

UNIVERSITÉ CLERMONT AUVERGNE  
ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES  
LABORATOIRE DE MATHÉMATIQUES BLAISE PASCAL – UMR 6620

---

THÈSE DE DOCTORAT – Spécialité MATHÉMATIQUES APPLIQUÉES  
présentée par ARTHUR BOTTOIS  
pour obtenir le grade de Docteur de l'Université Clermont Auvergne

---

---

QUELQUES PROBLÈMES DE CONTRÔLABILITÉ EXACTE  
POUR L'ÉQUATION DES ONDES ET APPROXIMATION

CONTROLLABILITY PROBLEMS FOR THE WAVE EQUATION  
AND NUMERICAL APPROXIMATION

---

Sous la direction de  
ARNAUD MÜNCH et NICOLAE CÎNDEA

Thèse soutenue le 26 septembre 2022 devant le jury composé de :

CARLOS CASTRO, Université Polytechnique de Madrid	Président du jury
YANNICK PRIVAT, Université de Strasbourg	Rapporteur
LIONEL ROSIER, Université du Littoral-Côte-d'Opale	Rapporteur
FARID AMMAR-KHODJA, Université Bourgogne-Franche-Comté	Examineur
LUCIE BAUDOUIN, LAAS-CNRS	Examinatrice
JULIE VALEIN, Université de Lorraine	Examinatrice
ARNAUD MÜNCH, Université Clermont-Auvergne	Directeur
NICOLAE CÎNDEA, Université Clermont-Auvergne	Directeur

# Table des matières

<b>Résumé</b>	<b>3</b>
<b>Abstract</b>	<b>5</b>
<b>1 Introduction</b>	<b>7</b>
1.1 Contrôlabilité de l'équation des ondes linéaire . . . . .	7
1.1.1 Généralités sur le contrôle des systèmes d'évolution . . . . .	7
1.1.2 Un cadre abstrait pour la contrôlabilité . . . . .	9
1.1.3 Dualité entre contrôlabilité et observabilité . . . . .	12
1.1.4 Contrôlabilité d'un système hyperbolique du second ordre . . . . .	15
1.1.5 Généralités sur l'équation des ondes linéaire . . . . .	16
1.1.6 Contrôlabilité de l'équation des ondes linéaire avec contrôle interne . . . . .	17
1.1.7 Contrôlabilité de l'équation des ondes linéaire avec contrôle au bord . . . . .	25
1.1.8 Contrôlabilité de l'équation des ondes linéaire avec contrôle ponctuel . . . . .	29
1.1.9 Quelques méthodes d'approximation du contrôle . . . . .	32
1.2 Contrôlabilité d'une équation des ondes semi-linéaire . . . . .	50
1.2.1 Contrôlabilité d'une équation des ondes semi-linéaire avec contrôle interne . . . . .	50
1.3 Contrôlabilité du système de l'élasticité linéaire . . . . .	51
1.3.1 Contrôlabilité du système de l'élasticité linéaire avec contrôle au bord . . . . .	51
1.3.2 Système de l'élasticité linéaire sous forme vitesse-contrainte . . . . .	53
1.4 Compléments à l'Article A . . . . .	54
1.4.1 Version alternative de la preuve du Théorème 2.1 . . . . .	54
1.4.2 Compléments à la Section 3 . . . . .	60
<b>2 Résultats de la thèse</b>	<b>66</b>
A Optimisation du support de contrôle non cylindrique pour la contrôlabilité à zéro de l'équation des ondes 1D . . . . .	66
B Contrôle ponctuel pour l'équation des ondes 1D – Approximation numérique et optimisation du support . . . . .	100
C Construction de contrôles exacts pour une équation des ondes semi-linéaire . . . . .	129
D Contrôlabilité de l'élasticité linéaire vu comme un système d'ordre 1 et approximation numérique par une formulation mixte espace-temps stabilisée . . . . .	161
<b>Perspectives</b>	<b>190</b>

# Résumé

Dans cette thèse, on étudie quelques problèmes de contrôlabilité exacte pour l'équation des ondes. On considère tour à tour les problèmes suivants.

- ▷ Contrôlabilité exacte pour l'équation des ondes linéaire avec contrôle interne (cf. Art. A).
- ▷ Contrôlabilité exacte pour l'équation des ondes linéaire avec contrôle ponctuel (cf. Art. B).
- ▷ Contrôlabilité exacte pour une équation des ondes semi-linéaire avec contrôle interne (cf. Art. C).
- ▷ Contrôlabilité exacte pour le système de l'élasticité linéaire avec contrôle au bord (cf. Art. D).

L'accent est particulièrement mis sur l'approximation numérique du contrôle. Pour chacun des quatre problèmes, on développe une méthode robuste pour approcher le contrôle. La méthode principalement utilisée est basée sur une approche variationnelle espace-temps dont la philosophie est décrite ci-dessous.

Au Chapitre 1, on commence par rappeler certains concepts clef pour l'étude de la contrôlabilité des systèmes d'évolution linéaires et autonomes. On rappelle notamment la dualité entre contrôlabilité du système d'état et observabilité du système adjoint, ainsi que la Méthode d'Unité de Hilbert (HUM). On évoque ensuite quelques méthodes d'approximation du contrôle pour des systèmes hyperboliques.

La méthode HUM permet d'exprimer le contrôle de norme  $L^2$  minimale à l'aide du point de minimum de la fonctionnelle conjuguée  $\mathcal{J}^*(\varphi^0)$ . Il est aujourd'hui établi que la minimisation directe de  $\mathcal{J}^*$  conduit généralement à un problème mal posé. Parmi les nombreux remèdes proposés, on opte ici pour une approche variationnelle espace-temps à l'image de celle développée dans [8]. Schématiquement, l'idée est de passer d'un problème de minimisation sans contrainte portant sur la donnée initiale  $\varphi^0(x)$  à un problème de minimisation avec contrainte portant sur la trajectoire  $\varphi(x, t)$ . En traitant la contrainte via un multiplicateur de Lagrange, on se ramène alors à une formulation mixte. Par ailleurs, la résolution de cette formulation mixte permet d'obtenir simultanément le contrôle de norme  $L^2$  minimale et la solution contrôlée associée. On signale également que dans le cas où le support de contrôle varie au cours du temps (cf. Art. A&B), le caractère espace-temps de la formulation permet d'effectuer une discrétisation adaptée au support.

Le Chapitre 2 regroupe les principaux résultats de la thèse sous la forme d'une compilation de quatre articles.

L'Article A traite de la contrôlabilité exacte pour l'équation des ondes linéaire unidimensionnelle avec contrôle interne. On autorise en outre le support de contrôle à être non cylindrique. L'article contient deux contributions principales. On démontre d'une part une inégalité d'observabilité uniforme par rapport au support de contrôle. On rappelle que l'inégalité d'observabilité est vérifiée si le support de contrôle satisfait le critère géométrique de contrôle (GCC) usuel (cf. Sect. 1.1.6 ou [24]). La notion d'uniformité permet ici de relier l'ordre de grandeur de la constante d'observabilité à un paramètre  $\varepsilon$  mesurant à quel point le critère (GCC) est "confortablement" satisfait. On signale également qu'une preuve remaniée de l'inégalité d'observabilité uniforme se trouve à la Sect. 1.4.1. La seconde contribution de l'article est l'analyse théorique et numérique d'un problème d'optimisation du support de contrôle. Pour un état initial fixé, on cherche le support de contrôle qui minimise l'énergie à dépenser pour contrôler à zéro. Ce problème s'apparente à un problème d'optimisation de forme. On montre l'existence d'un minimum global et on met en place un algorithme de descente dont on illustre le comportement par quelques simulations numériques. Les contrôles sont ici approchés via la formulation variationnelle espace-temps décrite précédemment.

L'Article B traite de la contrôlabilité exacte pour l'équation des ondes linéaire unidimensionnelle avec contrôle ponctuel. On autorise en outre le support de contrôle à être non cylindrique. L'article contient deux contributions principales. Dans un premier temps, on adapte au cas d'un contrôle ponctuel la formulation variationnelle espace-temps introduite dans [8] et utilisée dans l'Art. A. On réalise une brève étude de la convergence de cette méthode d'approximation du contrôle. La seconde contribution de l'article est l'analyse (essentiellement numérique) d'un problème d'optimisation du support de contrôle.

Il s'agit de la version analogue du problème traité dans l'Art. A. On met en place un algorithme de descente dont on illustre le comportement par quelques simulations numériques.

L'Article C traite de la contrôlabilité exacte pour une équation des ondes semi-linéaire avec contrôle interne. La contribution principale de l'article est la mise en place d'un algorithme permettant d'approcher un contrôle exact. Pour ce faire, on introduit une fonctionnelle type moindres carrés dont les zéros sont précisément les contrôles exacts du système semi-linéaire. Par une méthode de descente, on construit ainsi une suite de fonctions  $(u_k)_{k \in \mathbb{N}}$  qui converge fortement vers un contrôle exact  $u$  du système semi-linéaire. À chaque itération, la direction de descente est obtenue en calculant le contrôle d'un système linéarisé. Il est à noter que cette méthode s'apparente à un algorithme de type Newton. Ceci se traduit par le fait que la convergence de la suite devient quadratique après un nombre fini d'itérations. L'article se conclut par quelques illustrations numériques de l'algorithme moindres carrés en une et deux dimensions d'espace. On compare notamment les performances de l'algorithme avec deux méthodes de point fixe.

L'Article D traite de la contrôlabilité exacte pour le système de l'élasticité linéaire avec contrôle au bord. La contribution principale de l'article est la mise en place d'une méthode permettant d'approcher le contrôle de norme  $L^2$  minimale. Pour ce faire, on adapte les idées développées dans [26] pour l'équation des ondes. Dans un premier temps, on réécrit le système de l'élasticité sous forme vitesse-contrainte. Ceci permet de se ramener à un système d'ordre un en espace et en temps, ainsi que d'exprimer le contrôle directement à l'aide de l'état adjoint. On adapte ensuite la formulation variationnelle espace-temps également issue de [26]. Par ailleurs, afin de s'affranchir de la condition *inf-sup* qui pèse sur le type d'élément fini utilisé dans la discrétisation de la formulation mixte, on ajoute un terme de stabilisation en exploitant le fait que le multiplicateur de Lagrange n'est autre que la solution contrôlée. Ceci permet ainsi d'utiliser des éléments finis linéaires lors de la discrétisation. On conclut l'article par quelques simulations numériques en deux dimensions d'espace.

# Abstract

In this thesis, we study some exact controllability problems for the wave equation. We consider successively the following problems.

- ▷ Exact controllability for the linear wave equation with distributed control (cf. Art. A).
- ▷ Exact controllability for the linear wave equation with pointwise control (cf. Art. B).
- ▷ Exact controllability for a semi-linear wave equation with distributed control (cf. Art. C).
- ▷ Exact controllability for the linear elasticity system with boundary control (cf. Art. D).

We focus particularly on the numerical approximation of the control. For each of the four problems, we develop a robust method to approximate the control. The method mainly used is based on a space-time variational approach briefly described below.

In Chapter 1, we start by recalling the main concepts for the study of the controllability of linear autonomous evolution systems. We recall in particular the duality between controllability of the state system and observability of the adjoint system, as well as the Hilbert Uniqueness Method (HUM). We then discuss some numerical methods for the approximation of the control of hyperbolic systems.

The HUM method allows to express the control of minimal  $L^2$ -norm in terms of the minimum point of the so-called conjugate functional  $\mathcal{J}^*(\varphi^0)$ . It is by now established that the direct minimization of  $\mathcal{J}^*$  usually leads to an ill-posed problem. Among the many proposed remedies, we choose here a space-time variational approach in the spirit of the one developed in [8]. In broad outline, the idea is to transform an unconstrained minimization problem on the initial datum  $\varphi^0(x)$  into a constrained minimization problem on the trajectory  $\varphi(x, t)$ . Then, using a Lagrange multiplier for the constraint, the problem is reduced to a mixed formulation. In addition, the resolution of this mixed formulation simultaneously yields the control of minimal  $L^2$ -norm and the associated controlled solution. We also emphasize that, in the case of a time-varying control domain (cf. Art. A&B), the space-time aspect of the formulation allows to build discretizations adapted to the control domain.

In Chapter 2, we gather the main results of the thesis. These results are transcribed in a series of four articles.

Article A deals with the exact controllability for the one-dimensional linear wave equation with distributed control. Furthermore, we consider non-cylindrical control domains. The article includes two main contributions. On one hand, we prove a uniform observability inequality with respect to the control domain. We recall that the observability inequality holds if the control domain satisfies the usual Geometric Control Condition (GCC) (cf. Sect. 1.1.6 or [24]). Here, the notion of uniformity allows to link the order of magnitude of the observability constant with a parameter  $\varepsilon$  quantifying how “well” the (GCC) criterion is satisfied. We also point out that an alternative proof of the uniform observability inequality can be found in Sect. 1.4.1. On the other hand, the second contribution of the article is the theoretical and numerical analysis of the problem of optimizing the control domain. For a fixed initial state, we try to find the control domain which minimizes the energy necessary to control to zero. This problem is related to shape optimization. We show the existence of a global minimum and we set up a descent algorithm whose behaviour is illustrated by several numerical experiments. In the present case, the control is approximated using the space-time variational formulation described previously.

Article B deals with the exact controllability for the one-dimensional linear wave equation with pointwise control. Furthermore, we consider non-cylindrical control domains. The article includes two main contributions. On one hand, we adapt the space-time variational formulation introduced in [8] and used in Art. A to the case of a pointwise control. We carry out a brief study of the convergence of this approximation method of the control. On the other hand, the second contribution of the article is the (mainly numerical) analysis of the problem of optimizing the control domain. It is the analogous version of the problem addressed in Art. A. We set up a descent algorithm and illustrate its behaviour with several numerical experiments.

Article C deals with the exact controllability for a semi-linear wave equation with distributed control. The main contribution of the article is the design of an algorithm to approximate an exact control. To do so, we introduce a least-squares functional whose zeros are precisely the exact controls of the semi-linear system. By a descent method, we then build a sequence of functions  $(u_k)_{k \in \mathbb{N}}$  strongly converging to an exact control  $u$  of the semi-linear system. At each iteration, the descent direction is obtained by computing the control of a linearized system. Note that this method is related to a Newton-type algorithm. This is reflected by the fact that the convergence of the sequence becomes quadratic after a finite number of iterations. We conclude the article with some numerical illustrations of the least-squares algorithm in one and two space dimensions. In particular, we compare our algorithm with two fixed-point methods.

Article D deals with the exact controllability for the linear elasticity system with boundary control. The main contribution of the article is the design of a method to approximate the control of minimal  $L^2$ -norm. To do so, we adapt the ideas developed in [26] for the wave equation. Firstly, we rewrite the elasticity system using the velocity-constraint variables. This change of variables turns the system into a system of order one in both space and time, and allows to express the control directly in terms of the adjoint state. We then adapt the space-time variational formulation also introduced in [26]. Furthermore, in order to get rid of the *inf-sup* condition on the type of finite element involved in the discretization of the mixed formulation, we add a stabilization term by using the fact that the Lagrange multiplier turns out to be the controlled state. This allows to use linear finite elements for the discretization. We conclude the article with some numerical illustrations in two space dimensions.

# Chapitre 1

## Introduction

### 1.1 Contrôlabilité de l'équation des ondes linéaire

#### 1.1.1 Généralités sur le contrôle des systèmes d'évolution

Les systèmes d'évolution permettent de modéliser l'évolution au cours du temps d'une ou plusieurs grandeurs intervenant dans un système physique.

**Définition 1** (Système d'évolution). Un système d'évolution peut s'écrire sous la forme générale

$$\begin{cases} \frac{d}{dt}z(t) = f(z(t), t), & t > 0, \\ z(0) = z^0. \end{cases} \quad (1.1a)$$

$$z(0) = z^0. \quad (1.1b)$$

La variable  $z = (z_1, \dots, z_k)$ ,  $k \geq 1$ , représente l'état du système et regroupe l'ensemble des grandeurs physiques à l'étude. Dans l'équation (1.1a), la fonction  $f$  définit la loi d'évolution des variables d'état. L'équation (1.1b) est une condition initiale et prescrit l'état du système au temps  $t = 0$ .

Pour un système d'évolution donné, l'objectif premier est de trouver un cadre théorique tel que le problème soit bien posé dans le sens suivant.

**Définition 2** (Système bien posé). Le système d'évolution (1.1) est dit bien posé au sens de Hadamard si pour toute donnée initiale  $z^0$  admissible, le système (1.1) admet une unique solution  $z$  qui dépend continûment de la donnée initiale  $z^0$ .

*Remarque 1.* Dans la suite, on étudiera des systèmes physiques définis sur un domaine borné  $\Omega$  de l'espace. De plus, l'expression de la fonction d'évolution  $f$  pourra contenir des opérateurs différentiels agissant sur la variable d'espace  $x$ . Pour  $t \geq 0$ , l'état  $z(t)$  appartient alors à un espace de fonctions définies sur  $\Omega$ , i.e.  $z(t) : x \in \Omega \mapsto z(x, t)$ . Dans ce cas, pour que le système (1.1) soit bien posé, il faut de surcroît imposer des conditions sur  $z(t)$  au bord de  $\Omega$ .

On suppose momentanément que le système (1.1) est bien posé. En l'état, le comportement du système (1.1) est entièrement déterminé par la donnée initiale  $z^0$ . Pour pouvoir considérer des problèmes de contrôle, on se donne une manière d'agir sur le système au cours du temps via une fonction dite de contrôle ou de commande.

**Définition 3** (Système d'évolution contrôlé). Un système d'évolution contrôlé peut s'écrire sous la forme générale

$$\begin{cases} \frac{d}{dt}z(t) = f(z(t), u(t), t), & t > 0, \\ z(0) = z^0. \end{cases} \quad (1.2a)$$

$$z(0) = z^0. \quad (1.2b)$$

La variable  $u = (u_1, \dots, u_l)$ ,  $l \geq 1$ , rassemble les différentes commandes appliquées au système. Le nombre de commandes est généralement pris inférieur ou égal au nombre de composantes de l'état, i.e.  $l \leq k$ . Enfin, le plus souvent, prendre  $u \equiv 0$  revient à considérer le système non contrôlé (1.1).

*Remarque 2.* Comme indiqué plus haut, lorsque l'on étudie des systèmes physiques définis sur un domaine borné  $\Omega$  de l'espace, il est nécessaire d'imposer des conditions aux limites sur l'état  $z(t) : x \in \Omega \mapsto z(x, t)$ . Pour contrôler un tel système, on peut agir via l'équation d'évolution, mais on peut également agir via les conditions aux limites.

Une nouvelle fois, l'objectif premier est de trouver un cadre théorique tel que le problème soit bien posé dans le sens suivant.

**Définition 4** (Système contrôlé bien posé). Le système d'évolution contrôlé (1.2) est dit bien posé au sens de Hadamard si pour toute donnée initiale  $z^0$  admissible et pour toute commande  $u$  admissible, le système (1.2) admet une unique solution  $z$  qui dépend continûment de la donnée initiale  $z^0$  et de la commande  $u$ .

On suppose momentanément que le système (1.2) est bien posé. L'objectif générique de la théorie du contrôle est alors le suivant.

*Pour tout état initial admissible et pour tout état cible admissible,  
déterminer une commande admissible permettant d'amener le système contrôlé  
de l'état initial à l'état cible.*

On peut chercher à atteindre l'état cible en temps fini ou de manière asymptotique. Le choix des espaces dans lesquels on s'autorise à prendre l'état initial, l'état cible et la commande est fondamental pour garantir l'existence de contrôles. On restreint parfois l'espace des états cibles à un seul élément, qui est le plus souvent un état d'équilibre du système. Ainsi, une fois cet état atteint, on y reste indéfiniment.

*Remarque 3.* Il est à noter qu'a priori, il n'y a pas d'unicité du contrôle. Pour obtenir l'unicité, on ajoute généralement une condition sur le contrôle, par exemple en exigeant que ce dernier soit minimal pour une certaine norme.

*Remarque 4.* On ne s'intéresse ici qu'à des problèmes dits de contrôlabilité. Cela signifie que les commandes  $u(t)$  admissibles ne sont supposées dépendre que du temps. En théorie du contrôle, la problématique de contrôlabilité se démarque principalement de la problématique de stabilisation. Dans un problème de stabilisation, la commande  $u(t, z(t))$  est supposée dépendre du temps mais aussi de l'état actuel du système. Cela crée une boucle de rétroaction qui apporte plus de robustesse au contrôle vis-à-vis de petites perturbations sur les données ou la dynamique. En revanche, la mise en place d'un cadre théorique dans lequel le système d'évolution contrôlé est bien posé peut s'avérer plus complexe. Enfin, en règle générale, la contrôlabilité est réalisée en temps fini, tandis que la stabilisation est réalisée de manière asymptotique. Il existe cependant un concept de stabilisation en temps fini, qui combine contrôlabilité et stabilisation par boucle de rétroaction.

Dans la suite, on se concentre sur des problèmes de contrôlabilité en temps fini. On fixe donc un horizon de contrôle  $T > 0$ . On peut alors distinguer plusieurs notions de contrôlabilité. Par souci de concision, on note  $z(t; z^0, u)$  la solution de (1.2) avec  $z^0$  comme donnée initiale et  $u$  comme contrôle.

