HAL
open science

Etude de l'équation de Korteweg-de Vries en variables lagrangiennes et sa contrôlabilité, stabilisation rapide d'une équation de Schrödinger et méthodes spectrales pour le calcul du contrôle optimal

Ludovick Gagnon

To cite this version:

Ludovick Gagnon. Etude de l'équation de Korteweg-de Vries en variables lagrangiennes et sa contrôlabilité, stabilisation rapide d'une équation de Schrödinger et méthodes spectrales pour le calcul du contrôle optimal. Mathématiques générales [math.GM]. Université Pierre et Marie Curie - Paris VI, 2016. Français. NNT : 2016PA066395 . tel-01826066

HAL Id: tel-01826066

https://theses.hal.science/tel-01826066

Submitted on 29 Jun 2018

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.

Université Pierre et Marie Curie
Paris 6

Laboratoire Jacques-Louis Lions
UMR 7598

Thèse de Doctorat de
l'Université Pierre et Marie Curie

Présentée et soutenue publiquement le 27 juin 2016

pour l'obtention du grade de

Docteur de l'Université Pierre et Marie Curie

Spécialité : Mathématiques Appliquées

par

Ludovick Gagnon

sous la direction de

Jean-Michel CORON

Étude de l'équation de Korteweg-de Vries en variables lagrangiennes et sa contrôlabilité, stabilisation rapide d'une équation de Schrödinger et méthodes spectrales pour le calcul du contrôle optimal

après avis des rapporteurs

Mme. Fatiha Alabau-Boussouira & M. Enrique Zuazua

devant le jury composé de

Mme.	Fatiha Alabau-Boussouira	Rapporteur
M.	Jean-Michel Coron	Directeur de Thèse
M.	Sylvain Ervedoza	Examinateur
M.	Olivier Glass	Examinateur
M.	Jean-Pierre Puel	Examinateur
M.	Pierre Rouchon	Examinateur
M.	Emmanuel Trélat	Examinateur
M.	Enrique Zuazua	Rapporteur

École Doctorale de Sciences Mathématiques
de Paris Centre ED 386

Faculté de Mathématiques
UFR 929

Ludovick Gagnon :

Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France.

Adresse électronique : gagnon@ann.jussieu.fr

Remerciements

Cette thèse, et mon parcours en général, n'aurait jamais eu lieu dans des circonstances aussi propices sans la contribution d'une quantité importante de personnes. Je m'estime excessivement chanceux et redevable d'avoir eu l'opportunité que ces gens se retrouvent sur mon chemin. Je profite de cet instant pour les remercier, bien qu'il m'est impossible de rendre la liste exhaustive.

Je tiens tout d'abord à exprimer ma profonde gratitude envers Jean-Michel Coron. S'il est incontestablement un mathématicien hors pair, c'est surtout son humilité, sa gentillesse sa patience et sa générosité qui m'ont le plus surpris. Accompagné de son ouverture d'esprit, de sa détermination et de son originalité, il m'a toujours encouragé et poussé dans toutes les démarches que j'ai entreprises. Il a également toujours été à l'écoute et de bon conseil. J'ai énormément appris avec lui durant ces trois années et je ne saurai jamais le remercier assez de m'avoir permis de me rendre au point où j'en suis.

Je veux également remercier Fatiha Alabau-Boussouira et Enrique Zuazua, non seulement d'avoir pris le temps de rapporter cette thèse, mais aussi pour les discussions que j'ai pu avoir avec eux ainsi que pour leur support.

Je remercie Sylvain Ervedoza, Olivier Glass, Jean-Pierre Puel, Pierre Rouchon et Emmanuel Trélat. Je suis très honoré que chacun d'entre eux ait accepté de faire partie de mon jury et je les remercie pour les discussions que j'ai pu avoir avec eux.

Je tiens à souligner la gentillesse générale de la communauté du contrôle. J'en profite pour remercier en particulier Morgan Morancey pour l'ambiance de travail durant la dernière année ainsi que pour son amitié.

Je remercie José Urquiza de l'Université Laval d'avoir toujours cru en moi. C'est grâce à lui que j'ai pu rencontrer Jean-Michel Coron et que j'ai pu faire ma thèse à Paris.

Je remercie également le laboratoire Jacques-Louis Lions et tous ses membres. L'ambiance du laboratoire a été à la fois très stimulante et amicale. Je tiens à remercier particulièrement les secrétaires, Catherine Drouet, Nadine Foucart, Malika Larcher et Salima Lounici, pour toute l'aide et la compréhension.

Je tiens également à remercier tous mes amis qui ont été présents durant cette thèse. Je tiens à remercier tous les gens qui ont été dans mon bureau durant mon passage, plus particulièrement Malik, Giacomo, Jean-Pierre, Casimir, Guillaume, Sylvain, Shuyang, Florian, Chen-Yu, Carlos et Juliette. Je veux également remercier ceux du laboratoire, et plus particulièrement Thibault, Sarah, Carlo, Pierre(s), Clément, Aurora, Mehdi, Eugénie, Long, Frédéric, Camille(s), Philippe, Ryadh, Guillaume, Shengquan, Diego, Antoine, Dena et j'en passe. Je veux également remercier mes amis du Québec qui ont toujours été présents malgré la distance Dom, Bod, Thierry, Jean-Philippe, Laurent(s), Vincent, Anick, Benoît, Jérôme, Nathalie, Jean, Sarah, Anika, Fred, Andrea, Max, Malik, Rose, Alex, P-Y, Louif, et, encore une fois, la liste est trop courte. Je remercie également tous les amis extérieurs aux groupes mentionnés précédemment, Swann, Charlotte, Dario, Steven... Je remercie spécialement Ivan pour ces deux années passées en colococation, pour le support moral et mathématique. Je remercie également Christèle pour tout ce qu'elle m'a apporté.

Je tiens finalement à remercier ma mère pour tout ce qu'elle a pu m'apporter. Elle est en très grande partie, avec mes grands-parents, la raison de ma présence à Paris pour faire ma thèse. Elle a compris rapidement mon désir à poursuivre en mathématiques et elle m'a toujours supporté dans cette direction. Elle m'a énormément apporté en présence, en compréhension et en compassion.

Les mots ne seront jamais suffisants afin d'exprimer ma gratitude pour tout ce que ma mère a pu faire pour moi. La seule chose que je peux faire est de lui dédier ce manuscrit.

Table des matières

Introduction **9**

1 Introduction générale 9

2 Quelques notions de contrôlabilité et de stabilisation en dimension finie et infinie 10

2.1 Rappel sur la théorie des semi-groupes 10

2.2 Contrôlabilité 13

2.3 Stabilisation 18

3 L'équation de Korteweg-de Vries 20

3.1 Historique de l'équation 21

3.2 Scattering de l'équation de Korteweg-de Vries 30

3.3 Contributions à la contrôlabilité lagrangienne de l'équation de Korteweg-de Vries 35

3.4 Contributions à l'approximation d'ordre supérieur du champ de vitesse associé à la solution de N-solitons 45

3.5 Perspectives 49

4 Stabilisation rapide d'une équation de Schrödinger linéaire 51

4.1 L'équation de Schrödinger 51

4.2 Contributions à la stabilisation du linéarisé de l'équation de Schrödinger bilinéaire 52

4.3 Backstepping 54
4.4 Méthode utilisée 59
4.5 Commentaires sur le résultat obtenu et perspectives 66
5 Contrôles numériques pour l'équation des ondes par une méthode spectrale 68
5.1 La problématique 68
5.2 Résultats antérieurs 72
5.3 Résultats obtenus 72
5.4 Discussion des résultats obtenus et perspectives 78

I L'équation de Korteweg-de Vries 79

1 Contrôlabilité lagrangienne de l'équation de Korteweg-de Vries 81
1.1 Introduction 81
1.2 Well-posedness and regular controls 90
1.2.1 Well-posedness of the KdV equation 90
1.2.2 Regular controls 92
1.3 Lagrangian controllability 96

2 Approximation d'ordre supérieur du champ de vitesse associé à la solution de N-solitons 107
2.1 Introduction 107
2.2 Derivation of the Korteweg-de Vries equation 111
2.3 A higher velocity field for the N-soliton solution 115
2.3.1 The N-solitons solution 115
2.3.2 Nonlaminar velocity field 117
2.3.3 Proof of Theorem 2.1.1 121

2.4 Numerical analysis of the particles trajectory 125

2.5 Conclusion 127

II Stabilisation d'une équation de Schrödinger 129

3 Stabilisation d'une équation de Schrödinger 131

3.1 Introduction 132

3.1.1 Main result 132

3.1.2 A finite dimensional example 134

3.1.3 The linear Schrödinger equation 139

3.1.4 A brief review of previous results 140

3.1.5 Structure of the article 141

3.2 Heuristic construction of the transformations 142

3.3 Definition and properties of the transformation 147

3.3.1 Riesz basis property 147

3.3.2 Definition and regularity of the feedback law 156

3.3.3 Wellposedness of the transformation 162

3.4 Invertibility of the transformation 163

3.4.1 Fredholm form 163

3.4.2 Operator equality 166

3.4.3 Invertibility 171

3.5 Well-posedness of the linearized system 174

3.5.1 Well-posedness of the linearized system 174

3.5.2 Proof of rapid stabilization 177

3.6 Admissibility of the feedback operator 178

III Calcul numérique des contrôles optimaux de l'équation des ondes 181

4 Calcul numérique des contrôles optimaux de l'équation des ondes 183

4.1 Introduction 183

4.2 Legendre Galerkin semi-discretization 188

4.2.1 Legendre Galerkin approximation 188

4.2.2 Discrete observability 191

4.3 Spectral filtering 193

4.4 Mixed Legendre-Galerkin method 197

4.5 Nitsche's method 201

4.5.1 Symmetric case 201

4.5.2 Non-symmetric case 206

4.5.3 Finite Element Method 209

4.6 Introduction of the numerical experiments 214

4.6.1 Numerical Experiments 215

A Annexe 223

A.1 Simplified Saint-Venant equation example 223

A.2 Quadratically close families 226

A.3 Rapid stabilization of the linearized system 230

A.4 Computation of $c_{N,T}$ and $C_{N,T}$ 232

Bibliography 248

Introduction

1 Introduction générale

Considérons un système physique évoluant au fil du temps (la température d'une pièce, l'élévation d'une vague d'un fluide dans un bac d'eau, une particule dans un champ électrique, un bras robotique, etc) sur lequel on a la possibilité d'agir à l'aide d'une force extérieure, appelée contrôle (un radiateur, le déplacement du bac, un laser, un moteur). La théorie du contrôle s'intéresse à caractériser l'influence de ce contrôle sur la dynamique du système. Distinguons deux notions relatives à la façon d'agir sur le système.

La première, la contrôlabilité, consiste à atteindre une cible (un état final du système, une trajectoire autonome du système) à partir de n'importe quel état initial du système grâce à un contrôle. Ce contrôle est déterminé à l'avance en fonction de l'état initial et de la cible recherchée. Si, de n'importe quel état initial, on peut atteindre n'importe quelle cible, le système est alors dit exactement contrôlable. Si, à partir de n'importe quel état initial, seulement une partie des cibles sont atteignables, le système est alors dit partiellement contrôlable. Le système est dit non contrôlable si on ne peut systématiquement pas atteindre les cibles souhaitées à l'aide du contrôle.

La seconde, la stabilisation, consiste à rapprocher des points d'équilibres la dynamique du système en ajustant continûment le contrôle en fonction de l'état du système. Si la notion de stabilité semble être une notion plus faible que la contrôlabilité, elle a l'avantage d'être plus robuste. En effet, une légère perturbation peut éloigner fortement l'état final d'un système contrôlé de la cible visée alors que la stabilisation, par la conception du contrôle, permettra toujours de se rapprocher du point d'équilibre.

En pratique, l'étude de la contrôlabilité et de la stabilisation d'un système physique passera par l'équation qui le modélise. Ce cadre mathématique permet le développement d'outils systématiques afin d'étudier le caractère contrôlable ou stabilisable d'une équation. Des résultats généraux peuvent également être dégagés en fonction des propriétés des équations considérées.

2 Quelques notions de contrôlabilité et de stabilisation en dimension finie et infinie

Débutons d'abord par une brève étude des notions de semi-groupe, de contrôlabilité et de stabilisation qui seront utiles dans la suite du document. La présentation sur la théorie des semi-groupes est inspirée de [42, 126, 35, 112] et celle de la théorie du contrôle de [42]. Nous renvoyons à ces références pour une présentation plus approfondie.

2.1 Rappel sur la théorie des semi-groupes

Introduisons quelques notations et définitions. Pour un espace de Hilbert H sur $\mathbb{C}$, nous noterons sa norme par $\|\cdot\|_H$ et son produit scalaire par $\langle\cdot,\cdot\rangle_H$. Pour deux espaces de Hilbert E et F, $\mathcal{L}(E,F)$ désignera l'espace des applications linéaires bornées de E dans F. Pour un opérateur linéaire $A : H \to H$ non borné, on définit le domaine de l'opérateur A par le sous-espace de H pour lequel A est à valeur dans H

$$D(A) := \{x \in H \,|\, Ax \in H\}.$$

On dit que $A : H \to H$ est fermé si son graphe $G(A) := \{(x, Ax) \,|\, x \in D(A)\}$ est fermé dans $H \times H$. Si $D(A)$ est dense dans H, alors on peut définir l'adjoint A^* de A comme l'opérateur linéaire satisfaisant

$$\langle Ax, y\rangle_H = \langle x, A^*y\rangle_H, \quad (x,y) \in D(A) \times D(A^*),$$

où $D(A^*) := \{y \in H \,|\, \exists C_y > 0$ telle que $|\langle Ax, y\rangle_H| \leq C_y\|x\|_H, \forall x \in D(A)\}$. On dit qu'un opérateur linéaire A est dissipatif si

$$\Re\langle Ax, x\rangle_H \leq 0, \quad \forall x \in D(A).$$

On dit que la famille d'opérateurs linéaires bornés $(S(t))_{t\geq 0}$ de H dans H est un semi-groupe fortement continu si

1. $S(t) \in \mathcal{L}(H)$, $\quad \forall t \geq 0$;
2. $S(0) = I$;
3. $S(t+s) = S(t) \circ S(s)$, $\quad \forall t, s \geq 0$;
4. $\lim_{t\to 0^+} S(t)x = x$, $\quad \forall x \in H$.

On dit que $(S(t))_{t\in\mathbb{R}}$ est un groupe fortement continu si les propriétés précédentes sont valides pour tout $t \in \mathbb{R}$. On dit que $S(t)$ est un semi-groupe de contractions si $\|S(t)\|_{\mathcal{L}(H)} \leq 1, \forall t \geq 0$, un groupe vérifiant toujours $\|S(t)\|_{\mathcal{L}(H)} = 1, \forall t \in \mathbb{R}$. On définit finalement le générateur infinitésimal A d'un semi-groupe fortement continu

$S(t)$ comme l'opérateur linéaire non borné A tel que

$$Ax = \lim_{t\to 0^+} \frac{S(t)x - x}{t}, \quad \forall x \in D(A),$$

où

$$D(A) := \left\{ x \in H \;\middle|\; \lim_{t\to 0^+} \frac{S(t)x - x}{t} \text{ existe} \right\}.$$

La limite à droite est remplacée par une limite conventionnelle dans la définition du générateur infinitésimal A d'un groupe fortement continu.

Ces définitions permettent d'introduire le Théorème de Lumer-Phillips,

Théorème 2.1 (Lumer-Phillips) *Soit $A : D(A) \to H$ fermé de domaine dense. Si A et A^* sont dissipatifs, alors A est le générateur d'un semi-groupe de contractions fortement continu.*

2.1.1 Solutions d'équations autonomes

Le Théorème de Lumer-Phillips donne une condition suffisante afin de définir une solution de

$$\begin{cases} x'(t) = Ax(t), & t \in [0, \infty), \\ x(0) = x_0. \end{cases} \tag{2.1}$$

Supposons que A est le générateur infinitésimal d'un semi-groupe de contractions fortement continu $S(t)$ et supposons que $x_0 \in D(A)$. Alors l'unique solution forte de (2.1) est donnée par

$$x(t) = S(t)x_0,$$

et $x \in C([0, \infty); D(A)) \cap C^1([0, \infty); H)$.

Les solutions faibles de (2.1) sont définies avec la théorie des semi-groupes de la facon suivante. Définissons $H_1 := D(A)$ et H_{-1} comme la complétion de H par la norme $\|(\beta I - A)^{-1}(.)\|_H$ avec $\beta \in \rho(A)$, où $\rho(A) := \mathbb{C} \setminus \sigma(A)$ est l'ensemble résolvant de A. On a les inclusions $H_1 \subset H \subset H_{-1}$. En considérant H comme l'espace pivot, on a alors $D(A^*)' \simeq H_{-1}$. La densité de H dans H_{-1} permet d'étendre l'opérateur linéaire A, défini sur H, à un opérateur linéaire $A_{-1} : H_{-1} \to H_{-1}$. Si A est le générateur infinitésimal du semi-groupe $S(t)$ sur H, alors celui-ci s'étend également à un semi-groupe fortement continu $S_{-1}(t) : H_{-1} \to H_{-1}$ ayant comme générateur infinitésimal A_{-1}. Ces indices seront omis par la suite puisque le cadre fonctionnel permet de déduire les opérateurs sous-entendus.

Pour $x_0 \in H$, l'unique solution faible de (2.1) est donnée par

$$x(t) = S(t)x_0, \tag{2.2}$$

et $x \in C([0,\infty); H) \cap C^1([0,\infty); H_{-1})$.

2.1.2 Solutions aux équations inhomogènes

Pour $T > 0$, nous nous intéresserons aux systèmes linéaires de la forme

$$\begin{cases} x'(t) = Ax(t) + Bu(t), & t \in [0,T], \\ x(0) = x_0, \end{cases} \tag{2.3}$$

où $x_0 \in H$ et $u \in L^2((0,T); U)$. Dans les problèmes de contrôlabilité, l'espace H sera l'espace des états et l'espace U, l'espace des contrôles, où H et U seront des espaces de Hilbert de dimension finie ou infinie.

Différentes régularités des solutions de (2.3) seront obtenues en fonction de la régularité de la donnée initiale et de l'image de B. Supposons que $x_0 \in D(A)$ et $B \in \mathcal{L}(U; D(A))$. Alors l'unique solution forte de (2.3) est donnée par

$$x(t) = S(t)x_0 + \int_0^t S(t-s)Bu(s)\mathrm{d}s, \tag{2.4}$$

et $x \in C([0,T]; D(A)) \cap C^1([0,T]; H)$. Si $x_0 \in H$ et $B \in \mathcal{L}(U; H)$, alors l'unique solution faible de (2.3) est donnée par (2.4) et $x \in C([0,T]; H)$. Dans ce cas, on dit que $x(t)$ est une solution de (2.3) si

$$(x(t), z)_H = (x_0, S(t)^* z)_H + \int_0^t (u(s), B^* S(t-s)^* z)_H \mathrm{d}s, \quad \forall z \in H.$$

Finalement, si $x_0 \in H_{-1}$ et $B \in \mathcal{L}(U; H_{-1})$, alors (2.4) définit une unique solution faible au sens des transpositions

$$\langle x(t), z\rangle_{H_{-1}, D(A^*)} = \langle x_0, S(t)^* z\rangle_{H_{-1}, D(A^*)} + \int_0^t \langle u(s), B^* S(t-s)^* z\rangle_{H_{-1}, D(A^*)} \mathrm{d}s, \quad \forall z \in D(A^*).$$

Le cadre fonctionnel des solutions de (2.3) en théorie du contrôle en dimension infinie est légèrement différent de celui présenté précédemment. Les données initiales x_0 appartiendront à l'espace d'état H mais les opérateurs de contrôle seront généralement non bornés dans l'espace des états, ceux-ci appartenant à l'espace plus faible $B \in \mathcal{L}(U, D(A^*)')$. Ce cadre fonctionnel permet en particulier de considérer des opérateurs de contrôle frontière. Il est toujours possible de définir une solution de (2.3) si l'opérateur linéaire B est dit admissible pour le semi-groupe $S(t)$ si, c'est-à-dire

$$\int_0^t S(t-s)Bu(s)\mathrm{d}s \in H, \quad t \geq 0.$$

Cette condition est équivalente à l'admissibilité de l'opérateur adjoint de B, c'est-à-

dire qu'il existe, pour tout $T > 0$, une constante $C_T > 0$ telle que

$$\int_0^T \|B^* S(t)^* z\|_U^2 \mathrm{d}t \leq C_T \|z\|_H^2, \quad \forall z \in D(A^*). \tag{2.5}$$

Dans la littérature, ces résultats d'admissibilité sont souvent appelés des résultats de régularité cachée de la solution.

2.2 Contrôlabilité

2.2.1 Contrôlabilité d'équations linéaires

Le cadre fonctionnel des solutions de (2.3) en dimension infinie étant établi, introduisons maintenant la notion de contrôlabilité exacte.

Definition 2.1 *On dit que* (2.3) *est exactement contrôlable en temps $T > 0$ si et seulement si, pour tout $x_0 \in H$ et $x_T \in H$, il existe un contrôle $u(t) \in L^2((0,T);U)$ tel que la solution $x(t)$ de* (2.3) *satisfasse $x(T) = x_T$.*

Pour des équations linéaires, cette définition est équivalente à la contrôlabilité à partir de l'état initial $x_0 = 0$, notion définie comme suit.

Definition 2.2 *On dit que* (2.3) *est exactement contrôlable en temps $T > 0$ à partir de $x_0 = 0$ si et seulement si, pour tout $x_T \in H$, il existe un contrôle $u(t) \in L^2((0,T);U)$ tel que la solution $x(t)$ de* (2.3) *satisfasse $x(T) = x_T$.*

En effet, si la contrôlabilité exacte implique trivialement la contrôlabilité à partir de zéro, la réciproque s'obtient en découplant la dynamique issue du contrôle à celle issue de la donnée initiale en écrivant $x(t) = x_1(t) + x_2(t)$ où x_1 est la solution de (2.3) avec $x_0 = 0$ et x_2 est la solution de (2.3) avec $u = 0$. Le contrôle amenant $x(t)$ de x_0 à x_T s'obtient en choisissant $x_T - x_2(T)$ comme cible pour la solution de $x_1(t)$.

Définissons également la contrôlabilité à zéro.

Definition 2.3 *On dit que* (2.3) *est exactement contrôlable à zéro en temps $T > 0$ si et seulement si, pour tout $x_0 \in H$, il existe un contrôle $u(t) \in L^2((0,T);U)$ tel que la solution $x(t)$ de* (2.3) *satisfasse $x(T) = 0$.*

Si l'équation est réversible en temps, alors la contrôlabilité à zéro est équivalente à la contrôlabilité exacte.

Un résultat classique permettant de caractériser la contrôlabilité de systèmes autonomes à coefficients constants de dimension finie est le critère de Kalman ([42]).

Théorème 2.2 (Critère de Kalman) *Soit $H = \mathbb{R}^n$ et $U = \mathbb{R}^m$. L'équation* (2.3) *est exactement contrôlable pour tout temps $T > 0$ si et seulement si*

$$rang[B, AB, \ldots, A^{n-1}B] = n.$$

Afin de caractériser la contrôlabilité en dimension infinie, considérons, sans perte de généralité, la contrôlabilité exacte à partir de zéro et introduisons l'application associant à un contrôle l'état final de la solution de (2.3)

$$\mathcal{F}_T(u) = x(T).$$

La définition de la contrôlabilité à zéro est équivalente à la surjectivité de $\mathcal{F}_T$. Un résultat classique d'analyse fonctionnelle ([24]) donne l'équivalence entre la surjectivité de $\mathcal{F}_T$ et l'existence d'une constante positive $c_T > 0$ telle que

$$c\|z_T\|_H^2 \leq \|\mathcal{F}_T^*(z_T)\|_{L^2((0,T);U)}^2, \quad \forall z_T \in H.$$

L'expression de $\mathcal{F}_T^*(z_T)$ est obtenue avec la formule de Duhamel. Pour tout $z_T \in D(A^*)$, on a

$$\begin{aligned}\langle \mathcal{F}_T(u), z_T\rangle_H =& \langle x(T), z_T\rangle_H \\ =& \int_0^T \langle S(T-t)Bu(t), z_T\rangle_H \mathrm{d}t \\ =& \int_0^T \langle u(t), B^*S(T-t)^* z_T\rangle_U \mathrm{d}t \\ =& \langle u, \mathcal{F}_T^*(z_T)\rangle_{L^2((0,T);U)}.\end{aligned}$$

La contrôlabilité de (2.3) est donc équivalente à l'existence d'une constante positive $c_T > 0$ telle que

$$c_T\|z_T\|_H^2 \leq \int_0^T \|B^*S(T-t)^* z_T\|_U^2, \quad \forall z_T \in D(A^*). \tag{2.6}$$

Cette dernière inégalité est appelée inégalité d'observabilité pour l'équation adjointe rétrograde

$$\begin{cases} z'(t) = -A^* z(t), & t \in [0,T], \\ z(T) = z_T. \end{cases} \tag{2.7}$$

Definition 2.4 *On dit que* (2.7) *est observable en temps $T > 0$ si et seulement si il existe une constante $c_T > 0$ telle que* (2.6) *soit satisfaite.*

Si $\|\cdot\|_H$ est une norme équivalente à l'énergie des solutions de (2.7), alors l'observabilité peut être interprétée comme l'estimation de l'énergie des solutions de (2.7)

solution par une observation B^*.

Le problème adjoint permet également d'obtenir un contrôle minimisant la norme $L^2((0,T);U)$. En effet, en considérant $z(t)$ solution de (2.7), on a

$$\begin{aligned}
0 &= \int_0^T < x'(t) - Ax(t) - Bu(t), z(t) >_H \, \mathrm{d}t \\
&= -\int_0^T < x(t), z'(t) - A^* z(t) >_H + < Bu(t), z(t) >_H \, \mathrm{d}t + < x(T), z_T >_H \\
&= -\int_0^T < Bu(t), z(t) >_H \, \mathrm{d}t + < x(T), z_T >_H \, . \qquad (2.8)
\end{aligned}$$

En choisissant $u(t) = B^* S(T-t)^* W(0,T)^{-1} x_T$, où x_T est la cible à atteindre et où $W(0,T)$ est le grammien d'observabilité

$$W(0,T) := \int_0^T S(T-t) B B^* S(T-t)^* \mathrm{dt},$$

qui est inversible si l'inégalité d'observabilité est satisfaite, on obtient

$$\begin{aligned}
< x(T), z_T >_H &= \int_0^T < Bu(t), z(t) >_H \, \mathrm{d}t \\
&= \int_0^T < S(T-t) B B^* S(T-t)^* W(0,T)^{-1} x_T, z_T >_H \, \mathrm{d}t \\
&= < \left(\int_0^T S(T-t) B B^* S(T-t)^* \mathrm{d}t \right) W(0,T)^{-1} x_T, z_T >_H \\
&= < x_T, z_T >_H \, .
\end{aligned}$$

Montrons que ce contrôle minimise la norme $L^2((0,T);U)$. Supposons qu'il existe un autre contrôle $\tilde{u} \neq u$ contrôlant (2.3) à x_T. Alors,

$$\begin{aligned}
\|\tilde{u}\|^2_{L^2((0,T);U)} &= \|\tilde{u} - u + u\|^2_{L^2((0,T);U)} \\
&= \|\tilde{u} - u\|_{L^2((0,T);U)} + 2(\tilde{u} - u, u)_{L^2((0,T);U)} + \|u\|^2_{L^2((0,T);U)}.
\end{aligned}$$

Or,

$$\begin{aligned}
(\tilde{u}, u)_{L^2((0,T);U)} &= (\tilde{u}, B^* S(T-t)^* W(0,T)^{-1} x_T)_{L^2((0,T);U)} \\
&= (S(T-t) B \tilde{u}, W(0,T)^{-1} x_T)_{L^2((0,T);H)} \\
&= (x_T, W(0,T)^{-1} x_T)_{L^2((0,T);H)} \\
&= (W(0,T) W(0,T)^{-1} x_T, W(0,T)^{-1} x_T)_{L^2((0,T);H)} \\
&= (B^* S(T-t)^* W(0,T)^{-1} x_T, B^* S(T-t)^* W(0,T)^{-1} x_T)_{L^2((0,T);U)} \\
&= \|u\|^2_{L^2((0,T);U)},
\end{aligned}$$

ce qui implique que $(\tilde{u}-u,u)_{L^2((0,T);U)}=0$ et donc que $\|u\|_{L^2((0,T);U)}<\|\tilde{u}\|_{L^2((0,T);U)}$.

Un contrôle de (2.3) est donc donné par $u(t)=B^*z(t)$ où $z(t)$ est la solution du problème rétrograde ayant comme donnée finale $z_T=W(0,T)^{-1}x_T$. Il est parfois pratique de reformuler l'expression de cette donnée initiale comme un problème de minimisation

$$\arg\min_{z\in D(A^*)} J(z_T),$$

où

$$J(z_T)=\frac{1}{2}\int_0^T \|B^*z(t)\|\mathrm{d}t- < x_T, z_T >_H .$$

Notons que les équivalences évoquées précédemment ne sont que différents angles d'attaques d'un même problème. Elles ne permettent pas en soi de déterminer la contrôlabilité d'une équation. Par contre, il est souvent plus aisé de travailler sur l'inégalité d'observabilité que la démonstration directe de la contrôlabilité d'une équation. Afin de montrer une inégalité d'observabilité, quelques méthodes sont classiques : la méthode des multiplicateurs, les inégalités de Carleman ou un raisonnement par l'absurde conduisant à une solution de l'équation adjointe rétrograde surdéterminée. La méthode des moments peut être également appliquée directement sur l'équation à contrôler ([42]).

Si la contrôlabilité d'une équation dépend de ses caractéristiques intrinsèques, par exemple l'hyperbolicité ou la dissipativité, notons le résultat abstrait d'obstruction à la contrôlabilité.

Théorème 2.3 ([125]) *Soit $T>0$ et supposons que $B\in\mathcal{L}(U,H)$ et est compact. Alors* (2.3) *n'est pas contrôlable en temps T.*

L'interprétation de ce résultat est que l'action du contrôle par B doit être suffisamment singulière afin de changer la dynamique du système engendrée par A.

2.2.2 Contrôlabilité d'équations non linéaires

Si plusieurs méthodes sont connues afin de caractériser la contrôlabilité d'équations linéaires, beaucoup reste encore à faire pour les équations non linéaires, tant en dimension finie qu'en dimension infinie. Présentons deux conditions suffisantes afin d'obtenir la contrôlabilité locale en temps petit d'une équation non linéaire. Considérons, par simplicité, le système de dimension finie non linéaire

$$\begin{cases} x'(t)=f(x(t),u(t)), \quad x\in[0,T], \\ x(0)=x_0, \end{cases} \tag{2.9}$$

où $x\in\mathbb{R}^n$ et $u\in\mathbb{R}^m$. Supposons de plus que f est $C^1(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R}^n)$ et que (2.9) admet un point d'équilibre (x_e,u_e), c'est-à-dire $f(x_e,u_e)=0$. La contrôlabilité locale

en temps petit est définie de la façon suivante.

Definition 2.5 *Soit* $(x_e, u_e) \in \mathbb{R}^n \times \mathbb{R}^m$ *un point d'équilibre de* (2.9)*. Le système* (2.9) *est dit localement contrôlable en temps petit autour du point d'équilibre* (x_e, u_e) *si, pour tout* $\varepsilon > 0$*, il existe* $r > 0$ *tel que, pour tout* $x_0 \in B_r(x_e) := \{x \in \mathbb{R}^n \mid \|x - x_e\|_{\mathbb{R}^n} < r\}$ *et* $x_T \in B_r(x_e) := \{x \in \mathbb{R}^n \mid \|x - x_e\|_{\mathbb{R}^n} < r\}$*, il existe une fonction mesurable* $u : [0, \varepsilon] \to \mathbb{R}^m$ *telle que* $\|u - u_e\|_{\mathbb{R}^m} < \varepsilon, \forall t \in [0, \varepsilon]$ *et telle que la solution de* (2.9) *satisfasse*

$$x(\varepsilon) = x_T.$$

L'une des premières approches afin de caractériser la contrôlabilité locale en temps petit autour d'un point d'équilibre (x_e, u_e) d'une équation non linéaire consiste à étudier la contrôlabilité du linéarisé autour de ce point d'équilibre

$$\begin{cases} x'(t) = \dfrac{\partial f}{\partial x}(x_e, u_e)x(t) + \dfrac{\partial f}{\partial u}(x_e, u_e)u(t), \quad x \in [0, T], \\ x(0) = x_0. \end{cases}$$

Si cette équation est contrôlable, alors la contrôlabilité locale en temps petit est obtenue grâce à un argument de point fixe.

Théorème 2.4 ([42]) *Soit* (x_e, u_e) *un point d'équilibre de* (2.9)*. Supposons que le linéarisé autour de ce point d'équilibre est contrôlable. Alors* (2.9) *est localement contrôlable en temps petit autour de ce point d'équilibre.*

Bien que le résultat obtenu soit seulement local, l'étude du linéarisé a l'avantage de mettre à disposition toutes les techniques développées dans le cadre linéaire.

Notons que si le linéarisé autour du point d'équilibre (x_e, u_e) n'est pas contrôlable, tout n'est pas perdu. Une approche afin d'obtenir la contrôlabilité locale autour du point d'équilibre (x_e, u_e) de (2.9) si le linéarisé autour de ce point d'équilibre n'est pas contrôlable est la méthode du retour.

Supposons, pour tout $T > 0$ et $\varepsilon > 0$, qu'il existe une trajectoire $(\overline{x}(t), \overline{u}(t))$ de (2.9), telle que $\overline{x}(0) = \overline{x}(T) = x_e$, telle que $\|\overline{u}\|_{L^\infty(0,T)} \leq \varepsilon$ et telle que le linéarisé autour de cette trajectoire

$$\begin{cases} x'(t) = \dfrac{\partial f}{\partial x}(\overline{x}(t), \overline{u}(t))x(t) + \dfrac{\partial f}{\partial u}(\overline{x}(t), \overline{u}(t))u(t), \quad x \in [0, T], \\ x(0) = x_0. \end{cases}$$

est contrôlable. Alors, toujours par un argument de point fixe, on obtient la contrôlabilité locale de (2.9) en temps T ([42]).

2.3 Stabilisation

Dans cette section, nous concentrerons notre attention sur des systèmes de dimension finie. Définissons d'abord quelques notions de stabilité pour le système dynamique

$$\begin{cases} x'(t) = f(x(t)), \quad x \in [0, \infty), \\ x(0) = x_0, \end{cases} \tag{2.10}$$

où $f \in C^1(\mathbb{R}^n; \mathbb{R}^n)$. Comme dans la section précédente, on dit que x_e est un point d'équilibre de (2.10) si $f(x_e) = 0$. On a les définitions suivantes pour un point d'équilibre de (2.10).

Definition 2.6 (Stabilité) *On dit qu'un point d'équilibre x_e de* (2.10) *est stable si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $\|x_0 - x_e\|_{\mathbb{R}^n} < \eta$ implique*

$$\|x(t) - x_e\|_{\mathbb{R}^n} < \varepsilon, \quad t \geq 0.$$

Definition 2.7 (Stabilité asymptotique globale) *On dit qu'un point d'équilibre x_e de* (2.10) *est globalement asymptotiquement stable si il est stable et si la solution $x(t)$ de* (2.10)*, pour toute condition initiale x_0, satisfait*

$$\lim_{t \to \infty} x(t) = x_e.$$

Considérons maintenant le système

$$\begin{cases} x'(t) = f(x(t), u(t)), \quad x \in [0, \infty), \\ x(0) = x_0, \end{cases} \tag{2.11}$$

où $u(t)$ est un feedback de la forme $u(t) = F(x(t))$. On suppose que F est continue et s'annule en x_e. Définissons les deux notions de stabilité de (2.11).

Definition 2.8 (Stabilité asymptotique globale à l'aide d'une loi de feedback) *On dit que* (2.11) *est globalement asymptotiquement stable (à l'aide d'un feedback) pour le point d'équilibre x_e s'il existe un feedback $u(t) = F(x(t))$ tel que* (2.11) *est globalement asymptotiquement stable.*

Definition 2.9 (Stabilité exponentielle rapide à l'aide d'une loi feedback) *On dit que* (2.11) *est stable exponentiellement rapidement à l'aide d'une loi feedback pour un point d'équilibre x_e si, pour tout $\lambda > 0$, il existe un feedback $u(t) = F(x(t))$ et une constante $C > 0$ tels que*

$$\|x(t) - x_e\| \leq Ce^{-\lambda t}\|x_0\|, \quad \forall x_0 \in H.$$

Puisque nous chercherons toujours la stabilisation de (2.11) à l'aide d'une loi de feedback, cette expression sera omise par la suite.

Un résultat classique pour des équations linéaires en dimension finie est que la contrôlabilité entraîne la stabilité exponentielle rapide. Cela peut se montrer par exemple avec le Théorème de placement de pôles ([42]).

Théorème 2.5 *Supposons que* (2.3) *est contrôlable. Alors, il existe une matrice* K *telle que le feedback* $u = Kx$ *rende le point d'équilibre* $(0, 0)$ *stable exponentiellement rapidement.*

Afin de traiter la stabilité asymptotique globale de (2.11), introduisons la notion de fonction de Lyapunov.

Definition 2.10 *Une fonction* $V \in C^1(\mathbb{R}^n; \mathbb{R})$ *est une fonction de Lyapunov pour le point* $\tilde{x}$ *de* (2.10) *si*

$$\begin{aligned} V(x) &\to \infty, \quad (|x| \to \infty), \\ V(x) &> 0, \quad \forall x \in \mathbb{R}^n \setminus \{\tilde{x}\}, \\ V(\tilde{x}) &= 0, \\ \langle \nabla V(x), f(x) \rangle &\leq 0, \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

On dira de plus que V *est une fonction de Lyapunov stricte si l'inégalité de la dernière équation est stricte sur* $\mathbb{R}^n \setminus \{\tilde{x}\}$.

Pour un choix de feedback donné, les deux théorèmes suivants permettent de caractériser la stabilité de points d'équilibres.

Théorème 2.6 (Théorème de Lyapunov) *Soit* x_e *un point d'équilibre de* (2.10)*. S'il existe une fonction de Lyapunov pour* $\tilde{x}$ *de* (2.10)*, alors* $\tilde{x}$ *est stable. Si la fonction de Lyapunov est stricte, alors* $\tilde{x}$ *est globalement asymptotiquement stable.*

Théorème 2.7 (Théorème de LaSalle) *Soit* V *une fonction de Lyapunov. Soit* I *le plus grand sous-ensemble de* $\{x \in \mathbb{R}^n \mid \langle \nabla V(x), f(x) \rangle = 0\}$ *qui soit invariant par rapport à la dynamique pour tout temps* $t \geq 0$*, c'est-à-dire les trajectoires issues de points de cet ensemble reste dans cet ensemble. Alors toute trajectoire* $x(t)$ *tend vers* I.

Notons que la stabilisation asymptotique globale résulte du Théorème de LaSalle en montrant que l'ensemble invariant I est en fait réduit à $\{\tilde{x}\}$.

3 L'équation de Korteweg-de Vries

Dans cette partie nous présentons les apports dans cette thèse résultats concernant l'équation de Korteweg-de Vries.

La Section 3.1 couvre l'historique de l'équation de Korteweg-de Vries. La Section 3.1.1 présente en détail un moyen d'obtenir l'équation de Korteweg-de Vries des équations d'Euler et la Section 3.14 présente l'Inverse Scattering Method, méthode permettant de résoudre l'équation de Korteweg-de Vries à partir de la donnée initiale du problème.

Le premier résultat concernant l'équation de Korteweg-de Vries est le résultat de contrôlabilité Lagrangienne en temps petit. On montre qu'il est possible de construire des solutions à N-solitons telles que les caractéristiques associées sortent en temps petit par la droite d'un domaine de longueur quelconque. Ces solutions sont construites en utilisant un nombre suffisant de solitons suffisamment grand entrant dans le domaine par la gauche et sortant par la droite. Les solitons sont ordonnées en ordre croissant de vitesse afin d'éviter des intéractions supplémentaires entre les solitons lorsque ceux-ci sont situés à l'intérieur du domaine. On montre ensuite que ces solutions, définies sur $\mathbb{R}$, peuvent être raccordées au repos à l'intérieur du domaine au temps initial et au temps final grâce à la contrôlabilité locale en temps petit.

La définition de la contrôlabilité Lagrangienne et l'énoncé du résultat sont présentés dans la sous-section 3.3.1. La connexion entre la contrôlabilité Lagrangienne et la contrôlabilité usuelle et les résultats antérieurs présents dans la littérature sont présentés dans la sous-section 3.3.2. Une explication de la méthode de preuve se trouve en section 3.3.3. Une discussion du résultat obtenu et son lien avec le problème ouvert de contrôlabilité de l'équation de Korteweg-de Vries en temps petit est présentée dans la section 3.3.4.

Le deuxième résultat est l'obtention d'un champ de vitesse d'ordre supérieur pour la solution à N-solitons. Le champ de vitesse d'ordre un obtenu dans la Section 3.1.1 ne permet pas d'obtenir une propriété importante de la trajectoire des particules soumises à une vague solitaire : les particules situées plus haut dans le fluide ont un déplacement plus important. En introduisant une fonction courant et en utilisant le champ de vitesse d'ordre un comme condition au bord des potentiels de vitesse harmoniques, on obtient l'expression d'un champ de vitesse d'ordre supérieur pour la solution à N-solitons permettant, numériquement, de retrouver le comportement souhaité des particules soumises à des vagues solitaires.

La présentation du problème ainsi que l'énoncé du résultat obtenu se trouve dans la section 3.4.1. Les résultats précédents présents dans la littérature sont présentés dans la section 3.4.2. La méthode utilisée afin d'obtenir le champ de vitesse d'ordre supérieur pour la solution à N-solitons est présentée dans la section 3.4.3. Des commentaires sur le résultat obtenu, notamment que la condition cinématique à la surface n'est, en général, pas satisfaite pour les champs de vitesses de l'équation de Korteweg-

de Vries, sont présentés dans la section 3.4.4.

Finalement, une perspective des travaux suivants ceux présentés dans cette partie se trouve dans la section 3.5.

3.1 Historique de l'équation

Les premiers travaux sur l'équation de Korteweg-de Vries sont dus à John Scott Russell. Ingénieur naval supervisant des travaux sur l'Union Canal en 1834, il fut témoin de l'émergence d'une onde solitaire sur le canal suite à l'arrêt brusque d'un bateau tiré par des chevaux.

> "I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped - not so the mass of water in the channel which it had put in motion ; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in windings of the channel. Such, in the month of August 1834, was my first chance interview with a singular and beautiful phenomenon which I have called the Wave of Translation."

Rapporté dans *Report on Waves* [119] dix ans plus tard, ce témoignage est accompagné d'expérimentations sur l'onde solitaire - la "Wave of Translation" - apparaissant dans des eaux de faibles profondeurs. Bien que le profil de cette vague, ou l'équation modélisant le phénomène observé, n'ait pu être déduit par Russell, l'expression empirique de la vitesse de déplacement de la vague figure dans le rapport

$$c^2 = g(h_0 + a), \tag{3.1}$$

où c est la vitesse de la vague, h_0 la profondeur de l'eau, a l'amplitude maximale de la vague et g la constante gravitationnelle. Cette expression rend compte d'une propriété importante de l'onde solitaire : il s'agit d'une vague gravitationnelle dont la vitesse augmente en fonction de son amplitude [119, 54]. Une trentaine d'années se sont ensuite écoulées avant qu'une confirmation théorique de l'expression de la vitesse (3.1) d'une onde solitaire ne soit donnée par Boussinesq en 1871 et par Lord Rayleigh en 1876 sous l'hypothèse que la longueur de la vague doit être beaucoup plus grande que la profondeur de l'eau et que la profondeur de l'eau est beaucoup plus grande que l'amplitude de la vague ([54]). L'approximation d'un profil de la vague solitaire

fut également obtenue

$$\zeta(x,t) = a\,\mathrm{sech}^2\left(B(x-ct)\right),$$

où

$$B = \sqrt{\frac{3a}{4h_0^2(h_0+a)}}.$$

L'équation gouvernant la dynamique de cette onde solitaire fut finalement découverte indépendamment par Boussinesq en 1877 ([23]) et par Korteweg et de Vries en 1895 ([86])

$$\frac{\partial\zeta}{\partial t} = \frac{3}{2}\sqrt{\frac{g}{h_0}}\left(\frac{2}{3}\kappa\frac{\partial\zeta}{\partial\xi} + \zeta\frac{\partial\zeta}{\partial\xi} + \frac{1}{3}\sigma\frac{\partial^3\zeta}{\partial\xi^3}\right),$$

où ξ est une variable se déplaçant quasiment à la même vitesse que la vague, σ est un paramètre incluant la tension surfacique et où κ est un paramètre arbitraire ([54]). Notons qu'un argument asymptotique multi-échelle afin d'obtenir l'équation de Korteweg-de Vries des équations d'Euler pour les fluides parfaits incompressibles est dorénavant préféré afin d'obtenir l'expression de l'équation de Korteweg-de Vries

$$\frac{1}{\sqrt{gh_0}}\partial_t\eta + \partial_x\eta + \frac{h_0^2}{6}\partial_x^3\eta + \frac{3}{2h_0}\eta\partial_x\eta = 0, \quad x\in\mathbb{R},\ t\in\mathbb{R}, \tag{3.2}$$

où l'effet de la tension surfacique a été négligée. L'expression de l'onde solitaire de (3.2) est donnée par

$$\eta(x,t) = a\,\mathrm{sech}^2\left(\beta(x-ct)\right),$$

où

$$\beta = \sqrt{\frac{3a}{4h_0^3}}.$$

L'existence d'une vague à profil fixe se propageant à vitesse constante solution de l'équation de Korteweg-de Vries provient de l'équilibre général entre la partie linéaire dispersive de l'équation avec la non-linéarité de type Burgers. En règle générale, la solution de (3.2) se disperse si elle se rapproche d'une singularité et un regroupement a lieu grâce au terme non linéaire si une dispersion trop forte est présente ([54]).

Il a fallu attendre le milieu des années 60 avant que de nouveaux progrès significatifs soient effectués sur l'équation de Korteweg-de Vries. Kruskal et Zabusky ont montré numériquement que le comportement asymptotique d'une perturbation initiale du profil de l'eau

$$\begin{cases} \frac{1}{\sqrt{gh_0}}\partial_t\eta + \partial_x\eta + \frac{h_0^2}{6}\partial_x^3\eta + \frac{3}{2h_0}\eta\partial_x\eta = 0, \quad x\in\mathbb{R},\ t\in\mathbb{R}^+. \\ \eta(x,0) = \eta^0(x), \end{cases} \tag{3.3}$$

est un nombre fini d'ondes solitaires se déplaçant vers la droite et des vagues oscillantes se dispersant vers la gauche. Ils ont également montré numériquement que

deux ondes solitaires se dépassant conservent leur profil respectif après l'interaction. Le seul changement notable de ce dépassement est un changement de phase de chacun des solitons : le plus rapide se situe devant la position qu'il aurait eue si aucune interaction n'avait eu lieu et le plus lent se situe derrière la position qu'il aurait eue. En raison de la similitude entre le comportement de l'onde solitaire et la dualité quasi-particule et onde, Kruskal et Zabusky ont baptisé solitons l'approximation des ondes solitaires par l'équation de Korteweg-de Vries.

Les résultats numériques de Kruskal et Zabusky furent confirmés théoriquement quelques années plus tard grâce à l'*Inverse Scattering Transform*, introduite par Gardner, Greene, Kruskal et Miura ([64, 65]). Cette méthode consiste à obtenir l'évolution en temps des solutions de (3.3) en obtenant d'abord des données scatterées à partir de la donnée initiale et en résolvant ensuite deux équations linéaires afin de retrouver l'évolution de la solution. Le comportement asymptotique des solutions de (3.3) fut ensuite obtenu par Segur en 1973 [120].

Depuis, l'équation de Korteweg-de Vries a reçu beaucoup d'attention. Le problème du caractère bien posé du problème de Cauchy a été intensivement étudié. Par ailleurs la stabilité des solitons, au sens où une donnée initiale proche du graphe d'un soliton reste près du graphe d'un soliton, a été analysée et beaucoup de problèmes de contrôlabilité ont été étudiés. Notons également que l'équation de Korteweg-de Vries sert également dans le contexte de vagues hydromagnétiques de plasma sans collision [66], de longues vagues dans les cristaux non-harmoniques [133], pour des plasmas ion-acoustique [130] et en cosmologie [93].

Avant de présenter les résultats de cette thèse concernant l'équation de Korteweg-de Vries, présentons comment l'équation de Korteweg-de Vries peut être déduite des équations d'Euler. La présentation de l'Inverse Scattering Method sera effectuée. Ces présentations permettront de préciser le comportement des solutions de l'équation de Korteweg-de Vries ainsi que la nature de validité de l'équation.

3.1.1 Justification de l'équation de Korteweg-de Vries

Détaillons comment l'équation de Korteweg-de Vries s'obtient des équations d'Euler. Considèrons un fluide approximativement bi-dimensionnel, c'est-à-dire que le profil est presque le même pour toute coupe transverse dans la direction de déplacement de la vague. Supposons que le fluide repose sur un fond plat. Une coupe de ce fluide au temps t est représentée par le domaine

$$\Omega_t = \big\{(x,y) \in \mathbb{R}^2 \,\big|\, 0 \leq y \leq h_0 + \eta(x,t)\big\},$$

où $y = h_0$ est la hauteur au repos du fluide par rapport au fond plat $y = 0$ et $\eta(x,t)$ est la déformation au point (x,t) de la surface au repos. On suppose que le fluide est parfait, c'est-à-dire que l'effet de la viscosité et de la conductivité thermique peuvent être négligés. En supposant que le fluide est incompressible et irrotationnel, le champ

de vitesse des particules du fluide $(u(x,y,t), v(x,y,t)), (x,y) \in \Omega_t$ est modélisé par les équations d'Euler

$$\begin{cases} u_t + uu_x + vu_y = -\frac{1}{\rho}P_x, & 0 \leq y \leq h_0 + \eta, \\ v_t + uv_x + vv_y = \frac{1}{\rho}P_y - g, & 0 \leq y \leq h_0 + \eta, \\ u_x + v_y = 0, & 0 \leq y \leq h_0 + \eta, \\ u_y - v_x = 0, & 0 \leq y \leq h_0 + \eta, \\ v = \eta_t + u\eta_x, & y = h_0 + \eta, \\ P = P_0, & y = h_0 + \eta, \\ v = 0, & y = 0. \end{cases} \tag{3.4}$$

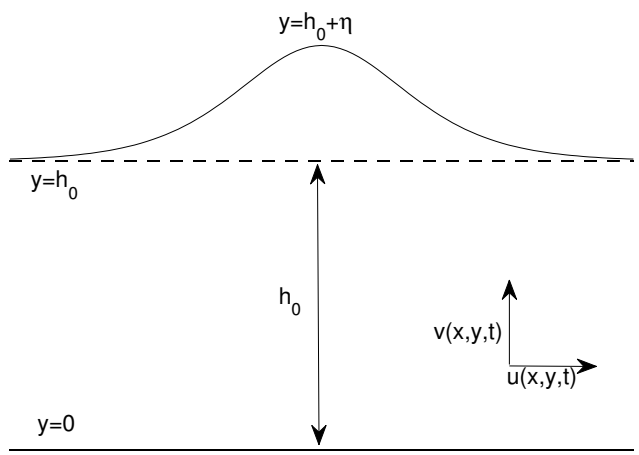

FIGURE 1 – Représentation du système de coordonnées pour l'équation de Korteweg-de Vries

Dans (3.4), les deux premières équations sont les équations d'Euler pour les fluides parfaits, où P est la pression, ρ est la densité, supposée constante, du fluide et g est la constante gravitationnelle. La troisième équation traduit que le fluide est incompressible alors que la quatrième signifie que le fluide est irrotationnel. La cinquième équation est une condition au bord cinématique traduisant que les particules constituant la surface restent à la surface. On suppose de plus que l'action de l'air et du fluide est découplée, résultant en une pression constante à la surface. Finalement, la dernière équation est la condition d'imperméabilité du fond plat. Notons que le cas d'une topographie plus général du fond peut être traité ([37]).

Introduisons le paramètre d'amplitude $\varepsilon := a/h_0$ et le paramètre de profondeur $\delta := h_0/\lambda$, où a est l'amplitude maximale des vagues et λ est la longueur moyenne des vagues. Le régime dans lequel l'équation de Korteweg-de Vries est valable est le

cas d'un fluide peu profond, c'est-à-dire pour ε et δ petits.

Afin d'obtenir l'équation de Korteweg-de Vries, obtenons d'abord une version adimensionnée de (3.4). Pour cela, considérons une perturbation de la pression

$$P = P_0 + \rho g(h_0 - y) + \rho g h_0 p,$$

qui exprime le changement de pression lorsqu'une vague se déplace. Considérons également le changement de variables

$$x \mapsto \lambda x, \quad y \mapsto h_0 y, \quad t \mapsto \frac{\lambda}{\sqrt{gh_0}} t, \quad u \mapsto \sqrt{gh_0} u, \quad v \mapsto \frac{h_0\sqrt{gh_0}}{\lambda} v.$$

Le changement $x \mapsto \lambda x$ signifie que la variable x est remplacée par la nouvelle variable λx. La version adimensionnée de (3.4) est donnée par

$$\begin{cases} u_t + uu_x + vu_y = -p_x, & 0 \leq y \leq 1 + \eta/h_0, \\ \delta^2 \left(v_t + uv_x + vv_y\right) = -p_y, & 0 \leq y \leq 1 + \eta/h_0, \\ u_x + v_y = 0, & 0 \leq y \leq 1 + \eta/h_0, \\ u_y - \delta^2 v_x = 0, & 0 \leq y \leq 1 + \eta/h_0, \\ v = \dfrac{1}{h_0}(\eta_t + u\eta_x), & y = 1 + \eta/h_0, \\ p = \dfrac{\eta}{h_0}, & y = 1 + \eta/h_0, \\ v = 0, & y = 0. \end{cases}$$

Puisque la hauteur des vagues joue un rôle important dans la formulation du problème, réexprimons la déformation de la surface par rapport au paramètre $a : \eta \mapsto a\eta$. Alors, le système adimensionné devient

$$\begin{cases} u_t + uu_x + vu_y = -p_x, & 0 \leq y \leq 1 + \varepsilon\eta, \\ \delta^2 \left(v_t + uv_x + vv_y\right) = -p_y, & 0 \leq y \leq 1 + \varepsilon\eta, \\ u_x + v_y = 0, & 0 \leq y \leq 1 + \varepsilon\eta, \\ u_y - \delta^2 v_x = 0, & 0 \leq y \leq 1 + \varepsilon\eta, \\ v = \varepsilon(\eta_t + u\eta_x), & y = 1 + \varepsilon\eta, \\ p = \varepsilon\eta, & y = 1 + \varepsilon\eta, \\ v = 0, & y = 0. \end{cases}$$

Puisque p et v dans les équations cinq et six sont de l'échelle de ε, nous effectuons le changement de variables sur celles-ci, ainsi que sur u par homogénéité

$$u \mapsto \varepsilon u, \quad v \mapsto \varepsilon v, \quad p \mapsto \varepsilon p,$$

pour obtenir

$$\begin{cases} u_t + \varepsilon\left(uu_x + vu_y\right) = -p_x, & 0 \le y \le 1 + \varepsilon\eta, \\ \delta^2\left(v_t + \varepsilon\left(uv_x + vv_y\right)\right) = -p_y, & 0 \le y \le 1 + \varepsilon\eta, \\ u_x + v_y = 0, & 0 \le y \le 1 + \varepsilon\eta, \\ u_y - \delta^2 v_x = 0, & 0 \le y \le 1 + \varepsilon\eta, \\ v = \eta_t + \varepsilon u\eta_x, & y = 1 + \varepsilon\eta, \\ p = \eta, & y = 1 + \varepsilon\eta, \\ v = 0, & y = 0. \end{cases}$$

Finalement, le changement de variables suivant enlève le paramètre de profondeur dans le système

$$x \mapsto \frac{\delta}{\sqrt{\varepsilon}}x, \quad t \mapsto \frac{\delta}{\sqrt{\varepsilon}}t, \quad v \mapsto \frac{1}{\delta\sqrt{\varepsilon}}v,$$

laissant donc

$$\begin{cases} u_t + \varepsilon uu_x + vu_y = -p_x, & 0 \le y \le 1 + \varepsilon\eta, \\ v_t + \varepsilon\left(uv_x + vv_y\right) = -p_y, & 0 \le y \le 1 + \varepsilon\eta, \\ \varepsilon u_x + v_y = 0, & 0 \le y \le 1 + \varepsilon\eta, \\ u_y - v_x = 0, & 0 \le y \le 1 + \varepsilon\eta, \\ v = \varepsilon\left(\eta_t + \varepsilon u\eta_x\right), & y = 1 + \varepsilon\eta, \\ p = \eta, & y = 1 + \varepsilon\eta, \\ v = 0, & y = 0. \end{cases} \tag{3.5}$$

La quatrième équation de (3.5) nous permet d'introduire un potentiel de vitesse ϕ satisfaisant

$$u(x, y, t) = \phi_x(x, y, t), \quad v(x, y, t) = \phi_y(x, y, t).$$

En écrivant (3.5) en fonction de ϕ, nous obtenons

$$\begin{cases} \varepsilon\phi_{xx} + \phi_{yy} = 0, & 0 < y < 1 + \varepsilon\eta, \\ \phi_y = 0, & y = 0, \\ \phi_y = \varepsilon(\eta_t + \varepsilon\phi_x\eta_x), & y = 1 + \varepsilon\eta, \\ \phi_t + \dfrac{\varepsilon}{2}\phi_x^2 + \dfrac{\phi_y^2}{2} + \eta = 0, & y = 1 + \varepsilon\eta. \end{cases} \tag{3.6}$$

Effectuons un développement asymptotique par rapport à ε du potentiel et du profil

$$\phi(x, y, t) = \sum_{k=0}^{\infty} \varepsilon^k \phi^k(x, y, t), \quad \eta(x, t) = \sum_{k=0}^{\infty} \varepsilon^k \eta^k(x, t).$$

Le premier ordre d'approximation satisfait

$$\begin{cases} \phi^0_{yy} = 0, & 0 < y < 1, \\ \phi^0_y = 0, & y = 0, \\ \phi^0_y = 0, & y = 1, \\ \phi^0_t + \dfrac{(\phi^0_y)^2}{2} + \eta^0 = 0, & y = 1. \end{cases}$$

Les trois premières équations montrent que ϕ^0 est indépendant de la variable y. Posons donc

$$\phi^0(x, y, t) = F(x, t).$$

La dernière équation implique une condition sur la surface

$$F_t + \eta^0 = 0, \quad y = 1.$$

L'équation d'ordre supérieur est

$$\begin{cases} \phi^0_{xx} + \phi^1_{yy} = 0, & 0 < y < 1, \\ \phi^1_y = 0, & y = 0, \\ \phi^1_y = \eta^0_t, & y = 1, \\ \phi^1_t + \dfrac{(\phi^0_x)^2}{2} + \eta^1 + \phi^0_y \phi^1_y = 0, & y = 1. \end{cases}$$

De la première et deuxième équation, en tenant compte de l'expression de η^0, on obtient

$$\phi^1_y = -yF_{xx}, \quad 0 \leq y \leq 1.$$

Ainsi, de la troisième équation, nous obtenons

$$\phi^1_y = -F_{xx}, \quad y = 1,$$

et donc

$$F_{tt} - F_{xx} = 0.$$

En prenant la dérivée en temps de cette équation, nous obtenons que le premier ordre du profil de surface satisfait une équation des ondes

$$\eta^0_{tt} - \eta^0_{xx} = 0, \quad x \in \mathbb{R}.$$

Nous obtenons donc que la dynamique du premier ordre s'obtient comme la somme de deux profils $\eta^0(x, t) = n_1(x - t) + n_2(x + t)$.

Connaissant la dynamique du profil du premier ordre, revenons à (3.6) avec un changement de variables permettant de suivre le profil se déplaçant vers la droite,

ainsi qu'un changement d'échelle sur le temps

$$\begin{cases} \xi = x - t, \\ T = \varepsilon t. \end{cases}$$

Notons

$$T_0 = T, \quad T_1 = \varepsilon T, \quad T_2 = \varepsilon^2 T \dots$$

Nous obtenons une nouvelle équation sur ϕ

$$\begin{cases} \varepsilon\phi_{\xi\xi} + \phi_{yy} = 0, & 0 < y < 1 + \varepsilon\eta, \\ \phi_y = 0, & y = 0, \\ \phi_y = \varepsilon(-\eta_\xi + \varepsilon\eta_T + \varepsilon\phi_\xi\eta_\xi), & y = 1 + \varepsilon\eta, \\ -\phi_\xi + \varepsilon\phi_T + \dfrac{\varepsilon}{2}\phi_\xi^2 + \dfrac{\phi_y^2}{2} + \eta = 0, & y = 1 + \varepsilon\eta. \end{cases} \tag{3.7}$$

Considérons à nouveau un développement en série du potentiel et du profil

$$\phi(\xi, y, T) = \sum_{k=0}^{\infty} \varepsilon^k \phi^k(\xi, y, T), \quad \eta(\xi, T) = \sum_{k=0}^{\infty} \varepsilon^k \eta^k(\xi, T).$$

Le premier ordre de ce système est donné par

$$\begin{cases} \phi_{yy}^0 = 0, & 0 < y < 1, \\ \phi_y^0 = 0, & y = 0, \\ \phi_y^0 = 0, & y = 1, \\ -\phi_\xi^0 + \dfrac{(\phi_y^0)^2}{2} + \eta^0 = 0, & y = 1. \end{cases} \tag{3.8}$$

Les trois premières équations impliquent que

$$\phi^0 = f(\xi, T_0, T_1, T_2, \dots), \quad 0 \le y \le 1, \tag{3.9}$$

et, avec la condition au bord en $y = 1$, nous avons

$$\frac{\partial \phi^0}{\partial \xi} = \eta^0(\xi, T), \quad 0 \le y \le 1. \tag{3.10}$$

Le système du second ordre est donnée par

$$\begin{cases} \phi^0_{\xi\xi} + \phi^1_{yy} = 0, & 0 < y < 1, \\ \phi^1_y = 0, & y = 0, \\ \phi^1_y = -\eta^0_\xi, & y = 1, \\ \phi^0_{T_0} - \phi^1_\xi + \dfrac{(\phi^0_x)^2}{2} + \eta^1 + \phi^0_y \phi^1_y = 0, & y = 1. \end{cases}$$

Des trois premières équations, ainsi que de (3.10), nous déduisons que

$$\phi^1_y = -y f_{\xi\xi}(\xi, T_0, T_1, T_2, \ldots).$$

Ainsi,

$$\phi^1 = -\frac{y^2}{2} f_{\xi\xi}(\xi, T_0, T_1, T_2, \ldots) + g(\xi, T_0, T_1, T_2, \ldots).$$

La dernière équation de (3.10) donne finalement à la surface

$$\eta^1 = -f_{T_0} - \frac{1}{2} f_{\xi\xi\xi} + \beta_\xi - \frac{1}{2} f_\xi^2, \quad y = 1. \tag{3.11}$$

Le troisième ordre d'approximation s'écrit

$$\begin{cases} \phi^1_{\xi\xi} + \phi^2_{yy} = 0, & 0 < y < 1, \\ \phi^2_y = 0, & y = 0, \\ \phi^2_y = \eta^0 f_{\xi\xi} - \eta^1_\xi + \eta^0_{T_0} + \phi^0_\xi \eta^0_\xi, & y = 1, \\ \phi^0_{T_1} + \phi^1_{T_0} - \phi^2_\xi + \phi^0_\xi \phi^1_\xi + \dfrac{(\phi^1_y)^2}{2} + \eta^2 + \phi^0_y \phi^2_y = 0, & y = 1. \end{cases}$$

Par les deux premières équations, on a

$$\phi^2 = -\frac{y^4}{24} f_{\xi\xi\xi\xi}(\xi, T_0, T_1, T_2, \ldots) - \frac{y^2}{2} g_{\xi\xi}(\xi, T_0, T_1, T_2, \ldots) + h(\xi, T_0, T_1, T_2, \ldots).$$

Finalement, en utilisant (3.9), (3.10), (3.11) et la troisième ligne du dernier système, on obtient

$$\frac{1}{3} f_{\xi\xi\xi\xi} + 3 f_\xi f_{\xi\xi} + 2 f_{T_0\xi} = 0,$$

qui, réécrit en fonction de η^0, donne la version adimensionnée de l'équation de Korteweg-de Vries

$$2\eta^0_{T_0} + \frac{1}{3}\eta^0_{\xi\xi\xi} + 3\eta^0 \eta^0_\xi = 0.$$

En revenant aux variables physiques, nous obtenons

$$\frac{1}{\sqrt{g h_0}} \partial_t \eta + \partial_x \eta + \frac{h_0^2}{6} \partial_x^3 \eta + \frac{3}{2h_0} \eta \partial_x \eta = 0, \quad x \in \mathbb{R},\ t \in \mathbb{R}. \tag{3.12}$$

Le régime dans lequel l'équation de Korteweg-de Vries est valide est lorsque $\xi = O(1), T_0 = O(1)$, c'est-à-dire, dans les variables physiques, lorsque

$$x\frac{\varepsilon^{3/2}}{\lambda\delta} = O(1), \quad x - \sqrt{gh_0}t = O(1).$$

De plus, le champ de vitesse d'ordre un pour des profils donnés par l'équation de Korteweg-de Vries (3.12) est obtenu grâce à (3.10) ainsi qu'avec la condition d'incompressibilité

$$\begin{cases} u(x,y,t) = \sqrt{\dfrac{g}{h_0}}\eta(x,t), \\ v(x,y,t) = -y\sqrt{\dfrac{g}{h_0}}\eta_x(x,t). \end{cases}$$

Notons finalement que l'équation de Korteweg-de Vries peut être vue comme la restriction des vagues se déplaçant de gauche à droite de l'équation de Boussinesq ([131])

$$\frac{1}{gh_0}\partial_t^2\eta - \partial_x^2\eta - \frac{h_0^2}{3}\partial_x^4\eta - \frac{3}{h_0}\eta\partial_x\eta = 0, \quad x \in \mathbb{R},\ t \in \mathbb{R},$$

modèle de propagation de vagues lorsque la longueur des vagues est plus importante que la profondeur de l'eau. Lorsque écrit sous forme d'un système d'équations portant sur la hauteur de l'eau et le champ de vitesse du fluide, on peut obtenir les équations de Saint-Venant du système de Boussinesq en supposant que la profondeur de l'eau h_0 est suffisamment petite pour annuler l'effet de la dispersion des vagues ([131]).

3.2 Scattering de l'équation de Korteweg-de Vries

L'équation de Korteweg-de Vries est l'un des exemples les plus simples d'une équation exactement résoluble. La dynamique de la solution $y(x,t)$ peut être obtenue en résolvant deux équations linéaires dépendant de la dynamique des données scatterées d'une donnée initiale y^0, si cette dernière est suffisamment régulière. Afin de donner une brève présentation de cette méthode, considérant l'équation de Korteweg-de Vries

$$\begin{cases} y_t + 6yy_x + y_{xxx} = 0, & (x,t) \in \mathbb{R} \times \mathbb{R}^+, \\ y(x,0) = y^0(x), & x \in \mathbb{R}, \end{cases} \tag{3.13}$$

obtenue de (3.2) par le changement de variables $\eta \mapsto y - 2h_0/3$, $t \mapsto \sqrt{gh_0}t$, $x \mapsto \sqrt[3]{\frac{6}{h_0^2}}x$. Nous renvoyons notamment à [54] pour une étude plus approfondie de la méthode présentée ici.

En considérant la transformation de Miura $y = \theta^2 + \theta_x$ et en remplaçant cette

expression dans (3.13), on obtient

$$\big(2\theta + \partial_x\big)\big(\theta_t + 6\theta^2\theta_x + \theta_{xxx}\big) = 0,$$

et, par conséquent, si θ est solution de

$$\theta_t + 6\theta^2\theta_x + \theta_{xxx} = 0,$$

alors $y = \theta^2 + \theta_x$ définit une solution de (3.13). Le changement de variable $\theta = \psi_x/\psi$ permet de retrouver une équation linéaire

$$\psi_{xx} + y\psi = 0,$$

ou, plus précisément, après le changement de variables $y(x,t) \mapsto \lambda + y(x+6\lambda t, t), \lambda \in \mathbb{R}$, l'équation de Sturm-Liouville

$$\psi_{xx}(x;t) + (\lambda + y(x;t))\psi(x;t) = 0. \tag{3.14}$$

La notation $(x;t)$ est utilisée ici afin d'illustrer que la variable de temps est considérée comme un paramètre de l'équation. Notons que l'équation (3.14) est également rencontrée fréquemment en mécanique quantique puisqu'elle correspond à l'équation stationnaire de Schrödinger, où la fonction d'onde ψ décrit la dynamique quantique d'une particule d'énergie λ soumise à un potentiel électrique $y(x;t)$.

Supposons que la donnée initiale de (3.13) appartient à l'espace de Faddeev $L^1_1(\mathbb{R})$, soit l'espace des fonctions mesurables à valeurs réelles telles que

$$\int_{-\infty}^{\infty} (1 + |x|)|y(x)|\mathrm{d}x < \infty.$$

Cette hypothèse est suffisante pour conclure que l'équation de Sturm-Liouville (3.14) admet un spectre continu $\lambda > 0$ et un nombre fini de valeurs propres $\lambda < 0$, notées $\lambda := -\kappa_n$. Notons que le cas $\lambda = 0$ n'arrive que si $y \equiv 0$. La résolution de (3.13) se fait de la manière suivante. D'une donnée initiale dans $L^1_1(\mathbb{R})$, on obtient les données scatterées associées aux fonctions propres en considérant le comportement asymptotique lorsque $x \to \infty$ des fonctions propres solutions de (3.14)

$$\psi_n(x) \simeq c_n e^{-\kappa_n x}, \quad x \to \infty.$$

On obtient également les données scatterées associées aux fonctions propres du spectre continu en considérant le comportement asymptotique

$$\tilde{\psi}(x) \simeq \begin{cases} e^{-ikx} + b(k)e^{ikx}, & x \to \infty, \\ a(k)e^{-ikx}, & x \to -\infty, \end{cases}$$

où $k = \sqrt{\lambda}$. La dynamique des données scatterées est régie par des équations aux dérivées ordinaires dont les solutions sont

$$\begin{cases} a(k;t) = a(k;0) \\ b(k;t) = b(k;0)e^{i8k^3t} \\ c_n(t) = c_n e^{4\kappa_n^3 t}. \end{cases}$$

La solution au temps $t > 0$ de (3.13) est alors obtenue en résolvant l'équation de Marchenko

$$K(x,z;t) + F(x+z;t) + \int_x^\infty K(x,y;t)F(y+z;t)\mathrm{dy} = 0,$$

où F est définie par

$$F(x) := \sum_{n=1}^{N} c_n^2(0)e^{8\kappa_n^3 t - \kappa_n x} + \frac{1}{2\pi}\int_{-\infty}^{\infty} b(k;0)e^{i8k^3t+ikx}\mathrm{dk}.$$

Pour K donné, la solution de (3.13) est finalement obtenue :

$$y(x,t) = 2\frac{d}{dx}K(x,x;t).$$

On désigne par les données de transmissions les données $c_n(t)$ et par les données de réflexions les données associées à $a(k;t)$ et $b(k;t)$. Ces noms font références au comportement de la solution de (3.13) issu de ces coefficients. Si seul un spectre discret de N valeurs propres existe pour des solutions de (3.14), alors la solution de (3.14) correspond à la solution de N-solitons. Un soliton de (3.13) est donné par

$$y(x,t) = \frac{\alpha^2}{2}\operatorname{sech}^2\left(\frac{-\alpha(x-s)+\alpha^3 t}{2}\right), \tag{3.15}$$

$\alpha > 0$ and $s \in \mathbb{R}$. Notons les propriétés suivantes d'un soliton.

1. L'amplitude d'un soliton est de $\alpha^2/2$ et est atteinte à $x = s + \alpha^2 t$;
2. la vitesse de déplacement est de α^2 ;
3. la largeur, définie par la distance dans la variable x où la hauteur d'un soliton est de $\alpha/4$, est
$$w(\alpha) := \frac{4}{\alpha}\ln(\sqrt{2\alpha}(1+\sqrt{1-\frac{1}{2\alpha}})). \tag{3.16}$$

De ces propriétés nous déduisons facilement qu'un soliton de plus grande amplitude se déplace plus rapidement et est plus étroit. Il est important de souligner qu'il n'existe pas d'approximation linéaire d'un soliton.

Soit $N \in \mathbb{N}$. La solution à N-solitons, telle que présentée par Hirota en 1971 ([80]),

est donnée par l'expression suivante

$$\begin{aligned} y(x,t) &= -2\frac{\partial^2}{\partial x^2}\ln F(x,t), \\ F &= 1+\sum_{n=1}^{N}\sum_{{}_NC_n} a(i_1,...,i_n)\prod_{j=1}^{n} f_{i_j}, \end{aligned} \tag{3.17}$$

où, pour $1 \leq i \leq N$

$$f_i(x,t) = \exp\left(-\alpha_i(x-s_i)+\alpha_i^3 t\right),\ s_i \in \mathbb{R}, \alpha_i > 0, \tag{3.18}$$

où

$$\begin{aligned} a(i_1,...,i_n) &= \prod_{k<l}^{(n)} a(i_k,i_l), \qquad \text{si } n \geq 2, \\ a(i_k,i_l) &= \left(\frac{\alpha_{i_k}-\alpha_{i_l}}{\alpha_{i_k}+\alpha_{i_l}}\right)^2, \\ a(i_k) &= 1, \end{aligned}$$

où $\sum_{{}_NC_n}$ est la somme sur tous les n indices $i_1,...,i_n$, où $i_1,...,i_n$ sont tous les choix possibles sans permutations de n indices parmi $\{1,\ldots,N\}$, et où

$$\prod_{k<l}^{(n)},$$

est le produit de toutes ces combinaisons de ces n éléments, satisfaisant $k<l$. Notons que, pour la solution à N-solitons, tous les coefficients α_i doivent être différents. Sinon, la solution à $(N-m)$-solitons sera obtenue où m est le nombre de fois que $\alpha_i = \alpha_j, 1 \leq i < j \leq N$. Afin d'expliquer pourquoi cette solution s'appelle la solution à N-solitons, considérons le cas $N=2$. La solution s'écrit

$$\frac{y}{2} = \frac{\alpha_1^2 f_1 + \alpha_2^2 f_2 + 2(\alpha_2-\alpha_1)^2 f_1 f_2 + ((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2(\alpha_2^2 f_1^2 f_2 + \alpha_1^2 f_2^2 f_1)}{(1+f_1+f_2+((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2 f_1 f_2)^2}.$$

Si $f_1 \simeq 1$, $f_2 \ll 1$, le comportement de y est alors donné par

$$\begin{aligned} y &\simeq 2\frac{\alpha_1^2 f_1}{(1+f_1)^2} \\ &= \frac{\alpha_1^2}{2}\operatorname{sech}^2\left(\frac{-\alpha_1(x-s_1)+\alpha_1^3 t}{2}\right), \end{aligned}$$

alors que, dans le cas où $f_2 \simeq 1$, $f_1 \gg 1$, nous avons

$$\begin{aligned} y &\simeq 2\frac{\alpha_2^2((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2 f_1^2 f_2}{(f_1+((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2 f_1 f_2)^2} \\ &= 2\frac{\alpha_2^2 \exp(\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2))f_2}{(1+\exp(\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2))f_2)^2} \\ &= \frac{\alpha_2^2}{2}\operatorname{sech}^2\left(\frac{-\alpha_2(x-s_2-\frac{1}{\alpha_2}\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2))+\alpha_2^3 t}{2}\right), \end{aligned}$$

c'est-à-dire, un soliton avec un changement de phase de $\frac{1}{\alpha_2}\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2)$.

Décrivons maintenant la dynamique générale de la solution à 2-solitons dans le cas où $0 < \alpha_1 < \alpha_2$. Nous ferons référence à chacun de ces solitons comme le soliton 1 et le soliton 2. Lorsque $t \to -\infty$, le soliton 1 est devant le soliton 2 et ce dernier possède un changement de phase $\frac{1}{\alpha_2}\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2)$. Lorsque $f_1 \simeq f_2 \simeq 1$, une interaction se produit entre les deux solitons. Si α_1 et α_2 sont du même ordre ($\alpha_2/\alpha_1 < 3$), alors les deux solitons restent distincts durant l'interaction. Le soliton 2 donne de sa masse au soliton 1 jusqu'au moment où le soliton 1 atteint l'amplitude initiale du soliton 2, et inversement. Si les deux solitons ne sont pas de la même amplitude, ils se fusionnent en un seul soliton pendant le moment de l'échange et le soliton 1 en ressort avec l'amplitude du soliton 2, et inversement [134]. Suite à cette interaction, les solitons conservent leur même forme, le seul effet de l'interaction étant un changement de phase du soliton 1, de $\frac{1}{\alpha_1}\ln(((\alpha_2-\alpha_1)/(\alpha_2+\alpha_1))^2)$, alors que le soliton 2 n'a plus de changement de phase. La Figure 2 illustre l'interaction entre deux solitons de même ordre ainsi que l'effet du changement de phase.

Le spectre discret quant à lui mène à un train de vagues oscillant autour de $y = 0$ et se dispersant vers la gauche. Seul un spectre discret existe si $y^0(x) \leq 0$. Le comportement du train de vagues est similaire à la solution de l'équation linéaire : les vagues de plus fortes amplitudes se situent à droite alors que les vagues de plus faibles amplitudes se situent à gauche. Il peut être montré que le sommet ou le creux de chacune des vagues décroît comme $t^{-1/3}$ et, par conséquent, que l'enveloppe du train de vagues décroît comme $t^{-p}, p > 0$ [120]. La décroissance n'est pas en $t^{-1/3}$ puisque des vagues peuvent être créées au fur et à mesure de la propagation du train de vagues mais la borne en $t^{-1/3}$ du sommet ou des creux des vagues assure la décroissance algébrique de l'enveloppe du train de vagues.

Finalement, pour toute donnée initiale appartenant à l'espace de Faddeev, on obtient le comportement asymptotique suivant : un nombre fini de solitons émergeant de la donnée initiale et se déplacent vers la droite alors qu'un train dispersif décroît vers la gauche [120]. Une borne inférieure du nombre N de solitons émergeant d'une

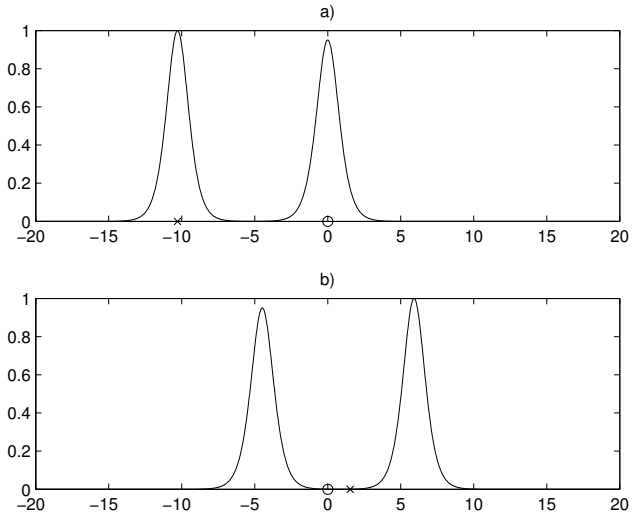

FIGURE 2 – Intéraction entre deux solitons. Le soliton à la gauche de l'image a) possède une vitesse de propagation plus grande que le soliton de droite. L'image b) représente les solitons après le dépassement. Le cercle et la croix dénote la position qu'aurait eue les solitons sans cette interaction

perturbation initiale est donnée par

$$N \leq 1 + \int_{-\infty}^{\infty} |x|(y^0(x))^+ \mathrm{d}x,$$

où $(y^0(x))^+$ est la partie positive de y^0.

