4 \geq d_2 \end{array} \right.$$

We separate the previous quadratic problem into two quadratic problems the first one for the abscesses, the second for the ordinates and we write

$$\begin{aligned} \min_{(y_1, y_3)} \quad & \frac{1}{2}(z_1^2 + z_3^2) - (z_1 y_1 + z_3 y_3) + \frac{1}{2}(y_1^2 + y_3^2) \\ \text{subject to} \quad & a_i \leq y_i \leq b_i, i = 1, 3 \\ & y_3 - y_1 \geq d_1, \end{aligned}$$

and

$$\begin{aligned} \min_{(y_2, y_4)} \quad & \frac{1}{2}(z_2^2 + z_4^2) - (z_2 y_2 + z_4 y_4) + \frac{1}{2}(y_2^2 + y_4^2) \\ \text{subject to} \quad & a_i \leq y_i \leq b_i, i = 2, 4 \\ & y_2 - y_4 \geq d_2, \end{aligned}$$

A Karush-Kuhn-Tucker (KKT) technique is used to solve the two quadratic problems above. We begin by the first one and we choose $a = a_1 = a_3$, and $b = b_1 = b_3$. We kept only three constraints and write the optimization problem to be solved as

$$\begin{aligned} \min_{(x_1, x_2)} \quad & c - (l_1 x_1 + l_2 x_2) + \frac{1}{2}(x_1^2 + x_2^2) \\ \text{subject to} \quad & x_2 \leq b, \\ & x_1 \geq a, \\ & x_2 - x_1 \geq d_1, \end{aligned}$$

where ,a, b, d_1 , c, and l_i , $i=1,2$, are real numbers.

We choose the following matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -b \\ a \\ d_1 \end{pmatrix},$$

Then the linear constraints write

$$Ax - v = b, \quad v \geq 0,$$

where $v = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)^T$ is the slack variables. The optimisation problem has an unique solution due to the convexity of the cost function in \mathbb{R}^2 . This solution satisfies the KKT conditions

$$\begin{aligned} v &= Ax - b \\ -l + x &= A^T \lambda, \\ v &\geq 0, \quad \lambda \geq 0, \\ v^T \lambda &= 0, \end{aligned}$$

where $l = (l_1, l_2, l_3)$, and $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is the vector of the Lagrange multipliers associated to the three constraints above. The unique optimal solution writes

$$x = l + A^T \lambda$$

which gives

$$\begin{cases} x_1 = l_1 + \lambda_2 - \lambda_3 \\ x_2 = l_2 - \lambda_1 + \lambda_3 \end{cases}$$

For computing the Lagrange multipliers λ_i , we solve a LCP problem which takes the form

$$\begin{aligned} v &= (-b + Al) + AA^T \lambda \\ v &\geq 0, \quad \lambda \geq 0 \\ v^T \lambda &= 0 \end{aligned}$$

Remark 5.6.1. *The gradient of the cost function is computed using the formula established in section 3.*

5.6.2 Numerical results

In this subsection we present some results about the optimization of a fishway using finite element method investigated in the previous sections to solve the state problem combined with a projected gradient algorithm for the resolution of the optimization problem. The domain fluid is considered as a rectangular channel consisting of ten pools. Inside each

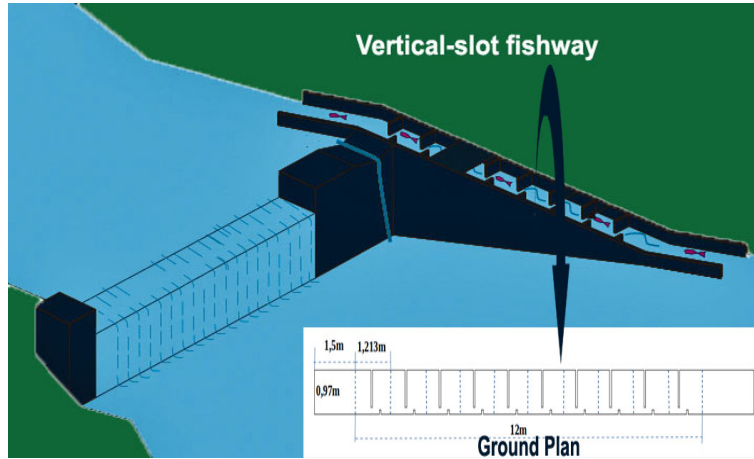


Figure 5.1: Schematic description of fishway structure

pool, two baffles in form of I or L are built and they are perpendicular to the lateral boundary. The length and width of baffles vary according to the nature of the studied fishway structures. In figure 5.1, we give a description of the fishway structure.