**Définition 5** (Notions de contrôlabilité). On note  $Z$  l'espace d'état et  $\mathcal{U}$  l'espace des contrôles. Le système d'évolution (1.2) est dit

▷ *exactement contrôlable* en temps  $T$  si

$$\forall z^0, z^T \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = z^T;$$

▷ *contrôlable aux trajectoires* en temps  $T$  si

$$\forall z^0, \tilde{z}^0 \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = z(T; \tilde{z}^0, 0);$$

▷ *contrôlable à zéro* en temps  $T$  si

$$\forall z^0 \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = 0;$$

▷ *approximativement contrôlable* en temps  $T$  si

$$\forall z^0, z^T \in Z, \quad \forall \varepsilon > 0, \quad \exists u \in \mathcal{U}, \quad \|z(T; z^0, u) - z^T\|_Z \leq \varepsilon.$$

*Remarque 5.* Il est clair que la contrôlabilité exacte implique les trois autres notions de contrôlabilité. En l'absence d'hypothèse sur la fonction d'évolution  $f$  dans (1.2), les réciproques sont généralement fausses. On verra cependant dans la suite que certaines réciproques sont vraies dans le cas où  $f$  est linéaire et autonome.



### 1.1.2 Un cadre abstrait pour la contrôlabilité

Dans cette section, on expose certains concepts clef de la contrôlabilité dans un cadre abstrait. La présentation est adaptée de [10, Sect. 2.3]. Comme précédemment, on fixe un horizon de contrôle  $T > 0$ . Pour faciliter l'exposé, on se place dans le cas d'un système d'évolution contrôlé linéaire et autonome. Cela signifie que la fonction d'évolution  $f$  présente dans le système (1.2) est supposée linéaire par rapport à l'état  $z(t)$ , linéaire par rapport au contrôle  $u(t)$  et indépendante du temps  $t$ . Ce cadre permet d'utiliser la théorie des semi-groupes (cf. [30, 34]).

On considère  $Z$  un espace de Hilbert réel muni du produit scalaire  $\langle \cdot, \cdot \rangle_Z$ . Il s'agit de l'espace d'état dans lequel évoluent les trajectoires du système. On commence par étudier le caractère bien posé du système non contrôlé suivant.

**Définition 6** (Système non contrôlé). Un système d'évolution linéaire et autonome peut s'écrire sous la forme générale

$$\begin{cases} \frac{d}{dt}z(t) = Az(t), & t \in (0, T), \\ z(0) = z^0. \end{cases} \quad (1.3a)$$

$$(1.3b)$$

L'opérateur linéaire d'évolution  $A : D(A) \subset Z \rightarrow Z$  est supposé fermé et de domaine  $D(A)$  dense dans  $Z$ . Muni de la norme  $\|\cdot\|_{D(A)}$  associée au produit scalaire

$$\langle z, \bar{z} \rangle_{D(A)} = \langle z, \bar{z} \rangle_Z + \langle Az, A\bar{z} \rangle_Z, \quad \forall z, \bar{z} \in D(A),$$

le domaine  $D(A)$  est un espace de Hilbert.

L'obtention de solutions pour le système (1.3) s'appuie sur la notion de  $C_0$ -semi-groupe.

**Définition 7** ( $C_0$ -semi-groupe). Soit  $(\mathbb{T}(t))_{t \geq 0} \subset \mathcal{L}(Z)$  une famille d'opérateurs linéaires continus. On dit que  $\mathbb{T}(t)$  est un  $C_0$ -semi-groupe sur  $Z$  si

$$\begin{aligned} \mathbb{T}(0) &= \text{Id}, \\ \mathbb{T}(t+s) &= \mathbb{T}(t)\mathbb{T}(s), \quad \forall t, s \geq 0, \\ \lim_{t \rightarrow 0^+} \mathbb{T}(t)z &= z, \quad \forall z \in Z. \end{aligned}$$

Si de plus  $\|\mathbb{T}(t)\| \leq 1$  pour  $t \geq 0$ , on dit que  $\mathbb{T}(t)$  est un  $C_0$ -semi-groupe de contractions. À tout  $C_0$ -semi-groupe  $\mathbb{T}(t)$ , on peut associer un opérateur linéaire  $A : D(A) \subset Z \rightarrow Z$  défini par

$$\begin{aligned} D(A) &:= \left\{ z \in Z; \lim_{t \rightarrow 0^+} \frac{\mathbb{T}(t)z - z}{t} \text{ existe} \right\}, \\ Az &:= \lim_{t \rightarrow 0^+} \frac{\mathbb{T}(t)z - z}{t}, \quad \forall z \in D(A). \end{aligned}$$

L'opérateur  $A$  est appelé le générateur infinitésimal du  $C_0$ -semi-groupe  $\mathbb{T}(t)$ . Il est fermé et de domaine dense. Enfin, pour tout  $z^0 \in D(A)$ , on a  $\frac{d}{dt}\mathbb{T}(t)z^0 = A\mathbb{T}(t)z^0$  pour  $t \geq 0$ .

*Remarque 6.* Dans le cas d'un système d'évolution linéaire non autonome

$$\begin{cases} \frac{d}{dt}z(t) = A(t)z(t), & t \in (0, T), \\ z(0) = z^0. \end{cases}$$

l'obtention de solutions s'appuie sur la notion de famille d'évolution.

Dans toute la suite, on suppose que l'opérateur  $A$  dans (1.3) est le générateur infinitésimal d'un  $C_0$ -semi-groupe  $\mathbb{T}(t)$  de contractions. Il existe de nombreux résultats pour caractériser de tels opérateurs. On peut par exemple citer le théorème de Hille-Yosida (cf. [30, Sect. 1.3]), ou encore le théorème de Lumer-Phillips (cf. [30, Sect. 1.4]) dont on rappelle un énoncé.

**Définition 8** (Opérateur adjoint). Soit  $A$  fermé et de domaine dense. On peut alors définir un opérateur linéaire  $A^* : D(A^*) \subset Z \rightarrow Z$  comme suit. Le domaine  $D(A^*)$  est l'ensemble des  $\zeta \in Z$  tels que l'application  $z \in D(A) \subset Z \mapsto \langle Az, \zeta \rangle_Z$  est continue. Pour tout  $\zeta \in D(A^*)$ ,  $A^*\zeta$  est l'unique élément de  $Z$  tel que

$$\langle Az, \zeta \rangle_Z = \langle z, A^*\zeta \rangle_Z, \quad \forall z \in D(A).$$

L'opérateur  $A^*$  est appelé l'adjoint de  $A$ . Il est fermé et de domaine dense.

**Définition 9** (Opérateur dissipatif). Soit  $A$  fermé et de domaine dense. L'opérateur  $A$  est dit dissipatif si

$$\langle Az, z \rangle_Z \leq 0, \quad \forall z \in D(A).$$

**Théorème 1** (Lumer-Phillips). [30, Sect. 1.4] Soit  $A$  fermé et de domaine dense. Si les opérateurs  $A$  et  $A^*$  sont tous deux dissipatifs, alors  $A$  est le générateur infinitésimal d'un  $C_0$ -semi-groupe  $\mathbb{T}(t)$  de contractions. En outre,  $A^*$  est le générateur infinitésimal du  $C_0$ -semi-groupe  $\mathbb{T}(t)^*$  de contractions.

**Proposition 1** (Solution non contrôlée). Pour tout  $z^0 \in D(A)$ , la fonction  $z(t) = \mathbb{T}(t)z^0$  est l'unique solution du système non contrôlé (1.3). Elle vérifie  $z \in C([0, T]; D(A)) \cap C^1([0, T]; Z)$  et  $\|z(t)\|_Z \leq \|z^0\|_Z$  pour  $t \in [0, T]$ . Une telle solution est qualifiée de solution forte. Pour tout  $z^0 \in Z$ , la fonction  $z(t) = \mathbb{T}(t)z^0$  vérifie  $z \in C([0, T]; Z)$  et  $\|z(t)\|_Z \leq \|z^0\|_Z$  pour  $t \in [0, T]$ . Par extension, cette fonction est qualifiée de solution généralisée.

On considère  $U$  un espace de Hilbert réel muni du produit scalaire  $\langle \cdot, \cdot \rangle_U$ . Il s'agit de l'espace dans lequel évoluent les commandes du système. On note  $\mathcal{U} := L^2(0, T; U)$ . On étudie ensuite le caractère bien posé du système contrôlé suivant.

**Définition 10** (Système contrôlé). Un système d'évolution contrôlé linéaire et autonome peut s'écrire sous la forme générale

$$\begin{cases} \frac{d}{dt}z(t) = Az(t) + Bu(t), & t \in (0, T), \\ z(0) = z^0. \end{cases} \quad (1.4a)$$

$$(1.4b)$$

L'opérateur linéaire  $B$  décrit la manière dont le contrôle  $u \in \mathcal{U}$  agit sur le système.

Il existe différentes approches pour définir les solutions du système (1.4) selon le choix de  $B$ .

**Proposition 2** (Solution contrôlée). On suppose  $B \in \mathcal{L}(U, Z)$ . Pour tous  $z^0 \in Z$  et  $u \in \mathcal{U}$ , la fonction

$$z(t) = \mathbb{T}(t)z^0 + \int_0^t \mathbb{T}(t-s)Bu(s) ds, \quad t \in [0, T], \quad (1.5)$$

vérifie  $z \in C([0, T]; Z)$  et  $\|z(t)\|_Z \leq \|z^0\|_Z + \sqrt{T}\|B\| \|u\|_{\mathcal{U}}$  pour  $t \in [0, T]$ . On qualifie cette fonction de solution généralisée du système contrôlé (1.4).

Pour aboutir à un cadre plus large que  $B \in \mathcal{L}(U, Z)$ , on utilise l'opérateur adjoint  $A^*$ . Muni de la norme  $\|\cdot\|_{D(A^*)}$  associée au produit scalaire

$$\langle \zeta, \bar{\zeta} \rangle_{D(A^*)} = \langle \zeta, \bar{\zeta} \rangle_Z + \langle A^*\zeta, A^*\bar{\zeta} \rangle_Z, \quad \forall \zeta, \bar{\zeta} \in D(A^*),$$

le domaine  $D(A^*)$  est un espace de Hilbert. On note alors  $D(A^*)'$  le dual de  $D(A^*)$  avec  $Z$  comme espace pivot, i.e.  $D(A^*) \subset Z \subset D(A^*)'$ . On suppose maintenant que  $B \in \mathcal{L}(U, D(A^*)')$ . Pour définir les solutions du système (1.4), on procède par transposition de (1.5).

**Définition 11** (Condition d'admissibilité). On dit que  $B \in \mathcal{L}(U, D(A^*)')$  est admissible pour le  $C_0$ -semi-groupe  $\mathbb{T}(t)$  si pour tout  $t \in [0, T]$ , il existe  $C_t > 0$  tel que

$$\int_0^t \|B^*\mathbb{T}(s)^*\zeta\|_U^2 ds \leq C_t \|\zeta\|_Z^2, \quad \forall \zeta \in D(A^*).$$

*Remarque 7* (Régularité cachée). La condition d'admissibilité sur  $B$  correspond à une condition dite de régularité cachée. Le plus souvent, cette condition ne peut être obtenue via des arguments classiques ou la théorie des semi-groupes.

**Proposition 3** (Solution contrôlée). On suppose  $B \in \mathcal{L}(U, D(A^*)')$  admissible pour le  $C_0$ -semi-groupe  $\mathbb{T}(t)$ . Pour tous  $z^0 \in Z$  et  $u \in \mathcal{U}$ , on dit qu'une fonction  $z \in C([0, T]; Z)$  est une solution généralisée du système contrôlé (1.4) si pour tout  $t \in [0, T]$ ,

$$\langle z(t), \zeta \rangle_Z = \langle z^0, \mathbb{T}(t)^*\zeta \rangle_Z + \int_0^t \langle u(s), B^*\mathbb{T}(t-s)^*\zeta \rangle_U ds, \quad \forall \zeta \in Z.$$

Pour tous  $z^0 \in Z$  et  $u \in \mathcal{U}$ , le système contrôlé (1.4) admet une unique solution généralisée  $z$  et il existe  $C_T > 0$  tel que

$$\|z(t)\|_Z \leq \|z^0\|_Z + C_T \|u\|_{\mathcal{U}}, \quad \forall t \in [0, T]. \quad (1.6)$$

Dans toute la suite, on suppose que l'opérateur  $B$  dans (1.4) est admissible pour le  $C_0$ -semi-groupe  $\mathbb{T}(t)$  généré par l'opérateur  $A$ . On rappelle à présent les notions de contrôlabilité introduites précédemment. On note  $z(t; z^0, u)$  la solution de (1.4) avec  $z^0$  comme donnée initiale et  $u$  comme contrôle.

**Définition 12** (Notions de contrôlabilité). Le système d'évolution (1.4) est dit

▷ *exactement contrôlable* en temps  $T$  si

$$\forall z^0, z^T \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = z^T;$$

▷ *contrôlable aux trajectoires* en temps  $T$  si

$$\forall z^0, \tilde{z}^0 \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = z(T; \tilde{z}^0, 0);$$

▷ *contrôlable à zéro* en temps  $T$  si

$$\forall z^0 \in Z, \quad \exists u \in \mathcal{U}, \quad z(T; z^0, u) = 0;$$

▷ *approximativement contrôlable* en temps  $T$  si

$$\forall z^0, z^T \in Z, \quad \forall \varepsilon > 0, \quad \exists u \in \mathcal{U}, \quad \|z(T; z^0, u) - z^T\|_Z \leq \varepsilon.$$

*Remarque 8.* Le système (1.4) étant linéaire, les notions de contrôlabilité aux trajectoires et à zéro se confondent. Clairement, la contrôlabilité aux trajectoires implique la contrôlabilité à zéro puisque  $z(T; 0, 0) = 0$ . Réciproquement, pour  $z^0, \tilde{z}^0 \in Z$ , il existe  $u \in \mathcal{U}$  tel que  $z(T; z^0 - \tilde{z}^0, u) = 0$ , soit  $z(T; z^0, u) = z(T; \tilde{z}^0, 0)$ .

Comme indiqué précédemment, la contrôlabilité exacte implique la contrôlabilité à zéro et la contrôlabilité approchée. En revanche, les réciproques sont généralement fausses. On peut néanmoins montrer que les notions de contrôlabilité exacte et à zéro sont équivalentes si le  $C_0$ -semi-groupe  $\mathbb{T}(t)$  généré par l'opérateur d'évolution  $A$  peut être prolongé en un  $C_0$ -groupe. Cette propriété se traduit par la réversibilité en temps du système libre (1.3).

**Définition 13** ( $C_0$ -groupe). Soit  $(\mathbb{T}(t))_{t \in \mathbb{R}} \subset \mathcal{L}(Z)$  une famille d'opérateurs linéaires continus. On dit que  $\mathbb{T}(t)$  est un  $C_0$ -groupe sur  $Z$  si

$$\begin{aligned} \mathbb{T}(0) &= \text{Id}, \\ \mathbb{T}(t+s) &= \mathbb{T}(t)\mathbb{T}(s), \quad \forall t, s \in \mathbb{R}, \\ \lim_{t \rightarrow 0} \mathbb{T}(t)z &= z, \quad \forall z \in Z. \end{aligned}$$

On remarque que  $\mathbb{T}(-t) = \mathbb{T}(t)^{-1}$  pour  $t \in \mathbb{R}$ . À tout  $C_0$ -groupe  $\mathbb{T}(t)$ , on peut associer un générateur infinitésimal  $A : D(A) \subset Z \rightarrow Z$  défini par

$$\begin{aligned} D(A) &:= \left\{ z \in Z; \quad \lim_{t \rightarrow 0} \frac{\mathbb{T}(t)z - z}{t} \text{ existe} \right\}, \\ Az &:= \lim_{t \rightarrow 0} \frac{\mathbb{T}(t)z - z}{t}, \quad \forall z \in D(A). \end{aligned}$$

**Proposition 4** (à zéro  $\Rightarrow$  exacte). *Soit  $A$  fermé et de domaine dense. On suppose que  $A$  est le générateur infinitésimal d'un  $C_0$ -groupe  $\mathbb{T}(t)$ . Si le système d'évolution (1.4) est contrôlable à zéro en temps  $T$ , il est alors exactement contrôlable en temps  $T$ .*

*Démonstration.* Soient  $z^0, z^T \in Z$ . Si le système (1.4) est contrôlable à zéro, il existe  $u \in \mathcal{U}$  tel que  $z(T; z^0 - \mathbb{T}(-T)z^T, u) = 0$ , soit  $z(T; z^0, u) = z(T; \mathbb{T}(-T)z^T, 0) = \mathbb{T}(T)\mathbb{T}(-T)z^T = z^T$ .  $\square$

On énonce également la caractérisation suivante, qui sera utile lors de l'étude ultérieure des systèmes hyperboliques du second ordre.

**Théorème 2** (Stone). [30, Sect. 1.10] *Soit  $A$  fermé et de domaine dense. L'opérateur  $A$  est le générateur infinitésimal d'un  $C_0$ -groupe  $\mathbb{T}(t)$  d'opérateurs unitaires (i.e.  $\mathbb{T}(t)^* = \mathbb{T}(t)^{-1}$ ) si et seulement si  $A$  est antiadjoint (i.e.  $A^* = -A$ ).*

### 1.1.3 Dualité entre contrôlabilité et observabilité

À partir de cette section, on se concentre sur la notion de contrôlabilité à zéro. Par dualité, on donne une caractérisation de la contrôlabilité à zéro par le biais d'une inégalité d'observabilité portant sur la solution d'un problème adjoint. Pour simplifier l'exposé, on suppose ici que l'opérateur d'évolution  $A$  est antiadjoint, c'est-à-dire que  $A^* = -A$ . D'après le théorème de Stone,  $A$  est donc le générateur infinitésimal d'un  $C_0$ -groupe  $\mathbb{T}(t)$  d'opérateurs unitaires, c'est-à-dire que  $\mathbb{T}(t)^* = \mathbb{T}(t)^{-1}$ . Ceci revient à supposer que le système libre (1.3) est réversible et conservatif. En effet, le fait que  $A$  soit antiadjoint implique que

$$\langle Az, z \rangle_Z = 0, \quad \forall z \in D(A).$$

Ainsi, si  $z$  est solution de (1.3), en définissant l'énergie  $\mathcal{E}(t) := \frac{1}{2} \|z(t)\|_Z^2$ , on a

$$\frac{d}{dt} \mathcal{E}(t) = \left\langle \frac{d}{dt} z(t), z(t) \right\rangle_Z = \langle Az(t), z(t) \rangle_Z = 0, \quad \forall t \in [0, T].$$

On suppose également que l'opérateur de contrôle  $B$  est admissible (voir Déf. 11).

**Définition 14** (Système contrôlé rétrograde). On considère le système d'évolution contrôlé rétrograde

$$\begin{cases} \frac{d}{dt} \xi(t) = A\xi(t) + Bu(t), & t \in (0, T), \\ \xi(T) = 0. \end{cases} \quad (1.7a)$$

$$(1.7b)$$

Pour tout  $u \in \mathcal{U}$ , le système contrôlé rétrograde (1.7) admet une unique solution  $\xi \in C([0, T]; Z)$  définie par

$$\langle \xi(t), \zeta \rangle_Z = - \int_t^T \langle u(s), B^* \mathbb{T}(t-s)^* \zeta \rangle_U ds, \quad \forall \zeta \in Z.$$

**Proposition 5** (contrôlabilité  $\Leftrightarrow$  observabilité). On note  $\xi(t; u)$  la solution de (1.7) avec  $u$  comme contrôle. On définit l'opérateur linéaire

$$\mathcal{F} : \begin{array}{ccc} \mathcal{U} & \rightarrow & Z \\ u & \mapsto & \xi(0; u) \end{array}.$$

Le système d'évolution (1.4) est contrôlable à zéro en temps  $T$  si et seulement si  $\mathcal{F}$  est surjective, i.e. si  $\mathcal{F}(\mathcal{U}) = Z$ . Or, ceci équivaut à l'existence d'une constante  $C > 0$  telle que

$$\|\zeta^0\|_Z \leq C \|\mathcal{F}^* \zeta^0\|_{\mathcal{U}}, \quad \forall \zeta^0 \in Z. \quad (1.8)$$

Cette inégalité est qualifiée d'inégalité d'observabilité.

*Démonstration.* Soit  $z^0 \in Z$ . Si l'on suppose que  $\mathcal{F}$  est surjective, il existe  $u \in \mathcal{U}$  tel que  $\xi(0; u) = z^0$ . On a alors  $\xi(t; u) = z(t; z^0, u)$ , d'où  $z(T; z^0, u) = 0$ . En effet, par définition des solutions de (1.7), on remarque que

$$\langle z^0, \zeta^0 \rangle_Z = \langle \xi(0; u), \zeta^0 \rangle_Z = - \int_0^T \langle u(s), B^* \mathbb{T}(-s)^* \zeta^0 \rangle_U ds, \quad \forall \zeta^0 \in Z.$$

Pour tout  $\zeta \in Z$ , en prenant  $\zeta^0 = \mathbb{T}(t)^* \zeta$ , on obtient

$$\langle z^0, \mathbb{T}(t)^* \zeta \rangle_Z = - \int_0^T \langle u(s), B^* \mathbb{T}(t-s)^* \zeta \rangle_U ds.$$

Finalement, on peut écrire

$$\begin{aligned} \langle \xi(t; u), \zeta \rangle_Z &= \int_0^t \langle u(s), B^* \mathbb{T}(t-s)^* \zeta \rangle_U ds - \int_0^T \langle u(s), B^* \mathbb{T}(t-s)^* \zeta \rangle_U ds \\ &= \int_0^t \langle u(s), B^* \mathbb{T}(t-s)^* \zeta \rangle_U ds + \langle z^0, \mathbb{T}(t)^* \zeta \rangle_Z = \langle z(t; z^0, u), \zeta \rangle_Z, \quad \forall \zeta \in Z. \end{aligned}$$

Réciproquement, si le système (1.4) est contrôlable à zéro, il existe  $u \in \mathcal{U}$  tel que  $z(T; z^0, u) = 0$ . Comme précédemment, on peut montrer que  $z(t; z^0, u) = \xi(t; u)$ , d'où  $\xi(0; u) = z^0$ . Pour une démonstration de l'équivalence entre la surjectivité de  $\mathcal{F}$  et l'inégalité (1.8), on renvoie à [33, Th. 4.15] ou à [5, Th. II.19].  $\square$

**Définition 15** (Système adjoint). On appelle système adjoint de (1.4) le système d'évolution

$$\begin{cases} \frac{d}{dt}\zeta(t) = A\zeta(t), & t \in (0, T), \\ \zeta(0) = \zeta^0. \end{cases} \quad (1.9a)$$

$$\zeta(0) = \zeta^0. \quad (1.9b)$$

Pour tout  $\zeta^0 \in Z$ , la solution du système (1.9) est donnée par  $\zeta(t) = \mathbb{T}(t)\zeta^0$ .