3.3 Contributions à la contrôlabilité lagrangienne de l'équation de Korteweg-de Vries

3.3.1 Problématique

Considérons l'équation de Korteweg-de Vries posée sur l'intervalle borné $[0, L]$, avec $T, L > 0$,

$$y_t + y_x + y_{xxx} + yy_x = 0, \qquad x \in [0, L] \quad t \in [0, T], \tag{3.19}$$
$$y(0, t) = u(t), \qquad t \in [0, T], \tag{3.20}$$
$$y(L, t) = v(t), \qquad t \in [0, T], \tag{3.21}$$
$$y_x(L, t) = w(t), \qquad t \in [0, T], \tag{3.22}$$
$$y(x, 0) = 0, \qquad x \in [0, L], \tag{3.23}$$

où $u(t), v(t)$ et $w(t) \in \mathbb{R}$ sont les contrôles frontières. Rappelons d'abord la définition de la contrôlabilité de (3.19)-(3.23).

Definition 3.1 *Soit $T, L > 0$. L'équation (3.19)-(3.22) est dite contrôlable en temps $T > 0$ si pour toute donnée initiale $y_0 \in L^2(0, L)$ et finale $y_T \in L^2(0, L)$, il existe des contrôles u, v, w telle que l'équation (3.19)-(3.22) partant de l'état initial $y(., 0) = y_0$ satisfasse $y(., T) = y_T$.*

Dans cette thèse, nous nous intéresserons à la notion de contrôlabilité lagrangienne de (3.19)-(3.23) suivante

Definition 3.2 (Contrôlabilité lagrangienne en temps petit) *L'équation (3.19)-(3.23) est dite contrôlable en coordonnées lagrangiennes en temps petit si et seulement si, pour tout $T, L > 0$, il existe des contrôles $u(t), v(t), w(t) \in \mathbb{R}$ tels que, si on considère $\hat{y}$ l'extension par les valeurs aux bords de la solution y de (3.19)-(3.23)*

$$\hat{y}(x,t) = \begin{cases} y(0,t), & \text{if } x \leq 0, \\ y(x,t), & \text{if } x \in [0, L], \\ y(L,t), & \text{if } x \geq L, \end{cases} \tag{3.24}$$

le flot Φ défini par

$$\begin{cases} \dfrac{\partial \Phi}{\partial t}(x,t) &= \hat{y}(\Phi(x,t), t), & x \in \mathbb{R}, t \in \mathbb{R}^+, \\ \Phi(x,0) &= x, & x \in \mathbb{R}, \end{cases} \tag{3.25}$$

satisfait $\Phi(x, T) \geq L$, $\forall x \in [0, L]$.

Remarque 3.1 *Afin de simplifier les notations, nous considérerons toujours implicitement cette extension par les valeurs aux bords de la solution de (3.19)-(3.23) lorsque nous parlerons de flot de (3.19)-(3.23).*

Plutôt que de modifier l'état final de la solution à l'aide de contrôles, la contrôlabilité lagrangienne consiste à modifier l'évolution d'un système dynamique, dépendant de la solution, par le biais de contrôles. Au vu des résultats de la Section 3.1.1, le système dynamique considéré ici correspond à la trajectoire horizontale des particules initialement situées sous la surface de l'eau dans la région $[0, L] \times [0, h_0]$.

Le résultat que j'ai obtenu pendant cette thèse est la contrôlabilité lagrangienne de (3.19)-(3.23) avec la condition supplémentaire que la solution soit nulle à l'instant final.

Théorème 3.1 ([61]) *Soit $T, L > 0$. Alors il existe une solution $y \in C([0,T]; H^2(0,L))$ de (3.19)-(3.23) telle que le flot associé Φ satisfait*

$$\begin{aligned} \frac{\partial \Phi}{\partial t}(x,0) &= 0, \qquad x \in [0,L], && (3.26)\\ \frac{\partial \Phi}{\partial t}(x,T) &= 0, \qquad x \in [0,L], && (3.27)\\ \Phi(x,t) &\geq L, \qquad x \in [0,L],\ t \in [T,\infty). && (3.28) \end{aligned}$$

3.3.2 Résultats antérieurs

Commençons par rappeler des résultats relatifs à la contrôlabilité lagrangienne et la contrôlabilité eulérienne de (3.19)-(3.23) par le rapprochement entre la contrôlabilité eulérienne et lagrangienne pour les équations d'Euler 2D. Cette présentation est en partie inspirée de [68].

Connexion entre la contrôlabilité eulérienne et lagrangienne : le cas de l'équation d'Euler 2D

Soit $\Omega \in \mathbb{R}^2$, un domaine ouvert, borné et à frontière régulière $\partial\Omega$. Supposons également, par simplicité, que Ω est simplement connexe. Considérons l'équation d'Euler incompressible

$$\begin{cases} \partial_t u + (u.\nabla)u + \nabla p = 0, & (x,t) \in \Omega \times [0,T], \\ \text{div } u = 0, & (x,t) \in \Omega \times [0,T], \end{cases} \tag{3.29}$$

où, suivant la notation de [41], la notation vectorielle pour le champ de vitesse $u : \Omega \times [0,T] \to \mathbb{R}^2$ est utilisée. On considère une partie Σ d'intérieur non vide de la frontière $\partial\Omega$ de Ω, partie où le contrôle agira, et une condition de non-pénétration est appliquée sur la partie restante,

$$u.n = 0, \quad (x,t) \in (\partial\Omega \setminus \Sigma) \times [0,T], \tag{3.30}$$

où n est le vecteur unité normal sortant de la frontière $\partial\Omega$.

La contrôlabilité eulérienne globale de cette équation est définie de la façon suivante,

Definition 3.3 *Soit $T > 0$ et soit $u_0, u_1 \in C^\infty(\overline{\Omega}; \mathbb{R}^2)$ tels que*

$$\begin{aligned} div\ u_0 = div\ u_1 &= 0, \quad \forall x \in \overline{\Omega}, && (3.31)\\ u_0.n = u_1.n &= 0, \quad \forall x \in (\partial\Omega \setminus \Sigma). && (3.32) \end{aligned}$$

On dit que (3.29) *est exactement contrôlable en variables eulériennes si et seulement si, pour tout* $u_0, u_1 \in C^\infty(\overline{\Omega}; \mathbb{R}^2)$ *satisfaisant* (3.31)-(3.32)*, il existe* $u \in C^\infty(\overline{\Omega} \times [0,T]; \mathbb{R}^2)$ *et* $p \in C^\infty(\overline{\Omega} \times [0,T]; \mathbb{R})$ *solution de* (3.29)-(3.30) *satisfaisant*

$$\begin{aligned} u(x,0) &= u_0(x), \quad \forall x \in \overline{\Omega} \\ u(x,T) &= u_1(x), \quad \forall x \in \overline{\Omega} \end{aligned}$$

La définition des contrôles ne figure pas dans cette définition afin d'éviter des difficultés techniques. Notons que les contrôles peuvent être, par exemple, la vitesse normale $u.n$ sur $\Sigma \times [0,T]$, en assurant que la moyenne est nulle afin que la condition d'incompressibilité soit respectée et le rotationnel rot u sur la partie rentrante du domaine

$$\Sigma_T^- := \{(x,t) \in \partial\Omega \times [0,T] \mid u(x,t).n(x) < 0\}$$

La contrôlabilité eulérienne exacte de l'équation d'Euler (3.29)-(3.30) a été établie par Jean-Michel Coron.

Théorème 3.2 ([41]) *Supposons que la région de contrôle* Σ *intersecte chacune des composantes connexes de* $\partial\Omega$*. Alors l'équation d'Euler* (3.29)-(3.30) *est exactement contrôlable en coordonnées eulérienne.*

La stratégie de la preuve repose sur la méthode du retour, introduite pour la première fois dans [39] pour obtenir la stabilisation asymptotique globale pour des systèmes en dimension finie sans terme de dérive et pour la première fois pour la contrôlabilité d'une équation aux dérivées partielles pour traiter la contrôlabilité de (3.29)-(3.30). Rappelons que la méthode du retour est une alternative afin d'étudier la contrôlabilité locale d'une équation lorsque le linéarisé autour d'un point d'équilibre, ici l'état nul, n'est pas contrôlable. Dans le cas de (3.29)-(3.30), le linéarisé autour de zéro est donné par

$$\begin{cases} \partial_t u + \nabla p = 0, & (x,t) \in \Omega \times [0,T], \\ \text{div } u = 0, & (x,t) \in \Omega \times [0,T], \end{cases}$$

et, en prenant le rotationnel, on obtient

$$\frac{\partial \text{rot } u}{\partial t} = 0.$$

La dynamique du rotationnel de u ne pouvant être modifiée, l'équation linéarisée n'est pas contrôlable.

Afin d'utiliser la méthode du retour, la trajectoire $\overline{u}$ est construite à partir des écoulements potentiels, c'est-à-dire, en cherchant la solution sous la forme

$$\overline{u}(x,t) = \alpha(t)\nabla\phi(x), \tag{3.33}$$

où ϕ est un potentiel harmonique $\phi \in C^\infty(\overline{\Omega})$ et où $\alpha \in C^\infty([0,T])$ s'annule en $t=0$ et $t=T$. Notons que $\overline{u}$ est une trajectoire des équations d'Euler en choisissant la pression de la forme

$$\overline{p}(x,t) = -\alpha'(t)\theta(x) - \frac{1}{2}\alpha^2(t)\left|\nabla\theta(x)\right|^2.$$

La condition de non pénétration se traduit sur le potentiel par

$$\frac{\partial\phi}{\partial n} = 0, \quad \forall x \in (\partial\Omega \setminus \Sigma).$$

Afin de caractériser les trajectoires $(\overline{u}, \overline{p})$ permettant d'obtenir la contrôlabilité locale autour de ces trajectoires, écrivons l'équation sur la vorticité du champ de vitesse de (3.29), $\omega := \text{rot } u$

$$\partial_t\omega + (u.\nabla)\omega = 0.$$

Supposons pour l'instant que le linéarisé autour de $\overline{u}$ s'écrit

$$\partial_t\omega + (\overline{u}.\nabla)\omega = 0.$$

La vorticité étant constante le long du flot engendré par $\overline{u}$, il suffit de construire $\overline{u}$ de sorte que la vorticité initialement située à l'intérieur du domaine sorte et que la vorticité rentrante est nulle. Le potentiel de vitesse doit alors satisfaire les conditions suivantes

$$\Delta\phi = 0, \quad \forall x \in \overline{\Omega}, \tag{3.34}$$

$$\nabla\phi \neq 0, \quad \forall x \in \overline{\Omega}, \tag{3.35}$$

$$\frac{\partial\phi}{\partial n} = 0, \quad \forall x \in (\partial\Omega \setminus \Sigma), \tag{3.36}$$

$$\Phi(x,T) \notin \overline{\Omega}, \tag{3.37}$$

pour le flot Φ défini par la trajectoire de référence $\overline{u}$

$$\begin{cases} \dfrac{\partial\Phi}{\partial t}(x,t) = \overline{u}(\Phi(x,t),t), & x \in \overline{\Omega}, t \in [0,T], \\ \Phi(x,0) = x, & x \in \overline{\Omega}, \end{cases}$$

La contrôlabilité eulérienne a donc été ramenée à un problème de contrôlabilité lagrangienne.

Il a été montré dans [41] qu'il existe un tel potentiel satisfaisant (3.34)-(3.37). En imposant que le rotationnel soit nul pour le fluide rentrant par la portion contrôlée du domaine, on obtient que l'état final a un rotationnel nul. Il est alors facile d'amener cet état à 0. La contrôlabilité locale à zéro est obtenue par un argument perturbatif et par un argument de point fixe. La contrôlabilité globale à zéro est obtenue par

un argument de changement d'échelle grâce au terme quadratique de l'équation non linéaire. En effet, si (u, p) est solution de l'équation, alors $(\lambda u(x, \lambda t), \lambda^2 p(x, \lambda t))$ est également solution de l'équation sur l'intervalle de temps $[0, T/\lambda]$. La contrôlabilité locale à zéro implique donc la contrôlabilité globale à zéro. La contrôlabilité globale pour tout état final est finalement obtenue par réversibilité en temps de l'équation.

Une notion plus forte de contrôlabilité lagrangienne pour l'équation d'Euler (3.29) a été établie après le résultat de contrôlabilité eulérienne globale.

Definition 3.4 (Contrôlabilité lagrangienne approchée dans C^k) *Soit $T > 0$, $k \in \mathbb{N}$ et $\varepsilon > 0$. Étant donné deux courbes de Jordan γ_1 et γ_2 dans Ω et un état initial u_0, existe-t-il un contrôle tel que le flot Φ définie par u solution de* (3.29) *satisfasse, à une reparamétrisation de la courbe près*

$$\|\Phi(T, \gamma_1) - \gamma_2\|_{C^k} \leq \varepsilon.$$

La construction de la trajectoire (3.33) permet d'obtenir un résultat de contrôlabilité lagrangienne si on permet à la zone du fluide entourée par les courbes de Jordan γ_1 et γ_2 de sortir du domaine Ω. Or, Olivier Glass et Thierry Horsin ont établi le résultat de contrôlabilité lagrangienne en s'assurant que la zone à contrôler reste à l'intérieur du domaine Ω. En notant, $|\mathrm{Int}(\gamma)|$ la mesure de Lebesgue de l'intérieur de la courbe de Jordan γ, le résultat est le suivant.

Théorème 3.3 ([71]) *Soit $T > 0$, $k \in \mathbb{N}$ et $\varepsilon > 0$. Soit $u_0 \in C^\infty(\overline{\Omega}; \mathbb{R}^2)$ satisfaisant,*

$$\begin{aligned} div(u_0) &= 0, \quad x \in \Omega, \\ u_0.n &= 0, \quad x \in (\partial\Omega \setminus \Sigma). \end{aligned}$$

Étant donné deux courbes de Jordan γ_1 et γ_2 telles que

$$|Int(\gamma_1)| = |Int(\gamma_2)| \tag{3.38}$$

et telles que γ_1 et γ_2 soient homotopiques. Alors, il existe $u \in C^\infty([0, T] \times \overline{\Omega}; \mathbb{R}^2)$ solution de (3.29) *telle que le flot associé à u satisfasse*

$$\Phi(t, \gamma_1) \subset \Omega, \quad \forall t \in [0, T],$$

et, à une reparamétrisation de la courbe près

$$\|\Phi(T, \gamma_1) - \gamma_2\|_{C^k} \leq \varepsilon.$$

Notons que la condition (3.38) vient du fait que (3.29) modélise des fluides incompressibles. Si ce résultat de contrôlabilité lagrangienne est plus fort, une similitude dans la preuve reste : la construction de solutions sous la forme d'écoulements potentiels

ϕ satisfaisant (3.34)-(3.36) et, cette fois-ci,

$$\Phi(x,t) \in \overline{\Omega}, \quad \forall t \in [0,T].$$

Autres résultats de la littérature

Quelques résultats de contrôlabilité lagrangienne sont trouvés dans la littérature. En plus de la contrôlabilité lagrangienne pour les équations d'Euler en deux dimensions, Glass et Horsin ont montré un résultat similaire à celui présenté ici pour l'équation d'Euler en trois dimensions ([72]). Horsin a montré dans [81] la contrôlabilité lagrangienne pour l'équation de la chaleur en une dimension. En considérant deux intervalles fermés, Horsin a montré qu'il existe un contrôle frontière telle que le flot de la solution associée partant du premier intervalle atteint le second. Il a également été montré dans le même article que pour des domaines radiaux (respectivement un domaine convexe) de dimension supérieure, il est possible de déplacer deux ensembles fermés réguliers avec le flot gradient grâce à un contrôle en temps arbitrairement petit (respectivement un temps suffisamment grand). Finalement, la contrôlabilité lagrangienne de l'équation de Burgers visqueuse a été montré par Horsin dans le cas où les deux intervalles fermés ne sont pas trop éloignés ([82]).

Nous soulignons également le résultat de Rosier sur la contrôlabilité locale autour de trajectoires régulières de l'équation de Korteweg-de Vries écrite en coordonnées lagrangiennes [94].

Nous retrouvons plusieurs résultats de contrôlabilité eulérienne pour l'équation de Korteweg-de Vries. Lorsque seul le contrôle w est utilisé ($u, v \equiv 0$), la contrôlabilité locale de (3.19)-(3.23) a été obtenue par Rosier ([116]) si $L \notin \mathcal{N}$, soit l'ensemble des longueurs critiques telles que l'équation linéarisée autour de zéro est contrôlable. Lorsque $L \in \mathcal{N}$, il existe un sous-espace de $L^2(0,L)$ non atteignable pour l'équation linéarisé autour de zéro. En utilisant la méthode d'expansion en série de puissance ([42]) afin d'atteindre ce sous-espace, la contrôlabilité locale autour de zéro de (3.19)-(3.23) a été obtenu par Coron et Crépeau ([45]), Cerpa ([27]) et Cerpa et Crépeau ([30]) lorsque la dimension du sous-espace non atteignable est de dimension 1, 2 et arbitraire respectivement. Si seulement le contrôle u est utilisé ($v, w \equiv 0$), alors (3.19)-(3.23) est localement contrôlable à zéro ([69]). Si seulement est utilisé v ($u, w \equiv 0$), alors (3.19)-(3.23) est localement contrôlable autour de zéro si L n'appartient pas à l'espace de longueurs critiques $\mathcal{O}$ ([70]). Si les contrôles v et w sont utilisés ($u \equiv 0$), alors il a été démontré dans [116] que (3.19)-(3.23) est contrôlable en temps petit. Toute autre combinaison de deux contrôles mène également à la contrôlabilité en temps petit ([69]). Finalement, la contrôlabilité aux trajectoires régulières de l'équation de Korteweg-de Vries a été obtenue en utilisant les trois contrôles ([135]). Nous renvoyons à [118, 28] pour de nombreux résultats sur la contrôlabilité ainsi que sur la stabilisabilité de (3.19)-(3.23).

On trouve deux résultats de contrôlabilité globale pour l'équation de Korteweg-de

Vries. Rosier a montré dans [117] que (3.19)-(3.23) est globalement contrôlable pour un temps $T > 0$ dépendant de la taille de la donnée initiale. La stratégie consiste simplement à attendre que l'énergie du système décroisse afin d'obtenir un état pour lequel la contrôlabilité locale peut être utilisée. En considérant un quatrième contrôle comme terme source de l'équation de Korteweg-de Vries

$$y_t + y_x + y_{xxx} + yy_x = a(t), \quad x \in [0, L], t \in [0, T], \tag{3.39}$$

Chapouly a montré dans [32] la contrôlabilité globale en temps petit de (3.39), (3.20), (3.21), (3.22) et (3.23) en utilisant la contrôlabilité de l'équation de Burgers non visqueuse ([33]).

3.3.3 Méthodes utilisées

La preuve de la contrôlabilité lagrangienne de (3.19)-(3.23) est similaire à la construction de la trajectoire de référence pour la méthode du retour de l'équation d'Euler (3.29) : des vagues rentrent d'un bord, en $x = 0$, se propagent à l'intérieur du domaine $[0, L]$ et sortent ensuite en $x = L$, entraînant avec elles les particules initialement situées à l'intérieur du domaine. Cette trajectoire de l'équation de Korteweg-de Vries n'est pas construite à l'aide d'un potentiel de vitesse mais à partir de l'expression explicite de la solution à N-solitons. Afin de construire cette trajectoire, notée $\tilde{y}$ dans cette introduction, étendons le domaine $[0, L]$ à $\mathbb{R}$ tout entier et considérons, afin d'alléger la présentation, l'équation de Korteweg-de Vries (3.13) plutôt que (3.19) sur $\mathbb{R}$, équation pour laquelle l'expression de la solution à N-solitons a déjà été donnée. Le changement de variables $x \mapsto x - t$ et $y \mapsto 6y$ permet de passer de (3.13) à (3.19) en conservant la donnée initiale $y_0 = 0$.

Un calcul direct montre que le déplacement d'une particule soumise à un soliton est de l'ordre de $2/\alpha$ et s'effectue à une vitesse de α^2. Un seul soliton ne suffit donc pas à obtenir la contrôlabilité lagrangienne en temps arbitrairement petit pour un domaine de longueur arbitraire. Nous montrerons, en considérant N suffisamment grand et une amplitude α_i des solitons suffisamment grande, qu'il est possible de construire une solution à N-solitons permettant d'obtenir un déplacement arbitrairement grand de particules en temps arbitrairement petit. Notons que les solitons sont suffisamment espacés et ordonnés en ordre croissant de taille afin d'empêcher des interactions lorsque la solution évolue au fil du temps. Les interactions initiales possibles entre les solitons sont estimées au départ.

La trajectoire de référence $\tilde{y}$ permet d'obtenir un déplacement arbitraire en temps petit mais elle ne permet pas d'obtenir la contrôlabilité lagrangienne puisque la solution est non-nulle dans le domaine $[0, L]$ au temps initial et au temps final. Par contre, en choisissant les paramètres d'amplitudes α_i suffisamment grands, la trajectoire $\tilde{y}$ sera aussi petite que souhaitée dans un voisinage de temps de $t = 0$, période où les solitons ne sont pas encore entrés dans le domaine, ainsi que dans un voisinage

de temps de $t = T$, période où les solitons sont sortis, entraînant avec eux les particules. La contrôlabilité lagrangienne est donc obtenue en utilisant la contrôlabilité eulérienne locale en temps petit sur ces intervalles de temps. Pour $\varepsilon > 0$ aussi petit que souhaité, la contrôlabilité lagrangienne se décompose donc en trois phases (par abus de notations, nous notons $\tilde{y}$ la trajectoire de référence pour (3.19)).

Phase 1 ($t \in [0, \varepsilon]$) **:** les N solitons de la solution $\tilde{y}$ sont situés à la gauche du domaine $[0, L]$ et se déplacent vers $x = 0$. Pendant cette période, les contrôles u, v et w servent à amener la solution y de (3.19)-(3.23) du repos à l'état $\tilde{y}(x, \varepsilon)$, $x \in [0, L]$;

Phase 2 ($t \in [\varepsilon, T - \varepsilon]$) **:** durant cette période $y = \tilde{y}$ et les solitons entrent dans le domaine $[0, L]$, entraînent les particules, et sortent avec les particules à la droite de $[0, L]$. Les contrôles u, v et w sont construits pendant cette période avec les traces correspondantes de $\tilde{y}$;

Phase 3 ($t \in [T - \varepsilon, T]$) **:** lors de cette période finale, les solitons sont situés à la droite de $x = L$ et l'état $\tilde{y}(x, T - \varepsilon), x \in [0, L]$, est suffisamment petit afin d'utiliser encore une fois la contrôlabilité locale. La solution y est donc amenée au repos au temps $t = T$.

Deux points techniques restent à préciser. Afin de définir le flot point par point, nous utiliserons un théorème de contrôlabilité eulérienne local en temps petit régulier [135], plutôt que les théorèmes classiques dans les espaces $L^2(0, L)$ ou $H^{-1}(0, L)$. La solution construite aura donc la régularité $H^2(0, L)$ en espace, régularité suffisante afin d'obtenir la définition du flot point par point. Finalement, puisqu'il n'existe pas de principe du maximum pour l'équation de Korteweg-de Vries, nous accepterons que les particules puissent se déplacer à la gauche de $x = 0$ pendant les temps de contrôlabilité. Grâce aux estimés du contrôle au bord $x = 0$ en fonction de la donnée initiale, nous pourrons estimer le déplacement vers la gauche des particules. Nous considérerons donc $\tilde{y}$ comme étant la solution à N-solitons permettant d'obtenir le déplacement des particules d'un domaine de longueur $L + \delta$, où δ majore le déplacement des particules pour les deux phases de contrôlabilité eulérienne.

3.3.4 Commentaires sur le résultat obtenu

Tel que mentionné précédemment, la construction de la solution menant à la contrôlabilité lagrangienne, avec les conditions $\tilde{y}(., 0) = \tilde{y}(., T) = 0$ rappelle la construction de la trajectoire pour la méthode du retour pour l'équation d'Euler. Contrairement au cas de l'équation d'Euler, le nombre important de dérivées dans la partie linéaire semble empêcher l'utilisation de cette trajectoire afin d'utiliser la méthode du retour : lorsque α tend vers l'infini, le terme linéaire ne peut être vu comme une perturbation du terme quadratique. Notons que dans la construction de la trajectoire $\tilde{y}$, les interactions entre les solitons ont été évitées afin de simplifier l'analyse du déplacement des particules. Une technique afin d'obtenir la contrôlabilité

globale de (3.19)-(3.23) serait plutôt de permettre aux solitons d'interagir afin de minimiser l'apport de la partie linéaire de l'équation. Cette approche semble toutefois difficile. En effet, pour qu'un plateau se forme d'interactions de solitons par exemple, cela nécessiterait des solitons d'ordres différents, impliquant des vitesses d'ordres différentes et des solitons de plus en plus étroits.

À ce jour, la contrôlabilité globale de l'équation de Korteweg-de Vries demeure toujours une question ouverte :

Problème ouvert 3.1 *Soit $T > 0$ et $L > 0$. Pour toute donnée initiale $y_0 \in L^2(0, L)$ et toute donnée finale $y_T \in L^2(0, L)$, existe-t-il des contrôles u, v et w tels que la solution (3.19)-(3.22) avec $y(., 0) = y_0$ satisfasse*

$$y(., T) = y_T?$$

La définition de la contrôlabilité lagrangienne de l'équation de Korteweg-de Vries diffère en deux points avec la définition usuelle. Plutôt que de bouger une portion du fluide à l'intérieur du domaine, nous le faisons sortir d'une extrémité du domaine. La raison principale de cette définition provient du sens de propagation privilégié de l'équation de Korteweg-de Vries : un changement de variable $\xi = x - t$ a été introduit dans la section 3.1.1 pour suivre les vagues se déplaçant de gauche à droite. L'inverse scattering nous dit de plus que seules des oscillations se propagent vers la gauche. Il serait donc éventuellement possible de faire bouger le fluide vers la gauche mais seulement en temps long. De plus, une restriction sur une donnée initiale nulle a été imposée dans la définition de la contrôlabilité lagrangienne de (3.19)-(3.23). Pour le moment, seule la contrôlabilité lagrangienne en temps long à partir d'une donnée initiale quelconque peut être obtenue en laissant d'abord la solution évoluer librement jusqu'à ce que ce qui reste à l'intérieur du domaine $[0, L]$ soit suffisamment petit pour utiliser la contrôlabilité locale eulérienne. Rappelons que le comportement asymptotique de la solution assure que cela arrivera à partir d'un certain temps. En estimant le déplacement des particules vers la gauche durant cette période, on peut ensuite utiliser le résultat de contrôlabilité lagrangienne à partir de zéro afin de pousser les particules à la droite du domaine $[0, L]$. Notons cependant que, contrairement à la démonstration du Théorème 3.1, le déplacement des particules à la gauche du domaine $[0, L]$ peut être très grand, ce qui est peu souhaitable. Une stratégie différente pour la contrôlabilité lagrangienne de (3.19)-(3.22) à partir d'une donnée initiale quelconque est donc souhaitable.

3.4 Contributions à l'approximation d'ordre supérieur du champ de vitesse associé à la solution de N-solitons

3.4.1 Problématique

Le champ de vitesse obtenu dans la section 3.1.1 est un champ de vitesse d'ordre un de l'équation de Korteweg-de Vries. Si, au niveau de la modélisation, l'équation de Korteweg-de Vries est vue comme un modèle satisfaisant pour la propagation de vagues dans des canaux de faibles amplitudes ([36, 37]), le champ de vitesse d'ordre un associé comporte certains défauts. En particulier, le déplacement horizontal des particules ne dépend pas de la profondeur à laquelle se trouve la particule. Dans le cas d'une onde solitaire, il a été démontré expérimentalement dans [34] que cela ne reflète pas le comportement des particules soumises à une onde solitaire : le déplacement des particules situées plus haut dans le fluide est plus grand. La figure 3 illustre les expérimentations du mouvement de particules pour différentes vagues solitaires.

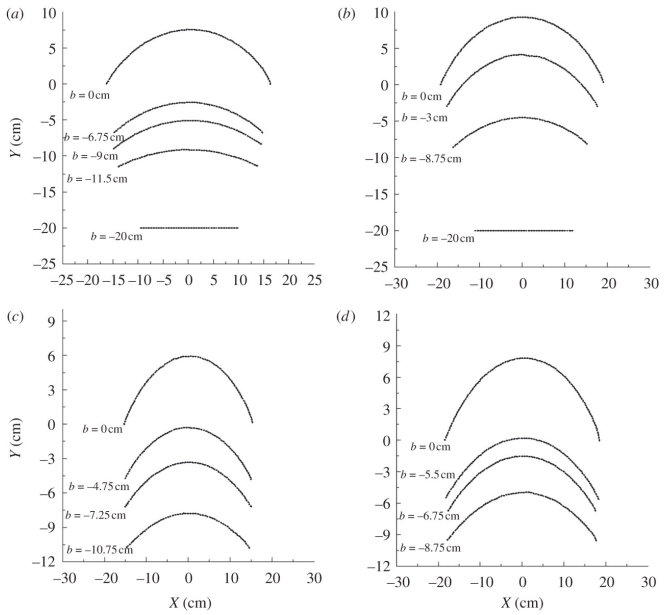

FIGURE 3 – La trajectoire de particules d'eau obtenue de mesures expérimentales du mouvement de pastilles de polystyrène à différente profondeur b dans quatre cas (a) h_0=20cm, a=7.07cm ; (b) h_0=20cm, a=8.56cm ; (c) h_0=30cm, a=5.46cm ; (d) h_0=30cm, a=7.56cm.

En suivant les travaux de Henry et de Constantin dans le cas de solutions périodiques et d'un soliton respectivement ([36, 78]), j'ai obtenu une meilleure approximation du champ de vitesse pour la solution à N-solitons.

Théorème 3.4 ([60]) *Étant donné N solitons, $N \in \mathbb{N}$, de paramètre d'amplitude a_i, $1 \leq i \leq N$. Pour des paramètres d'amplitudes $\varepsilon_i := a_i/h_0$ suffisamment petits*

$$\sum_{i=1}^{N} \sqrt{\frac{3a_i}{4h_0^3}}(h_0 + a_i) < \frac{\pi}{4}, \tag{3.40}$$

la trajectoire des particules décrite par le champ de vitesse d'ordre supérieur satisfait les assertions suivantes :

(a) Une particule située sur le fond de l'eau se déplace en ligne droite à vitesse positive ;

(b) Toutes les particules sous la surface possèdent une vitesse horizontale positive ;

De plus, la trajectoire de particules ne rencontrant pas d'interactions de solitons satisfont les assertions suivantes :

(c) Pour chacune de ces particules, il existe $t_1, \ldots, t_{2N-1}$ tels que la particule se déplace vers le haut pour les temps,

$$t \in (-\infty, t_1) \bigcup \left(\bigcup_{k=1}^{N-1} (t_{2k}, t_{2k+1}) \right),$$

et vers le bas pour les temps,

$$t \in \left(\bigcup_{k=1}^{N-1} (t_{2k-1}, t_{2k}) \right) \bigcup (t_{2N-1}, \infty);$$

(d) La position de chacune de ces particules au temps $t_1, t_3, \ldots, t_{2N-1}$ correspond au changement de phase typique de la solution à N-solitons.

3.4.2 Résultats précédents

Dans la littérature, nous retrouvons deux méthodes permettant d'obtenir une approximation d'ordre supérieur du champ de vitesse de l'équation de Korteweg-de Vries. En introduisant une fonction courant et en utilisant le fait que la fonction de courant et le potentiel de vitesse sont des conjugués harmoniques ayant comme condition aux bords le champ de vitesse d'ordre un, un champ de vitesse d'ordre supérieur a été obtenu par Constantin ([36]) et Henry ([78]) dans le cas d'un soliton et de solutions périodiques (solutions cnoïdales) de l'équation de Korteweg-de Vries respectivement. Des résultats qualitatifs sur la trajectoire des particules de ces champs de vitesse ont également été obtenus.

Il a été démontré dans le livre de Whitham ([131]), en déduisant l'équation de Korteweg-de Vries des équations d'Euler d'une manière différente que celle présentée

dans cette introduction, que l'équation de Korteweg-de Vries est en fait une approximation du second ordre. Ces travaux ont également montré que la composante horizontale du champ de vitesse du second ordre sur le fond s'écrit

$$u(x,0,t) = \sqrt{\frac{g}{h_0}}\left(\eta(x,t) - \frac{1}{4h_0}\eta^2(x,t) + \frac{h_0^2}{3}\eta_{xx}(x,t)\right).$$

Une analyse détaillée de ces travaux ont permis à Borluk and Kalisch d'obtenir, à partir de cette expression, un champ de vitesse d'ordre deux à l'intérieur du fluide ([16]). Des simulations numériques ont montré que la trajectoire des particules issues de ce champ de vitesse était plus précise dans le cas d'un soliton, de la solution à 2 solitons et pour les solutions périodiques. Ces travaux sont inspirés de travaux similaires effectués pour l'équation de Boussinesq ([2, 12]).

Notons finalement que, même si l'expression explicite de l'onde solitaire pour les équations d'Euler n'est pas connue, une étude qualitative de la trajectoire des particules pour une onde solitaire de cette équation a été établie [37], rejoignant les résultats obtenus dans [36] pour la trajectoire des particules induites par un soliton.

3.4.3 Méthode utilisée

Il est possible d'obtenir une meilleure approximation du déplacement des particules sans changer d'équation modélisant le phénomène en cherchant une approximation d'ordre supérieur du champ de vitesse. À cette fin, introduisons la fonction courant ψ associée à (3.4) satisfaisant

$$\psi_y = u, \quad \psi_x = -v.$$

Rappelons que pour (3.4), le potentiel de vitesse satisfait

$$\phi_x = u, \quad \phi_y = v.$$

Pour des fluides irrotationnels, le potentiel de vitesse et la fonction de courant sont des conjugués harmoniques. On identifie donc $\mathbb{R}^2$ à $\mathbb{C}$ et on introduit le potentiel de vitesse complexe

$$f(x,y,t) = \phi(x,y,t) + i\psi(x,y,t), \quad x+iy \in \mathbb{C},$$

où le temps t est vu comme un paramètre. En particulier

$$f'(x,y,t) = u(x,y,t) - iv(x,y,t), \quad x+iy \in \mathbb{C}.$$

Plutôt que d'utiliser les relations (3.8), le champ de vitesse d'ordre supérieur est obtenu en imposant le champ de vitesse d'ordre un comme condition au bord $y = 0$

du laplacien défini sur le demi-plan supérieur $\{x + iy \in \mathbb{C} \,\big|\, y \geq 0\}$

$$\phi_y(x, 0, t) = 0, \quad \phi_x(x, 0, t) = \sqrt{\frac{g}{h_0}}\eta(x, t).$$

La première équation reflète la condition de non-pénétration du fond alors que la seconde correspond au terme d'ordre zéro dans le développement en série par rapport à la hauteur y de la composante horizontale champ de vitesse ([36, 78]).

Les composantes du champ de vitesse s'expriment donc grâce à la formule de d'Alembert

$$\begin{aligned} u(x, y, t) &= \frac{1}{2}\sqrt{\frac{g}{h_0}}(\eta(x + iy, t) + \eta(x - iy, t)), \\ v(x, y, t) &= \frac{i}{2}\sqrt{\frac{g}{h_0}}(\eta(x + iy, t) - \eta(x - iy, t)). \end{aligned}$$

En remplaçant η par la solution à N-solitons, on obtient alors l'approximation d'ordre supérieur du champ de vitesse associé à cette solution. L'expression du champ de vitesse ainsi obtenu étant assez longue, mentionnons seulement que son expression est donnée par la somme de termes faisant intervenir différents produits des fonctions f_i, présentées en (3.18), ainsi que des coefficients dépendant de α_i. Le Théorème 3.4 s'obtient alors en étudiant l'expression explicite obtenue. Les conditions aux bords donnent directement la première assertion du théorème alors que l'hypothèse (3.40) permet d'assurer que la composante horizontale du champ de vitesse est positive sous la surface de l'eau.

Les deux dernières assertions du Théorème 3.4 s'obtiennent en étudiant le signe de la composante verticale du champ de vitesse. Cela se fait d'une manière semblable à l'interprétation de la solution à N-solitons présentée en section 3.2. En supposant que chacun des solitons sont éloignés les uns des autres, c'est-à-dire en supposant que si $f_k \simeq 1$, alors $f_i \ll 1, i < k$ et $f_i \gg 1, i > k$, l'étude du signe de v peut alors être effectuée. Un changement de signe de la composante verticale v a lieu au point où le soliton atteint son maximum, la composante v étant positive à droite du maximum du soliton et négative à gauche. Il est important de noter que la position du maximum des solitons prend en compte le changement de phase induit par l'interaction des solitons. Nous notons que bien qu'il ait été intéressant de décrire la trajectoire des particules issue du champ de vitesse d'ordre supérieur lors d'une interaction de solitons, la quantité importante de termes significatifs dans l'expression de la composante verticale du champ de vitesse empêche son étude.

3.4.4 Commentaires sur le résultat obtenu

Suite à la démonstration du Théorème 3.4, une simulation numérique de la trajectoire de particules issue du champ de vitesse d'ordre supérieur dans le cas de l'interaction entre deux solitons est présentée. La trajectoire obtenue numériquement concorde avec la trajectoire expérimentale. Il est également montré numériquement que la condition (3.40) est suffisante mais pas nécessaire pour que le champ de vitesse soit bien défini. Une étude numérique montre que le champ de vitesse est bien défini si et seulement si

$$\sqrt{\frac{3a_i}{4h_0^3}}(h_0 + a_i) < \frac{\pi}{4}, \quad \forall 1 \leq i \leq N, \tag{3.41}$$

soit l'hypothèse physique pour l'existence des solitons. Finalement, une étude numérique montre que le champ de vitesse d'ordre supérieur, ainsi que le champ de vitesse d'ordre un, ne satisfont pas la condition cinématique, c'est-à-dire que les particules initialement situées sous la surface de l'eau restent sous la surface de l'eau. Si cela peut paraître surprenant, cela s'explique par un certain découplage entre la solution de l'équation de Korteweg-de Vries et le champ de vitesse : ce dernier est en fonction de ce premier mais pas vice versa comme dans les équations de Saint-Venant par exemple. Cela s'explique également par la formulation non linéaire de la condition de cinématique, ce qui implique que celle-ci n'est pas assurée par tous les termes du développement en série. Les résultats numériques sont donc qualitatifs. Des résultats quantitatifs peuvent être retrouvés en considérant un paramètre d'amplitude ε petit ou en reparamétrisant les résultats obtenus.

3.5 Perspectives

La suite logique des travaux présentés précédemment est l'étude de la contrôlabilité lagrangienne pour le champ de vitesse d'ordre supérieur associé à la solution de N-solitons. En considérant la restriction sur la hauteur des solitons (3.41), la contrôlabilité lagrangienne, pour la profondeur h_0 fixée, devrait être obtenue pour un domaine de longueur arbitraire $L > 0$ en temps long. De plus, en utilisant des estimations plus fines sur le déplacement des particules soumises à un soliton, la contrôlabilité lagrangienne en temps petit et pour un domaine de longueur arbitraire $L > 0$ devrait être obtenue en choisissant la profondeur h_0 suffisamment grande. Un résultat préliminaire a déjà été publié dans l'acte de conférence [62].

Plusieurs autres équations admettent également des solitons ou des approximations d'ondes solitaires. Parmi celles-ci, on compte certaines équations non linéaires de Schrödinger, l'équation de Sin-Gordon, l'équation de Korteweg-de Vries généralisée avec une non-linéarité $((u^p)_x, p \leq 3$ et l'équation de Boussinesq. Il serait donc intéressant d'étudier la contrôlabilité lagrangienne de ces équations grâce aux solitons ou aux ondes solitaires.

En général, beaucoup reste à faire en contrôlabilité lagrangienne. En dehors du problème de la caractérisation de la contrôlabilité lagrangienne de plusieurs équations, on peut se demander, notamment dans le cas de l'équation de Korteweg-de Vries, s'il est possible de diminuer le nombre de contrôles afin d'obtenir la contrôlabilité lagrangienne. De nouvelles approches afin de résoudre ce problème seront nécessaires. Le problème de stabilisation lagrangienne n'a également pas encore été traité. Étant donné une région du fluide, est-il possible de construire des lois de feedbacks de sorte que cette région reste approximativement au même endroit suite à des perturbations d'un état initial ?

4 Stabilisation rapide d'une équation de Schrödinger linéaire

Avant de présenter la stabilisation rapide de l'équation de Schrödinger, introduisons quelques notions de mécanique quantique. Cette présentation très sommaire a pour but unique de donner l'interprétation physique des solutions de l'équation de Schrödinger. Nous référons notamment à [113] pour une présentation plus approfondie de la mécanique quantique.

Le résultat de stabilisation rapide du linéarisé autour de la première fonction propre de l'équation de Schrödinger bilinéaire est ensuite présenté. Une équation de Schrödinger stable est considérée comme l'image d'une transformation ayant comme domaine de définition l'équation de Schrödinger linéaire à stabiliser. La stabilisation rapide du linéarisé s'obtient en montrant l'inversibilité de la transformation grâce à la contrôlabilité du linéarisé.

Plus précisément, la transformation considérée est un opérateur de noyau, permettant d'obtenir un système d'équations que le noyau doit satisfaire afin de lier l'équation de Schrödinger à stabiliser à l'équation de Schrödinger stable. Une condition d'unicité est introduite afin d'exclure les termes non-locaux apparaissant dans le système. Ces termes non-locaux proviennent du fait que le contrôle considéré est distribué. L'existence de la transformation, son caractère bien posée et son inversibilité s'obtient essentiellement de l'étude des coefficients de Fourier du noyau de l'opérateur.

L'énoncé du résultat se trouve dans la section 4.2. Une couverture de la méthode du backstepping en dimension finie et infinie ainsi que les résultats présents dans la littérature est effectuée dans la section 4.3. Une explication plus en profondeur de la preuve du résultat se trouve dans la section 4.4. Les perspectives quant à l'application du résultat pour l'équation linéarisé à l'équation bilinéaire ainsi que la généralisation dans un cadre abstrait de la méthode utilisée sont présentées dans la section 4.5.

4.1 L'équation de Schrödinger

En mécanique classique, la deuxième loi de Newton $F = ma$ permet de décrire la trajectoire $x(t)$ d'un objet à partir d'une position initiale. De cette trajectoire peut ensuite être obtenu d'autres propriétés de cet objet telles que sa vitesse ou son accélération.

En mécanique quantique, l'évolution des propriétés d'une particule non-relativiste dans $\mathbb{R}^n$ est encodée dans la fonction d'onde $\psi(t,.) \in L^2(\mathbb{R}^n;\mathbb{C})$, de norme 1, solution

de l'équation de Schrödinger

$$\begin{cases} i\hbar\psi_t = H(t)\psi, & (x,t) \in \mathbb{R}^n \times [0,T], \\ \psi(0,x) = \psi_0(x), & x \in \mathbb{R}^n, \end{cases}$$

où ψ_0 est l'état initial de la particule, $H(t)$ est le hamiltonien, dépendant du temps, du système, et $\hbar$ la constante de Planck réduite. Contrairement à la mécanique classique, l'état de la particule ne peut être déterminé précisément par la fonction d'onde, seulement une probabilité de trouver la particule dans un certain état peut être obtenue de la fonction d'onde. La probabilité que la particule soit située dans une région ω à un moment t s'obtient en calculant

$$\int_\omega |\psi(t,x)|^2 \, \mathrm{d}x.$$

Les autres moments de la particules correspondent à

$$\int_\omega p(x) \, |\psi(t,x)|^2 \, \mathrm{d}x,$$

où $p(x)$ est le polynôme correspondant au moment désiré. Lorsque la particule est observée physiquement, la position se retrouve dans un certain état et il n'existe plus d'incertitude sur son état. On dit alors que la fonction d'onde *collapse* dans l'état observé. Si le système n'a pas changé par la suite, de nouvelles observations de la particule redonneront le même état.

L'équation de Schrödinger étant une équation linéaire par rapport à la fonction d'onde, l'étude des fonctions propres du hamiltonien est particulièrement intéressante. Si l'hamiltonien admet un spectre discret seulement, les états de la particule correspondent à une superposition des états propres. La valeur propre associée à une fonction propre correspond à l'énergie de la particule lorsque trouvée dans cet état.

4.2 Contributions à la stabilisation du linéarisé de l'équation de Schrödinger bilinéaire

4.2.1 Problématique

Considérons, pour $T > 0$, l'équation de Schrödinger bilinéaire,

$$\begin{cases} i\partial_t\psi = -\Delta\psi - u(t)\mu(x)\psi, & (x,t) \in (0,1) \times (0,T), \\ \psi(t,0) = \psi(t,1) = 0, & t \in (0,T), \\ \psi(0,x) = \psi_0(x), & x \in (0,1). \end{cases} \tag{4.1}$$

Dans (4.1), ψ est la fonction d'onde d'une particule confinée dans un puits infini de potentiel unidimensionnel. La valeur du puits en dehors de l'intervalle $[0,1]$ vaut l'infini, empêchant virtuellement la particule de se situer en dehors de ce segment et permettant donc l'imposition des conditions de Dirichlet aux bords. La particule est soumise à un champ électrique à l'intérieur du domaine, où $u \in L^2((0,T);\mathbb{R})$ est l'amplitude du champ électrique et $\mu \in H^3((0,1);\mathbb{R})$ est le moment dipolaire de la particule.

Les valeurs propres et fonctions propres du laplacien de Dirichlet sur l'intervalle $[0,1]$ sont

$$\lambda_k := (k\pi)^2, \quad \varphi_k(x) := \sqrt{2}\sin(k\pi x), \quad k \in \mathbb{N}^*.$$

Les états propres de (4.1) sont donnés par $\Phi_k(x,t) := e^{-i\lambda_k t}\varphi_k(x)$ et nous nous intéressons à la stabilisation de la trajectoire de (4.1) $(\psi, u) = (\Phi_k, 0)$. En définissant $\mathcal{A} := \Delta$, $D(\mathcal{A}) = H^2 \cap H_0^1$, on introduit l'espace $H^s_{(0)} := D(\mathcal{A}^{s/2})$ équipé du produit scalaire,

$$\langle \phi, \psi \rangle_{H^s_{(0)}} := \sum_{k=1}^{+\infty} \lambda_k^s \langle \phi, \varphi_k \rangle \overline{\langle \psi, \varphi_k \rangle},$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire $L^2((0,1);\mathbb{C})$. L'espace $H^s_{(0)}$ est munit de la norme $\|\cdot\|_{H^s_{(0)}}$ associée au produit scalaire $H^s_{(0)}$. Remarquons que pour les valeurs entières de s, ces espaces peuvent être explicitement représentés. En particulier, $H^2_{(0)} = H^2 \cap H_0^1((0,1);\mathbb{C})$ et

$$H^3_{(0)} = \left\{ \phi \in H^3 \cap H_0^1((0,1);\mathbb{C}) \; ; \; \phi''(0) = \phi''(1) = 0 \right\}.$$

Définissons la stabilisation exponentielle rapide de (4.1) localement autour de $(\Phi_k, 0)$.

Definition 4.1 *On dit que* (4.1) *est rapidement exponentiellement stabilisable localement autour de* $(\Phi_k, 0)$ *si, pour tout* $\lambda > 0$*, il existe un feedback* $u(t) = K(\psi(t,\cdot))$*,* $C > 0$ *et* $\varepsilon > 0$ *tel que si* $\|\psi_0 - \Phi_k(\cdot, 0)\|_{H^3_{(0)}} < \varepsilon$*, alors la solution de* (4.1) *issue de cette donnée initiale satisfait*

$$\|\psi - \Phi_k\|_{H^3_{(0)}} \leq Ce^{-\lambda t}\|\psi_0\|_{H^3_{(0)}}, \quad \forall t \geq 0.$$

Le linéarisé de (4.1) autour des états propres est donné par

$$\begin{cases} i\partial_t \psi = -\Delta\psi - u(t)\mu(x)\Phi_k, & (x,t) \in (0,1) \times (0,T), \\ \psi(t,0) = \psi(t,1) = 0, & t \in (0,T), \\ \psi(0,x) = \psi_0(x), & x \in (0,1). \end{cases} \tag{4.2}$$

L'hypothèse suivante sur le champ électrique μ sera effectuée dans le chapitre 3.

Hypothèse 4.1 ([7]) *La fonction μ appartient à $H^3((0,1);\mathbb{R})$ et il existe $c>0$ telle que*

$$\frac{c}{k^3} \leq |\langle \mu\varphi_1, \varphi_k\rangle|, \quad \forall k \in \mathbb{N}^*.$$

Avec Jean-Michel Coron et Morgan Morancey, notre résultat sur la stabilisation exponentielle rapide de (4.2) est le suivant.

Théorème 4.1 ([48]) *Supposons que μ satisfait l'hypothèse 4.1. Alors, pour tout $\lambda > 0$, il existe $C > 0$ et une loi de feedback $u(t) = K(\psi(t,\cdot))$ telle que pour tout $\Psi_0 \in \left\{\phi \in H^3_{(0)}\ ;\ \Re\langle\phi,\varphi_1\rangle = 0\right\}$ la solution associée de (4.2) autour du premier état propre ($k=1$) satisfait*

$$\|\Psi(t,\cdot)\|_{H^3_{(0)}} \leq Ce^{-\lambda t}\|\Psi_0\|_{H^3_{(0)}}.$$

Ce théorème sera montré dans le chapitre 3 en utilisant des arguments inspirés du backstepping en dimension infinie.

4.3 Backstepping

Avant de couvrir les résultats présents dans la littérature, présentons brièvement la méthode du backstepping en dimension finie ainsi que ses connexions avec la stabilisation d'équations en dimension infinie. Nous référons à [88] pour une présentation plus détaillée de la méthode du backstepping en dimension finie.

Considérons d'abord le système

$$\begin{cases} x'(t) = f(x(t)) + g(x(t))\xi(t), & t \in [0,\infty) \\ \xi'(t) = u(t), & t \in [0,\infty) \end{cases} \tag{4.3}$$

où $f(0) = 0$, $u \in \mathbb{R}$ et où l'état du système est $(x,\xi) \in \mathbb{R}^{n+1}$ avec $x \in \mathbb{R}^n$, $\xi \in \mathbb{R}$. Le point $(x,\xi,u) = (0,0,0)$ est un point d'équilibre du système. Le backstepping consiste à obtenir, grâce à la structure particulière de (4.3), la stabilité globale asymptotique de (4.3) à partir de la stabilité globale asymptotique du sous-système

$$x'(t) = f(x) + g(x)u(t). \tag{4.4}$$

Supposons que (4.4) satisfait l'hypothèse suivante :

Hypothèse 4.2 *Il existe une loi de feedback $u(t) = \alpha(x(t))$ continûment différentiable de (4.4) telle que le point d'équilibre $(x,u) = (0,0)$ soit globalement asymptotiquement stable ainsi qu'une fonction de Lyapunov $V(x)$ pour le point $(x,u) = (0,0)$ telle que*

$$\frac{\partial V}{\partial x}[f(x) + g(x)u(t)] \leq -W(x) \leq 0, \quad x \in \mathbb{R}^n,$$

où W est semi-définie positive.

La stabilité globale asymptotique de (4.4) entraîne alors la stabilité globale asymptotique de (4.3).

Lemme 4.3 ([88]) *Supposons que le système* (4.4) *satisfasse l'hypothèse 4.2 et considérons le système augmenté* (4.3)*. Alors,*

1. *Si W est définie positive, alors*

$$V_\alpha(x,\xi) = V(x) + \frac{1}{2}[\xi - \alpha(x)]^2,$$

est une fonction de Lyapunov pour (4.3) *pour le feedback donné par*

$$u = -c(\xi - \alpha(x)) + \frac{\partial \alpha}{\partial x}(x)[f(x) + g(x)\xi] - \frac{\partial V}{\partial x}g(x), \quad c > 0. \tag{4.5}$$

2. *Si W est semi-définie positive, alors* (4.5) *est un feedback de telle sorte que $V_\alpha(x,\xi) \leq -W_\alpha(x,\xi) \leq 0$ tel que $W_\alpha(x,\xi) > 0$ dès que $W(x) > 0$ ou $\xi \neq \alpha$. Cela garantit la majoration globale et la convergence de $(x(t),\xi(t))^t$ vers le plus grand ensemble invariant M_α contenu dans l'ensemble*

$$I_\alpha = \left\{(x,\xi) \in \mathbb{R}^{n+1} \,|\, W(x) = 0, \xi = \alpha(x)\right\}.$$

Esquissons la preuve de ce résultat. Introduisons la variable d'erreur mesurant la différence entre de la variable ξ et le feedback permettant de stabiliser le sous-système (4.4)

$$z = \xi - \alpha(x).$$

On peut alors réécrire (4.3) en fonction de (x,z)

$$\begin{aligned} x'(t) =& f(x(t)) + g(x(t))[z(t) + \alpha(x(t))], \\ z'(t) =& u(t) - \frac{\partial \alpha}{\partial x}(x(t))[f(x(t)) + g(x(t))[z(t) + \alpha(x(t))]]. \end{aligned}$$

La dérivée de la fonction de Lyapunov V_α s'écrit,

$$\dot{V_\alpha} \leq -W(x) + z[u - \frac{\partial \alpha}{\partial x}(x)(f + g(\alpha + x)) + \frac{\partial V}{\partial x}g]. \tag{4.6}$$

Les deux parties du Lemme 4.3 sont obtenues par le Théorème de Lyapunov et le Théorème de Lasalle respectivement en prenant des feedbacks u de telle sorte que le membre de droite de (4.6) soit défini positif et semi-défini positif respectivement.

Notons que cette méthode peut être utilisée itérativement afin de traiter des

systèmes de la forme

$$\begin{cases} x'(t) = f(x(t)) + g(x(t))\xi_1(t), \\ \xi_1'(t) = \xi_2(t), \\ \qquad \vdots \\ \xi_{k-1}'(t) = \xi_k(t), \\ \xi_k'(t) = u(t). \end{cases} \tag{4.7}$$

où l'état est $(x, \xi_1, \ldots, \xi_k) \in \mathbb{R}^{n+k}$ avec $x \in \mathbb{R}^n$ et $(\xi_1, \ldots, \xi_k) \in \mathbb{R}^n$.

Corollaire 4.4 ([88]) *Supposons que le système* (4.4) *satisfasse l'hypothèse 4.2 avec une loi de feedback α de classe C^k et considérons le système augmenté* (4.7).

Posons, avec $\alpha_0(x) := \alpha(x)$

$$V_a(x, \xi_1, \ldots, \xi_k) = V(x) + \frac{1}{2}\sum_{i=1}^{k}[\xi_k - \alpha_{i-1}(x, \xi_1, \ldots, \xi_{i-1})]^2,$$

où, pour $1 \leq i \leq k-1$, α_i est le feedback du sous-système d'état $(x, \xi_1, \ldots, \xi_i)$ obtenu du Lemme 4.3 en considérant la variable ξ_{i+1} comme un contrôle fictif. Tout choix de feedback impliquant $V_a' \leq -W_a(x, \xi_1, \ldots, \xi_{k-1}) \leq 0$ avec

$$W_a(x, \xi_1, \ldots, \xi_{k-1}) = 0,$$

seulement si $W(x) = 0$ et $\xi_i \neq \alpha_{i-1}(x, \xi_1, \ldots, \xi_{i-1}), i = 1, ..., k$ garantit la majoration globale et la convergence de $(x, \xi_1, \ldots, \xi_k)$ vers le plus grand ensemble invariant M_α contenu dans l'ensemble

$$E_\alpha = \left\{(x, \xi_1, \ldots, \xi_k) \in \mathbb{R}^{n+k} \,|\, W(x) = 0, \xi_i = \alpha(x, \xi_1, \ldots, \xi_{i-1}), i = 1, \ldots, k\right\}.$$

De plus, si W est définie positive, alors V_α est une fonction de Lyapunov pour (4.7) *et le feedback considéré.*

Le backstepping a ensuite trouvé ses applications en dimension infinie. En dehors de [46, 98], le backstepping en dimension infinie fait référence à des constructions impliquant une transformation de Volterra du deuxième type transformant le système à stabiliser en le système stable. Le premier de ces exemples est trouvé dans [4, 20] où les auteurs ont considéré le problème de stabilisation pour une classe d'EDP parabolique

$$u_t(t, x) = \varepsilon u_{xx}(t, x) + Bu_x(t, x) + \lambda(x)u(t, x), \quad x \in (0, 1), t \in \mathbb{R}^+,$$

avec des conditions de Dirichlet ou de Neumann aux bords, homogène en $x = 0$ et avec un contrôle en $x = 1$. Afin de trouver un feedback stabilisant cette classe

d'équation, les auteurs proposent d'étudier la semi-discrétisation par différence finie de l'équation

$$\begin{aligned} u_0 &= 0, \\ u_i' &= \varepsilon \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + B\frac{u_{i+1} - u_i}{h} + \lambda_i u_i, \quad i = 1, \dots, n, \\ u_{n+1} &= \alpha_n(u_1, \dots, u_n), \end{aligned}$$

où u_i sont les valeurs de la solution aux noeuds, $\lambda_i = \lambda(ih)$ et α_n est le feedback discrétisé. Appliquer directement la méthode du backstepping sur le sysème semi-discrétisé changerait considérablement les valeurs propres du système et, par conséquent, la nature du système ([20]). Afin de préserver la nature parabolique du système, les auteurs proposent plutôt de considérer un changement de variables

$$\begin{aligned} w_0 &= u_0, \\ w_i &= u_i - \alpha_{i-1}(u_1, \dots, u_{i-1}), \quad i = 1, \dots, n, \\ w_{n+1} &= 0, \end{aligned} \tag{4.8}$$

amenant la solution à stabiliser vers la solution d'un système stable

$$\begin{aligned} w_0 &= 0, \\ w_i' &= \varepsilon \frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + B\frac{w_{i+1} - w_i}{h} + cw_i, \quad i = 1, \dots, n, \\ w_{n+1} &= 0. \end{aligned}$$

Les solutions de ce dernier système correspondent à la semi-discrétisation de l'équation

$$w_t(t,x) = \varepsilon w_{xx}(t,x) + Bw_x(t,x) + cw(t,x), \quad x \in (0,1), t \in \mathbb{R}^+,$$

avec des conditions de Dirichlet homogène aux deux extrémités, qui est stable exponentiellement rapidement lorsque $c > -\varepsilon\pi^2 - B^2/(4\varepsilon)$. Dans [20], on fait référence à la transformation (4.8) comme une "*backstepping-style coordinate transformation*" en raison de la forme triangulaire du système ainsi que la définition récursive de α_{i-1}. Contrairement à la méthode de backstepping en dimension finie, il ne s'agit plus de chercher de nouveaux feedbacks à partir d'un feedback et d'une fonction de Lyapunov connus mais plutôt de trouver un feedback à partir d'une transformation amenant à un système stable. Toujours dans [20], il est montré que la transformation (4.8) est la version discrétisée d'une transformation de Volterra du deuxième type

$$w(x,t) = u(x,t) - \int_0^x k(x,y)u(y,t)\mathrm{dy},$$

où l'expression du feedback est obtenue par une simple évaluation en $x = 1$

$$\alpha(t) = \int_0^1 k(1,y)u(y,t)\mathrm{dy}.$$

4.3.1 Résultats précédents

Suite aux travaux [4, 20], la stabilisation de plusieurs équations aux dérivées partielles a été obtenue en considérant une transformation de Volterra du deuxième type associant une équation à stabiliser à une équation stable. Nous référons à [89] pour une présentation générale de ces résultats. Cette technique a également été utilisée afin de stabiliser des équations non linéaires [29, 52].

Plus récemment, la transformation de Volterra du deuxième type a été remplacée par la transformation de Fredholm afin d'obtenir la stabilisation exponentielle rapide locale de l'équation de Korteweg-de Vries [50], de l'équation de Kuramoto-Shivashinksi [51] et d'un système hyperbolique [49]. La méthode pour montrer la stabilisation par une transformation de Volterra du deuxième type ou par une transformation de Fredholm repose sur un argument similaire. Une équation sur le noyau définissant la transformation est obtenue de l'équation à stabiliser et de l'équation cible. Une fois l'expression du noyau obtenue de cette équation, la stabilisation de l'équation est obtenue par l'inversibilité de la transformation. L'existence et l'unicité du noyau de la transformation de Fredholm a été démontré grâce à l'hypothèse de la contrôlabilité du linéarisé dans [50, 51, 49]

La contrôlabilité locale de (4.1) pour des données initiales dans $H^3_{(0)}$ proche du premier état propre a été démontrée par Beauchard et Laurent dans [7] (voir aussi [5, 6] pour des résultats précédents) en montrant la contrôlabilité de l'équation linéarisée (4.2) avec l'hypothèse 4.1. Ce cadre fonctionnel est nécessaire puisque, du résultat abstrait de Ball, Marsden et Slemrod ([3]), il a été démontré par Turinici ([127]) que le complément de l'espace des états atteignables pour des conditions initiales dans $\mathcal{S} \cap H^2_{(0)}$, pour des contrôles L^2, est dense dans $\mathcal{S} \cap H^2_{(0)}$, où $\mathcal{S}$ est la sphère de rayon 1 de $L^2((0,1);\mathbb{C})$.

Ce résultat de contrôlabilité locale a été étendu par la suite dans [10] avec des hypothèses plus faibles mais avec un temps de contrôle minimal, dans un cadre plus général avec un temps infini de contrôlabilité ([110]) et dans le cas de la contrôlabilité globale simultanée d'un nombre fini de particules [103, 104]. Plus récemment, la contrôlabilité locale d'un modèle de Schrödinger-Poisson a été obtenu dans [8] et pour l'équation (3.1.1) posé dans un espace de dimension plus petit ou égal à 3 avec un contrôle dépendant également de l'espace ([115]).

4.4 Méthode utilisée

Présentons l'idée de la preuve en dimension finie, inspirée du changement de variables (4.8), telle que présentée dans [44].

4.4.1 Illustration en dimension finie

Considérons le problème de la stabilité exponentielle rapide du point d'équilibre $(0,0)$ du système linéaire

$$\begin{cases} x'(t) = Ax(t) + Bu(t), & t \in [0,T], \\ x(0) = x_0, \end{cases} \tag{4.9}$$

où $A \in M_{n\times n}$ et $B \in M_{n\times 1}$. Faisons l'hypothèse que (4.9) est contrôlable. Introduisons le système cible

$$\begin{cases} z'(t) = (A - \lambda I)z(t), & t \in [0,T], \\ z(0) = z_0, \end{cases} \tag{4.10}$$

qui est stable exponentiellement rapidement pour tout $\lambda > \max(\Re(\lambda_i))$, où λ_i sont les valeurs propres de A. Considérons l'application linéaire

$$z = Tx,$$

qui, à une solution du système à stabiliser, associe la solution du système cible. Considérons également une transformation exprimant le feedback en fonction de la solution de (4.9)

$$u = Kx + v.$$

En utilisant (4.9) et (4.10), les transformations T et K doivent satisfaire

$$\begin{aligned} Tx'(t) =& TAx(t) + TBu(t) \\ (A - \lambda I)z(t) =& TAx(t) + TB(Kx(t) + v(t)) \\ (AT - \lambda T)x(t) =& TAx(t) + TBKx(t) + TBv(t). \end{aligned}$$

On impose à cette dernière équation $TB = B$. Une interprétation possible de cette condition est qu'elle permet de retrouver l'opérateur de contrôle agissant sur v. En considérant $v \equiv 0$, on obtient

$$AT - \lambda T = TA + BK, \tag{4.11}$$

$$TB = B. \tag{4.12}$$

L'existence et l'unicité des transformations T et K est montrée grâce à la contrôlabilité de (4.9).

Théorème 4.2 ([44]) *Soit $A \in M_{n\times n}$ et $B \in M_{n\times 1}$ telles que* (4.9) *est contrôlable. Alors il existe une unique transformation $(T, K) \in GL_n(\mathbb{R})\times\mathbb{R}^{1\times n}$ satisfaisant* (4.11)-(4.12).

L'inversibilité de T implique la stabilisation exponentielle rapide de (4.9). En effet, en prenant $z_0 = Tx_0$, on a

$$\begin{aligned}\|x(t)\| &= \|T^{-1}z(t)\| \\ &\leq \|T^{-1}\|\|z(t)\| \\ &\leq C\|T^{-1}\|e^{-\lambda t}\|z_0\| \\ &= C\|T^{-1}\|e^{-\lambda t}\|Tx_0\| \\ &\leq C\|T^{-1}\|\|T\|e^{-\lambda t}\|x_0\|.\end{aligned}$$

Notons également que si B n'est pas qu'un vecteur colonne, l'existence de T est toujours assurée mais pas l'unicité.

Théorème 4.3 ([44]) *Soit $A \in M_{n\times n}$ et $B \in M_{n\times m}$, $m \leq n$ telles que (A, B) est contrôlable. Alors il existe une transformation $(T, K) \in GL_n(\mathbb{R}) \times \mathbb{R}^{1\times n}$ satisfaisant* (4.11)-(4.12).

4.4.2 Équation de Schrödinger linéarisée

Le Théorème 4.1 sera démontré pour l'équation linéaire

$$\begin{cases} i\partial_t\Psi = -\Delta\Psi - v(t)\mu(x)\varphi_1(x), & (x,t) \in (0,1)\times(0,T), \\ \Psi(t,0) = \Psi(t,1) = 0, & t \in (0,T), \\ \Psi(0,x) = \Psi_0(x), & x \in (0,1). \end{cases} \tag{4.13}$$

La différence entre (4.2) et (4.13) est un changement de phase permettant de considérer un état stationnaire comme point d'équilibre du système. Les détails pour passer de la preuve de la stabilisation exponentielle rapide de (4.2) à la stabilisation exponentielle rapide de (4.13) sont esquissés dans appendice A.3.

La stabilisation exponentielle rapide de (4.13) suit les grandes lignes de la preuve en dimension finie. Puisque le feedback doit être à valeur réelle, on définit la transformation T et K pour la partie réelle et imaginaire de la solution $\Psi = \Psi^1 + i\Psi^2$

$$\begin{aligned} T : H^3_{(0)} \times H^3_{(0)} &\longrightarrow H^3_{(0)} \times H^3_{(0)} \\ \begin{pmatrix}\Psi^1(t,\cdot)\\ \Psi^2(t,\cdot)\end{pmatrix} &\longmapsto \int_0^1 \begin{pmatrix} k_{11}(x,y) & k_{12}(x,y) \\ k_{21}(x,y) & k_{22}(x,y)\end{pmatrix}\begin{pmatrix}\Psi^1(t,y)\\ \Psi^2(t,y)\end{pmatrix} dy, \\ K : H^3_{(0)} \times H^3_{(0)} &\longrightarrow \mathbb{R} \end{aligned} \tag{4.14}$$

$$\begin{pmatrix}\Psi^1(t,\cdot)\\ \Psi^2(t,\cdot)\end{pmatrix} \longmapsto \int_0^1 \alpha^1(y)\Psi^1(t,y)+\alpha^2(y)\Psi^2(t,y)dy. \tag{4.15}$$

Ces transformations sont définies telles que $\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix} := T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}$ est la solution stable exponentiellement rapidement dans $H^3_{(0)}$ de

$$\begin{cases} \partial_t\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix} = \begin{pmatrix}0 & -\Delta\\ \Delta & 0\end{pmatrix}\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix} - \lambda\begin{pmatrix}\xi^1\\ \xi^2\end{pmatrix}, & (t,x)\in(0,1)\times(0,T),\\ \xi^1(t,0)=\xi^1(t,1)=0, \quad \xi^2(t,0)=\xi^2(t,1)=0, & t\in(0,T),\\ \xi^1(0,x)=\xi^1_0(x), \quad \xi^2(0,x)=\xi^2_0(x), & x\in(0,1), \end{cases} \tag{4.16}$$

avec $\begin{pmatrix}\xi^1_0\\ \xi^2_0\end{pmatrix} = T\begin{pmatrix}\Psi^1_0\\ \Psi^2_0\end{pmatrix}$.

À l'aide de (4.13) et de (4.16), on obtient formellement l'équation non locale sur les noyaux

$$\begin{cases} (\Delta_y k_{11} - \Delta_x k_{22} - \lambda k_{12})(x,y) - \alpha^2(y)\left(\int_0^1 (\mu\varphi_1)(z)k_{12}(x,z)dz\right) = 0,\ (x,y)\in(0,1)^2,\\ (\Delta_y k_{12} + \Delta_x k_{21} + \lambda k_{11})(x,y) + \alpha^1(y)\left(\int_0^1 (\mu\varphi_1)(z)k_{12}(x,z)dz\right) = 0,\ (x,y)\in(0,1)^2,\\ (\Delta_y k_{21} + \Delta_x k_{12} - \lambda k_{22})(x,y) - \alpha^2(y)\left(\int_0^1 (\mu\varphi_1)(z)k_{22}(x,z)dz\right) = 0,\ (x,y)\in(0,1)^2,\\ (\Delta_y k_{22} - \Delta_x k_{11} + \lambda k_{21})(x,y) + \alpha^1(y)\left(\int_0^1 (\mu\varphi_1)(z)k_{22}(x,z)dz\right) = 0,\ (x,y)\in(0,1)^2,\\ k_{ij}(x,0)=k_{ij}(x,1)=0, \quad x\in(0,1),\\ k_{ij}(0,y)=k_{ij}(1,y)=0, \quad y\in(0,1), \end{cases}$$

où les conditions aux bords ont été imposées afin de satisfaire les conditions aux bords de (4.13) et de (4.16) ainsi que pour annuler les termes de bords apparaissant dans les intégrations par parties. En raison des termes non locaux, l'existence et l'unicité d'une solution de ce système est difficile à obtenir. Afin de simplifier ce système, nous utilisons, comme en dimension finie, la condition $TB = B$

$$T\begin{pmatrix}0\\ (\mu\varphi_1)(x)\end{pmatrix} = \int_0^1 \begin{pmatrix}k_{12}(x,y)(\mu\varphi_1)(y)\\ k_{22}(x,y)(\mu\varphi_1)(y)\end{pmatrix} dy = \begin{pmatrix}0\\ (\mu\varphi_1)(x)\end{pmatrix}.$$

En utilisant cette nouvelle identité, nous cherchons des noyaux satisfaisant

$$\begin{cases} (\Delta_y k_{11} - \Delta_x k_{22} - \lambda k_{12})(x,y) = 0, \\ (\Delta_y k_{12} + \Delta_x k_{21} + \lambda k_{11})(x,y) = 0, \\ (\Delta_y k_{21} + \Delta_x k_{12} - \lambda k_{22})(x,y) - \alpha^2(y)(\mu\varphi_1)(x) = 0, \\ (\Delta_y k_{22} - \Delta_x k_{11} + \lambda k_{21})(x,y) + \alpha^1(y)(\mu\varphi_1)(x) = 0, \\ k_{ij}(x,0) = k_{ij}(x,1) = 0, \\ k_{ij}(0,y) = k_{ij}(1,y) = 0, \end{cases} \tag{4.17}$$

ainsi que la condition $TB = B$

$$\begin{cases} \displaystyle\int_0^1 k_{12}(x,y)(\mu\varphi_1)(y)dy = 0, \\ \displaystyle\int_0^1 k_{22}(x,y)(\mu\varphi_2)(y)dy = \mu\varphi_1. \end{cases} \tag{4.18}$$

Dans les travaux précédents sur la transformation de Fredholm [50, 51, 49], la condition $TB = B$ était implicitement imposée, puisqu'il s'agissait de termes de bords apparaissant dans les intégrations par parties effectuées afin d'obtenir l'équation des noyaux. L'existence et l'unicité des noyaux a été démontrée grâce à la condition $TB = B$ ainsi qu'avec l'hypothèse de contrôlabilité de l'équation. Le résultat présenté ici est le premier cas en dimension infinie où la condition $TB = B$ est imposée explicitement. Nous verrons que la condition $TB = B$, bien qu'aidant à simplifier l'expression de l'équation des noyaux, apporte un lot de complications, l'opérateur linéaire B étant non borné dans l'espace des états.

Afin de construire la transformation T, écrivons les noyaux en séries de Fourier

$$k_{ij}(x,y) = \sum_{n=1}^{+\infty} f_n^{ij}(x)\varphi_n(y), \qquad \alpha^j(y) = \sum_{n=1}^{+\infty} \alpha_n^j \varphi_n(y). \tag{4.19}$$

Cela mène à

$$\begin{cases} f_n^{11\prime\prime}(x) + \lambda_n f_n^{22}(x) - \lambda f_n^{21}(x) = \alpha_n^1(\mu\varphi_1)(x), \\ f_n^{12\prime\prime}(x) - \lambda_n f_n^{21}(x) - \lambda f_n^{22}(x) = \alpha_n^2(\mu\varphi_1)(x), \\ f_n^{21\prime\prime}(x) - \lambda_n f_n^{12}(x) + \lambda f_n^{11}(x) = 0, \\ f_n^{22\prime\prime}(x) + \lambda_n f_n^{11}(x) + \lambda f_n^{12}(x) = 0, \\ f_n^{ij}(0) = f_n^{ij}(1) = 0. \end{cases} \tag{4.20}$$

et la condition $TB = B$ (4.18) devient

$$\begin{cases} \displaystyle\sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle f_n^{12}(x) = 0, \\ \displaystyle\sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle f_n^{22}(x) = (\mu\varphi_1)(x). \end{cases} \tag{4.21}$$

Le système (4.20) est résolu en considérant séparément $\alpha_n^1(\mu\varphi_1)$ et $\alpha_n^2(\mu\varphi_1)$: on note $(g_n^{11}, g_n^{12}, g_n^{21}, g_n^{22})^t$ la solution avec comme membre de droite de (4.20) $(\beta_n^1(\mu\varphi_1), 0, 0, 0)^t$ et $(h_n^{11}, h_n^{12}, h_n^{21}, h_n^{22})^t$ la solution avec comme membre de droite de (4.20) $(0, \beta_n^2(\mu\varphi_1), 0, 0)^t$. Les coefficients β_n^i serviront à renormaliser la base que nous construirons. Les noyaux de T (4.20) s'écrivent donc

$$k_{ij}(x, y) = \sum_{n=1}^{+\infty} \left(\frac{\alpha_n^1}{\beta_n^1} g_n^{ij}(x) + \frac{\alpha_n^2}{\beta_n^2} h_n^{ij}(x) \right) \varphi_n(y).$$

La résolution de (4.20)-(4.21) passera par la construction d'une base de Riesz dans $H_{(0)}^3$ formée de solutions de (4.20) permettant par la suite de décomposer dans cette base le membre de droite de (4.21) afin de retrouver les coefficients α_n^i. Un critère suffisant pour qu'un ensemble soit une base de Riesz d'un espace est qu'il soit libre et quadratiquement proche de la base orthonormée de l'espace ([132, Théorème 15]). On dit que deux ensembles $\{e_n\}_{n\in\mathbb{N}}$ et $\{g_n\}_{n\in\mathbb{N}}$ sont quadratiquement proches dans un espace de Hilbert H si

$$\sum_{n\in\mathbb{N}} \|e_n - g_n\|_H^2 < \infty.$$

La première étape dans la construction des transformations T et K est de montrer que l'ensemble

$$\mathcal{B} := \left\{ \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix}, \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} \right\}_{\backslash\in\mathbb{N}},$$

est une base de Riesz de $X_{(0)}^3 := H_{(0)}^3 \times H_{(0)}^3$. L'expression des séries de Fourier de chacune des fonctions apparaissant dans l'expression de $\mathcal{B}$ permet de montrer que $\mathcal{B}$ est quadratiquement proche de la base orthonormale de $X_{(0)}^3$

$$\left\{ \begin{pmatrix} \varphi_n/\lambda_n^{3/2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi_n/\lambda_n^{3/2} \end{pmatrix} \right\}_{n\in\mathbb{N}},$$

alors que l'hypothèse de contrôlabilité est utilisée afin de montrer que $\mathcal{B}$ est libre.

La base de Riesz $\mathcal{B}$ permet ensuite d'obtenir l'expression des coefficients α_n^i de la

relation $TB = B$

$$\begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix} = \sum_{n=1}^{+\infty} \frac{\alpha_n^1 \langle \mu\varphi_1, \varphi_n \rangle}{\beta_n^1} \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + \frac{\alpha_n^2 \langle \mu\varphi_1, \varphi_n \rangle}{\beta_n^2} \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix}. \tag{4.22}$$

La fonction $\mu\varphi_1$ n'appartenant pas à $H^3_{(0)}$, les résultats sur les bases de Riesz ne peuvent pas être appliqués directement afin d'obtenir l'existence de coefficients α_n^i tels que (4.22) soit satisfait avec la régularité nécessaire afin que les transformations T et K soient définies dans les espaces voulus. En effet, (4.22) n'a de sens que dans $X^s_{(0)}$ pour $s < 5/2$ en raison de l'incompatibilité des traces des dérivées secondes. Il est important de noter que cette incompatibilité des traces des dérivées secondes est nécessaire : si $\mu\varphi_1 \in H^3_{(0)}$, non seulement l'hypothèse 4.1 ne serait pas satisfaite et l'équation (4.2) ne serait pas contrôlable, mais aussi la transformation T obtenue serait compacte, et donc d'inverse non bornée de $X^3_{(0)}$ dans lui-même.

Les coefficients α_n^i de telle sorte que (4.22) soit satisfait dans $(H^3 \cap H_0^1)^2$ sont obtenus grâce à la décomposition de $\mu\varphi_1$ introduite par Puel dans [114] afin d'étendre le résultat d'admissibilité de l'opérateur de contrôle de (4.2) obtenu dans [7] à des domaines spatiaux de dimensions supérieures. La fonction $\mu\varphi_1$ peut être décomposée en la somme de deux fonctions g et h où h est un relèvement de la trace de la dérivée seconde de $\mu\varphi_1$. Ainsi, g appartient à $H^3_{(0)}$ et le bilaplacien de h est nul. De plus,

$$\langle \mu\varphi_1, \varphi_n \rangle = \frac{4}{n^3\pi^2} \left((-1)^{n+1}\mu'(1) - \mu'(0) \right) - \frac{\sqrt{2}}{(n\pi)^3} \int_0^1 (\mu\varphi_1)'''(x) \cos(k\pi x) \mathrm{d}x,$$

où la partie en n^{-3} provient de la fonction h alors que la partie en $n^{-3} \times l^2(\mathbb{R})$ provient de g. Les coefficients α_n^i sont obtenus essentiellement en deux temps. Un choix initial des coefficients α_n^i est fait afin de traiter la partie singulière se comportant en n^{-3}. Il est ensuite montré que le reste est une fonction $H^3_{(0)}$ pour lequel on peut utiliser les résultats sur la base de Riesz afin d'obtenir l'existence de coefficients. La régularité des coefficients α_n^i est suffisante afin de démontrer que la transformation T est bien définie de $X^3_{(0)}$ dans lui-même.

S'en suit alors la preuve de l'inversibilité de T, effectuée en deux étapes. La première consiste à montrer que T s'écrit comme la somme d'un opérateur inversible de $X^3_{(0)}$ vers lui-même et d'un opérateur compact. Cette décomposition de T suit la décomposition des coefficients de Fourier α_n^i. La partie associée à la partie inversible, notée $\tilde{T}$, est celle de α_n^2 ayant un comportement en n^{-3}, alors que la partie compacte correspond à celle associée aux coefficients en $n^{-3} \times l^2(\mathbb{R})$. Cela implique en particulier que l'indice de T, défini par

$$\text{ind } T = \dim(\text{Ker } T) + \text{codim}(\text{Im } T),$$

est zéro. Il est donc suffisant pour montrer l'inversibilité de T que $\text{Ker } T^* = \{0\}$.

La preuve de ce dernier point, comme la preuve de la base de Riesz, provient de la contrôlabilité, ou, plus précisément, de l'observabilité du problème adjoint. Considérons formellement l'équation satisfaite par (T^*, K^*)

$$A^*T^* + K^*B^* - T^*A^* + \lambda T^* = 0, \tag{4.23}$$
$$B^*T^* = B^*. \tag{4.24}$$

Supposons de plus qu'il existe $\phi \in \operatorname{Ker} T^*$ non-nul. De (4.23), on obtient que

$$T^*A^*\phi = 0,$$

ce qui implique que $\operatorname{Ker} T^*$, de dimension finie, est stable par A^*. En itérant l'argument, on déduit que

$$T^*(A^*)^k\phi = 0, \quad k \in \mathbb{N}.$$

Finalement, en utilisant (4.24), on a

$$B^*T^*(A^*)^k\phi = B^*(A^*)^k\phi = 0, \quad k \in \mathbb{N},$$

ce qui implique que $\phi = 0$ par observabilité.