We conduct two 3-D tests of fishway design mostly used for the modelling of fishway passage. Firstly, we begin by a 3-D channel with I shaped baffle (Figure 5.2 (right)). Next, we consider a L shaped baffle design (Figure 5.2 (left)). The numerical simulation for the two tests is done with the same physical and numerical conditions. All initial and boundary conditions are taken constants. the inflow velocity is $u_1 = 0.1 m^2 s^{-1}$, the target velocity is $c = 0.8 m.s^{-1}$. The vorticity parameter is $\sigma = 0$. The viscosity is 0.01. The exterior function is $f = 0$. For fish passage with comfortable conditions, we put $d_1 = 0.1$ and $d_2 = 0.05$.

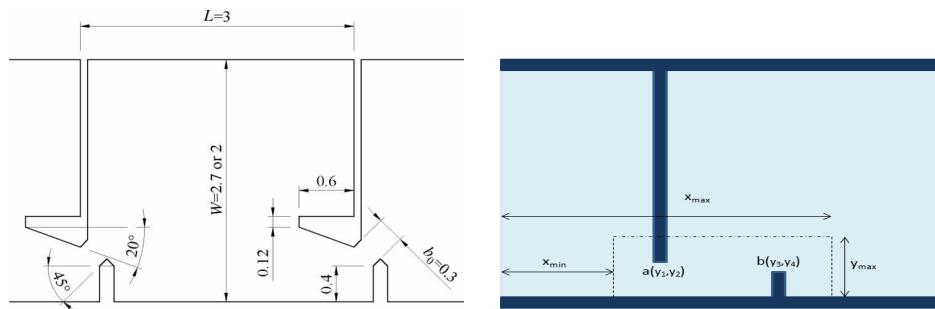


Figure 5.2: Fish pass geometry L shaped baffle (left) and I shaped baffle (right) for the central pool

5.6.2.1 Slot Fishway (3-D "I" shaped baffle-rectangular slots)

The state penalized problem (5.2.42)-(5.2.43) is solved by the finite element method (5.5.28) combined with a spectral projected gradient algorithm for the resolution of the

shape optimization problem (5.4.2).

This test consists to consider a 3-D geometry of pools with two rectangular baffles as shown in the Figure 5.3.

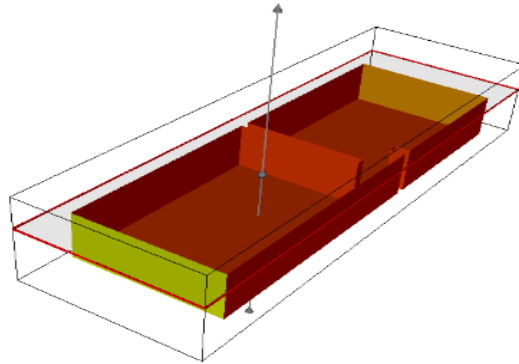


Figure 5.3: Configuration of central pool

The coming results gives a description of the velocity for ten pools related to the natures of the studied shape of fishway structures.