*Remarque 9.* Plus classiquement, le système adjoint de (1.4) s'écrit en fonction de  $A^*$  sous la forme du système rétrograde

$$\begin{cases} \frac{d}{dt}\zeta(t) = -A^*\zeta(t), & t \in (0, T), \\ \zeta(T) = \zeta^T. \end{cases} \quad (1.10a)$$

$$\zeta(T) = \zeta^T. \quad (1.10b)$$

Pour tout  $\zeta^T \in Z$ , la solution du système (1.10) est donnée par  $\zeta(t) = \mathbb{T}(T-t)^*\zeta^T$ . Le choix de (1.9) est ici une conséquence de l'hypothèse  $A^* = -A$ , qui confère de la réversibilité au système.

**Proposition 6** (Inégalité d'observabilité). On note  $\zeta(t; \zeta^0)$  la solution de (1.9) avec  $\zeta^0$  comme donnée initiale. L'opérateur linéaire adjoint  $\mathcal{F}^*$  est donné par

$$\mathcal{F}^* : \begin{array}{ccc} Z & \rightarrow & \mathcal{U} \\ \zeta^0 & \mapsto & v(\zeta^0) \end{array}, \quad \text{avec } v(\zeta^0) := -B^*\zeta(\cdot; \zeta^0).$$

Ainsi, le système d'évolution (1.4) est contrôlable à zéro en temps  $T$  si et seulement si il existe  $C > 0$  tel que

$$\|\zeta^0\|_Z^2 \leq C\|v(\zeta^0)\|_{\mathcal{U}}^2, \quad \forall \zeta^0 \in Z. \quad (1.11)$$

On appelle alors constante d'observabilité la borne inférieure  $C_{\text{obs}}$  des constantes  $C > 0$  telles que l'inégalité d'observabilité (1.11) est vérifiée.

*Démonstration.* Soient  $u \in \mathcal{U}$  et  $\zeta^0 \in Z$ . En utilisant la définition des solutions de (1.7), il vient

$$\langle \mathcal{F}u, \zeta^0 \rangle_Z = \langle \xi(0; u), \zeta^0 \rangle_Z = - \int_0^T \langle u(s), B^*\mathbb{T}(-s)^*\zeta^0 \rangle_{\mathcal{U}} ds = \langle u, v(\zeta^0) \rangle_{\mathcal{U}} = \langle u, \mathcal{F}^*\zeta^0 \rangle_{\mathcal{U}}.$$

□

*Remarque 10.* L'estimation (1.11) correspond à un résultat d'observabilité pour le système adjoint

$$\begin{cases} \frac{d}{dt}\zeta(t) = A\zeta(t), & t \in (0, T), \\ \zeta(0) = \zeta^0, \\ v(t) = -B^*\zeta(t), & t \in (0, T). \end{cases}$$

L'inégalité (1.11) signifie que la donnée initiale  $\zeta^0$  du système adjoint est déterminée de manière unique par l'observation  $v = -B^*\zeta$ .

**Définition 16** (Contrôle de norme minimale). On note  $z(t; z^0, u)$  la solution de (1.4) avec  $z^0$  comme donnée initiale et  $u$  comme contrôle. Si le système d'évolution (1.4) est contrôlable à zéro en temps  $T$ , à tout  $z^0 \in Z$ , on peut associer l'ensemble des contrôles à zéro

$$\mathcal{U}(z^0) := \left\{ u \in \mathcal{U}; \quad z(T; z^0, u) = 0 \right\}.$$

Avec l'estimation (1.6), on vérifie que  $\mathcal{U}(z^0)$  est un sous-espace affine fermé de  $\mathcal{U}$ . On note  $\hat{u}(z^0)$  le projeté orthogonal de 0 sur  $\mathcal{U}(z^0)$ . Il s'agit alors du contrôle de norme  $\mathcal{U}$  minimale, i.e.

$$\hat{u}(z^0) := \arg \min_{u \in \mathcal{U}(z^0)} \|u\|_{\mathcal{U}}.$$

**Définition 17** (Coût de contrôle). Si le système d'évolution (1.4) est contrôlable à zéro en temps  $T$ , on définit l'opérateur linéaire

$$\hat{u} : \begin{array}{ccc} Z & \rightarrow & \mathcal{U} \\ z^0 & \mapsto & \hat{u}(z^0) \end{array}$$

qui à une donnée initiale  $z^0$  associe le contrôle  $\hat{u}(z^0)$  de norme  $\mathcal{U}$  minimale. On vérifie que l'opérateur  $\hat{u}$  est continu. On appelle alors coût de contrôle la norme d'opérateur  $C_{\text{opt}} := \|\hat{u}\|$ . On a ainsi

$$\|\hat{u}(z^0)\|_{\mathcal{U}} \leq C_{\text{opt}}\|z^0\|_Z, \quad \forall z^0 \in Z.$$

On peut également montrer (cf. [10, Th. 2.44]) que  $C_{\text{opt}} = \sqrt{C_{\text{obs}}}$ , où  $C_{\text{obs}}$  est la constante d'observabilité introduite en (1.11).

## Méthode HUM

On présente ici une méthode exposée dans [25] et connue sous le nom de méthode HUM pour *Hilbert Uniqueness Method*. Cette méthode permet de caractériser le contrôle de norme minimale et est également à la base de plusieurs méthodes numériques d'approximation du contrôle.

**Définition 18** (Opérateur HUM). On note  $\xi(t; u)$  la solution de (1.7) avec  $u$  comme contrôle. De même, on note  $\zeta(t; \zeta^0)$  la solution de (1.9) avec  $\zeta^0$  comme donnée initiale. On définit l'opérateur linéaire  $\Lambda := \mathcal{F}\mathcal{F}^*$ , qui s'exprime

$$\Lambda : \begin{array}{l} Z \rightarrow Z \\ \zeta^0 \mapsto \xi(0; v(\zeta^0)) \end{array}, \quad \text{avec } v(\zeta^0) := -B^*\zeta(\cdot; \zeta^0).$$

Pour  $z^0 \in Z$  fixé, la méthode HUM consiste à résoudre l'équation

$$\Lambda\zeta^0 = z^0. \quad (1.12)$$

Si  $\zeta^0 \in Z$  est solution de (1.12), on a alors  $\xi(0; v(\zeta^0)) = z^0$ , d'où  $\xi(t; v(\zeta^0)) = z(t; z^0, v(\zeta^0))$  et  $z(T; z^0, v(\zeta^0)) = 0$ , ce qui signifie que  $v(\zeta^0)$  est un contrôle à zéro. En outre, on peut voir que l'équation (1.12) est équivalente au problème variationnel

$$a(\zeta^0, \bar{\zeta}^0) = \ell(\bar{\zeta}^0), \quad \forall \bar{\zeta}^0 \in Z, \quad (1.13)$$

où

$$\begin{aligned} a : Z \times Z &\rightarrow \mathbb{R}, & a(\zeta^0, \bar{\zeta}^0) &:= \langle v(\zeta^0), v(\bar{\zeta}^0) \rangle_{\mathcal{U}}, \\ \ell : Z &\rightarrow \mathbb{R}, & \ell(\bar{\zeta}^0) &:= \langle z^0, \bar{\zeta}^0 \rangle_Z. \end{aligned}$$

En effet, par définition de  $\Lambda$ , on a l'identité

$$\langle \Lambda\zeta^0, \bar{\zeta}^0 \rangle_Z = \langle \mathcal{F}\mathcal{F}^*\zeta^0, \bar{\zeta}^0 \rangle_Z = \langle \mathcal{F}^*\zeta^0, \mathcal{F}\bar{\zeta}^0 \rangle_{\mathcal{U}} = \langle v(\zeta^0), v(\bar{\zeta}^0) \rangle_{\mathcal{U}}, \quad \forall \zeta^0, \bar{\zeta}^0 \in Z.$$

**Proposition 7** (Contrôle HUM). *On suppose que l'inégalité d'observabilité (1.11) est vérifiée. Pour  $z^0 \in Z$ , le contrôle de norme  $\mathcal{U}$  minimale pour le système d'évolution (1.4) est donné par  $\hat{u}(z^0) = v(\hat{\zeta}^0)$ , où  $\hat{\zeta}^0 \in Z$  est l'unique solution du problème variationnel (1.13).*

*Démonstration.* La forme bilinéaire  $a$  et la forme linéaire  $\ell$  sont continues. La continuité de  $a$  provient de l'admissibilité de  $B$ . En outre, si l'inégalité d'observabilité (1.11) est vérifiée, la forme  $a$  est coercive puisque

$$a(\zeta^0, \zeta^0) \geq C_{\text{obs}}^{-1} \|\zeta^0\|_Z^2, \quad \forall \zeta^0 \in Z.$$

D'après le théorème de Lax-Milgram (cf. [5, Cor. V.8]), le problème (1.13) admet alors une unique solution  $\hat{\zeta}^0 \in Z$ . On constate enfin que  $\hat{u}(z^0) = v(\hat{\zeta}^0)$  est le contrôle de norme  $\mathcal{U}$  minimale. En effet, si  $u \in \mathcal{U}$  est un contrôle à zéro, par définition de  $\hat{\zeta}^0$  et de  $z(t; z^0, u)$ , il vient

$$\|v(\hat{\zeta}^0)\|_{\mathcal{U}}^2 = \langle z^0, \hat{\zeta}^0 \rangle_Z = \langle u, v(\hat{\zeta}^0) \rangle_{\mathcal{U}} \leq \|u\|_{\mathcal{U}} \|v(\hat{\zeta}^0)\|_{\mathcal{U}}.$$

□

**Corollaire 1** (Contrôle HUM). *On suppose que l'inégalité d'observabilité (1.11) est vérifiée. Pour  $z^0 \in Z$ , le contrôle de norme  $\mathcal{U}$  minimale pour le système d'évolution (1.4) est donné par  $\hat{u}(z^0) = v(\hat{\zeta}^0)$ , où  $\hat{\zeta}^0 \in Z$  est l'unique point de minimum de la fonctionnelle*

$$\mathcal{J}^*(\zeta^0) := \frac{1}{2}a(\zeta^0, \zeta^0) - \ell(\zeta^0), \quad \forall \zeta^0 \in Z.$$

*Démonstration.* Comme la forme bilinéaire  $a$  est symétrique, le problème variationnel (1.13) peut être vu comme une caractérisation des points critiques de la fonctionnelle  $\mathcal{J}^*$ . En outre, l'inégalité d'observabilité (1.11) implique la coercivité et la stricte convexité de  $\mathcal{J}^*$ . On en déduit donc que l'unique solution  $\hat{\zeta}^0 \in Z$  de (1.13) est également l'unique point de minimum de  $\mathcal{J}^*$ . □

*Remarque 11.* La fonctionnelle  $\mathcal{J}^*$  est souvent qualifiée de fonctionnelle conjuguée. Ceci provient du fait que l'expression de  $\mathcal{J}^*$  peut être obtenue en appliquant le théorème de dualité de Fenchel-Rockafellar (cf. [32, Sect. 31]) au problème de minimisation sous contrainte de la fonctionnelle coût

$$\mathcal{K}(u) := \frac{1}{2} \|u\|_{\mathcal{U}}^2, \quad \forall u \in \mathcal{U}.$$

En effet, le théorème de dualité fournit l'égalité

$$\inf_{u \in \mathcal{U}(z^0)} \mathcal{K}(u) = - \inf_{\zeta^0 \in Z} \mathcal{J}^*(\zeta^0).$$

À titre d'exemple, on renvoie à [19, Sect. 6.4] pour le cas de l'équation des ondes avec contrôle au bord, et à la Sect. 1.1.6 pour le cas de l'équation des ondes avec contrôle interne.

*Remarque 12.* Il est à noter que le problème de minimisation de  $\mathcal{J}^*$  est sans contrainte. En outre, la forme bilinéaire  $a$  est symétrique définie positive. Ainsi, du point de vue numérique, le problème de minimisation  $\inf_Z \mathcal{J}^*$  peut être résolu par l'algorithme du gradient conjugué.

### 1.1.4 Contrôlabilité d'un système hyperbolique du second ordre

Dans cette section, on s'intéresse à la contrôlabilité à zéro d'un système hyperbolique du second ordre. Le cadre choisi englobe l'équation des ondes linéaire (traitée dans les Art. A&B) et le système de l'élasticité linéaire (traité dans l'Art. D). Afin de se ramener au cadre théorique exposé précédemment, on réécrit le système d'ordre deux sous la forme du système (1.4) d'ordre un.

On considère  $H$  et  $U$  deux espaces de Hilbert réels munis des produits scalaires  $\langle \cdot, \cdot \rangle_H$  et  $\langle \cdot, \cdot \rangle_U$ . On note  $\mathcal{U} := L^2(0, T; U)$ . On considère également un opérateur  $A_0 : D(A_0) \subset H \rightarrow H$  autoadjoint et strictement positif. On peut alors construire l'espace  $D(A_0^{1/2})$  comme le complété de  $D(A_0)$  pour la norme associée au produit scalaire

$$\langle y, \bar{y} \rangle_{D(A_0^{1/2})} = \langle A_0 y, \bar{y} \rangle_H, \quad \forall y, \bar{y} \in D(A_0).$$

On introduit ensuite l'espace  $D(A_0^{1/2})'$ , dual de  $D(A_0^{1/2})$  avec  $H$  comme espace pivot. On a ainsi

$$D(A_0) \subset D(A_0^{1/2}) \subset H \subset D(A_0^{1/2})',$$

avec injections denses et continues. On se donne enfin un opérateur  $B_0 \in \mathcal{L}(U, D(A_0^{1/2})')$ .

**Définition 19** (Système hyperbolique d'ordre deux). On considère le système d'évolution contrôlé

$$\begin{cases} \frac{d^2}{dt^2} y(t) + A_0 y(t) = B_0 u(t), & t \in (0, T), \\ y(0) = y^0, \quad \frac{d}{dt} y(0) = y^1. \end{cases} \quad (1.14a)$$

$$(1.14b)$$

*Remarque 13.* L'équation des ondes linéaire correspond au cas  $A_0 = -\Delta$ , tandis que le système de l'élasticité linéaire correspond au cas  $A_0 = -\mu\Delta - (\lambda + \mu)\nabla \operatorname{div}$ , où  $\mu$  et  $\lambda$  sont les coefficients de Lamé.

Pour réécrire le système (1.14) d'ordre deux sous la forme du système (1.4) d'ordre un, on pose

$$z = (y, \frac{d}{dt} y) \quad \text{et} \quad z^0 = (y^0, y^1),$$

puis on définit l'espace d'état

$$Z := D(A_0^{1/2}) \times H,$$

ainsi que l'opérateur d'évolution  $A : D(A) \subset Z \rightarrow Z$  par

$$D(A) := D(A_0) \times D(A_0^{1/2}) \quad \text{et} \quad A := \begin{pmatrix} 0 & \operatorname{Id} \\ -A_0 & 0 \end{pmatrix}.$$

On remarque alors que l'opérateur  $A$  est antiadjoint. En effet, on constate que

$$\langle Az, \bar{z} \rangle_Z = \langle z_2, \bar{z}_1 \rangle_{D(A_0^{1/2})} - \langle A_0 z_1, \bar{z}_2 \rangle_H = \langle z_2, A_0 \bar{z}_1 \rangle_H - \langle z_1, \bar{z}_2 \rangle_{D(A_0^{1/2})} = \langle z, -A\bar{z} \rangle_Z, \quad \forall z, \bar{z} \in D(A).$$

En outre, comme l'opérateur  $A_0$  est inversible, on en déduit que l'opérateur  $A$  est également inversible, d'inverse

$$A^{-1} = \begin{pmatrix} 0 & -A_0^{-1} \\ \text{Id} & 0 \end{pmatrix}.$$

D'autre part, on renomme  $Z_1 := D(A)$  et on introduit l'espace

$$Z_{-1} := H \times D(A_0^{1/2})',$$

dual de  $Z_1$  avec  $Z$  comme espace pivot. On a donc  $Z_1 \subset Z \subset Z_{-1}$ , avec injections denses et continues. Enfin, on définit l'opérateur de contrôle  $B \in \mathcal{L}(U, Z_{-1})$  par

$$B := \begin{pmatrix} 0 \\ B_0 \end{pmatrix}.$$

On a ainsi  $B^* \in \mathcal{L}(Z_1, U)$ .

### 1.1.5 Généralités sur l'équation des ondes linéaire

Soit  $\Omega$  un domaine borné non vide de  $\mathbb{R}^d$ , avec  $d \in \{1, 2, 3\}$ . On note  $\Gamma := \partial\Omega$  le bord de  $\Omega$ . Sous sa forme la plus simple, l'équation des ondes linéaire, aussi appelée équation de d'Alembert, s'écrit

$$\begin{cases} \partial_t^2 y(x, t) - c^2 \Delta y(x, t) = 0, & x \in \Omega, t > 0, & (1.15a) \\ y(x, t) = 0, & x \in \Gamma, t > 0, & (1.15b) \\ (y, \partial_t y)(x, 0) = (y^0, y^1)(x), & x \in \Omega. & (1.15c) \end{cases}$$

Cette équation modélise la propagation d'une onde à la vitesse  $c > 0$  dans un milieu homogène, linéaire et isotrope. Dans toute la suite, on considère, sans perte de généralité, que  $c = 1$ . La quantité  $y(x, t)$  représente l'amplitude de la perturbation liée à l'onde au point  $x$  et au temps  $t$ . L'équation (1.15b) est une condition aux limites. Elle prescrit le comportement de l'onde sur le bord de  $\Omega$ . Il s'agit là d'une condition de type Dirichlet, qui sera la condition principalement utilisée par la suite. L'équation (1.15c) est une condition initiale et prescrit l'état de l'onde au temps  $t = 0$ . L'équation des ondes est un exemple archétypal d'équation aux dérivées partielles hyperbolique. De ce fait, l'étude de cette équation peut être vue comme une première étape avant l'étude éventuelle de systèmes plus complexes.

*Remarque 14.* Par exemple, si  $d = 1$  et  $\Omega = (0, 1)$ , le système (1.15) modélise le comportement d'une corde en tension entre les extrémités fixes  $x = 0$  et  $x = 1$ . La quantité  $y(x, t)$  représente alors la déformation transversale de la corde par rapport à sa position d'équilibre (voir Fig. 1.1). La fonction  $y^0(x)$  correspond ici à la déformation initiale, tandis que la fonction  $y^1(x)$  correspond à la vitesse transversale de déformation initiale.

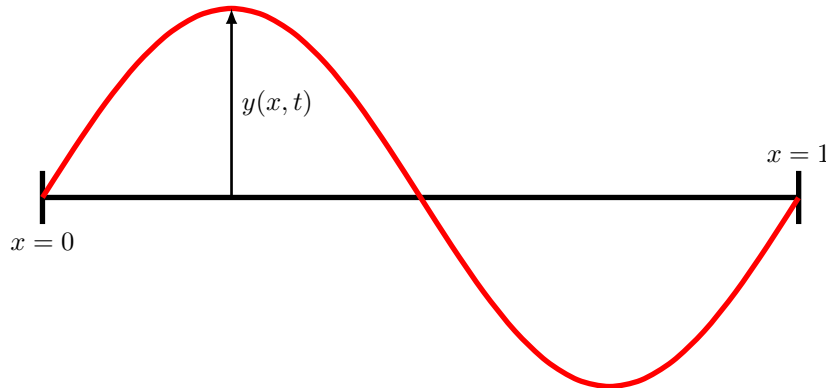


FIGURE 1.1 – Corde vibrante.

En l'état, le comportement de l'onde modélisée par le système (1.15) est entièrement déterminé par la donnée initiale  $(y^0, y^1)$ . Pour pouvoir considérer des problèmes de contrôle, on fixe un horizon  $T > 0$ . Il faut ensuite spécifier une manière d'agir sur la propagation de l'onde. Dans la suite, on va alternativement considérer les trois types de contrôle suivants.



**Contrôle interne** Soit  $\omega \subset \Omega$  un ouvert non vide. On note  $\mathbb{1}_\omega : \Omega \rightarrow \{0, 1\}$  la fonction caractéristique de  $\omega$ . On suppose ici que le contrôle  $u$  peut agir dans l'ouvert  $\omega$ , ce qui revient à considérer le système contrôlé

$$\begin{cases} \partial_t^2 y(x, t) - \Delta y(x, t) = u(x, t) \mathbb{1}_\omega(x), & x \in \Omega, t \in (0, T), & (1.16a) \\ y(x, t) = 0, & x \in \Gamma, t \in (0, T), & (1.16b) \\ (y, \partial_t y)(x, 0) = (y^0, y^1)(x), & x \in \Omega. & (1.16c) \end{cases}$$

**Contrôle au bord** Soit  $\gamma \subset \Gamma$  un ouvert non vide. On note  $\mathbb{1}_\gamma : \Gamma \rightarrow \{0, 1\}$  la fonction caractéristique de  $\gamma$ . On suppose ici que le contrôle  $u$  peut agir sur la partie  $\gamma$  du bord, ce qui revient à considérer le système contrôlé

$$\begin{cases} \partial_t^2 y(x, t) - \Delta y(x, t) = 0, & x \in \Omega, t \in (0, T), & (1.17a) \\ y(x, t) = u(x, t) \mathbb{1}_\gamma(x), & x \in \Gamma, t \in (0, T), & (1.17b) \\ (y, \partial_t y)(x, 0) = (y^0, y^1)(x), & x \in \Omega. & (1.17c) \end{cases}$$

**Contrôle ponctuel** Soit  $\gamma : [0, T] \rightarrow \Omega$  une courbe. On note  $\delta_{\gamma(t)}(x)$  la distribution de Dirac en  $x = \gamma(t)$ . On suppose ici que le contrôle  $u$  peut agir sur le point mobile  $x = \gamma(t)$ , ce qui revient à considérer le système contrôlé

$$\begin{cases} \partial_t^2 y(x, t) - \Delta y(x, t) = u(t) \delta_{\gamma(t)}(x), & x \in \Omega, t \in (0, T), & (1.18a) \\ y(x, t) = 0, & x \in \Gamma, t \in (0, T), & (1.18b) \\ (y, \partial_t y)(x, 0) = (y^0, y^1)(x), & x \in \Omega. & (1.18c) \end{cases}$$

Ici et dans toute la suite,  $\Delta := \sum_{i=1}^d \partial_{x_i}^2$  désigne l'opérateur laplacien. On pourra aussi noter  $\square := \partial_t^2 - \Delta$  l'opérateur d'alembertien.