Il s'agit d'une preuve générale. La preuve utilisée dans le chapitre 3 diffère légèrement en utilisant plutôt que la stabilité de $\operatorname{Ker} T^*$ par A^* implique l'existence d'un vecteur propre dans $\operatorname{Ker} T^*$ de A^*. En utilisant l'adjoint de la condition $TB = B$, il est ensuite montré que ce vecteur propre doit forcément être nul, illustrant une fois de plus l'importance de l'hypothèse de la contrôlabilité.

Cette preuve formelle est rendue rigoureuse en donnant un sens à l'égalité entre opérateurs

$$TA + BK = AT - \lambda T. \tag{4.25}$$

Le cadre fonctionnel de cette équation n'est pas trivial puisque B est un opérateur non borné dans $X^3_{(0)}$. Le choix de fonctions dans $X^5_{(0)}$ par exemple permet de définir AT et TA dans $X^3_{(0)}$ mais l'opérateur BK reste toujours à valeur dans $H^3_{(0)} \times (H^3 \cap H^1_0)$, ce qui n'est pas le cadre fonctionnel recherché. L'idée consiste plutôt à consider, à l'aide de l'égalité $TB = B$, $A + BK$ comme un seul opérateur. Il sera montré que l'égalité (4.25) a un sens dans le domaine de $A + BK$

$$\begin{aligned} D(A+BK) &:= \left\{ \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)} \;\middle|\; (A+BK)\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)} \right\} \\ &= \left\{ \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in (H^5 \cap H^3_{(0)}) \times H^5_{(0)} \;\middle|\; \Delta\Psi^1 - K\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \mu\varphi_1 \in H^3_{(0)} \right\}, \end{aligned}$$

et que $D(A + BK)$ est dense dans $X^3_{(0)}$.

Il sera ensuite montré que $A + BK$ est le générateur infinitésimal d'un groupe sur $X^3_{(0)}$. La démonstration de ce dernier point repose sur l'inversibilité de T. Celle-ci

permet de définir une norme équivalente à la norme $X^3_{(0)}$ et de montrer facilement la dissipativité de $A+BK$ à l'aide du système cible. En effet, pour $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in D(A+BK)$, on a

$$\begin{aligned}\langle (A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\rangle_T :=& \langle T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\ =& \langle AT\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\rangle_{X^3_{(0)}} - \lambda \left\|\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|_T^2\\ =& -\lambda \left\|\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|_T^2 \leq 0.\end{aligned}$$

Notons qu'il est difficile de montrer la dissipativité sans passer par le système cible. En effet, un calcul direct montre que

$$\langle (A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\rangle_{X^3_{(0)}} = \langle BK\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\rangle_{X^3_{(0)}},$$

terme pour lequel il est difficile d'établir le signe.

La stabilisation exponentielle rapide de (4.2) suit ensuite facilement en montrant que les solutions $T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}$ sont stables exponentiellement rapidement. L'inversibilité de T permet de conclure donc que les solutions de (4.2) sont stables exponentiellement rapidement.

4.5 Commentaires sur le résultat obtenu et perspectives

Le résultat sur l'équation linéarisée ne s'étend pas directement à l'équation non linéaire en raison de la définition du feedback K. En effet, ce feedback n'est pas bien défini pour des fonctions appartenant seulement à $X^3_{(0)}$ en raison du comportement en n^3 du coefficient de Fourier de α_n^2. La définition du groupe engendré par le générateur infinitésimal $A+BK$ est passée par la définition d'un domaine qui prenait en compte les conditions au bord que devaient satisfaire les fonctions pour que leur image par $A+BK$ soit dans $X^3_{(0)}$. Les solutions de (4.2) avec $u = K\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}$ sont données par

$$\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix} = e^{(A+BK)t}\begin{pmatrix}\Psi^1_0\\ \Psi^2_0\end{pmatrix}, \tag{4.26}$$

empêchant donc à priori de dissocier le feedback d'une solution de (4.2).

À la fin du chapitre 3, nous montrerons en fait que

$$K\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix} \in L^2(0,T), \tag{4.27}$$

où $\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}$ est solution de $A+BK$ et que la solution peut s'écrire

$$\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix} = e^{At}\begin{pmatrix}\Psi^1_0\\ \Psi^2_0\end{pmatrix} - \int_0^T e^{A(T-s)}K\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}K\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{d}s. \tag{4.28}$$

Afin d'obtenir ce résultat, on obtient d'abord, pour des données initiales dans $D(A+BK)$, que les expressions (4.26) et (4.28) sont égales dans $C([0,T];X^3_{(0)})\cap C^1([0,T];X^1_{(0)})$. Une estimation de

$$\int_0^T e^{A(T-s)}K\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}K\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{d}s,$$

par rapport aux données initiales est ensuite obtenue afin de passer à la limite afin de définir cette expression pour des solutions définies à partir de données initiales appartenant à $X^3_{(0)}$. Finalement, un résultat de régularité caché similaire à l'admissibilité de l'opérateur B présentée dans [7] permet d'obtenir (4.27).

Ce résultat ne permet pas de définir directement un feedback pour l'équation non linéaire. Une voie possible afin d'obtenir la stabilisation de l'équation non linéaire est de considérer le système couplé avec l'équation linéaire

$$\begin{cases}\Psi_t = -\Delta\Psi - K(\chi)\mu\Psi, & (x,t)\in(0,1)\times(0,T),\\ \chi_t = (A+BK)\chi - a(\Psi-\chi), & (x,t)\in(0,1)\times(0,T),\\ \Psi(t,0)=\Psi(t,1)=\chi(t,0)=\chi(t,1)=0, & t\in(0,T),\\ \Psi(0,x)=\Psi_0(x), \quad \chi(0,x)=\chi_0(x) & x\in(0,1),\end{cases}$$

où $a>0$. L'idée est de montrer la stabilisation exponentielle rapide de l'équation non linéaire en montrant la stabilisation exponentielle rapide de χ et de l'erreur $\Psi-\chi$ en utilisant l'hypothèse que la donnée initiale Ψ_0 est près du premier état fondamental. Le cadre fonctionnel rend toutefois la démonstration difficile.

Un problème important serait d'étudier la généralisation de la preuve du Théorème 4.2 en dimension infinie sous les hypothèses notamment que l'opérateur A soit diagonalisable et de contrôlabilité de l'équation considérée. Les étapes présentés dans le chapitre 3 afin d'obtenir le Théorème 4.1 semblent être un bon point de départ. En suivant ces étapes, un des points cruciaux afin d'obtenir ce résultat abstrait serait l'existence d'une base de Riesz solution du système des noyaux.

Si un tel résultat est montré, alors on obtiendrait l'équivalent du résultat en dimension finie que la contrôlabilité implique la stabilité pour des équations du type $x'(t)=Ax(t)+Bu(t)$ pour A et B sont des matrices indépendantes du temps.

5 Contrôles numériques pour l'équation des ondes par une méthode spectrale

Dans cette partie un travail en collaboration avec José Urquiza est présenté sur des moyens d'obtenir la minoration uniforme de l'inégalité d'observabilité de l'équation des ondes semi-discrétisée avec une méthode de Legendre-Galerkin.

En général, la convergence des schémas numériques n'est pas suffisante afin d'obtenir la convergence des contrôles numériques pour l'équation des ondes. Il est générale ment accepté que la minoration uniforme de l'écart des racines des valeurs propres ainsi qu'une minoration uniforme de l'inégalité d'observabilité et une majoration uniforme de l'inégalité d'admissibilité permet de démontrer la convergence des contrôles numériques vers le contrôle optimal. Dans un premier temps, nous montrons numériquement qu'une semi-discrétisation avec une méthode spectrale de Legendre-Galerkin de l'équa tion des ondes permet d'obtenir une minoration uniforme de l'écart des racines des valeurs propres. De plus, en filtrant la trace normale au bord grâce à des filtres spectraux, nous obtenons numériquement la minoration uniforme de l'inégalité d'observabilité. Dans un second temps, une semi-discrétisation avec une méthode spectrale de Legendre-Galerkin d'une formulation mixte et permet d'obtenir la minoration uniforme de l'écart des racines des valeurs propres ainsi que de la constante d'observabilité discrète. Finalement, la méthode de Nitsche est utilisée afin d'obtenir la minoration uniforme de l'écart des racines des valeurs propres ainsi que de la constante d'observabilité discrète. Dans ce cas, l'inégalité d'observabilité n'est pas directement la version discrète du cas continu mais celle obtenue lors de la caractérisation de la contrôlabilité du schéma discret. Un terme pénalisé apparait alors naturellement dans la formulation de l'inégalité d'observabilité. Un article est en préparation sur le sujet ([63]).

La problématique est présentée plus précisément dans le cas d'une semi-discrétisa tion par éléments finis $\mathbb{P}_0^1$ dans la section 5.1. Un tour d'horizon sur les méthodes connues dans la littérature afin d'obtenir la convergence des contrôles numériques est présenté dans la section 5.3. Les résultats numériques concernant chacune des méthodes sont présentés, dans l'ordre, dans les sous-section 5.3.1, 5.3.2 et 5.3.3.

5.1 La problématique

Pour $T > 0$, considérons l'équation des ondes unidimensionnelle

$$\begin{cases} y_{tt} - y_{xx} = 0, & (x,t) \in (-1,1) \times (0,T), \\ y(-1,t) = 0, & t \in (0,T), \\ y(1,t) = w(t), & t \in (0,T), \\ y(x,0) = y_0(x), y_t(x,0) = y_1(x), & x \in (-1,1), \end{cases} \tag{5.1}$$

où $(y_0, y_1) \in L^2(-1,1) \times H^{-1}(-1,1)$ et $w \in L^2(0,T)$. Il est bien connu que (5.1) est contrôlable si et seulement si $T \geq 4$, soit le temps que les solutions de (5.1) se propageant à vitesse constante de 1 atteigne la région de contrôlabilité. La contrôlabilité est liée au problème d'observabilité du problème adjoint

$$\begin{cases} u_{tt} - u_{xx} = 0, & (x,t) \in (-1,1) \times (0,T), \\ u(-1,t) = u(1,t) = 0, & t \in (0,T), \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & x \in (-1,1), \end{cases} \tag{5.2}$$

où $(u_0, u_1) \in H_0^1(-1,1) \times L^2(-1,1)$. Rappelons que les solutions de (5.2) préservent l'énergie des données initiales au fil du temps

$$E(u(0)) \equiv E(u(t)) := \frac{1}{2} \int_{-1}^{1} \left[|u_t(x,t)|^2 + |u_x(x,t)|^2 \right] \, \mathrm{d}x, \quad t \in [0,T].$$

La contrôlabilité de (5.1) est équivalente à l'inégalité d'observabilité

$$c_T E(u(0)) \leq \int_0^T |u_x(1,t)|^2 \, \mathrm{d}t, \tag{5.3}$$

qui est satisfaite pour $c_T > 0$ si et seulement si $T \geq 4$. Rappelons qu'en dépit de la régularité des données initiales, la dérivée normale est bien définie pour les solutions de (5.2)

$$\int_0^T |u_x(1,t)|^2 \, \mathrm{d}t \leq C_T E(u(0)), \tag{5.4}$$

où $C_T > 0$ pour tout temps $T > 0$. Il s'agit de l'admissibilité de l'opérateur B^* (2.5) pour (5.2).

Le contrôle minimisant la norme $L^2(0,T)$ est donné par l'expression

$$w(t) = \widetilde{u}_x(1,t),$$

où $\widetilde{u}$ est solution de (5.2) issue des données initiales satisfaisant

$$(u_0, u_1) = \underset{(u_0,u_1) \in H_0^1 \times L^2}{\arg\min} \; J(u_0, u_1),$$

où

$$J(u_0, u_1) = \frac{1}{2} \int_0^T |u_x(1,t)|^2 \, \mathrm{d}t - < (y_1, -y_0), (u_0, u_1) >_{(H^{-1} \times L^2, H_0^1 \times L^2)} . \tag{5.5}$$

L'existence d'un unique minimum de (5.5) est assurée par (5.3) et (5.4).

Une question que l'on peut considérer une fois que la contrôlabilité de l'équation est obtenue est de savoir s'il est possible d'approcher les contrôles numériquement.

Cela a une importance particulière pour des domaines spatiaux de dimensions supé rieures avec des géométries pour lesquelles nous n'avons pas d'expression explicites des fonctions propres. Or, même en dimension 1, la convergence des contrôles n'est pas nécessairement assurée par la convergence du schéma numérique utilisé. Ce phénomène a été illustré pour la première fois dans le contexte d'une approximation par éléments finis ou par différences finies dans [83].

Illustrons ce problème dans le cas d'un intervalle spatial $[0,1]$ pour avec une méthode d'éléments finis. Notons, pour $N \in \mathbb{N}$, $h = 1/N$ et $x_i = ih$ pour $0 \leq i \leq N$. Décomposons la solution dans la base de $\mathbb{P}_0^1$

$$u^h(x,t) = \sum_{k=1}^{N-1} \alpha_k(t)\psi_k(x),$$

où, pour $1 \leq k \leq N-1$

$$\psi_k(x) = \begin{cases} \dfrac{x - x_{i-1}}{h}, & x \in [x_{i-1}, x_i], \\ \dfrac{x_{i+1} - x}{h}, & x \in [x_i, x_{i+1}], \\ 0, & \text{sinon.} \end{cases}$$

Les coefficients α_k sont obtenus en résolvant la solution faible de

$$\int_0^1 u_{tt}^h(x,t)v^h \,\mathrm{d}x + \int_0^1 \nabla u^h(x,t)\nabla v^h \,\mathrm{d}x = 0, \quad \forall v^h \in \mathbb{P}_0^1. \tag{5.6}$$

En choisissant successivement $v^h = \psi_k$, on obtient le système linéaire

$$\alpha''(t) + K\alpha(t) = 0,$$

où $\alpha = (\alpha_1, \ldots, \alpha_{N-1})^t$ et où la matrice K est donnée par

$$K(i,j) = \begin{cases} 2/h^2, & \text{si } i = j, \\ -1/h^2, & \text{si } |i-j| = 1, \\ 0, & \text{sinon.} \end{cases}$$

Il est généralement reconnu que la convergence des contrôles est assurée si, en plus de considérer un schéma numérique convergeant, il existe une borne inférieure $\omega > 0$ uniforme par rapport au paramètre de discrétisation de l'écart des racines des valeurs propres

$$\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} \geq \omega > 0, \tag{5.7}$$

ainsi qu'une minoration $c_T > 0$ uniforme par rapport au paramètre de discrétisation de l'inégalité d'observabilité et une majoration $C_T > 0$ uniforme de l'inégalité de

continuité

$$c_T E(u^h(0)) \leq \int_0^T \left|u_x^h(1,t)\right|^2 \mathrm{d}t, \tag{5.8}$$

$$\int_0^T \left|u_x^h(1,t)\right|^2 \mathrm{d}t \leq C_T E(u^h(0)). \tag{5.9}$$

Dans le cas d'une semi-discrétisation par éléments finis $\mathbb{P}_0^1$, il n'est pas possible d'obtenir une borne inférieure uniforme par rapport au paramètre de discrétisation telle que (5.7) soit satisfaite. La figure 4 illustre le comportement de l'écart des valeurs propres pour la méthode des éléments finis $\mathbb{P}_0^1$. Des exemples de données initiales du

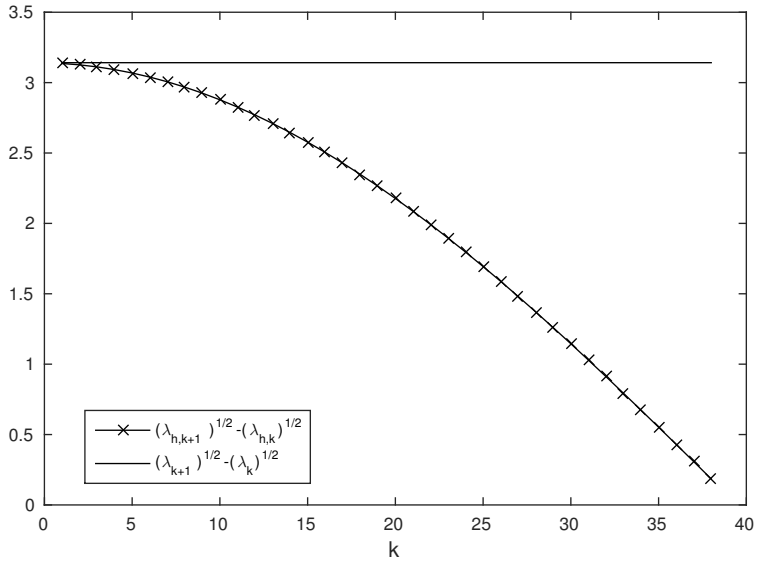

FIGURE 4 – Comparaison entre le comportement de l'écart des racines des valeurs propres discrètes et celui du cas continu pour $N = 40$

problème à contrôler menant à des contrôles numériques divergeant en norme $L^2(0,T)$ ont également été donnés dans [1].

En général, l'une des conditions (5.7)-(5.9) n'est pas satisfaite en raison du mauvais comportement des hautes fréquences : soit les valeurs propres discrètes ne constituent pas de bonnes approximations des valeurs propres du cas continu pour un ratio de l'ordre de $N/\pi < k < N$, soit les fonctions propres associées aux valeurs propres pour un ratio élevé de l'ordre de $N/\pi < k < N$ engendrent une dégénérescence de l'uniformité de la constante d'observabilité ou de la constante de continuité.

Une technique afin d'obtenir la convergence des contrôles numériques consiste à éliminer l'effet des hautes fréquences indésirables en projetant la solution de (5.6) sur l'espace engendré par les basses fréquences ([83, 136, 21]). Cette méthode possède

cependant le désavantage de devoir calculer les fonctions propres discrètes, ce qui est particulièrement coûteux lorsque la dimension du domaine spatial augmente. Il est donc souhaitable de trouver des schémas numériques permettant d'obtenir des bornes uniformes de (5.7)-(5.9).

5.2 Résultats antérieurs

Différentes méthodes ont été proposées par Glowinski ([74, 73]) afin d'obtenir la convergence des contrôles numériques lorsque la semi-discrétisation est effectuée avec des schémas numériques de différences finies ou d'éléments finis. Dans ce cas, c'est généralement la condition (5.7) qui n'est pas satisfaite uniformément, résultant en l'existence de suites de solutions discrètes telles que leur vitesse de propagation tend vers zéro.

Il a été montré dans [106] qu'une méthode bi-grille permet d'obtenir l'uniformité des constantes d'observabilité et de continuité $c_{h,T}$ and $C_{h,T}$ ainsi que la convergence des contrôles numériques. Dans [26], une formulation mixte de l'équation adjointe a également permis d'obtenir les constantes d'observabilité et de continuité uniformément par rapport au paramètre de discrétisation. La convergence du contrôle numérique a été démontrée. Ce résultat a été ensuite généralisé à une famille de semi-discrétisation par différences finies dans [105]. Le point commun de ces travaux est l'utilisation d'espaces d'approximations différents pour la composante d'élévation u et la composante de vitesse u_t de la solution du problème adjoint.

5.3 Résultats obtenus

Les résultats obtenus dans cette thèse concernent la semi-discrétisation de l'équation des ondes avec une méthode de Legendre-Galerkin. Les polynômes de Legendre sont définis par la relation de récurrence

$$\begin{aligned}
&L_0(x) = 1,\\
&L_1(x) = x,\\
&(k+1)L_{k+1}(x) = (2k+1)xL_k(x) - kL_{k-1}(x), \quad k \geq 2.
\end{aligned}$$

L'espace d'approximation $\mathbb{P}_N^0 \subset H_0^1$ est alors construit à l'aide de la base

$$\widetilde{L_k}(x) = c_k(1-x^2)L_k(x), \quad c_k = \frac{\sqrt{k+1/2}}{k(k+1)}, \quad k \geq 1.$$

La méthode de Legendre-Galerkin consiste à chercher une solution u^N de

$$\int_{-1}^{1} u_{tt}^N(x,t)v^N(x)\,\mathrm{d}x + \int_{-1}^{1} u_x^N(x,t)v_x^N(x)\,\mathrm{d}x = 0, \quad \forall v^N \in \mathbb{P}_N^0(-1,1), \tag{5.10}$$

de la forme

$$u^N(x,t) = \sum_{k=1}^{N-1} \alpha_{N,k}(t)\widetilde{L_k}(x).$$

L'étude du spectre de (5.10) ainsi que le comportement des constantes d'observabilité et de continuité associée à (5.10) a été effectuée dans [21]. Une borne inférieure sur l'écart entre les racines des valeurs propres a été montrée. La figure 5 montre le comportement des racines des valeurs propres discrètes.

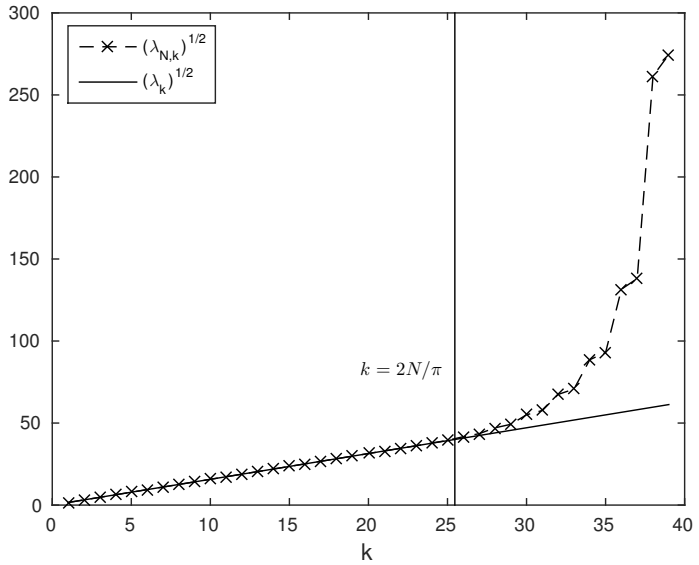

FIGURE 5 – Comparaison entre le comportement des racines des valeurs propres discrètes et les racines des valeurs propres du cas continu pour $N = 40$

Afin d'étudier le comportement de la constante d'observabilité discrète et de la constante de continuité discrète, introduisons

$$\Delta_{N,k} = \frac{\left|\phi'_{N,k}(1)\right|^2}{\displaystyle\int_{-1}^{1} \left|\phi'_{N,k}(x)\right|^2 \ \mathrm{dx}},$$

où $\phi_{N,k}$ est la k-ième fonction propre du problème discret. On voit aisément que la minoration uniforme de la constante d'observabilité discrète est équivalente à l'existence d'une borne inférieure uniforme par rapport au paramètre de discrétisation de $\Delta_{N,k}$ et que la majoration uniforme de la constante de continuité discrète est équivalente à l'existence d'une borne supérieure uniforme par rapport au paramètre de discrétisation de $\Delta_{N,k}$. Il a été démontré dans [21] que $c_{N,T} \to 0$ et $C_{N,T} \to \infty$ lorsque $N \to \infty$. Notons que dans le cas continu, $\Delta_k = 1, \forall k \in \mathbb{N}$ La figure (6) illustre le comportement des $\Delta_{N,k}$ pour $N = 40$.

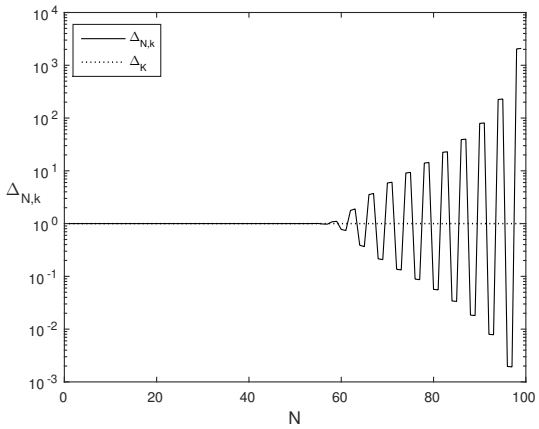

FIGURE 6 – Comparaison entre le comportement des $\Delta_{N,k}$ et la valeur du cas continu pour $N = 40$

Trois méthodes sont proposées dans le chapitre 4 afin d'obtenir l'uniformité des constantes d'observabilité discrètes avec la méthode de Legendre-Galerkin tout en préservant la minoration uniforme de l'écart entre les racines des valeurs propres discrètes. Afin d'alléger la présentation, nous ne présenterons que les résultats numéri ques portant sur la constante d'observabilité et de continuité discrète.

5.3.1 Filtres spectraux

La première méthode consiste à filtrer la trace normale apparaissant dans l'inégalité d'observabilité et l'inégalité de continuité à l'aide de filtres spectraux. Les filtres spectraux σ sont définis par :

Definition 5.1 *On dit que $\sigma : \mathbb{R} \to \mathbb{R}^+$ est un filtre d'ordre p ($p \geq 1$) si,*

(i) σ *est* $p-1$ *continûment différentiable sur* $\mathbb{R}$ *pour* $p \geq 1$,

(ii) $\sigma(\eta) = 0$, *si* $|\eta| \geq 1$,

(iii) $\sigma(0) = 1$, $\sigma^{(j)}(0) = 0$ *pour* $1 \leq j \leq p-1$.

En définissant

$$\mathcal{F}_N u^N(x) = \sum_{k=0}^{N} \sigma\left(\frac{k}{N}\right) \alpha_{N,k} L_k(x),$$

on s'intéresse aux inégalités suivantes

$$c_{N,T} E(u^N(0)) \leq \int_0^T \left|\left(\mathcal{F}_N u^N\right)_x (1,t)\right|^2 dt,$$

$$\int_0^T \left|\left(\mathcal{F}_N u^N\right)_x (1,t)\right|^2 dt \leq C_{N,T} E(u^N(0)).$$

Les filtres permettent d'amenuir l'impact des hautes fréquences sur la solution numéri que et ainsi d'obtenir une minoration uniforme de la constante d'observabilité. Il ne semble pas possible par contre de majorer uniformément la constante de continuité. La figure 7 illustre le comportement de la constante d'observabilité discrète ainsi que le comportement de la constante de continuité lorsque N augmente pour différents filtres spectraux.

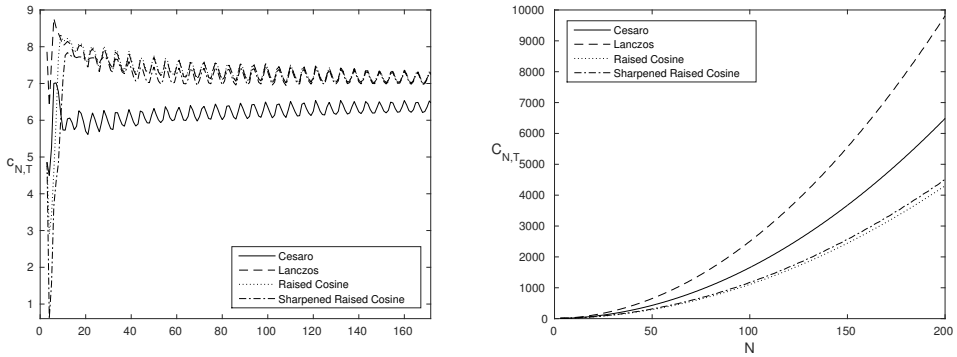

FIGURE 7 – Le comportement $c_{N,T}$ (à gauche) and $C_{N,T}$ (à droite) pour différents filtres spectraux et $T = 8$

5.3.2 Formulation mixte

La seconde méthode repose sur une formulation mixte de l'équation des ondes

$$\int_{-1}^{1} u_t^N(x,t)\widetilde{v}^M(x)\,\mathrm{d}x = \int_{-1}^{1} z^M(x,t)\widetilde{v}^M(x)\,\mathrm{d}x, \qquad \forall \quad \widetilde{v}^M \in \mathbb{P}^M$$

$$\frac{d}{dt} < z^M(t), \hat{v}^N >_{H^{-1},H_0^1} = -\int_{-1}^{1} u_x^N(x,t)\hat{v}_x^N(x)\,\mathrm{d}x, \qquad \forall \quad \hat{v}^N \in \mathbb{P}_0^N.$$

où $\mathbb{P}^M = \mathrm{span}\{L_0(x), \ldots, L_M(x)\}$. On cherche les solutions de la forme

$$\begin{aligned} u^N(x,t) &= \sum_{k=1}^{N-1} \alpha_{N,k}(t), \widetilde{L_k}(x), \\ z^M(x,t) &= \sum_{k=0}^{M} \beta_{M,k}(t) L_k(x). \end{aligned}$$

On s'intéresse au comportement des constantes d'observabilité et de continuité des inégalités suivantes

$$c_{N,T} E(u^N(0), z^M(0)) \leq \int_0^T \left| u_x^N(1,t) \right|^2 dt,$$

$$\int_0^T \left|u_x^N(1,t)\right|^2 dt \leq C_{N,T} E(u^N(0), z^M(0)),$$

où

$$E(u^N(t), z^M(t)) = \frac{1}{2}\int_{-1}^{1} \left|z^M(x,t)\right|^2 + \left|u^N(x,t)\right|^2 \mathrm{dx},$$
$$E(u^N(t), z^M(t)) = E(u^N(0), z^M(0)).$$

Cette formulation permet d'obtenir numériquement une minoration uniforme de la constante d'observabilité discrète. Toutefois, il ne semble pas possible d'obtenir une majoration uniforme sur la constante de continuité discrète. La figure 8 illustre le comportement de la constante d'observabilité discrète ainsi que le comportement de la constante de continuité lorsque N augmente pour la formulation mixte.

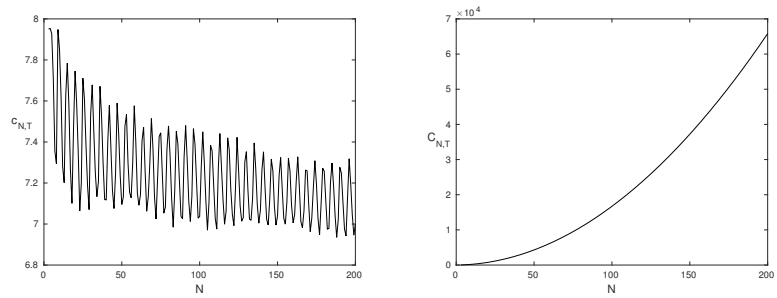

FIGURE 8 – Le comportement $c_{N,T}$ (à gauche) and $C_{N,T}$ (à droite) pour la formulation mixte et $T = 8$

5.3.3 Méthode de Nitsche

Finalement, nous considérons une formulation de Nitsche pour l'équation des ondes

$$\int_{-1}^{1} u_{tt}^N(x,t)\psi^N(x)\,\mathrm{d}x + a^N(u^N, \psi^N) = 0, \quad \forall v^N \in \mathbb{P}_L^N, \tag{5.11}$$

où

$$a^N(u^N, \psi^N) = \int_{-1}^{1} u_x^N(x,t)\psi_x^N(x)\,\mathrm{d}x - u_x^N(1,t)\psi^N(1) - u^N(1,t)\psi_x^N(1) + \gamma N^2 u^N(1,t)\psi^N(1),$$

et où l'espace d'approximation est $\mathbb{P}_L^N = \mathrm{span}\{\widehat{L_1}(x), \ldots, \widehat{L_N}(x)\} \subset H_L^1(-1,1)$ où $H_L^1(-1,1) = \{\psi \in H^1(-1,1) \,|\, \psi(-1) = 0\}$ avec

$$\widehat{L_k}(x) = L_k(x) - (-1)^k, \quad k \geq 1.$$

La solution est cherchée de la forme

$$u^N(x,t) = \sum_{k=1}^{N} \alpha_{N,k}(t)\widehat{L_k}(x).$$

La particularité de la méthode de Nitsche est la possibilité d'imposer faiblement les conditions aux bords de Dirichlet. Dans le cas considéré ici, seulement la condition de Dirichlet au bord $x = 1$ a été imposée faiblement. Des résultats similaires sont obtenus si les deux conditions aux bords sont imposées faiblement. La méthode de Nitsche non symétrique sera également présentée dans cette thèse mais nous limitons l'introduction au cas symétrique, les résultats du cas non symétrique étant également similaires.

La méthode de Nitsche symétrique ne permet pas d'obtenir des bornes uniformes sur les constantes d'observabilité et de continuité discrètes associées à (5.3) et (5.4). Elle permet cependant d'obtenir une minoration uniforme d'une inégalité discrète différente. Reprenons les calculs (2.8) caractérisant la contrôlabilité d'une équation. Considérons l'équation des ondes à contrôler semi-discrétisée par une méthode de Nitsche

$$\begin{aligned}\int_{-1}^{1} y_{tt}^N(x,t)\psi^N(x)\,\mathrm{d}x + \int_{-1}^{1} y_x^N(x,t)\psi_x^N(x)\,\mathrm{d}x - y_x^N(1,t)\psi^N(1) - y^N(1,t)\psi_x^N(1)\\ +\gamma N^2 y^N(1,t)\psi^N(1) = v(t)\left(-\psi_x^N(1) + \gamma N^2\psi^N(1)\right), \quad \forall \psi^N \in \mathbb{P}_L^N.\end{aligned}$$

En remplaçant $\psi^N(x)$ par la solution u^N, on obtient

$$\int_0^T v(t)\left(-u_x^N(1,t) + \gamma N^2 u^N(1,t)\right)\,\mathrm{d}t = \int_{-1}^{1} u_1^N(x)y_0^N(x) - u_0^N(x)y_1^N(x)\,\mathrm{d}x,$$

ce qui suggère un nouveau choix d'inégalité d'observabilité et de continuité pour la nouvelle fonctionnelle

$$\begin{aligned}J^N(u_0^N, u_1^N) \;=\; & \frac{1}{2}\int_0^T \left|-u_x^N(1,t) + \gamma N^2 u^N(1,t)\right|^2 \mathrm{d}t\\ & - < (y_1, -y_0), (u_0^N, u_1^N) >_{H^{-1}\times L^2, H_0^1\times L^2}.\end{aligned}$$

De cette nouvelle formulation, nous obtenons une minoration uniforme de la constante d'observabilité. Par contre, il ne semble pas possible d'obtenir une majoration uniforme de la constante de continuité discrète. La figure 9 illustre le comportement des constantes d'observabilité et de continuité discrète lorsque N augmente. Notons que la nouvelle inégalité d'observabilité discrète a été obtenue grâce au fait que l'espace d'approximation $\mathbb{P}_L^N$ n'est pas un sous-espace de H_0^1, comme il est généralement le cas. Il n'est pas rare de voir des termes pénalisés dans des inégalités d'observabilité afin d'obtenir une borne uniforme sur les constantes d'observabilité et de continuité

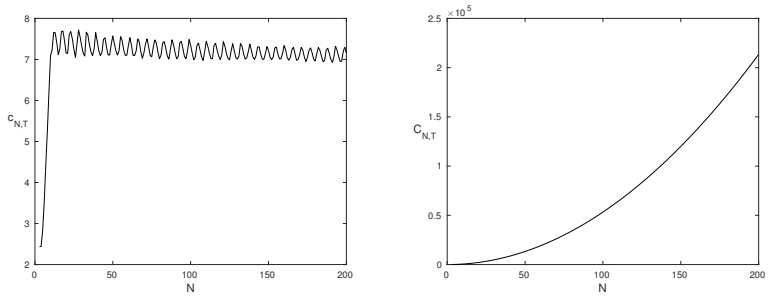

FIGURE 9 – Le comportement des constantes $c_{N,T}$ (gauche) et $C_{N,T}$ (droite) de la formulation de Nitsche symétrique pour $N = 3, .., 250$, $\gamma = 0.8$ et $T = 8$

discrètes mais il s'agit, à notre connaissance, de la première fois que la nouvelle expression de ces inégalités s'obtient directement de la caractérisation de la contrôlabilité de l'équation semi-discrétisée.

Une étude sur la convergence des contrôles numériques est finalement présentée. L'exemple d'une donnée initiale engendrant un contrôle numérique divergeant pour le schéma de Legendre-Galerkin non-filtré est présenté alors que les schémas pour lesquels seule une minoration uniforme de la constante d'observabilité a été obtenue engendrent des contrôles convergeant.

5.4 Discussion des résultats obtenus et perspectives

Les schémas numériques présentés dans cette thèse ne permettent d'obtenir numé riquement qu'une minoration de la constante d'observabilité ainsi que de l'écart des racines des valeurs propres. Une preuve théorique de ces résultats est relativement difficile en raison de l'expression implicite des valeurs propres des formulations considérées (voir [21] pour l'étude du spectre de (5.10)). De plus, à notre connaissance, il n'existe pas de résultat dans la littérature sur la convergence des contrôles numériques sans l'uniformité de la constante de continuité. Cette borne uniforme permet d'extraire une sous-suite de contrôles numériques convergeant. L'uniformité de la constante d'observabilité permet essentiellement de montrer ensuite que la limite de cette sous-suite est le contrôle de norme minimal $L^2(0, T)$ du problème continu.

Suite aux expérimentations numériques sur la convergence des contrôles sans l'uniformité de la constante de continuité, il serait intéressant de voir si d'autres conditions sur le schéma numérique permettent de montrer la convergence des contrôles numériques vers les contrôles continus. Il serait également intéressant de voir si les méthodes proposées ici permettent d'obtenir l'uniformité de la constante d'observabilité pour l'équation des ondes avec un domaine spatial de dimension supérieur ainsi que pour l'équation des poutres pour laquelle des problèmes similaires apparaissent (voir notamment [92]).

Première partie

L'équation de Korteweg-de Vries

Chapitre 1

Contrôlabilité lagrangienne de l'équation de Korteweg-de Vries

Contents

1.1 Introduction **81**
1.2 Well-posedness and regular controls **90**
1.2.1 Well-posedness of the KdV equation 90
1.2.2 Regular controls 92
1.3 Lagrangian controllability **96**

1.1 Introduction

The Korteweg-de Vries (KdV) equation defined on the real line,

$$y_t + y_x + y_{xxx} + yy_x = 0, \quad x \in \mathbb{R}, t \in \mathbb{R}, \tag{1.1.1}$$

first derived, independently, by Boussinesq in 1877 ([23]) and by Korteweg and de Vries in 1895 ([86]), is obtained as a first order approximation of the free-surface solution of the full governing equations for a homogeneous, non-viscous and irrotational shallow fluid (see, for example, [131, p. 460] for a complete derivation of the KdV from the governing equations). In the context of water waves, solutions of (1.1.1) correspond to the free-surface of approximately two-dimensional waves (the motion of the waves are assumed to be parallel to the crest) travelling from left to right. More recently, the KdV equation has found applications in the context of collisionless plasma hydromagnetic waves [66], long waves in anharmonic crystals [133], ion-acoustic plasma [130] and cosmology [93].

In this paper, we are interested by the small-time Lagrangian controllability of

the Korteweg-de Vries equation starting from rest,

$$y_t + y_x + y_{xxx} + yy_x = 0, \qquad x \in [0, L], \quad t \in [0, T], \tag{1.1.2}$$
$$y(0, t) = u(t), \qquad t \in [0, T], \tag{1.1.3}$$
$$y(L, t) = v(t), \qquad t \in [0, T], \tag{1.1.4}$$
$$y_x(L, t) = w(t), \qquad t \in [0, T], \tag{1.1.5}$$
$$y(x, 0) = 0, \qquad x \in [0, L], \tag{1.1.6}$$

with boundary controls $u(t), v(t)$ and $w(t) \in \mathbb{R}$ and $T, L > 0$. The Lagrangian controllability of (1.1.2)-(1.1.6) is defined as follow,

Definition 1.1.1 (Small-time Lagrangian Controllability) *Equations (1.1.2)-(1.1.6) are small-time Lagrangian controllable if and only if, for all $T, L > 0$, there exists $u(t), v(t)$ and $w(t) \in \mathbb{R}$ such that, if we consider $\hat{y}$ the extension of the solution y of (1.1.2)-(1.1.6) by*

$$\hat{y}(x, t) = \begin{cases} y(0, t), & \text{if } x \leq 0, \\ y(x, t), & \text{if } x \in [0, L], \\ y(L, t), & \text{if } x \geq L, \end{cases} \tag{1.1.7}$$

the flow Φ defined by

$$\begin{cases} \dfrac{\partial \Phi}{\partial t}(x, t) &= \hat{y}(\Phi(x, t), t), & x \in \mathbb{R}, t \in \mathbb{R}^+, \\ \Phi(x, 0) &= x, & x \in \mathbb{R}, \end{cases} \tag{1.1.8}$$

satisfies $\Phi(x, T) \geq L$, $\forall x \in [0, L]$.

For sake of simplicity, we will always refer, in the following, to this construction when speaking of the flow of a solution of (1.1.2)-(1.1.6).

Let us describe a physical interpretation of the Lagrangian controllability. Since we consider approximately two-dimensional waves, consider the Cartesian coordinates (x, z) such that x is the horizontal direction in which the waves travel and $z = 0$ denotes the flat bottom of the fluid, $z = h_0$ the height of the fluid at rest and $z = h_0 + y(x, t)$ the free-surface where y is solution of (1.1.2)-(1.1.6). The flow defined by (1.1.8) is, up to a physical constant, the horizontal component (independent of z) of the first order approximation of the velocity field of (1.1.1) ([131, p. 460]). Therefore, one may interpret the problem of the Lagrangian controllability of (1.1.2)-(1.1.6) as the problem of moving the particles initially located in the region $[0, L] \times [0, h_0]$ at time $t = 0$ to the right of L at time $t = T$ by means of waves created by the boundary controls. A possible application of the Lagrangian controllability is the displacement of polluted water in channels to a waste water treatment plant.

Remark 1.1.1 *It is a natural condition to impose that the flow exit to the right and not to the left since solutions of the KdV equation correspond to waves travelling from left to right. This assumption in the derivation of the KdV equation (1.1.1) is transposed in the asymptotic behaviour, given by the Inverse Scattering Method ([120]), of solutions of (1.1.1) for smooth initial data : a finite number of solitons travelling to the right and a decaying wave train to the left. This asymmetric behaviour is different, for example, than the Euler equation for which there would be no geometrical restriction on where the flow should leave for a similar problem. In fact, the solution constructed for the return method by Coron and Glass to show the global Eulerian controllability of the 2-D and 3-D Euler equation respectively ([41],[67]) also provides a proof that the flow exit the domain for the 2-D and 3-D Euler equation.*

The main result of this paper is the small-time Lagrangian controllability of (1.1.2)-(1.1.6), with the additional property that the solution is at rest at time $t = T$.

Theorem 1.1.1 *Let $T, L > 0$. Then, there exists $y \in C([0,T]; H^2(0,L))$ solution of (1.1.2)-(1.1.6) such that the associated flow Φ satisfies*

$$\frac{\partial \Phi}{\partial t}(x,0) = 0, \qquad x \in [0,L], \tag{1.1.9}$$

$$\frac{\partial \Phi}{\partial t}(x,T) = 0, \qquad x \in [0,L], \tag{1.1.10}$$

$$\Phi(x,t) \geq L, \qquad x \in [0,L],\ t \in [T,\infty). \tag{1.1.11}$$

Theorem 1.1.1 follows from an explicit solution of (1.1.1) satisfying (1.1.11) and a smallness condition on the state in the neighborhood of $t = 0$ and $t = T$.

Theorem 1.1.2 *Let $L, T, \delta > 0$ and $(\epsilon_1, \epsilon_2) \in (0, T/2)^2$. Then there exists a positive solution $y \in C^\infty([0,T] \times \mathbb{R})$ of (1.1.1) such that*

$$\|y(.,t)\|_{H^2(0,L)} < \delta, \quad \forall t \in (0,\epsilon_1) \cup (T - \epsilon_2, T),$$

and such that the flow $\Phi(x,t)$ associated to y satisfies

$$\Phi(x,t) \geq L, \quad \forall (x,t) \in [0,L] \times (T - \epsilon_2, T).$$

This solution is constructed by means of the N-solitons solution. For later conveniences, let us express the soliton solution of KdV for the change of variables $x \mapsto x - t$, $y \mapsto 6\eta$. The Korteweg-de Vries equation becomes

$$\eta_t + 6\eta\eta_x + \eta_{xxx} = 0, \quad x \in \mathbb{R}, t \in \mathbb{R}, \tag{1.1.12}$$

and solitons of (1.1.12) are given, for $\alpha > 0$ and $s \in \mathbb{R}$, by

$$\eta(x,t) = \frac{\alpha^2}{2}\operatorname{sech}^2\left(\frac{-\alpha(x-s)+\alpha^3 t}{2}\right). \tag{1.1.13}$$

We note the following for solitons given by (1.1.13)

1. The amplitude is given by $\alpha^2/2$ and is reached at $x = s + \alpha^2 t$;
2. The traveling speed is α^2;
3. The distance between the x coordinates where the height of the soliton is $\alpha/4$, defined as the width, is

$$w(\alpha) := \frac{4}{\alpha}\ln(\sqrt{2\alpha}(1+\sqrt{1-\frac{1}{2\alpha}})). \tag{1.1.14}$$

From these properties, one remarks that taller solitons travel faster and are narrower. Moreover, the width of a soliton tend to infinity as its amplitude tends to zero.

We point out that Theorem 1.1.1 is not the consequence of the passage of a single soliton inside the domain $[0, L]$. Indeed, consider the flow defined on the whole real line,

$$\frac{\partial \Phi}{\partial t}(x,t) = \eta(\Phi(x,t),t), \quad x \in \mathbb{R}, t \in \mathbb{R},$$

where η is a soliton with $\alpha > 0$ and $s \in \mathbb{R}$, solution of (1.1.12). Then, from (1.1.13), one obtains that the total displacement $|\Phi(x,\infty) - \Phi(x,-\infty)|$ is of order $2/\alpha$ and, since the speed of propagation of a soliton is α^2, one cannot obtain both a large displacement and an arbitrarily small time.

We rather use the N-solitons solution to prove Theorem 1.1.1. Expressed in the closed form by Hirota in 1971 [80], the N-solitons solution of (1.1.12) takes the form

$$\begin{aligned} \eta &= -2\left(\ln F\right)_{xx}, \\ F &= 1 + \sum_{n=1}^{N}\sum_{{}_NC_n} a(i_1,...,i_n)\prod_{j=1}^{n} f_{i_j}, \end{aligned} \tag{1.1.15}$$

where, for $1 \le i \le N$, $N \in \mathbb{N}$,

$$f_i(x,t) = \exp\left(-\alpha_i(x-s_i)+\alpha_i^3 t\right), \; s_i \in \mathbb{R}, \alpha_i > 0,$$

where

$$a(i_1,...,i_n) = \prod_{k<l}^{n} a(i_k,i_l), \qquad \text{if } n \ge 2,$$

$$a(i_k, i_l) = \left(\frac{\alpha_{i_k} - \alpha_{i_l}}{\alpha_{i_k} + \alpha_{i_l}}\right)^2,$$
$$a(i_k) = 1,$$

where $\sum_{{}_NC_n}$ is the sum over all the n indexes $i_1, ..., i_n$, taken, without permutations, from $\{1, \ldots, N\}$.

To explain why this solution is called the N-solitons solution, let us consider the case where $N = 2$. The solution has the form

$$\frac{\eta}{2} = \frac{\alpha_1^2 f_1 + \alpha_2^2 f_2 + 2(\alpha_2 - \alpha_1)^2 f_1 f_2 + ((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2(\alpha_2^2 f_1^2 f_2 + \alpha_1^2 f_2^2 f_1)}{(1 + f_1 + f_2 + ((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2 f_1 f_2)^2}.$$

If $f_1 \simeq 1$, $f_2 \ll 1$, the behaviour of η is then given by

$$\begin{aligned}\eta &\simeq 2\frac{\alpha_1^2 f_1}{(1 + f_1)^2} \\ &= \frac{\alpha_1^2}{2}\operatorname{sech}^2\left(\frac{-\alpha_1(x - s_1) + \alpha_1^3 t}{2}\right),\end{aligned}$$

while, in the case where $f_2 \simeq 1$, $f_1 \gg 1$, we have,

$$\begin{aligned}\eta &\simeq 2\frac{\alpha_2^2((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2 f_1^2 f_2}{(f_1 + ((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2 f_1 f_2)^2} \\ &= 2\frac{\alpha_2^2 \exp(\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)) f_2}{(1 + \exp(\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)) f_2)^2} \\ &= \frac{\alpha_2^2}{2}\operatorname{sech}^2\left(\frac{-\alpha_2(x - s_2 - \frac{1}{\alpha_2}\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)) + \alpha_2^3 t}{2}\right),\end{aligned}$$

that is, a soliton with a phase shift of $\frac{1}{\alpha_2}\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)$.

Let us now describe the behaviour of the 2-solitons solution in the case where $0 < \alpha_1 < \alpha_2$ and let us denote them soliton 1 and 2 respectively. When $t \to -\infty$, the solitons behaves like 2 distinct solitons, soliton 1 being ahead of soliton 2 and the latter having a phase shift of $\frac{1}{\alpha_2}\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)$. When $f_1 \simeq f_2 \simeq 1$, interactions occur between soliton 1 and 2 and, during this period, their combined amplitude decreases. Moreover, if α_1 and α_2 are of same magnitude, they exchange their amplitudes and velocities [134]. After the interaction, they behaves as distinct solitons and are left unchanged in shape, the only notable effect is the phase shift of soliton 1, of $\frac{1}{\alpha_1}\ln(((\alpha_2 - \alpha_1)/(\alpha_2 + \alpha_1))^2)$, while the soliton 2 no longer has one. Those effects are the results of the nonlinearity of the equation. Figure 1 illustrates the phase shift produced when two solitons interacts for the 2-solitons solution. The frame is fixed at the speed of the slower soliton.

Let us sketch the three steps of the proof of Theorem 1.1.2. Consider the N-

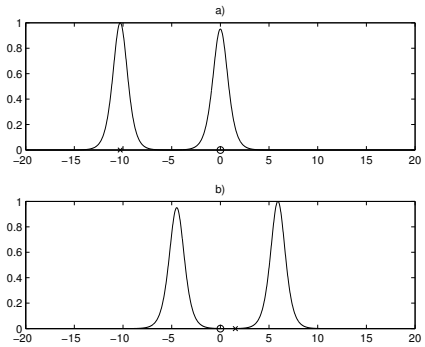

FIGURE 1 – An interaction between two solitons. The cross (circle) represents the position of the maximum of faster (slower) soliton. Figure a) is the state of the solution before the collison and b) is after the collision

solitons solution.

Step 1. At time $t = 0$, the solitons are located at the left of $x = 0$. They are ordered in increased order of height to prevent further interactions. During this step, only the tail of the N-solitons solution is located inside the domain $[0, L]$. An estimation of the norm of the tail with respect of α_i and N shows that the larger the α_i are, the smaller the norm of the tail is.

Step 2. During the time interval $(0, T)$, the N solitons travel inside the domain. A lower bound of the flow is provided with respect of α_i and N. This lower bound is estimated only by the displacement induced by the N solitons, the displacement due to the interactions being negligible.

Step 3. At time $t = T$, the solitons are located at the right of $x = L$. Only the tail of the N-solitons solution is located in $[0, L]$. A similar estimate than in step 1 proves that the norm of the tail is small for α_i large.

With the required estimates at hand, one chooses α_i and N large enough to obtain Theorem 1.1.2.

To prove Theorem 1.1.1, one brings the solution constructed in the proof of Theorem 1.1.2 to rest at time $t = 0$ and $t = T$ with the (Eulerian) local controllability of (1.1.2)-(1.1.6). One notices that while the regularity of the solution constructed in Theorem 1.1.2 is sufficient to define the flow pointwisely, the usual L^2 local controllability of (1.1.2)-(1.1.6) is not. We therefore use the local controllability result of Zhang. To state the result, let us consider a non-zero initial data

$$y(x, 0) = y_0(x), \quad x \in [0, L] \tag{1.1.16}$$

and let us denote the set of equations (1.1.2), (1.1.3), (1.1.4), (1.1.5) and (1.1.16) by (1.1.2)-(1.1.16). In [135], the following was proven.

Theorem 1.1.3 *Let $T > 0$ and $s \geq 0$ be given and $[0, L] \subset (\alpha_1, \beta_1)$. Suppose that*

$$w \equiv w(x,t) \in C^\infty((\alpha_1, \beta_1) \times (-\epsilon, T + \epsilon)),$$

for some $\epsilon > 0$, satisfies

$$w_t + w_x + w_{xxx} + ww_x = 0, \qquad (x,t) \in (\alpha_1, \beta_1) \times (-\epsilon, T + \epsilon).$$

Then there exists $\delta > 0$ such that for any $y_0, y_T \in H^s(0, L)$ satisfying

$$\|y_0 - w(.,0)\|_{H^s(0,L)} \leq \delta \quad \textit{and} \quad \|y_T - w(.,T)\|_{H^s(0,L)} \leq \delta,$$

one can find control inputs $(u, v, w) \in H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T)$ such that (1.1.2)-(1.1.16) has a solution

$$y \in C([0,T]; H^s(0,L)) \cap L^2((0,T); H^{s+1}(0,L)),$$

satisfying

$$y(x,0) = y_0(x) \quad \textit{and} \quad y(x,T) = y_T(x),$$

on the interval $(0, L)$.

Using Theorem 1.1.3 with $s = 2$ to connect the solution of Theorem 1.1.2 to rest is sufficient to define the flow pointwisely, hence the statement of Theorem 1.1.1 in $C([0,T]; H^2(0,L))$.

During the control phases, from the lack of maximum principle for the KdV equation, one cannot insure that the flow does not exit the domain $[0, L]$ from the left. Therefore, we use a stability estimate of the controls of Theorem 1.1.3 with respect with the initial data to estimate the flow during the control phases. The following corollary is a consequence of the results in [135].

Corollary 1.1.1 *Let $T > 0$ and $s \geq 0$. Then, there exists $\delta > 0$ such that for any $y_0, y_T \in H^s(0,1)$ satisfying*

$$\|y_0\| < \delta \quad \textit{and} \quad \|y_T\| < \delta$$

there exists control inputs $(u, v, w) \in H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T)$ such that (1.1.2)-(1.1.16) has a solution

$$y \in C([0,T]; H^s(0,L)) \cap L^2((0,T); H^{s+1}(0,L)),$$

satisfying

$$y(x,0) = y_0(x) \quad \textit{and} \quad y(x,T) = y_T(x),$$

on the interval $(0, L)$*. Moreover,*

$$\|u\|^2_{H^{\frac{s+1}{3}}(0,T)} + \|v\|^2_{H^{\frac{s+1}{3}}(0,T)} + \|w\|^2_{H^{\frac{s}{3}}(0,T)} \leq c\left(\|y_0\|^2_{H^s(0,L)} + \|y_T\|^2_{H^s(0,L)}\right)$$

where $c > 0$ *is independant of* y_0 *and* y_T*.*

Corollary 1.1.1 is the stability estimate of Theorem 1.1.3 in the simpler case $w \equiv 0$ for which the Banach Fixed Point Theorem can be used. We comment the proof of the estimate in Section 2 for sake of completeness.

Aside of Theorem 1.1.3, several other results of Eulerian controllability for (1.1.2)-(1.1.16) are found in literature. When only the control w is used ($u, v \equiv 0$), the local controllability around the equilibrium of (1.1.2)-(1.1.16) was obtained by Rosier ([116]) if $L \notin \mathcal{N}$, that is, when the linearized equation around the equilibrium is exactly controllable. If $L \in \mathcal{N}$, there exists an unreachable state subspace for the linearized equation around the equilibrium. Using the power expansion method ([42]) to reach this subspace, the local controllability around the equilibrium of (1.1.2)-(1.1.16) was obtained by Coron and Crépeau ([45]), Cerpa ([27]) and Cerpa and Crépeau ([30]) when the dimension of the unreachable states subspace is of dimension 1, 2 and of arbitrarily dimension, respectively. If one only uses u as a control ($v, w \equiv 0$), then (1.1.2)-(1.1.16) is locally controllable to zero ([69]). If one only uses v ($u, w \equiv 0$), then (1.1.2)-(1.1.16) is locally controllable around the equilibrium if L doesn't belong to a countable set of critical lengths $\mathcal{O}$ ([70]). If one only uses v and w ($u \equiv 0$), then it was shown in [116] that the system is small-time controllable. Any other combination of two controls also leads to the small-time controllability ([69]). Good surveys on the small-time eulerian controllability and stability of (1.1.2)-(1.1.16) can be found in [118, 28]

It is important to note here that, exception made of Theorem 1.1.3, none of the previously mentioned results of small-time Eulerian controllability of (1.1.2)-(1.1.16) allows to define (1.1.8) pointwisely, the regularity of the solution of (1.1.2)-(1.1.16) obtained with these results being at most in $C([0,T]; L^2(0,L)) \cap L^2([0,T]; H^1(0,L))$.

One finds in the literature two results of global Eulerian controllability for the KdV equation. Rosier proved in [117] that (1.1.2)-(1.1.16) is globally controllable, that is, that there are no smallness restrictions on the initial or final data. However, the minimal time of controllability $T > 0$ depends on the initial and final data and may be large. By considering

$$y_t + y_x + y_{xxx} + yy_x = a(t),\ x \in [0, L], t \in [0, T], \tag{1.1.17}$$

instead of (1.1.2), Chapouly proved in [32] the small-time global Eulerian controllability of (1.1.17), (1.1.3), (1.1.4), (1.1.5) and (1.1.16) using the controllability of the non viscous Burgers equation ([33]) where $a(t)$ is used as a fourth control.

To date, the following challenging open problem still holds.

Open Problem 1.1.1 *Let $T, L > 0$. For any $y_0 \in L^2(0, L)$ and $y_T \in L^2(0, L)$, does there exist $u(t), v(t)$ and $w(t) \in \mathbb{R}$ such that the solution y of (1.1.2)-(1.1.16) satisfies*

$$y(x, T) = y_T(x)?$$

Few results on Lagrangian controllability are found in the literature. Glass and Horsin showed for the 2-D ([71]) and the 3-D ([72]) Euler equation that, for two given smooth contractible sets of particles surrounding the same volume of fluids and any initial velocity field, it is possible to find a boundary control and a time interval such that the corresponding solution of the Euler equation makes the first set reaches approximately the second. Horsin proved in [81] for the heat equation posed on $[0, L], L > 0$ that if one consider any two closed interval of $[0, L]$, then there exists a boundary control such that the flow induced by the solution of the heat equation starting from the first interval reaches the second. In the same article, he proved in the case of a radial domain (respectively a convex domain) of higher dimension, one can move two regular closed sets with the flow induced by minus the gradient of the solution by a control action on a part of the domain in arbitrarily small-time (respectively sufficiently large time). Finally, Horsin proved, in the case of the Burgers viscous equation that the previously mentioned result holds locally, that is, for two intervals not too far apart ([82]).

Considering the controllability of a PDE written in Lagrangian coordinates, one finds the work of Rosier on the Korteweg-de Vries equation written in Lagrangian coordinates with wave-maker controls. He proved in [94] the local controllability around regular trajectories.

One notes that the global controllability results of nonlinear equations are usually obtained by considering either the linear or nonlinear part of the equation as a perturbation. Fabre proved the approximate controllability of variations of the Navier-Stokes equation using the latter approach by truncating the nonlinearity ([57]). Still using the latter method, Fernandez-Cara and Zuazua [59] proved that, for the semilinear heat equation, the equation is null controllable if the nonlinear term is of controlled growth (see also [11], Problem 5.5, for a review and open problems of exact controllability of the semi-linear wave equation). Finally, Lions and Zuazua proved the controllability for some fluid systems using a Galerkin's approximation ([97]).

When considering the former approach, one may mimic the following finite dimensional result. Consider the finite dimensional system $y' = F(y) + Bu$ where F is quadratic ($F(\lambda y) = \lambda^2 F(y)$) and assume that there exists a trajectory $(\overline{y}, \overline{u})$, satisfying $\overline{y}(0) = \overline{y}(T) = 0$, of the system such that the linearized system around this trajectory is controllable. Then, by performing a scaling, one can show that $y' = F(y) + Ay + Bu$ is globally controllable $\forall A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$ ([43]). It is by using this technique that Chapouly proved the that the small-time global Eulerian controllability of (1.1.17)-(1.1.16) ([32]). This technique was also used by Coron and Glass to show, respectively, the global Eulerian controllability of the 2-D and 3-D Euler

equation ([41, 67]). Coron [40] and Coron and Fursikov [47] proved the global Eulerian controllability of the 2-D Navier-Stokes equations, using the global Eulerian controllability of the 2-D Euler equation, in the case where the whole boundary is used to control the interior or in the case of a Navier slip boundary condition.

One remarks that if the result holds in finite dimension for any linear operator $A \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, the presence of high order derivatives in the linear term and boundary layers issues may prevent one to apply this result in the infinite dimension framework. The first example where the global Eulerian controllability was obtain despite the presence of a boundary layer is due to Marbach, using the Hopf-Cole transformation and the maximum principle to show the small-time global null controllability of viscous Burgers equation ([100]).

The novelty of this paper is that we make full use of both the linearity and the nonlinearity of the KdV equation to obtain Theorem 1.1.1, as solitons don't exists if the linear or nonlinear term is dropped from (1.1.12). It is thus, to our knowledge, the first global controllability result obtained for a nonlinear equation without considering the linear or nonlinear part as a perturbation.

The outline of the paper is the following. In Section 2, we state the well-posedness results for the linear and nonlinear KdV equation and review Corollary 1.1.1. Section 3 is devoted to the proof of the main result.

1.2 Well-posedness and regular controls

1.2.1 Well-posedness of the KdV equation

Consider the linear KdV equation

$$\begin{cases} y_t + y_x + y_{xxx} = 0, & x \in [0, L], \quad t \in [0, T], \\ y(0,t) = u(t), y(L,t) = v(t), & t \in (0, T), \\ y_x(L,t) = w(t), & t \in (0, T), \\ y(x, 0) = y_0(x), & x \in (0, L). \end{cases} \tag{1.2.1}$$

The well-posedness of (1.2.1) was obtained in [13], exhibiting the smoothing effects of the solution y with respect to the initial value and the boundary data.

Theorem 1.2.1 *Let $L, T > 0$. Let $y_0 \in H^s(0, L)$ and $(u, v, w) \in H^{\frac{s+1}{3}}(0, T) \times H^{\frac{s+1}{3}}(0, T) \times H^{\frac{s}{3}}(0, T)$. Then, the problem (1.2.1) has a unique solution in*

$$C([0, T]; H^s(0, L)) \cap L^2((0, T); H^{s+1}(0, L)).$$

Moreover, there exists $C > 0$ such that

$$\|y\|^2_{C(0,T;H^s(0,L))\cap L^2((0,T);H^{s+1}(0,L))} \leq C\Big(\|y_0\|^2_{H^s(0,L)} + \|u\|^2_{H^{s/3}(0,T)} + \|v\|^2_{H^{s/3}(0,T)} + \|w\|^2_{H^{s/3}(0,T)}\Big).$$

Let us state the global well-posedness of the non-linear KdV equation (1.1.2)-(1.1.16) obtained in [13] (we refer to [58] for a sharper result on the compatibility conditions). In order to state the well-posedness result, one needs to consider the s-compatibility conditions.

Definition 1.2.1 (s-compatibility conditions) *Let $T, L, s > 0$. A four-tuple $(\tilde{y}, u, v, w) \in H^s(0,L) \times H^{(s+1)/3}(0,T) \times H^{(s+1)/3}(0,T) \times H^{s/3}(0,T)$ is said to be s-compatible if*

$$\tilde{y}_k(0) = u^{(k)}(0), \quad \tilde{y}_k(L) = w^{(k)}(0), \tag{1.2.2}$$

hold for :

1. *$k = 0, ..., \lfloor s/3 \rfloor - 1$ when $s - 3\lfloor s/3 \rfloor \leq 1/2$;*
2. *$k = 0, ..., \lfloor s/3 \rfloor$ when $3/2 \geq s - 3\lfloor s/3 \rfloor > 1/2$;*

and

$$\tilde{y}_k(0) = u^{(k)}(0), \quad \tilde{y}_k(L) = w^{(k)}(0), \tilde{y}'_k(L) = h_3^{(k)}(0), \tag{1.2.3}$$

holds for $k = 0, ..., \lfloor s/3 \rfloor$ when $s - 3\lfloor s/3 \rfloor > 3/2$, where

$$\begin{cases} \tilde{y_0}(x) := \tilde{y}(x), \\ \tilde{y_k}(x) := -\tilde{y}^{(3)}_{k-1}(x) - \tilde{y}'_{k-1}(x) - \displaystyle\sum_{j=0}^{k-1} (\tilde{y}_j(x)\tilde{y}_{k-j-1}(x))', \quad k \in \mathbb{N}. \end{cases}$$

We assume that (1.2.2) is vacuous if $\lfloor s/3 \rfloor - 1 < 0$.

Let $\epsilon > 0$ and

$$\mu_1(s) := \begin{cases} \epsilon + (5s+9)/18 & \text{if } 0 \leq s \leq 3, \\ (s+1)/3 & \text{if } 3 \leq s, \end{cases}$$

$$\mu_2(s) := \begin{cases} \epsilon + (5s+3)/18 & \text{if } 0 \leq s \leq 3, \\ (s+1)/3 & \text{if } 3 \leq s. \end{cases}$$

Well-posedness follows from [13, Theorem 1.3, p.1396] (see [14] for the well-posedness for any $s > -1$),

Theorem 1.2.2 *For any $s \geq 0$, for any $T, L > 0$ and for any s-compatible $(\tilde{y}, u, v, w) \in H^s(0,L) \times H^{\mu_1(s)}(0,T) \times H^{\mu_1(s)}(0,T) \times H^{\mu_2(s)}(0,T)$, (1.1.2)-(1.1.6) is*

well-posed in

$$C([0,T];H^s(0,L))\bigcap L^2([0,T];H^{s+1}(0,L))$$

1.2.2 Regular controls

Consider the linear KdV equation defined on the real line

$$\begin{cases} z_t + z_x + z_{xxx} = 0, & (x,t) \in \mathbb{R}^2 \\ z(x,0) = z_0(x), & x \in \mathbb{R}. \end{cases} \tag{1.2.4}$$

The initial value control problem was solved for (1.2.4) in [135]

Theorem 1.2.3 ([135], Theorem 3.1, p.554) *Let $s \geq 0$ and $T > 0$ be given. There exists a bounded linear operator $G : H^s(0,L) \times H^s(0,L) \to H^s(\mathbb{R})$ such that for any $y_0, y_T \in H^s(0,L)$ if one chooses $z_0 = G(y_0, y_T) \in H^s(\mathbb{R})$, then the corresponding solution z of (1.2.4) satisfies*

$$z(x,0) = y_0(x), \qquad z(x,T) = y_T(x),$$

on the interval $(0,L)$ and

$$\|z_0\|_{H^s(\mathbb{R})} \leq c\left(\|y_0\|_{H^s(0,L)} + \|y_T\|_{H^s(0,L)}\right) \tag{1.2.5}$$

where $c > 0$ is independant of y_0 and y_T.

In fact, it was proven that the solution z constructed in Theorem 1.2.3 is $C^\infty(\mathbb{R} \times (0,T))$.

A direct corollary of Theorem 1.2.3 is the exact controllability of the linear KdV equation by using the trace of z as the boundary controls

$$\begin{cases} y_t + y_x + y_{xxx} = 0, & x \in [0,L], \quad t \in [0,T], \\ y(0,t) = u(t), y(L,t) = v(t) & t \in [0,T], \\ y_x(L,t) = w(t) & t \in [0,T], \\ y(x,0) = y_0(x), & x \in [0,L], \end{cases} \tag{1.2.6}$$

The corollary that we state here is slightly different than Corollary 3.4 in [135, p. 559]

Corollary 1.2.1 *Let $s \geq 0$ and $T > 0$ be given. For any $y_0, y_T \in H^s(0,L)$, there exists $(u,v,w) \in H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T)$, depending linearly on y_0, y_T, such that (1.2.6) has a solution*

$$y \in C([0,T];H^s(0,L)) \cap L^2((0,T);H^{s+1}(0,L))$$

satisfying

$$y(x,0) = y_0(x), \qquad y(x,T) = y_T(x),$$

in the interval $(0,L)$. *Moreover, there exists* $C > 0$, *independant of* y_0 *and* y_T *such that*

$$\|u\|^2_{H^{\frac{s+1}{3}}(0,T)} + \|v\|^2_{H^{\frac{s+1}{3}}(0,T)} + \|w\|^2_{H^{\frac{s}{3}}(0,T)} \leq C\left(\|y_0\|^2_{H^s(0,L)} + \|y_T\|^2_{H^s(0,L)}\right) \tag{1.2.7}$$

The operator G constructed in the proof of Theorem 1.2.3 relies on the extension of the initial data y_0 of (1.2.6), defined on $[0,L]$, to the initial of (1.2.4), defined on $\mathbb{R}$, and of compact support. Since there exists infinitely many such extensions, the uniqueness of the controls of (1.2.6) is not guaranteed. The linearity of the controls (u,v,w) with respect to y_0, y_T stated in Corollary 1.2.1 is obtained either by always choosing the same extension of the initial datas, either by considering the controls (u,v,w) of minimal $H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T)$-norm since the projection is a linear operator . Moreover, the stability estimate (1.2.7) is obtained, thanks to the fact that the initial data z_0 constructed in the proof of Theorem 1.2.3 is compactly supported, from a sharp Kato smoothing effect of (1.2.4) (see for instance [84]) and from (1.2.5). The regularity of the controls was stated in [118]. Hence, there exists a linear continuous mapping from the initial data to the controls in the appropriate spaces.

The last result needed to prove Corollary 1.1.1 is the following, which is the generalization of [116, Proposition 4.1]

Proposition 1.2.1 *Let* $s \geq 0$. *Let* $y \in L^2((0,T);H^{s+1}(0,L))$. *Then,* $yy_x \in L^1((0,T);H^s(0,L))$ *and the map* $y \in L^2((0,T);H^{s+1}(0,L)) \mapsto yy_x \in L^1((0,T);H^s(0,L))$ *is continuous. Moreover, there exists* $K > 0$ *such that, if* $(y,z) \in L^2((0,T);H^{s+1}(0,L))^2$, *then*

$$\begin{aligned}\|yy_x - zz_x\|_{L^1((0,T);H^s(0,L))} \leq \quad & K\left(\|y\|_{L^2((0,T);H^{s+1}(0,L))} + \|z\|_{L^2((0,T);H^{s+1}(0,L))}\right) \\ & \cdot\|y-z\|_{L^2((0,T);H^{s+1}(0,L))}.\end{aligned} \tag{1.2.8}$$

Proof :

First, consider the case $n \in \mathbb{N} \cup \{0\}$. Let $y, z \in L^2((0,T);H^{n+1}(0,L))$. Using the Sobolev embedding of $H^{n+1}(0,L)$ in $L^\infty(0,L)$, we have

$$\begin{aligned}\|yy_x - zz_x\|_{L^1((0,T);H^n(0,L))} &\leq \int_0^T \|(y-z)y_x\|_{H^n(0,L)} + \|z(y_x - z_x)\|_{H^n(0,L)}\,\mathrm{dt} \\ &\leq C\int_0^T \|y-z\|_{L^\infty(0,L)}\|y_x\|_{H^n(0,L)} + \|y-z\|_{H^{n+1}(0,L)}\|y\|_{H^n(0,L)} \\ &\quad + \|z\|_{L^\infty(0,L)}\|y_x - z_x\|_{H^n(0,L)} + \|z\|_{H^{n+1}(0,L)}\|y-z\|_{H^n(0,L)}\,\mathrm{dt} \\ &\leq K\|y-z\|_{L^2((0,T);H^{n+1}(0,L))} \cdot \left(\|y\|_{L^2((0,T);H^{n+1}(0,L))} + \|z\|_{L^2((0,T);H^{n+1}(0,L))}\right)\end{aligned}$$

The result for $s \geq 0$ follows by interpolation ([124]).

$\square$

We obtain Corollary 1.1.1 by proving that the nonlinear equation is locally controllable with the Banach Fixed Point Theorem, the continuity of the controls with respect of the initial data being preserved by the continuity of the linear operators considered in the argument. The Banach Fixed Point Theorem argument to obtain the local controllability of a nonlinear equation from the controllability of the linearized equation is classical (see [42]).

Proof :

Let $y_0, y_T \in H^s(0,L)$ such that $\|y_0\|_{H^s(0,L)} \leq r$ and $\|y_T\|_{H^s(0,L)} \leq r$ with $r > 0$ to be chosen later on. Consider y^1, y^2, y^3 the solutions of the following problems

$$\begin{cases} y^1_t + y^1_x + y^1_{xxx} = 0, & x \in [0,L], \quad t \in [0,T], \\ y^1(0,t) = 0, & t \in [0,T], \\ y^1(L,t) = 0, & t \in [0,T], \\ y^1_x(L,t) = 0, & t \in [0,T], \\ y^1(x,0) = y_0(x), & x \in [0,L], \end{cases} \tag{1.2.9}$$

$$\begin{cases} y^2_t + y^2_x + y^2_{xxx} = 0, & x \in [0,L], \quad t \in [0,T], \\ y^2(0,t) = u(t), & t \in [0,T], \\ y^2(L,t) = v(t), & t \in [0,T], \\ y^2_x(L,t) = w(t), & t \in [0,T], \\ y^2(x,0) = 0, & x \in [0,L], \end{cases} \tag{1.2.10}$$

$$\begin{cases} y^3_t + y^3_x + y^3_{xxx} = f, & x \in [0,L], \quad t \in [0,T], \\ y^3(0,t) = 0, & t \in [0,T], \\ y^3(L,t) = 0, & t \in [0,T], \\ y^3_x(L,t) = 0, & t \in [0,T], \\ y^3(x,0) = 0, & x \in [0,L]. \end{cases}$$

Consider the continuous maps,

$$\begin{array}{rccl} \psi_1 : & H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T) & \to & L^2((0,T); H^{s+1}(0,L)) \\ & (u,v,w) & \mapsto & y^2, \end{array}$$

and

$$\begin{array}{rccl} \psi_2 : & L^1((0,T); H^s(0,L)) & \to & L^2((0,T); H^{s+1}(0,L)) \\ & f & \mapsto & y^3, \end{array}$$

Let

$$\begin{array}{rccl} \Gamma : & H^s(0,L) & \to & H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T) \\ & y_T & \mapsto & (u,v,w), \end{array}$$

be the continuous map associating to y_T, the controls $(u,v,w) \in H^{\frac{s+1}{3}}(0,T) \times$

$H^{\frac{s+1}{3}}(0,T) \times H^{\frac{s}{3}}(0,T)$ given by Corollary 1.2.1, such that y^2, the solution of the backward equation (1.2.10) starting from $y^2(.,T) = y_T$, reaches $y^2(.,0) = 0$. Let

$$\begin{array}{rccl} F: & L^2((0,T);H^{s+1}(0,L)) & \to & L^2((0,T);H^{s+1}(0,L)) \\ & y & \mapsto & F(y), \end{array}$$

with

$$F(y) := y^1 + \psi_1 \circ \Gamma\left(y_T - y^1(.,T) + \psi_2(yy_x)(.,T)\right) + \psi_2(-yy_x).$$

The map F was constructed such that it is well-defined, continuous and that every fixed point of F is a solution of (1.1.2)-(1.1.16) satisfying $y(.,T) = y_T$. Therefore, it is sufficient to prove that, for a closed ball $B(0,R) \subset L^2((0,T);H^{s+1}(0,L))$, $F(B(0,R)) \subset B(0,R)$, and that there exists $C \in (0,1)$ such that, $\forall (y,z) \in B(0,R)^2$,

$$\|F(y) - F(z)\|_{L^2((0,T);H^{s+1}(0,L))} \leq C\|y - z\|_{L^2((0,T);H^{s+1}(0,L))},$$

to show the existence of a fixed point of F by the Banach Fixed Point Theorem.

Let K_1, K_Γ be the norms of ψ_1, Γ and let K_2, K_2' be the norm of ψ_2 in $L^2((0,T);H^{s+1}(0,L))$ and $C((0,T);H^s(0,L))$ respectively. Furthermore, let K, be the constant from Proposition 1.2.1, C_1 denotes the constant in the stability estimate for solutions of (1.2.9),

$$\|y^1\|_{L^2(0,T;H^{s+1}(0,L))} + \|y^1\|_{C(0,T;H^s(0,L))} \leq C_1\|y_0\|_{H^s(0,L)}.$$

Then, for $y \in L^2((0,T);H^{s+1}(0,L))$,

$$\begin{aligned} \|F(y)\|_{L^2((0,T);H^{s+1}(0,L))} \leq & \|y^1\|_{L^2(0,T;H^{s+1}(0,L))} + K_1K_\Gamma\left(\|y_T\|_{H^s(0,L)}\right. \\ & \left. + \|y^1(.,T)\|_{H^s(0,L)} + \|\psi_2(yy_x)(.,T)\|_{H^s(0,L)}\right) \\ & + \|\psi_2(-yy_x)\|_{L^2(0,T;H^{s+1}(0,L))} \\ \leq & C_1\|y_0\|_{H^s(0,L)} + K_1K_\Gamma\left(\|y_T\|_{H^s(0,L)}\right. \\ & \left. + \|y^1\|_{C(0,T);H^s(0,L))} + \|\psi_2(yy_x)\|_{C(0,T;H^s(0,L))}\right) \\ & + K_2\|yy_x\|_{L^1(0,T;H^s(0,L))} \\ \leq & C_1\|y_0\|_{H^s(0,L)} + K_1K_\Gamma\left(\|y_T\|_{H^s(0,L)}\right. \\ & \left. + C_1\|y_0\|_{H^s(0,L)} + K_2'\|yy_x\|_{L^1(0,T;H^s(0,L))}\right) \\ & + K_2\|yy_x\|_{L^1(0,T;H^s(0,L))} \\ \leq & C_1\|y_0\|_{H^s(0,L)} + K_1K_\Gamma\left(\|y_T\|_{H^s(0,L)}\right. \\ & \left. + C_1\|y_0\|_{H^s(0,L)} + K_2'K\|y\|^2_{L^2(0,T;H^{s+1}(0,L))}\right) \\ & + K_2K\|y\|^2_{L^2(0,T;H^{s+1}(0,L))} \\ \leq & C_1r + K_1K_\Gamma\left(r + C_1r + K_2'KR^2\right) + K_2KR^2, \end{aligned}$$

where all the constant are independent of r and R. We impose on r and R that

$$C_1 r + K_1 K_\Gamma \left(r + C_1 r + K_2' K R^2\right) + K_2 K R^2 < R. \tag{1.2.11}$$

Furthermore,

$$\begin{aligned}
\|F(y) - F(z)\|_{L^2((0,T);H^{s+1}(0,L))} =& \|\psi_1 \circ \Gamma\left(\psi_2(yy_x - zz_x)(.,T)\right) \\
&+ \psi_2(-yy_x + zz_x)\|_{L^2((0,T);H^{s+1}(0,L))} \\
\leq& K_1 K_\Gamma \|\psi_2(yy_x - zz_x)(.,T)\|_{H^s(0,L)} \\
&+ \|\psi_2(-yy_x + zz_x)\|_{L^2((0,T);H^{s+1}(0,L))} \\
\leq& K_1 K_\Gamma \|\psi_2(yy_x - zz_x)\|_{C(0,T;H^s(0,L))} \\
&+ K_2 \| - yy_x + zz_x\|_{L^1((0,T);H^s(0,L))} \\
\leq& K_1 K_\Gamma K_2' \|yy_x - zz_x\|_{L^1((0,T);H^s(0,L))} \\
&+ K_2 \| - yy_x + zz_x\|_{L^1((0,T);H^2(0,L))} \\
\leq& 2R(K_1 K_\Gamma K_2' K + K_2 K) \\
&\cdot \|y - z\|_{L^2((0,T);H^{s+1}(0,L))}.
\end{aligned}$$

We obtain that F is a contraction by choosing $R > 0$ small enough so that

$$2R(K_1 K_\Gamma K_2' K + K_2 K) < 1.$$

By taking into account (1.2.11), we then choose

$$r = \frac{R}{2(C_1 + K_1 K_\Gamma + K_1 K_\Gamma C_1)}.$$

Thus there exists a fixed point of F.