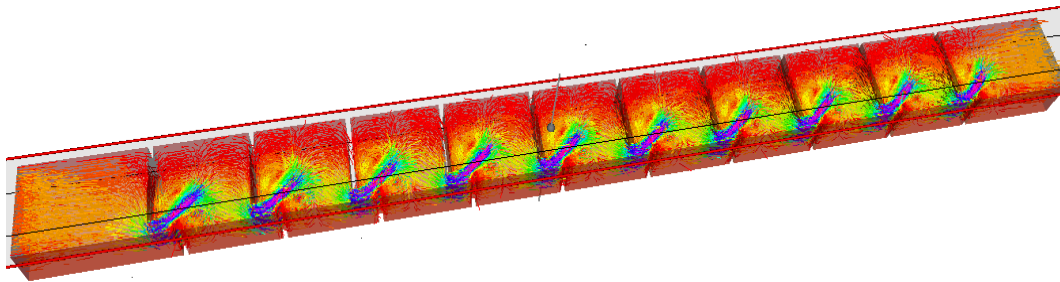


Figure 5.4: Initial and random shape and corresponding velocity for ten pools

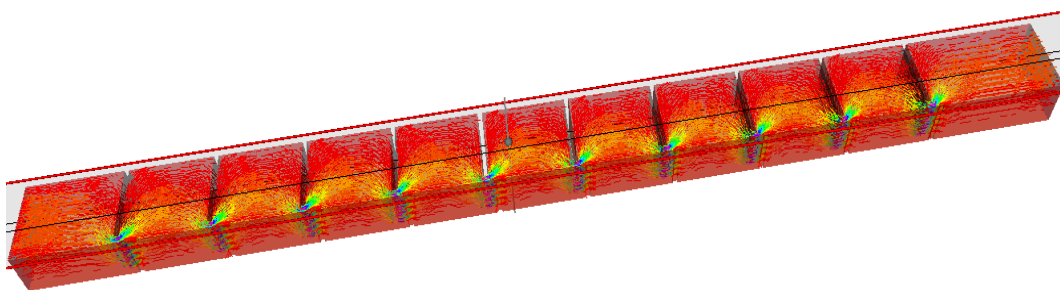


Figure 5.5: Optimal shape and corresponding velocity for ten pools

We observe that the recirculation areas near of slots are removed for the optimal shape (Figure 5.5) compared to the initial random structure (Figure 5.4). The velocity is almost

null in the areas near from the boundaries behind slots for the optimal shape which provides a rest zones for fish before to begin the passage from the current pool to the next one. These results appears clearly in the following examples for the central pool (Figure 5.6 (right))

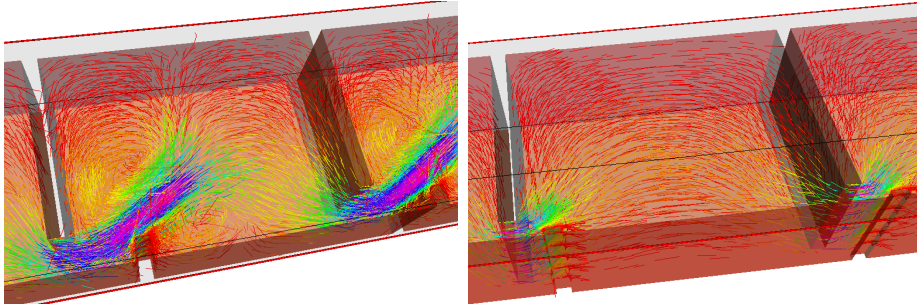


Figure 5.6: Initial and random shape (left) Optimal shape (right) and corresponding velocity for the central pool

5.6.2.2 Slot Fishway (3-D "L" shaped baffle-oblique slots)

In this subsection we give results of numerical simulations in 3-D dimension with a L shape baffle. We consider a channel of ten pools. Each pool has two vertical slots. The small one is oblique and the longest is in the form of L and they are vertical to the lateral boundary. The different values corresponding to the geometry of pools are shown in the figure 5.2) (left). An example of the first pool geometry is given in the Figure 5.7.

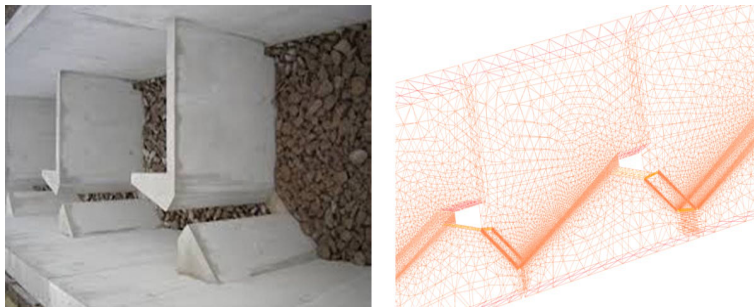


Figure 5.7: 3-D example of fishway with porous medium (left), Mesh of the central pool (right)

The following results represent the velocity for ten pools in the initial random shape (Figure 5.8) and the optimal shape (Figure 5.9). The nature of the studied structure takes part of the minimization of turbulence by the use of the form L of baffles which are considered as a break waves of water.