*Remarque 15.* On traitera parfois des variantes des systèmes ci-dessus. Par exemple, dans le cas du système (1.16), on considérera que le contrôle  $u$  peut agir dans l'ouvert mobile  $(\omega_t)_{t \in (0, T)} \subset \Omega$ . Le terme  $\mathbb{1}_\omega(x)$  dans (1.16a) sera alors remplacé par  $\mathbb{1}_q(x, t)$  avec  $q := \{(x, t) \in \Omega \times (0, T); x \in \omega_t\}$ . On signale également que le cas du contrôle ponctuel ne sera traité qu'en une dimension d'espace, i.e. pour  $d = 1$  et  $\Omega = (0, 1)$ .

*Remarque 16.* On a vu que la question de la contrôlabilité à zéro se ramène à la démonstration d'une inégalité d'observabilité. Pour l'équation des ondes, on verra que cette inégalité est vérifiée si l'on suppose que la zone de contrôle satisfait une condition d'optique géométrique portant sur les lignes caractéristiques de l'équation. Dans le cas du contrôle interne, l'inégalité d'observabilité (1.22) sera ainsi vérifiée si toute ligne caractéristique, dont les éventuelles réflexions sur le bord du domaine  $\Omega$  sont régies par les lois de l'optique géométrique, rencontre l'ouvert de contrôle  $\omega$  en temps inférieur à  $T$ . Il existe des conditions analogues pour les cas du contrôle au bord et du contrôle ponctuel. L'équation des ondes étant une équation à vitesse de propagation finie, la condition géométrique ci-dessus peut s'interpréter en disant que  $T$  doit être suffisamment grand pour que chaque "composante" de l'onde ait le temps d'atteindre la zone de contrôle. Par conséquent, on ne pourra pas contrôler l'équation des ondes en temps arbitrairement petit.

### 1.1.6 Contrôlabilité de l'équation des ondes linéaire avec contrôle interne

Soit  $\Omega$  un domaine borné non vide de  $\mathbb{R}^d$ , avec  $d \in \{1, 2, 3\}$ . Le bord  $\Gamma := \partial\Omega$  est supposé de classe  $C^2$ . Alternativement, on peut supposer  $\Omega$  convexe et  $\Gamma$  lipschitzien. Pour  $T > 0$ , on pose  $Q := \Omega \times (0, T)$  et  $\Sigma := \Gamma \times (0, T)$ . On se donne également  $q \subset Q$  un ouvert non vide. On considère le système d'état

$$\begin{cases} \partial_t^2 y - \Delta y = u \mathbb{1}_q & \text{dans } Q, & (1.19a) \\ y = 0 & \text{sur } \Sigma, & (1.19b) \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega, & (1.19c) \end{cases}$$

ainsi que le système adjoint

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = 0 & \text{dans } Q, & (1.20a) \\ \varphi = 0 & \text{sur } \Sigma, & (1.20b) \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega. & (1.20c) \end{cases}$$

La fonction  $\mathbb{1}_q : Q \rightarrow \{0, 1\}$  désigne la fonction caractéristique de  $q$ . Par souci de concision, on notera parfois  $\mathbf{y}^0 = (y^0, y^1)$  et  $\boldsymbol{\varphi}^0 = (\varphi^0, \varphi^1)$ . De même, on pourra noter  $y(x, t; \mathbf{y}^0, u)$  la solution de (1.19) avec  $\mathbf{y}^0$  comme donnée initiale et  $u$  comme contrôle, ainsi que  $\varphi(x, t; \boldsymbol{\varphi}^0)$  la solution de (1.20) avec  $\boldsymbol{\varphi}^0$  comme donnée initiale.

*Remarque 17* (Régularité des solutions). Le système d'état (1.19) et le système adjoint (1.20) sont bien posés.

▷ Pour  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $u \in L^2(q)$ , il existe une unique solution  $y$  de (1.19) vérifiant

$$(y, \partial_t y) \in C([0, T]; H_0^1(\Omega) \times L^2(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|y\|_{L^\infty(H_0^1)} + \|\partial_t y\|_{L^\infty(L^2)} \leq C_{\Omega, T} \left( \|(y^0, y^1)\|_{H_0^1 \times L^2} + \|u\|_{L^2(q)} \right).$$

▷ Pour  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , il existe une unique solution  $\varphi$  de (1.20) vérifiant

$$(\varphi, \partial_t \varphi) \in C([0, T]; H_0^1(\Omega) \times L^2(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|\varphi\|_{L^\infty(H_0^1)} + \|\partial_t \varphi\|_{L^\infty(L^2)} \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}.$$

▷ Pour  $(\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , il existe une unique solution  $\varphi$  de (1.20) vérifiant

$$(\varphi, \partial_t \varphi) \in C([0, T]; L^2(\Omega) \times H^{-1}(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|\varphi\|_{L^\infty(L^2)} + \|\partial_t \varphi\|_{L^\infty(H^{-1})} \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{L^2 \times H^{-1}}.$$

**Définition 20** (Contrôlabilité à zéro). Le système (1.19) est dit contrôlable à zéro en temps  $T$  si quel que soit  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , il existe  $u \in L^2(q)$  tel que la solution correspondante  $y$  de (1.19) vérifie

$$(y, \partial_t y)(\cdot, T) = (0, 0) \quad \text{dans } \Omega.$$

**Proposition 8** (Inégalité directe). *Les assertions suivantes sont vérifiées.*

▷ Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $\varphi$  la solution de (1.20) correspondante. Il existe  $C_{\Omega, T} > 0$  tel que

$$\iint_q |\partial_t \varphi|^2 \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

▷ Soient  $(\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  et  $\varphi$  la solution de (1.20) correspondante. Il existe  $C_{\Omega, T} > 0$  tel que

$$\iint_q |\varphi|^2 \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{L^2 \times H^{-1}}^2.$$

*Démonstration.* Si  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , on a

$$\iint_q |\partial_t \varphi|^2 \leq \iint_Q |\partial_t \varphi|^2 \leq T \|\partial_t \varphi\|_{L^\infty(L^2)}^2.$$

Puis, comme  $\varphi$  est continu par rapport à la donnée initiale, il vient

$$\iint_q |\partial_t \varphi|^2 \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

Si  $(\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , on a

$$\iint_q |\varphi|^2 \leq \iint_Q |\varphi|^2 \leq T \|\varphi\|_{L^\infty(L^2)}^2.$$

Puis, comme  $\varphi$  est continu par rapport à la donnée initiale, il vient

$$\iint_q |\varphi|^2 \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{L^2 \times H^{-1}}^2.$$

□

On pose  $H := L^2(\Omega)$  et  $\mathcal{U} := L^2(q)$ . On pose également  $H_0^2(\Omega) := H^2(\Omega) \cap H_0^1(\Omega)$  et on considère les opérateurs

$$A_0 : \begin{array}{ccc} H_0^2(\Omega) & \rightarrow & L^2(\Omega) \\ y & \mapsto & -\Delta y \end{array} \quad \text{et} \quad B_0 : \begin{array}{ccc} L^2(q) & \rightarrow & L^2(0, T; L^2(\Omega)) \\ u & \mapsto & u \mathbb{1}_q \end{array}.$$

En reprenant les notations des Sect. 1.1.4 et 1.1.3, on a

$$Z_1 = H_0^2(\Omega) \times H_0^1(\Omega), \quad Z = H_0^1(\Omega) \times L^2(\Omega), \quad Z_{-1} = L^2(\Omega) \times H^{-1}(\Omega),$$

ainsi que

$$\mathcal{Z}_1 = L^2(0, T; Z_1), \quad \mathcal{Z} = L^2(0, T; Z), \quad \mathcal{Z}_{-1} = L^2(0, T; Z_{-1}).$$

On rappelle que les opérateurs  $A : Z_1 \rightarrow Z$  et  $B \in \mathcal{L}(\mathcal{U}, \mathcal{Z})$  apparaissant dans (1.4) sont donnés par

$$A = \begin{pmatrix} 0 & \text{Id} \\ -A_0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 \\ B_0 \end{pmatrix}.$$

On rappelle également que  $v(\zeta^0) = -B^* \zeta(\cdot; \zeta^0)$ , où  $\zeta(t; \zeta^0)$  est la solution de (1.9) avec  $\zeta^0$  comme donnée initiale. On constate ici que

$$B^* : \begin{array}{ccc} \mathcal{Z} & \rightarrow & \mathcal{U} \\ (\zeta_1, \zeta_2) & \mapsto & \zeta_2|_q \end{array}.$$

*Remarque 18.* On énonce certaines relations entre la solution  $\varphi$  de (1.20) et la solution  $\zeta$  de (1.9).

- ▷ Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $\zeta^0 = (\varphi^0, \varphi^1)$ , on a  $\zeta = (\varphi, \partial_t \varphi)$  et  $v(\zeta^0) = -\partial_t \varphi|_q$ .
- ▷ Soit  $(\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ . En prenant  $\zeta^0 = A^{-1}(\varphi^0, \varphi^1)$ , on a  $\zeta = (-A_0^{-1} \partial_t \varphi, \varphi)$  et  $v(\zeta^0) = -\varphi|_q$ .

De même, on peut écrire les relations de dualité suivantes.

- ▷ Soient  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $z^0 = (y^0, y^1)$  et  $\zeta^0 = (\varphi^0, \varphi^1)$ , on a

$$\langle z^0, \zeta^0 \rangle_Z = \langle y^0, \varphi^0 \rangle_{H_0^1} + \langle y^1, \varphi^1 \rangle_{L^2}.$$

- ▷ Soient  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $(\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ . En prenant  $z^0 = (y^0, y^1)$  et  $\zeta^0 = A^{-1}(\varphi^0, \varphi^1)$ , on a

$$\langle z^0, \zeta^0 \rangle_Z = -\langle y^0, \varphi^1 \rangle_{H_0^1, H^{-1}} + \langle y^1, \varphi^0 \rangle_{L^2}.$$

**Proposition 9** (Inégalité d'observabilité). *Par dualité entre contrôlabilité et observabilité, les assertions suivantes sont vérifiées.*

- ▷ Le système (1.19) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq C \|\partial_t \varphi\|_{L^2(q)}^2, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega), \quad (1.21)$$

où  $\varphi$  est la solution de (1.20) associée à  $(\varphi^0, \varphi^1)$ .

- ▷ Le système (1.19) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\|(\varphi^0, \varphi^1)\|_{L^2 \times H^{-1}}^2 \leq C \|\varphi\|_{L^2(q)}^2, \quad \forall (\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega), \quad (1.22)$$

où  $\varphi$  est la solution de (1.20) associée à  $(\varphi^0, \varphi^1)$ .

**Proposition 10** (Contrôle de norme minimale). *En appliquant la méthode HUM, les assertions suivantes sont vérifiées.*

- ▷ On suppose que l'inégalité d'observabilité (1.21) est vérifiée. Pour  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , le contrôle de norme  $L^2(q)$  minimale pour le système (1.19) est donné par  $\hat{u} = -\partial_t \hat{\varphi}|_q$ , avec  $\hat{\varphi}$  la solution de (1.20) associée au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) := \frac{1}{2} \|\partial_t \varphi\|_{L^2(q)}^2 - \langle y^0, \varphi^0 \rangle_{H_0^1} - \langle y^1, \varphi^1 \rangle_{L^2}, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.20) associée à  $(\varphi^0, \varphi^1)$ .

- ▷ On suppose que l'inégalité d'observabilité (1.22) est vérifiée. Pour  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , le contrôle de norme  $L^2(q)$  minimale pour le système (1.19) est donné par  $\hat{u} = -\hat{\varphi}|_q$ , avec  $\hat{\varphi}$  la solution de (1.20) associée au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) := \frac{1}{2} \|\varphi\|_{L^2(q)}^2 + \langle y^0, \varphi^1 \rangle_{H_0^1, H^{-1}} - \langle y^1, \varphi^0 \rangle_{L^2}, \quad \forall (\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega),$$

où  $\varphi$  est la solution de (1.20) associée à  $(\varphi^0, \varphi^1)$ .

### Cas d'un contrôle à poids

On s'intéresse à présent au contrôle de norme  $\mathcal{U}$  minimale dans le cas où  $\mathcal{U}$  est un espace à poids. Plus précisément, pour  $\chi : Q \rightarrow [0, 1]$  tel que  $\text{Supp } \chi = \bar{q}$ , on définit les espaces  $L^2$  à poids

$$\mathcal{U} := \left\{ u : q \rightarrow \mathbb{R}; \quad \iint_q u^2 \chi^{-1} < +\infty \right\} \quad \text{et} \quad \mathcal{U}' := \left\{ v : q \rightarrow \mathbb{R}; \quad \iint_q v^2 \chi < +\infty \right\},$$

que l'on munit des produits scalaires

$$\langle u, \bar{u} \rangle_{\mathcal{U}} := \iint_q u \bar{u} \chi^{-1}, \quad \forall u, \bar{u} \in \mathcal{U} \quad \text{et} \quad \langle v, \bar{v} \rangle_{\mathcal{U}'} := \iint_q v \bar{v} \chi, \quad \forall v, \bar{v} \in \mathcal{U}'.$$

Il est à noter que si  $\chi = \mathbb{1}_q$ , on a simplement  $\mathcal{U} = \mathcal{U}' = L^2(q)$ . On introduit les espaces  $V := H_0^1(\Omega) \times L^2(\Omega)$  et  $W := L^2(\Omega) \times H^{-1}(\Omega)$ . Le produit de dualité entre  $V$  et  $W$  s'écrit

$$\langle \mathbf{y}^0, \boldsymbol{\varphi}^0 \rangle_{V,W} := -\langle \mathbf{y}^0, \boldsymbol{\varphi}^1 \rangle_{H_0^1, H^{-1}} + \langle \mathbf{y}^1, \boldsymbol{\varphi}^0 \rangle_{L^2}, \quad \forall \mathbf{y}^0 \in V, \forall \boldsymbol{\varphi}^0 \in W.$$

**Définition 21** (Contrôlabilité à zéro). Le système (1.19) est dit contrôlable à zéro en temps  $T$  si quel que soit  $\mathbf{y}^0 \in V$ , il existe  $u \in \mathcal{U}$  tel que la solution correspondante  $y$  de (1.19) vérifie

$$(y, \partial_t y)(\cdot, T) = (0, 0) \quad \text{dans } \Omega.$$

Pour  $u \in \mathcal{U}$ , on note  $\psi(x, t; u)$  la solution du système rétrograde

$$\begin{cases} \partial_t^2 \psi - \Delta \psi = u \mathbb{1}_q & \text{dans } Q, \\ \psi = 0 & \text{sur } \Sigma, \\ (\psi, \partial_t \psi)(\cdot, T) = (0, 0) & \text{dans } \Omega, \end{cases} \quad \begin{array}{l} (1.23a) \\ (1.23b) \\ (1.23c) \end{array}$$

et on définit l'opérateur linéaire

$$\mathcal{F} : \begin{array}{l} \mathcal{U} \rightarrow V \\ u \mapsto (\psi, \partial_t \psi)(\cdot, 0; u) \end{array}.$$

**Lemme 1.** *Le système (1.19) est contrôlable à zéro en temps  $T$  si et seulement si  $\mathcal{F}$  est surjectif.*

*Démonstration.* Soit  $\mathbf{y}^0 \in V$ . Si  $\mathcal{F}$  est surjectif, il existe  $u \in \mathcal{U}$  tel que  $(\psi, \partial_t \psi)(\cdot, 0; u) = \mathbf{y}^0$ . Par réversibilité de l'équation des ondes, on a alors  $(y, \partial_t y)(\cdot, T; \mathbf{y}^0, u) = (0, 0)$ . La réciproque se démontre de façon similaire.  $\square$

**Lemme 2.** *L'opérateur adjoint  $\mathcal{F}^*$  est donné par*

$$\mathcal{F}^* : \begin{array}{l} W \rightarrow \mathcal{U}' \\ \boldsymbol{\varphi}^0 \mapsto v(\boldsymbol{\varphi}^0) \end{array}, \quad \text{avec } v(\boldsymbol{\varphi}^0) := -\varphi|_q,$$

où  $\varphi$  est la solution de (1.20) associée à  $\boldsymbol{\varphi}^0$ .

*Démonstration.* Soient  $\boldsymbol{\varphi}^0 \in W$  et  $\varphi$  la solution de (1.20) correspondante. Pour  $u \in \mathcal{U}$ , on multiplie l'équation (1.23a) par  $\varphi$  et on intègre sur  $Q$ . Par intégration par parties, on obtient alors

$$\langle \mathcal{F}u, \boldsymbol{\varphi}^0 \rangle_{V,W} = -\langle u, \varphi \rangle_{L^2(q)} = \langle u, v(\boldsymbol{\varphi}^0) \rangle_{\mathcal{U}, \mathcal{U}'} = \langle u, \mathcal{F}^* \boldsymbol{\varphi}^0 \rangle_{\mathcal{U}, \mathcal{U}'}. \quad \square$$

**Proposition 11** (Inégalité d'observabilité). *Le système (1.19) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que*

$$\|\boldsymbol{\varphi}^0\|_W^2 \leq C \|v(\boldsymbol{\varphi}^0)\|_{\mathcal{U}'}^2, \quad \forall \boldsymbol{\varphi}^0 \in W. \quad (1.24)$$

On appelle alors constante d'observabilité la borne inférieure  $C_{\text{obs}}$  des constantes  $C > 0$  telles que l'inégalité (1.24) est vérifiée.

*Démonstration.* On sait par le Lem. 1 que le système (1.19) est contrôlable si et seulement si  $\mathcal{F}$  est surjectif. Or, d'après [33, Th. 4.15],  $\mathcal{F}$  est surjectif si et seulement s'il existe  $C > 0$  tel que

$$\|\varphi^0\|_W \leq C\|\mathcal{F}^*\varphi^0\|_{\mathcal{U}'}, \quad \forall \varphi^0 \in W.$$

On conclut avec le Lem. 2. □

**Corollaire 2.** On note  $\tilde{v}(\varphi^0) := -\partial_t \varphi|_q$ , où  $\varphi$  est la solution de (1.20) associée à  $\varphi^0$ . Le système (1.19) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\|\varphi^0\|_V^2 \leq C\|\tilde{v}(\varphi^0)\|_{\mathcal{U}'}^2, \quad \forall \varphi^0 \in V. \quad (1.25)$$

En outre, les constantes d'observabilité associées à (1.24) et (1.25) sont égales.