□

1.3 Lagrangian controllability

First, let us prove Theorem 1.1.1 assuming Theorem 1.1.2.

Proof :

Let $L, T > 0$, $(\epsilon_1, \epsilon_2) \in (0, T/2)^2$ and $\gamma > 0$. For every $(a, b) \in \mathbb{R}^2$, such that $a < 0 < L < b$ and $b - a > L + \gamma$, let us denote by $\widetilde{y_0}$ an extension of y_0 in $H^2(a, b)$ with homogeneous Dirichlet boundary conditions. It is well-known that this extension can be chosen so that there exists $C(a, b) > 0$ such that

$$\|\widetilde{y_0}\|_{H^2(a,b)} \leq C(a, b)\|y_0\|_{H^2(0,L)}. \tag{1.3.1}$$

By the classical stability estimate of solutions of (1.1.2)-(1.1.6), by Corollary 1.1.1 and the Sobolev embedding of $H^2(a,b)$ in $L^\infty(a,b)$, there exists a solution y solution of (1.1.2)-(1.1.6) starting from 0 to y_0 such that there exists $C_2 > 0$ such that

$$\|y\|_{L^\infty((a,b)\times(0,T))} \leq C_2 \|y_0\|_{H^2(0,L)}. \tag{1.3.2}$$

Therefore, let y^2 denotes the solution given by Theorem 1.1.2 such that $\|y^2(.,\epsilon_1)\|_{H^2(0,L)}$ and $\|y^2(.,\epsilon_2)\|_{H^2(0,L)}$ are smaller than the tolerance δ stated in Corollary 1.1.1, $C_2\|y^2(.,\epsilon_i)\|_{H_2(0,L)} \leq \gamma/2$, for $i = 1,2$, and such that

$$\Phi(x,T) \geq L + \delta, \forall x \in [0,L].$$

Thus, Corollary 1.1.1 implies that there exists $(u^1, v^1, w^1) \in H^1(0,T) \times H^1(0,T) \times H^{\frac{2}{3}}(0,T)$ such that the solution $y^1 \in C([0,\epsilon_1]; H^2(a,b))$ of

$$\begin{cases} y^1_t + y^1_x + y^1_{xxx} + y^1 y^1_x = 0, & x \in [a,b] \quad t \in [0,\epsilon_1], \\ y^1(a,t) = u^1(t), & t \in [0,\epsilon_1], \\ y^1(b,t) = v^1(t), & t \in [0,\epsilon_1], \\ y^1_x(b,t) = w^1(t), & t \in [0,\epsilon_1], \\ y^1(x,0) = 0, & x \in [a,b], \end{cases}$$

satisfies $y^1(.,\epsilon_1) = \widetilde{y^2}(.,\epsilon_1)$ and that there exists $(u^3, v^3, w^3) \in H^1(0,T) \times H^1(0,T) \times H^{\frac{2}{3}}(0,T)$ such that $y^3 \in C([T-\epsilon_2, T]; H^2(a,b))$, solution of

$$\begin{cases} y^3_t + y^3_x + y^3_{xxx} + y^3 y^3_x = 0, & x \in [a,b] \quad t \in [T-\epsilon_2, T], \\ y^3(a,t) = u^3(t), & t \in [T-\epsilon_2, T], \\ y^3(b,t) = v^3(t), & t \in [T-\epsilon_2, T], \\ y^3_x(b,t) = w^3(t), & t \in [T-\epsilon_2, T], \\ y^3(x, T-\epsilon_2) = \widetilde{y^2}(.,\epsilon_2), & x \in [a,b], \end{cases}$$

satisfies $y^3(.,T) = 0$.

Let

$$u(t) := \begin{cases} u^1(t), & t \in (0,\epsilon_1), \\ y^2(0,t), & t \in (\epsilon_1, T-\epsilon_2), \\ u^3(t), & t \in (T-\epsilon_2, T), \end{cases}$$

$$v(t) := \begin{cases} v^1(t), & t \in (0,\epsilon_1), \\ y^2(L,t), & t \in (\epsilon_1, T-\epsilon_2), \\ v^3(t), & t \in (T-\epsilon_2, T), \end{cases}$$

$$w(t) := \begin{cases} w^1(t), & t \in (0,\epsilon_1), \\ y^2_x(L,t), & t \in (\epsilon_1, T-\epsilon_2), \\ w^3(t), & t \in (T-\epsilon_2, T). \end{cases}$$

Then, y solution of

$$\begin{cases} y_t + y_x + y_{xxx} + yy_x = 0, & x \in [0, L] \quad t \in [0, T], \\ y(0, t) = u(t), & t \in [0, T], \\ y(L, t) = v(t), & t \in [0, T], \\ y_x(L, t) = w(t), & t \in [0, T], \\ y(x, 0) = 0, & x \in [0, L], \end{cases}$$

belongs to $C([0, T]; H^2(0, L))$ and, by construction of y, we obtain Theorem 1.1.1. □

We now conclude with the proof of Theorem 1.1.1

Proof :

Let $T, L, \delta > 0$ and $(\epsilon_1, \epsilon_2) \in (0, T/2)^2$. Let $\alpha_1 > 0$ and $\epsilon > 0$. We define

$$N := \left\lceil 4L\alpha_1^2 / \ln\left(\sqrt{2\alpha_1}\left(1 + \sqrt{1 - \frac{1}{2\alpha_1}}\right)\right)\right\rceil$$

where $\lceil x \rceil$ is the ceiling function. Let

$$0 < \alpha_N < ... < \alpha_1. \tag{1.3.3}$$

where

$$\alpha_1 - \alpha_N = \epsilon, \tag{1.3.4}$$

and, for $i = 1, ..., N$,

$$\begin{aligned} s_i :=& L - (\alpha_1 - \epsilon)^2(T - \epsilon_2) + \frac{N - i + 1}{N + 1}\left(-\alpha_1^2\epsilon_1 - L + (\alpha_1 - \epsilon)^2(T - \epsilon_2)\right) \\ =& -\alpha_1^2\epsilon_1\frac{N + 1 - i}{N + 1} + \frac{i}{N + 1}\left(L - (\alpha_1 - \epsilon)^2(T - \epsilon_2)\right) \end{aligned} \tag{1.3.5}$$

Finally, let $G := (\ln F)_{xx}$ and

$$F = 1 + \sum_{n=1}^{N}\sum_{{}_N C_n} a(i_1, ..., i_n)\prod_{j=1}^{n} f_{i_j},$$

where

$$f_i(x, t) = \exp(-\alpha_i(x - s_i) + \alpha_i^3 t).$$

The function G corresponds to N solitons, ordered, for $t \in [0, T]$, from the right with the fastest soliton, associated to α_1, to left with the slowest, associated to α_N. Moreover, they were constructed so that, for α_1 large, the N solitons are located to the left of the interval $[0, L]$ for time $t \in [0, \epsilon_1]$, that they pass inside the domain during

the time interval $(\epsilon_1, T-\epsilon_2)$ and are located to the right of $[0, L]$ for $t \in (T-\epsilon_2, T)$. We now prove that, for α_1 sufficiently large, the function G fulfills the requirements of Theorem 1.1.2.

One notes from (1.3.5) that, if

$$(\alpha_1 - \epsilon)^2 > L/(T-\epsilon_2), \tag{1.3.6}$$

a natural condition on the speed of the slowest soliton for it to cross the domain $[0, L]$ in the required time, then,

$$s_N < \ldots < s_1 < 0.$$

Assumption (1.3.6) on α_1 is therefore made for the rest of the proof.

Let us show that there exists α_1 sufficiently large such that $\|G(.,t)\|^2_{H^2} \leq \delta, t \in [0, \epsilon_1]$. The expression of G and its first two derivatives takes the form,

$$G = \frac{FF_{xx} - (F_x)^2}{F^2}, \tag{1.3.7}$$

$$G_x = \frac{F^2F_{xxx} - 3FF_xF_{xx} + 2(F_x)^3}{F^3}, \tag{1.3.8}$$

$$G_{xx} = \frac{-4F^2F_xF_{xxx} + F^3F_{xxxx} + 12F(F_x)^2F_{xx} - 3(FF_{xx})^2 - 6(F_x)^4}{F^4}. \tag{1.3.9}$$

Thus, we have the bounds, thanks to the fact that $F(x,t) \geq 1$, $\forall (x,t) \in \mathbb{R}^2$,

$$\begin{aligned}
|G| &\leq |F||F_{xx}| + |F_x|^2, \\
|G_x| &\leq |F|^2|F_{xxx}| + 3|F||F_x||F_{xx}| + 2|F_x|^3, \\
|G_{xx}| &\leq 4|F|^2|F_x||F_{xxx}| + |F|^3|F_{xxxx}| + 12|F||F_x|^2|F_{xx}| + 3|F|^2|F_{xx}|^2 + 6|F_x|^4.
\end{aligned}$$

By noting that the derivatives of F are given, for $k \in \mathbb{N}$, by

$$\frac{d^kF}{dx^k} = \sum_{n=1}^{N} \sum_{{}_NC_n} (-1)^k (\alpha_{i_1} + ... + \alpha_{i_n})^k a(i_1, ..., i_n) \prod_{j=1}^{n} f_{i_j},$$

we see, taking into account (1.3.3), (1.3.5), that for $t \in [0, \epsilon_1]$,

$$\left\| \frac{d^kF}{dx^k}(x,t) \right\|_{L^\infty(0,L)} = \left| \frac{d^kF}{dx^k}(0, \epsilon_1) \right|.$$

Therefore, taking into account that $a(i_1, ..., i_n)$ are at least bounded by (1.3.4), if

$$s_i + \alpha_i^2 \epsilon_1 < 0, \tag{1.3.10}$$

hold, that is a condition insuring that all the solitons are located to the left of $[0, L]$,

then there exists α_1 sufficiently large such that $\|G(.,t)\|^2_{H^2(0,L)} \leq \delta, t \in [0,\epsilon_1]$. Or,

$$\begin{aligned}\alpha_i^2\epsilon_1 + s_i <&\alpha_1^2\epsilon_1 + s_i\\ =&\frac{i}{N+1}(L-(\alpha_1-\epsilon)^2(T-\epsilon_2)+\alpha_1^2\epsilon_1)\\ =&\frac{i}{N+1}(L-\alpha_1^2(T-\epsilon_1-\epsilon_2)+2\alpha_1\epsilon(T-\epsilon_2)-\epsilon^2(T-\epsilon_2))\\ <&0,\end{aligned}$$

the last line holding for sufficiently large α_1. Thus, (1.3.10), a condition insuring that the N solitons are located in $x<0$ for $t \in [0,\epsilon_1]$, holds.

We now prove that for α_1 sufficiently large, $\|G(.,t)\|^2_{H^2} \leq \delta, t \in [0,\epsilon_1]$. From (1.3.7)-(1.3.9), one sees that, when the f_i's are large, the leading term $(f_1\cdots f_N)^{j+2}$ of $G^{(j)}$, for $j=0,1,2$, is only present to the denominator. Therefore, one has that if

$$s_i - L + \alpha_i^2(T-\epsilon_2) > 0, \tag{1.3.11}$$

then there exists α_1 such that $\|G(.,t)\|^2_{H^2(0,L)} \leq \delta, t \in [T-\epsilon_2, T]$. Or,

$$\begin{aligned}s_i - L + \alpha_i^2(T-\epsilon_2) >&s_i - L + \alpha_N^2(T-\epsilon_2)\\ =&-\alpha_1^2\epsilon_1\frac{N+1-i}{N+1}+\frac{i}{N+1}\left(L-(\alpha_1-\epsilon)^2(T-\epsilon_2)\right)\\ &-L+(\alpha_1-\epsilon)^2(T-\epsilon_2)\\ =&-\alpha_1^2\epsilon_1+(\alpha_1-\epsilon)^2(T-\epsilon_2)-L\\ &+\frac{i}{N+1}(\alpha_1^2\epsilon_1+L-(\alpha_1-\epsilon)^2(T-\epsilon_2))\\ =&\alpha_1^2(T-\epsilon_1-\epsilon_2)-2\epsilon\alpha_1(T-\epsilon_2)+\epsilon^2-L+\frac{i}{N+1}\\ &\cdot(-\alpha_1^2(T-\epsilon_1-\epsilon_2)+2\epsilon\alpha_1(T-\epsilon_2)-\epsilon^2(T-\epsilon_2)+L)\\ >&0,\end{aligned}$$

holds for sufficiently large α_1. One remarks that (1.3.11) represents the fact that the N solitons are located in $x>L$ for $t \in [T-\epsilon_2, T]$. Moreover, combining (1.3.10) and (1.3.11) asks for $\alpha_i^2(T-\epsilon_1-\epsilon_2) > L$, a natural condition on the travelling speed of the N solitons. Moreover, we have

$$L - \alpha_N^2(T-\epsilon_2) < s_N < \ldots < s_1 < -\alpha_1^2\epsilon_1$$

In order to prove that the flow Φ associated to G satisfies $\Phi(x,T) \geq L, (x,t) \in (0,L)\times(T-\epsilon_2,T)$, we use a rough lower bound on each solitons of G : a characteristic function of the same amplitude and the same width. Prior this estimate, let us first obtain rigorously that G is greater than N solitons. It is important to note here

that, from the definition of G, we have that every coefficient in front of $f_{i_1}\cdots f_{i_m}$, $1\leq m\leq 2N$ with, possibly, twice repeated indexes i_j, are positive.

For a fixed time t in $[0,T]$, in the neighbourhood of the first soliton $x\in((3s_1+s_2)/4+\alpha_1^2t,-\alpha_1^2\epsilon_1+\alpha_1^2t)$, we have,

$$\begin{aligned}
G(x,t)\geq&\frac{\alpha_1^2f_1(x,t)}{\Big(1+\sum_{n=1}^{N}\sum_{{}_NC_n}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}(x,t)\Big)^2}\\
=&\frac{\alpha_1^2f_1(x,t)}{\Big(1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}(x,t)+f_1(x,t)\Big)^2}\\
\geq&\frac{\alpha_1^2f_1(x,t)}{\Big(1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}\Big(\frac{3s_1+s_2}{4}+\alpha_1^2t,t\Big)+f_1(x,t)\Big)^2}\\
=&\frac{1}{\Big(1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}\Big(\frac{3s_1+s_2}{4}+\alpha_1^2t,t\Big)\Big)^2}\\
&\cdot\frac{\alpha_1^2f_1(x,t)}{\Big(1+\Big(\frac{1}{1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}((3s_1+s_2)/4+\alpha_1^2t,t)}\Big)f_1(x,t)\Big)^2}\\
=&\frac{1}{\Big(1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}\Big(\frac{3s_1+s_2}{4}+\alpha_1^2t,t\Big)\Big)}\\
&\cdot\frac{\alpha_1^2\Big(\frac{1}{1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}((3s_1+s_2)/4+\alpha_1^2t,t)}\Big)f_1(x,t)}{\Big(1+\Big(\frac{1}{1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}((3s_1+s_2)/4+\alpha_1^2t,t)}\Big)f_1(x,t)\Big)^2}.
\end{aligned}$$

Let

$$A_1(t):=\frac{1}{\Big(1+\sum_{n=1}^{N}\sum_{\substack{{}_NC_n\\ i_1\neq 1}}a(i_1,...,i_n)\prod_{j=1}^{n}f_{i_j}\Big(\frac{3s_1+s_2}{4}+\alpha_1^2t,t\Big)\Big)}.$$

Then, we obtain, in the region $x \in ((3s_1 + s_2)/4 + \alpha_1^2 t, -\alpha_1^2 \epsilon_1 + \alpha_1^2 t)$,

$$G(x,t) \geq \frac{\alpha_1^2 A_1(t)}{4} \operatorname{sech}^2 \left(\frac{-\alpha_1(x - \sigma_1(t) - \alpha_1^2 t)}{2} \right), \tag{1.3.12}$$

a soliton of amplitude $\alpha_1^2 A_1(t)/4$ of phase $\sigma_1(t) := s_1 + \frac{1}{\alpha_1} \ln(A_1(t))$.

For fixed t in $[0,T]$, we consider, for the k-th soliton, $2 \leq k \leq N$, the neighbourhood

$$(\xi_k^-(t), \xi_k^+(t)) = \begin{cases} \left(\dfrac{2s_k + s_{k+1}}{3} + \alpha_k^2 t, \dfrac{2s_k + s_{k-1}}{3} + \alpha_k^2 t \right), & 2 \leq k \leq N-1, \\ \left(\dfrac{4s_N - s_{N-1}}{3} + \alpha_N^2 t, \dfrac{2s_N + s_{N-1}}{3} + \alpha_N^2 t \right), & k = N. \end{cases}$$

One notices, from the definition of s_N, that $L - \alpha_N^2 (T - \epsilon_2) < (4s_N - s_{N-1})/3 < s_N$.

Let

$$A_k(t) = \left(\frac{1}{1 + \dfrac{1 + \sum_{n=1}^N \sum_{{}_NC_{n,n\neq k-1,k}} a(i_1,\dots,i_n) \prod_{\substack{j=1 \\ i_j \leq k-1}}^n f_{i_n}(\xi_k^+(t),t) \prod_{\substack{j=1 \\ i_j > k-1}}^n f_{i_n}(\xi_k^-(t),t)}{a(1,\dots,k-1) f_1 \cdots f_{k-1}(\xi_k^+(t),t)}} \right),$$

where ${}_NC_{n,n\neq k-1,k}$ denotes the usual ${}_NC_n$ but excluding the case where $i_1, \dots, i_n$ equal $1, \dots, k-1$ or $1, \dots, k$. Then, the same steps as in the case of the first soliton yield

$$\begin{aligned} G(x,t) \geq & \frac{\alpha_k^2 a(1,\dots,k-1) a(1,\dots,k) f_1^2 \cdots f_{k-1}^2 f_k(x,t)}{\left(1 + \sum_{n=1}^N \sum_{{}_NC_n} a(i_1,\dots,i_n) \prod_{j=1}^n f_{i_j}(x,t)\right)^2} \\ = & \frac{\alpha_k^2 \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k(x,t)}{\left(\dfrac{1 + \sum_{n=1}^N \sum_{\substack{{}_NC_n \\ n\neq k-1,k}} a(i_1,\dots,i_n) \prod_{j=1}^n f_{i_j}(x,t)}{a(1,\dots,k-1) f_1 \cdots f_{k-1}(x,t)} + 1 + \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k(x,t) \right)^2} \\ \geq & \frac{\alpha_k^2 \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k}{\left(A_k(t)^{-1} + \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k \right)^2} \\ = & A_k(t) \frac{\alpha_k^2 A_k(t) \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k}{\left(1 + A_k(t) \frac{a(1,\dots,k)}{a(1,\dots,k-1)} f_k\right)^2}, \end{aligned}$$

Thus, for $x \in (\xi_k^-(t), \xi_k^+(t))$, we have

$$G(x,t) \geq \frac{\alpha_k^2 A_k(t)}{4} \operatorname{sech}^2\left(\frac{-\alpha_k(x-\sigma_k(t)-\alpha_k^2 t)}{2}\right), \tag{1.3.13}$$

a soliton of amplitude $\alpha_k^2 A_k(t)/4$ and of phase $\sigma_k := s_k + \frac{1}{\alpha_k}\ln(\frac{A_k(t)a(1,\ldots,k)}{a(1,\ldots,k-1)})$.

We prove the following for $A_k(t)$

Lemma 1.3.1 *For $1 \leq k \leq N$, we have, for $t \in [0,T]$, $A_k(t) \to 1$ as $\alpha_1 \to \infty$.*

Proof :

We first consider $k = 1$. Let us remark that the term f_1 alone was removed from $A_1(t)$. Thus, let us show that $f_1 f_2(\frac{2s_1+s_2}{3} + \alpha_1^2 t, t)$ converges to zero as α_1 tends to infinity, since it is the biggest term in the sum.

$$\begin{aligned} f_1 f_2\left(\frac{2s_1+s_2}{3} + \alpha_1^2 t, t\right) &= \exp\left(-\alpha_1\left(\frac{2s_1+s_2}{3} + \alpha_1^2 t - s_1 - \alpha_1^2 t\right)\right) \\ &\quad \cdot \exp\left(-\alpha_2\left(\frac{2s_1+s_2}{3} + \alpha_1^2 t - s_2 - \alpha_2^2 t\right)\right) \\ &= \exp\left((2\alpha_2 - \alpha_1)\left(\frac{s_2 - s_1}{3}\right) - \alpha_2(\alpha_1^2 - \alpha_2^2)t\right). \end{aligned}$$

Thus, since $s_2 - s_1 < 0$ and, for large α_1, $\alpha_1^2 - \alpha_2^2 > 0$ and $2\alpha_2 - \alpha_1 > 0$, since $\alpha_1 - \alpha_N = \epsilon$, then the last expression converges to zero as α_1 tends to infinity, and so do all the other terms in the sum of $A_1(t)$.

We consider next the case $2 \leq k \leq N-1$. We stress here that, since $f_1 \cdots f_{k-1}$ and $f_1 \cdots f_k$ were removed from the sum in $A_k(t)$, no terms of the form $f_{i_1} \cdots f_{i_n}$, where $(i_1, \ldots, i_n) \in \{1, ..., k-1\}^n$, appears at the numerator of

$$\frac{1 + \sum_{n=1}^N \sum_{{}_NC_{n,n\neq k-1,k}} a(i_1,...,i_n) \prod_{\substack{j=1 \\ i_j \leq k-1}}^n f_{i_n}(\xi_k^+(t),t) \prod_{\substack{j=1 \\ i_j > k-1}}^n f_{i_n}(\xi_k^-(t),t)}{a(1,...,k-1) f_1 \cdots f_{k-1}(\xi_k^+(t),t)}.$$

Therefore, it is sufficient to show that the terms of the form $f_k(\xi_k^-(t),t) f_{k-1}^{-1}(\xi_k^+(t),t)$ and $f_k(\xi_k^-(t),t) f_{k+1}(\xi_k^-(t),t)$ tend to zero, as all the other terms will converge to zero as well. We have,

$$\begin{aligned} f_k(\xi_k^-(t),t) f_{k-1}^{-1}(\xi_k^+(t),t) &= \exp\left(-\alpha_k\left(\frac{-s_k + s_{k+1}}{3}\right)\right. \\ &\quad \left. + \alpha_{k-1}\left(\frac{2s_k - 2s_{k-1}}{3}\right) + (\alpha_k^2 - \alpha_{k-1}^2)t\right) \\ &= \exp\left((\alpha_{k-1} - \alpha_k)\left(\frac{-s_k + s_{k+1}}{3}\right)\right. \end{aligned}$$

$$+\alpha_{k-1}\left(\frac{3s_k-2s_{k-1}-s_{k+1}}{3}+(\alpha_k^2-\alpha_{k-1}^2)t\right)\Bigg).$$

Or,

$$\begin{aligned}3s_k-2s_{k-1}-s_{k+1}=&\frac{3k-2k+2-k-1}{3}\left(-\alpha_1^2(T-\epsilon_1-\epsilon_2)+2\epsilon\alpha_1(T-\epsilon_2)\right.\\&\left.-\epsilon^2(T-\epsilon_2)+L\right)\\=&\frac{1}{3}\left(-\alpha_1^2(T-\epsilon_1-\epsilon_2)+2\epsilon\alpha_1(T-\epsilon_2)-\epsilon^2(T-\epsilon_2)+L\right),\end{aligned}$$

which is negative for α_1 sufficiently large, as well as $\alpha_k^2-\alpha_{k-1}^2$ and $s_{k+1}-s_k$, while $\alpha_{k-1}-\alpha_k>0$. Therefore, $f_k(\xi_k^-(t),t)f_{k-1}^{-1}(\xi_k^+(t),t)$ tends to zero as α_1 tends to infinity. Moreover,

$$\begin{aligned}f_k(\xi_k^-(t),t)f_{k+1}(\xi_k^-(t),t)=&\exp\left(-\alpha_k\left(\frac{s_{k+1}-s_k}{3}\right)\right.\\&\left.-\alpha_{k+1}\left(\frac{2s_k-2s_{k+1}}{3}\right)+(\alpha_k^2-\alpha_{k+1}^2)t\right)\\=&\exp\left((\alpha_k-2\alpha_{k+1})\left(\frac{s_k-s_{k+1}}{3}\right)-\alpha_{k+1}(\alpha_k^2-\alpha_{k+1}^2)t\right),\end{aligned}$$

which tends to zero as α_1 tends to infinity.

Finally, let us consider the case where $k=N$. For the same reasons as in the case $2\le k\le N-1$, it is sufficient to consider the case $f_N(\xi_k^-(t),t)f_{N-1}^{-1}(\xi_k^+(t),t)$. We have,

$$\begin{aligned}f_N(\xi_k^-(t),t)f_{N-1}^{-1}(\xi_k^+(t),t)=&\exp\left(-\alpha_N\left(\frac{s_N-s_{N-1}}{3}\right)\right.\\&\left.+\alpha_{N-1}\left(4\frac{s_N-s_{N-1}}{3}+(\alpha_N^2-\alpha_{N-1}^2)t\right)\right)\\=&\exp\left((4\alpha_{N-1}-\alpha_N)\left(\frac{s_N-s_{N-1}}{3}\right)\right.\\&\left.+\alpha_{N-1}(\alpha_N^2-\alpha_{N-1}^2)t\right),\end{aligned}$$

which tends to zero as α_1 tends to infinity since $s_N-s_{N-1}<0$, $4\alpha_{N-1}-\alpha_N>0$, and $\alpha_N^2-\alpha_{N-1}^2<0$.

□

We can now prove that G satisfies $\Phi(x,T)\ge L,(x,t)\in(0,L)\times(T-\epsilon_2,T)$. We obtain the rough estimate on G, from (1.3.12), (1.3.13), and from the definition of

the width of a soliton,

$$G(x,t) \geq \sum_{k=1}^{N} \frac{A_k(t)\alpha_k}{8} \mathbb{1}(x)_{[\alpha_k^2 t+\sigma_k(t)-w(\alpha_k)/2, \alpha_k^2 t+\sigma_k(t)+w(\alpha_k)/2]},$$

where $w(\alpha)$ is the width defined by (1.1.14). From Lemma 1.3.1, we suppose that α_1 is large enough so $A_k(t) \leq 1/2$, $\forall t \in [0,T]$ and $1 \leq k \leq N$. We have,

$$G(x,t) \geq \sum_{k=1}^{N} \frac{\alpha_k}{16} \mathbb{1}(x)_{[\alpha_k^2 t+\sigma_k(t)-w(\alpha_k)/2, \alpha_k^2 t+\sigma_k(t)+w(\alpha_k)/2]},$$

and, therefore,

$$\begin{aligned}
\frac{\partial \Phi}{\partial t}(x,t) =& G(\Phi(x,t),t) \\
\geq& \sum_{k=1}^{N} \frac{\alpha_k}{16} \mathbb{1}(\Phi(x,t))_{[\alpha_k^2 t+\sigma_k(t)-w(\alpha_k)/2, \alpha_k^2 t+\sigma_k(t)+w(\alpha_k)/2]}, && (1.3.14) \\
\geq& \sum_{k=1}^{N} \frac{\alpha_k}{16} \mathbb{1}(x)_{[\alpha_k^2 t+\sigma_k(t)-w(\alpha_k)/2, \alpha_k^2 t+\sigma_k(t)+w(\alpha_k)/2]}, && (1.3.15)
\end{aligned}$$

the last line coming from the fact that, since all the characteristic functions are positive, the displacement of Φ will always be on the right. Therefore, the displacement of Φ will be greater in (1.3.14) than in (1.3.15), since the characteristic functions follow, even for a brief moment, the displacement of Φ.

The displacement Φ under the flow of G is then easily estimated from (1.3.15) : $\ln\left(\sqrt{2\alpha_k}\left(1+\sqrt{1-\frac{1}{2\alpha_k}}\right)\right)/4\alpha_k^2$, that is, the height times the width divided by the speed α_k of each soliton. Consequently, each point $x \in [0,L]$ under the action of the flow of G will move of at least $\ln\left(\sqrt{2\alpha_k}\left(1+\sqrt{1-\frac{1}{2\alpha_k}}\right)\right)/4\alpha_k^2$ to the right each time a soliton passes through the domain $[0,L]$. Since

$$N = \left\lceil 4L\alpha_1^2 / \ln\left(\sqrt{2\alpha_1}\left(1+\sqrt{1-\frac{1}{2\alpha_1}}\right)\right)\right\rceil,$$

we obtain that $\Phi(x,T) \geq L, (x,t) \in (0,L) \times (T-\epsilon_2, T)$.

□

Chapitre 2

Approximation d'ordre supérieur du champ de vitesse associé à la solution de N-solitons

Contents

2.1 Introduction **107**
2.2 Derivation of the Korteweg-de Vries equation **111**
2.3 A higher velocity field for the N-soliton solution **115**
2.3.1 The N-solitons solution 115
2.3.2 Nonlaminar velocity field 117
2.3.3 Proof of Theorem 2.1.1 121
2.4 Numerical analysis of the particles trajectory **125**
2.5 Conclusion **127**

2.1 Introduction

The Korteweg-de Vries (KdV) equation,

$$\frac{1}{\sqrt{gh_0}}\partial_t\eta + \partial_x\eta + \frac{h_0^2}{6}\partial_x^3\eta + \frac{3}{2h_0}\eta\partial_x\eta = 0, \quad x \in \mathbb{R},\ t \in \mathbb{R}, \tag{2.1.1}$$

corresponds to the first order approximation of waves propagating from left to right of the full governing equations for homogeneous, non-viscous and irrotational fluid under the shallow water regime. The shallow water regime refers to the smallness of the amplitude parameter $\epsilon := a/h_0$ and shallowness parameter $\delta := h_0/\lambda$. Here, h_0 denotes the depth of the fluid, a the typical amplitude of the waves, λ the typical length of the waves and g the gravitational constant. The Cartesian coordinates (x, y)

are chosen such that the body of fluid is represented by the two-dimensional domain bounded by the flat bottom $y = 0$ and the fluid free surface $y = h_0 + \eta(x, t)$.

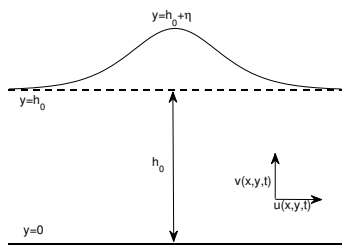

FIGURE 1 – A soliton solution of (2.1.1) represented in the Cartesian coordinates.

The first work on the Korteweg-de Vries equation are due to John Scott Russell. Following his observation of a solitary wave on the Union Canal in 1834, a travelling wave propagating at constant speed without losing its shape, Russell published a report of his experiments and observations in 1844 [119]. The Korteweg-de Vries equation was later derived independently by Boussinesq in 1877 [23] and by Korteweg and de Vries in 1895 [86].

The asymptotic behaviour of solutions of (2.1.1) arising from smooth and fast decaying initial disturbance obtained by the *Inverse Scattering Transform* (Gardner, Greene, Kruskal and Miura, ([64, 1967], [65, 1974])), confirmed the numerical experiments of Zabusky and Kruskal ([134, 1965]) : a finite number of solitons are travelling to the right while an oscillating and decaying wavetrain disperses to the left.

A soliton of (2.1.1), the approximation in the KdV regime of a solitary wave, is given by

$$\eta(x,t) = a\,\mathrm{sech}^2\left(\sqrt{\frac{3a}{4h_0^3}}\left(x - s - \sqrt{gh_0}\left(1 + \frac{a}{2h_0}\right)t\right)\right), \tag{2.1.2}$$

where $a > 0$, $s \in \mathbb{R}$. The following properties hold for (2.1.2)

1. The amplitude is given by a and is reached at $x = s + \sqrt{gh_0}(1 + \frac{a}{2h_0})t$;
2. The traveling speed is $\sqrt{gh_0}(1 + \frac{a}{2h_0})t$;
3. The taller the soliton, the narrower.

The behaviour of N solitons is given by the N-solitons solution, representing the KdV approximation of the interactions of N solitary waves propagating in the same direction. They behaves like N distinct solitons travelling at different constant speed except when interactions occur, in which case they are left unscathed from an interaction except for a phase shift. Moreover, if two solitons with the same order of amplitude interact, they exchange their mass and speed during the interaction.

The Korteweg-de Vries equation not only provides an approximation of the surface elevation in shallow water but admits as well an approximation of the underlying particle trajectories. To the first order approximation in the perturbation parameter, one obtains, with the divergence-free condition, the first order velocity field of (2.1.1),

$$\begin{cases} u(x,y,t) = \sqrt{\dfrac{g}{h_0}}\eta(x,t) \\ v(x,y,t) = -y\sqrt{\dfrac{g}{h_0}}\eta_x(x,t). \end{cases} \tag{2.1.3}$$

Since the KdV equation also holds to the second order, one recovers a second order approximation of the horizontal velocity field at the flat bottom ([131]),

$$u(x,0,t) = \sqrt{\frac{g}{h_0}}\left(\eta - \frac{1}{4h_0}\eta^2 + \frac{h_0^2}{3}\eta_{xx}\right). \tag{2.1.4}$$

Thus, one may study the particle trajectories, for a given velocity field, associated to solutions of KdV by solving,

$$\begin{cases} \dfrac{dX}{dt}(x,y,t) = u(X(x,y,t),Y(x,y,t),t), & t \in \mathbb{R}, \\ \dfrac{dY}{dt}(x,y,t) = v(X(x,y,t),Y(x,y,t),t), & t \in \mathbb{R}, \\ X(x,y,-\infty) = x, \quad Y(x,y,-\infty) = y, & x \in \mathbb{R},\ 0 \le y \le h_0 + \eta. \end{cases}$$

In the case of a soliton, the particle paths obtained with the first order velocity field (2.1.3) fail to capture one essential property of the particle trajectories observed experimentally for solitary waves : the higher in the fluid a particle is initially located, the larger its horizontal displacement [34]. It is a consequence of the independance on the y variable of the horizontal component of the velocity field. Since the soliton solution is considered as an accurate approximation of the solitary waves solution of the full governing equations ([37]), it is crucial to seek for higher approximations of the velocity field in order to derive better approximation of the particle trajectories. Figure 2 illustrates the experimental results of [34] for different solitary waves.

In the literature, two methods have been developed to recover a higher order approximation of the velocity field for the KdV equation. Introducing a stream function for the first order approximation and using the fact that the velocity potential and the stream function are harmonic conjugate, Constantin ([36]) and Henry ([78]) have obtained a higher order velocity field for a soliton and periodic solutions respectively. Qualitatives results on the particle paths were obtained in both articles.

A careful analysis of the derivation of the KdV equation in the second order approximation allowed Borluk and Kalisch to recover from (2.1.4) a nonlaminar velocity field underneath the surface ([16]). They obtained numerically a better path of the particles in the case of the soliton solution, the cnoidals solution and the 2-solitons

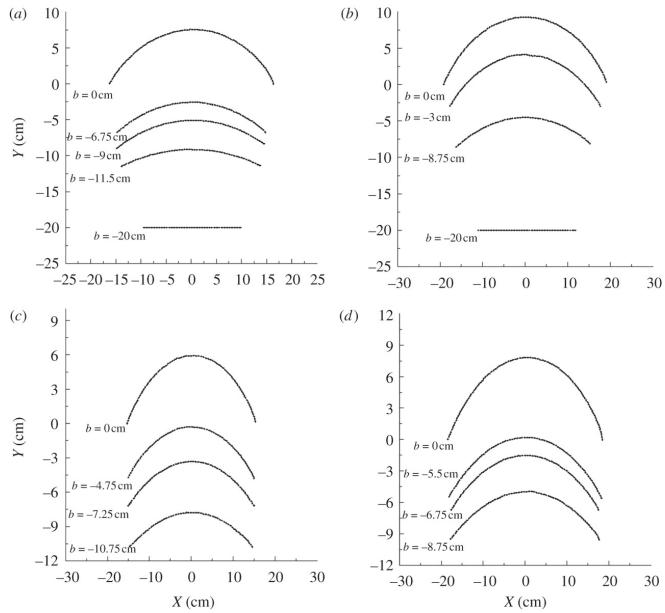

FIGURE 2 – The orbits of water particles obtained from the experimental measurements of the polystyrene beads motions at different water levels b in the four experimental wave cases (a) h_0=20cm, a=7.07cm ; (b) h_0=20cm, a=8.56cm ; (c) h_0=30cm, a=5.46cm ; (d) h_0=30cm, a=7.56cm.

solution. A similar work in the case of the Boussinesq equation had been done before in [2, 15].

The main result of this article, following the work of Constantin and Henry ([36, 78]), is the qualitative description of the particle trajectories for the N-solitons solution of KdV.

Theorem 2.1.1 *Let $N \in \mathbb{N}$ denotes the number of solitons and a_i, $1 \leq i \leq N$, denotes the amplitude parameter of each solitons. For sufficiently small amplitude parameters $\epsilon_i := a_i/h_0$, the particle trajectories described by the velocity field satisfy the following :*

(a) A particle on the flat bed move in straight line at a positive speed ;

(b) All the particles below the surface have a positive horizontal speed ;

Moreover, if a particle doesn't encounter soliton interactions, we have the following :

(c) For each such particles, there exists $t_1, \ldots, t_{2N-1}$ such that the particle move upward for,

$$t \in (-\infty, t_1) \bigcup \left(\bigcup_{k=1}^{N-1} (t_{2k}, t_{2k+1}) \right),$$

and downward for,

$$t \in \left(\bigcup_{k=1}^{N-1} (t_{2k-1}, t_{2k}) \right) \bigcup (t_{2N-1}, \infty);$$

(d) The particle's position at times $t_1, t_3, \ldots, t_{2N-1}$ follows the phase shift of the solitons.

We would like to emphasize on the nonlinear nature of this result. As it will be pointed out in Section 3, the phase shifts of the solitons are the result of the nonlinearity of the KdV equation. Therefore, one cannot obtain the right behaviour of the particles for the N-solitons solution with the superposition principle used on the 1-soliton velocity field obtained in [36]. Moreover, there exists no linear approximation of solitons.

The outline of the paper is the following. In Section 2, inspired from [36, 38], we recall the derivation of the Korteweg-de Vries to the first order in the perturbation parameter. In Section 3, we present the N-solitons solution and we use the method of [36, 78] in order to prove Theorem 2.1.1. This is done by noticing that this method is not limited to travelling waves. Finally, a numerical analysis of the particle trajectories is provided in Section 4. We show that the monotonicity in the y variable of the horizontal displacement is recovered. Moreover, the numerical investigation indicate the possible lost of the free surface kinematic boundary condition for velocity fields of KdV.

2.2 Derivation of the Korteweg-de Vries equation

The physical assumptions to derive the Korteweg-de Vries equation are the following. Consider approximately two-dimensional waves, that is, the wave profile is approximately the same for all cross sections taken with respect to the crest line. Assume irrotational waves in a homogeneous, non-viscous and incompressible fluid. Moreover, assume the impermeability of the bottom (bottom kinematic boundary condition), that the same particles form the surface (free surface kinematic boundary condition) and that the motion of the air and that of water is decoupled. Then, the

full governing equations are given by

$$\begin{cases} u_t + uu_x + vu_y = -\frac{1}{\rho}P_x, & 0 \le y \le h_0 + \eta, \\ v_t + uv_x + vv_y = \frac{1}{\rho}P_y - g, & 0 \le y \le h_0 + \eta, \\ u_x + v_y = 0, & 0 \le y \le h_0 + \eta, \\ u_y - v_x = 0, & 0 \le y \le h_0 + \eta, \\ v = \eta_t + u\eta_x, & y = h_0 + \eta, \\ P = P_0 & y = h_0 + \eta, \\ v = 0, & y = 0. \end{cases} \tag{2.2.1}$$

Let us first set the non-dimensional full governing equations. Consider the perturbation of the pressure,

$$P = P_0 + \rho g(h_0 - y) + \rho g h_0 p,$$

which express the change of the pressure as the waves move, and the change of variables,

$$x \mapsto \lambda x, \quad y \mapsto h_0 y, \quad t \mapsto \frac{\lambda}{\sqrt{gh_0}}t, \quad u \mapsto \sqrt{gh_0}u, \quad v \mapsto \frac{\lambda}{\sqrt{gh_0^3}}v,$$

where λ is the typical, or average, wavelength and where $x \mapsto \lambda x$ is to be understood as x replaced by the new variable λx. Introducing the shallowness parameter,

$$\delta = h_0/\lambda,$$

the nondimensional full governing equations reads

$$\begin{cases} u_t + uu_x + vu_y = -p_x, & 0 \le y \le 1 + \eta/h_0, \\ \delta^2\left(v_t + uv_x + vv_y\right) = -p_y, & 0 \le y \le 1 + \eta/h_0, \\ u_x + v_y = 0, & 0 \le y \le 1 + \eta/h_0, \\ u_y - \delta^2 v_x = 0, & 0 \le y \le 1 + \eta/h_0, \\ v = \dfrac{1}{h_0}(\eta_t + u\eta_x), & y = 1 + \eta/h_0, \\ p = \dfrac{\eta}{h_0} & y = 1 + \eta/h_0, \\ v = 0, & y = 0. \end{cases}$$

The amplitude of the waves plays a fundamental role in the formulation of the water-wave problem. Thus, let a denotes the typical, or average, amplitude of a wave and express the wave profile by $\eta \mapsto a\eta$. Then, defining the amplitude parameter by

$\epsilon = a/h_0$, the nondimensional full governing equations become,

$$\begin{cases} u_t + uu_x + vu_y = -p_x, & 0 \leq y \leq 1 + \epsilon\eta, \\ \delta^2 \left(v_t + uv_x + vv_y\right) = -p_y, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_x + v_y = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_y - \delta^2 v_x = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ v = \epsilon(\eta_t + u\eta_x), & y = 1 + \epsilon\eta, \\ p = \epsilon\eta & y = 1 + \epsilon\eta, \\ v = 0, & y = 0. \end{cases}$$

Since p and v scale like ϵ, we rescale them as well as u for consistancy,

$$u \mapsto \epsilon u, \quad v \mapsto \epsilon v, \quad p \mapsto \epsilon p.$$

Then,

$$\begin{cases} u_t + \epsilon \left(uu_x + vu_y\right) = -p_x, & 0 \leq y \leq 1 + \epsilon\eta, \\ \delta^2 \left(v_t + \epsilon \left(uv_x + vv_y\right)\right) = -p_y, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_x + v_y = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_y - \delta^2 v_x = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ v = \eta_t + \epsilon u\eta_x, & y = 1 + \epsilon\eta, \\ p = \eta & y = 1 + \epsilon\eta, \\ v = 0, & y = 0. \end{cases}$$

One removes the shallowness parameter and symmetrizes the equations by considering the change of variables

$$x \mapsto \frac{\delta}{\sqrt{\epsilon}} x, \quad t \mapsto \frac{\delta}{\sqrt{\epsilon}} t, \quad v \mapsto \frac{1}{\delta\sqrt{\epsilon}} v,$$

which yields,

$$\begin{cases} u_t + \epsilon \left(uu_x + vu_y\right) = -p_x, & 0 \leq y \leq 1 + \epsilon\eta, \\ v_t + \epsilon \left(uv_x + vv_y\right) = -p_y, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_x + v_y = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ u_y - v_x = 0, & 0 \leq y \leq 1 + \epsilon\eta, \\ v = \eta_t + \epsilon u\eta_x, & y = 1 + \epsilon\eta, \\ p = \eta & y = 1 + \epsilon\eta, \\ v = 0, & y = 0. \end{cases} \tag{2.2.2}$$

The fourth equation of (2.2.2) allows us to define a velocity potential $\phi(x, y, t)$ such that

$$u = \phi_x, \quad v = \phi_y.$$

By restricting ourselves to waves travelling from left to right and by considering a slow time scale,

$$\begin{cases} \xi = x - t \\ T = \epsilon t \end{cases}$$

we obtain the following for ϕ,

$$\begin{cases} \epsilon\phi_{\xi\xi} + \phi_{yy} = 0, & 0 < y < 1 + \epsilon\eta, \\ \phi_y = 0, & y = 0, \\ \phi_y = \epsilon(-\eta_\xi + \epsilon\eta_T + \epsilon\phi_\xi\eta_\xi), & y = 1 + \epsilon\eta, \\ -\phi_\xi + \epsilon\phi_T + \dfrac{\epsilon}{2}\phi_\xi^2 + \dfrac{\phi_y^2}{2} + \eta = 0, & y = 1 + \epsilon\eta. \end{cases} \tag{2.2.3}$$

Consider the following perturbative expansions in the amplitude parameter,

$$\phi(\xi, y, T) = \sum_{k=0}^{\infty} \epsilon^k \phi^k(\xi, y, T), \quad \eta(\xi, T) = \sum_{k=0}^{\infty} \epsilon^k \eta^k(\xi, T).$$

The zero order approximation yields,

$$\begin{cases} \phi_{yy}^0 = 0, & 0 < y < 1, \\ \phi_y^0 = 0, & y = 0, \\ \phi_y^0 = 0, & y = 1, \\ -\phi_\xi^0 + \dfrac{(\phi_y^0)^2}{2} + \eta^0 = 0, & y = 1. \end{cases} \tag{2.2.4}$$

The first three equations implies,

$$\phi_y^0 \equiv 0, \quad 0 \le y \le 1,$$

while the boundary condition on $y = 1$ allow us to deduce,

$$\frac{\partial \phi^0}{\partial \xi}(\xi, y, T) = \eta^0(\xi, T), \quad 0 \le y \le 1. \tag{2.2.5}$$

Using the first and second order of (2.2.3), one obtains that η_0 satisfies the Korteweg-de Vries equation

$$2\eta_{T_0}^0 + 3\eta^0\eta_\xi^0 + \frac{1}{3}\eta_{\xi\xi\xi}^0 = 0.$$

The region where the KdV balance occurs is when $\xi = O(1), T_0 = O(1)$, that is, in the physical variables, when

$$x\frac{\epsilon^{3/2}}{\lambda\delta} = O(1), \quad x - \sqrt{gh_0}t = O(1).$$

Coming back to the physical variables, one obtains the Korteweg-de Vries equation (2.1.1) and a laminar flow for solutions η of (2.1.1)

$$\begin{cases} u(x,y,t) = \sqrt{\dfrac{g}{h_0}}\eta(x,t) \\ v(x,y,t) = -y\sqrt{\dfrac{g}{h_0}}\eta_x(x,t). \end{cases}$$

2.3 A higher velocity field for the N-soliton solution

2.3.1 The N-solitons solution

Let us recall the N-solitons solution as introduced by Hirota in [80] (see also [131]) for the Korteweg-de Vries (2.1.1). Let

$$\eta = -\frac{4h_0^3}{3}\left(\ln F\right)_{xx}, \tag{2.3.1}$$

where F is given by the power serie expansion,

$$F = 1 + \epsilon F^{(1)} + \epsilon^2 F^{(2)} + ... \tag{2.3.2}$$

For $N \in \mathbb{N}$, let $a_i > 0, s_i \in \mathbb{R}, 1 \leq i \leq N$,

$$f_i(x,t) = \exp\left(-2\sqrt{\frac{3a_i}{4h_0^3}}\left(x - s_i - \sqrt{gh_0}\left(1 + \frac{a_i}{2h_0}\right)t\right)\right).$$

and

$$F^{(1)} = \sum_{n=1}^{N} f_i(x,t).$$

Then, by replacing (2.3.1) in (2.1.1), one obtains that only the first N order term in (2.3.2) are nonzero. Moreover, one may express F as

$$F = 1 + \sum_{n=1}^{N}\sum_{{}_NC_n}\alpha(i_1,...,i_n)\prod_{j=1}^{n} f_{i_j},$$

with

$$\alpha(i_1,...,i_n) = \prod_{k<l}^{(n)}\alpha(i_k,i_l), \qquad \text{if } n \geq 2,$$

$$\alpha(i_k, i_l) = \left(\frac{\sqrt{a_{i_k}} - \sqrt{a_{i_l}}}{\sqrt{a_{i_k}} + \sqrt{a_{i_l}}}\right)^2,$$
$$\alpha(i_k) = 1,$$

where $\sum_{{}_NC_n}$ is the sum over all the possible combinations of n indexes taken from $\{1, ..., N\}$ and where

$$\prod^{(n)},$$

is the product over these n indexes, provided $k < l$ when specified. It will be convenient to denote

$$\beta_i := \sqrt{\frac{3a_i}{4h_0^3}}, \quad c_i := \sqrt{gh_0}\left(1 + \frac{a_i}{2h_0}\right).$$

Let us explain why this solution is called the N-solitons solution. Consider $0 < a_N < ... < a_1$ and $s_N < ... < s_1$. From (2.3.1), η can be recasted as

$$\begin{aligned}\eta =& 4\left(\frac{\sum_{n=1}^N \sum_{{}_NC_n} \left(\sqrt{a_{i_1}} + ... + \sqrt{a_{i_n}}\right)^2 \alpha(i_1, ..., i_n) \prod^{(n)} f_{i_n}}{\left(1 + \sum_{n=1}^N \sum_{{}_NC_n} \alpha(i_1, ..., i_n) \prod^{(n)} f_{i_n}\right)^2}\right.\\ &\frac{+\sum_{n,m=1}^N \sum_{{}_NC_{n,m}} \left(\sqrt{a_{i_1}} + ... + \sqrt{a_{i_n}}\right) \left[\left(\sqrt{a_{i_1}} + ... + \sqrt{a_{i_n}}\right)\right.}{\left(1 + \sum_{n=1}^N \sum_{{}_NC_n} a(i_1, ..., i_n) \prod^{(n)} f_{i_n}\right)^2}\\ &\left.\frac{\left.- \left(\sqrt{a_{j_1}} + ... + \sqrt{a_{j_m}}\right)\right] \alpha(i_1, ..., i_n)\alpha(j_1, ..., j_m) \prod^{(n)} \prod^{(m)} f_{i_n} f_{j_m}}{\left(1 + \sum_{n=1}^N \sum_{{}_NC_n} \alpha(i_1, ..., i_n) \prod^{(n)} f_{i_n}\right)^2}\right).\end{aligned} \tag{2.3.3}$$

First, suppose that we are in the region where $f_1 \simeq 1$ and $f_k \ll 1$ for all $k \geq 2$. The behaviour of η is given by,

$$\begin{aligned}\eta &\simeq 4a_1 \frac{f_1}{(1 + f_1)^2}\\ &= a_1 \operatorname{sech}^2 \left(\beta_1 \left(x - s_1 - c_1 t\right)\right),\end{aligned} \tag{2.3.4}$$

that is, a soliton of height a_1 and phase s_1.

Suppose now, for $2 \leq k \leq N$, that we are in the region where $f_k \simeq 1$, $f_i \gg 1$ for $1 \leq i < k$ and $f_i \ll 1$ for $k < i \leq N$. Then,

$$\eta \simeq 4a_k \frac{\alpha(1, ..., k)\alpha(1, ..., k-1) f_1^2 \cdots f_{k-1}^2 f_k}{\left(\alpha(1, ..., k-1) f_1 \cdots f_{k-1} + \alpha(1, ..., k) f_1 \cdots f_k\right)^2}$$

$$= 4a_k \frac{\frac{\alpha(1,\ldots,k)}{\alpha(1,\ldots,k-1)} f_k}{\left(1 + \frac{\alpha(1,\ldots,k)}{\alpha(1,\ldots,k-1)} f_k\right)^2},$$

$$= a_k \operatorname{sech}^2\left(\beta_k\left(x - s_k - \frac{1}{2\beta_k}\ln\left(\prod_{i=1}^{k-1}\alpha(i,k)\right) - c_k t\right)\right). \tag{2.3.5}$$

that is, a soliton with a phase shift of $\frac{1}{2\beta_k}\ln\left(\prod_{i=1}^{k-1}\alpha(i,k)\right)$. From the last computations, one sees that the N-solitons solution behaves like N distinct solitons when they are far apart from each other. Moreover, interactions between solitons must occur since the solitons are travelling at different speed. Indeed, a direct computation shows that $a_i \neq a_j$, otherwise the solution becomes a $(N-k)$-solitons solution, k being the number of times $a_i = a_j$, $i < j$. Thus, up to a change of indexes in the last computations, one sees that they are left unchanged in shape after interactions, the only notable effect of the interaction being the phase shift. Figure 3 illustrates the phase shift resulting from the interaction of 2 solitons.

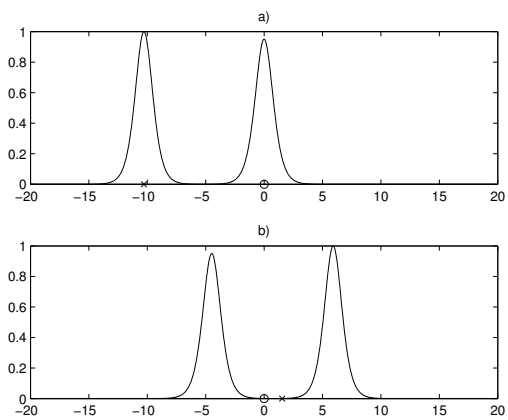

FIGURE 3 – Interaction between two solitons. The cross (resp. circle) represents the position of the maximum of the faster (resp. slower) soliton if no interaction would have occured. Figure a) is the state of the 2-solitons solution before the interaction and b) is the state after the interaction. The frame is fixed at the speed of the slower soliton.

2.3.2 Nonlaminar velocity field

The velocity potential obtained in the previous section is independent of the variable y. Therefore, even if the soliton solution of the KdV equation provides a good approximation of the solitary wave, one cannot describe accurately the underlying structure of the particles paths.

Let us explain how one may recover a nonparallel displacement of the particles. Consider that ϕ is the velocity potential of (2.2.1) satisfying,

$$\phi_x = u, \quad \phi_y = v.$$

We introduce the stream function ψ of (2.2.1) satisfying

$$\psi_y = u, \quad \psi_x = -v.$$

For irrotational flows, the velocity potential and stream function are harmonic conjugates. Consider the complex velocity potential f, with t acting as a parameter, such that

$$f(z,t) := \phi(x,y,t) + i\psi(x,y,t),\ z \in \mathbb{C},$$

with

$$f'(z,t) = u(x,y,t) - iv(x,y,t).$$

Instead of using (2.2.4) to obtain the expression of the velocity potential independant of the y variable, we rather solve the Laplacian equation of the velocity potential and of the stream function in the upper-half plane $\{z = x + iy \in \mathbb{C} | y > 0\}$ with the velocity field (2.1.3) as a boundary condition on $y = 0$.

$$\phi_y(x,0,t) = 0, \quad \phi_x(x,0,t) = \sqrt{\frac{g}{h_0}}\eta(x,t), \quad x \in \mathbb{R},\ t \in \mathbb{R} \tag{2.3.6}$$

If the first boundary condition is the natural non penetration condition, we use the independance of the y coordinate of the horizontal component of the velocity field to obtain the second boundary condition ([36, 78]).

The boundary condition implies

$$f'(x,0,t) = \sqrt{\frac{g}{h_0}}\eta(x,t).$$

By the reflection principle, the analytic function $f'(z)$ has an unique extension across $y = 0$. Therefore, f is defined up to a constant. Using Alembert's formula, we obtain the following expressions of u and v

$$u(x,y,t) = \frac{1}{2}\sqrt{\frac{g}{h_0}}\left(\eta(x+iy,t) + \eta(x-iy,t)\right),$$
$$v(x,y,t) = \frac{i}{2}\sqrt{\frac{g}{h_0}}\left(\eta(x+iy,t) - \eta(x-iy,t)\right).$$

In the case where η is the N-solitons solution, an explicit expression of the velocity

is obtained. Indeed, let,

$$\begin{aligned} F(x+iy,t) =& F^c(x,y,t) - iF^s(x,y,t), \\ F(x-iy,t) =& F^c(x,y,t) + iF^s(x,y,t), \end{aligned}$$

where,

$$\begin{aligned} F^c(x,y,t) :=& 1 + \sum_{n=1}^{N} \sum_{{}_NC_n} \alpha(i_1,...,i_n) \left(\prod_{j=1}^{n} f_{i_j}(x,t) \right) \cos\left(\sum_{j=1}^{n} 2\beta_{i_j} y \right), \\ F^s(x,y,t) :=& \sum_{n=1}^{N} \sum_{{}_NC_n} \alpha(i_1,...,i_n) \left(\prod_{j=1}^{n} f_{i_j}(x,t) \right) \sin\left(\sum_{j=1}^{n} 2\beta_{i_j} y \right). \end{aligned}$$

Then, one has,

$$\begin{aligned} u(x,y,t) =& \frac{1}{2}\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\eta(x+iy,t)+\eta(x-iy,t)\right) \\ =& \frac{1}{2}\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^c_{xx}-iF^s_{xx})(F^c-iF^s)-(F^c_x-iF^s_x)^2}{(F^c-iF^s)^2}\right. \\ & \left. + \frac{(F^c_{xx}+iF^s_{xx})(F^c+iF^s)-(F^c_x+iF^s_x)^2}{(F^c+iF^s)^2}\right) \\ =& \mathrm{Re}\left(\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^c_{xx}-iF^s_{xx})(F^c+iF^s)[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right.\right. \\ & \left.\left. - \frac{[(F^c_x-iF^s_x)(F^c+iF^s)]^2}{((F^c)^2+(F^s)^2)^2}\right)\right) \\ =& \mathrm{Re}\left(\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^c_{xx}-iF^s_{xx})(F^c+iF^s)[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right.\right. \\ & \left.\left. - \frac{[(F^c_x-iF^s_x)(F^c+iF^s)]^2}{((F^c)^2+(F^s)^2)^2}\right)\right) \\ =& \frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^cF^c_{xx}+F^sF^s_{xx})[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right. \\ & \left. - \frac{[(F^cF^c_x+F^sF^s_x)^2-(F^sF^c_x-F^cF^s_x)^2]}{((F^c)^2+(F^s)^2)^2}\right) \\ =& \frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{((F^c_x)^2+F^cF^c_{xx}+(F^s_x)^2+F^sF^s_{xx})[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right. \\ & \left. - 2\frac{(F^cF^c_x+F^sF^s_x)^2}{((F^c)^2+(F^s)^2)^2}\right). \end{aligned} \tag{2.3.7}$$

On the other hand,

$$
\begin{aligned}
v(x,y,t) =& \frac{i}{2}\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\eta(x+iy,t)-\eta(x-iy,t)\right)\\
=& \frac{i}{2}\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^c_{xx}-iF^s_{xx})(F^c-iF^s)-(F^c_x-iF^s_x)^2}{(F^c-iF^s)^2}\right.\\
&\left.-\frac{(F^c_{xx}+iF^s_{xx})(F^c+iF^s)-(F^c_x+iF^s_x)^2}{(F^c+iF^s)^2}\right)\\
=& -\operatorname{Im}\left(\frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^c_{xx}-iF^s_{xx})(F^c+iF^s)[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right.\right.\\
&\left.\left.+\frac{[(F^c_x+iF^s_x)(F^c-iF^s)]^2}{((F^c)^2+(F^s)^2)^2}\right)\right)\\
=& \frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^cF^s_{xx}-F^sF^c_{xx})[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right.\\
&\left.-2\frac{(F^cF^c_x+F^sF^s_x)(F^cF^s_x-F^sF^c_x)}{((F^c)^2+(F^s)^2)^2}\right)\\
=& \frac{4h_0^3}{3}\sqrt{\frac{g}{h_0}}\left(\frac{(F^cF^c_{xy}+F^sF^s_{xy})[(F^c)^2+(F^s)^2]}{((F^c)^2+(F^s)^2)^2}\right.\\
&\left.-2\frac{(F^cF^c_x+F^sF^s_x)(F^cF^c_y+F^sF^s_y)}{((F^c)^2+(F^s)^2)^2}\right),
\end{aligned}
\tag{2.3.8}
$$

the last line coming from the expression of the derivatives of F^c and F^s,

$$
\begin{aligned}
\frac{\partial^k F^c}{\partial x^k} &= \sum_{n=1}^{N}\sum_{{}_NC_n}(-2\beta_{i_1}-...-2\beta_{i_n})^k\alpha(i_1,...,i_n)\Big(\prod_{j=1}^{n}f_{i_j}(x,t)\Big)\cos\Big(\sum_{j=1}^{n}2\beta_{i_j}y\Big),\\
\frac{\partial^k F^s}{\partial x^k} &= \sum_{n=1}^{N}\sum_{{}_NC_n}(-2\beta_{i_1}-...-2\beta_{i_n})^k\alpha(i_1,...,i_n)\Big(\prod_{j=1}^{n}f_{i_j}(x,t)\Big)\sin\Big(\sum_{j=1}^{n}2\beta_{i_j}y\Big),\\
\frac{\partial F^c}{\partial y} &= \sum_{n=1}^{N}\sum_{{}_NC_n}(-2\beta_{i_1}-...-2\beta_{i_n})\alpha(i_1,...,i_n)\Big(\prod_{j=1}^{n}f_{i_j}(x,t)\Big)\sin\Big(\sum_{j=1}^{n}2\beta_{i_j}y\Big),\\
\frac{\partial F^s}{\partial y} &= \sum_{n=1}^{N}\sum_{{}_NC_n}(2\beta_{i_1}+...+2\beta_{i_n})\alpha(i_1,...,i_n)\Big(\prod_{j=1}^{n}f_{i_j}(x,t)\Big)\cos\Big(\sum_{j=1}^{n}2\beta_{i_j}y\Big).
\end{aligned}
$$

2.3.3 Proof of Theorem 2.1.1

Before proving Theorem 2.1.1, let us introduce some notations to simplify some of the expressions. We denote

$$\sum := \sum_{n=1}^{N} \sum_{{}_N C_n}, \quad \prod := \prod_{j=1}^{n},$$

$\sum^m$, for m sums, and $\prod^m$, for m products. We shorten the expression of $\alpha(i_1, ..., i_n)$ by α_{i_n} and define $\alpha_{i_n,j_m} := \alpha_{i_n}\alpha_{j_m}$ and so on for product over more than two terms. Finally, we use the slightly modified Einstein notation (β_{i_n}) for a sum from β_{i_1} to β_{i_n}, while β_{i_j} denotes a single coefficient. Since no confusion is possible, $(\beta_{i_n})(\beta_{j_m})$ denotes the product over the two sums.

Let us turn ourselves to the proof of Theorem 2.1.1

Proof :

(a) From (2.3.7)-(2.3.8), we recover the original velocity field when $y = 0$.

(b) The expression of the numerator of u is given by,

$$\begin{aligned}
&\sum (2\beta_{i_n})^2 \alpha_{i_n} \prod f \cos((2\beta_{i_n})y) \\
&+ \sum^{2} \left[3(2\beta_{i_n})^2 - (2\beta_{i_n})(2\beta_{j_m})\right] \alpha_{i_n,j_m} \prod^{2} f \cos((2\beta_{i_n})y) \cos((2\beta_{j_m})y) \\
&+ \sum^{3} \left[3(2\beta_{i_n})^2 - 2(2\beta_{i_n})(2\beta_{j_m})\right] \\
&\quad \times \alpha_{i_n,j_m,k_p} \prod^{3} f \cos((2\beta_{i_n})y) \cos((2\beta_{j_m})y) \cos((2\beta_{k_p})y) \\
&+ \sum^{4} \left[(2\beta_{i_n})^2 - (2\beta_{i_n})(2\beta_{j_m})\right] \\
&\quad \times \alpha_{i_n,j_m,k_p,l_q} \prod^{4} f \cos((2\beta_{i_n})y) \cos((2\beta_{j_m})y) \cos((2\beta_{k_p})y) \cos((2\beta_{l_q})y) \\
&+ \sum^{2} \left[(2\beta_{i_n})^2 + (2\beta_{i_n})(2\beta_{j_m})\right] \alpha_{i_n,j_m} \prod^{2} f \sin((2\beta_{i_n})y) \sin((2\beta_{j_m})y) \\
&+ \sum^{4} \left[(2\beta_{i_n})^2 - (2\beta_{i_n})(2\beta_{j_m})\right] \\
&\quad \times \alpha_{i_n,j_m,k_p,l_q} \prod^{4} f \sin((2\beta_{i_n})y) \sin((2\beta_{j_m})y) \sin((2\beta_{k_p})y) \sin((2\beta_{l_q})y) \\
&+ \sum^{3} \left[(2\beta_{i_n})^2 + 2(2\beta_{k_p})(2\beta_{l_q}) + 2(2\beta_{k_p})^2 - 4(2\beta_{i_n})(2\beta_{k_p})\right] \\
&\quad \times \alpha_{i_n,j_m,k_p} \prod^{3} f \cos((2\beta_{i_n})y) \sin((2\beta_{j_m})y) \sin((2\beta_{k_p})y)
\end{aligned}$$

$$+\sum^{4}\left[(2\beta_{i_n})^2+(2\beta_{i_n})(2\beta_{j_m})+(2\beta_{k_p})^2+(2\beta_{k_p})(2\beta_{l_q})-4(2\beta_{i_n})(2\beta_{k_p})\right]$$
$$\times\,\alpha_{i_n,j_m,k_p,l_q}\prod^{4} f\cos((2\beta_{i_n})y)\cos((2\beta_{j_m})y)\sin((2\beta_{k_p})y)\sin((2\beta_{l_q})y).$$

The coefficients of the first sum with respect to β_i are all positive. Moreover, the coefficients appearing in the other sums are non-negative. Indeed, consider for example the coefficient in front of $\cos((2\beta_{i_n})y)\cos((2\beta_{j_m})y)\sin((2\beta_{k_p})y)\sin((2\beta_{l_q})y)$:

$$\begin{aligned}
&(2\beta_{i_n})^2+(2\beta_{i_n})(2\beta_{j_m})+(2\beta_{k_p})^2+(2\beta_{k_p})(2\beta_{l_q})-4(2\beta_{i_n})(2\beta_{k_p})\\
&+(2\beta_{j_m})^2+(2\beta_{j_m})(2\beta_{i_n})+(2\beta_{l_q})^2+(2\beta_{l_q})(2\beta_{k_p})-4(2\beta_{j_m})(2\beta_{l_q})\\
&=((2\beta_{i_n})-(2\beta_{k_p}))^2+(2\beta_{i_n})((2\beta_{j_m})-(2\beta_{k_p}))+(2\beta_{k_p})((2\beta_{l_q})-(2\beta_{i_n}))\\
&+((2\beta_{j_m})-(2\beta_{l_q}))^2+(2\beta_{j_m})((2\beta_{i_n})-(2\beta_{i_n}))+(2\beta_{l_q})((2\beta_{k_p})-(2\beta_{j_m}))\\
&=((2\beta_{i_n})+(2\beta_{j_m})-(2\beta_{k_p})-(2\beta_{l_q}))^2.
\end{aligned}$$

Therefore, since

$$0\le y\le h_0+a,$$

where,

$$a:=\max_{x\in\mathbb{R}}\eta=\max_{1\le i\le n}a_i,$$

in order to prove (b), it is sufficient to impose,

$$(h_0+a)\sum_{i=1}^{N}\beta i\le\frac{\pi}{4}. \tag{2.3.9}$$

(c) Let us give an equivalent form of v. Consider first

$$\begin{aligned}
\left[F^cF^c_{xy}+F^sF^s_{xy}\right]&=4\sum(\beta_{i_n})^2\alpha_{i_n}\prod f\sin(2(\beta_{i_n})y)\\
&\quad+4\sum^{2}\left[(\beta_{j_m})^2-(\beta_{i_n})^2\right]\alpha_{i_n,j_m}\prod^{2} f\cos(2(\beta_{i_n})y)\sin(2(\beta_{j_m})y).
\end{aligned}$$

Then coefficient associated to $(f_{i_1}\dots f_{i_n})(f_{j_1}\dots f_{j_m})$ in this expression is given by

$$\begin{aligned}
4\alpha_{i_n,j_m}\big(&\left[(\beta_{j_m})^2-(\beta_{i_n})^2\right]\cos(2(\beta_{i_n})y)\sin(2(\beta_{j_m})y)\\
&+\left[(\beta_{i_n})^2-(\beta_{j_m})^2\right]\cos(2(\beta_{j_m})y)\sin(2(\beta_{i_n})y)\big),
\end{aligned}$$

which can be written as

$$4\alpha_{i_n,j_m}\left[(\beta_{j_m})^2-(\beta_{i_n})^2\right]\sin(2((\beta_{j_m})-(\beta_{i_n}))y).$$

Therefore, using the symmetry of the second sum, one has

$$\begin{aligned}\left[F^cF^c_{xy}+F^sF^s_{xy}\right] &= 4\sum(\beta_{i_n})^2\alpha_{i_n}\prod f\sin(2(\beta_{i_n})y)\\ &\quad+2\sum^{2}\left[(\beta_{j_m})^2-(\beta_{i_n})^2\right]\alpha_{i_n,j_m}\prod^{2} f\sin(2((\beta_{j_m})-(\beta_{i_n}))y),\end{aligned}$$

In the same fashion, one obtains the following expressions

$$\begin{aligned}-2\left[F^cF^c_{y}+F^sF^s_{y}\right] &= 4\sum(\beta_{i_n})\alpha_{i_n}\prod f_i\sin(2(\beta_{i_n})y)\\ &\quad+2\sum^{2}\left[(\beta_{j_m})-(\beta_{i_n})\right]\alpha_{i_n,j_m}\prod^{2} f\sin(2((\beta_{j_m})-(\beta_{i_n}))y),\\ \left[F^cF^c_{x}+F^sF^s_{x}\right] &= -2\sum(\beta_{i_n})\alpha_{i_n}\prod f\cos(2(\beta_{i_n})y)\\ &\quad-2\sum^{2}(\beta_{i_n})\alpha_{i_n,j_m}\prod^{2} f\cos(2((\beta_{i_n})-(\beta_{j_m}))y),\end{aligned}$$

$$\begin{aligned}-2\left[F^cF^c_{x}+F^sF^s_{x}\right]\left[F^cF^c_{y}+F^sF^s_{y}\right] &= -8\sum^{2}(\beta_{i_n})(\beta_{j_m})\\ &\quad\times\alpha_{i_n,j_m}\prod^{2} f\sin(2(\beta_{i_n})y)\cos(2(\beta_{j_m})y)\\ &\quad-4\sum^{3}(\beta_{i_n})\left[(\beta_{j_m})-(\beta_{k_p})\right]\\ &\quad\times\alpha_{i_n,j_m,k_p}\prod^{3} f\sin(2((\beta_{j_m})-(\beta_{k_p}))y)\cos(2(\beta_{i_n})y)\\ &\quad-8\sum^{3}(\beta_{i_n})(\beta_{j_m})\alpha_{i_n,j_m,k_p}\prod^{3} f\sin(2(\beta_{i_n})y)\cos(2((\beta_{j_m})-(\beta_{k_p}))y)\\ &\quad-4\sum^{4}(\beta_{i_n})\left[(\beta_{k_p})-(\beta_{l_q})\right]\\ &\quad\times\alpha_{i_n,j_m,k_p,l_q}\prod^{4} f\sin(2((\beta_{k_p})-(\beta_{l_q}))y)\cos(2((\beta_{i_n})-(\beta_{j_m}))y).\end{aligned}$$

Thus, the numerator of v boils down to

$$\begin{aligned}&4\sum(\beta_{i_n})^2\alpha_{i_n}\prod f\sin(2(\beta_{i_n})y)\\ &+2\sum^{2}\left[((\beta_{j_m})^2-(\beta_{i_n})^2)\right]\alpha_{i_n,j_m}\prod^{2} f\sin(2((\beta_{j_m})-(\beta_{i_n}))y)\\ &+8\sum^{2}(\beta_{i_n})[(\beta_{i_n})-(\beta_{j_m})]\alpha_{i_n,j_m}\prod^{2} f\sin(2(\beta_{i_n})y)\cos(2(\beta_{j_m})y)\\ &+4\sum^{3}[(\beta_{j_m})((\beta_{j_m})-(\beta_{i_n}))-(\beta_{k_p})((\beta_{k_p})-(\beta_{i_n}))]\end{aligned}$$

$$\times\, \alpha_{i_n,j_m,k_p} \prod^{3} f \sin(2((\beta_{j_m}) - (\beta_{k_p}))y) \cos(2(\beta_{i_n})y)$$
$$+ 4 \sum^{3} [(\beta_{i_n})((\beta_{i_n}) - 2(\beta_{j_m}))] \alpha_{i_n,j_m,k_p} \prod^{3} f \sin(2(\beta_{i_n})y) \cos(2((\beta_{j_m}) - (\beta_{k_p}))y)$$
$$+ 2 \sum^{4} [(\beta_{j_m})((\beta_{j_m}) - 2(\beta_{k_p})) - (\beta_{i_n})((\beta_{i_n}) - 2(\beta_{k_p}))]$$
$$\times\, \alpha_{i_n,j_m,k_p,l_q} \prod^{4} f \sin(2((\beta_{j_m}) - (\beta_{i_n}))y) \cos(2((\beta_{k_p}) - (\beta_{l_q}))y).$$

For sake of simplicity, assume that $0 < a_n < \ldots < a_1$. From the previous expression, the behaviour of v when $x \to \infty$ is governed by f_N,

$$v \simeq 4\sqrt{\frac{g}{h_0}} a_N f_N \sin(2\beta_N y)$$

which is positive.

On the other hand, the leading coefficient of v when $x \to -\infty$, as the coefficient in front $(f_1 \ldots f_N)^4$ vanishes at the numerator, is given by

$$v \simeq -4\sqrt{\frac{g}{h_0}} \alpha(1, ..., N)^3 \alpha(1, ..., N-1) a_N (f_1 \ldots f_{N-1})^4 f_N^3 \sin(2\beta_N y),$$

which is negative.

Let us consider the case $f_1 \simeq 1$, $f_i \ll 1, i > 1$. The general case is obtained by permutations of indexes. The leading term of v corresponds to the different powers of f_1,

$$v \simeq \sqrt{\frac{g}{h_0}} f_1 (1 - f_1^2)[4a_1 \sin(2\beta_1 y)] \tag{2.3.10}$$

which is positive in the neighborhood of $x = s_1 + c_1 t$ if and only if $x - c_1 t > s_1$. The neglected terms in the previous analysis don't affect the nature of the change of sign but rather, and slightly, its position.

In the case where $f_k \simeq 1$, $f_i \gg 1$, $i < k$ and $f_i \ll 1$, $i > k$, the leading terms of the numerator are given by

$$\sqrt{\frac{g}{h_0}} \frac{4h_0^3}{3} \left(4\alpha_{k-1}^3 \alpha_k (f_1 \ldots f_{k-1})^4 f_k \sin(2\beta_k y) \left[(\beta_k)(\beta_k - (\beta_{k-1})) - (\beta_{k+1})^2\right]\right.$$
$$\left. + 4\alpha_k^3 \alpha_{k-1} (f_1 \ldots f_{k-1})^4 f_k^3 \sin(2\beta_k y) \left[-(\beta_k)^2 + (\beta_{k-1})(\beta_k + (\beta_k))\right]\right)$$
$$= 4a_k \sqrt{\frac{g}{h_0}} \Big(1 - \Big(\frac{\alpha(1, ..., k)}{\alpha(1, ..., k-1)} f_k\Big)^2\Big) \frac{\alpha(1, ..., k)}{\alpha(1, ..., k-1)} (f_1 \ldots f_{k-1})^4 f_k \sin(2\beta_k y).$$

thus,

$$v \simeq 4a_k\sqrt{\frac{g}{h_0}}\frac{\left(1-\left(\frac{\alpha(1,...,k)}{\alpha(1,...,k-1)}f_k\right)^2\right)\frac{\alpha(1,...,k)}{\alpha(1,...,k-1)}(f_1\dots f_{k-1})^4 f_k\sin(2\beta_k y)}{\left[(F^c)^2+(F^s)^2\right]^2}.$$

Factoring $\alpha(1,...,k-1)^4(f_1\dots f_{k-1})^4$ at the denominator yields,

$$v \simeq 4a_k\sqrt{\frac{g}{h_0}}\left(1-\left(\frac{\alpha(1,...,k)}{\alpha(1,...,k-1)}f_k\right)^2\right)\frac{\alpha(1,...,k)}{\alpha(1,...,k-1)}f_k\sin(2\beta_k y). \tag{2.3.11}$$

We obtain that this last expression is positive in the neighbourhood of $x = s_k + c_k t$ if and only if $x - c_k t > s_k + \frac{1}{2\beta_k}\ln(\alpha(1,...,k)/\alpha(1,...,k-1))$. The phase shift here reproducing the phase shift in the N-solitons solution shown in (2.3.5). Here again, the neglected terms may induce a small difference between the phase shift in this analysis and the one corresponding to (2.3.5) but the nature of the change of sign of v is preserved.

From the previous analysis, we obtain c). Indeed, $v > 0$ when $x \to \infty$ and v is still positive from $x \to \infty$ to a small neighborhood of $x = s_1 + c_1 t$, representing the position of the crest of the first soliton. From the neighbourhood of $x = s_1 + c_1 t$ to the neighborhood of $x = s_2 + c_2 t + \frac{1}{2\beta_2}\ln(\alpha(1,2))$, only (2.3.10) and (2.3.11) play an important role in the sign of v. The functions f_i being monotonic, only one change of sign of v occurs in this region and v therefore become positive for values of x greater than the small neighborhood of $x = s_2 + c_2 t + \frac{1}{2\beta_2}\ln(\alpha(1,2))$. The analysis of (2.3.11) shows that there a change of sign in the neighborhood of the crest of the second soliton. The same analysis can be carried through all the crests of the N-solitons solution until the final crest, that is, at the left of the last soliton where v is always negative.

□

2.4 Numerical analysis of the particles trajectory

This section is devoted to the numerical analysis of the particle trajectories given by the higher approximation of the velocity field (2.3.7)-(2.3.8). First, we compare the results in the case of a single soliton with the numerical approximation of the particle trajectories of the first order approximation field (2.1.3). We use the experimental results in [34] as the reference trajectory. The simulation were done using a fourth order Runge-Kutta scheme. Figure 4 represents the numerical approximation of the particles trajectory with both velocity field with the parameters of experiment (c) of [34].

Table 5 compares the total displacement in the x variable and the maximal dis-

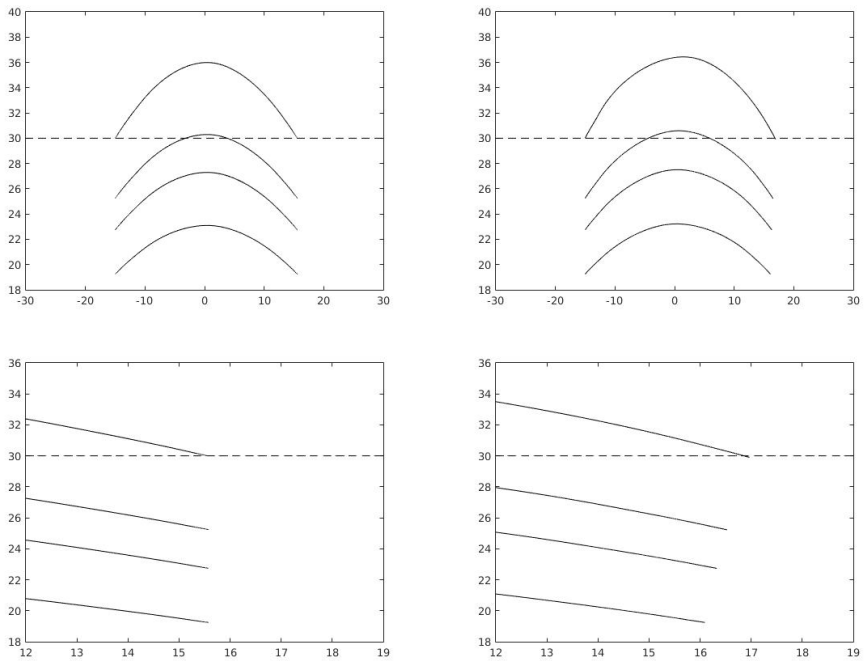

FIGURE 4 – Comparison of the numerical approximation of the particle trajectories for the first order velocity field (top left) and the higher order velocity field (top right). Zoom on the end of the particle trajectories for the first order velocity field (bottom left) and the higher order velocity field (bottom right). The depth of the fluid is 30 cm and the height of the solitary wave is 5.46 cm. The dashed line represents the undisturbed water surface.

placement in the y variable of the velocity fields with the experimental results of [34]. From the numerical results, one recovers an important feature of the particle trajectories of solitary waves : the higher the particles are initially located, the greater the horizontal displacement. However, one notices for both velocity fields that the vertical displacement of the particle located at the surface is greater that the height of the soliton. It comes from the fact that the free surface kinematic boundary conditions assuring that the particles stay underneath the water surface is not preserved in (2.1.3) and therefore in (2.3.7)-(2.3.8). Indeed, this boundary condition is a higher order term in the approximation parameter as one can see from the third line of (2.2.3). Numerical simulations show that the overshoot is proportional to ϵ. The numerical approximations of the particles trajectory are therefore relevant for small values of ϵ and may be interpreted qualitatively for larger values of ϵ.