The velocity is almost close from the uniform target velocity \mathbf{u}_d for the optimal shape

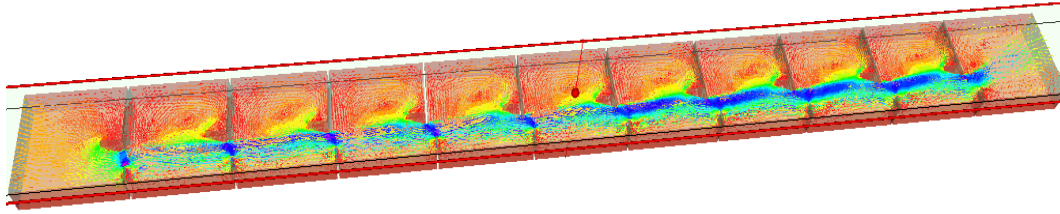


Figure 5.8: Non optimal shape and corresponding velocity for ten pools

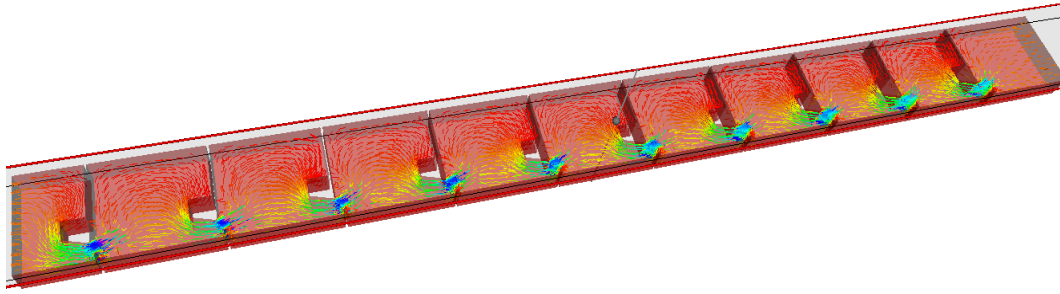


Figure 5.9: Optimal shape and corresponding velocity for ten pools

(Figure 5.9).

We give numerical results for the central pool in the ensuing figures. Turbulence appears clearly in the areas near from slots in the figure 5.10 (left). For the optimal shape (Figure 5.10(right)), the nature of velocity permits a suitable passage for fish and provides for them a rest in the zones where the velocity is almost null.

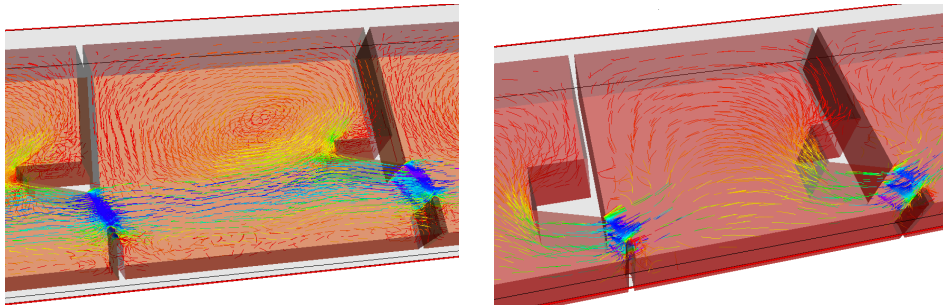


Figure 5.10: Non optimal shape and corresponding velocity for the central pool

5.7 Conclusion

In this work, we have studied the Navier Stokes Forchheimer equations. A penalty method has been introduced to examine different properties of a penalized problem along two parts the first one is the derivation of an error estimation and use it for the global error

estimates found in the second part. A finite element method has been used to discretize the fluid domain and derive an error estimate for the velocity and the pressure in term of norm H^1 and L^2 respectively. Numerical results were presented for Two 3-D examples for fishways structures. The obtained results stated the robustness and efficiency of the adopted technique.