*Démonstration.* On montre que les inégalités d'observabilité (1.24) et (1.25) sont équivalentes. On définit l'opérateur linéaire  $\mathcal{R}\varphi^0 := (\Delta^{-1}\varphi^1, \varphi^0)$ . Cet opérateur est une isométrie de  $W$  dans  $V$ , d'inverse  $\mathcal{R}^{-1}\psi^0 = (\psi^1, \Delta\psi^0)$ . Soit  $C > 0$  une constante telle que l'inégalité (1.25) est vérifiée. Pour  $\varphi^0 \in W$  fixé, on pose  $\psi^0 = \mathcal{R}\varphi^0$ . On note  $\varphi$  et  $\psi$  les solutions de (1.20) associées à  $\varphi^0$  et  $\psi^0$  respectivement. Ces solutions sont liées par les relations  $\varphi = \partial_t \psi$  et  $\psi = \Delta^{-1}\partial_t \varphi$ . Par suite, on a

$$\|\varphi^0\|_W^2 = \|\psi^0\|_V^2 \leq C\|\tilde{v}(\psi^0)\|_{\mathcal{U}'}^2 = C\|v(\varphi^0)\|_{\mathcal{U}'}^2.$$

La réciproque se traite de manière similaire. □

Pour  $\mathbf{y}^0 \in V$  fixé, on introduit l'ensemble des contrôles à zéro

$$\mathcal{U}(\mathbf{y}^0) := \left\{ u \in \mathcal{U}; \quad (y, \partial_t y)(\cdot, T; \mathbf{y}^0, u) = (0, 0) \right\}.$$

On définit également la fonctionnelle coût

$$\mathcal{K}(u) := \frac{1}{2}\|u\|_{\mathcal{U}}^2, \quad \forall u \in \mathcal{U}.$$

Le contrôle de norme  $\mathcal{U}$  minimale est ainsi solution du problème de minimisation

$$\inf_{u \in \mathcal{U}(\mathbf{y}^0)} \mathcal{K}(u).$$

On commence par réécrire ce problème de minimisation sous la forme

$$\inf_{u \in \mathcal{U}(\mathbf{y}^0)} \mathcal{K}(u) = \inf_{u \in \mathcal{U}} \left( \mathcal{K}(u) + \mathcal{I}(\mathcal{F}u) \right),$$

où la fonction indicatrice  $\mathcal{I} : V \rightarrow \mathbb{R} \cup \{+\infty\}$  est donnée par

$$\mathcal{I}(\psi^0) := \begin{cases} 0 & \text{si } \psi^0 = \mathbf{y}^0, \\ +\infty & \text{sinon,} \end{cases} \quad \forall \psi^0 \in V.$$

Le théorème de dualité de Fenchel-Rockafellar (cf. [32, Sect. 31]) implique ensuite

$$\inf_{u \in \mathcal{U}(\mathbf{y}^0)} \mathcal{K}(u) = \inf_{u \in \mathcal{U}} \left( \mathcal{K}(u) + \mathcal{I}(\mathcal{F}u) \right) = - \inf_{\varphi^0 \in W} \left( \mathcal{K}^*(\mathcal{F}^*\varphi^0) + \mathcal{I}^*(-\varphi^0) \right),$$

où  $\mathcal{K}^*$  et  $\mathcal{I}^*$  sont respectivement les conjuguées convexes de  $\mathcal{K}$  et  $\mathcal{I}$ . On calcule aisément

$$\mathcal{K}^*(v) = \frac{1}{2}\|v\|_{\mathcal{U}'}^2, \quad \forall v \in \mathcal{U}' \quad \text{et} \quad \mathcal{I}^*(\varphi^0) = \langle \mathbf{y}^0, \varphi^0 \rangle_{V,W}, \quad \forall \varphi^0 \in W.$$

On obtient donc finalement

$$\inf_{u \in \mathcal{U}(\mathbf{y}^0)} \mathcal{K}(u) = - \inf_{\varphi^0 \in W} \mathcal{J}^*(\varphi^0),$$

où la fonctionnelle conjuguée  $\mathcal{J}^*$  s'exprime

$$\mathcal{J}^*(\varphi^0) := \frac{1}{2}\|v(\varphi^0)\|_{\mathcal{U}'}^2 - \langle \mathbf{y}^0, \varphi^0 \rangle_{V,W}, \quad \forall \varphi^0 \in W.$$

Le problème  $\inf_{\mathcal{U}(\mathbf{y}^0)} \mathcal{K}$  est qualifié de problème primal, tandis que le problème  $\inf_W \mathcal{J}^*$  est qualifié de problème dual. Si le problème dual admet un unique point de minimum  $\hat{\varphi}^0 \in W$ , le problème primal admet alors un unique point de minimum  $\hat{u} \in \mathcal{U}(\mathbf{y}^0)$  donné par  $\hat{u} = v(\hat{\varphi}^0)_\chi$ . Par ailleurs, le problème dual étant sans contrainte, c'est ce dernier que l'on privilégie généralement pour approcher le contrôle  $\hat{u}$  de norme  $\mathcal{U}$  minimale.

**Lemme 3.** Soit  $\mathbf{y}^0 \in V$ . La fonctionnelle  $\mathcal{J}^*$  est Gateaux-différentiable. La dérivée de  $\mathcal{J}^*$  au point  $\varphi^0 \in W$  dans la direction  $\bar{\varphi}^0 \in W$  est donnée par

$$d\mathcal{J}^*(\varphi^0; \bar{\varphi}^0) = \langle v(\varphi^0), v(\bar{\varphi}^0) \rangle_{\mathcal{U}'} - \langle \mathbf{y}^0, \bar{\varphi}^0 \rangle_{V,W}.$$

**Proposition 12** (Condition d'optimalité). Soit  $\mathbf{y}^0 \in V$ . Un élément  $\varphi^0 \in W$  est un point critique de  $\mathcal{J}^*$  si et seulement si

$$\langle v(\varphi^0), v(\bar{\varphi}^0) \rangle_{\mathcal{U}'} = \langle \mathbf{y}^0, \bar{\varphi}^0 \rangle_{V,W}, \quad \forall \bar{\varphi}^0 \in W.$$

**Proposition 13** (Contrôle de norme minimale). On suppose que l'inégalité d'observabilité (1.24) est vérifiée. Pour  $\mathbf{y}^0 \in V$ , le contrôle de norme  $\mathcal{U}$  minimale est donné par  $\hat{u} = v(\hat{\varphi}^0)\chi$ , où  $\hat{\varphi}^0 \in W$  est l'unique point de minimum de  $\mathcal{J}^*$ .

*Démonstration.* On montre facilement que la fonctionnelle  $\mathcal{J}^*$  est continue et convexe. Avec l'inégalité d'observabilité (1.24), on peut écrire

$$\mathcal{J}^*(\varphi^0) \geq \frac{C_{\text{obs}}^{-1}}{2} \|\varphi^0\|_W^2 - \|\mathbf{y}^0\|_V \|\varphi^0\|_W, \quad \forall \varphi^0 \in W,$$

ce qui montre la coercivité de  $\mathcal{J}^*$ . La fonctionnelle  $\mathcal{J}^*$  est enfin strictement convexe. En effet, pour  $\lambda, \mu \in (0, 1)$  avec  $\lambda + \mu = 1$  et pour  $\varphi^0, \psi^0 \in W$  avec  $\varphi^0 \neq \psi^0$ , on a

$$\mathcal{J}^*(\lambda\varphi^0 + \mu\psi^0) = \lambda\mathcal{J}^*(\varphi^0) + \mu\mathcal{J}^*(\psi^0) - \frac{\lambda\mu}{2} \|v(\varphi^0 - \psi^0)\|_{\mathcal{U}'}^2$$

et

$$\|v(\varphi^0 - \psi^0)\|_{\mathcal{U}'}^2 \geq C_{\text{obs}}^{-1} \|\varphi^0 - \psi^0\|_W^2 > 0.$$

Ainsi, d'après la méthode directe du calcul des variations (cf. [5, Cor. III.20]),  $\mathcal{J}^*$  admet un unique point de minimum  $\hat{\varphi}^0 \in W$ . Par dualité,  $\hat{u} = v(\hat{\varphi}^0)\chi$  est alors le contrôle de norme  $\mathcal{U}$  minimale.  $\square$

On introduit l'opérateur HUM, qui s'exprime

$$\Lambda : \begin{array}{ccc} W & \rightarrow & V \\ \varphi^0 & \mapsto & (\psi, \partial_t \psi)(\cdot, 0; v(\varphi^0)\chi) \end{array}.$$

Par définition de  $\Lambda$ , on a l'identité

$$\langle \Lambda\varphi^0, \bar{\varphi}^0 \rangle_{V,W} = \langle v(\varphi^0), v(\bar{\varphi}^0) \rangle_{\mathcal{U}'}, \quad \forall \varphi^0, \bar{\varphi}^0 \in W.$$

Pour  $\mathbf{y}^0 \in V$ , la dérivée de  $\mathcal{J}^*$  au point  $\varphi^0 \in W$  dans la direction  $\bar{\varphi}^0 \in W$  s'écrit alors

$$d\mathcal{J}^*(\varphi^0; \bar{\varphi}^0) = \langle \Lambda\varphi^0 - \mathbf{y}^0, \bar{\varphi}^0 \rangle_{V,W}.$$

**Proposition 14** (Caractérisation des contrôles). Soient  $\mathbf{y}^0 \in V$  et  $u \in \mathcal{U}$ . On a alors  $u \in \mathcal{U}(\mathbf{y}^0)$  si et seulement si

$$\langle u, v(\varphi^0) \rangle_{\mathcal{U},\mathcal{U}'} = \langle \mathbf{y}^0, \varphi^0 \rangle_{V,W}, \quad \forall \varphi^0 \in W. \quad (1.26)$$

*Démonstration.* Soient  $\varphi^0 \in W$  et  $\varphi$  la solution de (1.20) correspondante. On multiplie l'équation (1.19a) par  $\varphi$  et on intègre sur  $Q$ . Par intégration par parties, on obtient alors

$$\langle (\mathbf{y}, \partial_t \mathbf{y})(\cdot, T; \mathbf{y}^0, u), (\varphi, \partial_t \varphi)(\cdot, T) \rangle_{V,W} = \langle \mathbf{y}^0, \varphi^0 \rangle_{V,W} - \langle u, v(\varphi^0) \rangle_{\mathcal{U},\mathcal{U}'}. \quad (1.27)$$

On conclut en utilisant la réversibilité de l'équation des ondes.  $\square$

*Remarque 19.* La caractérisation (1.26) des contrôles permet de retrouver le fait que le contrôle de norme  $\mathcal{U}$  minimale est donné par  $\hat{u} = v(\hat{\varphi}^0)\chi$ , où  $\hat{\varphi}^0 \in W$  est le point de minimum de  $\mathcal{J}^*$ . En effet, pour  $\mathbf{y}^0 \in V$  et  $u \in \mathcal{U}(\mathbf{y}^0)$ , avec la condition d'optimalité de  $\hat{\varphi}^0$ , on a

$$\|\hat{u}\|_{\mathcal{U}}^2 = \|v(\hat{\varphi}^0)\|_{\mathcal{U},\mathcal{U}'}^2 = \langle \mathbf{y}^0, \hat{\varphi}^0 \rangle_{V,W} = \langle u, v(\hat{\varphi}^0) \rangle_{\mathcal{U},\mathcal{U}'} = \langle u, \hat{u} \rangle_{\mathcal{U}} \leq \|u\|_{\mathcal{U}} \|\hat{u}\|_{\mathcal{U}}.$$

### Condition géométrique de contrôle

Dans cette section, on énonce un critère géométrique portant sur l'ouvert de contrôle  $q$  et l'horizon de contrôle  $T$ . Il s'agit d'une condition suffisante pour avoir les inégalités d'observabilité (1.21)-(1.22) et donc la contrôlabilité à zéro. Ce critère est souvent abrégé en GCC pour *Geometric Control Condition*. Il est basé sur la propagation des solutions de l'équation des ondes le long des lignes caractéristiques. Par souci de simplicité, on considère le cas unidimensionnel  $\Omega = (0, 1)$ . On rappelle que  $V := H_0^1(\Omega) \times L^2(\Omega)$ .

**Proposition 15** (Formule de d'Alembert). [14, Sect. 2.4] *On considère  $\Omega = (0, 1)$ . Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $f \in L^1(0, T; L^2(\Omega))$ . On construit  $(\tilde{\varphi}^0, \tilde{\varphi}^1) \in H_{\text{loc}}^1(\mathbb{R}) \times L_{\text{loc}}^2(\mathbb{R})$  en prolongeant  $(\varphi^0, \varphi^1)$  par imparité sur  $(-1, 1)$  puis par 2-périodicité sur  $\mathbb{R}$ , i.e.*

$$\tilde{\varphi}^0(x) := \begin{cases} \varphi^0(\mathbf{m}(x)) & \text{si } \mathbf{m}(x) \geq 0, \\ -\varphi^0(-\mathbf{m}(x)) & \text{si } \mathbf{m}(x) < 0, \end{cases} \quad \text{et} \quad \tilde{\varphi}^1(x) := \begin{cases} \varphi^1(\mathbf{m}(x)) & \text{si } \mathbf{m}(x) \geq 0, \\ -\varphi^1(-\mathbf{m}(x)) & \text{si } \mathbf{m}(x) < 0, \end{cases}$$

avec  $\mathbf{m}(x) := x - 2k$  pour  $x \in (2k - 1, 2k + 1]$ ,  $k \in \mathbb{Z}$ . On construit de même un prolongement  $\tilde{f} \in L^1(0, T; L_{\text{loc}}^2(\mathbb{R}))$ , i.e.

$$\tilde{f}(x, t) := \begin{cases} f(\mathbf{m}(x), t) & \text{si } \mathbf{m}(x) \geq 0, \\ -f(-\mathbf{m}(x), t) & \text{si } \mathbf{m}(x) < 0. \end{cases}$$

On peut alors montrer que la solution  $\varphi$  de l'équation des ondes inhomogène

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = f & \text{dans } Q, \\ \varphi = 0 & \text{sur } \Sigma, \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega, \end{cases}$$

est donnée par

$$\varphi(x, t) = \frac{1}{2} \left( \tilde{\varphi}^0(x+t) + \tilde{\varphi}^0(x-t) \right) + \frac{1}{2} \int_{x-t}^{x+t} \tilde{\varphi}^1(y) dy + \frac{1}{2} \int_0^t \int_{x-t+s}^{x+t-s} \tilde{f}(y, s) dy ds. \quad (1.27)$$

**Définition 22** (Caractéristiques). Soit  $x_0 \in \overline{\Omega}$ . On appelle lignes caractéristiques issues de  $x_0$  les ensembles  $C_{x_0}^+$  et  $C_{x_0}^-$  définis par

$$C_{x_0}^\pm := \left\{ (x, t) \in \mathbb{R}^2; \quad x = |\mathbf{m}(x_0 \pm t)| \right\}, \quad (1.28)$$

avec  $\mathbf{m}(x) := x - 2k$  pour  $x \in (2k - 1, 2k + 1]$ ,  $k \in \mathbb{Z}$ .

À la Fig. 1.2, on trace les deux lignes caractéristiques  $C_{x_0}^\pm$  issues d'un point  $x_0$ . On constate que ces courbes sont obtenues à partir des droites  $x = x_0 \pm t$  et de leurs réflexions successives sur le bord de  $\Omega$ . Pour  $(x, t) \in \overline{Q}$  fixé, on pose  $x_1 = \mathbf{m}(x - t)$  et  $x_2 = \mathbf{m}(x + t)$ , ainsi que  $s_1 = \text{sgn}(x_1)$  et  $s_2 = \text{sgn}(x_2)$ . On peut alors montrer que  $(x, t) \in C_{|x_1|}^{+s_1} \cap C_{|x_2|}^{-s_2}$ . Autrement dit, le point  $(x, t)$  se trouve à l'intersection d'une caractéristique issue de  $|x_1|$  et d'une caractéristique issue de  $|x_2|$  (cf. Fig. 1.3). Pour  $(\varphi^0, \varphi^1) \in V$ , on note  $\varphi$  la solution de l'équation des ondes homogène (1.20). En utilisant la formule de d'Alembert (voir Prop. 15), il vient

$$\varphi(x, t) = \frac{1}{2} \left( s_2 \varphi^0(|x_2|) + s_1 \varphi^0(|x_1|) \right) + \frac{1}{2} \int_{|x_1|}^{|x_2|} \varphi^1(y) dy.$$

La valeur de  $\varphi$  en  $(x, t)$  dépend donc uniquement des valeurs de la donnée initiale  $(\varphi^0, \varphi^1)$  sur le segment d'extrémités  $|x_1|$  et  $|x_2|$ . De plus, si  $\varphi^1 \equiv 0$ , on voit que la solution  $\varphi$  est obtenue en propageant  $\varphi^0$  le long des lignes caractéristiques.

**Définition 23** (GCC). On dit qu'un ouvert  $q \subset Q$  vérifie le critère (GCC) (pour *Geometric Control Condition*) si pour tout  $x_0 \in \overline{\Omega}$ , les lignes caractéristiques  $C_{x_0}^\pm$  issues de  $x_0$  rencontrent  $q$ , i.e.  $q \cap C_{x_0}^\pm \neq \emptyset$ .

*Remarque 20.* En dimension  $d > 1$ , la notion de ligne caractéristique dans la définition du critère (GCC) doit être remplacée par la notion de bicaractéristique généralisée (cf. [24, Sect. 1.3.2]).

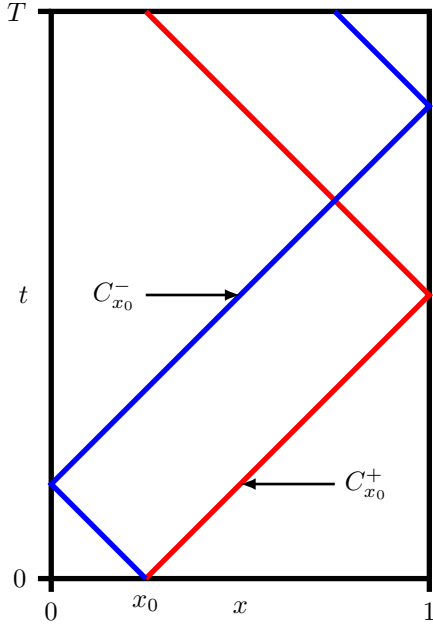


FIGURE 1.2 – Lignes caractéristiques issues de  $x_0$ .

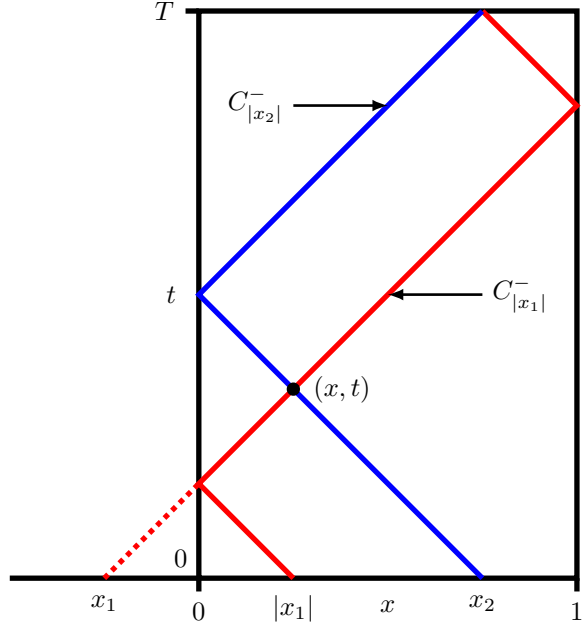


FIGURE 1.3 – Lignes caractéristiques passant par  $(x, t)$ .

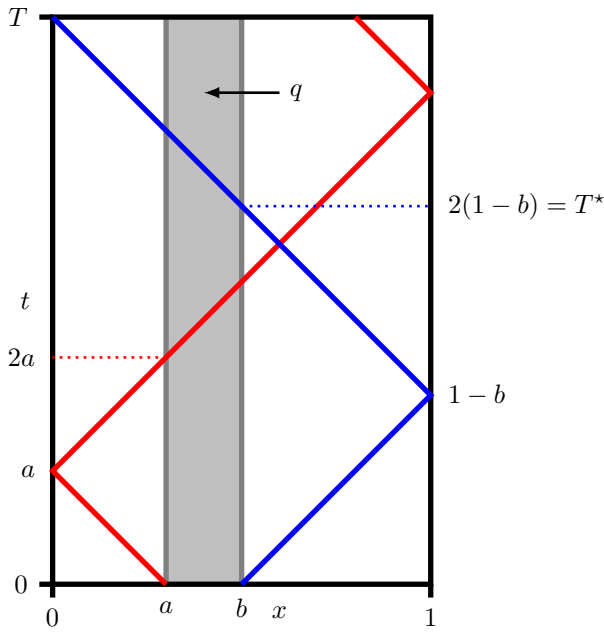


FIGURE 1.4 – Ouvert de contrôle  $q$  cylindrique.

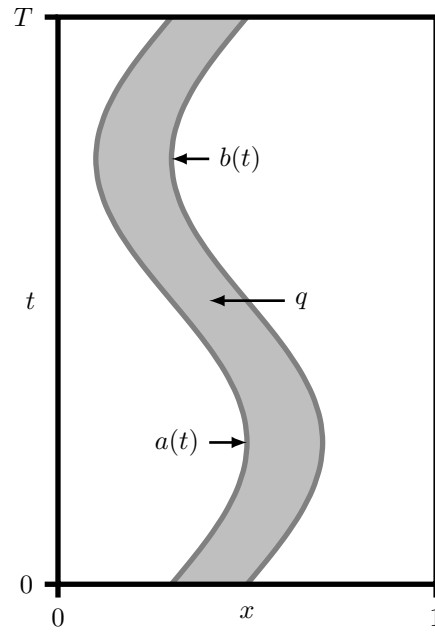


FIGURE 1.5 – Ouvert de contrôle  $q$  courbe.

*Remarque 21.* Un ouvert de contrôle  $q$  est dit cylindrique (cf. Fig. 1.4) s'il est de la forme  $q = \{(x, t) \in Q; x \in (a, b)\}$ , avec  $a, b \in \bar{\Omega}$  et  $a < b$ . On montre facilement qu'un ouvert de contrôle  $q$  cylindrique vérifie (GCC) si et seulement si  $T > T^* := 2 \max(a, 1 - b)$ . Un ouvert de contrôle  $q$  est dit courbe (cf. Fig. 1.5) s'il est de la forme  $q = \{(x, t) \in Q; x \in (a(t), b(t))\}$ , avec  $a, b \in C([0, T]; \bar{\Omega})$  et  $a < b$ .

*Remarque 22.* Soit  $q \subset Q$  un ouvert ne vérifiant pas (GCC). Plus précisément, on suppose qu'il existe un intervalle  $I \subset \bar{\Omega}$  tel que pour tout  $x_0 \in I$ , la ligne caractéristique  $C_{x_0}^-$  ne rencontre pas  $q$ . On note  $C_I^- = \bigcup_{x_0 \in I} C_{x_0}^-$ . On peut alors montrer que l'inégalité d'observabilité (1.21) (et donc (1.22)) n'est pas vérifiée. En effet, si on considère  $(\varphi^0, \varphi^1) \in V$  tel que  $\text{Supp } \varphi^0 \subset I$  et  $\varphi^1 = (\varphi^0)'$ , on observe que la



solution correspondante  $\varphi$  de (1.20) vérifie  $\text{Supp } \varphi \subset C_I^-$ . Comme  $q \cap C_I^- = \emptyset$  par hypothèse, on obtient donc  $\partial_t \varphi|_q \equiv 0$ .