Figure 6 illustrates the numerical approximation of the particles trajectory associated to the higher velocity field in the case of the 2-solitons in the case where

$$\beta_i(h_0 + a) < \frac{\pi}{4}, \forall i \in \{1, \dots, N\} \tag{2.4.1}$$

b (cm)	X First (cm)	Y First (cm)	X Hi. (cm)	Y Hi. (cm)	X Exp. (cm)	Y Exp. (cm)
30	30.57	6.00	31.97	6.43	30.63	5.94
25.25	30.57	5.05	31.53	5.36	30.10	4.45
22.75	30.57	4.55	31.33	4.77	30.17	3.93
19.25	30.57	3.85	31.09	3.97	29.42	2.96

FIGURE 5 – Total displacement (X) in the x variable and maximal displacement (Y) in the y variable with respect to the initial vertical position above the flat bottom of the particle b for the first order velocity field (First), the higher velocity field (Hi.) and the experimental results (Exp.).

holds but the sufficient condition (2.3.9) of Theorem 2.1.1 does not.

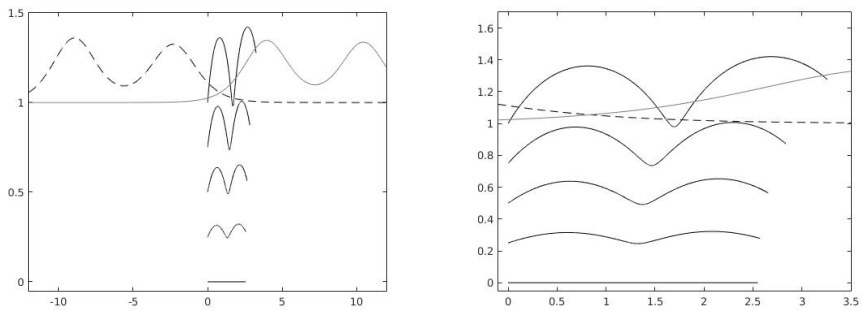

FIGURE 6 – Numerical approximation of the particle trajectories for the 2-solitons solution. The particles trajectories are in black, the initial position of the 2-solitons is in dashed black and the final position is in gray. The height of the soliton in front is 0.4cm and the soliton behind is 0.3cm. The depth of the water is 1cm.

Up to a renormalisation, one obtains a physically relevant particle trajectories in the case of an interaction between two solitons. Moreover, numerical investigation shows that (2.4.1) is a necessary and sufficient condition for (2.3.7)-(2.3.8) to be physically relevant for the relevance of the particles trajectory.

2.5 Conclusion

It was shown theoretically that, under hypothesis (2.3.9), when there is no interactions, the behaviour of the particles under the velocity field (2.3.7)-(2.3.8) is the one induced by N distinct solitons, including their phase shift. The numerical results allow us to conclude on three points which were difficult to prove theoretically for the velocity field (2.3.7)-(2.3.8) :

1. The higher the particles is initially located, the greater its horizontal displacement ;

2. Hypothesis (2.4.1), which is the theoretical bound on the amplitude parameter for solitons to exist ([131]), is necessary and sufficient for the physical relevance of (2.3.7)-(2.3.8) ;
3. The particles paths obtained from (2.3.7)-(2.3.8) are still relevant when interactions between solitons occur.

The numerical investigation also allowed us to highlight the possible lost of the free surface kinematic boundary condition for the first order velocity field and the higher order velocity field obtained from the first order velocity field. If one follows the derivation of the Korteweg-de Vries as in [131], one sees that this free surface kinematic boundary condition appears at the second order and might be recovered for second order velocity fields. However, even at this order of approximation, an overshoot is still present in some cases for second order velocity fields [16]. It is in contrast with the Saint-Venant equation for which the free surface kinematic boundary condition is insured. Quantitative results are recovered for the KdV velocity fields for small value of ϵ or by renormalisation of the particles trajectories.

Deuxième partie

Stabilisation d'une équation de Schrödinger

Chapitre 3

Stabilisation d'une équation de Schrödinger

Contents

3.1 Introduction 132

3.1.1 Main result 132

3.1.2 A finite dimensional example 134

3.1.3 The linear Schrödinger equation 139

3.1.4 A brief review of previous results 140

3.1.5 Structure of the article 141

3.2 Heuristic construction of the transformations 142

3.3 Definition and properties of the transformation 147

3.3.1 Riesz basis property 147

3.3.2 Definition and regularity of the feedback law 156

3.3.3 Wellposedness of the transformation 162

3.4 Invertibility of the transformation 163

3.4.1 Fredholm form 163

3.4.2 Operator equality 166

3.4.3 Invertibility 171

3.5 Well-posedness of the linearized system 174

3.5.1 Well-posedness of the linearized system 174

3.5.2 Proof of rapid stabilization 177

3.6 Admissibility of the feedback operator 178

3.1 Introduction

3.1.1 Main result

Let $T > 0$. Consider the Schrödinger equation

$$\begin{cases} i\partial_t\psi = -\Delta\psi - u(t)\mu(x)\psi, & (x,t) \in (0,1) \times (0,T), \\ \psi(t,0) = \psi(t,1) = 0, & t \in (0,T). \end{cases} \tag{3.1.1}$$

In (3.1.1), ψ is the wave function, of L^2-norm 1, of a particle confined in a $1-D$ infinite square potential well. The particle is subjected to an electric field inside of the domain, where $u \in L^2((0,T);\mathbb{R})$ is the amplitude of the eletric field and $\mu \in H^3((0,1);\mathbb{R})$ is the dipolar moment of the particle.

Before stating our main result, we set some notations. Let $\mathcal{A} : D(\mathcal{A}) \subset L^2(0,1) \to L^2(0,1)$ be defined by

$$\mathcal{A}\phi := -\Delta\phi, \quad D(\mathcal{A}) := H^2(0,1) \cap H_0^1(0,1).$$

The eigenvalues and eigenfunctions of $\mathcal{A}$ are given by

$$\lambda_k := (k\pi)^2, \quad \varphi_k(x) := \sqrt{2}\sin(k\pi x), \quad k \in \mathbb{N}^*.$$

The eigenstates of (3.1.1) ($u = 0$) are given by $\Phi_k(t,x) := e^{-i\lambda_k t}\varphi_k(x)$. The eigenstate Φ_1 associated to the smallest eigenvalue is called the ground state.

Define the space $H^s_{(0)} := D(\mathcal{A}^{s/2})$ equipped with the inner product

$$\langle \phi, \psi \rangle_{H^s_{(0)}} := \sum_{k=1}^{+\infty} \lambda_k^s \langle \phi, \varphi_k \rangle \overline{\langle \psi, \varphi_k \rangle},$$

where $\langle \cdot, \cdot \rangle$ is the $L^2((0,1);\mathbb{C})$-inner product. The space $H^s_{(0)}$ is endowed with the $\|.\|_{H^s_{(0)}}$-norm associated to the $H^s_{(0)}$-inner product. We underline that the spaces used in this article can also be explicitly described by $H^2_{(0)} = H^2 \cap H_0^1((0,1);\mathbb{C})$ and

$$\begin{aligned} H^3_{(0)} &= \left\{ \phi \in H^3 \cap H_0^1((0,1);\mathbb{C}) \; ; \; \phi''(0) = \phi''(1) = 0 \right\}, \\ H^5_{(0)} &= \left\{ \phi \in H^5 \cap H^3_{(0)}((0,1);\mathbb{C}) \; ; \; \phi^{(4)}(0) = \phi^{(4)}(1) = 0 \right\}. \end{aligned}$$

We denote by $\mathcal{S}$ the radius 1 sphere of $L^2((0,1);\mathbb{C})$.

Throughout this article, we will assume that μ satisfy the following assumption.

Hypothesis 3.1.1 *The function μ belongs to $H^3((0,1);\mathbb{R})$ and there exists $c>0$ satisfying*

$$\frac{c}{k^3} \leq |\langle \mu\varphi_1, \varphi_k\rangle|, \quad \forall k \in \mathbb{N}^*.$$

Remark 3.1.1 *A direct computations shows that, for every function $\mu \in H^3((0,1);\mathbb{R})$, we have*

$$\langle \mu\varphi_1, \varphi_k\rangle = \frac{4}{(k\pi)^3}\left((-1)^{k+1}\mu'(1) - \mu(0)\right) - \frac{\sqrt{2}}{(k\pi)^3}\int_0^1 (\mu\varphi_1)'''(x)\cos(k\pi x)dx$$

and therefore there exists $C>0$ such that

$$|\langle \mu\varphi_1, \varphi_k\rangle| \leq \frac{C}{k^3}, \quad \forall k \in \mathbb{N}^*.$$

Hypothesis 3.1.1 is not necessary to get local exact controllability of (3.1.1). However, as proved in [10], it is a necessary and sufficient condition to get exact controllability of the following linearized equation around the ground state

$$\begin{cases} i\partial_t\Psi = -\Delta\Psi - v(t)\mu(x)\Phi_1(t,x), & (x,t) \in (0,1)\times(0,T), \\ \Psi(t,0) = \Psi(t,1) = 0, & t \in (0,T), \\ \Psi(0,x) = \Psi_0(x), & x \in (0,1). \end{cases} \tag{3.1.2}$$

Theorem 3.1.1 ([7]) *Let $T>0$ and assume that μ satisfies Hypothesis 3.1.1. Then, for every*

$$\Psi_0 \in \left\{\phi \in H^3_{(0)}\ ;\ \Re\langle\phi,\varphi_1\rangle = 0\right\} =: \mathcal{H}_0, \qquad \Psi_T \in \left\{\phi \in H^3_{(0)}\ ;\ \Re\langle\phi,\Phi_1(T)\rangle = 0\right\} =: \mathcal{H}_T, \tag{3.1.3}$$

there exists $v \in L^2((0,T);\mathbb{R})$ such that the solution Ψ of (3.1.2) with the initial condition $\Psi(0,.) = \Psi_0$ satisfies, $\Psi(T,.) = \Psi_T$.

Notice that condition (3.1.3) is not a restriction. Due to the linearization of the preservation of the norm, the solution of (3.1.2) satisfies $\Re\langle\Psi(t),\Phi_1(t)\rangle = 0$ for every $t \geq 0$.

The main result of this paper, using the ideas of the backstepping method in infinite dimension (see [89] for a detailed presentation), is the rapid stabilization of the linearized equation (3.1.2).

Theorem 3.1.2 *Let $T>0$. Assume that μ satisfies Hypothesis 3.1.1. Then, for every $\lambda>0$, there exists $C>0$ and a feedback law $v(t) = K(\Psi(t,\cdot))$ such that for*

every $\Psi_0 \in \mathcal{H}_0$ *the associated solution of (3.1.2) satisfies*

$$\|\Psi(t,\cdot)\|_{H^3_{(0)}} \leq Ce^{-\lambda t}\|\Psi_0\|_{H^3_{(0)}}.$$

For the sake of simplicity we will focus, for the rest of this article, on the rapid stabilization of the following linear Schrödinger equation

$$\begin{cases} i\partial_t\Psi = -\Delta\Psi - v(t)\mu(x)\varphi_1(x), & (x,t) \in (0,1)\times(0,T), \\ \Psi(t,0) = \Psi(t,1) = 0, & t \in (0,T), \\ \Psi(0,x) = \Psi_0(x), & x \in (0,1). \end{cases} \tag{3.1.4}$$

The only difference between (3.1.2) and (3.1.4) is that the control term "$v(t)\mu(x)\Phi_1(t,x)$" has been replaced by "$v(t)\mu(x)\varphi_1$". Using again [7, Proposition 4], we get the analogous of Theorem 3.1.1 that is exact controllability with $L^2((0,T);\mathbb{R})$ controls of system (3.1.4) but now in the state space $H^3_{(0)}$.

We prove the following rapid stabilization result.

Theorem 3.1.3 *Let* $T > 0$. *Assume that* μ *satisfies Hypothesis 3.1.1. Then, for every* $\lambda > 0$, *there exists* $C > 0$ *and a feedback law* $v(t) = K(\psi(t,\cdot))$ *such that for every* $\Psi_0 \in H^3_{(0)}$ *the associated solution of (3.1.4) satisfies*

$$\|\Psi(t,\cdot)\|_{H^3_{(0)}} \leq Ce^{-\lambda t}\|\Psi_0\|_{H^3_{(0)}}.$$

Remark 3.1.2 *The final goal would be to achieve local rapid stabilization of the bilinear problem (3.1.1) toward the ground state* Φ_1. *To avoid dealing with a moving target, notice that*

$$\|\psi(t,\cdot) - \Phi_1(t,\cdot)\|_{H^3_{(0)}} = \|e^{i\lambda_1 t}\psi(t,\cdot) - \varphi_1\|_{H^3_{(0)}}.$$

Thus it is simpler to look at the system satisfied by $e^{i\lambda_1 t}\psi(t,\cdot)$. *In the same spirit, we will develop the proof of Theorem 3.1.3 in this article and detail in Appendix A.3 how we can modify the proof to obtain Theorem 3.1.2. The obtained feedback law does not allow us, for now, to obtain rapid stabilization of (3.1.1).*

3.1.2 A finite dimensional example

Let us explain the general idea of the proof of Theorem 3.1.3 in a finite dimensional setting. Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^n$. Consider the control system

$$x'(t) = Ax(t) + Bu(t) \tag{3.1.5}$$

where, at time t, the state is $x(t) \in \mathbb{R}^n$ and the control is $u(t) \in \mathbb{R}$. We assume that the system is controllable, which is equivalent to the Kalman rank condition

$$rank(B, AB, \ldots, A^{n-1}B) = n \tag{3.1.6}$$

(see e.g. [42, Theorem 1.16]). It is well-known that the controllability allows one to use the pole-shifting theorem [42, Theorem 10.1]) to design a feedback law $u(t) = Kx(t)$ to obtain the exponential stability with an arbitrary exponential decay rate of (3.1.5). Let us present a different approach, more suitable for PDEs, to this result. Let $\lambda \in \mathbb{R}$. Consider the target system

$$y'(t) = (A - \lambda I)y(t) + Bv(t) \tag{3.1.7}$$

where, at time t, $y(t) \in \mathbb{R}^n$ is the state and $v(t) \in \mathbb{R}$ is the control. A straightforward computation shows the existence of $C > 0$ and R such that, for $v \equiv 0$, the solutions to (3.1.7) satisfy

$$\|y(t)\| \leq e^{-(\lambda - \|A\|)t}\|y(0)\|.$$

In order to design a feedback law for (3.1.5), consider the linear mapping

$$\begin{pmatrix} x \\ v \end{pmatrix} \mapsto \begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} Tx \\ Kx + v \end{pmatrix}. \tag{3.1.8}$$

Note that, if T is invertible,

$$\begin{aligned} \|x(t)\| = \|T^{-1}y(t)\| &\leq \|T^{-1}\|e^{-(\lambda-\|A\|)t}\|y(0)\| \\ &\leq \|T^{-1}\|e^{-(\lambda-\|A\|)t}\|Tx(0)\| \leq \|T^{-1}\|\|T\|e^{-(\lambda-\|A\|)t}\|x(0)\|. \end{aligned}$$

Therefore, the exponential stability of (3.1.5) with an arbitrary exponential decay rate is reduced to find a transformation (T, K) such that T is continuous and invertible. The transformation (3.1.8) maps (3.1.5) into

$$y'(t) = Tx'(t) = T\left(Ax(t) + BKx(t) + Bv(t)\right) = \left(TA + TBK\right)x(t) + TBv(t),$$

Hence the transformation (T, K) transformation (3.1.8) maps (3.1.5) into (3.1.7) if and only if

$$TA + BK = AT - \lambda T, \tag{3.1.9}$$

$$TB = B. \tag{3.1.10}$$

One has the following theorem.

Theorem 3.1.4 ([44]) *There exists one and only one $(T, K) \in GL_n(\mathbb{R}) \times \mathbb{R}^{1\times n}$ satisfying (3.1.9)-(3.1.10).*

The proof of Theorem 3.1.4 provided in [44] relies on the companion matrix of A and is not suitable in the infinite dimensional setting. We present here a different proof of this result on which the proof of Theorem 3.1.3 will be based.

Let $\{\lambda_i, e_i\}_{1\leq i\leq n}$ be the eigenvalues and eigenvectors of A. The eigenvalues are assumed to be simple. Then, (3.1.9)-(3.1.10) become

$$(\lambda_i + \lambda - A)Te_i = -BKe_i, \tag{3.1.11}$$

$$TBe_i = Be_i. \tag{3.1.12}$$

If $\lambda > 0$ is such that $\lambda_i + \lambda \neq \lambda_k$, $\forall\, 1 \leq i, k \leq n$, then $(\lambda_i + \lambda - A)$ is invertible. We begin by proving that $\{\widetilde{T}e_i\}_{1\leq i\leq n}$, satisfying

$$(\lambda_i + \lambda - A)\widetilde{T}e_i = -B \tag{3.1.13}$$

is a basis of $\mathbb{R}^n$. We proceed in this way since the transformation K is also to be determined. Note that $Te_i = (\widetilde{T}e_i)Ke_i$, where Ke_i is a scalar.

We prove that $\{\widetilde{T}e_i\}_{1\leq i\leq n}$ is a basis of $\mathbb{R}^n$ by contradiction. Suppose there exists $\{a_i\}_{1\leq i\leq n} \neq 0$ such that

$$\sum_{i=1}^{n} a_i\widetilde{T}e_i = 0.$$

Applying A to this equation and using (3.1.13), we obtain

$$\sum_{i=1}^{n} a_i(\lambda_i + \lambda)\widetilde{T}e_i = -\left(\sum_{i=1}^{n} a_i\right) B.$$

Here we also assume that $\lambda_i + \lambda \neq 0$. Applying successively A, we end up, for each $p \in \mathbb{N}^*$, with

$$\sum_{i=1}^{n} a_i(\lambda_i + \lambda)^p\widetilde{T}e_i = -\sum_{k=1}^{p}\left(\sum_{i=1}^{n} a_i(\lambda_i + \lambda)^{p-k}\right) A^{k-1}B. \tag{3.1.14}$$

Note that, for all $j \in \mathbb{N}$, each coefficient

$$\sum_{i=1}^{n} a_i(\lambda_i + \lambda)^j, \tag{3.1.15}$$

appears in (3.1.14) for all $p \geq j + 1$ in front of $A^{p-j-1}B$. Therefore, suppose there exists $j \in \mathbb{N}$ such that (3.1.15) is not equal to zero. It implies that $\{A^{k-1}B\}_{k\in\mathbb{N}^*} \subset \text{span}\{\widetilde{T}e_i\}$. Or, the controllability assumption implies that $\{A^{k-1}B\}_{k\in\mathbb{N}^*} = \mathbb{R}^n$. Since

we have assumed that $\{\widetilde{T}e_i\}_{1\leq i\leq n}$ is not a basis, it implies that

$$\sum_{i=1}^{n} a_i(\lambda_i+\lambda)^j Ke_i = 0, \tag{3.1.16}$$

for all $j \in \mathbb{N}$. In this case, consider entire function

$$G : z \in \mathbb{C} \mapsto \sum_{i=1}^{n} a_i e^{(\lambda_i+\lambda)z}.$$

From (3.1.16), we obtain for all $j \in \mathbb{N}$,

$$G^{(j)}(0) = \sum_{i=1}^{n} a_i(\lambda_i+\lambda)^j = 0.$$

Therefore $G \equiv 0$. Let

$$i_0 := \operatorname{argmax}\{\Re(\lambda_i)\,;\, 1 \leq i \leq n \text{ such that } a_i \neq 0\}$$

Then

$$a_{i_0} + \sum_{\substack{i=1\\ i\neq i_0}}^{n} a_i e^{(\lambda_i-\lambda_{i_0})z} = 0.$$

By considering z real and letting z go to infinity, we obtain that $a_{i_0} = 0$, which is in contradiction with the definition of i_0. Therefore $a_i = 0$ for all $1 \leq i \leq n$, which proves that $\{\widetilde{T}e_i\}_{1\leq i\leq n}$ is a basis of $\mathbb{R}^n$.

The transformation is then obtained using (3.1.12). Indeed, let

$$B = \sum_{i=1}^{n} b_i e_i.$$

Then,

$$\begin{aligned} & TB = B \\ \iff \quad & T\sum_{i=1}^{n} b_i e_i = B \\ \iff \quad & \sum_{i=1}^{n} b_i Te_i = B \\ \iff \quad & \sum_{i=1}^{n} b_i Ke_i \widetilde{T}e_i = B. \end{aligned} \tag{3.1.17}$$

Since $\{\widetilde{T}e_i\}_{1\leq i\leq n}$ is a basis of $\mathbb{R}^n$, there exists $\{Ke_i\}_{1\leq i\leq n}$ such that the last equation is verified, allowing to define T and K. Note that one can ensure that the transformations are real-valued since $\overline{T}$ and $\overline{K}$ is also a solution of this equation.

We finally prove that T is invertible by showing that Ker $T^* = \{0\}$. Otherwise, let $x \in$ Ker T^*. Using (3.1.11)-(3.1.12) to define T^* and K^*, we obtain

$$T^*A^*x = A^*T^*x + K^*B^*T^*x + \lambda T^*x = 0.$$

Therefore $A^*x \in$ Ker T^*. Repeating the same computations for $(A^*)^k x$ shows that $(A^*)^{k+1}x \in$ Ker T^* for all $k \in \mathbb{N}$. Or,

$$B^*(A^*)^k x = B^*T^*(A^*)^k x = 0, \quad \forall k \in \mathbb{N},$$

which, by observability, implies $x = 0$.

If the functional setting in the infinite dimensional case complicate the proof, the strategy remains the same. Riesz basis results will be used to prove the existence of a basis of the state space and the invertibility of the transformation will be proved using the observability of the adjoint equation.

The main technical difficulty of this paper lies in the decomposition of B (3.1.17) in the basis of the state space. Indeed, the control operator B is admissible but not bounded in the state space. A careful analysis of the Fourier components of the control operator B allow us to define a transformation T which is bounded from the state space into itself but the feedback transformation won't be bounded from the state space into $\mathbb{R}$. Even so, the transformation T will be proved to be invertible and the closed-loop linear equation will be proved to be well-posed in the state space. It is important to note that this technical difficulty is in fact essential for the invertibility of T. Indeed, in our case, if B were to be bounded, then the transformation T would be compact and thus not continuously invertible. However, the unboundedness of K from the state space into $\mathbb{R}$ prevent us to prove the well-posedness of the nonlinear equation.

Let us underline that the uniqueness condition $TB = B$, which was used implicitly in similar previous works, will be crucial not only to obtain the existence and uniqueness of the transformation, but also to deal with the analogous of (3.1.9).

3.1.3 The linear Schrödinger equation

The strategy to prove the rapid stabilization of the linear equation (3.1.4) is inspired by the backstepping method. Recast the equations for $\Psi^1 + i\Psi^2 = \Psi$ as

$$\begin{cases} \partial_t \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = \begin{pmatrix} 0 & -\Delta \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} + v(t) \begin{pmatrix} 0 \\ (\mu\varphi_1)(x) \end{pmatrix}, & (t,x) \in (0,1) \times (0,T), \\ \Psi^1(t,0) = \Psi^1(t,1) = 0, \quad \Psi^2(t,0) = \Psi^2(t,1) = 0, & t \in (0,T), \\ \Psi^1(0,x) = \Psi_0^1(x), \quad \Psi^2(0,x) = \Psi_0^2(x), & x \in (0,1). \end{cases} \tag{3.1.18}$$

where Ψ_0^1 and Ψ_0^2 are the real and imaginary part of Ψ^0 respectively. Based on the previous work of the first author and Q. Lü ([50, 51]), instead of Volterra transformations of the second kind of the backstepping approach, we seek for transformations (T, K) of the form

$$\begin{aligned} T : H^3_{(0)} \times H^3_{(0)} &\longrightarrow H^3_{(0)} \times H^3_{(0)} \\ \begin{pmatrix} \Psi^1(t,\cdot) \\ \Psi^2(t,\cdot) \end{pmatrix} &\longmapsto \int_0^1 \begin{pmatrix} k_{11}(x,y) & k_{12}(x,y) \\ k_{21}(x,y) & k_{22}(x,y) \end{pmatrix} \begin{pmatrix} \Psi^1(t,y) \\ \Psi^2(t,y) \end{pmatrix} dy, \end{aligned} \tag{3.1.19}$$

$$\begin{aligned} K : H^3_{(0)} \times H^3_{(0)} &\longrightarrow \mathbb{R} \\ \begin{pmatrix} \Psi^1(t,\cdot) \\ \Psi^2(t,\cdot) \end{pmatrix} &\longmapsto \int_0^1 \alpha^1(y)\Psi^1(t,y) + \alpha^2(y)\Psi^2(t,y)dy, \end{aligned} \tag{3.1.20}$$

such that $\begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} := T \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix}$ is solution to

$$\begin{cases} \partial_t \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} = \begin{pmatrix} 0 & -\Delta \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} - \lambda \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}, & (t,x) \in (0,1) \times (0,T), \\ \xi^1(t,0) = \xi^1(t,1) = 0, \quad \xi^2(t,0) = \xi^2(t,1) = 0, & t \in \times(0,T), \\ \xi^1(0,x) = \xi_0^1(x), \quad \xi^2(0,x) = \xi_0^2(x), & x \in \times(0,1), \end{cases} \tag{3.1.21}$$

with $\begin{pmatrix} \xi_0^1 \\ \xi_0^2 \end{pmatrix} = T \begin{pmatrix} \Psi_0^1 \\ \Psi_0^2 \end{pmatrix}$ and T is invertible in the state space. The decomposition in real and imaginary part of the solution of (3.1.4) is made in order to ensure that the feedback $v(t) = K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix}$ is real-valued. From now, we will thus consider couples of functions in $H^3_{(0)}$ that are real-valued. For the sake of simplicity, we set, for $s \geq 0$,

$$X^s_{(0)} := \left(H^s_{(0)} \cap L^2((0,1);\mathbb{R})\right)^2, \tag{3.1.22}$$

with the product topology. Note that

$$\left\| \begin{pmatrix} \xi^1(t) \\ \xi^2(t) \end{pmatrix} \right\|_{X^3_{(0)}} \leq e^{-\lambda t} \left\| \begin{pmatrix} \xi^1(t) \\ \xi^2(t) \end{pmatrix} \right\|_{X^3_{(0)}}, \quad t \in [0, +\infty). \tag{3.1.23}$$

The kernels are defined through the equations they must satisfy for (T, K) to map solutions of (3.1.18) to solutions of (3.1.21). This is done formally in Section 3.2 together with a more detailed presentation of this strategy.

3.1.4 A brief review of previous results

The controllability properties for the Schrödinger equation were mostly studied in the usual (in opposition to the bilinear model presented here) linear setting. For the control of the linear Schrödinger equation with internal control (localized on a subdomain), we refer to the survey [91] and the references therein. In this more classical setting we also mention [99, 90, 53] concerning stabilization. For rapid stabilization, using a backstepping approach, of a linear Schrödinger equation but with a boundary control we refer to [87].

Exact controllability of the bilinear Schrödinger equation.
Following the general result [3], it is proved in [127] that the complement of the reachable set, from any initial condition in $\mathcal{S} \cap H^2_{(0)}$, with L^2 controls is dense in $\mathcal{S} \cap H^2_{(0)}$. This prevents exact controllability in this space.

In fact, this lack of controllability is mainly due to the functional setting used as already mentioned in Theorem 3.1.1 from [7] (see also [5, 6] for previous results). This local controllability result has been extended with weaker assumptions but with a minimal control time [10], in a more general setting in infinite time [110] and also in the case of simultaneous controllability of a finite number of particles [103, 104]. More recently, local exact controllability has been established in [8] for a Schrödinger-Poisson model in 2D (see also [101] for approximate controllability) and for the analogous of (3.1.1) with a control depending on time and space in dimension less or equal than 3 [115].

Approximate controllability and stabilization of the bilinear Schrödinger equation.
The above mentioned results of exact controllability are mostly limited to the one dimensional case. In a more general setting, the available results deal with approximate controllability. Using geometric control techniques on appropriate Galerkin approximations, approximate controllability in different settings has been proved [22, 18, 17, 31]. For a detailed presentation, see the survey [19].

However, most of these results are not suitable to prove approximate controllability in higher norms (typically $H^3_{(0)}$) and thus approximate controllability for bilinear Schrödinger equations has also been studied from the Lyapunov functional point of view [102, 9, 108, 109]. Though it enabled global controllability results, this strategy usually gives no indication on the decay rate.

Rapid stabilization.
The strategy used in this article is inspired from backstepping techniques. Initially developed to design, in a recursive way, more effective feedback laws for globally asymptotic stable finite dimensional system for which a feedback law and a Lyapunov function are already known (we refer to [88, 123, 42] for a comprehensive introduction in finite dimension and [46, 98] for an application of this method to partial differential equations), the backstepping approach was later used in the infinite dimensional setting to design feedback laws by mapping the system to stabilize to a target stable system. To our knowledge, this strategy was first introduced in the context of partial differential equations to design a feedback law for heat equation [4] and, later on, for a class of parabolic PDE [20]. The backstepping-like change of coordinates of the semi-discretized equation maps the discrete solution to stabilize to a stable solution. The corresponding continuous mapping obtained is a Volterra transformation of the second kind, seen as an infinite dimensional backstepping transformation from the triangular domain of definition of the transformation. This backstepping strategy in infinite dimensions led to a series of work (see [89] for a global presentation). The backstepping approach is not limited to linear equations, as shown in [29, 52]. Though using Volterra transformations of the second kind provides easily invertible transformations it is also limited. Fredholm transformations, from which this paper is inspired, has already been used [50, 51, 49] for rapid stabilization using boundary feedback laws.

3.1.5 Structure of the article

In Sec. 3.2, we give a detailed presentation of the strategy used to construct the transformation (T, K) and give a formal expression of this transformation. In Sec. 3.3 we prove that this formal transformation T is well defined and is continuous in the state space $X^3_{(0)}$. Then, we prove in Sec. 3.4 that the previous transformation is indeed invertible in the state space. These properties of T will follow using Hypothesis 3.1.1 i.e. exact controllability of the linearized system. We end the proof of Theorem 3.1.3 in Sec. 3.5 by proving that the constructed feedback K leads to a well-posed system and that T actually maps the solution of (3.1.18) to the exponentially stable solutions of (3.1.21). In Appendix A.1 we study in a similar way a simplified Saint-Venant equation which exhibits the same phenomenon but on which we explicitly compute the transformation (T, K).

3.2 Heuristic construction of the transformations

We recall that we look for a transformation (T, K) of the form (3.1.19)- (3.1.20). Let us derive the set of equations for (T, K) to map solutions of (3.1.18) to solutions of (3.1.21). First, to ensure that the boundary conditions of (3.1.21) are satisfied, we assume that

$$k_{ij}(0, y) = k_{ij}(1, y) = 0, \quad \forall y \in (0, 1), \quad \forall i, j \in \{1, 2\}. \tag{3.2.1}$$

Using the fact that $\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)}$, and imposing the conditions

$$k_{ij}(x, 0) = k_{ij}(x, 1) = 0, \quad \forall x \in (0, 1), \quad \forall i, j \in \{1, 2\}, \tag{3.2.2}$$

formal computations, denoting Δ_x and Δ_y the Dirichlet Laplacian with respect to x and y variables respectively, lead to

$$\begin{aligned}
&\partial_t \xi^1(t, x) + \Delta \xi^2(t, x) + \lambda \xi^1(t, x) \\
&= \int_0^1 k_{11}(x, y) \left(-\Delta_y \Psi^2(t, y)\right) + k_{12}(x, y) \left(\Delta_y \Psi^1(t, y) + v(t)(\mu\varphi_1)(y)\right) dy \\
&+ \int_0^1 \Delta_x k_{21}(x, y) \Psi^1(t, y) + \Delta_x k_{22}(x, y) \Psi^2(t, y) dy \\
&+ \int_0^1 \lambda k_{11}(x, y) \Psi^1(t, y) + \lambda k_{12}(x, y) \Psi^2(t, y) dy \\
&= \int_0^1 \left(\Delta_x k_{21} + \Delta_y k_{12} + \lambda k_{11}\right)(x, y) \Psi^1(t, y) dy \\
&+ \int_0^1 \left(\Delta_x k_{22} - \Delta_y k_{11} + \lambda k_{12}\right)(x, y) \Psi^2(t, y) dy \\
&+ v(t) \int_0^1 (\mu\varphi_1)(z) k_{12}(x, z) dz.
\end{aligned}$$

The boundary conditions (3.2.2) were imposed to avoid boundary terms in the integrations by parts. Using the expression of the feedback leads to

$$\begin{aligned}
&\partial_t \xi^1(t, x) + \Delta \xi^2(t, x) + \lambda \xi^1(t, x) \\
&= \int_0^1 \left[\left(\Delta_y k_{12} + \Delta_x k_{21} + \lambda k_{11}\right)(x, y) + \alpha^1(y) \left(\int_0^1 (\mu\varphi_1)(z) k_{12}(x, z) dz\right)\right] \Psi^1(t, y) dy \\
&- \int_0^1 \left[\left(\Delta_y k_{11} - \Delta_x k_{22} - \lambda k_{12}\right)(x, y) - \alpha^2(y) \left(\int_0^1 (\mu\varphi_1)(z) k_{12}(x, z) dz\right)\right] \Psi^2(t, y) dy.
\end{aligned} \tag{3.2.3}$$

In the same way one gets

$$\begin{aligned}
&\partial_t \xi^2(t,x) - \Delta \xi^1(t,x) + \lambda \xi^2(t,x) \\
&= \int_0^1 \left[\left(\Delta_y k_{22} - \Delta_x k_{11} + \lambda k_{21}\right)(x,y) + \alpha^1(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{22}(x,z) dz \right) \right] \Psi^1(t,y) dy \\
&- \int_0^1 \left[\left(\Delta_y k_{21} + \Delta_x k_{12} - \lambda k_{22}\right)(x,y) - \alpha^2(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{22}(x,z) dz \right) \right] \Psi^2(t,y) dy.
\end{aligned} \tag{3.2.4}$$

If we want (ξ^1, ξ^2) to satisfy (3.1.21) then we need to find the functions k_{ij} and α^j satisfying

$$\begin{cases}
\left(\Delta_y k_{11} - \Delta_x k_{22} - \lambda k_{12}\right)(x,y) - \alpha^2(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{12}(x,z) dz \right) = 0, \ (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{12} + \Delta_x k_{21} + \lambda k_{11}\right)(x,y) + \alpha^1(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{12}(x,z) dz \right) = 0, \ (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{21} + \Delta_x k_{12} - \lambda k_{22}\right)(x,y) - \alpha^2(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{22}(x,z) dz \right) = 0, \ (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{22} - \Delta_x k_{11} + \lambda k_{21}\right)(x,y) + \alpha^1(y) \left(\int_0^1 \left(\mu\varphi_1\right)(z) k_{22}(x,z) dz \right) = 0, \ (x,y) \in (0,1)^2, \\
k_{ij}(x,0) = k_{ij}(x,1) = 0, \quad x \in (0,1), \\
k_{ij}(0,y) = k_{ij}(1,y) = 0, \quad y \in (0,1).
\end{cases} \tag{3.2.5}$$

A fundamental extra condition. One could try to solve rightaway (3.2.5) and prove the invertibility of the transformation T but the non-local terms yield a tedious task. To overcome this difficulty, one introduces, as in the finite dimensional setting, what will be referred throughout this article as the $TB = B$ condition,

$$T \begin{pmatrix} 0 \\ (\mu\varphi_1)(x) \end{pmatrix} = \int_0^1 \begin{pmatrix} k_{12}(x,y)(\mu\varphi_1)(y) \\ k_{22}(x,y)(\mu\varphi_1)(y) \end{pmatrix} dy = \begin{pmatrix} 0 \\ (\mu\varphi_1)(x) \end{pmatrix}.$$

Plugging this into (3.2.5) we obtain that we now seek for a solution to

$$\begin{cases}
\left(\Delta_y k_{11} - \Delta_x k_{22} - \lambda k_{12}\right)(x,y) = 0, & (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{12} + \Delta_x k_{21} + \lambda k_{11}\right)(x,y) = 0, & (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{21} + \Delta_x k_{12} - \lambda k_{22}\right)(x,y) - \alpha^2(y)(\mu\varphi_1)(x) = 0, & (x,y) \in (0,1)^2, \\
\left(\Delta_y k_{22} - \Delta_x k_{11} + \lambda k_{21}\right)(x,y) + \alpha^1(y)(\mu\varphi_1)(x) = 0, & (x,y) \in (0,1)^2, \\
k_{ij}(x,0) = k_{ij}(x,1) = 0, & x \in (0,1), \\
k_{ij}(0,y) = k_{ij}(1,y) = 0, & y \in (0,1),
\end{cases} \tag{3.2.6}$$

together to the $TB = B$ condition

$$\begin{cases} \displaystyle\int_0^1 k_{12}(x,y)(\mu\varphi_1)(y)dy = 0, & x \in (0,1), \\ \displaystyle\int_0^1 k_{22}(x,y)(\mu\varphi_1)(y)dy = \mu\varphi_1, & x \in (0,1). \end{cases} \tag{3.2.7}$$

Remark 3.2.1 *In [50, 51], the authors were dealing with a boundary control for the Korteweg-de Vries equation and for the Kuramoto-Sivashinsky equation. In their case, the $TB = B$ condition writes*

$$k_y(x,L) = 0, \quad x \in (0,L),$$

for the former and

$$k_{yy}(x,L) = 0, \quad x \in (0,L),$$

for the latter, where k is the kernel of the Fredholm transformation in each case. Contrary to our framework, this boundary condition appeared naturally from the integration by parts performed in order to obtain the equation on the kernel. It was not seen as a particular boundary condition although, a careful analysis of their work shows that the $TB = B$ condition was used to prove the uniqueness of the transformation (T,K). The common ground between their work and this article is the additional regularity that the kernel needs in order to satisfy the $TB = B$ condition. Notice that in what we present in this article, the relation between the kernels k_{ij} and α^j is more intricate and considerably modify the analysis.

Formal decomposition The global strategy to construct a solution of (3.2.6)-(3.2.7) is the following. First assume that α^1 and α^2 are known. This enables us to compute the kernels k_{ij} satisfying (3.2.6) as functions of α^1 and α^2. Then we prove that we find α^1 and α^2 such that (3.2.7) is satisfied.

We decompose the functions in the following form

$$k_{ij}(x,y) = \sum_{n=1}^{+\infty} f_n^{ij}(x)\varphi_n(y), \qquad \alpha^j(y) = \sum_{n=1}^{+\infty} \alpha_n^j \varphi_n(y). \tag{3.2.8}$$

This leads to

$$\begin{cases} f_n^{11\prime\prime}(x) + \lambda_n f_n^{22}(x) - \lambda f_n^{21}(x) = -\alpha_n^1(\mu\varphi_1)(x), \\ f_n^{12\prime\prime}(x) - \lambda_n f_n^{21}(x) - \lambda f_n^{22}(x) = \alpha_n^2(\mu\varphi_1)(x), \\ f_n^{21\prime\prime}(x) - \lambda_n f_n^{12}(x) + \lambda f_n^{11}(x) = 0, \\ f_n^{22\prime\prime}(x) + \lambda_n f_n^{11}(x) + \lambda f_n^{12}(x) = 0, \\ f_n^{ij}(0) = f_n^{ij}(1) = 0. \end{cases} \tag{3.2.9}$$

The $TB = B$ condition (3.2.7) becomes

$$\begin{cases} \displaystyle\sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle f_n^{12}(x) = 0, \\ \displaystyle\sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle f_n^{22}(x) = (\mu\varphi_1)(x). \end{cases} \tag{3.2.10}$$

As mentioned, we begin by assuming that the feedback law is known. We consider two sequences $(\beta_n^1)_{n\in\mathbb{N}^*}$ and $(\beta_n^2)_{n\in\mathbb{N}^*}$ to be precised later on.

Let $(g_n^{11}, g_n^{12}, g_n^{21}, g_n^{22})^t$ be the solution of system (3.2.9) with right-hand side $(-\beta_n^1(\mu\varphi_1), 0, 0, 0)^t$ and let $(h_n^{11}, h_n^{12}, h_n^{21}, h_n^{22})^t$ be the solution of system (3.2.9) with right-hand side $(0, \beta_n^2(\mu\varphi_1), 0, 0)^t$. System (3.2.9) being linear it comes that

$$f_n^{ij} = \frac{\alpha_n^1}{\beta_n^1} g_n^{ij} + \frac{\alpha_n^2}{\beta_n^2} h_n^{ij}. \tag{3.2.11}$$

Decomposing g_n^{ij} in Fourier series, if we denote by A_{nk} the following matrix

$$A_{nk} = \begin{pmatrix} -\lambda_k & 0 & -\lambda & \lambda_n \\ 0 & -\lambda_k & -\lambda_n & -\lambda \\ \lambda & -\lambda_n & -\lambda_k & 0 \\ \lambda_n & \lambda & 0 & -\lambda_k \end{pmatrix}, \tag{3.2.12}$$

system (3.2.9) leads to

$$A_{nk} \begin{pmatrix} \langle g_n^{11}, \varphi_k \rangle \\ \langle g_n^{12}, \varphi_k \rangle \\ \langle g_n^{21}, \varphi_k \rangle \\ \langle g_n^{22}, \varphi_k \rangle \end{pmatrix} = \begin{pmatrix} -\beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Then

$$\begin{cases} \displaystyle g_n^{11}(x) = \sum_{k=1}^{+\infty} \frac{\lambda_k \left(\lambda^2 + \lambda_k^2 - \lambda_n^2\right)}{\delta_{nk}(\lambda)} \beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle \varphi_k(x), \\ \displaystyle g_n^{12}(x) = \sum_{k=1}^{+\infty} \frac{-2\lambda\lambda_k\lambda_n}{\delta_{nk}(\lambda)} \beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle \varphi_k(x), \\ \displaystyle g_n^{21}(x) = \sum_{k=1}^{+\infty} \frac{\lambda \left(\lambda^2 + \lambda_k^2 + \lambda_n^2\right)}{\delta_{nk}(\lambda)} \beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle \varphi_k(x), \\ \displaystyle g_n^{22}(x) = \sum_{k=1}^{+\infty} \frac{-\lambda_n \left(\lambda^2 - \lambda_k^2 + \lambda_n^2\right)}{\delta_{nk}(\lambda)} \beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle \varphi_k(x), \end{cases} \tag{3.2.13}$$

where

$$\delta_{nk}(\lambda) = \det(A_{nk}) = \left(\lambda^2 + (\lambda_k - \lambda_n)^2\right)\left(\lambda^2 + (\lambda_k + \lambda_n)^2\right). \tag{3.2.14}$$

The same computations can be carried out for h_n^{ij}, leading to the following relations

$$\begin{aligned} h_n^{11} &= \frac{\beta_n^2}{\beta_n^1} g_n^{12}, \qquad & h_n^{12} &= -\frac{\beta_n^2}{\beta_n^1} g_n^{11}, \\ h_n^{21} &= \frac{\beta_n^2}{\beta_n^1} g_n^{22}, \qquad & h_n^{22} &= -\frac{\beta_n^2}{\beta_n^1} g_n^{21}. \end{aligned} \tag{3.2.15}$$

For the sake of simplicity, we denote by c_{nk}^{ij} and d_{nk}^{ij} the coefficients such that

$$g_n^{ij}(x) = \sum_{k=1}^{+\infty} c_{nk}^{ij} \beta_n^1 \langle \mu\varphi_1, \varphi_k\rangle \varphi_k(x), \quad h_n^{ij}(x) = \sum_{k=1}^{+\infty} d_{nk}^{ij} \beta_n^2 \langle \mu\varphi_1, \varphi_k\rangle \varphi_k(x). \tag{3.2.16}$$

Summary of the construction. Finally, using the definition of the transformation (3.1.19), the decompositions (3.2.8), (3.2.11) and the relations (3.2.15) it comes that the transformation T we are looking for is defined by

$$\begin{aligned} T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} &= \begin{pmatrix} \displaystyle\sum_{n=1}^{+\infty}\left(\frac{\alpha_n^1}{\beta_n^1} g_n^{11} + \frac{\alpha_n^2}{\beta_n^2} h_n^{11}\right)\langle \Psi^1, \varphi_n\rangle + \sum_{n=1}^{+\infty}\left(\frac{\alpha_n^1}{\beta_n^1} g_n^{12} + \frac{\alpha_n^2}{\beta_n^2} h_n^{12}\right)\langle \Psi^2, \varphi_n\rangle \\ \displaystyle\sum_{n=1}^{+\infty}\left(\frac{\alpha_n^1}{\beta_n^1} g_n^{21} + \frac{\alpha_n^2}{\beta_n^2} h_n^{21}\right)\langle \Psi^1, \varphi_n\rangle + \sum_{n=1}^{+\infty}\left(\frac{\alpha_n^1}{\beta_n^1} g_n^{22} + \frac{\alpha_n^2}{\beta_n^2} h_n^{22}\right)\langle \Psi^2, \varphi_n\rangle \end{pmatrix} \\ &= \sum_{n=1}^{+\infty}\left[\frac{\alpha_n^1\langle \Psi^2, \varphi_n\rangle}{\beta_n^1} - \frac{\alpha_n^2\langle \Psi^1, \varphi_n\rangle}{\beta_n^1}\right]\begin{pmatrix} g_n^{12}\\ g_n^{22}\end{pmatrix} + \left[\frac{\alpha_n^1\langle \Psi^1, \varphi_n\rangle}{\beta_n^2} + \frac{\alpha_n^2\langle \Psi^2, \varphi_n\rangle}{\beta_n^2}\right]\begin{pmatrix} h_n^{12}\\ h_n^{22}\end{pmatrix}. \end{aligned} \tag{3.2.17}$$

In the same spirit, the $TB = B$ condition (3.2.10) becomes

$$\begin{pmatrix} 0\\ \mu\varphi_1\end{pmatrix} = \sum_{n=1}^{+\infty} \frac{\alpha_n^1\langle \mu\varphi_1, \varphi_n\rangle}{\beta_n^1}\begin{pmatrix} g_n^{12}\\ g_n^{22}\end{pmatrix} + \frac{\alpha_n^2\langle \mu\varphi_1, \varphi_n\rangle}{\beta_n^2}\begin{pmatrix} h_n^{12}\\ h_n^{22}\end{pmatrix}. \tag{3.2.18}$$

This ends the heuristic of the construction of the transformation and the feedback law. Indeed, in the following section we prove that for a suitable choice of the sequences $(\beta_n^1)_{n\in\mathbb{N}^*}$ and $(\beta_n^2)_{n\in\mathbb{N}^*}$ then

$$\mathcal{B} := \left\{\begin{pmatrix} g_n^{12}\\ g_n^{22}\end{pmatrix}, \begin{pmatrix} h_n^{12}\\ h_n^{22}\end{pmatrix} \ ; \ n \in \mathbb{N}^*\right\}, \tag{3.2.19}$$

is a Riesz basis of $X^2_{(0)}$ (see Proposition 3.3.1). Then from (3.2.18) we get the feedback laws from the expansion of $\begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix}$ in the basis $\mathcal{B}$. Finally, we prove that coefficients α^1_n and α^2_n decay sufficiently rapidly as n goes to infinity so that the transformation T given by (3.2.17) is indeed continuous from $X^3_{(0)}$ to $X^3_{(0)}$.

Remark 3.2.2 *From (3.2.18) it already appears that the behaviour of the coefficients $\langle \mu\varphi_1, \varphi_n \rangle$ will play a crucial role.*

3.3 Definition and properties of the transformation

In this section, we make rigorous the heuristic developed in the previous section. In subsection 3.3.1, we prove that for a suitable choice of β^1_n and β^2_n then $\mathcal{B}$ defined in (3.2.19) is a Riesz basis of $X^2_{(0)}$ where the functions g^{ij}_n and h^{ij}_n are defined by (3.2.13) and (3.2.15). This enables us to define the feedback law and the transformation T using the relations (3.2.18) and (3.2.17). However, this does not give enough regularity to prove that $T : X^3_{(0)} \to X^3_{(0)}$. We prove the extra regularity we need on the feedback laws in subsection 3.3.2. This leads, in subsection 3.3.3, to the expected regularity for the transformation T.

3.3.1 Riesz basis property

Let us recall some results on Riesz basis.

Definition 3.3.1 *Let H be an Hilbert space and $\{g_n\}_{n\in\mathbb{N}^*} \subset H$. We say that $\{g_n\}_{n\in\mathbb{N}^*}$ is ω-independent if*

$$\sum_{n\in\mathbb{N}^*} a_n g_n = 0, \text{ with } \{a_n\}_{n\in\mathbb{N}^*} \subset \mathbb{R} \quad \Longrightarrow \quad a_n = 0, \forall n \in \mathbb{N}^*.$$

Theorem 3.3.1 *[132, Theorem 15] Let H be a separable Hilbert space and let $\{e_n\}_{n\in\mathbb{N}^*}$ be an orthonormal basis for H. If $\{g_n\}_{n\in\mathbb{N}^*}$ is an ω-independent sequence quadratically close to $\{e_n\}_{n\in\mathbb{N}^*}$, that is*

$$\sum_{n\in\mathbb{N}^*} \|e_n - g_n\|^2_H < +\infty,$$

then $\{g_n\}_{n\in\mathbb{N}^*}$ is a Riesz basis for H.

Theorem 3.3.2 *Let H be a separable Hilbert space and let $\{e_n\}_{n\in\mathbb{N}^*}$ be an orthonormal basis for H. If $\{g_n\}_{n\in\mathbb{N}^*}$ is dense in H and is quadratically close to $\{e_n\}_{n\in\mathbb{N}^*}$, i.e.*

$$\sum_{n\in\mathbb{N}^*} \|e_n - g_n\|_H^2 < +\infty,$$

then $\{g_n\}_{n\in\mathbb{N}^*}$ is a Riesz basis for H.

Let us provide a proof of Theorem 3.3.2, stated as a remark in [75, Remark 2.1, p. 318].

Proof :

Let us prove Theorem 3.3.2 by contradiction. Suppose that the $\{g_n\}_{n\in\mathbb{N}^*}$ is dense in H and is quadratically close to $\{e_n\}_{n\in\mathbb{N}^*}$ but that $\{g_n\}_{n\in\mathbb{N}^*}$ is not a Riesz basis. Therefore, by Theorem 3.3.1, there must exist a non-zero sequence $\{a_n\}_{n\in\mathbb{N}^*} \subset \mathbb{R}$ such that

$$\sum_{n\in\mathbb{N}^*} a_n g_n = 0.$$

Since $\{g_n\}_{n\in\mathbb{N}^*}$ is quadratically close to $\{e_n\}_{n\in\mathbb{N}^*}$, there exists $N \in \mathbb{N}^*$ such that

$$\sum_{n\geq N+1} \|e_n - g_n\|_H^2 < 1.$$

Therefore, from [132, Theorem 13], the $\{e_n\}_{1\leq n\leq N} \cup \{g_n\}_{n\geq N+1}$ is a Riesz basis of H. This implies that there exists $k \leq N$ such that $a_k \neq 0$. However, with the density assumption, we have

$$H = \overline{\text{span}\{g_n \,|\, n \in \mathbb{N}^*\}} = \overline{\text{span}\{g_n \,|\, n \in \mathbb{N}^* \setminus \{k\}\}}.$$

Thus,

$$\text{codim}(\text{span}\{g_n \,|\, n \geq N+1\}) \leq N-1.$$

However

$$H/\overline{\text{span}\{g_n \,|\, n \geq N+1\}},$$

is isomorphic to $\text{span}\{e_n \,|\, n \leq N\}$, which is of dimension N, leading to a contradiction.

□

We will use the previous criterions to prove the following proposition.

Proposition 3.3.1 *Assume that*

$$\lambda \notin \mathcal{N} := \{\lambda_n; n \in \mathbb{N}^*\} \cup \left\{ \lambda \in \mathbb{R}; \exists (k,n) \in (\mathbb{N}^*)^2 \text{ such that } \frac{\lambda_n^2 - \lambda^2}{(\lambda_n^2 + \lambda^2)^2} = \frac{\lambda_k^2 - \lambda^2}{(\lambda_k^2 + \lambda^2)^2} \right\}. \tag{3.3.1}$$

Then, with a suitable choice of β_n^i,

$$\mathcal{B} := \left\{ \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix}, \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} \; ; \; n \in \mathbb{N}^* \right\}$$

is a Riesz basis of $X^2_{(0)}$ *and*

$$\tilde{\mathcal{B}} := \left\{ \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix}, \begin{pmatrix} h_n^{12}/\lambda_n^{1/2} \\ h_n^{22}/\lambda_n^{1/2} \end{pmatrix} \; ; \; n \in \mathbb{N}^* \right\}$$

is a Riesz basis of $X^3_{(0)}$.

Remark 3.3.1 *Assumption* $\lambda \notin \mathcal{N}$ *is a technical assumption in order to prove the existence of the Riesz basis* $\mathcal{B}$ *and* $\tilde{\mathcal{B}}$. *As we shall see in the following sections, the transformations* (T, K) *are constructed with these Riesz basis and, therefore, will only be defined for* $\lambda \notin \mathcal{N}$.

However, as $\mathcal{N}$ *is a countable set, this does not modify the statement of Theorem 3.1.3. Indeed if* $\lambda \in \mathcal{N}$, *one can choose* $\epsilon > 0$ *such that* $\lambda + \epsilon \notin \mathcal{N}$ *and obtain an exponential stabilization of* (3.1.4) *with an exponential decay rate at least equal to* $\lambda + \epsilon$ *and thus at least equal to* λ. *In all what follows we assume that* $\lambda \notin \mathcal{N}$ *even when not explicitly specified.*

To apply the previous criterion for the Riesz basis, we prove that $\mathcal{B}$ (resp. $\tilde{\mathcal{B}}$) is quadratically close to the orthonormal basis of $X^2_{(0)}$ (resp. $X^3_{(0)}$) given by

$$\left\{ \begin{pmatrix} \varphi_n/\lambda_n^{s/2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \varphi_n/\lambda_n^{s/2} \end{pmatrix} \; ; \; n \in \mathbb{N}^* \right\}, \quad s \in \{2, 3\}.$$

Thus, we choose β_n^1 and β_n^2 such that g_n^{12} and h_n^{22} are close to $\frac{\varphi_n}{\lambda_n}$ i.e. $\langle g_n^{12}, \varphi_n \rangle = \langle h_n^{22}, \varphi_n \rangle = \frac{1}{\lambda_n}$, that is,

$$\begin{aligned} \beta_n^1 &:= -\frac{\lambda(\lambda^2 + 4\lambda_n^2)}{2\lambda_n^3 \langle \mu\varphi_1, \varphi_n \rangle}, \\ \beta_n^2 &:= -\frac{\lambda(\lambda^2 + 4\lambda_n^2)}{(\lambda^2 + 2\lambda_n^2)\lambda_n \langle \mu\varphi_1, \varphi_n \rangle}. \end{aligned} \tag{3.3.2}$$

Notice that, from Hypothesis 3.1.1 and Remark 3.1.1, there exists $C > 0$ such that

$$\frac{n}{C} \leq -\beta_n^1 \leq Cn, \quad \frac{n}{C} \leq -\beta_n^2 \leq Cn, \quad \forall n \in \mathbb{N}^*. \tag{3.3.3}$$

During the proof of Proposition 3.3.1, we will use the following lemma. Its proof is purely technical and postponed to Appendix A.2.

Lemma 3.3.1 *With the above definition, for $s = 2$ and $s = 3$, one gets,*

$$\sum_{n\in\mathbb{N}^*} \left\| \begin{pmatrix} \varphi_n/\lambda_n^{s/2} \\ 0 \end{pmatrix} - \begin{pmatrix} g_n^{12}/\lambda_n^{(s-2)/2} \\ g_n^{22}/\lambda_n^{(s-2)/2} \end{pmatrix} \right\|^2_{X^s_{(0)}} < +\infty,$$

$$\sum_{n\in\mathbb{N}^*} \left\| \begin{pmatrix} 0 \\ \varphi_n/\lambda_n^{s/2} \end{pmatrix} - \begin{pmatrix} h_n^{12}/\lambda_n^{(s-2)/2} \\ h_n^{22}/\lambda_n^{(s-2)/2} \end{pmatrix} \right\|^2_{X^s_{(0)}} < +\infty.$$

We are now ready to prove Proposition 3.3.1.

Proof : Let $s = 2$ or $s = 3$. In view of Theorem 3.3.1 and Lemma 3.3.1, assume that there exists a sequence $\{a_n, b_n\}_{n\in\mathbb{N}^*} \subset (\ell^2(\mathbb{R}))^2$ such that

$$\sum_{n\in\mathbb{N}^*} a_n \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + b_n \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } X^s_{(0)}. \tag{3.3.4}$$

First step : we apply negative powers of the Laplacian to characterize elements of $S := \overline{\operatorname{span}\mathcal{B}}$.

Let us write (3.2.9) for g_n^{ij} as

$$\mathcal{A}g_n = \underbrace{\left[\lambda_n \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\right]}_{=:J_n} g_n + \underbrace{\begin{pmatrix} \beta_n^1\mu\varphi_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=:F_n},$$

where $g_n = (g_n^{11}, g_n^{12}, g_n^{21}, g_n^{22})^t$. Since $\det(J_n) = (\lambda^2 + \lambda_n^2)^2 \neq 0$, we have

$$\mathcal{A}^{-1}g_n = J_n^{-1}g_n - \mathcal{A}^{-1}J_n^{-1}F_n.$$

Therefore, using (3.2.15), we obtain

$$\begin{aligned}
\mathcal{A}^{-1}g_n^{12} &= \frac{1}{\lambda_n^2+\lambda^2}\left[\lambda g_n^{22} + \lambda_n\frac{\beta_n^1}{\beta_n^2}h_n^{22}\right], \\
\mathcal{A}^{-1}g_n^{22} &= \frac{-1}{\lambda_n^2+\lambda^2}\left[\lambda_n\frac{\beta_n^1}{\beta_n^2}h_n^{12} + \lambda g_n^{12} + \lambda_n\beta_n^1\mathcal{A}^{-1}(\mu\varphi_1)\right], \\
\mathcal{A}^{-1}h_n^{12} &= \frac{1}{\lambda_n^2+\lambda^2}\left[-\lambda_n\frac{\beta_n^2}{\beta_n^1}g_n^{22} + \lambda h_n^{22}\right], \\
\mathcal{A}^{-1}h_n^{22} &= \frac{1}{\lambda_n^2+\lambda^2}\left[\lambda_n\frac{\beta_n^2}{\beta_n^1}g_n^{12} - \lambda h_n^{12} - \lambda\beta_n^2\mathcal{A}^{-1}(\mu\varphi_1)\right].
\end{aligned} \tag{3.3.5}$$

Thus, applying $\mathcal{A}^{-1}$ to (3.3.4) leads to

$$\begin{pmatrix} c_0\mathcal{A}^{-1}(\mu\varphi_1) \\ 0 \end{pmatrix} = \sum_{n\in\mathbb{N}^*} \frac{-\lambda a_n + \lambda_n \frac{\beta_n^2}{\beta_n^1} b_n}{\lambda_n^2+\lambda^2} \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} - \frac{\lambda_n \frac{\beta_n^1}{\beta_n^2} a_n + \lambda b_n}{\lambda_n^2+\lambda^2} \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix}, \tag{3.3.6}$$

where

$$c_0 := \sum_{n\in\mathbb{N}^*} \frac{\lambda_n \beta_n^1 a_n + \lambda \beta_n^2 b_n}{\lambda_n^2+\lambda^2}. \tag{3.3.7}$$

Applying $\mathcal{A}^{-1}$ to (3.3.6) we obtain

$$\begin{pmatrix} c_1\mathcal{A}^{-1}(\mu\varphi_1) \\ c_0\mathcal{A}^{-2}(\mu\varphi_1) \end{pmatrix} = \sum_{n\in\mathbb{N}^*} \frac{(\lambda_n^2-\lambda^2)a_n + 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} b_n}{(\lambda_n^2+\lambda^2)^2} \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + \frac{-2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} a_n + (\lambda_n^2-\lambda^2) b_n}{(\lambda_n^2+\lambda^2)^2} \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix}, \tag{3.3.8}$$

where

$$c_1 := \sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n \beta_n^1 a_n - (\lambda_n^2-\lambda^2)\beta_n^2 b_n}{(\lambda_n^2+\lambda^2)^2}. \tag{3.3.9}$$

In order to iterate (3.3.6) and (3.3.8), notice that applying $\mathcal{A}^{-1}$ to the relations (3.3.5) leads to

$$\begin{aligned}
\mathcal{A}^{-2} g_n^{12} &= \frac{1}{(\lambda_n^2+\lambda^2)^2}\Big[(\lambda_n^2-\lambda^2)g_n^{12} - 2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} h_n^{12} - 2\lambda\lambda_n \beta_n^1 \mathcal{A}^{-1}(\mu\varphi_1)\Big], \\
\mathcal{A}^{-2} g_n^{22} &= \frac{1}{(\lambda_n^2+\lambda^2)^2}\Big[(\lambda_n^2-\lambda^2)g_n^{22} - 2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} h_n^{22}\Big] - \frac{\lambda_n \beta_n^1}{\lambda_n^2+\lambda^2}\mathcal{A}^{-2}(\mu\varphi_1), \\
\mathcal{A}^{-2} h_n^{12} &= \frac{1}{(\lambda_n^2+\lambda^2)^2}\Big[(\lambda_n^2-\lambda^2)h_n^{12} + 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} g_n^{12} + (\lambda_n^2-\lambda^2)\beta_n^2 \mathcal{A}^{-1}(\mu\varphi_1)\Big], \\
\mathcal{A}^{-2} h_n^{22} &= \frac{1}{(\lambda_n^2+\lambda^2)^2}\Big[(\lambda_n^2-\lambda^2)h_n^{22} + 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} g_n^{22}\Big] - \frac{\lambda \beta_n^2}{\lambda_n^2+\lambda^2}\mathcal{A}^{-2}(\mu\varphi_1).
\end{aligned} \tag{3.3.10}$$

Applying successively $\mathcal{A}^{-2}$ to (3.3.6) we obtain, for every $p \in \mathbb{N}$ (with the convention that the sum from 1 to 0 is equal to 0),

$$\begin{aligned}
&\begin{pmatrix} c_0\mathcal{A}^{-(2p+1)}(\mu\varphi_1) + \sum_{j=1}^p \left(\sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n\beta_n^1 k_n^j - (\lambda_n^2-\lambda^2)\beta_n^2 l_n^j}{(\lambda_n^2+\lambda^2)^{2j+1}}\right)\mathcal{A}^{-2(p-j)-1}(\mu\varphi_1) \\ \sum_{j=1}^p \left(\sum_{n\in\mathbb{N}^*} \frac{\lambda_n\beta_n^1 k_n^j + \lambda\beta_n^2 l_n^j}{(\lambda_n^2+\lambda^2)^{2j}}\right)\mathcal{A}^{-2(p-j+1)}(\mu\varphi_1) \end{pmatrix} \\
&= \sum_{n\in\mathbb{N}^*} \frac{1}{(\lambda_n^2+\lambda^2)^{2p+1}} \left[k_n^{p+1}\begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + l_n^{p+1}\begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix}\right],
\end{aligned} \tag{3.3.11}$$

where the coefficients (k_n^j, l_n^j) are defined by

$$\begin{pmatrix} k_n^{j+1} \\ l_n^{j+1} \end{pmatrix} = \begin{pmatrix} (\lambda_n^2 - \lambda^2) & 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} \\ -2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} & (\lambda_n^2 - \lambda^2) \end{pmatrix} \begin{pmatrix} k_n^j \\ l_n^j \end{pmatrix}, \quad \text{with } \begin{pmatrix} k_n^1 \\ l_n^1 \end{pmatrix} = \begin{pmatrix} -\lambda a_n + \lambda_n \frac{\beta_n^2}{\beta_n^1} b_n \\ -\lambda_n \frac{\beta_n^1}{\beta_n^2} a_n - \lambda b_n \end{pmatrix}. \tag{3.3.12}$$

Applying successively $\mathcal{A}^{-2}$ to (3.3.8) we obtain, for every $p \in \mathbb{N}^*$,

$$\begin{aligned} &\begin{pmatrix} c_1 \mathcal{A}^{-(2p+1)}(\mu\varphi_1) + \sum_{j=1}^p \left(\sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n \beta_n^1 \tilde{k}_n^j - (\lambda_n^2 - \lambda^2)\beta_n^2 \tilde{l}_n^j}{(\lambda_n^2+\lambda^2)^{2(j+1)}} \right) \mathcal{A}^{-2(p-j)-1}(\mu\varphi_1) \\ c_0 \mathcal{A}^{-2(p+1)} \sum_{j=1}^p \left(\sum_{n\in\mathbb{N}^*} \frac{\lambda_n \beta_n^1 \tilde{k}_n^j + \lambda \beta_n^2 \tilde{l}_n^j}{(\lambda_n^2+\lambda^2)^{2j+1}} \right) \mathcal{A}^{-2(p-j+1)}(\mu\varphi_1) \end{pmatrix} \\ &= \sum_{n\in\mathbb{N}^*} \frac{1}{(\lambda_n^2+\lambda^2)^{2(p+1)}} \left[\tilde{k}_n^{p+1} \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + \tilde{l}_n^{p+1} \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} \right], \end{aligned} \tag{3.3.13}$$

where the coefficients $(\tilde{k}_n^j, \tilde{l}_n^j)$ are defined by

$$\begin{pmatrix} \tilde{k}_n^{j+1} \\ \tilde{l}_n^{j+1} \end{pmatrix} = \begin{pmatrix} (\lambda_n^2 - \lambda^2) & 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} \\ -2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} & (\lambda_n^2 - \lambda^2) \end{pmatrix} \begin{pmatrix} \tilde{k}_n^j \\ \tilde{l}_n^j \end{pmatrix}, \quad \text{with } \begin{pmatrix} \tilde{k}_n^0 \\ \tilde{l}_n^0 \end{pmatrix} = \begin{pmatrix} a_n \\ b_n \end{pmatrix}. \tag{3.3.14}$$

Assuming that $c_0 \neq 0$, equality (3.3.6) implies that

$$\begin{pmatrix} \mathcal{A}^{-1}(\mu\varphi_1) \\ 0 \end{pmatrix} \in \overline{\operatorname{span} \mathcal{B}}.$$

Then using (3.3.8) we obtain

$$\begin{pmatrix} 0 \\ c_0 \mathcal{A}^{-2}(\mu\varphi_1) \end{pmatrix} + \begin{pmatrix} c_1 \mathcal{A}^{-1}(\mu\varphi_1) \\ 0 \end{pmatrix} \in \overline{\operatorname{span} \mathcal{B}} \quad \Longrightarrow \quad \begin{pmatrix} 0 \\ \mathcal{A}^{-2}(\mu\varphi_1) \end{pmatrix} \in \overline{\operatorname{span} \mathcal{B}}.$$

Iterating with the relations (3.3.11) and (3.3.13) it comes that, for every $p \in \mathbb{N}^*$,

$$\begin{pmatrix} \mathcal{A}^{-(2p-1)}(\mu\varphi_1) \\ 0 \end{pmatrix} \in \overline{\operatorname{span} \mathcal{B}} \quad \text{and} \quad \begin{pmatrix} 0 \\ \mathcal{A}^{-2p}(\mu\varphi_1) \end{pmatrix} \in \overline{\operatorname{span} \mathcal{B}}. \tag{3.3.15}$$

Notice that if $c_0 = 0$ and $c_1 \neq 0$, the same argument can be repeated to obtain (3.3.15). Actually, one gets (3.3.15) as soon as there exists a non-zero coefficient in the left side of (3.3.11) or in the left side of (3.3.13). Thus it comes that

either (3.3.15) holds or, for every $j \in \mathbb{N}^*$,

$$\begin{aligned}
0 &= \sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n\beta_n^1 k_n^j - (\lambda_n^2-\lambda^2)\beta_n^2 l_n^j}{(\lambda_n^2+\lambda^2)^{2j+1}},\\
0 &= \sum_{n\in\mathbb{N}^*} \frac{\lambda_n\beta_n^1 k_n^j + \lambda\beta_n^2 l_n^j}{(\lambda_n^2+\lambda^2)^{2j}},\\
0 &= \sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n\beta_n^1 \tilde{k}_n^j - (\lambda_n^2-\lambda^2)\beta_n^2 \tilde{l}_n^j}{(\lambda_n^2+\lambda^2)^{2(j+1)}},\\
0 &= \sum_{n\in\mathbb{N}^*} \frac{\lambda_n\beta_n^1 \tilde{k}_n^j + \lambda\beta_n^2 \tilde{l}_n^j}{(\lambda_n^2+\lambda^2)^{2j+1}}.
\end{aligned} \tag{3.3.16}$$

Second step : we prove that if (3.3.15) holds, then $\mathcal{B}$ is a Riesz basis.

Let $\begin{pmatrix} d^1 \\ d^2 \end{pmatrix} \in X^s_{(0)}$ such that

$$\left\langle \begin{pmatrix} d^1 \\ d^2 \end{pmatrix}, \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} \right\rangle = 0, \quad \forall \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} \in \overline{\operatorname{span}\mathcal{B}}.$$

Using (3.3.15), we obtain that, for every $p \in \mathbb{N}^*$,

$$\begin{aligned}
0 &= \left\langle \begin{pmatrix} d^1 \\ d^2 \end{pmatrix}, \begin{pmatrix} \mathcal{A}^{-(2p-1)}(\mu\varphi_1) \\ 0 \end{pmatrix} \right\rangle_{X^s_{(0)}}\\
&= \langle d^1, \mathcal{A}^{-(2p-1)}(\mu\varphi_1)\rangle_{H^s_{(0)}}\\
&= \sum_{k\in\mathbb{N}^*} \lambda_k^s \langle d^1, \varphi_k\rangle\langle \varphi_k, \mathcal{A}^{-(2p-1)}(\mu\varphi_1)\rangle\\
&= \sum_{k\in\mathbb{N}^*} \lambda_k^{s+1} \frac{\langle \mu\varphi_1, \varphi_k\rangle\langle d^1, \varphi_k\rangle}{\lambda_k^{2p}}.
\end{aligned}$$

Let $d_k^1 = \langle d^1, \varphi_k\rangle$ and define

$$G : z \in \mathbb{C} \mapsto \sum_{k\in\mathbb{N}^*} \lambda_k^{s-1}\langle \mu\varphi_1, \varphi_k\rangle d_k^1 e^{z/\lambda_k^2}.$$

From uniform convergence on compact sets, it comes that G is an entire function and the previous relation imply that, for every $p \in \mathbb{N}^*$,

$$G^{(p-1)}(0) = \sum_{k\in\mathbb{N}^*} \lambda_k^{s+1} \frac{\langle \mu\varphi_1, \varphi_k\rangle d_k^1}{\lambda_k^{2p}} = 0.$$

Thus, $G \equiv 0$. If $d^1 \neq 0$, let

$$n_0 := \min\{n \in \mathbb{N}^*; d^1_n \neq 0\}.$$

It comes that

$$0 = \lambda_{n_0}^{s-1}\langle \mu\varphi_1, \varphi_{n_0}\rangle d^1_{n_0} + \sum_{k>n_0} \lambda_k^{s-1}\langle \mu\varphi_1, \varphi_k\rangle d^1_k \exp\left(z\left(\frac{1}{\lambda_k^2} - \frac{1}{\lambda_{n_0}^2}\right)\right).$$

As $\langle \mu\varphi_1, \varphi_{n_0}\rangle \neq 0$, considering z real and letting z go to infinity implies $d^1_{n_0} = 0$. Therefore, from the definition of n_0, $d^1 = 0$.

Using the exact same strategy for

$$\begin{pmatrix} 0 \\ \mathcal{A}^{-2p}(\mu\varphi_1) \end{pmatrix} \in \overline{\operatorname{span}\mathcal{B}}, \quad \forall p \in \mathbb{N}^*,$$

implies $d^2 = 0$. Thus, it comes that $d^1 = d^2 = 0$ i.e. $\overline{\operatorname{span}\mathcal{B}} = X^s_{(0)}$. From Theorem 3.3.2, we obtain that $\mathcal{B}$ is a Riesz basis of $X^2_{(0)}$ (resp. $\tilde{\mathcal{B}}$ is a Riesz basis of $X^3_{(0)}$).

Third step : we prove that in the remaining case (3.3.16), one has $a_n = b_n = 0$ for any $n \in \mathbb{N}^*$. We use the assumption $\lambda \notin \mathcal{N}$.

Let us define

$$\tilde{G} : z \in \mathbb{C} \mapsto \sum_{n\in\mathbb{N}^*} \frac{1}{(\lambda_n^2+\lambda^2)^2} \left\langle \begin{pmatrix} 2\lambda\lambda_n\beta_n^1 \\ -(\lambda_n^2-\lambda^2)\beta_n^2 \end{pmatrix}, \exp(zM) \begin{pmatrix} a_n \\ b_n \end{pmatrix} \right\rangle, \tag{3.3.17}$$

with

$$M := \frac{1}{(\lambda_n^2+\lambda^2)^2} \begin{pmatrix} (\lambda_n^2-\lambda^2) & 2\lambda\lambda_n \frac{\beta_n^2}{\beta_n^1} \\ -2\lambda\lambda_n \frac{\beta_n^1}{\beta_n^2} & (\lambda_n^2-\lambda^2) \end{pmatrix}. \tag{3.3.18}$$

Notice that the matrix appearing in this definition is the one used in the recurrence relations (3.3.12) and (3.3.14). This matrix is diagonalized as follows

$$M = \frac{1}{(\lambda_n^2+\lambda^2)^2} \begin{pmatrix} -i\frac{\beta_n^2}{\beta_n^1} & i\frac{\beta_n^2}{\beta_n^1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (\lambda_n^2-\lambda^2)+2i\lambda\lambda_n & 0 \\ 0 & (\lambda_n^2-\lambda^2)-2i\lambda\lambda_n \end{pmatrix} \begin{pmatrix} \frac{i}{2}\frac{\beta_n^1}{\beta_n^2} & \frac{1}{2} \\ \frac{-i}{2}\frac{\beta_n^1}{\beta_n^2} & \frac{1}{2} \end{pmatrix}.$$

From this diagonalization we deduce that

$$\begin{aligned}\tilde{G}(z) = \sum_{n\in\mathbb{N}^*} \frac{\exp\left(\frac{\lambda_n^2-\lambda^2}{(\lambda_n^2+\lambda^2)^2}z\right)}{(\lambda_n^2+\lambda^2)^2} \Big[& 2\lambda\lambda_n\beta_n^1 \cos\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) a_n + 2\lambda\lambda_n\beta_n^2 \sin\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) b_n \\ & + (\lambda_n^2-\lambda^2)\beta_n^1 \sin\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) a_n - (\lambda_n^2-\lambda^2)\beta_n^2 \cos\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) b_n \Big].\end{aligned}$$

Again, $\tilde{G}$ is an entire function from the uniform convergence on compact sets. The recurrence relation (3.3.14) implies that, for every $p \in \mathbb{N}$,

$$\begin{aligned}\tilde{G}^{(p)}(0) &= \sum_{n\in\mathbb{N}^*} \frac{1}{(\lambda_n^2+\lambda^2)^2} \left\langle \begin{pmatrix} 2\lambda\lambda_n\beta_n^1 \\ -(\lambda_n^2-\lambda^2)\beta_n^2 \end{pmatrix}, M^p \begin{pmatrix} a_n \\ b_n \end{pmatrix} \right\rangle \\ &= \sum_{n\in\mathbb{N}^*} \frac{2\lambda\lambda_n\beta_n^1\tilde{k}_n^p - (\lambda_n^2-\lambda^2)\beta_n^2\tilde{l}_n^p}{(\lambda_n^2+\lambda^2)^{2(p+1)}} \\ &= 0.\end{aligned}$$

Thus we obtain $\tilde{G} \equiv 0$. From the assumption $\lambda \notin \mathcal{N}$ it comes that the terms $\frac{\lambda_n^2-\lambda^2}{(\lambda_n^2+\lambda^2)^2}$ in the exponentials of $\tilde{G}$ are all different. As n increases, these terms are first increasing and then decreasing. Assume that the sequences $(a_n)_{n\in\mathbb{N}^*}$ and $(b_n)_{n\in\mathbb{N}^*}$ are both not equal to zero. Thus,

$$n_0 := \operatorname{argmax}\left\{ \frac{\lambda_n^2-\lambda^2}{(\lambda_n^2+\lambda^2)^2} \; ; \; n \in \mathbb{N}^* \text{ such that } a_n \neq 0 \text{ or } b_n \neq 0 \right\}.$$

is uniquely defined. As $\tilde{G} = 0$, it comes that, for every $z \in \mathbb{C}$,

$$\begin{aligned}&2\lambda\lambda_{n_0}\beta_{n_0}^1 \cos\left(\frac{2\lambda\lambda_{n_0}}{(\lambda_{n_0}^2+\lambda^2)^2}z\right) a_{n_0} + 2\lambda\lambda_{n_0}\beta_{n_0}^2 \sin\left(\frac{2\lambda\lambda_{n_0}}{(\lambda_{n_0}^2+\lambda^2)^2}z\right) b_{n_0} \\ &+ (\lambda_{n_0}^2-\lambda^2)\beta_{n_0}^1 \sin\left(\frac{2\lambda\lambda_{n_0}}{(\lambda_{n_0}^2+\lambda^2)^2}z\right) a_{n_0} - (\lambda_{n_0}^2-\lambda^2)\beta_{n_0}^2 \cos\left(\frac{2\lambda\lambda_{n_0}}{(\lambda_{n_0}^2+\lambda^2)^2}z\right) b_{n_0} \\ &= -\sum_{n\neq n_0} \frac{\exp\left(\left(\frac{\lambda_n^2-\lambda^2}{(\lambda_n^2+\lambda^2)^2} - \frac{\lambda_{n_0}^2-\lambda^2}{(\lambda_{n_0}^2+\lambda^2)^2}\right) z\right)}{(\lambda_n^2+\lambda^2)^2} \Big[2\lambda\lambda_n\beta_n^1 \cos\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) a_n \\ &+ 2\lambda\lambda_n\beta_n^2 \sin\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) b_n + (\lambda_n^2-\lambda^2)\beta_n^1 \sin\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) a_n \\ &- (\lambda_n^2-\lambda^2)\beta_n^2 \cos\left(\frac{2\lambda\lambda_n}{(\lambda_n^2+\lambda^2)^2}z\right) b_n \Big].\end{aligned}$$

Notice that for every term appearing in the right-hand side (excluding the terms where $a_n = b_n = 0$)

$$\frac{\lambda_n^2-\lambda^2}{(\lambda_n^2+\lambda^2)^2} - \frac{\lambda_{n_0}^2-\lambda^2}{(\lambda_{n_0}^2+\lambda^2)^2} < 0.$$

Thus applying the previous relation to $z = \frac{(\lambda_{n_0}^2+\lambda^2)^2}{\lambda\lambda_{n_0}} k\pi$ with $k \in \mathbb{N}^*$ and letting k

goes to infinity implies

$$2\lambda\lambda_{n_0}\beta^1_{n_0}a_{n_0} - (\lambda^2_{n_0} - \lambda^2)\beta^2_{n_0}b_{n_0} = 0. \tag{3.3.19}$$

In the same fashion choosing $z = \frac{(\lambda^2_{n_0}+\lambda^2)^2}{2\lambda\lambda_{n_0}}\left(2k\pi + \frac{\pi}{2}\right)$ with $k \in \mathbb{N}^*$ and letting k goes to infinity implies

$$2\lambda\lambda_{n_0}\beta^2_{n_0}b_{n_0} + (\lambda^2_{n_0} - \lambda^2)\beta^1_{n_0}a_{n_0} = 0. \tag{3.3.20}$$

Finally, the relations (3.3.19) and (3.3.20) imply $a_{n_0} = b_{n_0} = 0$. This contradicts the definition of n_0 and thus we obtain that for every $n \in \mathbb{N}^*$, $a_n = b_n = 0$. Theorem 3.3.1 thus implies the Riesz basis property.