Bibliography

- [1] C.H. Clay,: Design of fishways and other Fish Facilities. Lewis Publishers, CRC Press, Boca Raton, p 248, 1995.
- [2] C. Katopodis, N. Rajaratnam, S. Wu, D. Towell: Denil Fishways of varying. ASCE J. Hydraul. Eng. 123 624–631, 1997.
- [3] C.T. Kelley: Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition. SIAM J. Optim., 10 43–55, 1999.
- [4] J. Cea: RAIRO-Mathematical modelling and numerical analysis, tome 20, number 3, 371–402, 1986.
- [5] J. Moser: A rapidly convergent iteration method and nonlinear partial differential equations. I. Ann. Scuola Norm. Sup. Pisa, number 20, 265-315, 1966.
- [6] L.J. Alvarez-Vazquez, A. Martinez, M.E. Vazquez Mendez, M.A. Vilar: An optimal shape problem related to the realistic design of river fishways, Ecological engineering, 32 293–300,2008.
- [7] J.L. Lions. Optimal Control of Systems governed by partial differential equations, North Holland, 1970.
- [8] J.L Lions. Optimal control of systems governed by partial differential equations. Berlin Heidelberg: Springer-Verlag, 1971
- [9] N. Rajaratnam, G. Van de Vinne, C. Katopodis. Hydraulics of vertical slot fishways. ASCE J. Hydraul. Eng. number 130, 909-917, 1986.
- [10] T. Kato, G. Ponce. Commutator estimates and the Euler and Navier Stockes equations. Comm. Pure Appl. Math., number 31, 89–106, 1988.
- [11] J. Chorda, M. M. Maubourguet, H. Roux, M. Larinier, L. Tarrade, L. David. Two-dimensional free surface flow numerical model for vertical slot fishways. Journal of Hydraulic Research, vol. 48. pp. 141-151. ISSN 1814-2079, 2010.
- [12] A.Matsumura, T. Nishida, The initial value problem of the equations of motion of viscous and heat-conductive gases. Journal of Mathematic of Kyoto University, 1980.

- [13] M. J. Castro, A. Pardo, C. Pares and E. F. Toro, On some fast well-balanced first order solvers for nonconservative systems, *Math. Comp.*, 79(271), 1427–1472, 2010.
- [14] E. Audusse, M.O. Bristeau, and A. Decoene. Numerical simulations of 3D free surface flows by a multilayer Saint-Venant model. *International journal for numerical methods in fluids* (2007).
- [15] G.F Carey, R. Krishnan, Penalty finite element method for the Navier Stokes equations. *Computer methods in applied mechanics and engineering* 42, 183-224, 1984.
- [16] W. Li, W. Wang, Q. Jiu, Existence and uniqueness of the weak solutions for the steady incompressible Navier Stokes equations with damping, *African diaspora journal of mathematics*, volume 12, number 2, pp 57-72, 2011.
- [17] J. Su, Z. Chen, Z. Wang, G. Xi, An adaptive finite element method for shape optimization in stationary incompressible flow with damping, *International journal of numerical analysis and modelling* , series B, volume 5, number 1-2, pp 79-96, 2014.
- [18] M. Louaked, N. Seloula, S. Trabelsi, Approximation of the unsteady Brinkman-Forchheimer equations by the pressure stabilization method, *Numerical Methods for Partial Differential Equations*, July 2017 Wiley.
- [19] Jie Shen, On error estimates of the penalty method for unsteady Navier Stokes equations. *SIAM Journal on Numerical Analysis*, 32(2):386:403, 1995.
- [20] Alexandre Joel Chorin, Numerical solution for the Navier Stokes equations. *Mathematics of computation*, 22(104):745-762, 1968.
- [21] Alexandre Joel Chorin, On the convergence of discrete approximations to the Navier Stokes equations. *Mathematics of computation*, 23(106):341-353, 1969.
- [22] G.P. Galdi, An introduction to the Mathematical Theory of the Navier Stokes Equations. Vol. Linearized Steady Problems. Springer-Verlag. New York. Berlin. etc. 1994.
- [23] R. Courant, Variational methods for the solution of problems of equilibrium and vibrations, *Bull. Amer. Math. Soc.* 49 1-23, 1943.
- [24] Y. He, Optimal error estimate of the penalty finite element method for the time-dependent Navier-Stokes equations. *Mathematics of computation*. 74 1201-1016, 2005.
- [25] F. Hecht, O. Pironneau, A. Le Hyaric, K. Ohtsuka, FreeFem++ Manual, available at <http://www.freefem.org>.
- [26] Z. Gao, Y. Ma, H. Zhuang, Optimal shape design for the time-dependent Navier Stokes flow. *International journal for numerical methods in fluids*. 57, 1505-1526, 2008.
- [27] HC Brinkman. A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles. *Applied Scientific Research*, 1(1):27–34, 1949.