**Proposition 16** (GCC  $\Rightarrow$  observabilité). *Si l'ouvert de contrôle  $q \subset Q$  vérifie (GCC), il existe alors  $C > 0$  tel que les inégalités d'observabilité (1.21)-(1.22) sont vérifiées.*

*Remarque 23.* Il existe plusieurs manières de démontrer les inégalités d'observabilité (1.21)-(1.22) à partir de (GCC). Dans le cas d'un ouvert de contrôle cylindrique, ces inégalités peuvent être montrées en 1D par

- ▷ la méthode des multiplicateurs (voir [21]),
- ▷ une méthode spectrale (voir [20, Th. 1]),
- ▷ une méthode propagative (voir [20, Th. 2]),

et en  $dD$  par analyse microlocale (voir [3]). Dans le cas d'un ouvert de contrôle non cylindrique, ces inégalités peuvent être montrées en 1D par une méthode propagative (voir [8]) et en  $dD$  par analyse microlocale (voir [24]).

### 1.1.7 Contrôlabilité de l'équation des ondes linéaire avec contrôle au bord

Soit  $\Omega$  un domaine borné non vide de  $\mathbb{R}^d$ , avec  $d \in \{1, 2, 3\}$ . Le bord  $\Gamma := \partial\Omega$  est supposé de classe  $C^2$ . Alternativement, on peut supposer  $\Omega$  convexe et  $\Gamma$  lipschitzien. Pour  $T > 0$ , on pose  $Q := \Omega \times (0, T)$  et  $\Sigma := \Gamma \times (0, T)$ . On se donne également  $\gamma \subset \Gamma$  un ouvert non vide et on note  $\sigma := \gamma \times (0, T)$ . On considère le système d'état

$$\begin{cases} \partial_t^2 y - \Delta y = 0 & \text{dans } Q, & (1.29a) \\ y = u \mathbb{1}_\gamma & \text{sur } \Sigma, & (1.29b) \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega, & (1.29c) \end{cases}$$

ainsi que le système adjoint

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = 0 & \text{dans } Q, & (1.30a) \\ \varphi = 0 & \text{sur } \Sigma, & (1.30b) \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega. & (1.30c) \end{cases}$$

La fonction  $\mathbb{1}_\gamma : \Gamma \rightarrow \{0, 1\}$  désigne la fonction caractéristique de  $\gamma$ . Par souci de concision, on notera parfois  $\mathbf{y}^0 = (y^0, y^1)$  et  $\boldsymbol{\varphi}^0 = (\varphi^0, \varphi^1)$ . De même, on pourra noter  $y(x, t; \mathbf{y}^0, u)$  la solution de (1.29) avec  $\mathbf{y}^0$  comme donnée initiale et  $u$  comme contrôle, ainsi que  $\varphi(x, t; \boldsymbol{\varphi}^0)$  la solution de (1.30) avec  $\boldsymbol{\varphi}^0$  comme donnée initiale.

*Remarque 24* (Régularité des solutions). Le système d'état (1.29) et le système adjoint (1.30) sont bien posés.

- ▷ Pour  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  et  $u \in L^2(\sigma)$ , il existe une unique solution  $y$  de (1.29) vérifiant

$$(y, \partial_t y) \in C([0, T]; L^2(\Omega) \times H^{-1}(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|y\|_{L^\infty(L^2)} + \|\partial_t y\|_{L^\infty(H^{-1})} \leq C_{\Omega, T} \left( \|(y^0, y^1)\|_{L^2 \times H^{-1}} + \|u\|_{L^2(\sigma)} \right).$$

- ▷ Pour  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , il existe une unique solution  $\varphi$  de (1.30) vérifiant

$$(\varphi, \partial_t \varphi) \in C([0, T]; H_0^1(\Omega) \times L^2(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|\varphi\|_{L^\infty(H_0^1)} + \|\partial_t \varphi\|_{L^\infty(L^2)} \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}.$$

**Définition 24** (Contrôlabilité à zéro). Le système (1.29) est dit contrôlable à zéro en temps  $T$  si quel que soit  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , il existe  $u \in L^2(\sigma)$  tel que la solution correspondante  $y$  de (1.29) vérifie

$$(y, \partial_t y)(\cdot, T) = (0, 0) \quad \text{dans } \Omega.$$

**Proposition 17** (Inégalité directe). [25, Cor. I.4.1] L'assertion suivante est vérifiée.

▷ Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $\varphi$  la solution de (1.30) correspondante. Il existe  $C_{\Omega, T} > 0$  tel que

$$\int_{\sigma} |\partial_{\nu} \varphi|^2 \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

*Démonstration.* D'après [25, Lem. I.3.1], si le bord  $\Gamma$  est de classe  $C^2$ , il existe un champ de vecteurs  $\theta \in [C^1(\overline{\Omega})]^d$  tel que  $\theta(x) = \nu(x)$  pour tout  $x \in \Gamma$ . Avec ce choix de  $\theta$ , l'identité (1.31) fournit

$$\int_{\sigma} |\partial_{\nu} \varphi|^2 \leq \int_{\Sigma} |\partial_{\nu} \varphi|^2 \leq C_{\Omega}(T+1) \left( \|\varphi\|_{L^{\infty}(H_0^1)}^2 + \|\partial_t \varphi\|_{L^{\infty}(L^2)}^2 \right) \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

La dernière inégalité provient de la continuité de  $\varphi$  par rapport à la donnée initiale. Pour le cas avec  $\Omega$  convexe et  $\Gamma$  lipschitzien, on renvoie à [25, Rem. I.4.3].  $\square$

**Proposition 18** (Inégalité directe en 1D). On suppose ici que  $\Omega = (0, 1)$ . L'assertion suivante est vérifiée.

▷ Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $\varphi$  la solution de (1.30) correspondante. Pour tous  $t_1, t_2 \in \mathbb{R}$  avec  $t_1 \leq t_2$ , on a

$$\int_{t_1}^{t_2} |\partial_x \varphi(0, t)|^2 dt \leq (t_2 - t_1 + 2) \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

*Démonstration.* En écrivant l'identité (1.31) avec  $\theta(x) = 1 - x$ , il vient

$$\frac{1}{2} \int_{t_1}^{t_2} |\partial_x \varphi(0, t)|^2 dt = \frac{1}{2} \int_{t_1}^{t_2} \int_{\Omega} \left( |\partial_x \varphi|^2 + |\partial_t \varphi|^2 \right) - \left[ \int_{\Omega} (1-x) \partial_x \varphi \partial_t \varphi \right]_{t_1}^{t_2}.$$

Or, la conservation de l'énergie (cf. [25, Lem. I.3.5]) donne

$$\int_{\Omega} \left( |\partial_x \varphi(x, t)|^2 + |\partial_t \varphi(x, t)|^2 \right) dx = \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2, \quad \forall t \in \mathbb{R}.$$

On conclut en remarquant que

$$\left| \int_{\Omega} (1-x) \partial_x \varphi(x, t) \partial_t \varphi(x, t) dx \right| \leq \frac{1}{2} \int_{\Omega} \left( |\partial_x \varphi(x, t)|^2 + |\partial_t \varphi(x, t)|^2 \right) dx, \quad \forall t \in \mathbb{R}.$$

$\square$

**Lemme 4.** [25, Cor. I.3.1] Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $\varphi$  la solution de (1.30) correspondante. Pour tout champ de vecteurs  $\theta \in [C^1(\overline{\Omega})]^d$ , on a l'identité

$$\frac{1}{2} \int_{\Sigma} (\theta \cdot \nu) |\partial_{\nu} \varphi|^2 = \frac{1}{2} \iint_Q \operatorname{div}(\theta) \left( |\partial_t \varphi|^2 - |\nabla \varphi|^2 \right) + \iint_Q (\nabla \theta \nabla \varphi) \cdot \nabla \varphi + \left[ \int_{\Omega} (\theta \cdot \nabla \varphi) \partial_t \varphi \right]_0^T. \quad (1.31)$$

*Démonstration.* L'identité se démontre par la méthode des multiplicateurs. Il s'agit de multiplier l'équation (1.30a) par  $\theta \cdot \nabla \varphi$ , d'intégrer sur  $Q$  et de procéder par intégrations par parties.  $\square$

On pose  $H := H^{-1}(\Omega)$ ,  $U := L^2(\gamma)$  et  $\mathcal{U} := L^2(\sigma)$ . On considère les opérateurs

$$A_0 : \begin{array}{ccc} H_0^1(\Omega) & \rightarrow & H^{-1}(\Omega) \\ y & \mapsto & -\Delta y \end{array} \quad \text{et} \quad B_0 : \begin{array}{ccc} L^2(\gamma) & \rightarrow & H^{-2}(\Omega) \\ u & \mapsto & A_0 D u \end{array},$$

où  $D : L^2(\gamma) \rightarrow L^2(\Omega)$  est l'opérateur défini par

$$\langle D u, \zeta \rangle_{L^2} = -\langle u, \partial_{\nu} (A_0^{-1} \zeta) \rangle_{L^2(\gamma)}, \quad \forall \zeta \in L^2(\Omega).$$

En reprenant les notations des Sect. 1.1.4 et 1.1.3, on a

$$Z_1 = H_0^1(\Omega) \times L^2(\Omega), \quad Z = L^2(\Omega) \times H^{-1}(\Omega), \quad Z_{-1} = H^{-1}(\Omega) \times H^{-2}(\Omega).$$

On rappelle que les opérateurs  $A : Z_1 \rightarrow Z$  et  $B \in \mathcal{L}(U, Z_{-1})$  apparaissant dans (1.4) sont donnés par

$$A = \begin{pmatrix} 0 & \operatorname{Id} \\ -A_0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 \\ B_0 \end{pmatrix}.$$

On rappelle également que  $v(\zeta^0) = -B^*\zeta(\cdot; \zeta^0)$ , où  $\zeta(t; \zeta^0)$  est la solution de (1.9) avec  $\zeta^0$  comme donnée initiale. On constate ici que

$$B^* : \begin{array}{ccc} Z_1 & \rightarrow & U \\ (\zeta_1, \zeta_2) & \mapsto & -\partial_\nu(A_0^{-1}\zeta_2)|_\gamma \end{array} .$$

En effet, pour  $u \in U$  et  $\zeta \in Z_1$ , on a

$$\langle Bu, \zeta \rangle_{Z_{-1}, Z_1} = -\langle A^{-1}Bu, A\zeta \rangle_Z = \langle A_0^{-1}B_0u, \zeta_2 \rangle_{L^2} = \langle Du, \zeta_2 \rangle_{L^2} = -\langle u, \partial_\nu(A_0^{-1}\zeta_2) \rangle_{L^2(\gamma)} = \langle u, B^*\zeta \rangle_U .$$

*Remarque 25.* On énonce une relation entre la solution  $\varphi$  de (1.30) et la solution  $\zeta$  de (1.9).

▷ Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $\zeta^0 = A(\varphi^0, \varphi^1)$ , on a  $\zeta = (\partial_t\varphi, -A_0\varphi)$  et  $v(\zeta^0) = -\partial_\nu\varphi|_\sigma$ . De même, on peut écrire la relation de dualité suivante.

▷ Soient  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  et  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $z^0 = (y^0, y^1)$  et  $\zeta^0 = A(\varphi^0, \varphi^1)$ , on a

$$\langle z^0, \zeta^0 \rangle_Z = \langle y^0, \varphi^1 \rangle_{L^2} - \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1} .$$

**Proposition 19** (Inégalité d'observabilité). *Par dualité entre contrôlabilité et observabilité, l'assertion suivante est vérifiée.*

▷ Le système (1.29) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq C \|\partial_\nu\varphi\|_{L^2(\sigma)}^2, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega), \quad (1.32)$$

où  $\varphi$  est la solution de (1.30) associée à  $(\varphi^0, \varphi^1)$ .

**Proposition 20** (Contrôle de norme minimale). *En appliquant la méthode HUM, l'assertion suivante est vérifiée.*

▷ On suppose que l'inégalité d'observabilité (1.32) est vérifiée. Pour  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , le contrôle de norme  $L^2(\sigma)$  minimale pour le système (1.29) est donné par  $\hat{u} = -\partial_\nu\hat{\varphi}|_\sigma$ , avec  $\hat{\varphi}$  la solution de (1.30) associée au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) := \frac{1}{2} \|\partial_\nu\varphi\|_{L^2(\sigma)}^2 - \langle y^0, \varphi^1 \rangle_{L^2} + \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1}, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.30) associée à  $(\varphi^0, \varphi^1)$ .

À la proposition suivante, on établit l'inégalité d'observabilité (1.32) pour un bord de contrôle  $\gamma \subset \Gamma$  défini par une condition géométrique, ainsi que pour  $T$  assez grand. La démonstration repose sur la méthode des multiplicateurs. Plus précisément, si  $x_0$  ci-dessous est pris dans  $\mathbb{R}^d \setminus \bar{\Omega}$ , le bord de contrôle  $\gamma = \Gamma_0$  correspond alors au complémentaire dans  $\Gamma$  de la partie du bord éclairée depuis  $x_0$ .

**Proposition 21** (Inégalité d'observabilité). *[25, Th. I.5.1] Pour  $x_0 \in \mathbb{R}^d$  fixé, on définit le champ de vecteurs  $\theta \in [C^1(\bar{\Omega})]^d$  par  $\theta(x) = x - x_0$ . On introduit également la partie du bord*

$$\Gamma_0 := \left\{ x \in \Gamma; \quad \theta(x) \cdot \nu(x) > 0 \right\} .$$

On pose  $R_0 := \max_{x \in \bar{\Omega}} |\theta(x)|$  et  $T^* := 2R_0$ . Pour tout  $T > T^*$ , l'inégalité suivante est vérifiée,

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq \frac{R_0}{T - T^*} \int_{\Sigma_0} |\partial_\nu\varphi|^2, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.30) associée à  $(\varphi^0, \varphi^1)$ .

*Démonstration.* Avec le choix de  $\theta$ , l'identité (1.31) s'écrit

$$\begin{aligned} \frac{1}{2} \int_{\Sigma} (\theta \cdot \nu) |\partial_\nu\varphi|^2 &= \frac{d}{2} \iint_Q (|\partial_t\varphi|^2 - |\nabla\varphi|^2) + \iint_Q |\nabla\varphi|^2 + \left[ \int_{\Omega} (\theta \cdot \nabla\varphi) \partial_t\varphi \right]_0^T \\ &= \frac{1}{2} \iint_Q (|\nabla\varphi|^2 + |\partial_t\varphi|^2) + \frac{d-1}{2} \iint_Q (|\partial_t\varphi|^2 - |\nabla\varphi|^2) + \left[ \int_{\Omega} (\theta \cdot \nabla\varphi) \partial_t\varphi \right]_0^T . \end{aligned}$$

Comme  $\theta \cdot \nu \leq 0$  sur  $\Sigma \setminus \Sigma_0$ , on a

$$\int_{\Sigma} (\theta \cdot \nu) |\partial_{\nu} \varphi|^2 \leq \int_{\Sigma_0} (\theta \cdot \nu) |\partial_{\nu} \varphi|^2 \leq R_0 \int_{\Sigma_0} |\partial_{\nu} \varphi|^2.$$

En multipliant l'équation (1.30a) par  $\varphi$  et en intégrant par parties sur  $Q$ , on montre que

$$\iint_Q (|\partial_t \varphi|^2 - |\nabla \varphi|^2) = \left[ \int_{\Omega} \varphi \partial_t \varphi \right]_0^T.$$

On en déduit alors que

$$\frac{R_0}{2} \int_{\Sigma_0} |\partial_{\nu} \varphi|^2 \geq \frac{1}{2} \iint_Q (|\nabla \varphi|^2 + |\partial_t \varphi|^2) + \left[ \int_{\Omega} (\theta \cdot \nabla \varphi + \frac{d-1}{2} \varphi) \partial_t \varphi \right]_0^T.$$

D'une part, la conservation de l'énergie (cf. [25, Lem. I.3.5]) donne

$$\iint_Q (|\nabla \varphi|^2 + |\partial_t \varphi|^2) = T \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

D'autre part, pour tout  $t \in [0, T]$ , on a

$$\left| \int_{\Omega} (\theta \cdot \nabla \varphi + \frac{d-1}{2} \varphi) \partial_t \varphi \right| \leq \frac{1}{2R_0} \int_{\Omega} |\theta \cdot \nabla \varphi + \frac{d-1}{2} \varphi|^2 + \frac{R_0}{2} \int_{\Omega} |\partial_t \varphi|^2.$$

Or, sachant que

$$\int_{\Omega} (\theta \cdot \nabla \varphi) \varphi = \frac{1}{2} \int_{\Omega} \theta \cdot \nabla |\varphi|^2 = -\frac{1}{2} \int_{\Omega} \operatorname{div} \theta |\varphi|^2 = -\frac{d}{2} \int_{\Omega} |\varphi|^2,$$

il vient

$$\begin{aligned} \int_{\Omega} \left| \theta \cdot \nabla \varphi + \frac{d-1}{2} \varphi \right|^2 &= \int_{\Omega} |\theta \cdot \nabla \varphi|^2 + (d-1) \int_{\Omega} (\theta \cdot \nabla \varphi) \varphi + \frac{(d-1)^2}{4} \int_{\Omega} |\varphi|^2 \\ &= \int_{\Omega} |\theta \cdot \nabla \varphi|^2 - \frac{(d-1)(d+1)}{4} \int_{\Omega} |\varphi|^2 \leq R_0^2 \int_{\Omega} |\nabla \varphi|^2. \end{aligned}$$

Par suite, on a

$$\left| \left[ \int_{\Omega} (\theta \cdot \nabla \varphi + \frac{d-1}{2} \varphi) \partial_t \varphi \right]_0^T \right| \leq R_0 \sup_{t \in [0, T]} \int_{\Omega} (|\nabla \varphi|^2 + |\partial_t \varphi|^2) = R_0 \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2,$$

d'où

$$\frac{R_0}{2} \int_{\Sigma_0} |\partial_{\nu} \varphi|^2 \geq \frac{T - T^*}{2} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

□

À la proposition suivante, on établit l'inégalité d'observabilité (1.32) en 1D, pour  $\Omega = (0, 1)$  et un contrôle localisé en  $x = 1$ . La démonstration repose sur une méthode spectrale. On signale par ailleurs que les méthodes spectrales sont généralement basées sur l'inégalité d'Ingham (cf. [23]).

**Proposition 22** (Inégalité d'observabilité en 1D). *On suppose ici que  $\Omega = (0, 1)$ . Pour tout  $T \geq 2$ , l'inégalité suivante est vérifiée,*

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq \frac{1}{2 \lfloor T/2 \rfloor} \int_0^T |\partial_x \varphi(1, t)|^2 dt, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.30) associée à  $(\varphi^0, \varphi^1)$ .

*Démonstration.* Pour tous  $p, q \geq 1$ , on a

$$\begin{cases} \int_0^2 \cos(p\pi t) \cos(q\pi t) dt = \delta_{p,q}, \\ \int_0^2 \sin(p\pi t) \sin(q\pi t) dt = \delta_{p,q}, \\ \int_0^2 \cos(p\pi t) \sin(q\pi t) dt = 0, \end{cases} \quad \text{et} \quad \begin{cases} \int_0^1 \cos(p\pi x) \cos(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \\ \int_0^1 \sin(p\pi x) \sin(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \end{cases}$$

où  $\delta_{p,q} = 0$  si  $p \neq q$  et  $\delta_{p,p} = 1$ . Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . On décompose

$$\varphi^0(x) = \sum_{p=1}^{+\infty} a_p \sin(p\pi x) \quad \text{et} \quad \varphi^1(x) = \sum_{p=1}^{+\infty} b_p \sin(p\pi x).$$

La solution  $\varphi$  de (1.30) s'écrit alors

$$\varphi(x, t) = \sum_{p=1}^{+\infty} \left( a_p \cos(p\pi t) + \frac{b_p}{p\pi} \sin(p\pi t) \right) \sin(p\pi x).$$

On calcule

$$\int_0^2 |\partial_x \varphi(1, t)|^2 dt = \sum_{p=1}^{+\infty} \left( (p\pi)^2 a_p^2 + b_p^2 \right) = 2 \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

Par suite, avec la 2-périodicité en temps de  $\varphi$ , il vient

$$\int_0^T |\partial_x \varphi(1, t)|^2 dt \geq \int_0^{2\lfloor T/2 \rfloor} |\partial_x \varphi(1, t)|^2 dt = \lfloor T/2 \rfloor \int_0^2 |\partial_x \varphi(1, t)|^2 dt = 2 \lfloor T/2 \rfloor \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

□

### Condition géométrique de contrôle

Dans cette section, on énonce un critère géométrique portant sur le bord de contrôle  $\gamma$  et l'horizon de contrôle  $T$ . Il s'agit d'une condition suffisante pour avoir l'inégalité d'observabilité (1.32) et donc la contrôlabilité à zéro. Ce critère est souvent abrégé en GCC pour *Geometric Control Condition*. Il est basé sur la propagation des solutions de l'équation des ondes le long des lignes caractéristiques. Par souci de simplicité, on considère le cas unidimensionnel  $\Omega = (0, 1)$ . On renvoie à la Sect. 1.1.6 et à la Déf. 22 pour la définition des lignes caractéristiques  $C_{x_0}^\pm$ .

**Définition 25** (GCC). On dit qu'un ouvert  $\gamma \subset \Gamma$  vérifie le critère (GCC) (pour *Geometric Control Condition*) si pour tout  $x_0 \in \Omega$ , les lignes caractéristiques  $C_{x_0}^\pm$  issues de  $x_0$  rencontrent  $\sigma = \gamma \times (0, T)$ , i.e.  $\sigma \cap C_{x_0}^\pm \neq \emptyset$ .

*Remarque 26.* En dimension  $d > 1$ , la notion de ligne caractéristique dans la définition du critère (GCC) doit être remplacée par la notion de bicaractéristique généralisée (cf. [24, Sect. 1.3.2]).

*Remarque 27.* Ainsi, pour un contrôle agissant en  $\gamma = \{0\}$  ou  $\gamma = \{1\}$ , le critère (GCC) est vérifié à partir de  $T = 2$ , tandis que pour un contrôle agissant en  $\gamma = \{0, 1\}$ , le critère (GCC) est vérifié à partir de  $T = 1$ .