□

3.3.2 Definition and regularity of the feedback law

So far, we have obtained from (3.2.9) the expression of the kernels k^{ij} with respect to the Fourier coefficients α^i_n of the feedback

$$K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} := \sum_{n=1}^{+\infty} \alpha^1_n\langle\Psi^1,\varphi_n\rangle + \alpha^2_n\langle\Psi^2,\varphi_n\rangle. \tag{3.3.21}$$

The regularity of the kernels and, consequently, the regularity of T, will be directly related to the decay rate of those coefficients as $n \to +\infty$. It remains to use the $TB = B$ condition (3.2.18) to construct K and T.

As $\mu\varphi_1 \in H^2_{(0)}$ and $\mathcal{B}$ is a Riesz basis of $X^2_{(0)}$ (see Proposition 3.3.1), it comes that there exist sequences $(a^1_n)_{n\in\mathbb{N}^*}$, $(a^2_n)_{n\in\mathbb{N}^*} \in \ell^2(\mathbb{N}^*,\mathbb{R})$ such that

$$\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix} = \sum_{n=1}^{+\infty} a^1_n\begin{pmatrix}g^{12}_n\\ g^{22}_n\end{pmatrix} + a^2_n\begin{pmatrix}h^{12}_n\\ h^{22}_n\end{pmatrix}, \quad \text{in } X^2_{(0)}. \tag{3.3.22}$$

Getting back to the $TB = B$ condition (3.2.18), we define

$$\alpha^1_n = \frac{\beta^1_n}{\langle\mu\varphi_1,\varphi_n\rangle}a^1_n, \qquad \alpha^2_n = \frac{\beta^2_n}{\langle\mu\varphi_1,\varphi_n\rangle}a^2_n. \tag{3.3.23}$$

Then, following the heuristic of Sec. 3.2, the transformation T is defined by (3.2.17) and the feedback law K is defined by (3.3.21).

Unfortunately, the regularity of the coefficients is only $(a^j_n)_{n\in\mathbb{N}^*} \in \ell^2(\mathbb{N}^*,\mathbb{R})$ for the $X^2_{(0)}$ Riesz basis and will not be sufficient to prove that T is continuous in $X^3_{(0)}$.

Remark 3.3.2 *Recall that we assumed controllability of the linearized system i.e. Hypothesis 3.1.1 . From [7, Remark 2], it follows that $\mu'(0) \pm \mu'(1) \neq 0$ and then that*

$(\mu\varphi_1) \notin H^3_{(0)}$. Thus, $\begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix}$ cannot be decomposed in the Riesz basis of $X^3_{(0)}$. This would have led to more regularity for the sequences $(\alpha_n^j)_n$.

It is fortunate since, as it will be noticed in Remark 3.4.1, with this extra regularity, the obtained transformation T would have been compact in $X^3_{(0)}$ and thus not invertible in $X^3_{(0)}$.

Performing a suitable decomposition of the function $\mu\varphi_1$, we prove that the coefficients of the feedback law satisfy the following regularity.

Proposition 3.3.2 *We define the sequence $(h_n)_{n\in\mathbb{N}^*}$ by*

$$h_n := \frac{4}{n^3\pi^2}\left((-1)^{n+1}\mu'(1) - \mu'(0)\right).$$

Then the sequences $(\alpha_n^1)_{n\in\mathbb{N}^*}$ and $(\alpha_n^2)_{n\in\mathbb{N}^*}$ defined in (3.3.23) satisfy

$$\left(\frac{\alpha_n^1}{n^3}\right) \in \ell^2(\mathbb{N}^*, \mathbb{R}), \quad \left(\frac{1}{n^3}\left(\alpha_n^2 - \lambda_n\beta_n^2\frac{h_n}{\langle\mu\varphi_1, \varphi_n\rangle}\right)\right) \in \ell^2(\mathbb{N}^*, \mathbb{R}).$$

Remark 3.3.3 *As a corollary, it comes that, for every $j \in \{1, 2\}$,*

$$\left(\frac{\alpha_n^j}{n^3}\right)_{n\in\mathbb{N}^*} \in \ell^\infty(\mathbb{N}^*, \mathbb{R}).$$

Proof :

First step : splitting of the problem. We start using the ideas of J.-P. Puel developed in [114] to extend the regularization result of K. Beauchard and C. Laurent [7, Lemma 1] to higher dimensions. Let

$$h : x \in (0,1) \mapsto -\frac{\pi\sqrt{2}}{3}\left((\mu'(0) + \mu'(1))x^3 - 3\mu'(0)x^2 + (2\mu'(0) - \mu'(1))x\right).$$

It comes that $g := \mu\varphi_1 - h \in H^3_{(0)}$. The Fourier coefficients of h are given by

$$\langle h, \varphi_k\rangle = h_k = \frac{4}{k^3\pi^2}\left((-1)^{k+1}\mu'(1) - \mu'(0)\right), \quad \forall k \in \mathbb{N}^*, \tag{3.3.24}$$

which is the leading term in the asymptotic expansion of $\langle\mu\varphi_1, \varphi_k\rangle$ given in [7, Remark 2]. Let us split the left-hand side of (3.3.22) in two parts

$$\begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix} = \begin{pmatrix} 0 \\ g \end{pmatrix} + \begin{pmatrix} 0 \\ h \end{pmatrix}.$$

As $g \in H^3_{(0)}$, using the Riesz basis of $X^3_{(0)}$ we get the existence of sequences $(\delta^1_n)_{n\in\mathbb{N}^*}$ and $(\delta^2_n)_{n\in\mathbb{N}^*}$ such that

$$\begin{pmatrix} 0 \\ g \end{pmatrix} = \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle \left[\frac{\lambda_n^{1/2}\delta^1_n}{\beta^1_n} \begin{pmatrix} g^{12}_n/\lambda_n^{1/2} \\ g^{22}_n/\lambda_n^{1/2} \end{pmatrix} + \frac{\lambda_n^{1/2}\delta^2_n}{\beta^2_n} \begin{pmatrix} h^{12}_n/\lambda_n^{1/2} \\ h^{22}_n/\lambda_n^{1/2} \end{pmatrix} \right]. \tag{3.3.25}$$

The coefficients of the decomposition in a Riesz basis being ℓ^2 sequences, Hypothesis 3.1.1 and the behaviour of coefficients β^j_n in (3.3.3) imply that, for every $j \in \{1, 2\}$,

$$\left(\frac{\delta^j_n}{n^3} \right)_{n\in\mathbb{N}^*} \in \ell^2(\mathbb{N}^*, \mathbb{R}). \tag{3.3.26}$$

Second step : decomposition of h. Using the Riesz basis $\mathcal{B}$ of $X^2_{(0)}$, we get coefficients $(\rho^1_n)_{n\in\mathbb{N}^*}$ and $(\rho^2_n)_{n\in\mathbb{N}^*}$ such that

$$\begin{pmatrix} 0 \\ h \end{pmatrix} = \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle \left[\frac{\rho^1_n}{\beta^1_n} \begin{pmatrix} g^{12}_n \\ g^{22}_n \end{pmatrix} + \frac{\rho^2_n}{\beta^2_n} \begin{pmatrix} h^{12}_n \\ h^{22}_n \end{pmatrix} \right]. \tag{3.3.27}$$

Recall that the basis $\mathcal{B}$ is obtained as a perturbation of the Fourier basis. To highlight this, we define the sequences $(\gamma^j_n)_{n\in\mathbb{N}^*}$ such that

$$\begin{aligned} \rho^1_n &= \lambda_n \beta^1_n \gamma^1_n, \\ \rho^2_n &= \lambda_n \beta^2_n \frac{\langle h, \varphi_n \rangle}{\langle \mu\varphi_1, \varphi_n \rangle} + \lambda_n \beta^2_n \gamma^2_n, \end{aligned} \tag{3.3.28}$$

where the factor $\lambda_n \beta^i_n$ is just for convenience.

From (3.2.16) we get, for every $k \in \mathbb{N}^*$

$$\begin{aligned} \begin{pmatrix} 0 \\ \langle h, \varphi_k \rangle \end{pmatrix} &= \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle \langle \mu\varphi_1, \varphi_k \rangle \left[\rho^1_n \begin{pmatrix} c^{12}_{nk} \\ c^{22}_{nk} \end{pmatrix} + \rho^2_n \begin{pmatrix} d^{12}_{nk} \\ d^{22}_{nk} \end{pmatrix} \right] \\ &= \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle \langle \mu\varphi_1, \varphi_k \rangle \left(\frac{\langle h, \varphi_n \rangle}{\langle \mu\varphi_1, \varphi_n \rangle} \lambda_n \beta^2_n \right) \begin{pmatrix} d^{12}_{nk} \\ d^{22}_{nk} \end{pmatrix} \\ &+ \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n \rangle \langle \mu\varphi_1, \varphi_k \rangle \left[\lambda_n \beta^1_n \gamma^1_n \begin{pmatrix} c^{12}_{nk} \\ c^{22}_{nk} \end{pmatrix} + \lambda_n \beta^2_n \gamma^2_n \begin{pmatrix} d^{12}_{nk} \\ d^{22}_{nk} \end{pmatrix} \right]. \end{aligned}$$

which, using (3.3.2), is equivalent to

$$- \begin{pmatrix} \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_k \rangle \langle h, \varphi_n \rangle \lambda_n \beta^2_n d^{12}_{nk} \\ \sum_{n=1, n\neq k}^{+\infty} \langle \mu\varphi_1, \varphi_k \rangle \langle h, \varphi_n \rangle \lambda_n \beta^2_n d^{22}_{nk} \end{pmatrix}$$

$$= \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_n\rangle \langle \mu\varphi_1, \varphi_k\rangle \left[\lambda_n \beta_n^1 \gamma_n^1 \begin{pmatrix} c_{nk}^{12} \\ c_{nk}^{22} \end{pmatrix} + \lambda_n \beta_n^2 \gamma_n^2 \begin{pmatrix} d_{nk}^{12} \\ d_{nk}^{22} \end{pmatrix} \right],$$

for every $k \in \mathbb{N}^*$. Equivalently, if we define

$$\begin{pmatrix} \tilde{h}_1 \\ \tilde{h}_2 \end{pmatrix} = -\sum_{k=1}^{+\infty} \begin{pmatrix} \left(\sum_{n=1}^{+\infty} \langle h, \varphi_n\rangle \langle \mu\varphi_1, \varphi_k\rangle \lambda_n \beta_n^2 d_{nk}^{12}\right) \varphi_k \\ \left(\sum_{n=1, n\neq k}^{+\infty} \langle h, \varphi_n\rangle \langle \mu\varphi_1, \varphi_k\rangle \lambda_n \beta_n^2 d_{nk}^{22}\right) \varphi_k \end{pmatrix},$$

we obtain from (3.2.16)

$$\begin{aligned} \begin{pmatrix} \tilde{h}_1 \\ \tilde{h}_2 \end{pmatrix} &= \sum_{n=1}^{+\infty} \sum_{k=1}^{+\infty} \langle \mu\varphi_1, \varphi_n\rangle \left[\lambda_n \gamma_n^1 \begin{pmatrix} \langle \mu\varphi_1, \varphi_k\rangle \beta_n^1 c_{nk}^{12} \\ \langle \mu\varphi_1, \varphi_k\rangle \beta_n^1 c_{nk}^{22} \end{pmatrix} + \lambda_n \gamma_n^2 \begin{pmatrix} \langle \mu\varphi_1, \varphi_k\rangle \beta_n^2 d_{nk}^{12} \\ \langle \mu\varphi_1, \varphi_k\rangle \beta_n^2 d_{nk}^{22} \end{pmatrix} \right] \varphi_k \\ &= \sum_{n=1}^{+\infty} \lambda_n^{3/2} \langle \mu\varphi_1, \varphi_n\rangle \left[\gamma_n^1 \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} + \gamma_n^2 \begin{pmatrix} h_n^{12}/\lambda_n^{1/2} \\ h_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right] \varphi_k. \end{aligned} \tag{3.3.29}$$

Finally, if $\begin{pmatrix} \tilde{h}_1 \\ \tilde{h}_2 \end{pmatrix} \in X_{(0)}^3$ (which will be proved in the next two steps) it comes that, for every $j \in \{1, 2\}$, thanks to Hypothesis 3.1.1,

$$\left(\lambda_n^{3/2} \langle \mu\varphi_1, \varphi_n\rangle \gamma_n^j\right)_n \in \ell^2(\mathbb{N}^*, \mathbb{R}),$$

which implies that $(\gamma_n^j)_n \in \ell^2(\mathbb{N}^*, \mathbb{R})$. Recalling that $\mu\varphi_1 = g + h$, we obtain

$$\begin{aligned} \frac{\alpha_n^1}{n^3} &= \frac{\rho_n^1}{n^3} + \frac{\delta_n^1}{n^3} = \frac{\lambda_n \beta_n^1}{n^3} \gamma_n^1 + \frac{\delta_n^1}{n^3}, \\ \frac{\alpha_n^2}{n^3} &= \frac{\rho_n^2}{n^3} + \frac{\delta_n^2}{n^3} = \frac{\lambda_n \beta_n^2}{n^3} \gamma_n^2 + \frac{\lambda_n \beta_n^2}{n^3} \frac{\langle h, \varphi_n\rangle}{\langle \mu\varphi_1, \varphi_n\rangle} + \frac{\delta_n^2}{n^3}, \end{aligned}$$

which will end the proof of Proposition 3.3.2.

Third step : $H_{(0)}^3$ regularity of $\tilde{h}_1$. We start by proving that $\tilde{h}_1 \in H_{(0)}^3$ i.e.

$$\sum_{k=1}^{+\infty} \left| k^3 \langle \tilde{h}_1, \varphi_k\rangle \right|^2 = \sum_{k=1}^{+\infty} \left| k^3 \sum_{n=1}^{+\infty} \langle \mu\varphi_1, \varphi_k\rangle \langle h, \varphi_n\rangle \lambda_n \beta_n^2 d_{nk}^{12} \right|^2 < +\infty. \tag{3.3.30}$$

From (3.3.3) and (3.3.24) it comes that there exists $C > 0$ such that

$$|\langle h, \varphi_n\rangle \lambda_n \beta_n^2| \leq C, \quad \forall n \in \mathbb{N}^*.$$

Using Hypothesis 3.1.1 it thus amounts to prove that

$$M_{12} := \sum_{k=1}^{+\infty} \left| \sum_{n=1}^{+\infty} d_{nk}^{12} \right|^2 = \sum_{k=1}^{+\infty} \left| \sum_{n=1}^{+\infty} \frac{\lambda_k(\lambda^2 + \lambda_k^2 - \lambda_n^2)}{(\lambda^2 + (\lambda_k - \lambda_n)^2)\,(\lambda^2 + (\lambda_k + \lambda_n)^2)} \right|^2 < +\infty. \tag{3.3.31}$$

Notice that

$$M_{12} \leq 2 \sum_{k=1}^{+\infty} \left[\left(\frac{\lambda^2 \lambda_k}{\lambda^2(\lambda^2 + 4\lambda_k^2)} \right)^2 + \left(\sum_{\substack{n=1 \\ n \neq k}}^{+\infty} d_{nk}^{12} \right)^2 \right].$$

As

$$\sum_{k=1}^{+\infty} \left(\frac{\lambda^2 \lambda_k}{\lambda^2(\lambda^2 + 4\lambda_k^2)} \right)^2 < +\infty,$$

we only deal with the second term. Straightforward computations lead to

$$\sum_{\substack{n=1 \\ n \neq k}}^{+\infty} d_{nk}^{12} = \frac{1}{2} \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \left(\frac{\lambda_k - \lambda_n}{\lambda^2 + (\lambda_k - \lambda_n)^2} + \frac{\lambda_k + \lambda_n}{\lambda^2 + (\lambda_k + \lambda_n)^2} \right). \tag{3.3.32}$$

The second term of this sum is easily dealt with. As $\lambda^2 + (\lambda_k + \lambda_n)^2 \geq (\lambda_k + \lambda_n)^2$ it comes that

$$\begin{aligned}
\sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \frac{\lambda_k + \lambda_n}{\lambda^2 + (\lambda_k + \lambda_n)^2} &\leq \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \frac{1}{\lambda_k + \lambda_n} \\
&\leq \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \int_n^{n+1} \frac{1}{\lambda_k + \pi^2 x^2} \mathrm{d}x \\
&\leq \int_1^{+\infty} \frac{1}{\lambda_k + \pi^2 x^2} \mathrm{d}x \\
&= \frac{1}{\pi^2 k} \int_{\frac{1}{k}}^{+\infty} \frac{1}{1 + y^2} \mathrm{d}y \\
&\leq \frac{1}{2k\pi}.
\end{aligned}$$

Thus,

$$\sum_{k=1}^{+\infty} \left(\sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \frac{\lambda_k + \lambda_n}{\lambda^2 + (\lambda_k + \lambda_n)^2} \right)^2 < +\infty. \tag{3.3.33}$$

We now turn back to the first term of the right-hand side of (3.3.32). We have

$$\left|\sum_{\substack{n=1\\n\neq k}}^{+\infty}\frac{\lambda_k-\lambda_n}{\lambda^2+(\lambda_k-\lambda_n)^2}\right|\leq\sum_{\substack{n=1\\n\neq k}}^{+\infty}\frac{1}{|\lambda_k-\lambda_n|}=\sum_{n=1}^{k-1}\frac{1}{\lambda_k-\lambda_n}+\sum_{n=k+1}^{+\infty}\frac{1}{\lambda_n-\lambda_k}.$$

Straightforward computations lead to

$$\begin{aligned}\sum_{n=1}^{k-1}\frac{1}{\lambda_k-\lambda_n}&=\frac{1}{2k\pi^2}\sum_{n=1}^{k-1}\left(\frac{1}{k-n}+\frac{1}{k+n}\right)\\&=\frac{1}{2k\pi^2}\left(\sum_{j=1}^{k-1}\frac{1}{j}+\sum_{j=k+1}^{2k-1}\frac{1}{j}\right)\\&=\frac{1}{2k\pi^2}\left(\sum_{j=1}^{2k-1}\frac{1}{j}-\frac{1}{k}\right).\end{aligned}$$

In the same spirit, for N large enough,

$$\sum_{n=k+1}^{N}\frac{1}{\lambda_k-\lambda_n}=\frac{1}{2k\pi^2}\sum_{j=1}^{2k}\frac{1}{j}-\frac{1}{2k\pi^2}\sum_{j=N-k}^{N+k}\frac{1}{j}$$

and thus,

$$\sum_{n=k+1}^{+\infty}\frac{1}{\lambda_n-\lambda_k}=\frac{1}{2k\pi^2}\sum_{j=1}^{2k}\frac{1}{j}.$$

Finally,

$$\left|\sum_{\substack{n=1\\n\neq k}}^{+\infty}\frac{\lambda_k-\lambda_n}{\lambda^2+(\lambda_k-\lambda_n)^2}\right|\leq\sum_{\substack{n=1\\n\neq k}}^{+\infty}\frac{1}{|\lambda_k-\lambda_n|}\leq\frac{1}{2k\pi^2}\left(2\sum_{j=1}^{2k-1}\frac{1}{j}-\frac{1}{2k}\right),\tag{3.3.34}$$

which belongs to $\ell^2(\mathbb{N}^*)$ as $\sum_{j=1}^{2k-1}\frac{1}{j}\underset{k\to+\infty}{\sim}\ln(2k-1)$. From (3.3.32), this proves that $\tilde{h}_1\in H^3_{(0)}$.

Fourth step : $H^3_{(0)}$ regularity of $\tilde{h}_2$. We end the proof of Proposition 3.3.2 by proving that $\tilde{h}_2\in H^3_{(0)}$ i.e.

$$\sum_{k=1}^{+\infty}\left|k^3\langle\tilde{h}_2,\varphi_k\rangle\right|^2=\sum_{k=1}^{+\infty}\left|k^3\sum_{\substack{n=1\\n\neq k}}^{+\infty}\langle\mu\varphi_1,\varphi_k\rangle\langle h,\varphi_n\rangle\lambda_n\beta_n^2d_{nk}^{22}\right|^2<+\infty.\tag{3.3.35}$$

From (3.3.3) and (3.3.24) it comes that there exists $C > 0$ such that

$$|\langle h, \varphi_n \rangle \lambda_n \beta_n^2| \leq C, \quad \forall n \in \mathbb{N}^*.$$

Using Hypothesis 3.1.1 it thus amounts to prove that

$$\sum_{k=1}^{+\infty} \left| \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} d_{nk}^{22} \right|^2 = \sum_{k=1}^{+\infty} \left| \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \frac{\lambda(\lambda^2 + \lambda_k^2 + \lambda_n^2)}{(\lambda^2 + (\lambda_k - \lambda_n)^2)\,(\lambda^2 + (\lambda_k + \lambda_n)^2)} \right|^2 < +\infty. \tag{3.3.36}$$

Notice that

$$\begin{aligned} \frac{\lambda(\lambda^2 + \lambda_k^2 + \lambda_n^2)}{(\lambda^2 + (\lambda_k - \lambda_n)^2)\,(\lambda^2 + (\lambda_k + \lambda_n)^2)} &= \frac{\lambda}{2} \left(\frac{1}{\lambda^2 + (\lambda_k + \lambda_n)^2} + \frac{1}{\lambda^2 + (\lambda_k - \lambda_n)^2} \right) \\ &\leq \frac{\lambda}{2} \left(\frac{\lambda_k + \lambda_n}{\lambda^2 + (\lambda_k + \lambda_n)^2} + \frac{|\lambda_k - \lambda_n|}{\lambda^2 + (\lambda_k - \lambda_n)^2} \right) \\ &\leq \frac{\lambda}{2} \left(\frac{\lambda_k + \lambda_n}{\lambda^2 + (\lambda_k + \lambda_n)^2} + \frac{1}{|\lambda_k - \lambda_n|} \right). \end{aligned}$$

From (3.3.33) and (3.3.34), it comes that

$$\sum_{k=1}^{+\infty} \left| \sum_{\substack{n=1 \\ n \neq k}}^{+\infty} \frac{\lambda(\lambda^2 + \lambda_k^2 + \lambda_n^2)}{(\lambda^2 + (\lambda_k - \lambda_n)^2)\,(\lambda^2 + (\lambda_k + \lambda_n)^2)} \right|^2 < +\infty.$$

This ends the proof of Proposition 3.3.2.

□

3.3.3 Wellposedness of the transformation

The regularity of the coefficient obtained in the previous section is sufficient to define a continuous operator T in the state space.

Proposition 3.3.3 *The transformation T defined on $X^3_{(0)}$ by (3.2.17) and (3.3.23) is linear continuous in $X^3_{(0)}$.*

Proof : Let $\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)}$. From (3.2.17), using the Riesz basis property of Propo-

sition 3.3.1, it comes that $T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}$ belongs to $X^3_{(0)}$ if the following sequences,

$$\left(\frac{\lambda_n^{1/2}\alpha_n^1\langle\Psi^2,\varphi_n\rangle}{\beta_n^1}-\frac{\lambda_n^{1/2}\alpha_n^2\langle\Psi^1,\varphi_n\rangle}{\beta_n^1}\right)_{n\in\mathbb{N}^*},$$
$$\left(\frac{\lambda_n^{1/2}\alpha_n^1\langle\Psi^1,\varphi_n\rangle}{\beta_n^2}+\frac{\lambda_n^{1/2}\alpha_n^2\langle\Psi^2,\varphi_n\rangle}{\beta_n^2}\right)_{n\in\mathbb{N}^*},$$

belong to $\ell^2(\mathbb{N}^*,\mathbb{R})$. From (3.3.3) and Remark 3.3.3, it comes that, for every $i,j,k\in\{1,2\}$,

$$\begin{aligned}\sum_{n=1}^{+\infty}\left|\frac{\lambda_n^{1/2}\alpha_n^j\langle\Psi^i,\varphi_n\rangle}{\beta_n^k}\right|^2 &\leq C\sum_{n=1}^{+\infty}\left|\frac{\alpha_n^j}{n^3}n^3\langle\Psi^i,\varphi_n\rangle\right|^2\\ &\leq C\sum_{n=1}^{+\infty}\left|n^3\langle\Psi^i,\varphi_n\rangle\right|^2\\ &= C\|\Psi^i\|^2_{H^3_{(0)}}.\end{aligned}$$

Finally, we obtain

$$\left\|T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|_{X^3_{(0)}}\leq C\left\|\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|_{X^3_{(0)}}.$$

□

3.4 Invertibility of the transformation

This section aims at proving the invertibility of T. As a first step, we prove in subsection 3.4.1 that T is a Fredholm operator. In subsection 3.4.1, we prove that the analogous of (3.1.9) in finite dimension holds on a certain functional space. This will be used in subsection (3.4.3) to obtain the invertibility of T.

3.4.1 Fredholm form

The goal of this subsection is the proof of the following result.

Proposition 3.4.1 *There exists $\widetilde{T}:X^3_{(0)}\to X^3_{(0)}$ invertible such that $T-\widetilde{T}$ is a compact operator.*

The proof of this proposition rely on the study of the feedback law done in Proposition 3.3.2.

Proof : Let $\widetilde{T}$ be the transformation defined by (3.2.17) where the coefficients α_n^1 and α_n^2 are respectively replaced by

$$\tilde{\alpha}_n^1 = 0, \qquad \tilde{\alpha}_n^2 = \frac{\langle h, \varphi_n \rangle}{\langle \mu\varphi_1, \varphi_n \rangle} \lambda_n \beta_n^2.$$

From (3.3.28) and the fact that $(\gamma_n^j)_n \in \ell^2(\mathbb{N}^*, \mathbb{R})$, it follows that defining $\hat{\alpha}_n^j = \alpha_n^j - \tilde{\alpha}_n^j$, we get that $\left(\frac{\hat{\alpha}_n^j}{n^3}\right)_n \in \ell^2(\mathbb{N}^*, \mathbb{R})$. Recall that $\langle h, \varphi_n \rangle$ is given in (3.3.24).

The computations done in the proof of Proposition 3.3.3 show that $\widetilde{T}$ is a linear continuous operator of $X^3_{(0)}$. We prove that $\widetilde{T}$ is invertible. For any $\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)}$,

$$\widetilde{T}\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = \sum_{n=1}^{+\infty} -\frac{\tilde{\alpha}_n^2 \lambda_n^{1/2}}{\beta_n^1} \langle \Psi^1, \varphi_n \rangle \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} + \frac{\tilde{\alpha}_n^2 \lambda_n^{1/2}}{\beta_n^2} \langle \Psi^2, \varphi_n \rangle \begin{pmatrix} h_n^{12}/\lambda_n^{1/2} \\ h_n^{22}/\lambda_n^{1/2} \end{pmatrix}.$$

From (3.3.3), we have that $\left(\frac{\lambda_n^{1/2}}{\beta_n^j}\right)_n \in \ell^\infty(\mathbb{N}^*, \mathbb{R})$. Thus,

$$\Psi \in H^3_{(0)} \Longleftrightarrow \left(\frac{\tilde{\alpha}_n^2 \lambda_n^{1/2}}{\beta_n^j} \langle \Psi, \varphi_n \rangle \right)_{n \in \mathbb{N}^*} \in \ell^2(\mathbb{N}^*, \mathbb{R}).$$

Then, the Riesz basis property of Proposition 3.3.1 ends the proof of the invertibility of $\widetilde{T}$.

Finally, we prove that $T - \widetilde{T}$ is compact using the Hilbert-Schmidt criterion, i.e., we prove that

$$\sum_{n=1}^{+\infty} \left\| (T - \widetilde{T}) \begin{pmatrix} \varphi_n/\lambda_n^{3/2} \\ 0 \end{pmatrix} \right\|^2_{X^3_{(0)}} + \left\| (T - \widetilde{T}) \begin{pmatrix} 0 \\ \varphi_n/\lambda_n^{3/2} \end{pmatrix} \right\|^2_{X^3_{(0)}} < +\infty. \tag{3.4.1}$$

From (3.2.17) and the definition of $\widetilde{T}$ it comes that

$$\begin{aligned} \left(T - \widetilde{T}\right) \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = & \sum_{n=1}^{+\infty} \left[\frac{\hat{\alpha}_n^1 \langle \Psi^2, \varphi_n \rangle}{\beta_n^1} - \frac{\hat{\alpha}_n^2 \langle \Psi^1, \varphi_n \rangle}{\beta_n^1} \right] \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} \\ & + \sum_{n=1}^{+\infty} \left[\frac{\hat{\alpha}_n^1 \langle \Psi^1, \varphi_n \rangle}{\beta_n^2} + \frac{\hat{\alpha}_n^2 \langle \Psi^2, \varphi_n \rangle}{\beta_n^2} \right] \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix}. \end{aligned}$$

Thus,

$$\left\| (T - \widetilde{T}) \begin{pmatrix} \varphi_n/\lambda_n^{3/2} \\ 0 \end{pmatrix} \right\|^2_{X^3_{(0)}}$$

$$= \left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n^{3/2}\beta_n^1} \right] \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} + \left[\frac{\hat{\alpha}_n^1}{\lambda_n^{3/2}\beta_n^2} \right] \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} \right\|_{X^3_{(0)}}^2$$

$$= \left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right] \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} + \left[\frac{\hat{\alpha}_n^1}{\lambda_n\beta_n^2} \right] \begin{pmatrix} h_n^{12}/\lambda_n^{1/2} \\ h_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 .$$

The first term is estimated in the following way

$$\begin{aligned}
&\left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right] \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 \\
&\leq 2 \left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right] \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} - \varphi_n/\lambda_n^{3/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 + 2 \left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right] \begin{pmatrix} \varphi_n/\lambda_n^{3/2} \\ 0 \end{pmatrix} \right\|_{X^3_{(0)}}^2 \\
&\leq 2 \left[\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right]^2 \left\| \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} - \varphi_n/\lambda_n^{3/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 + 2 \left[\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right]^2 .
\end{aligned}$$

We proved in Lemma 3.3.1 that

$$\sum_{n=1}^{+\infty} \left\| \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} - \varphi_n/\lambda_n^{3/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 < +\infty.$$

As $\left(\frac{\hat{\alpha}_n^2}{n^3}\right)_n \in \ell^2(\mathbb{N}^*, \mathbb{R})$, this gives that

$$\sum_{n=1}^{+\infty} \left\| \left[-\frac{\hat{\alpha}_n^2}{\lambda_n\beta_n^1} \right] \begin{pmatrix} g_n^{12}/\lambda_n^{1/2} \\ g_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 < +\infty.$$

The term

$$\sum_{n=1}^{+\infty} \left\| \left[\frac{\hat{\alpha}_n^1}{\lambda_n\beta_n^2} \right] \begin{pmatrix} h_n^{12}/\lambda_n^{1/2} \\ h_n^{22}/\lambda_n^{1/2} \end{pmatrix} \right\|_{X^3_{(0)}}^2 ,$$

is dealt with in the exact same way, ending the proof of Proposition 3.4.1.

□

Remark 3.4.1 *Notice that the key point in proving the invertibility of $\widetilde{T}$ is that for any $n \in \mathbb{N}^*$, $\langle h, \varphi_n \rangle \neq 0$.*

We also underline that the compactness of $T - \widetilde{T}$ comes from the fact that $\left(\frac{\hat{\alpha}_n^j}{n^3}\right)_{n\in\mathbb{N}^*} \in \ell^2(\mathbb{N}^*, \mathbb{R})$. If we had $\mu\varphi_1 \in H^3_{(0)}$ (this is not possible due to Hypothesis 3.1.1, see [7,

Remark 2]), then from (3.2.18) we would have obtained that $\left(\frac{\alpha_n^j}{n^3}\right)_{n\in\mathbb{N}^*} \in \ell^2(\mathbb{N}^*,\mathbb{R})$. *This would have led to the compactness of the transformation* T *preventing its invertibility.*

3.4.2 Operator equality

We prove that the formal operator equality

$$T(A+BK) = AT - \lambda T \tag{3.4.2}$$

holds true on an appropriate functional space. Recall that K is defined by (3.3.21).

Remark 3.4.2 *Notice that due to the regularity of the coefficients* α_n^j *obtained in Proposition 3.3.2, the operator* K *is not defined on* $X^3_{(0)}$. *Otherwise, taking* $\begin{pmatrix} 0 \\ \psi^2 \end{pmatrix} \in X^3_{(0)}$ *with* $\psi^2 \notin H^s_{(0)}$ *for any* $s > 3$ *would imply*

$$\left| K \begin{pmatrix} 0 \\ \psi^2 \end{pmatrix} \right| = \left| \sum_{n\in\mathbb{N}^*} \alpha_n^2 \langle \psi^2, \varphi_n \rangle \right| = \left| \sum_{n\in\mathbb{N}^*} \frac{\alpha_n^2}{n^3} (n^3 \langle \psi^2, \varphi_n \rangle) \right| < +\infty,$$

and therefore $\left(\frac{\alpha_n^2}{n^3}\right)_{n\in\mathbb{N}^*} \in l^2(\mathbb{N}^*,\mathbb{R})$ *which is in contradiction with Proposition 3.3.2.*

Due to the previous remark, the functional setting for (3.4.2) to hold needs to be specified.

Recall that

$$(A+BK)\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = \begin{pmatrix} 0 & -\Delta \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} + K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix}.$$

We define

$$\begin{aligned} D(A+BK) &:= \left\{ \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)} \; ; \; (A+BK)\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in X^3_{(0)} \right\} \\ &= \left\{ \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in (H^5 \cap H^3_{(0)}) \times H^5_{(0)} \; ; \; \Delta\Psi^1 + K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \mu\varphi_1 \in H^3_{(0)} \right\}. \end{aligned}$$

The previous characterization is equivalent to

$$\Delta^2\Psi^1(x) + 2K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \mu'(x)\varphi_1'(x) = 0, \quad x \in \{0,1\}. \tag{3.4.3}$$

Let us point out that K is well defined on $D(A+BK)$. Indeed, let $\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in D(A+$

BK). Then, Proposition 3.3.2 implies

$$\begin{aligned}\left|K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right| &= \left|\sum_{n\in\mathbb{N}^*}\alpha_n^1\langle\Psi^1,\varphi_n\rangle+\alpha_n^2\langle\Psi^2,\varphi_n\rangle\right|\\ &\le\left(\sum_{n\in\mathbb{N}^*}\left|\frac{\alpha_n^1}{n^3}\right|^2\right)^{1/2}\|\Psi^1\|_{H^3_{(0)}}+C\sum_{n\in\mathbb{N}^*}\lambda_n^{3/2}\left|\langle\Psi^2,\varphi_n\rangle\right|\\ &\le\left(\sum_{n\in\mathbb{N}^*}\left|\frac{\alpha_n^1}{n^3}\right|^2\right)^{1/2}\|\Psi^1\|_{H^3_{(0)}}+C\left(\sum_{n\in\mathbb{N}^*}\frac{1}{\lambda_n}\right)^{1/2}\left(\sum_{n\in\mathbb{N}^*}\lambda_n^4\left|\langle\Psi^2,\varphi_n\rangle\right|^2\right)^{1/2}<+\infty\end{aligned}$$

We now prove the density of $D(A+BK)$.

Lemma 3.4.1 *The domain $D(A+BK)$ is dense in $X^3_{(0)}$.*

Proof : Let us prove that $D(A+BK)^\perp=\{0\}$ in $X^3_{(0)}$. Let $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\in D(A+BK)^\perp$.

First step : we prove that $\Psi^1=0$. For all $k\in\mathbb{N}^*$, consider $\phi^1=\varphi_k$.

If $\alpha_k^1=0$, then

$$K\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}=\sum_{n\in\mathbb{N}^*}\alpha_n^2\langle\phi^2,\varphi_n\rangle.$$

and therefore $\phi^2=0$ is sufficient for (3.4.3) to be satisfied. Thus $\begin{pmatrix}\phi^1\\ 0\end{pmatrix}\in D(A+BK)$. This leads to

$$0=\left\langle\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix},\begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle_{X^3_{(0)}}=\left\langle\Psi^1,\varphi_k\right\rangle_{H^3_{(0)}},$$

and therefore $\psi_1\equiv 0$. Otherwise, if $\alpha_k^1\neq 0$, from the asymptotic behaviour of α_n^2 (see Proposition 3.3.2), there exists $N\in\mathbb{N}^*$ such that $\forall n\ge N$, $\alpha_n^2\neq 0$. Therefore, setting $\phi^2=-\frac{\alpha_k^1}{\alpha_n^2}\varphi_n$ implies that

$$K\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}=\alpha_k^1+\alpha_n^2\left(-\frac{\alpha_k^1}{\alpha_n^2}\right)=0$$

which means that (3.4.3) is satisfied and $\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\in D(A+BK)$. Thus,

$$\begin{aligned}0=\left\langle\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix},\begin{pmatrix}\varphi_k\\ -\frac{\alpha_k^1}{\alpha_n^2}\varphi_n\end{pmatrix}\right\rangle_{X^3_{(0)}}&=\left\langle\Psi^1,\varphi_k\right\rangle_{H^3_{(0)}}-\frac{\alpha_k^1}{\alpha_n^2}\left\langle\Psi^2,\varphi_n\right\rangle_{H^3_{(0)}}\\ &=\left\langle\Psi^1,\varphi_k\right\rangle_{H^3_{(0)}}-\frac{\lambda_n^{3/2}\alpha_k^1}{\alpha_n^2}\left(\lambda_n^{3/2}\left\langle\Psi^2,\varphi_n\right\rangle\right)\end{aligned}$$

$$\underset{n\to+\infty}{\longrightarrow} \left\langle \Psi^1, \varphi_k \right\rangle_{H^3_{(0)}}.$$

Indeed, $(n^3\langle \Psi^2, \varphi_n\rangle)_n \in l^2(\mathbb{N}^*)$ and Remark 3.3.3 implies that $\left(\frac{\lambda_n^{3/2}\alpha_k^1}{\alpha_n^2}\right)_n \in l^\infty(\mathbb{N}^*)$. We conclude that, in any case,

$$\left\langle \Psi^1, \varphi_k \right\rangle_{H^3_{(0)}} = 0, \quad \forall k \in \mathbb{N}^*,$$

and thus $\Psi^1 = 0$.

Second step : we prove that $\Psi^2 = 0$. For all $k \in \mathbb{N}^*$, let $\phi^2 = \varphi_k$.

If there exists $n \in \mathbb{N}^*$ such that $\alpha_n^1 \neq 0$, then setting $\phi^1 = -\frac{\alpha_k^2}{\alpha_n^1}\varphi_n$ implies that $K\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} = 0$ and therefore $\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in D(A+BK)$. Thus,

$$0 = \left\langle \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}-\frac{\alpha_k^2}{\alpha_n^1}\varphi_n\\ \varphi_k\end{pmatrix} \right\rangle_{X^3_{(0)}} = \left\langle \Psi^2, \varphi_k \right\rangle_{H^3_{(0)}}.$$

Otherwise, for every $\phi^1 \in H^3_{(0)}$,

$$K\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} = \alpha_k^2.$$

Therefore, we consider ϕ^1 solution to

$$\begin{cases}\Delta\phi^1 = +\alpha_k^2(\mu\varphi_1)\\ \phi^1(0) = \phi^1(1) = 0.\end{cases}$$

Since $\mu\varphi_1 \in H^3 \cap H_0^1$, then $\phi^1 \in H^5 \cap H^3_{(0)}$ and, since (3.4.3) is satisfied, $\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in D(A+BK)$. Moreover, as $\Psi^1 = 0$,

$$0 = \left\langle \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \right\rangle_{X^3_{(0)}} = \left\langle \Psi^2, \varphi_k \right\rangle_{H^3_{(0)}}.$$

This proves that $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = 0$.

□

Let us now turn our attention to the kernel system. More precisely, we prove the following.

Proposition 3.4.2 *For* $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in D(A+BK)$,

$$T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = (A-\lambda I)T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \quad \textit{in } X^3_{(0)}.$$

Proof : Let $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in D(A+BK)$. Then

$$(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = \begin{pmatrix}-\Delta\Psi^2\\ \Delta\Psi^1 + K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mu\varphi_1\end{pmatrix}.$$

Therefore

$$\begin{aligned}
T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} =& \sum_{n\in\mathbb{N}^*}\left[\frac{\alpha_n^1}{\beta_n^1}\left\langle \Delta\Psi^1 + K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mu\varphi_1, \varphi_n\right\rangle + \frac{\alpha_n^2}{\beta_n^1}\left\langle \Delta\Psi^2, \varphi_n\right\rangle\right]\begin{pmatrix}g_n^{12}\\ g_n^{22}\end{pmatrix}\\
&+\left[-\frac{\alpha_n^1}{\beta_n^2}\left\langle \Delta\Psi^2, \varphi_n\right\rangle + \frac{\alpha_n^2}{\beta_n^2}\left\langle \Delta\Psi^1 + K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mu\varphi_1, \varphi_n\right\rangle\right]\begin{pmatrix}h_n^{12}\\ h_n^{22}\end{pmatrix}\\
=& K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\sum_{n\in\mathbb{N}^*}\left[\frac{\alpha_n^1}{\beta_n^1}\begin{pmatrix}g_n^{12}\\ g_n^{22}\end{pmatrix} + \frac{\alpha_n^2}{\beta_n^2}\begin{pmatrix}h_n^{12}\\ h_n^{22}\end{pmatrix}\right]\left\langle \mu\varphi_1, \varphi_n\right\rangle\\
&+\sum_{n\in\mathbb{N}^*} -\lambda_n\left[\frac{\alpha_n^1}{\beta_n^1}\left\langle \Psi^1, \varphi_n\right\rangle + \frac{\alpha_n^2}{\beta_n^1}\left\langle \Psi^2, \varphi_n\right\rangle\right]\begin{pmatrix}g_n^{12}\\ g_n^{22}\end{pmatrix}\\
&+\lambda_n\left[\frac{\alpha_n^1}{\beta_n^2}\left\langle \Psi^2, \varphi_n\right\rangle - \frac{\alpha_n^2}{\beta_n^2}\left\langle \Psi^1, \varphi_n\right\rangle\right]\begin{pmatrix}h_n^{12}\\ h_n^{22}\end{pmatrix}\\
=& \left(\sum_{n\in\mathbb{N}^*}\alpha_n^1\langle\Psi^1, \varphi_n\rangle + \alpha_n^2\langle\Psi^2, \varphi_n\rangle\right)\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\\
&+\sum_{n\in\mathbb{N}^*} -\lambda_n\left[\frac{\alpha_n^1}{\beta_n^1}\left\langle \Psi^1, \varphi_n\right\rangle + \frac{\alpha_n^2}{\beta_n^1}\left\langle \Psi^2, \varphi_n\right\rangle\right]\begin{pmatrix}g_n^{12}\\ g_n^{22}\end{pmatrix}\\
&+\lambda_n\left[\frac{\alpha_n^1}{\beta_n^2}\left\langle \Psi^2, \varphi_n\right\rangle - \frac{\alpha_n^2}{\beta_n^2}\left\langle \Psi^1, \varphi_n\right\rangle\right]\begin{pmatrix}h_n^{12}\\ h_n^{22}\end{pmatrix}, \text{ in } X^3_{(0)},
\end{aligned}$$

using the $TB = B$ condition (3.2.18) and the expression of K given in (3.3.21). Moreover, by the definition of g_n^{ij}, h_n^{ij} given in (3.2.13) and by the relations (3.2.15), we have, on one hand,

$$\begin{aligned}
\left\langle (A-\lambda I)T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle_{(L^2)^2} =& \left\langle T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, A^*\begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle_{(L^2)^2} - \lambda\left\langle T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle_{(L^2)^2}\\
=& \left\langle T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}-\lambda\varphi_k\\ \lambda_k\varphi_k\end{pmatrix}\right\rangle_{(L^2)^2}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{n\in\mathbb{N}^*} \left[\frac{\alpha_n^1}{\beta_n^1}\langle \Psi^2, \varphi_n\rangle - \frac{\alpha_n^2}{\beta_n^1}\langle \Psi^1, \varphi_n\rangle\right] \left(\lambda_k\langle g_n^{22}, \varphi_k\rangle - \lambda\langle g_n^{12}, \varphi_k\rangle\right) \\
&\quad + \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle + \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle\right] \left(\lambda_k\langle h_n^{22}, \varphi_k\rangle - \lambda\langle h_n^{12}, \varphi_k\rangle\right) \\
&= \sum_{n\in\mathbb{N}^*} \left[\frac{\alpha_n^1}{\beta_n^1}\langle \Psi^2, \varphi_n\rangle - \frac{\alpha_n^2}{\beta_n^1}\langle \Psi^1, \varphi_n\rangle\right] \lambda_n\langle g_n^{11}, \varphi_k\rangle \\
&\quad + \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle + \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle\right] \lambda_n\langle h_n^{11}, \varphi_k\rangle \\
&= \sum_{n\in\mathbb{N}^*} \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle - \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle\right] \lambda_n\langle h_n^{12}, \varphi_k\rangle \\
&\quad - \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle + \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle\right] \lambda_n\langle g_n^{12}, \varphi_k\rangle \\
&= \left\langle T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle_{(L^2)^2}.
\end{aligned}
$$

On the other hand, using again (3.2.13) and (3.2.15),

$$
\begin{aligned}
\left\langle (A-\lambda I)T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}0\\ \varphi_k\end{pmatrix}\right\rangle_{(L^2)^2} &= \left\langle T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}-\lambda_k\varphi_k\\ -\lambda\varphi_k\end{pmatrix}\right\rangle_{(L^2)^2} \\
&= \sum_{n\in\mathbb{N}^*} \left[\frac{\alpha_n^1}{\beta_n^1}\langle \Psi^2, \varphi_n\rangle - \frac{\alpha_n^2}{\beta_n^1}\langle \Psi^1, \varphi_n\rangle\right] \left(-\lambda\langle g_n^{22}, \varphi_k\rangle - \lambda_k\langle g_n^{12}, \varphi_k\rangle\right) \\
&\quad + \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle + \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle\right] \left(-\lambda\langle h_n^{22}, \varphi_k\rangle - \lambda_k\langle h_n^{12}, \varphi_k\rangle\right) \\
&= \sum_{n\in\mathbb{N}^*} \left[\frac{\alpha_n^1}{\beta_n^2}\langle \Psi^2, \varphi_n\rangle - \frac{\alpha_n^2}{\beta_n^2}\langle \Psi^1, \varphi_n\rangle\right] \lambda_n\langle h_n^{22}, \varphi_k\rangle \\
&\quad + \left[\frac{\alpha_n^1}{\beta_n^1}\langle \Psi^1, \varphi_n\rangle + \frac{\alpha_n^2}{\beta_n^1}\langle \Psi^2, \varphi_n\rangle\right] - \lambda_n\langle g_n^{22}, \varphi_k\rangle \\
&\quad + \sum_{n\in\mathbb{N}^*} \left[\alpha_n^1\langle \Psi^1, \varphi_n\rangle + \alpha_n^2\langle \Psi^2, \varphi_n\rangle\right] \langle \mu\varphi_1, \varphi_k\rangle \\
&= \left\langle T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}, \begin{pmatrix}0\\ \varphi_k\end{pmatrix}\right\rangle_{(L^2)^2}.
\end{aligned}
$$

Indeed, (3.2.13) and (3.2.15) imply

$$
\begin{aligned}
-\lambda_k\langle h_n^{12}, \varphi_k\rangle - \lambda\langle h_n^{22}, \varphi_k\rangle &= \lambda_n\langle h_n^{21}, \varphi_k\rangle + \beta_n^2\langle \mu\varphi_1, \varphi_k\rangle \\
&= -\lambda_n\frac{\beta_n^2}{\beta_n^1}\langle g_n^{22}, \varphi_k\rangle + \beta_n^2\langle \mu\varphi_1, \varphi_k\rangle.
\end{aligned}
$$

Consequently, by the definition of $D(A+BK)$ and by the continuity of T from $X^3_{(0)}$ into itself, $T(A+BK) = (A-\lambda I)T$ holds in $X^3_{(0)}$. Notice that all the previous sums are finite due to the regularity assumptions on the functions of $D(A+BK)$.

□

We conclude this section by noting that, following Proposition 3.4.2, if $\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in D(A+BK)$, then $T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in X^5_{(0)}$. Indeed,

$$AT\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = T(A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - \lambda T\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix},$$

where the right-hand side of the last equation is in $X^3_{(0)}$.

3.4.3 Invertibility

Let us now turn to the invertibility of T.

Proposition 3.4.3 *The operator T is invertible from $X^3_{(0)}$ into itself.*

Proof : In the previous section, we have proven that $\widetilde{T}$ is invertible and $T-\widetilde{T}$ is a compact operator. Consequently, the index of T is equal to zero. Thus, it is sufficient to prove that Im $T^* = X^3_{(0)}$ to show the invertibility of T.

Let us rewrite $TA + BK = AT - \lambda T$ under the form

$$\begin{aligned} T(A+BK+\lambda I) &= AT \\ T(A+BK+\lambda I+\rho J) &= AT + \rho TJ \end{aligned}$$

where $\rho > 0$ and

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Assume for the moment that ρ is such that $(A+BK+\lambda I+\rho J)$ is an invertible operator from $D(A+BK)$ to $X^3_{(0)}$. Then,

$$A^{-1}T = T(A+BK+\lambda I+\rho J)^{-1} + \rho A^{-1}TJ(A+BK+\lambda I+\rho J)^{-1}$$

Consider $\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} \in$ Ker T^*. Then, for all $\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} \in X^3_{(0)}$, we have

$$0 = \langle A^{-1}T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} - T(A+BK+\lambda I+\rho J)^{-1}\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, \begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}}$$

$$
\begin{aligned}
&- \rho\langle A^{-1}TJ(A+BK+\lambda I+\rho J)^{-1}\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, \begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\
=&\langle\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, T^*(A^{-1})^*\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}} - \langle(A+BK+\lambda I+\rho J)^{-1}\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, T^*\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\
&- \rho\langle J(A+BK+\lambda I+\rho J)^{-1}\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, T^*(A^{-1})^*\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\
=&\langle I-\rho J(A+BK+\lambda I+\rho J)^{-1}\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, T^*(A^{-1})^*\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{X^3_{(0)}}
\end{aligned}
$$

Since $(A+BK+\lambda I+\rho J)$ is invertible, we have, by choosing $\rho \notin \sigma\left(I-\rho J(A+BK+\lambda I+\rho J)^{-1}\right)$, that $I-\rho J(A+BK+\lambda I+\rho J)^{-1}$ is an isomorphism from $X^3_{(0)}$ into itself. Therefore, $T^*(A^*)^{-1}\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$. Thus the space Ker T^*, which is of finite dimension, is stable by $(A^*)^{-1}$. Therefore, there exists $\nu \in \mathbb{C}$ and $\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} \in$ Ker T^* such that

$$
(A^*)^{-1}\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix} = \begin{pmatrix}-\Delta^{-1}\chi^2\\ \Delta^{-1}\chi^1\end{pmatrix} = \nu\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}. \tag{3.4.4}
$$

Hence, for every $j \in \mathbb{N}^*$,

$$
\nu^2\langle\chi^1,\varphi_j\rangle = \nu\langle-\Delta^{-1}\chi^2,\varphi_j\rangle = \frac{1}{\lambda_j}\nu\langle\chi^2,\varphi_j\rangle = \lambda_j\langle\Delta^{-1}\chi^1,\varphi_j\rangle = -\frac{1}{\lambda_j^2}\langle\chi^1,\varphi_j\rangle.
$$

Therefore,

$$
(\nu^2+\frac{1}{\lambda_j^2})\langle\chi^1,\varphi_j\rangle = 0,
$$

and, since $\chi^1 \neq 0$, we obtain that there exists one and only one $k \in \mathbb{N}^*$ such that

$$
\nu = \pm i\frac{1}{\lambda_k}, \quad \chi^1 = c_k\varphi_k, \quad c_k \in \mathbb{C}\setminus\{0\}.
$$

Furthermore, from (3.4.4), we obtain $\chi^2 = \mp i c_k\varphi_k$. Finally, we have, by the $TB = B$ condition (3.2.18),

$$
\begin{aligned}
\pm i\overline{c_k}\lambda_k^3\langle\mu\varphi_1,\varphi_k\rangle =&\langle\mu\varphi_1,\chi^2\rangle_{H^1_0,H^5_{(0)}}\\
=&\langle\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}, \begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{(H^1_0)^2,(H^5_{(0)})^2}\\
=&\langle T\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}, \begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{(H^1_0)^2,(H^5_{(0)})^2}\\
=&\langle\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}, T^*\begin{pmatrix}\chi^1\\ \chi^2\end{pmatrix}\rangle_{(H^1_0)^2,(H^5_{(0)})^2}
\end{aligned}
$$

$$=0.$$

Since $\langle \mu\varphi_1, \varphi_k\rangle \neq 0$, we conclude that c_k must be zero, which implies that Ker $T^* = \{0\}$.

It remains to prove that $(A + BK + \lambda I + \rho J)$ is an invertible operator from $D(A + BK)$ to $X^3_{(0)}$. Let us provide the sketch of the proof of this fact. Consider

$$B_n = \begin{pmatrix} 0 \\ (\mu\varphi_1)_n \end{pmatrix},$$

where

$$(\mu\varphi_1)_n = \sum_{i=1}^{n} \langle \mu\varphi_1, \varphi_i\rangle \varphi_i(x).$$

Note that, for all $n \in \mathbb{N}$, $(\mu\varphi_1)_n \in H^3_{(0)}$. The key argument for the invertibility of $(A + BK + \lambda I + \rho J)$ is the fact that $(\mu\varphi_1)_n$ is an uniformly bounded sequence in $H^2_{(0)}$. Indeed,

$$\begin{aligned} \|(\mu\varphi_1)_n\|^2_{H^2_{(0)}} &= \sum_{i=1}^{n} \lambda_i^2 \left|\langle \mu\varphi_1, \varphi_i\rangle\right|^2 \\ &\leq \sum_{i=1}^{\infty} \lambda_i^2 \left|\langle \mu\varphi_1, \varphi_i\rangle\right|^2 \\ &\leq C \sum_{i=1}^{\infty} \frac{1}{\lambda_i} < +\infty \end{aligned}$$

Let us introduce the space

$$H := (H^3_{(0)} \cap H^4) \times H^4_{(0)},$$

equipped with the inner product

$$\langle \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rangle_H := \langle \begin{pmatrix} \psi_1^{(4)} \\ \psi_2^{(4)} \end{pmatrix}, \begin{pmatrix} \phi_1^{(4)} \\ \phi_2^{(4)} \end{pmatrix} \rangle_{(L^2)^2} + \langle \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rangle_{(L^2)^2}$$

Let $\delta \in (1/2, 1)$. Then there exists $C > 0$ such that, for every $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \in H$, we have

$$\left| K \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \right| \leq C \left(\|\psi_1\|_{H^3_{(0)}} + \|\psi_2\|_{H^{3+\delta}_{(0)}} \right).$$

We first prove with the Lax-Milgram Theorem that, for $\rho > 0$ sufficiently large and

independent of n, there exists, for every $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in H^3_{(0)}$, a solution $\begin{pmatrix} \psi_1^n \\ \psi_2^n \end{pmatrix} \in H$ of

$$\langle (A + B_n K + \lambda I + \rho J) \begin{pmatrix} \psi_1^n \\ \psi_2^n \end{pmatrix}, J \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rangle_H = \langle \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, J \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rangle_H,$$

for $\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \in H$. The coercivity of the bilinear form comes from the boundedness of $(\mu\varphi_1)_n$ in H^2, from the Cauchy-Schwarz inequality and from fact that

$$\|\psi_2\|_{H^{3+\delta}_{(0)}} \leq \epsilon \|\psi_2\|_{H^4_{(0)}} + C_\epsilon \|\psi_2\|_{L^2},$$

where $\epsilon > 0$ is small, where $C_\epsilon > 0$ and where $\rho > 0$ is chosen sufficently large so that $\rho - C_\epsilon > 0$.

Moreover, from the a priori estimate given by the Lax-Milgram Theorem, we obtain that the sequence $\begin{pmatrix} \psi_1^n \\ \psi_2^n \end{pmatrix}$ converges weakly to $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ in H. Finally, using the fact that $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} - (\lambda I + \rho J) \begin{pmatrix} \psi_1^n \\ \psi_2^n \end{pmatrix} \in X^3_{(0)}$, allows us to prove that the limit $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ is a solution of $(A + BK + \lambda I + \rho J)$ and belongs to $D(A + BK)$, which concludes the proof.

□

3.5 Well-posedness of the linearized system

This section ends the proof of Theorem 3.1.3. Due to Remark 3.4.2, the feedback K is not well defined for functions in $X^3_{(0)}$. In subsection 3.5.1, we give a meaning to the solution $\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix}$ of (3.1.18) by proving that $A + BK$ generates a C^0-semigroup. Finally we conclude to the exponential stability using the operator equality of Proposition 3.4.2 and the invertibility of the transformation T.

3.5.1 Well-posedness of the linearized system

Let us first show,

Proposition 3.5.1 *The operator $(A + BK)$ defined on $D(A + BK)$ generates a C^0-semigroup on $X^3_{(0)}$. Thus, there exists a unique solution $C([0, T]; X^3_{(0)})$ of* (3.1.18) *with $v(t) = K \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix}$.*

Proof : We prove that $(A+BK)$ is the infinitesimal generator of a C^0-semigroup on $X^3_{(0)}$.

First step. The density of $D(A+BK)$ in $X^3_{(0)}$ was proven in Lemma 3.4.1.

Second step. Let us prove that $(A+BK)$ is closed. Let $\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix} \in D(A+BK)$ such that

$$\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix} \underset{n\to+\infty}{\longrightarrow} \begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, \text{ in } X^3_{(0)},$$
$$(A+BK)\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix} \underset{n\to+\infty}{\longrightarrow} \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, \text{ in } X^3_{(0)}.$$

We have

$$(A+BK)\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix} = \begin{pmatrix}-\Delta\psi^2_n\\ \Delta\psi^1_n + K\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix}\mu\varphi_1\end{pmatrix}.$$

Hence

$$-\Delta\psi^2_n \underset{n\to+\infty}{\longrightarrow} -\Delta\psi^2, \text{ in } H^1_0$$
$$-\Delta\psi^2_n \underset{n\to+\infty}{\longrightarrow} \phi^1, \text{ in } H^3_{(0)},$$

and, consequently, $-\Delta\psi^2 = \phi^1$, $\psi^2 \in H^5_{(0)}$ and $\psi^2_n \underset{n\to+\infty}{\longrightarrow} \psi^2$ in $H^5_{(0)}$. Therefore, $K\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}$ is well-defined and

$$\left|K\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix} - K\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\right| = \left|\sum_{k\in\mathbb{N}^*}\alpha^1_k\langle\psi^1_n-\psi^1,\varphi_k\rangle + \alpha^2_k\langle\psi^2_n-\psi^2,\varphi_k\rangle\right|$$
$$\leq \left(\sum_{k\in\mathbb{N}^*}\left(\frac{\alpha^1_k}{k^3}\right)^2\right)^{1/2}\|\psi^1_n-\psi^1\|_{H^3_{(0)}} + \left(\sum_{k\in\mathbb{N}^*}\left(\frac{\alpha^2_k}{k^4}\right)^2\right)^{1/2}\|\psi^2_n-\psi^2\|_{H^4_{(0)}} \underset{n\to+\infty}{\longrightarrow} 0.$$

We obtain

$$\Delta\psi^1_n + K\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix}\mu\varphi_1 \underset{n\to+\infty}{\longrightarrow} \phi^2 \quad \text{ in } H^3_{(0)},$$
$$\Delta\psi^1_n + K\begin{pmatrix}\psi^1_n\\ \psi^2_n\end{pmatrix}\mu\varphi_1 \underset{n\to+\infty}{\longrightarrow} \Delta\psi^1 + K\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\mu\varphi_1 \quad \text{ in } H^1_0.$$

We conclude that $\Delta\psi^1 + K\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\mu\varphi_1 = \phi^2$ and therefore $\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix} \in D(A+BK)$.

Third step. Let us now prove the dissipativity of $(A+BK)$. Since T is invertible

from $X^3_{(0)}$ into itself, we define the norm $\|\cdot\|_T := \|T\cdot\|_{X^3_{(0)}}$, which is equivalent to the $X^3_{(0)}$ norm. We denote $\langle\cdot,\cdot\rangle_T$ the associated inner product.

Consider $\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix} \in D(A+BK)$. From Proposition 3.4.2, we have, since $T\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix} \in X^5_{(0)}$,

$$\begin{aligned}\langle (A+BK)\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, \begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\rangle_T =& \langle T(A+BK)\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, T\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\ =& \langle AT\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, T\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\rangle_{X^3_{(0)}} - \lambda\left\|\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\right\|_T^2\\ =& -\lambda\left\|\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}\right\|_T^2 \leq 0.\end{aligned}$$

Consider now the dissipativity of $(A+BK)^*$. First, let $\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix} \in D(A+BK)$ and $\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in X^3_{(0)}$. We have

$$\begin{aligned}\langle (A+BK)\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_T =& \langle T(A+BK)\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_{X^3_{(0)}}\\ =& \langle \begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}, (T^{-1}A^*T - \lambda I)\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_T.\end{aligned}$$

which implies that $D((A+BK)^*) = \left\{\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in X^3_{(0)} \,\middle|\, T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in X^5_{(0)}\right\}$. Then, for $\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \in D((A+BK)^*)$, we have

$$\begin{aligned}\langle (A+BK)^*\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_T =& \langle (T^{-1}A^*T - \lambda I)\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_T\\ =& \langle A^*T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}, T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\rangle_{X^3_{(0)}} - \lambda\left\|T\begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix}\right\|_{X^3_{(0)}}^2\\ \leq& 0.\end{aligned}$$

Thus from the Lumer-Philipps theorem (see e.g. [112, Corollary 4.4]) we obtain that $A+BK$ generates a C^0-semigroup on $X^3_{(0)}$.

□

3.5.2 Proof of rapid stabilization

This section is dedicated to the proof of the main result, the rapid stabilization stated in Theorem 3.1.3.

Proof : To begin, let us assume that $\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix} \in D(A+BK)$. Then, from Proposition 3.5.1, we get that the associated solution of (3.1.18) with the feedback law $v(t) = K\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}$ is given by

$$\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix} = e^{t(A+BK)}\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}. \tag{3.5.1}$$

Let $\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix} := T\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}$. Using the equality of operators given in Proposition 3.4.2, it comes that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix} &= T\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}\\ &= T(A+BK)\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}\\ &= (AT-\lambda T)\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}\\ &= A\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix} - \lambda\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix}.\end{aligned}$$

Thus, we obtain the exponential stability of $\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix}$ i.e.

$$\left\|\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix}\right\|_{X^3_{(0)}} \le e^{-\lambda t}\left\|\begin{pmatrix}\xi_0^1\\ \xi_0^2\end{pmatrix}\right\|_{X^3_{(0)}} = e^{-\lambda t}\left\|T\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|_{X^3_{(0)}}.$$

From the continuity and invertibility of T in $X^3_{(0)}$ (see Proposition 3.4.3) it comes that

$$\begin{aligned}\left\|\begin{pmatrix}\Psi^1(t)\\ \Psi^2(t)\end{pmatrix}\right\|_{X^3_{(0)}} &\le \|T^{-1}\|\left\|\begin{pmatrix}\xi^1(t)\\ \xi^2(t)\end{pmatrix}\right\|_{X^3_{(0)}}\\ &\le \|T^{-1}\|\,\|T\|\,e^{-\lambda t}\left\|\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|_{X^3_{(0)}}.\end{aligned} \tag{3.5.2}$$

Finally, let $\Psi_0 \in H^3_{(0)}$. Then, $\begin{pmatrix}\Re(\Psi_0)\\ \Im(\Psi_0)\end{pmatrix} \in X^3_{(0)}$. The stability estimate (3.5.2) and the

density proved in Lemma 3.4.1 ends the proof of Theorem 3.1.3.

□

3.6 Admissibility of the feedback operator

In the previous section, it was shown that $A + BK$ is the infinitesimal generator of a group on $X^3_{(0)}$ and that the solution of (3.1.18) is given by

$$\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = e^{(A+BK)t} \begin{pmatrix} \Psi^1_0 \\ \Psi^2_0 \end{pmatrix}.$$

In this section, we prove, despite the fact that K is not bounded on $X^3_{(0)}$, that K is admissible for the group e^{At} defined on $X^3_{(0)}$ with the domain $X^5_{(0)}$.

First, we prove that the Duhamel's formula makes sense for initial datas in $D(A+BK)$.

Lemma 3.6.1 *For any* $(\Psi^1_0, \Psi^2_0) \in D(A + BK)$*, the solution of* (3.1.18) *may be written as*

$$\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = e^{At} \begin{pmatrix} \Psi^1_0 \\ \Psi^2_0 \end{pmatrix} - \int_0^t e^{A(t-s)} K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix} ds, \tag{3.6.1}$$

in $C([0,T]; X^3_{(0)}) \cap C^1([0,T]; X^1_{(0)})$

Proof : From the semi-group theory, we know that, for $(\Psi^1_0, \Psi^2_0) \in D(A + BK)$, the solution of (3.1.18),

$$\begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = e^{(A+BK)t} \begin{pmatrix} \Psi^1_0 \\ \Psi^2_0 \end{pmatrix},$$

is in $C^1([0,T]; X^3_{(0)}) \cap C([0,T]; D(A + BK))$. Therefore, from the definition of K, we have

$$K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \in C([0,T]).$$

From the admissibility of B for the group e^{At} on $X^3_{(0)}$ ([7, Lemma 1]), we obtain

$$\int_0^t e^{A(t-s)} K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix} \mathrm{ds} \in C([0,T]; X^3_{(0)}).$$

Moreover, since $(\Psi^1_0, \Psi^2_0) \in X^3_{(0)}$, we obtain that

$$\begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} := e^{At} \begin{pmatrix} \Psi^1_0 \\ \Psi^2_0 \end{pmatrix} - \int_0^t e^{A(t-s)} K \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} \begin{pmatrix} 0 \\ \mu\varphi_1 \end{pmatrix} \mathrm{ds}, \tag{3.6.2}$$

is in $C([0,T];X^3_{(0)} \cap C^1([0,T];X^1_{(0)})$. Hence,

$$\begin{aligned}\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}_t - \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix}_t =& (A+BK)\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - Ae^{At}\begin{pmatrix}\Psi^1_0\\ \Psi^2_0\end{pmatrix}\\ &- A\int_0^t e^{A(t-s)}K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{ds} - K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\\ =& A\left(\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix}\right)\end{aligned}$$

Therefore,

$$\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix},$$

solves

$$\begin{cases}\partial_t\left(\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix}\right) = \begin{pmatrix}0 & -\Delta\\ \Delta & 0\end{pmatrix}\left(\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} - \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix}\right), \ (t,x)\in(0,1)\times(0,T),\\ (\Psi^1-\theta^1)(t,0) = (\Psi^1-\theta^1)(t,1) = 0, \quad (\Psi^2-\theta^2)(t,0) = (\Psi^2-\theta^2)(t,1) = 0,\\ (\Psi^1-\theta^1)(0,x) = 0, \quad (\Psi^2-\theta^2)(0,x) = 0,\end{cases}$$

which implies that

$$\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = \begin{pmatrix}\theta^1\\ \theta^2\end{pmatrix},$$

in $C([0,T];X^3_{(0)}) \cap C^1([0,T];X^1_{(0)})$.

□

From Lemma 3.6.1, we deduce the admissibility of K.

Lemma 3.6.2 *For* $(\Psi^1_0,\Psi^2_0) \in X^3_{(0)}$, *the solution* $(\Psi^1,\Psi^2) \in C([0,T];X^3_{(0)})$ *of* (3.1.18) *satisfies,*

$$\int_0^t e^{A(t-s)}K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix} ds \in C([0,T];X^3_{(0)}).$$

Moreover,

$$K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} \in L^2(0,T).$$

Proof :

Let us first consider $(\Psi^1_0,\Psi^2_0) \in D(A+BK)$. From (3.6.1), we have

$$\left\|\int_0^t e^{A(t-s)}K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{ds}\right\|^2_{C(0,T;X^3)} \le \left\|\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|^2_{C(0,T;X^3)}$$

$$
\begin{aligned}
&+\left\|e^{At}\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|^2_{C(0,T;X^3)}\\
\leq&\left\|e^{(A+BK)t}\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|^2_{C(0,T;X^3)}\\
&+\left\|e^{At}\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|^2_{C(0,T;X^3)}\\
\leq&C\left\|\begin{pmatrix}\Psi_0^1\\ \Psi_0^2\end{pmatrix}\right\|^2_{X^3}.
\end{aligned}
$$

Therefore, by density,

$$
\int_0^t e^{A(t-s)}K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{ds}\in C([0,T];X^3_{(0)}).
$$

Finally, for $(\Psi_0^1,\Psi_0^2)\in X^3_{(0)}$, we have

$$
\begin{aligned}
&\left\|\int_0^t e^{-As}K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix}\mathrm{ds}\right\|^2_{C(0,T;X^3)}\\
&=\sup_{t\in[0,T]}\sum_{k\in\mathbb{N}}\lambda_k^3\left|\int_0^t\left\langle e^{-As}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix},\begin{pmatrix}\varphi_k\\ 0\end{pmatrix}\right\rangle K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&\quad+\sup_{t\in[0,T]}\sum_{k\in\mathbb{N}}\lambda_k^3\left|\int_0^t\left\langle e^{-As}\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix},\begin{pmatrix}0\\ phi_k\end{pmatrix}\right\rangle K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&=\sup_{t\in[0,T]}\sum_{k\in\mathbb{N}}\lambda_k^3\left|\int_0^t\left\langle\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix},\begin{pmatrix}\sin(\lambda_k s)\\ \cos(\lambda_k s)\end{pmatrix}\varphi_k\right\rangle K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&\quad+\sup_{t\in[0,T]}\sum_{k\in\mathbb{N}}\lambda_k^3\left|\int_0^t\left\langle\begin{pmatrix}0\\ \mu\varphi_1\end{pmatrix},\begin{pmatrix}\sin(\lambda_k s)\\ \cos(\lambda_k s)\end{pmatrix}\varphi_k\right\rangle K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&\geq\sup_{t\in[0,T]}\sum_{k\in\mathbb{N}}C\left|\int_0^t\sin(\lambda_k s)K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&\quad+\sup_{t\in[0,T]}C\sum_{k\in\mathbb{N}}\left|\int_0^t\cos(\lambda_k s)K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\mathrm{ds}\right|^2\\
&\geq C\sup_{t\in[0,T]}\left\|K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|^2_{L^2(0,t)}=C\left\|K\begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix}\right\|^2_{L^2(0,T)},
\end{aligned}
$$

the last line holding by the Ingham inequality ([77]) and Riesz Basis results.

Troisième partie

Calcul numérique des contrôles optimaux de l'équation des ondes

Chapitre 4

Calcul numérique des contrôles optimaux de l'équation des ondes

Contents

4.1 Introduction **183**
4.2 Legendre Galerkin semi-discretization **188**
4.2.1 Legendre Galerkin approximation 188
4.2.2 Discrete observability 191
4.3 Spectral filtering **193**
4.4 Mixed Legendre-Galerkin method **197**
4.5 Nitsche's method **201**
4.5.1 Symmetric case 201
4.5.2 Non-symmetric case 206
4.5.3 Finite Element Method 209
4.6 Introduction of the numerical experiments **214**
4.6.1 Numerical Experiments 215

4.1 Introduction

Let us consider the one-dimensonal wave equation with homogeneous Dirichlet boundary conditions,

$$\begin{cases} u_{tt}(x,t) - u_{xx}(x,t) = 0, & \text{for } (x,t) \in (-1,1) \times (0,T) \\ u(-1,t) = u(1,t) = 0, & \text{for } t \in (0,T) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{for } x \in (-1,1) \end{cases} \tag{4.1.1}$$

With initial data $(u_0, u_1) \in H_0^1(-1,1) \times L^2(-1,1)$, the energy of the solution,

$$E(u(t)) := \frac{1}{2}\int_{-1}^{1} \left[|u_t(x,t)|^2 + |u_x(x,t)|^2\right] \, \mathrm{dx},$$

is known to be conserved along time : $E(u(t)) = E(u(0))$, $t \geq 0$.

We are interested in the *boundary observability* of the wave equation. We say that it is *observable* in time $T > 0$ if there exists $c_T > 0$ such that

$$c_T E(u(0)) \leq \int_0^T |u_x(1,t)|^2 \, \mathrm{dt}, \quad \forall (u_0, u_1) \in H_0^1(-1,1) \times L^2(-1,1). \tag{4.1.2}$$

Note that the right hand side of this inequality is finite for every time $T > 0$. Indeed, for every $T > 0$, there exists $C_T > 0$ such that

$$\int_0^T |u_x(1,t)|^2 \, \mathrm{dt} \leq C_T E(u(0)), \quad \forall (u_0, u_1) \in H_0^1(-1,1) \times L^2(-1,1). \tag{4.1.3}$$

This is often referred as the *direct* or *hidden regularity* inequality ([95]).

Whereas the direct inequality holds for every $T > 0$, the observability inequality (4.1.2) holds only if T is sufficiently large, $T \geq 4$ in this case (see for instance [85]). The time $T = 4$ is the minimal time for all waves travelling within the spatial domain $(-1,\ 1)$ at propagation speed 1 to reach the region of observation which is $x = 1$. In fact, if $T = 4$, inequalities (4.1.2) and (4.1.3) hold with $c_T = C_T = 8$ so that both become equalities.

Inequalities (4.1.2) and (4.1.3) play a fundamental role in the Dirichlet boundary controllability of the wave equation

$$\begin{cases} y_{tt}(x,t) - y_{xx}(x,t) = 0, & \text{for } (x,t) \in (-1,1) \times (0,T), \\ y(-1,t) = 0, & \text{for } t \in (0,T), \\ y(1,t) = v(t), & \text{for } t \in (0,T), \\ y(x,0) = y_0(x), y_t(x,0) = y_1(x), & \text{for } x \in (-1,1). \end{cases} \tag{4.1.4}$$

We say that (4.1.4) is *null controllable* in time $T > 0$ if, for all initial data $(y_0, y_1) \in L^2(-1,1) \times H^{-1}(-1,1)$, there exists a control $v \in L^2(0,T)$ that drives the solution y from (y_0, y_1) to

$$y(.,T) = y_t(.,T) = 0. \tag{4.1.5}$$

Because of the time-reversibility of the wave equation, null controllability is equivalent to controllability, which is the property to drive the solution of (4.1.4) from any initial state (at $t = 0$) to any final state (at $t = T$) in $L^2(-1,1) \times H^{-1}(-1,1)$.

The equivalence between the observability of (4.1.1) and the controllability of

(4.1.4) was shown in [95] using the so-called Hilbert Uniqueness Method (HUM). Moreover, when controllable, a characterization of the control of minimal $L^2(0,T)$-norm was given.

If v is a control candidate, it must satisfy

$$\int_0^T v(t)\tilde{u}_x(1,t)\mathrm{dt} =< (y_1, -y_0), (\tilde{u}_0, \tilde{u}_1) >, \quad \forall(\tilde{u}_0, \tilde{u}_1) \in H_0^1 \times L^2, \tag{4.1.6}$$

where $\tilde{u}$ is solution of (4.1.1) with initial data $(\tilde{u}_0, \tilde{u}_1)$. To see this, one formally multiplies (4.1.4) by $\tilde{u}(x,t)$ and make some integrations by parts. Note that here and in the sequel, H_0^1 stands for $H_0^1(-1,1)$, and the same notational simplification is adopted for spaces like $L^2(-1,1)$ and $H^{-1}(-1,1)$. Moreover, $< .,. >$ stands for the duality product between $H^{-1} \times L^2$ and $H_0^1 \times L^2$.

The control of minimal $L^2(0,T)$-norm is simply given by

$$v(t) = u_x(1,t), \tag{4.1.7}$$

where u is the solution of (4.1.1) for suitably determined initial data $(u_0, u_1) \in H_0^1 \times L^2$.

How are (u_0, u_1) determined ? Define the linear, symmetric and positive operator

$$\Lambda_T : H_0^1 \times L^2 \quad \to \quad H^{-1} \times L^2 \tag{4.1.8}$$

$$< \Lambda_T(u_0, u_1), (\tilde{u}_0, \tilde{u}_1) > \quad = \quad \int_0^T u_x(1,t)\tilde{u}_x(1,t)\mathrm{dt}, \ \forall(\tilde{u}_0, \tilde{u}_1) \in H_0^1 \times L^2. \tag{4.1.9}$$

The direct inequality (4.1.3) ensures that Λ_T is well defined and continuous, while the observability inequality (4.1.2) ensures that Λ_T is positive definite and thus invertible. Thus, if T is sufficiently large for the observability inequality to hold, one can solve : find $(u_0, u_1) \in H_0^1 \times L^2$ such that

$$\Lambda_T(u_0, u_1) = (y_1, -y_0). \tag{4.1.10}$$

Λ_T is referred as a *observability gramian* (for system (4.1.1)) in the control literature (see for instance [122]) or as the *HUM operator* (for control problem (4.1.4)) in [95]. Note also that, due to the symmetry of Λ_T, the solution (u_0, u_1) of (4.1.10) can also be viewed as the minimizer over $H_0^1 \times L^2$ of the quadratic functional

$$J(u_0, u_1) = \frac{1}{2}\int_0^T |u_x(1,t)|^2 \,\mathrm{dt} \quad - \ < (y_1, -y_0), (u_0, u_1) > . \tag{4.1.11}$$

From (4.1.10) (or, equivalently, (4.1.11)), one also deduces an upper bound for the

control of the form

$$||v||_{L^2(0,T)} \leq \frac{C}{c_T}||y_0, y_1||_{L^2\times H^{-1}}.$$

In the present work, we consider space discretizations of the wave equation (4.1.1) using a Legendre Galerkin method, with approximation vector spaces made of polynomials of degree N or less. Let us denote by $u^N(x,t)$ the resulting approximation of u.

We are interested with the discrete version of the observability inequality (4.1.2) that may be written :

$$c_{N,T}E(u^N(0)) \leq \int_0^T \left|u_x^N(1,t)\right|^2 \mathrm{dt}\ . \tag{4.1.12}$$

In particular, we would like the discrete observability inequality to hold uniformly with respect to the discretization parameter N, that is with

$$c_{N,T} = c_T > 0.$$

This uniform observability property is an important condition for proving the convergence (as $N \to \infty$) of the control approximations resulting from the approximation method chosen for (4.1.1) (see for instance [137, 55, 56]). An approximation v_N of the control is obtained by minimizing the discrete analogue of the functional J,

$$\begin{aligned} J_N(u_0^N, u_1^N) &= \frac{1}{2}\int_0^T \left|u_x^N(1,t)\right|^2 \mathrm{dt} \\ &\quad - < (y_1, -y_0), (u_0^N, u_1^N) >_{H^{-1}\times L^2, H_0^1\times L^2}, \end{aligned} \tag{4.1.13}$$

over the approximation spaces for the initial data (u_0^N, u_1^N), and then by setting $v_N(t) = u_x^N(1,t)$, where u^N is the solution of (4.2.2) with the initial data (u_0^N, u_1^N) minimizing (4.1.13). A uniform observability property thus ensures that J_N is uniformly coercive and provides a uniform upper bound (with respect to the discretization parameter N) for $||v_N||_{L^2(0,T)}$.

Unfortunately, the usual convergence of solutions for approximations of wave equations like (4.1.1) does not guarantee the convergence of the control approximations. In particular, this does not guarantee the uniformity of (4.1.12). This was first proved in the context of finite-difference and finite element approximations in [83]. Similarly, the standard Legendre Galerkin semi-discretization does not provide a uniform observability inequality ([21]). This pathology is caused by the high frequency components of the solutions, the highest numerical eigenmodes not being suitable approximations of the corresponding eigenmodes of the continuous problem. We refer the reader to [137, 55] for interpretations in terms of numerical wave propagation and

for a sharp analysis of the phenomena involved in the context of finite-difference and finite-element approximations of the one-dimensional (and two-dimensional) wave equation.