- [28] Vivette Girault and Pierre-Arnaud Raviart. Finite element methods for Navier-Stokes equations: theory and algorithms, volume 5. Springer Science and Business Media, 2012.
- [29] Philipp Forchheimer, Wasserbewegung durch boden. *Z. Ver. Deutsch. Ing.*, 45 (1782):1788. 1901.
- [30] Thomas JR Hughes, Leopoldo P Franca, and Marc Balestra. A new finite element formulation for computational fluid dynamics: V. circumventing the Babuska-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering*, 59(1):85–99, 1986.
- [31] JL Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. dunod, gauthier-villars, paris, 1969. Google Scholar, 2002.
- [32] Mohammed Louaked, Nour Seloula, Shuyu Sun, Saber Trabelsi, et al. A pseudocompressibility method for the incompressible Brinkman-Forchheimer equations. *Differential and Integral Equations*, 28(3/4):361–382, 2015.
- [33] Lawrence E Payne, Jong C Song, and Brian Straughan. Continuous dependence and convergence results for Brinkman and Forchheimer models with variable viscosity. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 455, pages 2173–2190. The Royal Society, 1999.
- [34] Jie Shen. On a new pseudocompressibility method for the incompressible Navier-Stokes equations. *Applied numerical mathematics*, 21(1):71–90, 1996.
- [35] Jie Shen. On error estimates of the projection methods for the Navier-Stokes equations: second-order schemes. *Mathematics of Computation of the American Mathematical Society*, 65(215):1039–1065, 1996.
- [36] R. Temam, Une méthode d’approximation de la solution des équations de Navier-Stokes, *Bull. Soc. Math. France*, 98, 115-152, 1968.
- [37] R. Temam, Sur l’approximation de la solution des équations de Navier Stokes par la méthode des pas fractionnaires (ii), *Archive for Rational Mechanics and Analysis*, 33(5): 377-385, 1969.
- [38] Roger Temam. *Navier-Stokes equations*, volume 2. North-Holland Amsterdam, 1984.
- [39] Caidi Zhao and Yuncheng You. Approximation of the incompressible convective Brinkman-Forchheimer equations. *Journal of Evolution Equations*, 12(4):767–788, 2012.
- [40] J.C Umavathi, Odelu ojjela, and K. Vajravelu. Numerical analysis of natural convective flow and heat transfer of nanofluids in a vertical rectangular duct using Darcy-Forchheimer-Brinkman model. *International journal of thermal sciences*. 111:511 524,2017.

- [41] Bixiang Wang and Siyu Lin. Existence of global attractors for the three-dimensional Brinkman-Forchheimer equation. *Mathematical methods in the applied sciences*, 31(12):1479-1495, 2008.
- [42] Andreas Prohl. *Projection and quasi-compressibility methods for solving the incompressible Navier Stokes equations*. Springer, 1997.
- [43] Varga K Kalantarov and Sergey Zelik. Smooth attractors for the Brinkman-Forchheimer equations with fast growing nonlinearities. *arXiv preprint arXiv:1101.4070*, 2011.
- [44] Wanchai Jiajan. *Solution to incompressible Navier Stokes equations by using finite element method*. 2010.
- [45] Karol W Hajduk and James C Robinson. Energy equality for the 3d critical convective Brinkman Forchheimer equations. *Journal of differential equations*, 263(11):7141:7161, 2017.
- [46] Konstantinos Chrysafinos. Error estimates for time-discretizations for the velocity tracking problem for Navier Stokes flows by penalty methods. *Discrete and Continuous Dynamical Systems Series B*, 6(5):1077, 2006.
- [47] Mehdi Badra, Jean-Marie Buchot, and Laetitia Thevenet. Méthode de pénalisation pour le contrôle frontière des équations de Navier Stokes. *Journal Européen des systèmes automatisés-JESA-APII*, 45(7):595, 2011.
- [48] Luigi C Berselli and stefano spirito. On the construction of suitable weak solutions to the 3d Navier Stokes equations in a bounded domain by an artificial compressibility method. *arXiv preprint arXiv:1504.07800*, 2015
- [49] Max D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, philadelphia, USA, 2002.
- [50] Bressan, A. *Hyperbolic Systems of Conservation Laws - The One-dimensional Cauchy Problem*. Oxford University Press. 2000.
- [51] Godlewski, E. and P.A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Applied Mathematical Sciences. Springer. 1996.
- [52] Dubois, F. and P.G. LeFloch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *Journal of Differential Equations*. 1988.
- [53] Lars Hörmander, *Linear Partial Differential Operators*, Springer-Verlag, 1963.
- [54] Toro, E. F., *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer-Verlag. 1999.
- [55] Adhémar Jean Claude Barré de Saint-Venant, *Théorie du mouvement non permanent des eaux, avec application aux crues des rivières et a l'introduction de marées dans leurs lits*,