**Proposition 23** (GCC  $\Rightarrow$  observabilité). *Si le bord de contrôle  $\gamma \subset \Gamma$  vérifie (GCC), il existe alors  $C > 0$  tel que l'inégalité d'observabilité (1.32) est vérifiée.*

*Remarque 28.* En 1D comme en  $dD$ , l'inégalité d'observabilité (1.32) peut être démontrée à partir de (GCC) par analyse microlocale (voir [3, 24]).

### 1.1.8 Contrôlabilité de l'équation des ondes linéaire avec contrôle ponctuel

Soient  $\Omega := (0, 1)$  et  $\Gamma := \partial\Omega = \{0, 1\}$ . Pour  $T > 0$ , on pose  $Q := \Omega \times (0, T)$  et  $\Sigma := \Gamma \times (0, T)$ . On se donne également  $\gamma : [0, T] \rightarrow \Omega$  une courbe  $C^1$  par morceaux telle que  $1 - |\gamma'|$  ne change de signe qu'un nombre fini de fois sur  $[0, T]$ . On fixe enfin  $\eta \in C^1(\mathbb{R})$  tel que  $\eta : \mathbb{R} \rightarrow [0, 1]$  et  $\text{Supp } \eta = [0, T]$ . Plus précisément, pour  $\tau > 0$  fixé, on impose  $\eta(t) = 1$  si  $t \in [\tau, T - \tau]$  et  $\eta(t) = 0$  si  $t \notin (0, T)$ . On considère le système d'état

$$\begin{cases} \partial_t^2 y - \Delta y = u\eta\delta_\gamma & \text{dans } Q, \\ y = 0 & \text{sur } \Sigma, \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega, \end{cases} \quad \begin{array}{l} (1.33a) \\ (1.33b) \\ (1.33c) \end{array}$$

ainsi que le système adjoint

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = 0 & \text{dans } Q, \\ \varphi = 0 & \text{sur } \Sigma, \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega. \end{cases} \quad \begin{array}{l} (1.34a) \\ (1.34b) \\ (1.34c) \end{array}$$

La fonction  $\delta_{\gamma(t)}(x)$  désigne la distribution de Dirac en  $x = \gamma(t)$ . Par souci de concision, on notera parfois  $\mathbf{y}^0 = (y^0, y^1)$  et  $\boldsymbol{\varphi}^0 = (\varphi^0, \varphi^1)$ . De même, on pourra noter  $y(x, t; \mathbf{y}^0, u)$  la solution de (1.33) avec  $\mathbf{y}^0$  comme donnée initiale et  $u$  comme contrôle, ainsi que  $\varphi(x, t; \boldsymbol{\varphi}^0)$  la solution de (1.34) avec  $\boldsymbol{\varphi}^0$  comme donnée initiale. En outre,  $\delta_{\gamma}^{\text{op}}(\varphi)$  désignera la fonction  $\varphi(\gamma(t), t)$ .

*Remarque 29* (Régularité des solutions). Le système d'état (1.33) et le système adjoint (1.34) sont bien posés.

▷ Pour  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  et  $u \in H^{-1}(0, T)$ , il existe une unique solution  $y$  de (1.33) vérifiant

$$(y, \partial_t y) \in C([0, T]; L^2(\Omega) \times H^{-1}(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|y\|_{L^\infty(L^2)} + \|\partial_t y\|_{L^\infty(H^{-1})} \leq C_{\Omega, T} \left( \|(y^0, y^1)\|_{L^2 \times H^{-1}} + \|u\|_{H^{-1}(0, T)} \right).$$

▷ Pour  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , il existe une unique solution  $\varphi$  de (1.34) vérifiant

$$(\varphi, \partial_t \varphi) \in C([0, T]; H_0^1(\Omega) \times L^2(\Omega)).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|\varphi\|_{L^\infty(H_0^1)} + \|\partial_t \varphi\|_{L^\infty(L^2)} \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}.$$

**Définition 26** (Contrôlabilité à zéro). Le système (1.33) est dit contrôlable à zéro en temps  $T$  si quel que soit  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , il existe  $u \in H^{-1}(0, T)$  tel que la solution correspondante  $y$  de (1.33) vérifie

$$(y, \partial_t y)(\cdot, T) = (0, 0) \quad \text{dans } \Omega.$$

**Proposition 24** (Inégalité directe). [7, Prop. 4.1] L'assertion suivante est vérifiée.

▷ Soient  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et  $\varphi$  la solution de (1.34) correspondante. Il existe  $C_{\gamma, T} > 0$  tel que

$$\int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt \leq C_{\gamma, T} \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

*Démonstration.* On suppose que  $\gamma \in C^1([0, T])$  et que  $1 - |\gamma'| \geq 0$  sur  $[0, T]$ . En échangeant le rôle de  $x$  et de  $t$  dans l'équation des ondes, la formule de d'Alembert fournit

$$\varphi(\gamma(t), t) = \frac{1}{2} \int_{t-\gamma(t)}^{t+\gamma(t)} \partial_x \varphi(0, s) ds, \quad \forall (x, t) \in Q.$$

Par suite, on a

$$\begin{aligned} \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt &\leq \frac{1}{2} \int_0^T |\partial_x \varphi(0, t + \gamma(t))|^2 |1 + \gamma'(t)|^2 dt + \frac{1}{2} \int_0^T |\partial_x \varphi(0, t - \gamma(t))|^2 |1 - \gamma'(t)|^2 dt \\ &\leq \frac{\varsigma_\gamma}{2} \int_0^T |\partial_x \varphi(0, t + \gamma(t))|^2 |1 + \gamma'(t)| dt + \frac{\varsigma_\gamma}{2} \int_0^T |\partial_x \varphi(0, t - \gamma(t))|^2 |1 - \gamma'(t)| dt, \end{aligned}$$

où  $\varsigma_\gamma = 1 + \|\gamma'\|_{L^\infty(0, T)}$ . En utilisant les changements de variable  $s = t \pm \gamma(t)$  dans les intégrales ci-dessus, il vient

$$\int_0^T |\partial_x \varphi(0, t \pm \gamma(t))|^2 |1 \pm \gamma'(t)| dt \leq \int_{\pm\gamma(0)}^{T \pm \gamma(T)} |\partial_x \varphi(0, s)|^2 ds.$$

On a finalement

$$\int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt \leq \varsigma_\gamma \int_{-\gamma(0)}^{T+\gamma(T)} |\partial_x \varphi(0, s)|^2 ds \leq \varsigma_\gamma (T + \gamma(T) + \gamma(0) + 2) \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

La dernière inégalité provient de l'inégalité directe pour l'équation des ondes avec contrôle au bord (voir Prop. 18). Le cas  $1 - |\gamma'| \leq 0$  sur  $[0, T]$  se traite de façon analogue. Enfin, le cas général se démontre en introduisant une subdivision de  $[0, T]$  adaptée au caractère  $C^1$  par morceaux de  $\gamma$  et au signe de  $1 - |\gamma'|$ .  $\square$

On pose  $H := H^{-1}(\Omega)$ ,  $\mathcal{U} := H^{-1}(0, T)$  et  $\mathcal{U}' := H_0^1(0, T)$ . On considère les opérateurs

$$A_0 : \begin{array}{ccc} H_0^1(\Omega) & \rightarrow & H^{-1}(\Omega) \\ y & \mapsto & -\Delta y \end{array} \quad \text{et} \quad B_0 : \begin{array}{ccc} H^{-1}(0, T) & \rightarrow & L^2(0, T; H^{-2}(\Omega)) \\ u & \mapsto & A_0 Du \end{array},$$

où  $D : H^{-1}(0, T) \rightarrow L^2(Q)$  est l'opérateur défini par

$$\langle Du, \zeta \rangle_{L^2(Q)} = \langle u, \delta_\gamma^{\text{op}}(A_0^{-1}\zeta)\eta \rangle_{\mathcal{U}, \mathcal{U}'}, \quad \forall \zeta \in L^2(Q).$$

En reprenant les notations des Sect. 1.1.4 et 1.1.3, on a

$$Z_1 = H_0^1(\Omega) \times L^2(\Omega), \quad Z = L^2(\Omega) \times H^{-1}(\Omega), \quad Z_{-1} = H^{-1}(\Omega) \times H^{-2}(\Omega),$$

ainsi que

$$\mathcal{Z}_1 = L^2(0, T; Z_1), \quad \mathcal{Z} = L^2(0, T; Z), \quad \mathcal{Z}_{-1} = L^2(0, T; Z_{-1}).$$

On rappelle que les opérateurs  $A : Z_1 \rightarrow Z$  et  $B \in \mathcal{L}(\mathcal{U}, \mathcal{Z}_{-1})$  apparaissant dans (1.4) sont donnés par

$$A = \begin{pmatrix} 0 & \text{Id} \\ -A_0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 \\ B_0 \end{pmatrix}.$$

On rappelle également que  $v(\zeta^0) = -B^*\zeta(\cdot; \zeta^0)$ , où  $\zeta(t; \zeta^0)$  est la solution de (1.9) avec  $\zeta^0$  comme donnée initiale. On constate ici que

$$B^* : \begin{array}{ccc} \mathcal{Z}_1 & \rightarrow & \mathcal{U}' \\ (\zeta_1, \zeta_2) & \mapsto & \delta_\gamma^{\text{op}}(A_0^{-1}\zeta_2)\eta \end{array}.$$

En effet, pour  $u \in \mathcal{U}$  et  $\zeta \in \mathcal{Z}_1$ , on a

$$\begin{aligned} \langle Bu, \zeta \rangle_{\mathcal{Z}_{-1}, \mathcal{Z}_1} &= -\langle A^{-1}Bu, A\zeta \rangle_{\mathcal{Z}} = \langle A_0^{-1}B_0u, \zeta_2 \rangle_{L^2(Q)} \\ &= \langle Du, \zeta_2 \rangle_{L^2(Q)} = \langle u, \delta_\gamma^{\text{op}}(A_0^{-1}\zeta_2)\eta \rangle_{\mathcal{U}, \mathcal{U}'} = \langle u, B^*\zeta \rangle_{\mathcal{U}, \mathcal{U}'}. \end{aligned}$$

*Remarque 30.* On énonce une relation entre la solution  $\varphi$  de (1.34) et la solution  $\zeta$  de (1.9).

▷ Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $\zeta^0 = A(\varphi^0, \varphi^1)$ , on a  $\zeta = (\partial_t \varphi, -A_0 \varphi)$  et  $v(\zeta^0) = \delta_\gamma^{\text{op}}(\varphi)\eta$ .

De même, on peut écrire la relation de dualité suivante.

▷ Soient  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  et  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . En prenant  $z^0 = (y^0, y^1)$  et  $\zeta^0 = A(\varphi^0, \varphi^1)$ , on a

$$\langle z^0, \zeta^0 \rangle_{\mathcal{Z}} = \langle y^0, \varphi^1 \rangle_{L^2} - \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1}.$$

**Proposition 25** (Inégalité d'observabilité). *Par dualité entre contrôlabilité et observabilité, l'assertion suivante est vérifiée.*

▷ Le système (1.33) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq C \|\delta_\gamma^{\text{op}}(\varphi)\eta\|_{H_0^1(0, T)}^2, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega), \quad (1.35)$$

où  $\varphi$  est la solution de (1.34) associée à  $(\varphi^0, \varphi^1)$ .

**Proposition 26** (Contrôle de norme minimale). *En appliquant la méthode HUM, l'assertion suivante est vérifiée.*

▷ On suppose que l'inégalité d'observabilité (1.35) est vérifiée. Pour  $(y^0, y^1) \in L^2(\Omega) \times H^{-1}(\Omega)$ , le contrôle de norme  $H^{-1}(0, T)$  minimale pour le système (1.33) est donné par  $\hat{u} = -\frac{d^2}{dt^2}(\delta_\gamma^{\text{op}}(\hat{\varphi})\eta)$ , avec  $\hat{\varphi}$  la solution de (1.34) associée au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) := \frac{1}{2} \|\delta_\gamma^{\text{op}}(\varphi)\eta\|_{H_0^1(0, T)}^2 - \langle y^0, \varphi^1 \rangle_{L^2} + \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1}, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.34) associée à  $(\varphi^0, \varphi^1)$ .

**Proposition 27** (Inégalité d'observabilité). *[ $\gamma$ , Prop. 4.1] On suppose qu'il existe  $c_1, c_2 > 0$  et des intervalles ouverts  $(I_k)_{k \in \llbracket 1, K \rrbracket} \subset [0, T]$  tels que pour tout  $k \in \llbracket 1, K \rrbracket$ ,*

$$\gamma \in C^1(I_k), \quad c_1 \leq |\gamma'| \leq c_2 \text{ sur } I_k \quad \text{et} \quad 1 - |\gamma'| \text{ ne change pas de signe sur } I_k.$$

Pour  $k \in \llbracket 1, K \rrbracket$ , on pose

$$U_k := \begin{cases} I_k - \gamma(I_k) & \text{si } \gamma \text{ décroissante sur } I_k, \\ I_k + \gamma(I_k) & \text{si } \gamma \text{ croissante sur } I_k, \end{cases} \quad \text{et} \quad V_k := \begin{cases} I_k + \gamma(I_k) & \text{si } \gamma \text{ décroissante sur } I_k, \\ I_k - \gamma(I_k) & \text{si } \gamma \text{ croissante sur } I_k. \end{cases}$$

On suppose enfin que les  $V_k$  sont disjoints et qu'il existe

▷ un intervalle  $\tilde{U}$  tel que  $|\tilde{U}| > 2$  et  $\tilde{U} \subset \overline{\bigcup_{k=1}^K U_k}$  ;

▷ un intervalle  $\tilde{V}$  tel que  $|\tilde{V}| \leq |\tilde{U}|$  et  $\bigcup_{k=1}^K V_k \subset \tilde{V}$ .

Avec les hypothèses ci-dessus, il existe  $C_{\gamma,T} > 0$  tel que

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq C_{\gamma,T} \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution de (1.34) associée à  $(\varphi^0, \varphi^1)$ .

*Démonstration.* La preuve se décompose en trois étapes. On commence par montrer qu'il existe  $C > 0$  et  $r < 1$  tels que pour tout  $k \in \llbracket 1, K \rrbracket$ ,

$$\int_{U_k} |\partial_x \varphi(0, t)|^2 dt - r \int_{V_k} |\partial_x \varphi(0, t)|^2 dt \leq C \int_{I_k} \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt.$$

Il est à noter que l'on a  $r \rightarrow 1$  si  $c_1 \rightarrow 0$  ou  $c_2 \rightarrow +\infty$ . On montre ensuite que

$$(1-r) \int_0^2 |\partial_x \varphi(0, t)|^2 dt \leq C \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt.$$

On conclut en utilisant l'inégalité d'observabilité pour l'équation des ondes avec contrôle au bord (voir Prop. 22). La constante  $C_{\gamma,T}$  de la proposition étant de la forme  $C/(1-r)$ , on remarque que  $C_{\gamma,T} \rightarrow +\infty$  si  $c_1 \rightarrow 0$  ou  $c_2 \rightarrow +\infty$ .  $\square$

### 1.1.9 Quelques méthodes d'approximation du contrôle

Dans cette section, on présente plusieurs méthodes classiques permettant d'approcher le contrôle à zéro de norme minimale pour un système hyperbolique d'ordre deux de la forme (1.14). Dans la suite, on suppose que l'inégalité d'observabilité (1.11) est vérifiée. Par ailleurs, on rappelle que  $v(\zeta^0) := -B^* \zeta(\cdot; \zeta^0)$ , où  $\zeta(t; \zeta^0)$  est la solution de (1.9) associée à la donnée initiale  $\zeta^0 \in Z$ . On suppose enfin que l'opérateur de contrôle  $B$  dans (1.4) est admissible (voir Déf. 11) pour le  $C_0$ -groupe  $\mathbb{T}(t)$  d'opérateurs unitaires généré par l'opérateur d'évolution  $A$ .

#### Minimisation directe de la fonctionnelle conjuguée

On a vu à la Sect. 1.1.3 que le contrôle à zéro de norme  $\mathcal{U}$  minimale pour le système d'évolution (1.4) est donné par

$$\hat{u} = v(\hat{\zeta}^0),$$

où  $\hat{\zeta}^0 \in Z$  est l'unique point de minimum de la fonctionnelle conjuguée

$$\mathcal{J}^*(\zeta^0) := \frac{1}{2} a(\zeta^0, \zeta^0) - \ell(\zeta^0), \quad \forall \zeta^0 \in Z.$$

Le problème de minimisation de  $\mathcal{J}^*$  est sans contrainte. En outre, la forme bilinéaire  $a$  est symétrique définie positive. On peut donc approcher le point de minimum de  $\mathcal{J}^*$  de manière directe par la méthode du gradient conjugué (cf. Algo. 1).

*Remarque 31.* En pratique, la fonctionnelle  $\mathcal{J}^*$  doit être discrétisée avant d'être minimisée. Cependant, cette discrétisation mène généralement à un problème mal posé. Ceci signifie que le contrôle discret  $\hat{u}_h$  obtenu en minimisant la fonctionnelle discrétisée  $\mathcal{J}_h^*$  ne converge pas forcément vers le contrôle  $\hat{u}$  du problème continu. Ce phénomène est du à l'émergence d'oscillations parasites quand le paramètre de discrétisation  $h$  tend vers 0. Pour une étude précise de ce phénomène, on renvoie à [18] ou à [22, 36, 13]. Pour pallier à ce problème, de nombreux remèdes ont été proposés. On peut par exemple utiliser

- ▷ une régularisation de Tikhonov (cf. [18]) ;
- ▷ une méthode multigrille (cf. [1, 16]) ;
- ▷ une méthode éléments finis mixte (cf. [17, 11, 9]) ;



---

**Algorithm 1** Minimisation de  $\mathcal{J}^*$  par la méthode du gradient conjugué

---

- 1) *Initialisation.* On fixe  $\varepsilon > 0$  et on se donne  $\zeta_0 \in Z$ .  
On calcule  $g_0 \in Z$  solution de

$$\langle g_0, \bar{\zeta} \rangle_Z = a(\zeta_0, \bar{\zeta}) - \ell(\bar{\zeta}), \quad \forall \bar{\zeta} \in Z.$$

On pose  $p_0 = g_0$ .

- 2) *Itération.* Pour  $n \geq 0$ , on pose

$$\alpha_n = \frac{\|g_n\|_Z^2}{a(p_n, p_n)}.$$

On met à jour

$$\zeta_{n+1} = \zeta_n - \alpha_n p_n.$$

On calcule  $g_{n+1} \in Z$  solution de

$$\langle g_{n+1}, \bar{\zeta} \rangle_Z = \langle g_n, \bar{\zeta} \rangle_Z - \alpha_n a(p_n, \bar{\zeta}), \quad \forall \bar{\zeta} \in Z.$$

- 3) *Test de convergence.* Si  $\|g_{n+1}\|_Z \leq \varepsilon \|g_0\|_Z$ , l'algorithme s'arrête et on prend  $\zeta_{n+1}$  comme approximation du point de minimum de  $\mathcal{J}^*$ . Sinon, on pose

$$\beta_n = \frac{\|g_{n+1}\|_Z^2}{\|g_n\|_Z^2}.$$

On met à jour

$$p_{n+1} = g_{n+1} + \beta_n p_n.$$

Effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).

---

- ▷ un schéma différences finies implicite (cf. [27, 2]);
- ▷ une méthode éléments finis stabilisée (cf. [6]).

Pour illustrer l'Algo. 1, on le met en œuvre dans le cas de l'équation des ondes linéaire 2D avec contrôle au bord. On renvoie à [18] pour plus de détails. On reprend ici les notations de la Sect. 1.1.7 avec  $d = 2$ . Pour tout  $\varphi^0 = (\varphi^0, \varphi^1) \in V := H_0^1(\Omega) \times L^2(\Omega)$ , la fonctionnelle conjuguée s'écrit

$$\mathcal{J}^*(\varphi^0) := \frac{1}{2} \int_{\Sigma} |\partial_\nu \varphi|^2 - \langle y^0, \varphi^1 \rangle_{L^2} + \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1},$$

où  $\varphi$  est la solution du système adjoint

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = 0 & \text{dans } Q, \\ \varphi = 0 & \text{sur } \Sigma, \end{cases} \quad (1.36a)$$

$$\varphi = 0 \quad \text{sur } \Sigma, \quad (1.36b)$$

$$(\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) \quad \text{dans } \Omega. \quad (1.36c)$$

On montre aisément que la dérivée de  $\mathcal{J}^*$  au point  $\varphi^0 \in V$  dans la direction  $\bar{\varphi}^0 \in V$  est donnée par

$$d\mathcal{J}^*(\varphi^0; \bar{\varphi}^0) = \langle \psi(\cdot, 0) - y^0, \bar{\varphi}^1 \rangle_{L^2} + \langle y^1 - \partial_t \psi(\cdot, 0), \bar{\varphi}^0 \rangle_{H^{-1}, H_0^1},$$

où  $\psi$  est la solution du système rétrograde

$$\begin{cases} \partial_t^2 \psi - \Delta \psi = 0 & \text{dans } Q, \end{cases} \quad (1.37a)$$

$$\begin{cases} \psi = -\partial_\nu \varphi & \text{sur } \Sigma, \end{cases} \quad (1.37b)$$

$$\begin{cases} (\psi, \partial_t \psi)(\cdot, T) = (0, 0) & \text{dans } \Omega, \end{cases} \quad (1.37c)$$

avec  $\varphi$  la solution de (1.36) associée à  $\varphi^0$ .