For all these standard space discretization methods, a remedy to recover the uniformity of the observability is to filter out the high frequency components of the solution. This method is often referred as *Fourier filtering*. It was shown in [83, 136, 21] that this method led to the uniformity of the observability inequality, for the numerical scheme considered in these papers. This method, however, has a major drawback : it is necessary to compute the eigenmodes (eigenvalues and associated eigenfunctions), which can be costly, especially if one wants to consider 2 or 3-dimensional spatial domains.

Based on numerical experiments in the case of a finite-element or a finite-difference semi-discretization of (4.1.1), other remedies were suggested by Glowinski in [74, 73]. Among them, let us cite a Tychonoff regularization, a mixed formulation of (4.1.1), a bi-grid method and a penalty method. It was shown in [107] that a bi-grid method applied to a finite-difference semi-discretization allows one to obtain the uniformity of the observability inequality and the convergence of the numerical controls. In [26], the same positive results were proved using a mixed finite element approximation.

In this paper we propose three remedies to recover the uniformity of the boundary observability inequality in the case of a spectral semi-discretization of (4.1.1). In the context of the Legendre-Galerkin semi-discretization of (4.1.1) considered in [21], we show numerically that a spectral filtering (a commonly used procedure in spectral methods), a mixed method and a Nitsche method allow one to recover the uniformity of $c_{N,T}$. To our knowledge, this is the first time that one obtains the uniformity of $c_{N,T}$ without the use of Fourier filtering in the case of a spectral semi-discretization of (4.1.1).

Unfortunately, these remedies are not fully satisfactory since they do not seem to guarantee the uniformity of the discrete analogue of the direct inequality (4.1.3),

$$\int_0^T \left|u_x^N(1,t)\right|^2 \mathrm{dt} \leq C_{N,T} E(u^N(0)), \tag{4.1.14}$$

that is with

$$C_{N,T} < C_T,$$

independently of the discretization parameter N. In this regard, the situation here is very different from what happens with classical finite element or finite difference methods, where the uniform direct inequality holds even without filtering out the highest frequencies ([83]). These negative results would prevent us from using the abstract results in [55, 56]) which need the uniformity for the direct inequality as well. Nevertheless, we will show through a numerical example that whenever the uniformity of the observability inequality holds (using the remedies cited above),

convergence of the associated control approximations holds as well.

The paper is organized as follows. In Section 2, we recall some results on the behaviour of $c_{N,T}$ and $C_{N,T}$ obtained in [21] with the Legendre-Galerkin approximation of (4.1.1). In Section 3, we show numerically that a modification of the discrete version of the functional $J(u_0, u_1)$ via a spectral filtering allows us to obtain an uniform and positive lower-bound of $c_{N,T}$. In Section 4, we obtain numerically the uniformity of $c_{N,T}$ when considering a mixed Legendre-Galerkin method. The uniform result on $c_{N,T}$ differs to the one obtained in [26] since a weaker version of the observability inequality was necessary. In Section 5, we introduce Nitsche's Method to append Dirichlet type boundary conditions. We show that if one considers a modified version of (4.1.12), then it is possible to recover an uniform and positive lower-bound of $c_{N,T}$. Finally, in Section 6, we present the results of our numerical investigation on the necessity of a uniform bound of $c_{N,T}$ and $C_{N,T}$ in (4.1.12)-(4.1.14) to obtain the convergence of the numerical controls. We provide an example of initial data such that the numerical controls are unbounded in $L^2(0,T)$ norm for the semi-discretization considered in [21], a formulation that doesn't satisfy the discrete version of (4.1.2)-(4.1.3) uniformly, while the formulations presented in this paper give convergent numerical controls.

4.2 Legendre Galerkin approximation and discrete observability properties

4.2.1 Legendre Galerkin approximation

First of all, let us recall that the solution of the wave equation admits the Fourier expansion

$$u(x,t) = \sum_{k=1}^{\infty} \left(\alpha_k \cos(\sqrt{\lambda_k} t) + \frac{\beta_k}{\sqrt{\lambda_k}} \sin(\sqrt{\lambda_k} t) \right) \phi_k(x),$$

where

$$\alpha_k = (u_0, \phi_k)_{L^2(-1,1)}; \quad \beta_k = (u_1, \phi_k)_{L^2(-1,1)},$$

and $(\lambda_k, \phi_k)_{k\geq 1}$ is the eigensystem of the eigenproblem : find $\lambda \in \mathbb{R}$, $\phi \in H_0^1$ such that

$$(\phi, \psi)_{H_0^1(-1,1)} = \lambda (\phi, \psi)_{L^2(-1,1)}, \quad \forall \psi \in H_0^1(-1,1). \tag{4.2.1}$$

with $(\phi_k, \phi_k)_{L^2(-1,1)} = 1$, that are

$$\lambda_k = (k\pi)^2 \quad \phi_k(x) = \sin(\sqrt{\lambda_k}\, x).$$

Let us now briefly describe the semi-discretization of (4.1.1) studied in [21]. Let

L_k denote the k-th order Legendre polynomial, given by the recurrence relation,

$$\begin{aligned}
&L_0(x) = 1,\\
&L_1(x) = x,\\
&(k+1)L_{k+1}(x) = (2k+1)xL_k(x) - kL_{k-1}(x), \quad k \geq 2.
\end{aligned}$$

Let $\mathbb{P}_N^0(-1,1)$ denote the space of polynomials of degree at most N vanishing at $x=-1$ and $x=1$. We notably have (see [121]) :

$$\mathbb{P}_N^0(-1,1) = \text{span}\{\widetilde{L}_1(x), \ldots, \widetilde{L}_{N-1}(x)\},$$

where

$$\widetilde{L}_k(x) = c_k(1-x^2)L_k(x), \quad c_k = \frac{\sqrt{k+1/2}}{k(k+1)}, \quad k \geq 1.$$

The semi-discretization of (4.1.1) with the Legendre-Galerkin spectral method is then : find

$$u^N(x,t) = \sum_{k=1}^{N-1} a_{N,k}(t)\widetilde{L}_k(x)$$

such that

$$\int_{-1}^{1} u_{tt}^N(x,t)\psi^N(x)\,\mathrm{dx} + \int_{-1}^{1} u_x^N(x,t)\psi_x^N(x)\,\mathrm{dx} = 0, \quad \forall \psi^N \in \mathbb{P}_N^0(-1,1). \quad (4.2.2)$$

Introducing $\mathbf{a}_N(t) := (a_{N,1}(t), \ldots, a_{N,N-1}(t))^t$, system (4.2.2) is a linear second order system which takes the form

$$M_N\mathbf{a}_N''(t) + K_N\mathbf{a}_N(t) = 0,$$

where M_N and K_N are positive definite symmetric matrices given by (see [121])

$$M_N(i,j) = \begin{cases} \frac{2}{(2i-1)(2i+3)} & \text{if i=j} \\ \frac{-1}{(2r+3)\sqrt{(2r+1)(2r+5)}} & \text{if } |i-j| = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$K_N(i,j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise,} \end{cases}$$

with $r = \min\{i,j\}$.

Taking $\psi^N = u_t^N(x,t)$ in (4.2.2), one obtains that the energy of the solution of (4.2.2) is preserved along time : $E(u^N(t)) = E(u^N(0)), \forall t \geq 0$.

Let $(\lambda_{N,k}, \phi_{N,k})$, $1 \leq k \leq N-1$ denote the eigenpair solutions of the eigenproblem :

find $\lambda \in \mathbb{R}$, $\phi \in \mathbb{P}_N^0(-1,1)$ such that

$$(\phi, \psi^N)_{H_0^1(-1,1)} = \lambda(\phi, \psi^N)_{L^2(-1,1)}, \quad \forall \psi \in \mathbb{P}_N^0(-1,1). \tag{4.2.3}$$

It was shown in [21] that the eigenvalues $\{\lambda_{N,k}\}$ of the discrete eigenproblem are the zeros of

$$\frac{J_\nu(x)}{J_{-\nu}(x)} = (-1)^N \tan\left(x - \frac{N\pi}{2}\right),$$

where $\nu = N + 1/2$ and where J_ν stands for the spherical Bessel's function of order ν. Figure 1 illustrates the behaviour of $\sqrt{\lambda_{N,k}}$, for $N = 40$, in comparison with $\sqrt{\lambda_k} = k\pi$. The first $2/\pi$ fraction of them are very well approximated, as proved in [128].

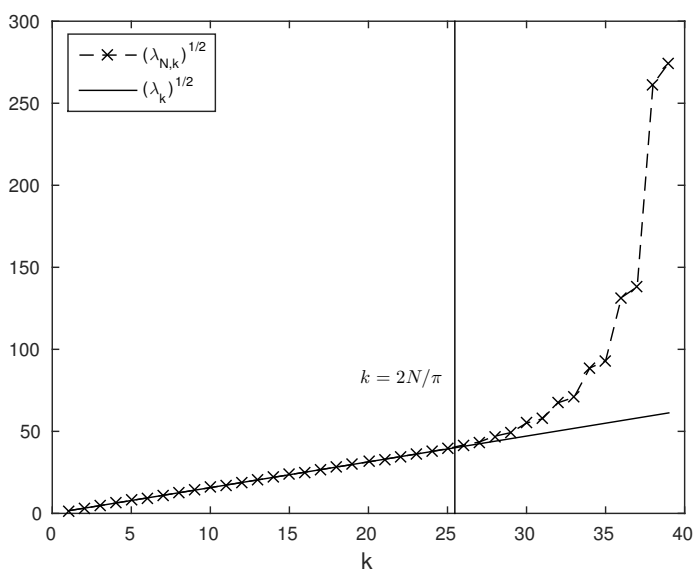

FIGURE 1 – Comparaison between the behaviour of the square roots of the discrete eigenvalues and the behaviour of the continuous ones for $N = 40$.

The Fourier expansion of the solution of (4.2.2) on the orthonormal basis formed by the discrete normalized (for the $L^2(-1,1)$ norm) eigenfunctions $\phi_{N,k}$ writes

$$u^N(x,t) = \sum_{k=1}^{N-1} \left(\alpha_{N,k} \cos(\sqrt{\lambda_{N,k}}t) + \frac{\beta_{N,k}}{\sqrt{\lambda_{N,k}}} \sin(\sqrt{\lambda_{N,k}}t) \right) \phi_{N,k}(x),$$

where, for $1 \leq k \leq N-1$,

$$\alpha_{N,k} = (u_0^N, \phi_{N,k})_{L^2(-1,1)}, \quad \beta_{N,k} = (u_1^N, \phi_{N,k})_{L^2(-1,1)}.$$

We also consider its Fourier truncation of order M, $1 \leq M \leq N-1$,

$$u^{N,M}(x,t) = \sum_{k=1}^{M} \left(\alpha_{N,k} \cos(\sqrt{\lambda_{N,k}}t) + \frac{\beta_{N,k}}{\sqrt{\lambda_{N,k}}} \sin(\sqrt{\lambda_{N,k}}t) \right) \phi_{N,k}(x).$$

4.2.2 Discrete observability

Considering the discrete observability and direct inequalities (4.1.2) and (4.1.3), we redefine the constant $c_{N,T}$ and $C_{N,T}$ as the best possible ones :

$$c_{N,T} = \inf_{(u_0^N, u_1^N) \in (\mathbb{P}_N^0)^2} \left(\int_0^T \left| u_x^N(1,t) \right|^2 dt \right) \; / \; E(u^N(0)),$$

$$C_{N,T} = \sup_{(u_0^N, u_1^N) \in (\mathbb{P}_N^0)^2} \left(\int_0^T \left| u_x^N(1,t) \right|^2 dt \right) \; / \; E(u^N(0)),$$

As mentioned in the introduction, the observability and direct inequalities do not hold uniformly, as we have for every $T > 0$ ([21]) :

$$c_{N,T} \to 0 \quad \text{and} \quad C_{N,T} \to \infty \quad \text{as} \quad N \to \infty.$$

Figure 2 shows the behaviour of $c_{N,T}$ and $C_{N,T}$ with respect to N for $T = 8$. We refer to the Appendix A.4 for the numerical method used to compute $c_{N,T}$ and $C_{N,T}$.

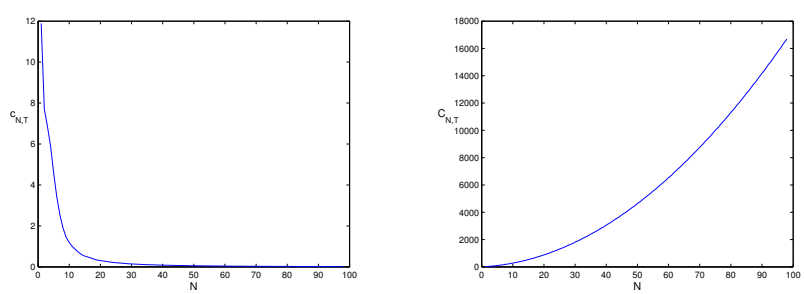

FIGURE 2 – Behaviour of $c_{N,T}$ (left) and $C_{N,T}$ (right) for $N = 3, ..., 250$ and $T = 8$

Nevertheless, if we consider a suitable Fourier truncation u_M^N, which filters out the high frequency components, then we can recover uniform observability and direct inequalities. More precisely, if $T > 4$ and if $M < \frac{2}{\pi}N$, that is if we keep only the $2/\pi$ fraction of the lowest frequency components, then

$$c_T \; E(u_M^N(0)) \; \leq \; \int_0^T \left| \frac{\partial u_M^N}{\partial x}(1,t) \right|^2 dt \; \leq \; C_T \; E(u_M^N(0))$$

where $c_T > 0$ and $C_T > 0$ are independent of N ([21]).

The proof of this result relies on two ingredients : First, the existence of a uniform gap between consecutive eigenvalue square roots,

$$\sqrt{\lambda_{N,k}} - \sqrt{\lambda_{N,k-1}} \geq \delta > 0 \tag{4.2.4}$$

which is proved in [21] to hold as long as $k < \frac{2}{\pi}N$; secondly, on the uniform lower and upper bounds of

$$\Delta_{N,k} = \frac{\left|\phi'_{N,k}(1)\right|^2}{\displaystyle\int_{-1}^{1} \left|\phi'_{N,k}(x)\right|^2 \ \mathrm{dx}},$$

that is

$$c \leq \Delta_{N,k} \leq C, \tag{4.2.5}$$

which also holds as long as $k < \frac{2}{\pi}N$. Note that $\Delta_{N,k}$ has to be compared with its counterpart in the continuous case, which is

$$\Delta_k = \frac{|\phi'_k(1)|^2}{\displaystyle\int_{-1}^{1} |\phi'_k(x)|^2 \ \mathrm{dx}}.$$

A straightforward computation shows that $\Delta_k = 1$, for every $k \geq 1$. Figure 3 illustrates the fact that values of $\Delta_{N,k}$ for $N = 40$ are indeed close to the exact values $\Delta_k = 1$ as long as $k < \frac{2}{\pi}N$;

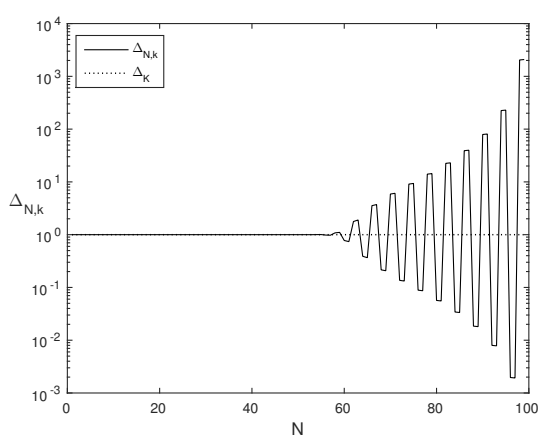

FIGURE 3 – Comparaison between $\Delta_{N,k}$, for $N = 40$, and their continuous counterparts $\Delta_k = 1$

If no filtering is applied, while the uniform gap property (4.2.4) between consecutive eigenvalue square roots seems to hold still (see Figure 1), Figure 3 shows that both the lower and upper uniform bound property on $\Delta_{N,k}$ in (4.2.5) do not hold

anymore, as proved in [21].

Let us point that the situation differs when using finite difference or finite element methods. In the cases, there is no uniform lower gap between consecutive eigenvalue square roots (caused by the highest numerical eigenvalues) causing by itself the non-uniform observability inequality, while a uniform upper bound on the equivalent $\Delta|N,k$ alone ensures a continuous boundary trace inequality.

4.3 Spectral filtering

We have just recalled that the lack of uniform observability is due to the high frequency components of the approximate solution of the wave equation. On another hand, spectral approximation methods, like the Legendre Galerkin method, are known to be very precise when exact solutions are very smooth but also to lose part of their efficiency when solutions are not that smooth. In the latter case, the high frequency components of the approximation are again seen as the main source of errors or instabilities, for instance giving rise to the famous Gibbs phenomenon for discontinuous exact solutions.

A way to recover precision is to dissipate the high modes using *spectral filters* (see for instance [25, 79]). These filters are called *spectral filters* because the filtered approximate solutions are obtained by modifying their spectral (here, polynomial) expansion. Given a spectral expansion

$$u^N(x) = \sum_{k=1}^{N-1} \alpha_{N,k} \widetilde{L}_k(x),$$

the filtered expansion $\mathcal{F}_N u^N$ is

$$\mathcal{F}_N u^N(x) = \sum_{k=1}^{N-1} \sigma\left(\frac{k-1}{N-1}\right) \alpha_{N,k} \widetilde{L}_k(x)$$

where σ is called a *filter function* or simply a *filter*. $\sigma(\eta)$ has values close to 1 for small values of η and close to 0 for η approaching 1.

More precisely, we say that $\sigma : [0,\,1] \to [0,\,1]$ is a *p-th order filter* ($p \geq 1$) if

(i) $\sigma \in C^{p-1}([0,\,1])$,

(ii) $\sigma(0) = 1$, $\sigma^{(j)}(0) = 0$ for $1 \leq j \leq p-1$,

(iii) $\sigma(1) = 0$, $\sigma^{(j)}(1) = 0$ for $1 \leq j \leq p-1$.

Legendre polynomials are the eigenfunctions of a Sturm-Liouville operator arranged in the increasing order of their associated eigenvalues or frequencies. The idea behind this filtering procedure is thus to keep almost unchanged the low frequency

components of the expansion while having a progressive damping on the higher modes.

Note that the filtering is generally performed on the expansion in terms of the Legendre polynomials $L_k(x)$, $k = 0, ..., N$, which are of increasing frequency with increasing k, instead of $\widetilde{L}_k(x)$. But as we have

$$\widetilde{L}_k(x) = \frac{1}{\sqrt{4k+2}}(L_{k-1}(x) - L_{k+1}(x)), \quad k \geq 1, \tag{4.3.1}$$

functions $\widetilde{L}_k(x)$ are still of increasing frequency components with increasing k.

In this work we consider several filter functions, all presented and discussed in [25] :

— **Cesáro filter** (first order)

$$\sigma(\eta) = 1 - \eta,$$

— **Lanczos filter** (first order)

$$\sigma(\eta) = \frac{\sin(\pi\eta)}{\pi\eta},$$

— **Raised cosine filter** (second order)

$$\sigma(\eta) = \frac{1 + \cos(\pi\eta)}{2},$$

— **Sharpened raised cosine filter** (8-th order)

$$\sigma(\eta) = \sigma_0(\eta)^4 \left(35 - 84\sigma_0(\eta) + 70\sigma_0(\eta)^2 - 20\sigma_0(\eta)^3\right),$$

where σ_0 denotes the raised cosine filter.

— **Vandeven filter** (p-th order)

$$\sigma(\eta) = 1 - \frac{(2p-1)!}{(p-1)!^2} \int_0^\eta (t(1-t))^{p-1}\,\mathrm{dt}$$

— **Exponential filter** (p-th order)

$$\sigma(\eta) = e^{-\alpha\eta^p}, \qquad \alpha > 0.$$

The exponential filter does not satisfy condition (iii) of the definition. However, taking $\alpha = -\log(\epsilon_m)$, where ϵ_m is the machine accuracy, the value of $\sigma(1)$ will be computationally interpreted as a 0. The exponential filter has the advantage to be of arbitrary order, like the Vandeven filter, but compared to the latter it has a much lower computational cost.

Vandeven and exponentiel filters in the context of Legendre spectral methods

were studied numerically in [79]. There, it was observed that they behave similarly, they improve accuracy in polynomial expansions of smooth functions and they may be essential to stability for first order wave equations. It was also observed that increasing the order of the filter improves its accuracy.

Figure 4 shows the graphs of the first four filters, whereas Figure 5 shows the graphs of the last two, for different values of p.

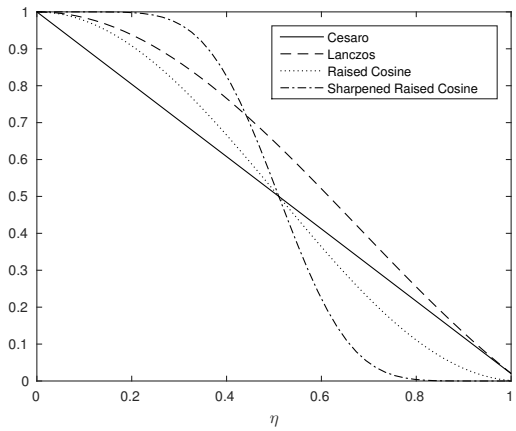

FIGURE 4 – Graphs of Cesàro, Lanczos, raised cosine and sharpened raised cosine filters

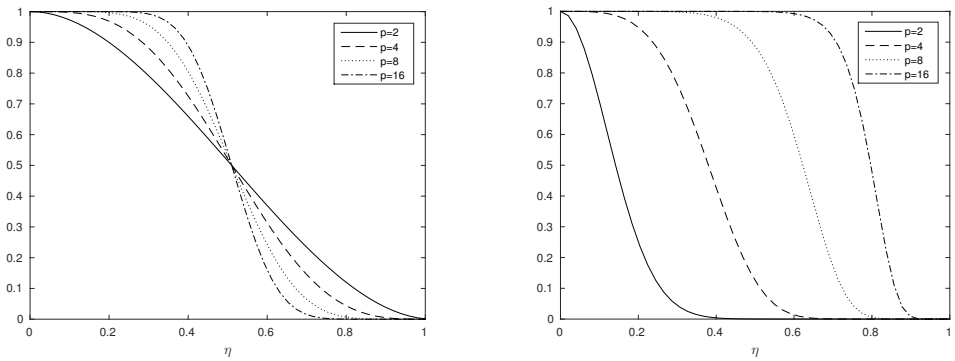

FIGURE 5 – Graphs of Vandeven (left) and exponential (right) filters for different values of p

Regarding the observability problem, we now consider the boundary observation of the filtered expansion $\mathcal{F}_N u^N$ of u^N, that is

$$(\mathcal{F}_N u^N)_x(1,t) = \sum_{k=1}^{N-1} \sigma\left(\frac{k-1}{N-1}\right) a_{N,k}(t)\widetilde{L}'_k(1).$$

$(\mathcal{F}_N u^N)_x(1,t)$ is expected to converge toward $u_x(1,t)$ (for an appropriate norm). The discrete observability problem then reduces to minimizing the functional

$$\begin{aligned} J_{N,\sigma}(u_0^N, u_1^N) &= \frac{1}{2}\int_0^T \left|(\mathcal{F}_N u^N)_x(1,t)\right|^2 \mathrm{dt} \\ &\quad - < (y_1, -y_0), (u_0^N, u_1^N) >_{H^{-1}\times L^2, H_0^1 \times L^2} \end{aligned} \tag{4.3.2}$$

over $(\mathbb{P}_N^0)^2$, and an approximation of the exact control is $v_N(t) = (\mathcal{F}_N u^N)_x(1,t)$, where u^N is the solution of (4.2.2) with the initial data $(u_0^N, u_1^N) \in (\mathbb{P}_N^0)^2$ minimizing (4.3.2).

Note that, if the spectral filtering $(\mathcal{F}_N u^N)_x(1,t)$ is very alike of the Fourier's filtering, the spectral filtering as a low computational cost : no eigenvalues have to be computed, the values $\sigma(k-1/N-1), k = 1, ..., N-1$ of the spectral filters may be stored and $(\mathcal{F}_N u^N)_x(1,t)$ and $\mathcal{F}_N u^N$ may be computed directly as the solution of (4.2.2) with initial data $(\mathcal{F}_N u_0^N, \mathcal{F}_N u_1^N)$.

It is sufficient to obtain uniform bounds of $c_{N,T}$ and $C_{N,T}$ appearing in

$$c_{N,T} E(u^N(0)) \leq \int_0^T \left|\left(\mathcal{F}_N u^N\right)_x (1,t)\right|^2 dt \tag{4.3.3}$$

$$\int_0^T \left|\left(\mathcal{F}_N u^N\right)_x (1,t)\right|^2 dt \leq C_{N,T} E(u^N(0)). \tag{4.3.4}$$

to show the convergence of the numerical controls.

Figure 6 illustrates the behaviour of $c_{N,T}$ and $C_{N,T}$ of (4.3.3)-(4.3.4) for different filters

It seems, according to the numerical results, that it is possible to obtain an uniform lower bound of $c_{N,T}$ for the spectral filters considered precedently. However, none of them, even for an appropriate value of p, seems to allow us to obtain an upper-bound independent of N of $C_{N,T}$. Figure 7 shows the best result concerning the uniform boundedness of the filtered $\Delta_{N,k}$ obtained with exponential smoothing.

$$\Delta_{N,k} = \frac{\left|(\mathcal{F}_N \phi_{N,k})'(1)\right|^2}{\displaystyle\int_{-1}^1 \left|\phi'_{N,k}(x)\right|^2 \ \mathrm{dx}}.$$

The boundedness of $C_{N,T}$, for small values of N, is obtained at the cost of the loss of quality of the bound of $c_{N,T}$. Finally, $C_{N,T}$ goes to infinity when N does.

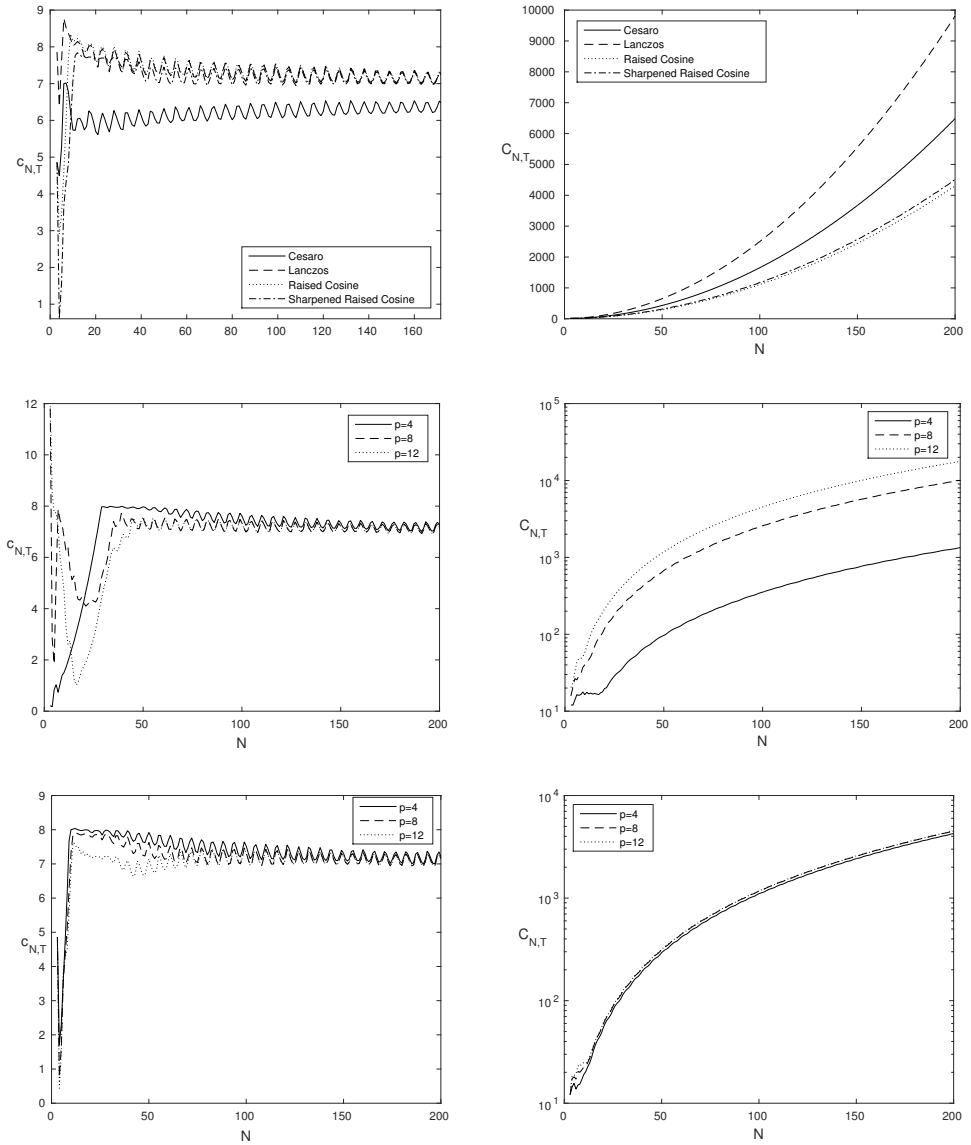

FIGURE 6 – Values of $c_{N,T}$ (left column) and $C_{N,T}$ (right column) for differents filters applied to (4.2.2) for $T = 8$

4.4 Mixed Legendre-Galerkin method

Consider the following mixed formulation of the wave equation (4.1.1).

Let $z = u_t$. We seek for a solution $(u, z) \in H_0^1(-1, 1) \times L^2(-1, 1)$ of

$$\int_{-1}^{1} u_t(x,t)\widetilde{v}(x)\,\mathrm{dx} = \int_{-1}^{1} z(x,t)\widetilde{v}(x)\,\mathrm{dx} \qquad \forall \widetilde{v} \in L^2(-1,1) \tag{4.4.1}$$

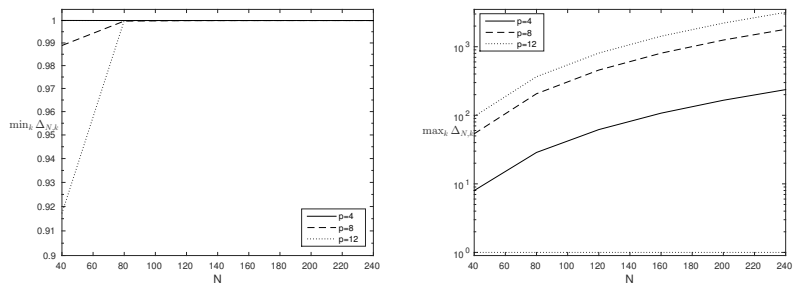

FIGURE 7 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) for the exponential smoothing applied to (4.2.2) for $p = 2$, $N = 3, .., 250$ and $T = 8$

$$\frac{d}{dt} < z(t), \hat{v} >_{H^{-1},H_0^1} = -\int_{-1}^{1} u_x(x,t)\hat{v}_x(x)\,\mathrm{dx} \qquad \forall \hat{v} \in H_0^1(-1,1). \tag{4.4.2}$$

The mixed Legendre-Galerkin method consists to approximate the solution (u, z) (4.4.1)-(4.4.2) by $(u^N, z^M) \in \mathbb{P}_0^N \times \mathbb{P}^M$, where $\mathbb{P}^M = \mathrm{span}\{L_0(x), \ldots, L_M(x)\}$, solutions of

$$\int_{-1}^{1} u_t^N(x,t)\widetilde{v}^M(x)\,\mathrm{dx} = \int_{-1}^{1} z^M(x,t)\widetilde{v}^M(x)\,\mathrm{dx} \qquad \forall \widetilde{v}^M \in \mathbb{P}^M \tag{4.4.3}$$

$$\frac{d}{dt} < z^M(t), \hat{v}^N >_{H^{-1},H_0^1} = -\int_{-1}^{1} u_x^N(x,t)\hat{v}_x^N(x)\,\mathrm{dx} \qquad \forall \hat{v}^N \in \mathbb{P}_0^N. \tag{4.4.4}$$

Writing

$$\begin{aligned} u^N(x,t) &= \sum_{k=1}^{N-1} a_{N,k}(t)\widetilde{L_k}(x) \\ z^M(x,t) &= \sum_{k=0}^{M} b_{M,k}(t)L_k(x), \end{aligned}$$

and by letting $\mathbf{a}_N(t) =: (a_{N,1}(t), \ldots, a_{N,N-1}(t))^t$ and $\mathbf{b}_M(t) =: (b_{M,0}(t), \ldots, b_{M,M}(t))^t$, (4.4.3)-(4.4.4) writes as

$$\begin{pmatrix} \mathbf{a}_N(t)' \\ \mathbf{b}_M(t)' \end{pmatrix} = \begin{pmatrix} D_{N,M} & 0 \\ 0 & D_{N,M}^t \end{pmatrix}^{-1} \begin{pmatrix} 0 & M_M \\ -K_N & 0 \end{pmatrix} \begin{pmatrix} \mathbf{a}_N(t) \\ \mathbf{b}_M(t) \end{pmatrix}$$

where $K_N \in M_{N-1\times N-1}(\mathbb{R}), M_M \in M_{M+1\times M+1}(\mathbb{R}), D_{N,M} \in M_{M+1,N-1}(\mathbb{R})$ are given

by

$$
\begin{aligned}
K_N(i,j) &= \begin{cases} 1 & \text{if i=j} \\ 0 & \text{otherwise} \end{cases} \\
M_M(i,j) &= \begin{cases} \dfrac{2}{2i-1} & \text{if i=j} \\ 0 & \text{otherwise} \end{cases} \\
D_{N,M}(i,j) &= \begin{cases} \dfrac{\sqrt{2}}{\sqrt{2i+1}(2i-1)} & \text{if i=j} \\ \dfrac{-\sqrt{2}}{\sqrt{2j+1}(2j+3)} & \text{if i=j+2} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Note that

$$
\begin{pmatrix} D_{N,M} & 0 \\ 0 & D^t_{N,M} \end{pmatrix}
$$

is invertible if and only if $N-1 = M+1$.

Associated to the mixed formulation (4.4.3)-(4.4.4) is the following eigenproblem : find $(\phi^N, \psi^M) \in \mathbb{P}^N_0 \times \mathbb{P}^M$ and $\lambda^N \in \mathbb{R}$ solution of

$$
(\psi^M, \widetilde{v}^M)_{L^2(-1,1)} = \lambda(\phi^N, \widetilde{v}^M)_{L^2(-1,1)}, \quad \widetilde{v}^M \in \mathbb{P}^M, \tag{4.4.5}
$$

$$
(\phi^N, \hat{v}^N)_{H^1_0(-1,1)} = \lambda(\psi^N, \hat{v}^N)_{L^2(-1,1)}, \quad \hat{v}^N \in \mathbb{P}^N_0. \tag{4.4.6}
$$

Figure 8 illustrates the behaviour of the square roots of the eigenvalues as well as the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}},\ k = 1, \ldots, N-2$. Numerical tests seems to show that there

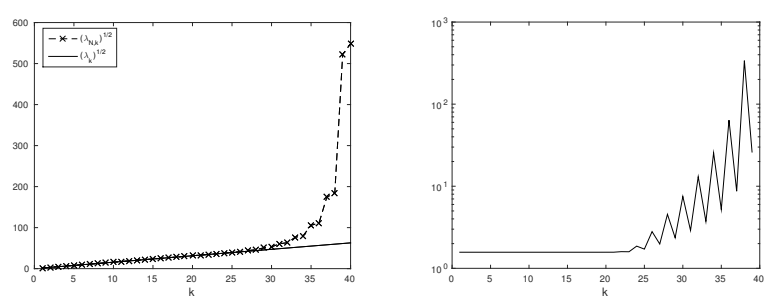

FIGURE 8 – Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.4.5)-(4.4.6) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > 0 k = 1, \ldots, N-2$ (right) for $N = 40$

exist a strictly positive and uniform lower bound of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}},\ k = 1, \ldots, N-2$.

Testing $z^M_t(x,t)$ in (4.4.3) and $u^N_t(x,t)$ in (4.4.4), we deduce that the energy of

the solutions of (4.4.3)-(4.4.4) is conserved along time

$$E(u^N(t), z^M(t)) = E(u^N(0), z^M(0)),$$

where

$$E(u^N(t), z^M(t)) = \frac{1}{2}\int_{-1}^{1} \left|z^M(x,t)\right|^2 + \left|u^N(x,t)\right|^2 \mathrm{dx}.$$

The behaviour of $c_{N,T}$ and $C_{N,T}$ appearing in

$$c_{N,T}E(u^N(0), z^M(0)) \leq \int_0^T \left|u_x^N(1,t)\right|^2 dt$$
$$\int_0^T \left|u_x^N(1,t)\right|^2 dt \leq C_{N,T}E(u^N(0), z^M(0)),$$

is illustrated in the figure 9. Once again, we are able to recover a strictly positive and

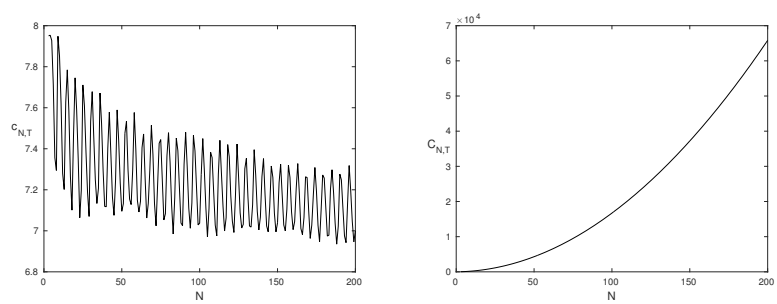

FIGURE 9 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.4.3)-(4.4.4) for $N = 3, .., 250$ and $T = 8$

uniform lower bound of $c_{N,T}$ but we are not able to bound uniformly $C_{N,T}$.

A mixed method was first suggested in [73] as a cure to the lack of a minimal uniform gap of $\sqrt{\lambda_{h,k}} - \sqrt{\lambda_{h,k-1}}$ in the context of finite element method or a finite difference method. It was proven in [26] that the mixed method provides the uniform gap of the discrete eigenvalues as well as the uniform bound on $c_{h,T}$ and $C_{h,T}$ appearing in the following discrete observability inequality and discrete trace inequality

$$c_{h,T}E(u^h(0), z^h(0)) \leq \int_0^T \left|u_x^h(1,t)\right|^2 + h^2\left|z_x^h(1,t)\right|^2 \mathrm{dt} \leq C_{h,T}E(u^h(0), z^h(0)), \quad (4.4.7)$$

It was in fact proved in [26] that (4.4.7) holds for $c_{h,T} = 2(T-2)$ and $C_{h,T} = 2(T+2)$. Note that the uniformness of $c_{h,T}$ and $C_{h,T}$ was only obtained at the cost of adding an extra term in the observation, vanishing as the $h \to 0$. The uniform gap and the minimal uniform bound on $c_{h,T}$ and $C_{h,T}$ result was later generalized by Münch in [105] for a family of mixed finite difference method.

Figure 10 illustrate the behaviour of $c_{h,T}$ and $C_{h,T}$.

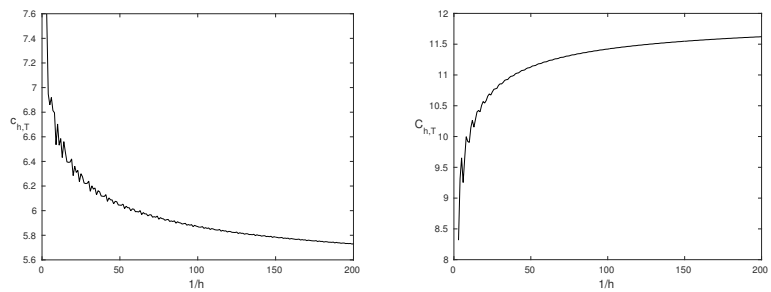

FIGURE 10 – Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) for the finite element mixed method for $1/h = 3, .., 250$ and $T = 4$

Note that, in view of the results of Castro and Micu in [26], figure 10 provide us a numerical example of the behaviour of uniformly bounded $c_{h,T}$ and $C_{h,T}$, thus confirming our assumptions on our previous results.

4.5 Nitsche's method

Nitsche's method was introduced by Nitsche in the 70's in [111] in order to weakly impose the boundary conditions without losing the coercivity of the bilinear elliptic form, a major drawback of the penalization's method. We consider here a modified form of the Nitsche's method, that is, only weakly imposing the end $x = 1$, a more appropriate formulation in the case of discrete observability. Numerical experiments were done in the case of weakly imposed Dirichlet conditions on both ends leading to similar numerical results.

4.5.1 Symmetric case

We consider $\mathbb{P}_L^N = \text{span}\{\widehat{L_1}(x), \ldots, \widehat{L_N}(x)\} \subset H_L^1(-1,1)$ where $H_L^1(-1,1) = \{\psi \in H^1(-1,1) \,|\, \psi(-1) = 0\}$ and where

$$\widehat{L_k}(x) = L_k(x) - (-1)^k, k \geq 1.$$

Introducing

$$\begin{aligned} a^N(u^N, \psi^N) &= \int_{-1}^{1} u_x^N(x,t)\psi_x^N(x)\,\mathrm{dx} - u_x^N(1,t)\psi^N(1) \\ &\quad -u^N(1,t)\psi_x^N(1) + \gamma N^2 u^N(1,t)\psi^N(1), \end{aligned}$$

we are looking for

$$u^N(x,t) = \sum_{k=1}^{N} a_{N,k}(t)\widehat{L_k}(x),$$

solution of

$$\int_{-1}^{1} u_{tt}^N(x,t)\psi^N(x)\,\mathrm{dx} + a^N(u^N,\psi^N) = 0,\ \forall v^N \in \mathbb{P}_L^N. \tag{4.5.1}$$

In [76], it was shown that for a suitable choice of γ, then the numerical solutions of (4.5.1) converge to the exact one when one considers a discontinuous Galerkin method. The restriction on γ ensures the coercivity of the bilinear elliptic form. In (4.5.1), the coercivity of the bilinear elliptic form is insured if $\gamma > 1/2$, a consequence of the inverse inequality ([129])

$$|p^N(1)|^2 \leq \frac{(N+1)^2}{2}|p^N(x)|^2_{H_0^1(-1,1)},\ \forall p^N(x) \in \mathbb{P}^N(-1,1)$$

Let's denote $a_N(u^N,\psi^N)$ the bilinear elliptic form

$$\begin{aligned} a^N(\phi^N,\psi^N) \ &:= \ \int_{-1}^{1} \phi_x^N(x)\psi_x^N(x)\,\mathrm{dx} - \phi_x^N(1)\psi^N(1) - \phi^N(1)\psi_x^N(1) \\ &\quad + \gamma N^2\phi^N(1)\psi^N(1). \end{aligned}$$

The coercivity easily follows,

$$\begin{aligned} a^N(\psi^N,\psi^N) &= |\psi^N|^2_{H_0^1} - 2\psi_x^N(1)\psi^N(1) + \gamma N^2\left|\psi^N(1)\right|^2 \\ &\geq |\psi^N|^2_{H_0^1} - 2\left|N^{-1}\psi_x^N(1)\right|\left|N\psi^N(1)\right| + \gamma N^2\left|\psi^N(1)\right|^2 \\ &\geq |\psi^N|^2_{H_0^1} - \frac{1}{\epsilon}\left|N^{-1}\psi_x^N(1)\right|^2 - \epsilon\left|N\psi^N(1)\right|^2 + \gamma N^2\left|\psi^N(1)\right|^2, \quad \forall\epsilon > 0 \\ &\geq (1-\frac{1}{2\epsilon})|\psi^N|^2_{H_0^1} + (\gamma-\epsilon)\left|N\psi^N(1)\right|^2, \quad \forall\epsilon > 0 \\ &\geq C\|\psi^N\|^2_{\mathbb{P}_L^N}, \end{aligned}$$

for $C > 0$ if $\gamma > \epsilon > 1/2$, where $\|.\|_{\mathbb{P}_L^N}$ denotes the norm on the $\mathbb{P}_L^N$ space

$$\|\psi^N\|^2_{\mathbb{P}_L^N} := \left|\psi^N\right|^2_{H_0^1(-1,1)} + N^2\left|\psi(1)\right|^2.$$

The system (4.5.1) writes as the linear system

$$M_N\mathbf{a}_N''(t) + K_N(\gamma)\mathbf{a}_N(t) = 0$$

where $\mathbf{a}_N(t) = (a_{N,1}(t), \ldots, a_{N,N-1}(t))^t$ and where $(M_N, K_\gamma(N)) \in M_{N\times N}(\mathbb{R})^2$ are given by

$$\begin{aligned}
M_N(i,j) &= \begin{cases} \frac{4i+4}{2i+1} & \text{if i=j} \\ -2 & \text{if } i+j \text{ odd} \\ 2 & \text{if } i+j \text{ even and } i \neq j \end{cases} \\
K_N(\gamma) &= K_{N,1} - K_{N,2} - K_{N,2}^t + \gamma N^2 K_{N,3}, \\
K_{N,1}(i,j) &= \begin{cases} r^2 + r & \text{if } i+j \text{ even} \\ 0 & \text{otherwise} \end{cases} \\
K_{N,2}(i,j) &= \begin{cases} j^2 + j & \text{if } i \text{ odd} \\ 0 & \text{otherwise} \end{cases} \\
K_{N,3}(i,j) &= \begin{cases} 4 & \text{if } i,j \text{ are odds} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Numerically, we have that the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}},\ k = 1, \ldots, N-1$ is uniformly bounded for every values of γ larger than $1/2$. Figure 11 illustrates the case $N = 40$ and $\gamma = 0.8$.

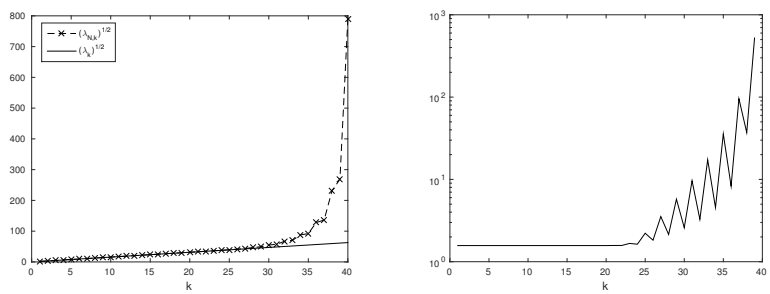

FIGURE 11 – Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.1) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 0.8$

By testing $u_t^N(x,t)$ in (4.5.1), we obtain that the energy of the numerical solutions of (4.5.1)

$$\begin{aligned}
E_{N,\gamma}(u^N(t)) = \frac{1}{2}\int_{-1}^{1} \left|u_t^N(x,t)\right|^2 + \left|u_x^N(x,t)\right|^2 \mathrm{dx} \\
+ \left(\frac{\gamma}{2} N^2 u^N(1,t) - u_x^N(1,t)\right) u^N(1,t),
\end{aligned}$$

is preserved along time,

$$E_{N,\gamma}(u^N(t)) = E_{N,\gamma}(u^N(0)).$$

The uniform results of $c_{N,T}$ rely on the particular form of the observability inequality

$$c_{N,T} E_{N,\gamma}(u^N(0)) \leq \int_0^T \left|u_x^N(1,t) - \gamma N^2 u^N(1,t)\right|^2 \mathrm{dt}. \tag{4.5.2}$$

Consequently, we consider the modified form of the continuity inequality

$$\int_0^T \left|u_x^N(1,t) - \gamma N^2 u^N(1,t)\right|^2 \mathrm{dt} \leq C_{N,T} E_{N,\gamma}(u^N(0)). \tag{4.5.3}$$

These inequalities arise when one considers, not the discrete analogue of the functional $J(u_0, u_1)$, but the discrete one following from (4.5.1). To make this point more precise, let's consider the analogue steps made in [95] to obtain the functional $J(u_0, u_1)$ in the case of Nitsche's method. To do so, let's formally consider $y^N(x,t)$, the numerical approximation of $y(x,t)$, solution of

$$\begin{aligned}\int_{-1}^1 y_{tt}^N(x,t)\psi^N(x)\,\mathrm{dx} + \int_{-1}^1 y_x^N(x,t)\psi_x^N(x)\,\mathrm{dx} - y_x^N(1,t)\psi^N(1) - y^N(1,t)\psi_x^N(1)\\ +\gamma N^2 y^N(1,t)\psi^N(1) = v(t)\left(-\psi_x^N(1) + \gamma N^2\psi^N(1)\right), \quad \forall v^N \in \mathbb{P}_L^N.\end{aligned} \tag{4.5.4}$$

Let's also consider, for the moment, that $v(t)$ is a control, that is $y^N(x,T) = y_t^N(x,T) = 0$. Now let, in (4.5.4), $\psi^N = u^N$, where $u^N(x,t)$ is the solution of (4.5.1), and let's integrate this relation in time. By integrating by parts twice in time the first integral, we obtain

$$\begin{aligned}\int_0^T\int_{-1}^1 y^N(x,t)u_{tt}^N(x,t)\,\mathrm{dxdt} + \int_0^T\int_{-1}^1 y_x^N(x,t)u_x^N(x,t)\,\mathrm{dxdt} - \int_0^T y_x^N(1,t)u^N(1,t)\,\mathrm{dt}\\ - \int_0^T y^N(1,t)u_x^N(1,t)\,\mathrm{dt} + \gamma N^2\int_0^T y^N(1,t)u^N(1,t)\,\mathrm{dt}\\ = \int_0^T v(t)\left(-u_x^N(1,t) + \gamma N^2 u^N(1,t)\right)\,\mathrm{dt} + \int_{-1}^1 u_0^N(x)y_1^N(x) - u_1^N(x)y_0^N(x)\,\mathrm{dx},\end{aligned}$$

The left-hand side corresponds to the left-hand side of (4.5.1) where $\psi^N = y^N$. Thus, we obtain

$$\int_0^T v(t)\left(-u_x^N(1,t) + \gamma N^2 u^N(1,t)\right)\,\mathrm{dt} = \int_{-1}^1 u_1^N(x)y_0^N(x) - u_0^N(x)y_1^N(x)\,\mathrm{dx}. \tag{4.5.5}$$

Then, we know, from the definition of controllability we introduced earlier, that $v(t)$ is a control of (4.5.4) if (4.5.5) is satisfied. We also know that the control $v(t) = u_x^N(1,t) - \gamma N^2 u^N(1,t)$ is given by the minimum of

$$\inf_{(u_0^N, u_1^N)\in(\mathbb{P}_L^N)^2} J^N(u_0^N, u_1^N),$$

where $J^N(u_0, u_1)$ is the quadratic functional

$$J^N(u_0^N, u_1^N) \;=\; \frac{1}{2}\int_0^T \left|-u_x^N(1,t) + \gamma N^2 u^N(1,t)\right|^2 \mathrm{dt}$$

$$- < (y_1, -y_0), (u_0^N, u_1^N) >_{H^{-1}\times L^2, H_0^1 \times L^2} \cdot$$

Uniform and strictly positive lower and upper bounds of $c_{N,T}$ and $C_{N,T}$ appearing in inequalities (4.5.2) and (4.5.3) allow us to show the convergence of numerical controls.

Note that this method of obtaining $J^N(u_0^N, u_1^N)$ is not possible in general since $y^N(x,t)$ is not included in the approximation space.

Figure 12 illustrate the behaviour of $c_{N,T}$ and $C_{N,T}$ in (4.5.2)-(4.5.3) for the particular case $\gamma = 0.8$.

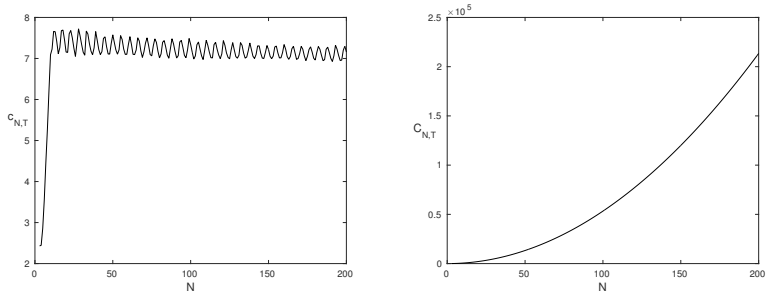

FIGURE 12 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.2)-(4.5.3) for $N = 3, .., 250$, $\gamma = 0.8$ and $T = 8$

We obtain numerically an uniform lower bound of $c_{N,T}$ while the values of $C_{N,T}$ tend to infinity when N does. To show numerically that the results on $c_{N,T}$ are stable by means of γ and to show that no values of γ allows one to recover an uniform bound on $C_{N,T}$ we illustrate in figure 13, for different values of N, the values of $c_{N,T}$ and $C_{N,T}$ when γ varies.

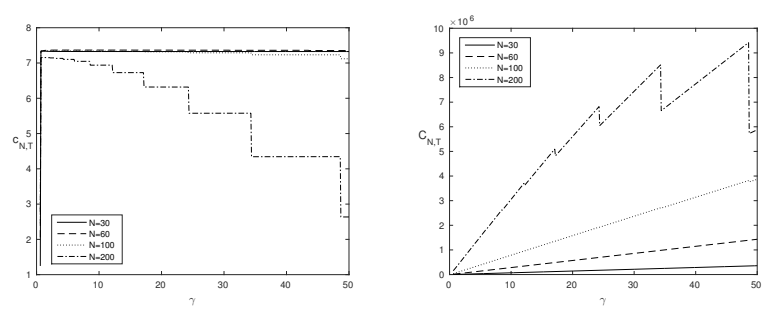

FIGURE 13 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.2)-(4.5.3) for $\gamma \in [0.8, 50]$, $N = 30, 60, 100, 200$ and $T = 8$

We conclude that one cannot expect to construct an appropriate function $\gamma(N)$, satisfying the coercivity condition $\gamma(N) > 1/2$, allowing to bound uniformly the values of $C_{N,T}$ when N goes to infinity.

Note that the term $\gamma N^2 u^N(1,t)$ is necessary to obtain the uniform boundedness of $c_{N,T}$. Indeed, considering the discrete analogue of $J(u_0, u_1)$, that is

$$J^N(u_0^N, u_1^N) = \frac{1}{2}\int_0^T \left|u_x^N(1,t)\right|^2 \mathrm{dt} - < (y_1, -y_0), (u_0^N, u_1^N) >_{H^{-1}\times L^2, H_0^1 \times L^2}, \quad (4.5.6)$$

then we are unable to obtain uniform boudedness on $c_{N,T}$ and $C_{N,T}$. Figure 14 illustrate the comparaison between the behaviour of the $\Delta_{N,k}$ when the term $\gamma N^2 u^N(1,t)$ is taken into account for $\gamma = 0.8$

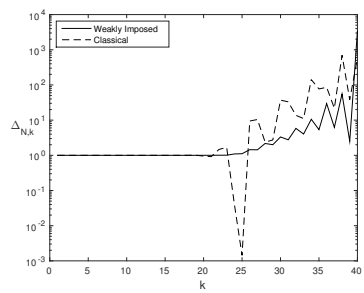

FIGURE 14 – Comparaison between the behaviour of the $\Delta_{N,k}$ when the term $\gamma N^2 u^N(1,t)$ is taken into account for $\gamma = 0.8$, $N = 40$ and $T = 8$

4.5.2 Non-symmetric case

The non-symmetric Nitsche's method consists to find $u^N(t) \in \mathbb{P}_L^N$ such that

$$\int_{-1}^{1} u_{tt}^N(x,t)v^N(x)\mathrm{dx} + \int_{-1}^{1} u_x^N(x,t)v_x^N(x)\mathrm{dx} - u_x^N(1,t)v^N(1)$$
$$+u^N(1,t)v_x^N(1) + \gamma N^2 u^N(1,t)v^N(1) = 0, \forall v^N \in \mathbb{P}_L^N. \quad (4.5.7)$$

where γ is still a strictly positive parameter at our disposal. In the non-symmetric Nitsche's method, the term $u^N(1,t)v_x^N(1)$ is added to the bilinear elliptic term rather than subtracted in order to insure the coercivity without the need of an inverse inequality on functions lying in the approximation space. Indeed, by denoting the bilinear elliptic form

$$\begin{aligned} a^N(u^N, \psi^N) \quad := \quad & \int_{-1}^{1} u_x^N(x,t)\psi_x^N(x)\,\mathrm{dx} - u_x^N(1,t)\psi^N(1) + u^N(1,t)\psi_x^N(1) \\ & + \gamma N^2 u^N(1,t)\psi^N(1), \end{aligned}$$

one directly has that

$$a^N(\psi^N, \psi^N) = \|\psi^N\|^2_{\mathbb{P}_L^N}.$$

Even if the coercivity of the bilinear elliptic form is rather easy issue, the question of the convergence of the numerical solutions of (4.5.7) is, to our knowledge, still open. The loss of symmetry of the bilinear elliptic form prevent us to use the classical energy proof to show the convergence of the numerical solutions. Alternative methods exist when the bilinear elliptic form is not symmetric, we can think of the one presented in [96] for instance, but difficulties arise quickly since the space of approximation $\mathbb{P}_L^1$ is not embedded and since the bilinear elliptic form depends of the discretization parameter. The convergence of the non-symmetric Nitsche's method was at least proved in the case of an elliptic system using inf-sup argument. We still think that the convergence of the numerical solutions of (4.5.7) is insured.

Let $\mathbf{a}_N(t) = (a_{N,1}(t), \ldots, a_{N,N}(t))^t$. Then, (4.5.7) writes as the linear system

$$M_N \mathbf{a}_N''(t) + K_N(\gamma)\mathbf{a}_N(t) = 0$$

where the matrixes $(M_N, K_N(\gamma)) \in M_{N\times N}(\mathbb{R})^2$, $K_N(\gamma) = K_1 - K_2 + K_2^t + \gamma N^2 K_3$, were given in the previous subsection.

Numerically, the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}}$, $k = 1, \ldots, N-1$ is uniformly bounded for every values of γ larger than 1. Figure 15 illustrate the case $N = 40$ and $\gamma = 1$.

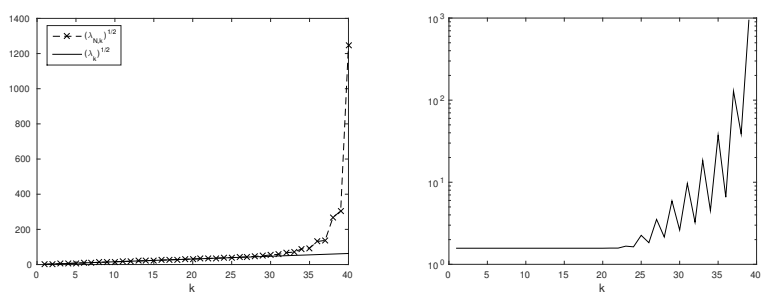

FIGURE 15 – Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 1$

The loss of symmetry of the bilinear elliptic form forces us to consider the following observability and continuity inequalities

$$c_{N,T} E(u^N(0)) \leq \int_0^T \left| u_x^N(1,t) + \gamma N^2 u^N(1,t) \right|^2 \mathrm{dt}, \tag{4.5.8}$$

$$\int_0^T \left| u_x^N(1,t) + \gamma N^2 u^N(1,t) \right|^2 \mathrm{dt} \leq C_{N,T} E(u^N(0)). \tag{4.5.9}$$

where $E(.)$ stands for the energy of the exact case (4.1.1). Even if the energy of the solutions of (4.5.7) is not preserved along time, $E(.)$ still is a norm on the approxima-

tion space $\mathbb{P}_L^N$. Thus, if these inequalities are uniformly bounded by strictly positive real numbers, these inequalities should suffice to show the convergence of the numerical controls. Note that the observation $u_x^N(1,t) + \gamma N^2 u^N(1,t)$ was obtained by the exact same way that in the previous subsection.

Figure 16 illustrate the behaviour of $c_{N,T}$ and $C_{N,T}$ for $\gamma = 1$.

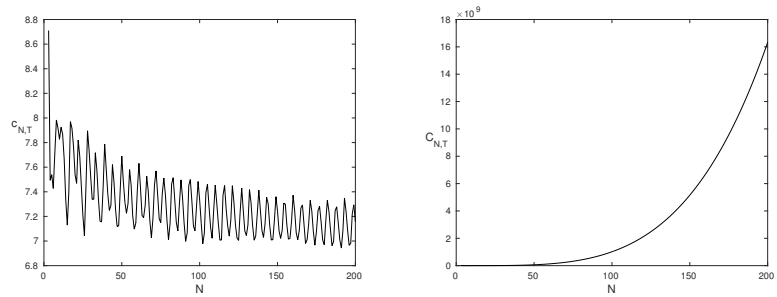

FIGURE 16 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.8)-(4.5.9) for $N = 3, .., 250$, $\gamma = 1$ and $T = 8$

The numerical result shows that we are able to obtain an uniform, by means of N, lower bound of $c_{N,T}$ for $\gamma = 1$ while it seems impossible to bound uniformly the values of $C_{N,T}$. As in the symmetric case, we confirm that the numerical results on the uniform bound of $c_{N,T}$ are stable by means of γ and that there exist no choice of $\gamma(N) > 0$ allowing us to obtain the uniform boundedness of $C_{N,T}$ (figure 17).

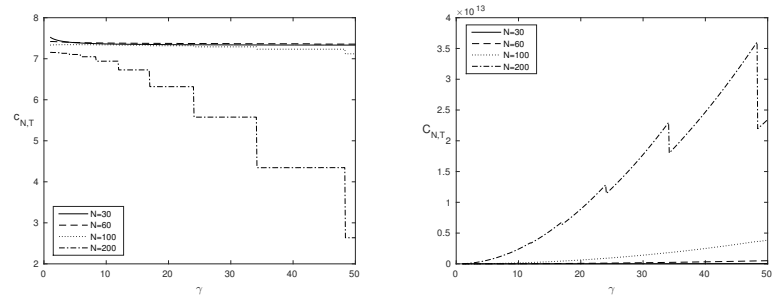

FIGURE 17 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.8)-(4.5.9) for $\gamma \in [1, 50]$, $N = 30, 60, 100, 200$ and $T = 8$

Note that the term $\gamma N^2 u^N(1,t)$ is necessary to obtain an uniform and strictly positive lower bound of $c_{N,T}$ and that removing this term of the observation does not allows one to recover an uniform upper bound of $C_{N,T}$. Figure 18 illustrate the case where $\gamma = 1$ but similar results are obtained with different values of γ.

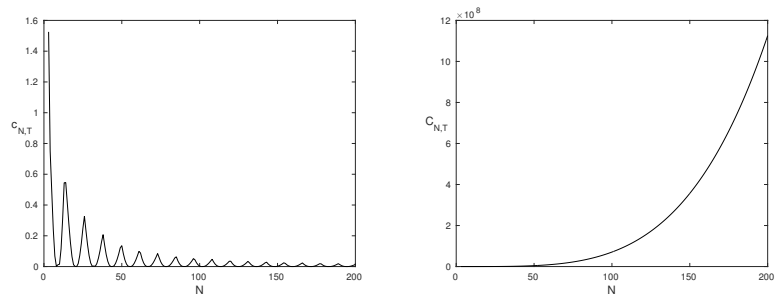

FIGURE 18 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) for $N = 3, .., 250$, $\gamma = 1$ and $T = 8$

4.5.3 Finite Element Method

Let's now compare the results obtained in the previous subsections with those obtained when we consider a finite element semi-discretization of the Nitsche's method. Let's recall that, for this case, we consider the space domain $(0, 1)$ for (4.1.4) and (4.1.1) and an uniform mesh $x_i = ih, i = 0, ..., N$ where $h = 1/N$, $N \geq 1$. We use the approximation space $\mathbb{P}_h^1 = \text{span}\{\psi_1(x), \ldots, \psi_N(x)\}$ where,

$$\psi_j(x) = \begin{cases} (x - x_{j-1})/h & x \in (x_{j-1}, x_j) \\ (x_{j+1} - x)/h & x \in (x_{j-1}, x_j) \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, \ldots, N-1$, and

$$\psi_N(x) = \begin{cases} (x - x_{N-1})/h & x \in (x_{N-1}, x_N) \\ 0 & \text{otherwise} \end{cases}$$

4.5.3.1 Symmetric case

The finite element $\mathbb{P}_h^1$ semi-discretization of Nitsche's method consists to find $u(t) \in \mathbb{P}_h^1$ such that

$$\int_0^1 u_{tt}^h(x,t)\psi^h(x)\,\mathrm{dx} + \int_0^1 u^h(x,t)_x\psi_x^h(x)\,\mathrm{dx} - u_x^h(1,t)\psi^h(1)$$
$$-u^h(1,t)\psi_x^h(1) + \gamma h^{-1}u^h(1,t)\psi^h(1) = 0, \quad \forall \psi^h \in \mathbb{P}_h^1, \qquad (4.5.10)$$

As in the spectral case, the term γh^{-1} is here to insure the coercivity of the bilinear form. We have the following inverse inequality on $\mathbb{P}_h^1$.

Lemma 4.5.1 *For $\psi^h \in \mathbb{P}_h^1$, we have*

$$\left|h^{-1/2}\psi_x^h(1)\right|^2 \leq |\psi^h|^2_{H_0^1(0,1)}$$

Proof :

Let $\psi^h \in \mathbb{P}_h^1$. We may write ψ^h as

$$\psi^h(x) = \sum_{i=1}^{N} a_i \psi_i(x).$$

which gives us that

$$\psi_x^h(x) = \frac{a_1}{h} \mathbb{1}_{(x_0,x_1)}(x) + \sum_{i=2}^{N} \frac{a_i - a_{i-1}}{h} \mathbb{1}_{(x_{i-1},x_i)}(x).$$

We then have

$$\begin{aligned}
\left|h^{-1/2}\psi_x^h(1)\right|^2 &= \left|\frac{a_N - a_{N-1}}{h^{1/2}}\right|^2 \\
&= \frac{1}{h}\left|a_N - a_{N-1}\right|^2 \\
&= \int_0^1 \left|a_N - a_{N-1}\right|^2 \mathbb{1}_{(x_N,x_{N-1})}(x)\,\mathrm{dx} \\
&= \int_0^1 \left|a_N - a_{N-1}\mathbb{1}_{(x_N,x_{N-1})}(x)\right|^2 \mathrm{dx}
\end{aligned}$$

Since $1/h \geq 1$, we finally obtain

$$\begin{aligned}
|h^{-1/2}\psi_x^h(1)|^2 &\leq \int_0^1 \left|\frac{a_N - a_{N-1}}{h}\mathbb{1}_{(x_N,x_{N-1})}(x)\right|^2 \mathrm{dx} \\
&\leq \int_0^1 \left|\frac{a_N - a_{N-1}}{h}\mathbb{1}_{(x_N,x_{N-1})}(x)\right|^2 \mathrm{dx} \\
&\quad + \int_0^1 \left|\frac{a_1}{h}\mathbb{1}_{[x_0,x_1]}(x) + \sum_{i=2}^{N-1} \frac{a_i - a_{i-1}}{h}\mathbb{1}_{[x_{i-1},x_i]}(x)\right|^2 \mathrm{dx} \\
&= \int_0^1 \left|\frac{a_1}{h}\mathbb{1}_{[x_0,x_1]}(x) + \sum_{i=2}^{N} \frac{a_i - a_{i-1}}{h}\mathbb{1}_{[x_{i-1},x_i]}(x)\right|^2 \\
&= |v|^2_{H_0^1(0,1)}
\end{aligned}$$

□

The convergence of the numerical scheme, following [76], is then obtained if $\gamma > 1$. Note that (4.5.10) may write as the following linear system

$$M_h \mathbf{a}_h''(t) + K_h(\gamma)\mathbf{a}_h(t) = 0.$$

where

$$M_h(i,j) = \begin{cases} \frac{2h}{3} & \text{if i=j}, i \neq N \\ \frac{h}{3} & \text{if i=j=N} \\ \frac{h}{6} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

and $K_h(\gamma) = K_{h,1} - K_{h,2} - K_{h,2}^t + \gamma h^{-1} K_{h,3}$ is given by

$$\begin{aligned} K_{h,1}(i,j) &= \begin{cases} \frac{2}{h} & \text{if } i = j, i \neq N \\ \frac{1}{h} & \text{if } i = j = N \\ \frac{-1}{h} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases} \\ K_{h,2}(i,j) &= \begin{cases} -\frac{1}{h} & \text{if } i = N, j = N-1 \\ \frac{1}{h} & \text{if } i = j = N \\ 0 & \text{otherwise} \end{cases} \\ K_{h,3}(i,j) &= \begin{cases} 1 & \text{if } i = j = N \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Figure 19 illustrate the loss of the uniform boundedness of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}},\ k = 1, \ldots, N-1$ in the case where $\gamma = 3.85$ and $N = 40$. Similar results are obtained with different values of γ.

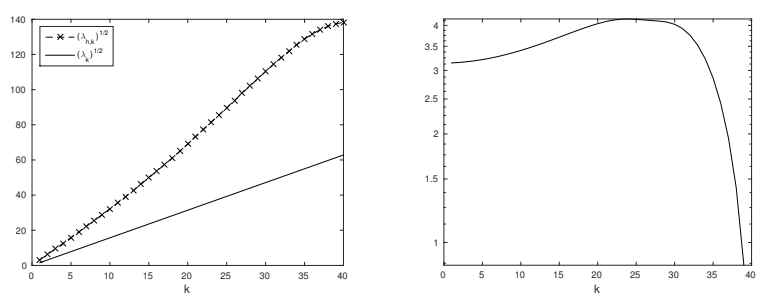

FIGURE 19 – Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 3$

In view of this numerical results, we know that there exists sequences of numerical solution such that the speed of propagation of these solutions tends to 0 as h tends to 0, hence resulting in the inability to obtain an uniform lower bound of $c_{h,T}$.

In order to compare the results obtained in the previous subsections, we consider the following two inequalities.

$$c_{h,T} E_{h,\gamma}(u^h(0)) \leq \int_0^T \left| u_x^h(1,t) - \gamma h^{-1} u^h(1,t) \right|^2 \mathrm{dt} \leq C_{h,T} E_{h,\gamma}(u^h(0)), \qquad (4.5.11)$$

where

$$E_{h,\gamma}(u^h(t)) := \frac{1}{2}\int_0^1 \left|u_t^h(x,t)\right|^2 + \left|u_x^h(x,t)\right|^2 \mathrm{dx} + \left(\frac{\gamma}{2}h^{-1}u^h(1,t) - u_x^h(1,t)\right)u^h(1,t),$$

and where $E_{h,\gamma}(u^h(t)) = E_{h,\gamma}(u^h(0))$. These inequalities may be obtained in the same fashion as in the spectral case.

As expected, the impossibility to obtain an uniform lower bound on the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > \omega, k = 1, \ldots, N-1$ prevent us to obtain an uniform lower bound of $c_{h,T}$. Figure 20 illustrate, for N fixed, values of $c_{h,T}$ and $C_{h,T}$ as γ varies.

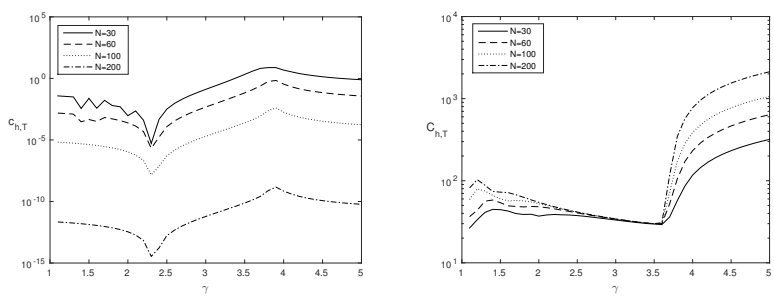

FIGURE 20 – Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) associated to (4.5.11) for $\gamma \in [1.1, 5]$, $N = 30, 60, 100, 200$ and $T = 4$

We see however that it seems possible to obtain an uniform upper bound of $C_{h,T}$. Figure 21 illustrate the case where $\gamma = 3$.

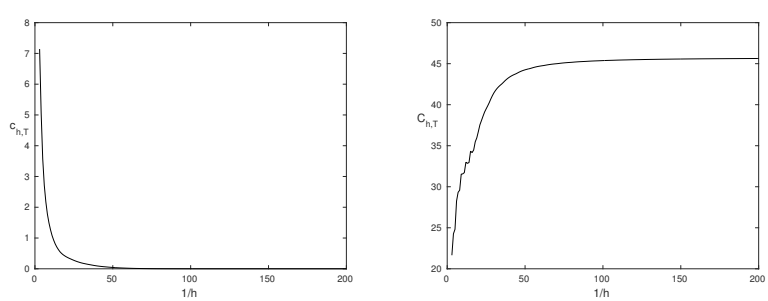

FIGURE 21 – Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.11) for $N = 3, .., 250$, $\gamma = 3$ and $T = 4$

4.5.3.2 Non-symmetric case

The non-symmetric case consist to find $u(t) \in \mathbb{P}_h^1$ solution of

$$\int_0^1 u_{tt}^h(x,t)\psi^h(x)\,\mathrm{dx} + \int_0^1 u_x^h(x,t)\psi_x^h(x)\mathrm{dx} - u_x^h(1,t)\psi^h(1) \qquad (4.5.12)$$
$$+u^h(1,t)\psi_x^h(1) + \gamma h^{-1}u^h(1,t)\psi^h(1) = 0, \forall \psi^h \in \mathbb{P}_h^1,$$

As in the spectral case, there is no need of an inverse inequalities to obtain the coercivity of the bilinear elliptic form. The system (4.5.12) may be written as

$$M_h \mathbf{a}_h''(t) + K_h(\gamma)\mathbf{a}_h(t) = 0.$$

where M_h and $K_h(\gamma) = K_{h,1} - K_{h,2} - K_{h,2}^t + \gamma h^{-1} K_{h,3}$, are given in the previous subsection.

The gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}},\ k = 1, \ldots, N-1$ is not uniformly lower bounded for the non-symmetric formulation. Figure 22 illustrate this loss of uniform boundedness in the case of $N = 40$ and $\gamma = 3$

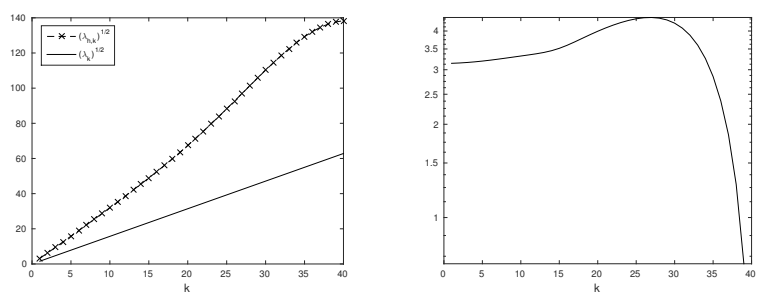

FIGURE 22 – Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 3$

The inequalities of interest writes

$$c_{h,T}E(u^h(0)) \leq \int_0^T \left|u_x^h(1,t) + \gamma h^{-1}u_x^h(1,t)\right|^2 \mathrm{dt} \leq C_{h,T}E(u^h(0)). \qquad (4.5.13)$$

where $E(.)$ stands for the energy of the continuous case (4.1.1).

The loss of the uniform lower bound of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > k = 1, \ldots, N-1$ prevent us to obtain an uniform lower bound of $c_{h,T}$. Figure 23 illustrate, for N fixed, values of $c_{h,T}$ and $C_{h,T}$ as γ varies. It is however difficult to see if it possible to obtain an uniform upper bound on $C_{h,T}$ in this case.

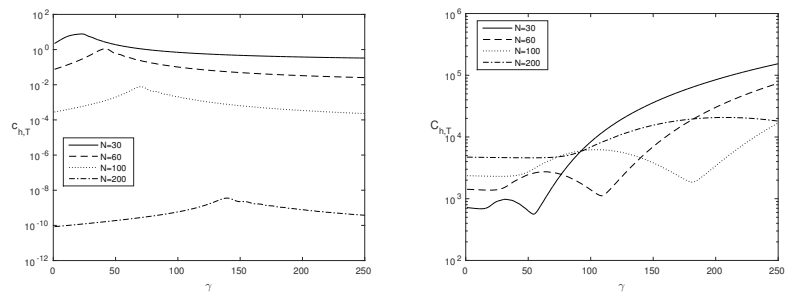

FIGURE 23 – Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) associated to (4.5.13) for $\gamma \in [1, 250]$, $N = 30, 60, 100, 200$ and $T = 4$

4.6 Introduction of the numerical experiments

In the litterature, to prove the convergence of control approximations, it is sufficient to show that the discrete versions of the observability and trace inequalities hold uniformly with respect to the discretization parameter. However, it has never been shown that these conditions were necessary. Micu in [26] presented an argument pointing out the fact that the trace inequality may be not necessary. Our goal for the rest of this paper is to present the results of the numerical investigation on the necessity of these discrete inequalities. In order to do so, we will compute several numerical control associated with different regularities of the initial condition (y_0, y_1) with all the numerical schemes presented in this paper. The interesting comparaison will be between numerical schemes which seems to provide the uniformity of the observability inequality with formulation (4.2.2), for which none of the inequality holds uniformly. The initial data are picked from other papers in order to compare the results with schemes that don't provide the uniformity of the inequalities.

To compare the numerical control with the continuous one, we will take advantage of the following : first, consider the decomposition of the initial data of (4.1.4) on the eigenfunctions,

$$y_0(x) = \sum_{k=1}^{\infty} \widehat{y_{0,k}} \sin(\sqrt{\lambda_k}(x+1)), \qquad y_1(x) = \sum_{k=1}^{\infty} \widehat{y_{1,k}} \sin(\sqrt{\lambda_k}(x+1)).$$

Since $J(u_0, u_1)$ is C^1, we know that the initial data $(\widetilde{u_0}, \widetilde{u_1})$ minimizing the quadratic functional will satisfy $DJ(\widetilde{u_0}, \widetilde{u_1}) = 0$, that is,

$$
\begin{aligned}
< DJ(\widetilde{u}_0, \widetilde{u}_1), (w_0, w_1) > &= \int_0^T \widetilde{u}_x(1,t) w_x(1,t) \mathrm{dt} \\
&\quad - < (y_1, -y_0), (w_0, w_1) >_{H^{-1}\times L^2, H_0^1 \times L^2} \\
&= 0,
\end{aligned}
$$

By decomposing $(\widetilde{u}_0, \widetilde{u}_1)$ and (w_0, w_1) on the eigenfunctions, we obtain

$$
\widehat{\widetilde{u_{0,k}}} = \frac{\widehat{y_{1,k}}}{k^2\pi^2}, \qquad \widehat{\widetilde{u_{1,k}}} = -\frac{\widehat{y_{0,k}}}{4}.
$$

Then, $(\widetilde{u}_0, \widetilde{u}_1)$ and the exact control are given by

$$
\begin{aligned}
u_0(x) &= \sum_{k=1}^{\infty} \frac{\widehat{y_{1,k}}}{k^2\pi^2} \sin(\sqrt{\lambda_k}(x+1)) \\
u_1(x) &= -\frac{y_0}{4} \\
v(t) &= u_x(1,t) = \sum_{k=1}^{\infty} \left(\frac{\widehat{y_{1,k}}}{2k\pi} \cos(\frac{k\pi}{2}t) - \frac{\widehat{y_{0,k}}}{4} \sin(\frac{k\pi}{2}t) \right) (-1)^k.
\end{aligned}
$$

4.6.1 Numerical Experiments

The first example that we consider is from the paper [73], used in the case of a central finite difference scheme of the wave equation.

$$
\begin{aligned}
y_0 &= \begin{cases} 16(x+1)^3 & \text{if } x \in [-1,0) \\ 16(1-x)^3 & \text{if } x \in [0,1], \end{cases} \\
y_1 &= 0
\end{aligned}
\tag{4.6.1}
$$

By decomposing (y_0, y_1) in Fourier series, we obtain the representation of (u_0, u_1) and $v(t)$.

This example was used in [73] to show an example of initial data for which the numerical control converge in the case where both trace and observability inequality hold uniformly ($\Delta t = h$) and the numerical control is of unbounded norm when the observability inequality fail to hold uniformly ($\Delta t < h$). This test was also used in [26] where, even for the case where the observability fail to hold uniformly ($\Delta t < h$), the numerical control converge to the continuous one.