- compte rendus hebdomadaires des séances de l'académie des sciences, vol. 73, p. 147–154 et 237–240, 1871.
- [56] Albert Cohen, Wolfgang Dahmen, Ron DeVore, Jalal Fadili, Olga Mula, James Nichols. Optimal reduced model algorithms for data-based state estimation. arXiv:1903.07938, 2019
- [57] Lauriane Schneider, Raphaël di Chiara Roupert, Gerhard Schäfer, Philippe Helluy. Highly gravity-driven flow of a NAPL in water-saturated porous media using the discontinuous Galerkin finite-element method with a generalised Godunov scheme. *Computational Geosciences*, Volume 19, Issue 4, pp 855–876, 2015.
- [58] Aymeric Maury, Grégoire Allaire, and François Jouve. Elasto-plastic Shape Optimization Using the Level Set Method, *SIAM J. Control Optim.*, 56(1), 556–581, 2018.
- [59] Xiaofeng Wang, Guang Xi, Zhiheng Wang. Aerodynamic optimization design of centrifugal compressor's impeller with Kriging model. *Proceeding of the institution of Mechanical Engineers Part A-Journal of Power and energy*, 220(A6):585-597, 200
- [60] Wenjing Yan, Yaling He and Yichen Ma. Shape inverse problem for the two-dimensional unsteady Stokes flow. *Numerical Methods for Partial Differential Equations*. 26(3),690-701, 2010.
- [61] Nicolas Focadel Z. Rao and H. Zidani. State-constrained optimal control problems of impulsive differential equations, *Applied Mathematics and Optimization*, 68 (1), pp. 1–19, 2013.
- [62] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems* . Elsevier Science Ltd, 1978.
- [63] O. C. Zienkiewicz . *Finite Element Method in Engineering Science*. McGraw-Hill Inc.,US, 1972.
- [64] B. Mohammadi, O. Pironneau, *Applied Shape optimization for fluids*, Clarendon Press,Oxford, 2001.

Optimisation de forme et applications aux ouvrages hydrauliques : Analyse mathématique et approximation numérique

Résumé : Nous nous intéressons à l'étude théorique et numérique de plusieurs modèles d'écoulement (Saint-Venant, multicouches, milieux poreux stationnaires et non stationnaires) et de leurs applications à l'optimisation de formes de certains ouvrages hydrauliques. Nous explorons le caractère bien posé des systèmes, nous dérivons un système adjoint lié à chaque modèle.

Une méthode de pénalisation est utilisée pour relaxer la contrainte d'incompressibilité.

Nous exprimons le gradient de forme en fonction de la vitesse comme variable d'état, des variables adjointes, et le vecteur unité normal au bord du domaine. Nous adoptons une méthode d'éléments finis discrète pour approcher la solution du problème pénalisé et établissons des estimations a priori afin de prouver la convergence de la solution approchée vers la solution du système non perturbé.

Le problème d'optimisation est implémenté en utilisant la méthode adjointe continue et la méthode d'éléments finis.

Mots clés : Équations des eaux peu profondes, Méthode de multicouches, Équations de milieux poreux, Méthode d'éléments finis.

Shape optimization and applications to hydraulic structures : Mathematical analysis and numerical approximation

Abstract : We are interested in the theoretical and numerical study of different flow models (shallow water system, multilayer, stationary and non stationary porous media) and their applications to the shape optimization of some hydraulic structures.

We explore the well-posedness of the models and derive the adjoint equations related to each system.

A penalty method is used to relax the incompressibility constraint. We express the shape gradient of the cost function in terms of the velocity value as a state variable, the adjoint variables and the unit normal vector to the boundary of the domain.

We propose a discrete finite element method to approximate the solution for the penalized problem and establish a priori estimates to prove the convergence of the approximate solution to the solution of the non perturbed problem. Error estimates for the velocity and the pressure are given.

The optimization procedure is implemented using the continuous adjoint method and the finite element method.

Keywords : Shallow water equations, Multilayer method, Porous media models, Finite element method.