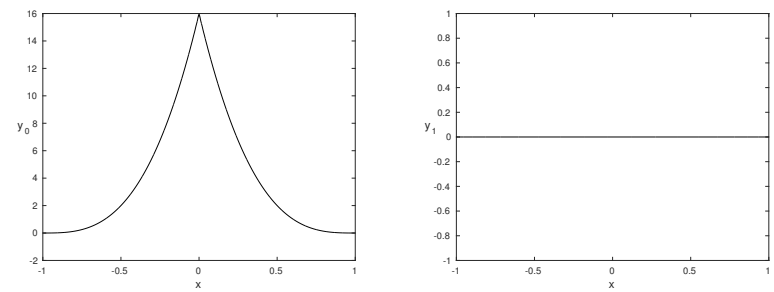

FIGURE 24 – y_0 (left), y_1 (right)

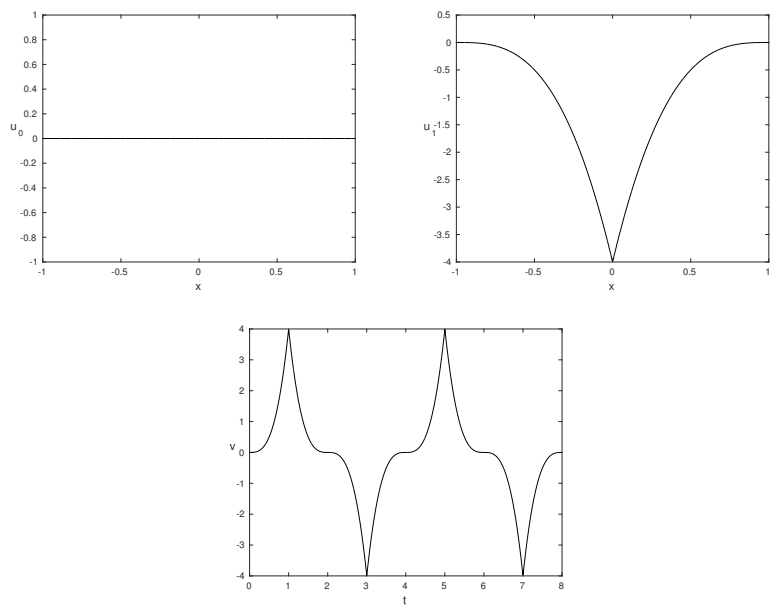

FIGURE 25 – Top : u_0 (left) and u_1 (right) - Bottom : $v(t)$

The convergence of the initial data (u_0^N, u_1^N) seems to be insured for all the scheme presented in this paper, including (4.2.2) for which none of the inequality hold uniformly.

Note that even if the convergence of the control is insured by the formulation (4.2.2), the qualitative result of the approximation of (u_0^N, u_1^N) is quite bad in view of the oscillation the propagate along the domain.

The second initial data that we will investigate is the following

$$\begin{aligned} y_0 &= x+1 \\ y_1 &= 0 \end{aligned} \tag{4.6.2}$$

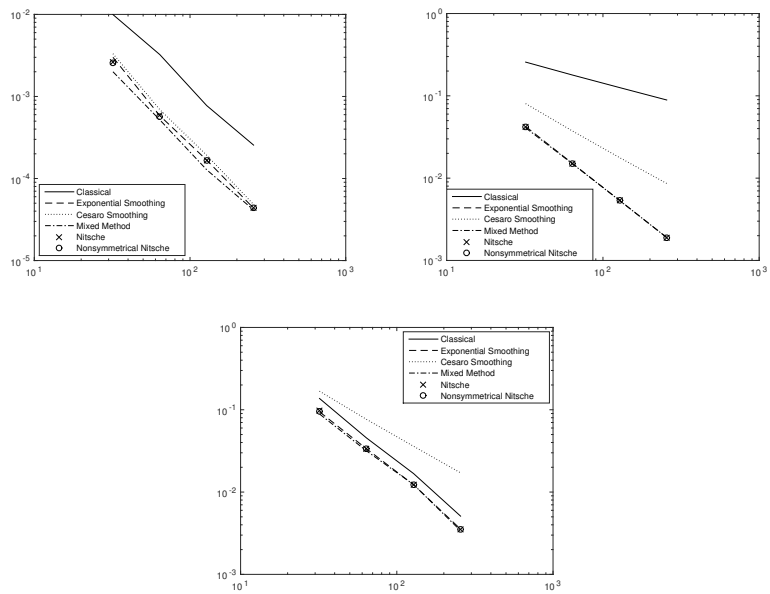

FIGURE 26 – Graphic of the convergence of initial datas u_0^N (top left), u_1^N (top right) and the convergence of the numerical control $v^N(t)$ (bottom) for : (4.2.2) , (4.2.2) filtred with exponential smoothing (p=6), (4.2.2) filtred with Cesàro smoothing, (4.4.3)-(4.4.4), (4.5.1) ($\gamma = 1$), (4.5.7) ($\gamma = 1$) and for values of N of $32, 64, 128, 256$

which gives us (u_0, u_1) and $v(t)$.

Note that the data y_0 has the weakest regularity possible for our problem. Indeed, in order to gain more regularity for the initial data (u_0, u_1), one has to consider (y_0, y_1) in $H^2 \bigcap H_0^1 \times H_0^1$.

We exhibed here an example of initial data for which the $L^2(0,T)$ norm of the control is unbounded for (4.2.2) and bounded for the formulation which preserve the observability inequality. Note also that the initial data that we presented here led to unbounded numerical for all the scheme when one was considering $(\Pi^N u_0, \Pi^N u_1)$ instead of $\Lambda^{-1}(u_0, u_1)$. During our investigation, we were not able to find a initial that leads to unbounded numerical control for the scheme only preserving the observability inequality uniformly.

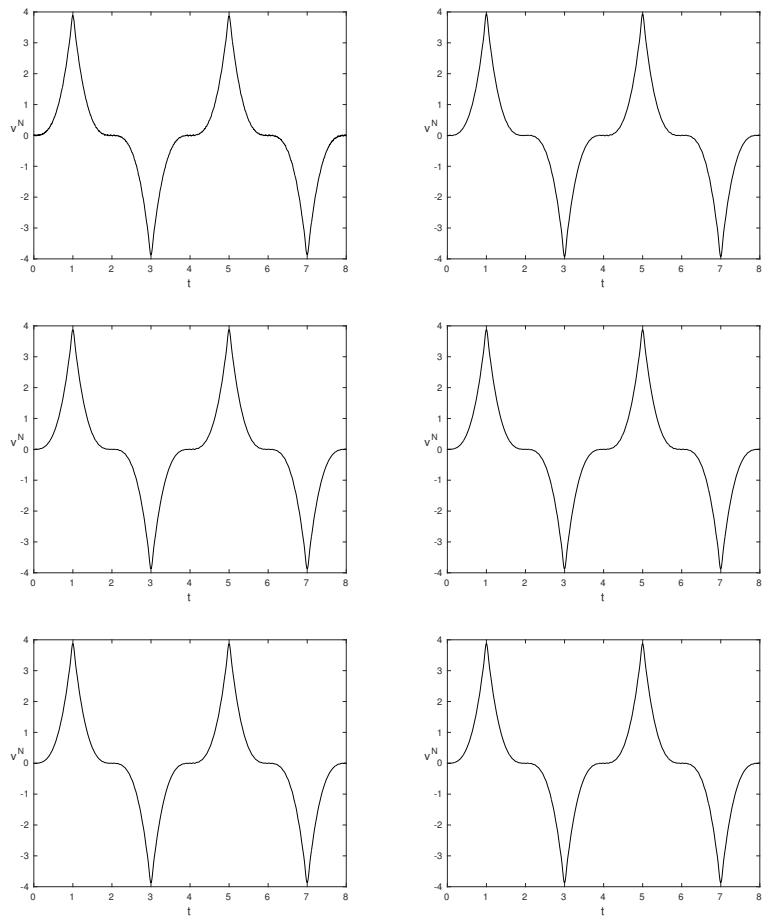

FIGURE 27 – Qualitative comparaison of the numerical control $v^N(t)$ for $N = 64$ with (from top to bottom) : (4.2.2) (left), Cesàro smoothing (right) - Exponential smoothing (p=6) (left), mixed (right) - Nitsche ($\gamma = 1$) (left), Nitsche non symmetric ($\gamma = 1$) (right)

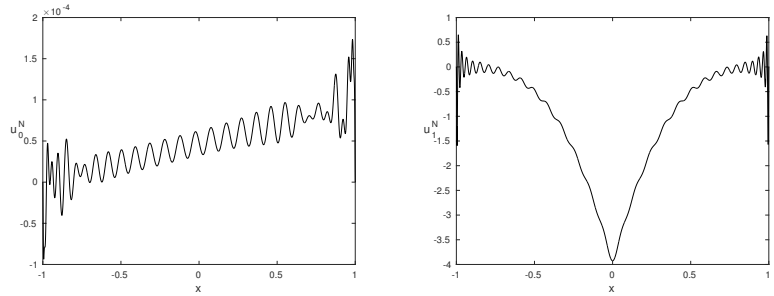

FIGURE 28 – u_0^N (left) and u_1^N (right) for $N = 64$ for (4.2.2)

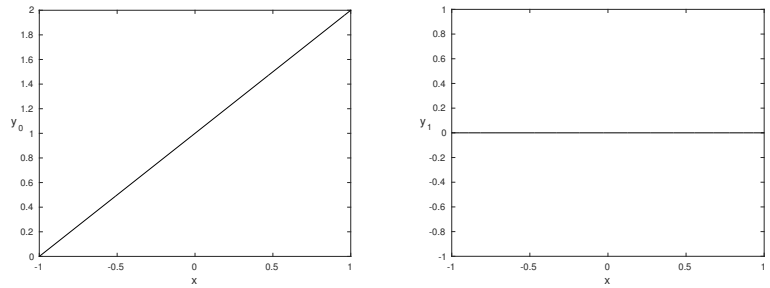

FIGURE 29 – y_0 (left), y_1 (right)

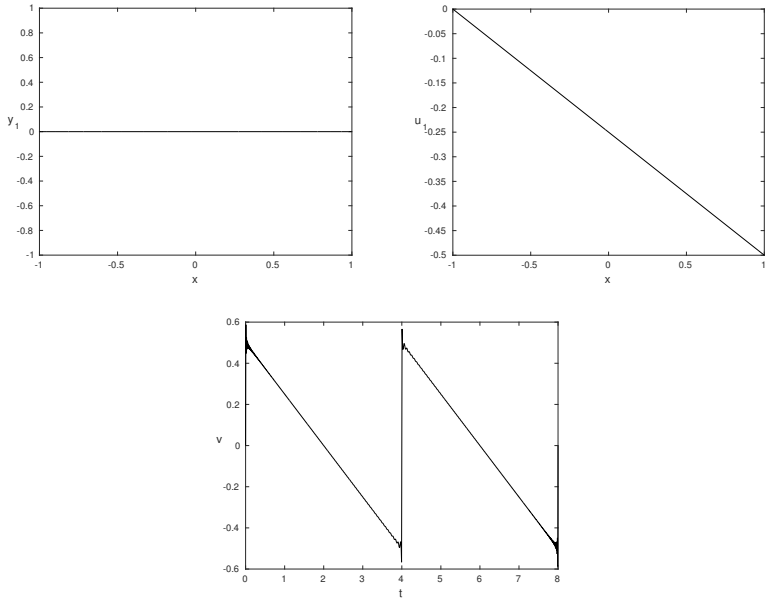

FIGURE 30 – Top : u_0 (left) and u_1 (right) - Bottom : $v(t)$

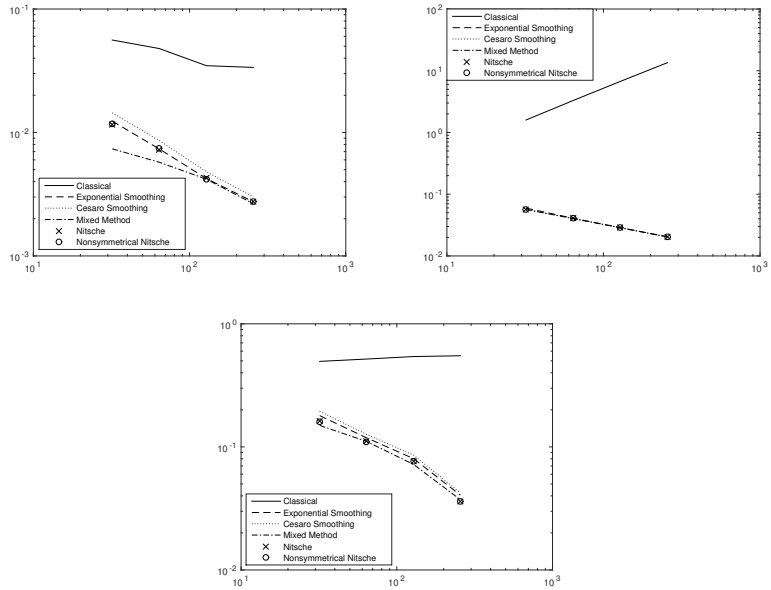

FIGURE 31 – Graphic of the convergence of initial datas u_0^N (top left), u_1^N (top right) and the convergence of the numerical control $v^N(t)$ (bottom) for : (4.2.2) , (4.2.2) filtred with exponential smoothing (p=6), (4.2.2) filtred with Cesàro smoothing, (4.4.3)-(4.4.4), (4.5.1) ($\gamma = 1$), (4.5.7) ($\gamma = 1$) and for values of N of $32, 64, 128, 256$

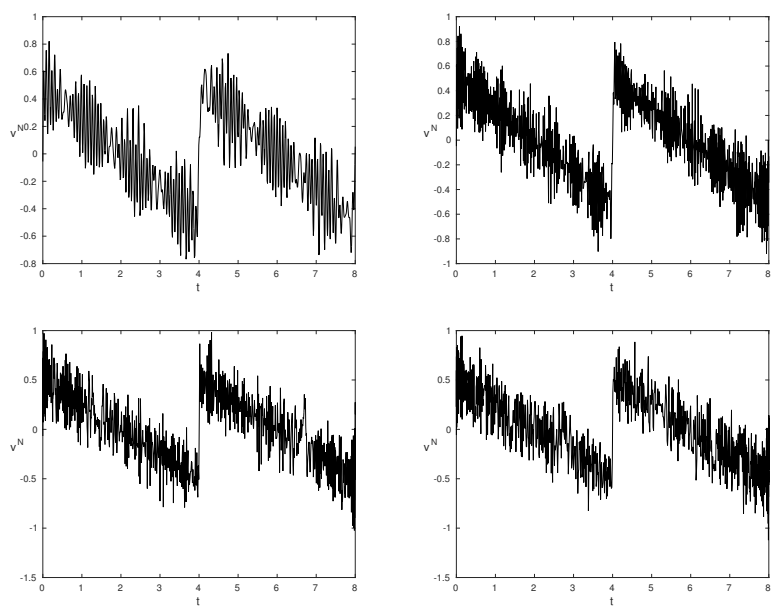

FIGURE 32 – Numeric control obtained with (4.2.2) for $N = 32$ (top left) $N = 64$ (top right) Bas $N = 128$ (bottom left) $N = 256$ (bottom right)

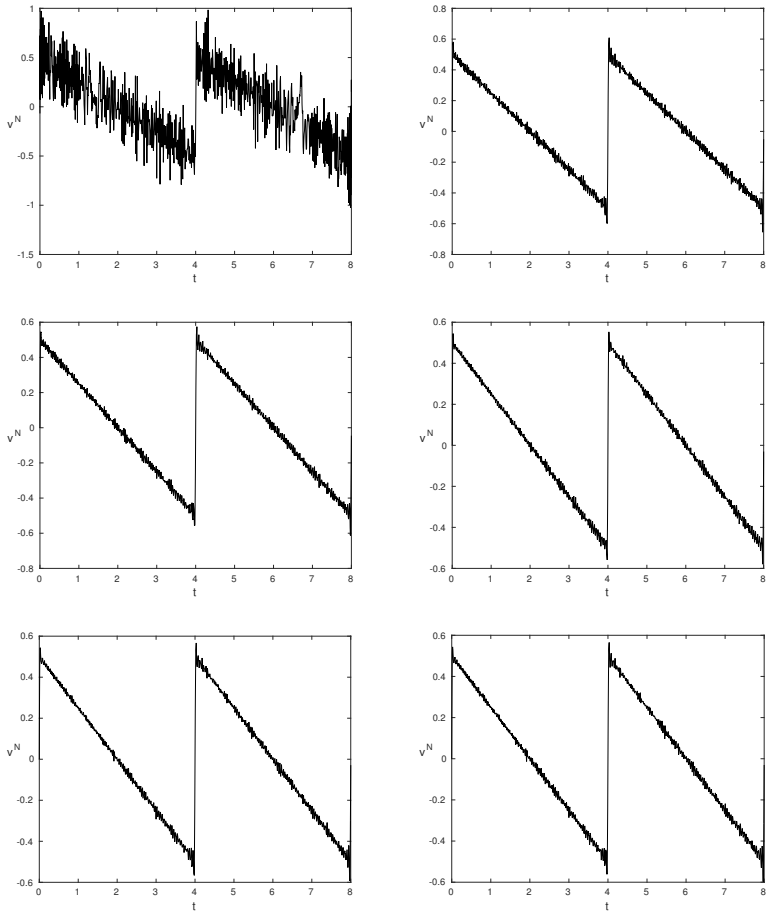

FIGURE 33 – Comparaison of numerical control for $N = 128$ with (from top to bottom) : (4.2.2) (left), Cesàro smoothing (right) - Exponential smoothing (p=6) (left), mixed (right) - Nitsche ($\gamma = 1$) (left), non symmetric Nitsche ($\gamma = 1$) (right)

Annexe A

Annexe

Contents

A.1 Simplified Saint-Venant equation example **223**
A.2 Quadratically close families **226**
A.3 Rapid stabilization of the linearized system **230**
A.4 Computation of $c_{N,T}$ and $C_{N,T}$ **232**

A.1 Simplified Saint-Venant equation example

Let us provide an explicit transformation (T, K) which stabilize exponentially rapidly the simplified Saint-Venant equation

$$\begin{cases} h_t + v_x = 0, & (x,t) \in (0,1) \times (0,T), \\ v_t + h_x = -u(t), & (x,t) \in (0,1) \times (0,T), \\ h(t,0) = v(t,1) = 0, & t \in (0,T), \\ h(0,x) = h_0(x), \quad v(0,x) = v_0(x), & x \in (0,1), \end{cases} \tag{A.1.1}$$

which is controllable for $T \geq 1$.

Let $H := h_x$ and $V := v_x$. Then, the equation on (H, V) writes

$$\begin{cases} H_t + V_x = 0, & (x,t) \in (0,1) \times (0,T), \\ V_t + H_x = 0, & (x,t) \in (0,1) \times (0,T), \\ H(t,1) = -u(t), & t \in (0,T), \\ V(t,0) = 0, & t \in (0,T), \\ H(0,x) = H_0(x), \quad V(0,x) = V_0(x), & x \in (0,1), \end{cases}$$

with $H_0 = (h_0)_x$ and $V_0 = (v_0)_x$. Consider now $R^1 := H + V$ and $R^2 := H - V$. Then

$$\begin{cases} R^1_t + R^1_x = 0, & (x,t) \in (0,1) \times (0,T), \\ R^2_t - R^2_x = 0, & (x,t) \in (0,1) \times (0,T), \\ (R^1 + R^2)(t,1) = -2u(t), & t \in (0,T), \\ (R^1 - R^2)(t,0) = 0, & t \in (0,T), \\ R^1(0,x) = R^1_0(x), \quad R^2(0,x) = R^2_0(x), & x \in (0,1), \end{cases}$$

with $R^1_0 := H_0 + V_0$ and $R^2_0 := H_0 - V_0$. Let us consider a transformation which maps (R^1, R^2) to a solution of a target stable system, that is, $\widetilde{R}^1 := e^{-\lambda x} R^1$ and $\widetilde{R}^2 := e^{\lambda x} R^2$, for $\lambda > 0$. A straightforward computation leads to

$$\begin{cases} \widetilde{R}^1_t + \widetilde{R}^1_x + \lambda \widetilde{R}^1 = 0, & (x,t) \in (0,1) \times (0,T), \\ \widetilde{R}^2_t - \widetilde{R}^2_x + \lambda \widetilde{R}^2 = 0, & (x,t) \in (0,1) \times (0,T), \\ (\widetilde{R}^1 + \widetilde{R}^2)(t,1) = -2e^{-\lambda} u(t) + 2\sinh(\lambda) R^2(t,1), & t \in (0,T) \\ (\widetilde{R}^1 - \widetilde{R}^2)(t,0) = 0, & t \in (0,T), \\ \widetilde{R}^1(0,x) = \widetilde{R}^1_0(x), \quad \widetilde{R}^2(0,x) = \widetilde{R}^2_0(x), & x \in (0,1). \end{cases} \tag{A.1.2}$$

Hence, the exponential stability of (A.1.2) is obtained if

$$-2e^{-\lambda} u(t) + 2\sinh(\lambda) R^2(t,1) = 0.$$

In terms of the original variables, it implies that

$$\begin{aligned} -2e^{-\lambda} u(t) + 2\sinh(\lambda) R^2(t,1) &= -2e^{-\lambda} u(t) + 2\sinh(\lambda)(h_x(t,1) + v_x(t,1)) \\ &= -2e^{-\lambda} u(t) + 2\sinh(\lambda)(v_x(t,1) - u(t)), \end{aligned}$$

that is, $u(t) = -\tanh(\lambda) v_x(t,1)$.

The target system of (A.1.1) is given by

$$\begin{cases} \widetilde{h}_t + \widetilde{v}_x + \lambda \widetilde{h} = 0, & (x,t) \in (0,1) \times (0,T), \\ \widetilde{v}_t + \widetilde{h}_x + \lambda \widetilde{v} = 0, & (x,t) \in (0,1) \times (0,T), \\ \widetilde{h}(t,0) = \widetilde{v}(t,1) = 0, & t \in (0,T), \\ \widetilde{h}(0,x) = \widetilde{h}_0(x), \, \widetilde{v}(0,x) = \widetilde{v}_0(x), & x \in (0,1), \end{cases} \tag{A.1.3}$$

One recovers an explicit expression of the transformation (T, K) by noting that

$$\begin{aligned} \widetilde{h}_x + \widetilde{v}_x &= e^{-\lambda x}(h_x + v_x), \\ \widetilde{h}_x - \widetilde{v}_x &= e^{\lambda x}(h_x - v_x), \end{aligned} \tag{A.1.4}$$

which boils down to

$$\begin{aligned}\widetilde{h}_x &= \cosh(\lambda x)h_x - \sinh(\lambda x)v_x\\ \widetilde{v}_x &= -\sinh(\lambda x)h_x + \cosh(\lambda x)v_x.\end{aligned}$$

Using the boundary conditions, one obtains the explicit transformation T

$$\begin{aligned}\widetilde{h}(x) &= \cosh(\lambda x)h(x) - \lambda\int_0^x \sinh(\lambda y)h(y)\,\mathrm{dy} - \sinh(\lambda x)v(x) + \lambda\int_0^x \cosh(\lambda y)v(y)\,\mathrm{dy}\\ &= \int_0^1 \left(\delta_{x=y}\cosh(\lambda y) - \lambda\mathbb{1}_{(0,x)}(y)\sinh(\lambda y)\right)h(y)\,\mathrm{dy}\\ &\quad + \int_0^1 \left(\lambda\mathbb{1}_{(0,x)}(y)\cosh(\lambda y) - \delta_{x=y}\sinh(\lambda y)\right)v(y)\,\mathrm{dy},\\ \widetilde{v}(x) &= \sinh(\lambda)h(1) - \sinh(\lambda x)h(x) - \int_x^1 \lambda\cosh(\lambda y)h(y) + \cosh(\lambda x)v(x)\\ &\quad + \lambda\int_x^1 \sinh(\lambda y)v(y)\,\mathrm{dy}\\ &= \int_0^1 \left(\delta_{y=1}\sinh(\lambda) - \delta_{x=y}\sinh(\lambda y) - \lambda\mathbb{1}_{(x,1)}(y)\cosh(\lambda y)\right)h(y)\,\mathrm{dy}\\ &\quad + \int_0^1 \left(\lambda\mathbb{1}_{(x,1)}(y)\sinh(\lambda y) + \delta_{x=y}\cosh(\lambda y)\right)v(y)\,\mathrm{dy}.\end{aligned}$$

Moreover, the explicit transformation K writes

$$u(t) = \tanh(\lambda)\int_0^1 \delta'_{y=1}v(t,y)\,\mathrm{dy}.$$

If one writes, in the same spirit as (3.2.6), the kernels equation for (A.1.1), then one obtains that the kernels of the transformations (T, K) exhibited here are the solution of this system. In fact, it is the solution, up to a multiplicative constant if one wants to satisfies the $TB = B$ condition, which is not an issue since all the equations are linear. Getting back to (A.1.4) one sees that the inverse of T can be computed explicitly performing similar computations.

Moreover, the Fourier coefficients of the kernels system associated to (A.1.1) have the same expression as (3.2.13), where the eigenvalues/eigenfunctions are replaced by those associated with (A.1.1) and the Fourier coefficients of the control operator $\mu\varphi_1$ are replaced by the one of (A.1.1), that is 1. Therefore, it is expected that a similar assumption as (3.3.1) has to be imposed to prove the existence of a Riesz basis. However, the explicit transformation exists for any $\lambda > 0$. It thus suggests that (3.3.1) is only technical.

It is also noticeable that the Fourier coefficients of the kernel α^2 are

$$\alpha_n^2 = (-1)^n (\pi n) \tanh(\lambda),$$

which is adequate with the perturbation argument used in Proposition 3.3.2 to obtain the Fourier coefficients from the $TB = B$ condition. One notice that, for (A.1.1), $\alpha^1 \equiv 0$. Whether $\alpha^1 \equiv 0$ or not in the case of the linearized Schrödinger equation cannot be verified with our analysis.

A.2 Quadratically close families

This section is devoted to the proof of Lemma 3.3.1.

Proof : Let $s = 2$ or 3. To simplify the notations, let,

$$\tilde{c}_{nk}^{ij} := c_{nk}^{ij} \beta_n^1 \langle \mu\varphi_1, \varphi_k \rangle, \quad \tilde{d}_{nk}^{ij} := d_{nk}^{ij} \beta_n^2 \langle \mu\varphi_1, \varphi_k \rangle,$$

where c_{nk}^{ij} and d_{nk}^{ij} are defined by (3.2.16) and β_n^j are defined by (3.3.2).

First step : let us prove that

$$\sum_{n\in\mathbb{N}^*} \left\| \begin{pmatrix} \varphi_n/\lambda_n^{s/2} \\ 0 \end{pmatrix} - \begin{pmatrix} g_n^{12}/\lambda_n^{(s-2)/2} \\ g_n^{22}/\lambda_n^{(s-2)/2} \end{pmatrix} \right\|_{X_{(0)}^s}^2 < +\infty.$$

First, by denoting $k = m + n$, we have

$$\begin{aligned}
&\sum_{n\in\mathbb{N}^*} \left\| \frac{\varphi_n}{\lambda_n^{s/2}} - \frac{g_n^{12}}{\lambda_n^{(s-2)/2}} \right\|_{H_{(0)}^s}^2 = \sum_{n\in\mathbb{N}^*} \sum_{k\in\mathbb{N}^*\setminus\{n\}} \left| \frac{\lambda_k^{s/2} \tilde{c}_{nk}^{12}}{\lambda_n^{(s-2)/2}} \right|^2 \\
&= \sum_{n\in\mathbb{N}^*} \left(\sum_{\substack{0<|m|<n \\ m\in\mathbb{N}^*}} + \sum_{\substack{n<m \\ m\in\mathbb{N}^*}} \right) \frac{\lambda_{n+m}^s}{\lambda_n^s} \left| \lambda_n \tilde{c}_{n,n+m}^{12} \right|^2 \\
&= \lambda^4 \sum_{n\in\mathbb{N}^*} \left(\sum_{\substack{0<|m|<n \\ m\in\mathbb{N}^*}} + \sum_{\substack{n<m \\ m\in\mathbb{N}^*}} \right) \frac{\lambda_{n+m}^s}{\lambda_n^s} \left| \frac{(\lambda^2 + 4\lambda_n^2)\lambda_{n+m}}{\delta_{n,n+m}(\lambda)\lambda_n} \right|^2 \left| \frac{\langle \mu\varphi_1, \varphi_{n+m} \rangle}{\langle \mu\varphi_1, \varphi_n \rangle} \right|^2 \\
&\leq C\lambda^4 \sum_{n\in\mathbb{N}^*} \left(\sum_{\substack{0<|m|<n \\ m\in\mathbb{N}^*}} + \sum_{\substack{m>n \\ m\in\mathbb{N}^*}} \right) \left(\frac{\lambda_{n+m}}{\lambda_n} \right)^{s-1} \left| \frac{\lambda^2 + 4\lambda_n^2}{\delta_{n,n+m}(\lambda)} \right|^2 . \qquad \text{(A.2.1)}
\end{aligned}$$

The two sums of (A.2.1) are dealt with separately.

Consider first the case where $0 < |m| < n$. We have,

$$\begin{aligned}\frac{(\lambda^2+4\lambda_n^2)}{\delta_{n,n+m}(\lambda)} &= \frac{1}{m^2n^2}\frac{\frac{\lambda^2}{n^4}+4\pi^4}{\left(\frac{\lambda^2}{n^4}+\left(\pi^2+\pi^2\left(1+\frac{m}{n}\right)\right)^2\right)\left(\frac{\lambda^2}{(mn)^2}+\left(\frac{m}{n}\pi^2+2\pi^2\right)^2\right)}\\ &\leq \frac{1}{m^2n^2}\frac{\lambda^4+4\pi^4}{\pi^8}, \end{aligned} \tag{A.2.2}$$

and

$$\lambda_{n+m} = \left(1+\frac{m}{n}\right)^2\lambda_n \leq 4\lambda_n. \tag{A.2.3}$$

Thus, the first term of the right-hand side of (A.2.1) yields

$$\sum_{n\in\mathbb{N}^*}\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-1}\left|\frac{(\lambda^2+4\lambda_n^2)}{\delta_{n,n+m}(\lambda)}\right|^2 \leq C\sum_{n\in\mathbb{N}^*}\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\frac{1}{m^4n^4} < +\infty. \tag{A.2.4}$$

Consider now the case where $m > n$. We have,

$$\begin{aligned}\frac{(\lambda^2+4\lambda_n^2)}{\delta_{n,n+m}(\lambda)} &= \frac{n^4}{m^8}\frac{\frac{\lambda^2}{n^4}+4\pi^4}{\left(\frac{\lambda^2}{m^4}+\left(\left(1+\frac{n}{m}\right)^2\pi^2+\left(\frac{n}{m}\right)^2\pi^2\right)^2\right)\left(\frac{\lambda^2}{m^4}+\left(\pi^2+2\frac{n}{m}\pi^2\right)^2\right)},\\ &\leq \frac{n^4}{m^8}\frac{\lambda^4+4\pi^4}{\pi^8} \end{aligned} \tag{A.2.5}$$

and

$$\frac{\lambda_{n+m}}{\lambda_n} = \frac{(n+m)^2}{n^2} = \frac{m^2}{n^2}\left(1+\frac{n}{m}\right)^2 \leq 4\frac{m^2}{n^2}. \tag{A.2.6}$$

Thus, the second term of the right-hand side of (A.2.1) yields,

$$\begin{aligned}\sum_{n\in\mathbb{N}^*}\sum_{\substack{n<m\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-1}\left|\frac{(\lambda^2+4\lambda_n^2)}{\delta_{n,n+m}(\lambda)}\right|^2 &\leq C\sum_{n\in\mathbb{N}^*}\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\left(\frac{n}{m}\right)^{8-2(s-1)}\frac{1}{m^8}\\ &\leq C\sum_{n\in\mathbb{N}^*}\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\frac{1}{m^4}\frac{1}{n^4} < +\infty. \end{aligned} \tag{A.2.7}$$

Then, inequalities (A.2.1), (A.2.4) and (A.2.7) imply that

$$\sum_{n\in\mathbb{N}^*}\left\|\frac{\varphi_n}{\lambda_n^{s/2}}-\frac{g_n^{12}}{\lambda_n^{(s-2)/2}}\right\|^2_{H^s_{(0)}} < +\infty. \tag{A.2.8}$$

The other sum is treated in a similar way. Indeed,

$$\begin{aligned}
&\sum_{n\in\mathbb{N}^*}\left\|\frac{g_n^{22}}{\lambda_n^{(s-2)/2}}\right\|_{H^s_{(0)}}^2 = \sum_{n\in\mathbb{N}^*}\sum_{k\in\mathbb{N}^*}\left|\frac{\lambda_k^{s/2}\tilde{c}_{nk}^{22}}{\lambda_n^{(s-2)/2}}\right|^2\\
&=\sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\right)\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^s\left|\lambda_n\tilde{c}_{n,n+m}^{22}\right|^2+\sum_{n\in\mathbb{N}^*}\left|\lambda_n\tilde{c}_{n,n}^{22}\right|^2\\
&=\frac{\lambda^2}{4}\sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\right)\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^s\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|^2\left|\frac{\langle\mu\varphi_1,\varphi_{n+m}\rangle}{\langle\mu\varphi_1,\varphi_n\rangle}\right|^2\\
&+\sum_{n\in\mathbb{N}^*}\left(\frac{\lambda}{2\lambda_n}\right)^2\\
&\le C+C\sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\right)\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|^2.
\end{aligned}\tag{A.2.9}$$

Consider first the case where $0<|m|<n$. We have

$$\begin{aligned}
\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right| &\le C\frac{n^3|m|}{n^2}\left|\frac{\lambda^2}{n^3m}-\pi^4\left(4+6\frac{m}{n}+4\left(\frac{m}{n}\right)^2+\left(\frac{m}{n}\right)^3\right)\right|\\
&\le Cn|m|\left(\lambda^2+15\pi^4\right).
\end{aligned}$$

Using (A.2.2), it comes that

$$\sum_{n\in\mathbb{N}^*}\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|^2\le C\sum_{n\in\mathbb{N}^*}\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\frac{1}{n^2m^2}.$$

This sum is clearly finite for $s=3$. For $s=2$, we notice that the previous series is the general term of a Cauchy product. Indeed,

$$\begin{aligned}
\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\frac{1}{n^2m^2} &= \sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\frac{1}{m^2}\frac{1}{(n+m)^2}\\
&\le \sum_{m=1}^{n-1}\frac{1}{m^2}\frac{1}{(n-m)^2}+\frac{1}{n^2}\sum_{m\in\mathbb{N}^*}\frac{1}{m^2}.
\end{aligned}$$

Thus,

$$\sum_{n\in\mathbb{N}^*}\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|^2<+\infty. \tag{A.2.10}$$

Consider now the case where $m>n$. We have

$$\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|=\left|\frac{m^4\left(\frac{\lambda^2}{m^4}-\left(1+\frac{n}{m}\right)^4\pi^4+\left(\frac{n}{m}\right)^4\right)}{n^2\pi^2}\right|\leq C\frac{m^4}{n^2},$$

and

$$\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\leq 1.$$

Using (A.2.5), it comes that,

$$\begin{aligned}\sum_{n\in\mathbb{N}^*}\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^{s-3}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2-\lambda_{n+m}^2+\lambda_n^2}{\lambda_n}\right|^2&\leq C\sum_{n\in\mathbb{N}^*}\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\left(\frac{m^4}{n^2}\frac{n^4}{m^8}\right)^2\\ &\leq C\sum_{n\in\mathbb{N}^*}\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\frac{1}{m^2}\frac{1}{n^2}<+\infty.\end{aligned} \tag{A.2.11}$$

Then, inequalities (A.2.9), (A.2.10) and (A.2.11) imply that

$$\sum_{n\in\mathbb{N}^*}\left\|\frac{g_n^{22}}{\lambda_n^{(s-2)/2}}\right\|_{H^s_{(0)}}^2<+\infty. \tag{A.2.12}$$

Together with (A.2.8) it ends the first step.

Second step : let us prove that

$$\sum_{n\in\mathbb{N}^*}\left\|\begin{pmatrix}0\\ \varphi_n/\lambda_n^{s/2}\end{pmatrix}-\begin{pmatrix}h_n^{12}/\lambda_n^{(s-2)/2}\\ h_n^{22}/\lambda_n^{(s-2)/2}\end{pmatrix}\right\|_{X^s_{(0)}}^2<+\infty.$$

This proof is very similar to the first step. Thus we give the expressions of the different sums but we do not detail every computation :

$$
\begin{aligned}
&\sum_{n\in\mathbb{N}^*}\left\|\frac{h_n^{12}}{\lambda_n^{(s-2)/2}}\right\|_{H^s_{(0)}}^2 = \sum_{n\in\mathbb{N}^*}\sum_{k\in\mathbb{N}^*}\left|\frac{\lambda_k^{s/2}\tilde{d}_{nk}^{12}}{\lambda_n^{(s-2)/2}}\right|^2\\
&= \sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\right)\left(\frac{\lambda_{n+m}}{\lambda_n}\right)^s\left|\lambda_n\tilde{d}_{n,n+m}^{12}\right|^2+\sum_{n\in\mathbb{N}^*}\left|\lambda_n\tilde{d}_{n,n}^{12}\right|^2\\
&= \lambda^2\sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{m>n\\ m\in\mathbb{N}^*}}\right)\frac{\lambda_{n+m}^s}{\lambda_n^s}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda_{n+m}\left(\lambda^2+\lambda_{n+m}^2-\lambda_n^2\right)}{\lambda^2+2\lambda_n^2}\right|^2\left|\frac{\langle\mu\varphi_1,\varphi_{n+m}\rangle}{\langle\mu\varphi_1,\varphi_n\rangle}\right|^2\\
&+\sum_{n\in\mathbb{N}^*}\left|\frac{\lambda^2\lambda_n}{\lambda^2+\lambda_n^2}\right|^2\\
&< +\infty,
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{n\in\mathbb{N}^*}\left\|\frac{\varphi_n}{\lambda_n^{s/2}}-\frac{h_n^{22}}{\lambda_n^{(s-2)/2}}\right\|_{H^s_{(0)}}^2 = \sum_{n\in\mathbb{N}^*}\sum_{k\in\mathbb{N}^*\setminus\{n\}}\left|\frac{\lambda_k^{s/2}\tilde{d}_{nk}^{22}}{\lambda_n^{(s-2)/2}}\right|^2\\
&= \sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{n<m\\ m\in\mathbb{N}^*}}\right)\frac{\lambda_{n+m}^s}{\lambda_n^s}\left|\lambda_n\tilde{d}_{n,n+m}^{22}\right|^2\\
&= \lambda^4\sum_{n\in\mathbb{N}^*}\left(\sum_{\substack{0<|m|<n\\ m\in\mathbb{N}^*}}+\sum_{\substack{n<m\\ m\in\mathbb{N}^*}}\right)\frac{\lambda_{n+m}^s}{\lambda_n^s}\left|\frac{\lambda^2+4\lambda_n^2}{\delta_{n,n+m}(\lambda)}\right|^2\left|\frac{\lambda^2+\lambda_{n+m}^2+\lambda_n^2}{\lambda^2+2\lambda_n^2}\right|^2\left|\frac{\langle\mu\varphi_1,\varphi_{n+m}\rangle}{\langle\mu\varphi_1,\varphi_n\rangle}\right|^2\\
&< +\infty.
\end{aligned}
$$

This completes the proof of Lemma 3.3.1.

□

A.3 Rapid stabilization of the linearized system

In this section we detail how Theorem 3.1.2 can be obtained from the results developed in this article.

Let $\Psi_0 \in \mathcal{H}_0$ and Ψ the associated solution for a control v. Then, if we define,

$\widetilde{\Psi}(t,\cdot) := \Psi(t,\cdot)e^{i\lambda_1 t}$, it comes that $\widetilde{\Psi}$ satisfies

$$\begin{cases} i\partial_t\widetilde{\Psi} = -\Delta\widetilde{\Psi} - \lambda_1\widetilde{\Psi} - v(t)\mu\varphi_1, & (x,t)\in(0,1)\times(0,T) \\ \widetilde{\Psi}(t,0) = \widetilde{\Psi}(t,1) = 0, & t\in(0,T) \\ \widetilde{\Psi}(0,\cdot) = \Psi_0, & x\in(0,1) \end{cases} \tag{A.3.1}$$

and $\|\widetilde{\Psi}(t)\|_{H^3_{(0)}} = \|\Psi(t)\|_{H^3_{(0)}}$.

Rapid stabilization of (A.3.1). Notice that system (A.3.1) is almost identical to (3.1.4) except that the spectrum of the underlying operator is shifted by λ_1. This modifies the state space. Indeed, for every $t \geq 0$,

$$0 = \Re\langle\Psi(t),\Phi_1(t)\rangle = \Re\langle\widetilde{\Psi}(t),\varphi_1\rangle.$$

Thus, $\widetilde{\Psi} \in \mathcal{H}_0$. Notice that due to Theorem 3.1.1, one gets that system (A.3.1) is exactly controllable in $\mathcal{H}_0$. The rest of the analysis is barely modified. For instance, the Riesz bases have instead the form

$$\mathcal{B} = \left\{ \begin{pmatrix} g_n^{12} \\ g_n^{22} \end{pmatrix} \ ; \ n \geq 2 \right\} \cup \left\{ \begin{pmatrix} h_n^{12} \\ h_n^{22} \end{pmatrix} \ ; \ n \geq 1 \right\},$$

where, for example,

$$g_n^{12}(x) = \sum_{k=1}^{+\infty} \frac{2\lambda(\lambda_k-\lambda_1)(\lambda_n-\lambda_1)}{(\lambda^2+(\lambda_k-\lambda_n)^2)(\lambda^2+(\lambda_k+\lambda_n-2\lambda_1)^2)}\beta_n^1\langle\mu\varphi_1,\varphi_k\rangle\varphi_k(x),$$

with β_n^1 chosen such that $\langle g_n^{12},\varphi_n\rangle = \frac{1}{\lambda_n}$.

This leads to the existence of a feedback law $\tilde{K}\begin{pmatrix}\Re(\widetilde{\Psi}) \\ \Im(\widetilde{\Psi})\end{pmatrix}$ such that

$$\left\|\begin{pmatrix}\Re(\widetilde{\Psi}(t)) \\ \Im(\widetilde{\Psi}(t))\end{pmatrix}\right\|_{X^3_{(0)}} \leq Ce^{-\lambda t}\left\|\begin{pmatrix}\Re(\Psi_0) \\ \Im(\Psi_0)\end{pmatrix}\right\|_{X^3_{(0)}}.$$

Rapid stabilization of (3.1.2). Due to the previous relation, $\widetilde{\Psi}(t) = \Psi(t)e^{i\lambda_1 t}$ it comes that if Ψ is the solution of (3.1.2) with the feedback law

$$v(t) = \tilde{K}\begin{pmatrix}\cos(\lambda_1 t)\Re(\Psi(t)) - \sin(\lambda_1 t)\Im(\Psi(t)) \\ \sin(\lambda_1 t)\Re(\Psi(t)) + \cos(\lambda_1 t)\Im(\Psi(t))\end{pmatrix},$$

then

$$\|\Psi(t)\|_{H^3_{(0)}} \leq Ce^{-\lambda t}\|\Psi_0\|_{H^3_{(0)}}.$$

A.4 Computation of $c_{N,T}$ and $C_{N,T}$

In order to explain how the values of $c_{N,T}$ and $C_{N,T}$ are computed, let's consider the first order matricial form of a semi-discretization of (4.1.1),

$$\begin{aligned}\begin{pmatrix} \mathbf{a}'_N(t) \\ \mathbf{a}''_N(t) \end{pmatrix} &= \begin{pmatrix} 0 & I_{N-1\times N-1} \\ -M_N^{-1}K_N & 0 \end{pmatrix}\begin{pmatrix} \mathbf{a}_N(t) \\ \mathbf{a}'_N(t) \end{pmatrix} \\ &=: A_N \begin{pmatrix} \mathbf{a}_N(t) \\ \mathbf{a}'_N(t) \end{pmatrix}.\end{aligned}$$

The coefficient in time of the numerical solution are given by

$$\begin{pmatrix} \mathbf{a}(t) \\ \mathbf{a}'(t) \end{pmatrix} = e^{A_N t}\begin{pmatrix} \mathbf{a}(0) \\ \mathbf{a}'(0) \end{pmatrix}.$$

Let's also write the observation term on the matricial form

$$u_x^N(1,t) = \left(C_N^t, \begin{pmatrix} \mathbf{a}_N(t) \\ \mathbf{a}'_N(t) \end{pmatrix}\right)_{\mathbb{R}^{2N}}.$$

We have then that the following expression of the observation

$$\begin{aligned}\int_0^T \left|u_x^N(1,t)\right|^2 \,\mathrm{dt} &= \left(\left(\int_0^T e^{A_N^t t}C_N^t C_N e^{At}\,\mathrm{dt}\right)\begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}, \begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}\right)_{\mathbb{R}^{2N}} \\ &=: \left(W^N(0,T)\begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}, \begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}\right)_{\mathbb{R}^{2N}}.\end{aligned}$$

On the other hand, the discrete energy writes

$$E(u^N(0)) = \frac{1}{2}\left(\begin{pmatrix} K_N & 0 \\ 0 & M_N \end{pmatrix}\begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}, \begin{pmatrix} \mathbf{a}_N(0) \\ \mathbf{a}'_N(0) \end{pmatrix}\right)_{\mathbb{R}^{2N}}.$$

The values of $c_{N,T}$ and $C_{N,T}$ are than obtained by considering respectively the minimal and maximal eigenvalue of

$$W^N(0,T) = \frac{\lambda}{2}\begin{pmatrix} K_N & 0 \\ 0 & M_N \end{pmatrix} \tag{A.4.1}$$

The following lemma ([34]) allows us to compute $W^N(0,T)$.

Lemma A.4.1 *Let*

$$\begin{aligned} F(t) &= \begin{pmatrix} F_{11}(t) & F_{12}(t) \\ 0 & F_{22}(t) \end{pmatrix} \\ &= \exp\left[\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} t\right]. \end{aligned}$$

Then,

$$F_{11}(t) = e^{A_{11}t}, F_{22}(t) = e^{A_{22}t} \ \textit{et} \ F_{12}(t) = \int_0^t e^{A_{11}(t-s)} A_{12} e^{A_{22}s} \, ds.$$

By letting $A_{11} = -A_N^t$, $A_{12} = C_N^t C_N$ et $A_{22} = A_N$, we have that

$$W(0,T) = F_{22}(T)^t F_{12}(T).$$

Table des figures

1 Représentation du système de coordonnées pour l'équation de Korteweg-de Vries 24

2 Intéraction entre deux solitons. Le soliton à la gauche de l'image a) possède une vitesse de propagation plus grande que le soliton de droite. L'image b) représente les solitons après le dépassement. Le cercle et la croix dénote la position qu'aurait eue les solitons sans cette interaction 35

3 La trajectoire de particules d'eau obtenue de mesures expérimentales du mouvement de pastilles de polystyrène à différente profondeur b dans quatre cas (a) h_0=20cm, a=7.07cm ; (b) h_0=20cm, a=8.56cm ; (c) h_0=30cm, a=5.46cm ; (d) h_0=30cm, a=7.56cm. 45

4 Comparaison entre le comportement de l'écart des racines des valeurs propres discrètes et celui du cas continu pour $N = 40$ 71

5 Comparaison entre le comportement des racines des valeurs propres discrètes et les racines des valeurs propres du cas continu pour $N = 40$ 73

6 Comparaison entre le comportement des $\Delta_{N,k}$ et la valeur du cas continu pour $N = 40$ 74

7 Le comportement $c_{N,T}$ (à gauche) and $C_{N,T}$ (à droite) pour différents filtres spectraux et $T = 8$ 75

8 Le comportement $c_{N,T}$ (à gauche) and $C_{N,T}$ (à droite) pour la formulation mixte et $T = 8$ 76

9 Le comportement des constantes $c_{N,T}$ (gauche) et $C_{N,T}$ (droite) de la formulation de Nitsche symétrique pour $N = 3, .., 250$, $\gamma = 0.8$ et $T = 8$ 78

1 An interaction between two solitons. The cross (circle) represents the position of the maximum of faster (slower) soliton. Figure a) is the state of the solution before the collison and b) is after the collision 86

1 A soliton solution of (2.1.1) represented in the Cartesian coordinates. 108

2 The orbits of water particles obtained from the experimental measurements of the polystyrene beads motions at different water levels b in the four experimental wave cases (a) h_0=20cm, a=7.07cm ; (b) h_0=20cm, a=8.56cm ; (c) h_0=30cm, a=5.46cm ; (d) h_0=30cm, a=7.56cm. 110

3 Interaction between two solitons. The cross (resp. circle) represents the position of the maximum of the faster (resp. slower) soliton if no interaction would have occured. Figure a) is the state of the 2-solitons solution before the interaction and b) is the state after the interaction. The frame is fixed at the speed of the slower soliton. 117

4 Comparison of the numerical approximation of the particle trajectories for the first order velocity field (top left) and the higher order velocity field (top right). Zoom on the end of the particle trajectories for the first order velocity field (bottom left) and the higher order velocity field (bottom right). The depth of the fluid is 30 cm and the height of the solitary wave is 5.46 cm. The dashed line represents the undisturbed water surface. 126

5 Total displacement (X) in the x variable and maximal displacement (Y) in the y variable with respect to the initial vertical position above the flat bottom of the particle b for the first order velocity field (First), the higher velocity field (Hi.) and the experimental results (Exp.). 127

6 Numerical approximation of the particle trajectories for the 2-solitons solution. The particles trajectories are in black, the initial position of the 2-solitons is in dashed black and the final position is in gray. The height of the soliton in front is 0.4cm and the soliton behind is 0.3cm. The depth of the water is 1cm. 127

1 Comparaison between the behaviour of the square roots of the discrete eigenvalues and the behaviour of the continuous ones for $N = 40$. 190

2 Behaviour of $c_{N,T}$ (left) and $C_{N,T}$ (right) for $N = 3, ..., 250$ and $T = 8$ 191

3 Comparaison between $\Delta_{N,k}$, for $N = 40$, and their continuous counterparts $\Delta_k = 1$ 192

4 Graphs of Cesàro, Lanczos, raised cosine and sharpened raised cosine filters . . . 195

5 Graphs of Vandeven (left) and exponential (right) filters for different values of p . 195

6 Values of $c_{N,T}$ (left column) and $C_{N,T}$ (right column) for differents filters applied to (4.2.2) for $T = 8$ 197

7 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) for the exponential smoothing applied to (4.2.2) for $p = 2$, $N = 3, .., 250$ and $T = 8$ 198

8 Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.4.5)-(4.4.6) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > 0 k = 1, \ldots, N-2$ (right) for $N = 40$ 199

9 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.4.3)-(4.4.4) for $N = 3, .., 250$ and $T = 8$ 200

10 Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) for the finite element mixed method for $1/h = 3, .., 250$ and $T = 4$ 201

11 Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.1) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 0.8$ 203

12 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.2)-(4.5.3) for $N = 3, .., 250$, $\gamma = 0.8$ and $T = 8$ 205

13 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.2)-(4.5.3) for $\gamma \in [0.8, 50]$, $N = 30, 60, 100, 200$ and $T = 8$ 205

14 Comparaison between the behaviour of the $\Delta_{N,k}$ when the term $\gamma N^2 u^N(1,t)$ is taken into account for $\gamma = 0.8$, $N = 40$ and $T = 8$ 206

15 Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{N,k+1}} - \sqrt{\lambda_{N,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 1$ 207

16 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.8)-(4.5.9) for $N = 3, .., 250$, $\gamma = 1$ and $T = 8$ 208

17 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.8)-(4.5.9) for $\gamma \in [1, 50]$, $N = 30, 60, 100, 200$ and $T = 8$ 208

18 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) for $N = 3, .., 250$, $\gamma = 1$ and $T = 8$ 209

19 Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 3$ 211

20 Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) associated to (4.5.11) for $\gamma \in [1.1, 5]$, $N = 30, 60, 100, 200$ and $T = 4$ 212

21 Values of $c_{N,T}$ (left) and $C_{N,T}$ (right) associated to (4.5.11) for $N = 3, .., 250$, $\gamma = 3$ and $T = 4$ 212

22 Comparaison between the behaviour of the square roots of the eigenvalues associated to (4.5.7) and the behaviour of the continuous ones (left) and behaviour of the gap $\sqrt{\lambda_{h,k+1}} - \sqrt{\lambda_{h,k}} > k = 1, \ldots, N-1$ (right) for $N = 40$ and $\gamma = 3$ 213

23 Values of $c_{h,T}$ (left) and $C_{h,T}$ (right) associated to (4.5.13) for $\gamma \in [1, 250]$, $N = 30, 60, 100, 200$ and $T = 4$ 214

24 Initial data (y_0, y_1) (4.6.1) 216

25 Données initiales (u_0, u_1) et contrôle $v(t)$ associés à (4.6.1) 216

26 Convergence des données initiales u_0^N, u_1^N et du contrôle $v^N(t)$ associées à (4.6.1) pour quelques formulations 217

27 Comparaison du contrôle numérique $v^N(t)$ associé à (4.6.1) pour quelques formulations pour N fixé à 64 218

28 Données initiales (u_0^N, u_1^N) associées à (4.6.1) pour la formulation (4.2.2) 218

29 Données initiales (y_0, y_1) (4.6.2) 219

30 Données initiales (u_0, u_1) et contrôle $v(t)$ associés à (4.6.2) 219

31 Convergence des données initiales u_0^N, u_1^N et du contrôle $v^N(t)$ associées à (4.6.2) pour quelques formulations 220

32 Contrôle numérique issu de la formulation (4.2.2) associé aux données initiales (4.6.2) 220

33 Comparaison du contrôle numérique $v^N(t)$ associé à (4.6.2) pour quelques formulations pour N fixé à 128 221

Bibliographie

[1] F. Alabau-Boussouira, R. Brockett, O. Glass, J. Le Rousseau, and E. Zuazua. *Control of Partial Differential Equations*. Springer Berlin Heidelberg, Paris, 2012. Lecture Notes in Mathematics.

[2] A. Ali and H. Kalisch. A dispersive model for undular bores. *Anal. Math. Phys.*, 2(4) :347–366, 2012.

[3] J. M. Ball, J.E. Marsden, and M. Slemrod. Controllability for distributed bilinear systems. *SIAM J. Control Optim.*, 20(4) :575–597, 1982.

[4] A. Balogh and M. Krstić. Infinite dimensional backstepping style feedback transformations for a heat equation with an arbitrary level of instability. *Eur. J. Control*, 8(3) :165–175, 2002.

[5] K. Beauchard. Local controllability of a 1-D Schrödinger equation. *J. Math. Pures Appl. (9)*, 84(7) :851–956, 2005.

[6] K. Beauchard and J.-M. Coron. Controllability of a quantum particle in a moving potential well. *J. Funct. Anal.*, 232(2) :328–389, 2006.

[7] K. Beauchard and C. Laurent. Local controllability of 1D linear and nonlinear Schrödinger equations with bilinear control. *J. Math. Pures Appl. (9)*, 94(5) :520–554, 2010.

[8] K. Beauchard and C. Laurent. Exact controllability of the 2D Schrödinger-Poisson system. Preprint, 2016.

[9] K. Beauchard and M. Mirrahimi. Practical stabilization of a quantum particle in a one-dimensional infinite square potential well. *SIAM J. Control Optim.*, 48(2) :1179–1205, 2009.

[10] K. Beauchard and M. Morancey. Local controllability of 1D Schrödinger equations with bilinear control and minimal time. *Math. Control Relat. Fields*, 4(2) :125–160, 2014.

[11] V. D. Blondel and A. Megretski. *Unsolved problems in Mathematical Systems and Control Theory*. Princeton University Press, 2004. Théorie et applications. [Theory and applications].

[12] J. L. Bona, M. Chen, and J.-C. Saut. Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. I. Derivation and linear theory. *J. Nonlinear Sci.*, 12(4) :283–318, 2002.

[13] J. L. Bona, S. M. Sun, and B.-Y. Zhang. A nonhomogeneous boundary-value problem for the Korteweg-de Vries equation posed on a finite domain. *Comm. Partial Differential Equations*, 28(7-8) :1391–1436, 2003.

[14] J. L. Bona, S. M. Sun, and B.-Y. Zhang. A non-homogeneous boundary-value problem for the Korteweg-de Vries equation posed on a finite domain. II. *J. Differential Equations*, 247(9) :2558–2596, 2009.

[15] J. L. Bona and R. Winther. The Korteweg-de Vries equation in a quarter plane, continuous dependence results. *Differential Integral Equations*, 2(2) :228–250, 1989.

[16] H. Borluk and H. Kalisch. Particle dynamics in the KdV approximation. *Wave Motion*, 49(8) :691–709, 2012.

[17] U. Boscain, M. Caponigro, T. Chambrion, and M. Sigalotti. A weak spectral condition for the controllability of the bilinear Schrödinger equation with application to the control of a rotating planar molecule. *Comm. Math. Phys.*, 311(2) :423–455, 2012.

[18] U. Boscain, M. Caponigro, and M. Sigalotti. Multi-input Schrödinger equation : controllability, tracking, and application to the quantum angular momentum. preprint, arXiv :1302.4173, 2013.

[19] U. Boscain, T. Chambrion, and M. Sigalotti. On some open questions in bilinear quantum control. preprint, arXiv :1304.7181, 2013.

[20] D. M. Bošković, A. Balogh, and M. Krstić. Backstepping in infinite dimension for a class of parabolic distributed parameter systems. *Math. Control Signals Systems*, 16(1) :44–75, 2003.

[21] T. Z. Boulmezaoud and J. M. Urquiza. On the eigenvalues of the spectral second order differentiation operator and application to the boundary observability of the wave equation. *J. Sci. Comput.*, 31(3) :307–345, 2007.

[22] N. Boussaid, M. Caponigro, and T. Chambrion. Regular propagators of bilinear quantum systems. working paper or preprint, June 2014.

[23] J. Boussinesq. Essai sur la théorie des aux courantes. *Mémoires présentés par divers savant à l'Acad. des Sci. Inst. Nat. France, XXIII*, pages 1–680, 1877.

[24] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[25] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer-Verlag, Berlin, 2006. Fundamentals in single domains.

[26] C. Castro and S. Micu. Boundary controllability of a linear semi-discrete 1-D wave equation derived from a mixed finite element method. *Numer. Math.*, 102(3) :413–462, 2006.

[27] E. Cerpa. Exact controllability of a nonlinear Korteweg-de Vries equation on a critical spatial domain. *SIAM J. Control Optim.*, 46(3) :877–899 (electronic), 2007.

[28] E. Cerpa. Control of a Korteweg-de Vries equation : a tutorial. *Math. Control Relat. Fields (Accepted)*, 2013.

[29] E. Cerpa and J.-M. Coron. Rapid stabilization for a Korteweg-de Vries equation from the left Dirichlet boundary condition. *IEEE Trans. Automat. Control*, 58(7) :1688–1695, 2013.

[30] E. Cerpa and E. Crépeau. Boundary controllability for the nonlinear Korteweg-de Vries equation on any critical domain. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(2) :457–475, 2009.

[31] T. Chambrion, P. Mason, M. Sigalotti, and U. Boscain. Controllability of the discrete-spectrum Schrödinger equation driven by an external field. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(1) :329–349, 2009.

[32] M. Chapouly. Global controllability of a nonlinear korteweg-de vries equation. *Commun. Contemp. Math.*, 11(3) :495–521, 2009.

[33] M. Chapouly. Global controllability of nonviscous and viscous Burgers-type equations. *SIAM J. Control Optim.*, 48(3) :1567–1599, 2009.

[34] Y.-Y. Chen, H.-C. Hsu, and H.-H. Hwung. Experimental study of the particle paths in solitary water waves. *Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 370(1964) :1629–1637, 2012.

[35] Y. Chitour and E. Trélat. Controllability of partial differential equations. In *Advanced topics in control systems theory*, volume 328 of *Lecture Notes in Control and Inform. Sci.*, pages 171–198. Springer, London, 2006.

[36] A. Constantin. Solitons from the Lagrangian perspective. *Discrete Contin. Dyn. Syst.*, 19(3) :469–481, 2007.

[37] A. Constantin and J. Escher. Particle trajectories in solitary water waves. *Bull. Amer. Math. Soc. (N.S.)*, 44(3) :423–431 (electronic), 2007.

[38] A. Constantin and R. S. Johnson. On the non-dimensionalisation, scaling and resulting interpretation of the classical governing equations for water waves. *J. Nonlinear Math. Phys.*, 15(suppl. 2) :58–73, 2008.

[39] J.-M. Coron. Global asymptotic stabilization for controllable systems without drift. *Math. Control Signals Systems*, 5(3) :295–312, 1992.

[40] J.-M. Coron. On the controllability of the 2-D incompressible Navier-Stokes equations with the Navier slip boundary conditions. *ESAIM Contrôle Optim. Calc. Var.*, 1 :35–75 (electronic), 1995/96.

[41] J.-M. Coron. On the controllability of the 2-D incompressible Navier-Stokes equations with the Navier slip boundary conditions. *ESAIM Contrôle Optim. Calc. Var.*, 1 :35–75 (electronic), 1996.

[42] J.-M. Coron. *Control and nonlinearity*, volume 136 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.

[43] J.-M. Coron. On the controllability of nonlinear partial differential equations. In *Proceedings of the International Congress of Mathematicians. Volume I*, pages 238–264, New Delhi, 2010. Hindustan Book Agency.

[44] J.-M. Coron. Stabilization of control systems and nonlinearities. In *Proceedings of the 8th International Congress on Industrial and Applied Mathematics*, pages 17–40. Higher Ed. Press, Beijing, 2015.

[45] J.-M. Coron and E. Crépeau. Exact boundary controllability of KdV equation with critical spatial length. *Int. J. Tomogr. Stat.*, 5(W07) :85–90, 2007.

[46] J.-M. Coron and B. d' Andréa-Novel. Stabilization of a rotating body beam without damping. *IEEE Trans. Automat. Control*, 43(5) :608–618, 1998.

[47] J.-M. Coron and A. V. Fursikov. Global exact controllability of the 2D Navier-Stokes equations on a manifold without boundary. *Russian J. Math. Phys.*, 4(4) :429–448, 1996.

[48] J.-M. Coron, L. Gagnon, and M. Morancey. Rapid stabilization of a linearized bilinear 1-d schrödinger equation. *Preprint*.

[49] J.-M. Coron, L. Hu, and G. Olive. Stabilization and controllability of first-order integro-differential hyperbolic equations. Preprint, arXiv :1511.01078, 2015.

[50] J.-M. Coron and Q. Lü. Local rapid stabilization for a Korteweg-de Vries equation with a Neumann boundary control on the right. *J. Math. Pures Appl. (9)*, 102(6) :1080–1120, 2014.

[51] J.-M. Coron and Q. Lü. Fredholm transform and local rapid stabilization for a Kuramoto–Sivashinsky equation. *J. Differential Equations*, 259(8) :3683–3729, 2015.

[52] J.-M. Coron, R. Vazquez, M. Krstić, and G. Bastin. Local exponential H^2 stabilization of a 2×2 quasilinear hyperbolic system using backstepping. *SIAM J. Control Optim.*, 51(3) :2005–2035, 2013.

[53] B. Dehman, P. Gérard, and G. Lebeau. Stabilization and control for the nonlinear Schrödinger equation on a compact surface. *Math. Z.*, 254(4) :729–749, 2006.

[54] P. G. Drazin and R. S. Johnson. *Solitons : an introduction*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1989.

[55] S. Ervedoza and E. Zuazua. The wave equation : control and numerics. In *Control of partial differential equations*, volume 2048 of *Lecture Notes in Math.*, pages 245–339. Springer, Heidelberg, 2012.

[56] S. Ervedoza and E. Zuazua. *Numerical approximation of exact controls for waves*. Springer Briefs in Mathematics. Springer, New York, 2013.

[57] C. Fabre. Uniqueness results for Stokes equations and their consequences in linear and nonlinear control problems. *ESAIM Control Optim. Calc. Var.*, 1 :267–302 (electronic), 1995/96.

[58] A. V. Faminskii. Global well-posedness of two initial-boundary-value problems for the Korteweg-de Vries equation. *Differential Integral Equations*, 20(6) :601–642, 2007.

[59] L. A. Fernández and E. Zuazua. Approximate controllability for the semilinear heat equation involving gradient terms. *J. Optim. Theory Appl.*, 101(2) :307–328, 1999.

[60] L. Gagnon. A higher order velocity field for the n-solitons solution of the Korteweg-de Vries equation. *Discrete Contin. Dyn. Syst. (accepted).*

[61] L. Gagnon. Lagrangian controllability of the 1-d Korteweg-de Vries equation. *Submitted.*

[62] L. Gagnon. Lagrangian controllability of the Korteweg-de Vries with a higher order velocity field for the N-solitons solution. *Proc. European Control Conference*, pages x–x+6, 2015.

[63] L. Gagnon and J. Urquiza. Some spectral formulations of the 1-d homogeneous dirichlet wave equation in order to recover the uniform discrete observability. *Preprint.*

[64] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura. Method for solving the korteweg-de Vries equation. *Phys. Rev. Letters*, 19 :1095–1097, 1967.

[65] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura. Korteweg-de Vries equation and generalization. VI. Methods for exact solution. *Comm. Pure Appl. Math.*, 15 :97–133, 1974.

[66] C. S. Gardner and G. K. Morikawa. The effect of temperature on the width of a small-amplitude, solitary wave in a collision-free plasma. *Comm. Pure Appl. Math.*, 18 :35–49, 1965.

[67] O. Glass. Exact boundary controllability of 3-D Euler equation. *ESAIM Control Optim. Calc. Var.*, 5 :1–44 (electronic), 2000.

[68] O. Glass. La méthode du retour en contrôlabilité et ses applications en mécanique des fluides [d'après Coron et al.]. *Astérisque*, (348) :Exp. No. 1027, vii, 1–16, 2012. Séminaire Bourbaki : Vol. 2010/2011. Exposés 1027–1042.

[69] O. Glass and S. Guerrero. Controllability of the Korteweg-de Vries equation from the right Dirichlet boundary condition. *Systems Control Lett.*, 59(7) :390–395, 2010.

[70] O. Glass and S. Guerrero. Controllability of the Korteweg-de Vries equation from the right Dirichlet boundary condition. *Systems Control Lett.*, 59(7) :390–395, 2010.

[71] O. Glass and T. Horsin. Approximate Lagrangian controllability for the 2-D Euler equation. Application to the control of the shape of vortex patches. *J. Math. Pures Appl. (9)*, 93(1) :61–90, 2010.

[72] O. Glass and T. Horsin. Prescribing the motion of a set of particles in a three-dimensional perfect fluid. *SIAM J. Control Optim.*, 50(5) :2726–2742, 2012.

[73] R. Glowinski. Ensuring well-posedness by analogy : Stokes problem and boundary control for the wave equation. *J. Comput. Phys.*, 103(2) :189–221, 1992.

[74] R. Glowinski, C. H. Li, and J.-L. Lions. A numerical approach to the exact boundary controllability of the wave equation. I. Dirichlet controls : description of the numerical methods. *Japan J. Appl. Math.*, 7(1) :1–76, 1990.

[75] I. C. Gohberg and M. G. Kreĭn. *Introduction to the theory of linear nonselfadjoint operators*. Translated from the Russian by A. Feinstein. Translations of Mathematical Monographs, Vol. 18. American Mathematical Society, Providence, R.I., 1969.

[76] M. J. Grote, A. Schneebeli, and D. Schötzau. Discontinuous Galerkin finite element method for the wave equation. *SIAM J. Numer. Anal.*, 44(6) :2408–2431 (electronic), 2006.

[77] A. Haraux. Séries lacunaires et contrôle semi-interne des vibrations d'une plaque rectangulaire. *J. Math. Pures Appl. (9)*, 68(4) :457–465 (1990), 1989.

[78] D. Henry. Steady periodic flow induced by the Korteweg-de Vries equation. *Wave Motion*, 46(6) :403–411, 2009.

[79] J. S. Hesthaven and R. M. Kirby. Filtering in Legendre spectral methods. *Math. Comp.*, 77(263) :1425–1452, 2008.

[80] R. Hirota. Exact Solution of the Korteweg-de Vries Equation for Multiple Collisions of Solitons. *Physical Review Letters*, 27(18) :1192–1194, 1971.

[81] T. Horsin. Application of the exact null controllability of the heat equation to moving sets. *C. R. Math. Acad. Sci. Paris*, 342(11) :849–852, 2006.

[82] T. Horsin. Local exact Lagrangian controllability of the Burgers viscous equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 25(2) :219–230, 2008.

[83] J. A. Infante and E. Zuazua. Boundary observability for the space semi-discretizations of the 1-D wave equation. *M2AN Math. Model. Numer. Anal.*, 33(2) :407–438, 1999.

[84] C. E. Kenig, G. Ponce, and L. Vega. Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle. *Comm. Pure Appl. Math.*, 46(4) :527–620, 1993.

[85] V. Komornik. *Exact controllability and stabilization*. RAM : Research in Applied Mathematics. Masson, Paris, 1994. The multiplier method.

[86] D. J. Korteweg and G. de Vries. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philos. Mag.*, 39 :422–443, 1895.

[87] M. Krstic, B.-Z. Guo, and A. Smyshlyaev. Boundary controllers and observers for the linearized Schrödinger equation. *SIAM J. Control Optim.*, 49(4) :1479–1497, 2011.

[88] M. Krstić, I. Kanellakopoulos, and P.V. Kokotovic. *Nonlinear and Adaptive Control Design*. Adaptative and Learning Systems for Signal Processing, Communications, and Control. John Wiley and Sons, July 1995. ISBN : 978-0-471-12732-1.

[89] M. Krstić and A. Smyshlyaev. *Boundary control of PDEs*, volume 16 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. A course on backstepping designs.

[90] I. Lasiecka, R. Triggiani, and X. Zhang. Global uniqueness, observability and stabilization of nonconservative Schrödinger equations via pointwise Carleman estimates. II. $L_2(\Omega)$-estimates. *J. Inverse Ill-Posed Probl.*, 12(2) :183–231, 2004.

[91] C. Laurent. Internal control of the Schrödinger equation. *Math. Control Relat. Fields*, 4(2) :161–186, 2014.

[92] L. León and E. Zuazua. Boundary controllability of the finite-difference space semi-discretizations of the beam equation. *ESAIM Control Optim. Calc. Var.*, 8 :827–862 (electronic), 2002. A tribute to J. L. Lions.

[93] J. E. Lidsey. Cosmology and the Korteweg-de Vries equation. *Phys. Rev. D*, 86 :123523, Dec 2012.

[94] R. Lionel. Control of the surface of a fluid by a wavemaker. *ESAIM Control Optim. Calc. Var.*, 10(3) :346–380 (electronic), 2004.

[95] J.-L. Lions. *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1*, volume 8 of *Recherches en Mathématiques Appliquées [Research in Applied Mathematics]*. Masson, Paris, 1988.

[96] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I-III*. Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[97] J.-L. Lions and E. Zuazua. Exact boundary controllability of Galerkin's approximations of Navier-Stokes equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 26(4) :605–621, 1998.

[98] W.-J. Liu and M. Krstić. Backstepping boundary control of Burgers' equation with actuator dynamics. *Systems Control Lett.*, 41(4) :291–303, 2000.

[99] E. Machtyngier and E. Zuazua. Stabilization of the Schrödinger equation. *Portugal. Math.*, 51(2) :243–256, 1994.

[100] F. Marbach. Small time global null controllability for a viscous Burgers' equation despite the presence of a boundary layer. *J. Math. Pures Appl. (9)*, 102(2) :364–384, 2014.

[101] F. Méhats, Y. Privat, and M. Sigalotti. On the controllability of quantum transport in an electronic nanostructure. *SIAM J. Appl. Math.*, 74(6) :1870–1894, 2014.

[102] M. Mirrahimi. Lyapunov control of a quantum particle in a decaying potential. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(5) :1743–1765, 2009.

[103] M. Morancey. Simultaneous local exact controllability of 1D bilinear Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 31(3) :501–529, 2014.

[104] M. Morancey and V. Nersesyan. Simultaneous global exact controllability of an arbitrary number of 1d bilinear Schrödinger equations. *J. Math. Pures Appl. (9)*, 103(1) :228–254, 2015.

[105] A. Münch. Famille de schémas implicites uniformément contrôlables pour l'équation des ondes 1-D. *C. R. Math. Acad. Sci. Paris*, 339(10) :733–738, 2004.

[106] M. Negreanu. Convergence of a semidiscrete two-grid algorithm for the controllability of the 1-d wave equation. *SIAM J. Numer. Anal.*, 46(6) :3233–3263, 2008.

[107] M. Negreanu and E. Zuazua. Convergence of a multigrid method for the controllability of a 1-d wave equation. *C. R. Math. Acad. Sci. Paris*, 338(5) :413–418, 2004.

[108] V. Nersesyan. Growth of Sobolev norms and controllability of the Schrödinger equation. *Comm. Math. Phys.*, 290(1) :371–387, 2009.

[109] V. Nersesyan. Global approximate controllability for Schrödinger equation in higher Sobolev norms and applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 27(3) :901–915, 2010.

[110] V. Nersesyan and H. Nersisyan. Global exact controllability in infinite time of Schrödinger equation. *J. Math. Pures Appl. (9)*, 97(4) :295–317, 2012.

[111] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36 :9–15, 1971. Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[112] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[113] A. C. Phillips. *Introduction to Quantum Mechanics*. John Wiley & Sons, 2013.

[114] J.-P. Puel. A regularity property for Schrödinger equations on bounded domains. *Rev. Mat. Complut.*, 26(1) :183–192, 2013.

[115] J.-P. Puel. Local exact bilinear control of the Schrödinger equation. Preprint, 2016.

[116] L. Rosier. Exact boundary controllability for the Korteweg-de Vries equation on a bounded domain. *ESAIM Control Optim. Calc. Var.*, 2 :33–55 (electronic), 1997.

[117] L. Rosier. Exact boundary controllability for the linear Korteweg-de Vries equation—a numerical study. In *Control and partial differential equations (Marseille-Luminy, 1997)*, volume 4 of *ESAIM Proc.*, pages 255–267 (electronic). Soc. Math. Appl. Indust., Paris, 1998.

[118] L. Rosier and B.-Y. Zhang. Control and stabilization of the Korteweg-de Vries equation : recent progresses. *J. Syst. Sci. Complex.*, 22(4) :647–682, 2009.

[119] J. S. Russell. Report on waves. *Report of the fourteenth meeting of the British Association for the Advancement of Science*, 1844.

[120] H. Segur. The Korteweg-de Vries equation and water waves. Solutions of the equation. I. *J. Fluid Mech.*, 59 :721–736, 1973.

[121] J. Shen. Efficient spectral-Galerkin method. I. Direct solvers of second- and fourth-order equations using Legendre polynomials. *SIAM J. Sci. Comput.*, 15(6) :1489–1505, 1994.

[122] E. D. Sontag. *Mathematical control theory*, volume 6 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1998. Deterministic finite-dimensional systems.

[123] E.D. Sontag. *Mathematical control theory*, volume 6 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1998. Deterministic finite-dimensional systems.

[124] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin ; UMI, Bologna, 2007.

[125] R. Triggiani. On the lack of exact controllability for mild solutions in Banach spaces. *J. Math. Anal. Appl.*, 50 :438–446, 1975.

[126] M. Tucsnak and G. Weiss. *Observation and control for operator semigroups*. Birkhäuser Advanced Texts : Basler Lehrbücher. Birkhäuser Verlag, Basel, 2009.

[127] G. Turinici. On the controllability of bilinear quantum systems. In *Mathematical models and methods for ab initio quantum chemistry*, volume 74 of *Lecture Notes in Chem.*, pages 75–92. Springer, Berlin, 2000.

[128] H. Vandeven. Family of spectral filters for discontinuous problems. *J. Sci. Comput.*, 6(2) :159–192, 1991.

[129] T. Warburton and J. S. Hesthaven. On the constants in *hp*-finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.*, 192(25) :2765–2773, 2003.

[130] H. Washimi and T. Taniuti. Propagation of ion-acoustic solitary waves of small amplitude. *Phys. Rev. Lett.*, 17 :996, 1966.

[131] G. B. Whitham. *Linear and nonlinear waves*. Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[132] R.M. Young. *An introduction to nonharmonic Fourier series*, volume 93 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.

[133] N. J. Zabusky. A synergetic approach to problems of nonlinear dispersive wave propagation and interaction. *Nonlinear Partial Differential Equations*, 1967.

[134] N. J. Zabusky and M. D. Kruskal. Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States. *Physical Review Letters*, 15 :240–243, August 1965.

[135] B.-Y. Zhang. Exact boundary controllability of the Korteweg-de Vries equation. *SIAM J. Control Optim.*, 37(2) :543–565, 1999.

[136] E. Zuazua. Boundary observability for the finite-difference space semi-discretizations of the 2-D wave equation in the square. *J. Math. Pures Appl. (9)*, 78(5) :523–563, 1999.

[137] E. Zuazua. Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Review*, 47(2) :197–243, 2005.

Étude de l'équation de Korteweg-de Vries en variables lagrangiennes et sa contrôlabilité, stabilisation rapide d'une équation de Schrödinger et méthodes spectrales pour le calcul du contrôle optimal

Résumé :
Cette thèse est consacrée à la contrôlabilité lagrangienne, l'étude du champ de vitesse de l'équation de Korteweg-de Vries, le problème de stabilisation rapide d'une équation aux dérivées partielles linéaires et aux méthodes numériques permettant d'obtenir la convergence des contrôles numériques vers les contrôles optimaux. Dans la première partie, on montre, à l'aide de la solution de N-solitons de l'équation de Korteweg-de Vries, qu'il est possible de faire sortir des particules du fluide à l'extérieur d'un domaine déterminé en temps arbitrairement petit. Une meilleure approximation du champ de vitesse associé à la solution de N-solitons est également présentée, permettant de retrouver en particulier une propriété typique des trajectoires des particules soumises à des ondes solitaires : les particules situées plus haut dans le fluide ont un plus grand déplacement. Dans la deuxième partie, la stabilisation rapide d'une équation de Schrödinger est obtenue grâce à une méthode inspirée du backstepping en dimension infinie. Une équation de Schrödinger stable est considérée comme l'image d'une transformation ayant comme domaine de définition les solutions de l'équation de Schrödinger à stabiliser. La stabilisation de l'équation de Schrödinger est obtenue en montrant l'inversibilité de la transformation. La nouveauté du travail présentée est l'introduction d'une condition d'unicité sur la transformation. Finalement, un filtre spectral, une formulation mixte et une formulation de Nitsche sont proposées comme technique afin d'obtenir numériquement l'observabilité uniforme de l'équation des ondes semi-discrétisée avec une méthode spectrale de Legendre-Galerkin. Une étude numérique de la convergence des contrôles numériques sans l'admissibilité uniforme de l'opérateur de contrôle est également présentée.
Mot clés : Contrôlabilité lagrangienne, équation Korteweg-de Vries, champ de vitesse, solution à N-solitons, équation de Schrödinger, stabilisation rapide, backstepping, contrôle numérique.

Abstract :
This thesis is devoted to the Lagrangian controllability and the analysis of the particle trajectories for the Korteweg-de Vries equation, to the rapid stabilization problem of the bilinear Schrödinger equation and to the convergence of the numerical controls of the wave equation. In the first part, we prove that the N-solitons solution of the Korteweg-de Vries equation allows one to move the particles outside an arbitrarily long domain in an arbitrarily small time. A higher approximation of the velocity field associated to the N-soliton is also presented, allowing to recover a typical property of solitary waves : the higher the particle is located in the fluid, the greater its displacement. These results are of a nonlinear nature since there exists no linear approximation of solitons. In the second part, inspired by the backstepping method, the rapid stabilization of a linearized Schrödinger equation is obtained. The proof consists to prove the invertibility of a transformation mapping the equation to stabilize to a stable linearized Schrödinger equation. The key ingredient of this proof is the introduction of a uniqueness condition. In the last part, a spectral filter, a mixed method and the Nitsche's method are proposed as a remedy to the lack of uniformness of the discrete observability constant for the Legendre-Galerkin semi-discretization of the wave equation. A numerical study of the convergence of the numerical controls is also presented.
Keywords : Lagrangian controllability, Korteweg-de Vries equation, velocity field, N-solitons solution, Schrödinger equation, rapid stabilization, backstepping, numerical controls.