**Discrétisation** On considère un maillage  $\mathcal{T}_h$  de  $\Omega \subset \mathbb{R}^2$ , de sorte que  $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} K$ . Le paramètre  $h > 0$  correspond à la taille de la plus grande maille de  $\mathcal{T}_h$ . On introduit ensuite deux sous-maillages  $\mathcal{T}_h^i$

et  $\mathcal{T}_h^b$  tels que  $\mathcal{T}_h = \mathcal{T}_h^i \sqcup \mathcal{T}_h^b$ . On construit  $\mathcal{T}_h^b$  comme l'ensemble des éléments  $K \in \mathcal{T}_h$  dont au moins un sommet se trouve sur le bord  $\Gamma := \partial\Omega$ , puis on pose  $\mathcal{T}_h^i := \mathcal{T}_h \setminus \mathcal{T}_h^b$ . Pour  $K \in \mathcal{T}_h$ , on note  $\mathbb{P}_1(K)$  l'espace des polynômes de degré un sur l'élément  $K$ . On définit enfin les espaces

$$\begin{aligned} V_h &:= \left\{ \phi \in C(\overline{\Omega}); \quad \phi|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h \right\}, \\ V_h^i &:= \left\{ \phi^i \in C(\overline{\Omega}); \quad \phi^i|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h^i, \quad \phi^i|_K \equiv 0, \forall K \in \mathcal{T}_h^b \right\}, \\ V_h^b &:= \left\{ \phi^b \in C(\overline{\Omega}); \quad \phi^b|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h^b, \quad \phi^b|_K \equiv 0, \forall K \in \mathcal{T}_h^i \right\}. \end{aligned}$$

On a alors  $V_h = V_h^i \oplus V_h^b$ . Concernant la discrétisation en temps, pour  $N \in \mathbb{N}^*$ , on pose  $\Delta t := T/N$  et  $t_n := n\Delta t$ ,  $n \in \llbracket 0, N \rrbracket$ . Dans la suite, pour tout  $(\phi_n)_{n \in \llbracket 0, N \rrbracket} \subset V_h$ , on notera

$$\delta_t^2 \phi_n := \frac{\phi_{n+1} - 2\phi_n + \phi_{n-1}}{\Delta t^2}.$$

*Remarque 32.* Le pas de temps  $\Delta t$  doit être pris suffisamment petit devant le paramètre  $h$ . En effet, ceci est nécessaire pour garantir la condition CFL liée aux schémas différences finis introduits plus bas.

**Approximation du système adjoint** Soit  $\varphi^0 = (\varphi^0, \varphi^1) \in V_h^i \times V_h^i$ . Pour tout  $n \in \llbracket 0, N \rrbracket$ , on note  $\varphi_n \in V_h^i$  l'approximation de  $\varphi(\cdot, t_n)$ , où  $\varphi$  est solution de (1.36). Cette approximation est calculée en utilisant le schéma numérique suivant. Pour  $n \in \llbracket 0, N \rrbracket$ , on définit  $\varphi_{n+1} \in V_h^i$  comme la solution du problème variationnel

$$\int_{\Omega} \delta_t^2 \varphi_n \phi^i + \int_{\Omega} \nabla \varphi_n \cdot \nabla \phi^i = 0, \quad \forall \phi^i \in V_h^i. \quad (1.38)$$

En outre, on initialise le schéma avec les relations

$$\varphi_0 = \varphi^0 \quad \text{et} \quad \frac{\varphi_1 - \varphi_{-1}}{2\Delta t} = \varphi^1.$$

**Approximation de la dérivée normale** Soit  $(\varphi_n)_{n \in \llbracket 0, N \rrbracket} \subset V_h^i$ . Pour tout  $n \in \llbracket 0, N \rrbracket$ , on note  $\delta_\nu \varphi_n \in V_h^b$  l'approximation de  $\partial_\nu \varphi(\cdot, t_n)$ , où  $\varphi$  est solution de (1.36). Cette approximation est calculée en utilisant le schéma numérique suivant. Pour  $n \in \llbracket 0, N \rrbracket$ , on définit  $\delta_\nu \varphi_n \in V_h^b$  comme la solution du problème variationnel

$$\int_{\Gamma} \delta_\nu \varphi_n \phi^b = \int_{\Omega} \nabla \varphi_n \cdot \nabla \phi^b + \int_{\Omega} \delta_t^2 \varphi_n \phi^b, \quad \forall \phi^b \in V_h^b. \quad (1.39)$$

**Approximation du système rétrograde** Soit  $(\delta_\nu \varphi_n)_{n \in \llbracket 0, N \rrbracket} \subset V_h^b$ . Pour tout  $n \in \llbracket 0, N \rrbracket$ , on note

$$\psi_n = \psi_n^i + \psi_n^b \in V_h = V_h^i \oplus V_h^b$$

l'approximation de  $\psi(\cdot, t_n)$ , où  $\psi$  est solution de (1.37). Cette approximation est calculée en utilisant le schéma numérique suivant. Pour  $n \in \llbracket 0, N \rrbracket$ , on définit  $\psi_{n-1}^b \in V_h^b$  par  $\psi_{n-1}^b = -\delta_\nu \varphi_{n-1}$  et  $\psi_{n-1}^i \in V_h^i$  comme la solution du problème variationnel

$$\int_{\Omega} \delta_t^2 \psi_n \phi^i + \int_{\Omega} \nabla \psi_n \cdot \nabla \phi^i = 0, \quad \forall \phi^i \in V_h^i. \quad (1.40)$$

En outre, on initialise le schéma avec les relations

$$\psi_N^i = 0 \quad \text{et} \quad \frac{\psi_{N+1}^i - \psi_{N-1}^i}{2\Delta t} = 0.$$

**Approximation de  $\mathcal{J}^*$**  Dans la suite, on note  $\theta_n = 1$  pour  $n \in \llbracket 1, N-1 \rrbracket$  et  $\theta_0 = \theta_N = 1/2$ . Pour tout  $\varphi^0 = (\varphi^0, \varphi^1) \in V_h^i \times V_h^i$ , on introduit la fonctionnelle discrétisée

$$\mathcal{J}_h^*(\varphi^0) := \frac{\Delta t}{2} \sum_{n=0}^N \theta_n \int_{\Gamma} |\delta_\nu \varphi_n|^2 - \langle y^0, \varphi^1 \rangle_{L^2} + \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1},$$

où  $(\delta_\nu \varphi_n)_{n \in \llbracket 0, N \rrbracket}$  est calculé via (1.39), à partir de la solution  $(\varphi_n)_{n \in \llbracket 0, N \rrbracket}$  de (1.38). On utilise ici le produit de dualité  $\langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1} := \langle (-\Delta)_h^{-1} y^1, \varphi^0 \rangle_{H_0^1}$ , où  $(-\Delta)_h^{-1} y^1 \in V_h^i$  est solution du problème variationnel

$$\int_{\Omega} \nabla(-\Delta)_h^{-1} y^1 \cdot \nabla \phi^i = \int_{\Omega} y^1 \phi^i, \quad \forall \phi^i \in V_h^i.$$

On vérifie alors que la dérivée de  $\mathcal{J}_h^*$  au point  $\varphi^0 \in V_h^i \times V_h^i$  dans la direction  $\bar{\varphi}^0 \in V_h^i \times V_h^i$  est donnée par

$$d\mathcal{J}_h^*(\varphi^0; \bar{\varphi}^0) = \langle \psi^0 - y^0, \bar{\varphi}^1 \rangle_{L^2} + \langle y^1 - \psi^1, \bar{\varphi}^0 \rangle_{H^{-1}, H_0^1},$$

avec

$$\psi^0 := \psi_0^i \in V_h^i \quad \text{et} \quad \psi^1 := \frac{\psi_1^i - \psi_{-1}^i}{2\Delta t} \in V_h^i, \quad (1.41)$$

où  $(\psi_n)_{n \in \llbracket 0, N \rrbracket}$  est calculé via (1.40), à partir des solutions  $(\varphi_n)_{n \in \llbracket 0, N \rrbracket}$  et  $(\delta_\nu \varphi_n)_{n \in \llbracket 0, N \rrbracket}$  de (1.38)-(1.39).

La preuve de l'expression de la dérivée repose essentiellement sur la formule d'intégration par parties discrète

$$\begin{aligned} \Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \delta_t^2 \phi_n \bar{\phi}_n &= \Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \phi_n \delta_t^2 \bar{\phi}_n + \int_{\Omega} \frac{\phi_{N+1} - \phi_{N-1}}{2\Delta t} \bar{\phi}_N - \int_{\Omega} \phi_N \frac{\bar{\phi}_{N+1} - \bar{\phi}_{N-1}}{2\Delta t} \\ &\quad - \int_{\Omega} \frac{\phi_1 - \phi_{-1}}{2\Delta t} \bar{\phi}_0 + \int_{\Omega} \phi_0 \frac{\bar{\phi}_1 - \bar{\phi}_{-1}}{2\Delta t}. \end{aligned}$$

Plus précisément, avec cette formule et les schémas (1.38) à (1.40), il vient

$$\begin{aligned} \Delta t \sum_{n=0}^N \theta_n \int_{\Gamma} \delta_\nu \varphi_n \delta_\nu \bar{\varphi}_n &= \Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \left( \nabla \delta_\nu \varphi_n \cdot \nabla \bar{\varphi}_n + \delta_\nu \varphi_n \delta_t^2 \bar{\varphi}_n \right) \\ &= -\Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \left( \nabla \psi_n^b \cdot \nabla \bar{\varphi}_n + \psi_n^b \delta_t^2 \bar{\varphi}_n \right) \\ &= \Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \left( -\nabla \psi_n \cdot \nabla \bar{\varphi}_n + \nabla \psi_n^i \cdot \nabla \bar{\varphi}_n - \psi_n^b \delta_t^2 \bar{\varphi}_n \right) \\ &= \Delta t \sum_{n=0}^N \theta_n \int_{\Omega} \left( \delta_t^2 \psi_n \bar{\varphi}_n - \psi_n \delta_t^2 \bar{\varphi}_n \right) \\ &= \int_{\Omega} \psi^0 \bar{\varphi}^1 - \int_{\Omega} \psi^1 \bar{\varphi}^0 = \langle \psi^0, \bar{\varphi}^1 \rangle_{L^2} - \langle \psi^1, \bar{\varphi}^0 \rangle_{H^{-1}, H_0^1}. \end{aligned}$$

**Algorithme du gradient conjugué** Pour minimiser la fonctionnelle discrétisée  $\mathcal{J}_h^*$ , on utilise l'algorithme suivant.

*Initialisation.* On fixe  $\varepsilon > 0$  et on se donne  $(\varphi_0^0, \varphi_0^1) \in V_h^i \times V_h^i$ . On calcule alors successivement

- ▷ la solution  $(\varphi_n)_n$  de (1.38) associée à  $(\varphi_0^0, \varphi_0^1)$ ;
- ▷ la solution  $(\delta_\nu \varphi_n)_n$  de (1.39) associée à  $(\varphi_n)_n$ ;
- ▷ la solution  $(\psi_n)_n$  de (1.40) associée à  $(\delta_\nu \varphi_n)_n$ ;
- ▷ le couple  $(\psi^0, \psi^1)$  calculé à partir de  $(\psi_n)_n$  via (1.41).

On pose ensuite

$$g_0^0 = (-\Delta)_h^{-1} (y^1 - \psi^1) \in V_h^i \quad \text{et} \quad g_0^1 = \psi^0 - y^0 \in V_h^i,$$

ainsi que  $(p_0^0, p_0^1) = (g_0^0, g_0^1)$ .

*Itération.* Pour  $k \geq 0$ , on calcule successivement

- ▷ la solution  $(\varphi_n)_n$  de (1.38) associée à  $(p_k^0, p_k^1)$ ;
- ▷ la solution  $(\delta_\nu \varphi_n)_n$  de (1.39) associée à  $(\varphi_n)_n$ ;
- ▷ la solution  $(\psi_n)_n$  de (1.40) associée à  $(\delta_\nu \varphi_n)_n$ ;
- ▷ le couple  $(\psi^0, \psi^1)$  calculé à partir de  $(\psi_n)_n$  via (1.41).

On pose ensuite

$$\widehat{p}_k^0 = -(-\Delta)_h^{-1} \psi^1 \in V_h^i \quad \text{et} \quad \widehat{p}_k^1 = \psi^0 \in V_h^i.$$

On calcule alors

$$\alpha_k = \frac{\|g_k^0\|_{H_0^1}^2 + \|g_k^1\|_{L^2}^2}{\langle \widehat{p}_k^0, p_k^0 \rangle_{H_0^1} + \langle \widehat{p}_k^1, p_k^1 \rangle_{L^2}},$$

puis on met à jour

$$(\varphi_{k+1}^0, \varphi_{k+1}^1) = (\varphi_k^0, \varphi_k^1) - \alpha_k (p_k^0, p_k^1) \quad \text{et} \quad (g_{k+1}^0, g_{k+1}^1) = (g_k^0, g_k^1) - \alpha_k (\widehat{p}_k^0, \widehat{p}_k^1).$$

*Test de convergence.* Si  $\|(g_{k+1}^0, g_{k+1}^1)\|_{H_0^1 \times L^2} \leq \varepsilon \|(g_0^0, g_0^1)\|_{H_0^1 \times L^2}$ , l'algorithme s'arrête et on prend  $(\varphi_{k+1}^0, \varphi_{k+1}^1)$  comme approximation du point de minimum de  $\mathcal{J}_h^*$ . Sinon, on pose

$$\beta_k = \frac{\|g_{k+1}^0\|_{H_0^1}^2 + \|g_{k+1}^1\|_{L^2}^2}{\|g_k^0\|_{H_0^1}^2 + \|g_k^1\|_{L^2}^2},$$

et on met à jour

$$(p_{k+1}^0, p_{k+1}^1) = (g_{k+1}^0, g_{k+1}^1) + \beta_k (p_k^0, p_k^1).$$

On effectue enfin  $k \leftarrow k + 1$  et on revient à l'étape *Itération*.

### Formulation mixte espace-temps

On a vu à la Sect. 1.1.3 que le contrôle à zéro de norme  $\mathcal{U}$  minimale pour le système d'évolution (1.4) est donné par

$$\widehat{u} = v(\widehat{\zeta}^0),$$

où  $\widehat{\zeta}^0 \in Z$  est l'unique point de minimum de la fonctionnelle conjuguée

$$\mathcal{J}^*(\zeta^0) := \frac{1}{2} a(\zeta^0, \zeta^0) - \ell(\zeta^0), \quad \forall \zeta^0 \in Z.$$

Cependant, la minimisation directe de  $\mathcal{J}^*$  mène généralement à un problème mal posé. Pour remédier à cela, l'idée est ici de remplacer le problème de minimisation sans contrainte  $\inf_{\zeta^0 \in Z} \mathcal{J}^*(\zeta^0)$  portant sur la donnée initiale  $\zeta^0$  par un problème de minimisation avec contrainte  $\inf_{\zeta \in \mathfrak{Z}_0} \mathcal{J}^*(\zeta)$  portant sur la trajectoire  $\zeta(t)$  du système adjoint (1.9).

**Définition 27** (Espace des trajectoires). On considère le système

$$\begin{cases} \frac{d}{dt} \zeta(t) = A\zeta(t) + f(t), & t \in (0, T), \\ \zeta(0) = \zeta^0. \end{cases} \quad (1.42a)$$

$$(1.42b)$$

On pose  $\mathcal{Z} := L^2(0, T; Z)$ . Pour  $\zeta^0 \in Z$  et  $f \in \mathcal{Z}$ , la solution de (1.42) est donnée par

$$\zeta(t) = \mathbb{T}(t)\zeta^0 + \int_0^t \mathbb{T}(t-s)f(s) ds, \quad t \in [0, T].$$

On note  $\zeta(t; \zeta^0, f)$  la solution de (1.42) avec  $\zeta^0$  comme donnée initiale et  $f$  comme second membre. On introduit ensuite les espaces

$$\mathfrak{Z} := \left\{ \zeta(\cdot; \zeta^0, f); \quad \zeta^0 \in Z, f \in \mathcal{Z} \right\} \quad \text{et} \quad \mathfrak{Z}_0 := \left\{ \zeta(\cdot; \zeta^0, 0); \quad \zeta^0 \in Z \right\}.$$

Pour  $\eta > 0$  fixé, on munit  $\mathfrak{Z}$  du produit scalaire

$$\langle \zeta, \bar{\zeta} \rangle_{\mathfrak{Z}} := \langle v(\zeta), v(\bar{\zeta}) \rangle_{\mathcal{U}} + \eta \langle f, \bar{f} \rangle_{\mathcal{Z}}, \quad \forall \zeta, \bar{\zeta} \in \mathfrak{Z},$$

où  $\zeta = \zeta(\cdot; \zeta^0, f)$ ,  $\bar{\zeta} = \bar{\zeta}(\cdot; \bar{\zeta}^0, \bar{f})$  et  $v(\zeta) = -B^* \zeta$ . Enfin, muni de la norme  $\|\cdot\|_{\mathfrak{Z}}$  associée à ce produit scalaire,  $\mathfrak{Z}$  est un espace de Hilbert.

*Remarque 33.* La définition de  $\langle \cdot, \cdot \rangle_{\mathfrak{Z}}$  ne dépend pas du choix de  $f$  et  $\bar{f}$ . En effet, pour tout  $\zeta \in \mathfrak{Z}$ , il existe un unique couple  $(\zeta^0, f) \in Z \times \mathcal{Z}$  tel que  $\zeta = \zeta(\cdot; \zeta^0, f)$ . Soient  $\zeta^0 \in Z$  et  $f \in C^1([0, T]; Z)$  tels que  $\zeta(\cdot; \zeta^0, f) = 0$ . En évaluant en  $t = 0$ , il vient  $\zeta^0 = 0$  dans  $Z$ . En outre, pour tout  $t \in [0, T]$ , on trouve  $f(t) = \frac{d}{dt}\zeta(t) - A\zeta(t) = 0$  dans  $Z$ . On conclut par densité.

*Remarque 34.* D'une part, le fait que  $\langle \cdot, \cdot \rangle_{\mathfrak{Z}}$  soit un produit scalaire sur  $\mathfrak{Z}$  est une conséquence directe de l'inégalité d'observabilité généralisée (1.43). D'autre part, le fait que  $\mathfrak{Z}$  soit un espace de Hilbert provient de la continuité de la solution  $\zeta(t; \zeta^0, f)$  de (1.42) par rapport à la donnée initiale  $\zeta^0 \in Z$  et au second membre  $f \in \mathcal{Z}$ .

**Lemme 5** (Principe de Duhamel). *On suppose que l'opérateur de contrôle  $B$  dans (1.4) est admissible (voir Déf. 11). Soit  $f \in \mathcal{Z}$  et soit  $\zeta$  la solution du système*

$$\begin{cases} \frac{d}{dt}\zeta(t) = A\zeta(t) + f(t), & t \in (0, T), \\ \zeta(0) = 0. \end{cases}$$

Il existe alors  $C > 0$  tel que

$$\|v(\zeta)\|_{\mathcal{U}} \leq C\|f\|_{\mathcal{Z}}.$$

*Démonstration.* Soit  $f \in \mathcal{Z}$ . La solution  $\zeta$  du système est donnée par

$$\zeta(s) = \int_0^s \mathbb{T}(s-t)^* f(t) dt, \quad s \in [0, T].$$

En utilisant l'admissibilité de  $B$ , pour tout  $t \in [0, T]$ , on obtient

$$\begin{aligned} \int_t^T \|B^*\mathbb{T}(T-s)^* f(t)\|_{\mathcal{U}}^2 ds &= \int_t^T \|B^*\mathbb{T}(T-s)^*\mathbb{T}(T-t)f(t)\|_{\mathcal{U}}^2 ds \\ &= \int_0^{T-t} \|B^*\mathbb{T}(s)^*\mathbb{T}(T-t)f(t)\|_{\mathcal{U}}^2 ds \\ &\leq C_T \|\mathbb{T}(T-t)f(t)\|_{\mathcal{Z}}^2 = C_T \|f(t)\|_{\mathcal{Z}}^2. \end{aligned}$$

Par suite, en utilisant l'inégalité intégrale de Minkowski, il vient

$$\begin{aligned} \|v(\zeta)\|_{\mathcal{U}} &= \|B^*\zeta\|_{\mathcal{U}} = \left( \int_0^T \left\| \int_0^s B^*\mathbb{T}(s-t)^* f(t) dt \right\|_{\mathcal{U}}^2 ds \right)^{\frac{1}{2}} \\ &\leq \int_0^T \left( \int_t^T \|B^*\mathbb{T}(s-t)^* f(t)\|_{\mathcal{U}}^2 ds \right)^{\frac{1}{2}} dt \\ &\leq \sqrt{TC_T} \|f\|_{\mathcal{Z}}. \end{aligned}$$

□

**Proposition 28** (Inégalité d'observabilité généralisée). *On suppose que l'inégalité d'observabilité (1.11) est vérifiée. Il existe alors  $C > 0$  tel que*

$$\|\zeta^0\|_{\mathcal{Z}}^2 \leq C \left( \|v(\zeta)\|_{\mathcal{U}}^2 + \|f\|_{\mathcal{Z}}^2 \right), \quad \forall \zeta \in \mathfrak{Z}, \quad \zeta = \zeta(\cdot; \zeta^0, f). \quad (1.43)$$

*Démonstration.* On décompose  $\zeta = \zeta_1 + \zeta_2$ , où  $\zeta_1$  et  $\zeta_2$  vérifient respectivement

$$\begin{cases} \frac{d}{dt}\zeta_1(t) = A\zeta_1(t), & t \in (0, T), \\ \zeta_1(0) = \zeta^0, \end{cases} \quad \text{et} \quad \begin{cases} \frac{d}{dt}\zeta_2(t) = A\zeta_2(t) + f(t), & t \in (0, T), \\ \zeta_2(0) = 0. \end{cases}$$

Par suite, d'après l'inégalité d'observabilité (1.11) et le Lem. 5, il existe  $C_1, C_2 > 0$  tels que

$$\|\zeta^0\|_{\mathcal{Z}}^2 \leq C_1 \|v(\zeta_1)\|_{\mathcal{U}}^2 \quad \text{et} \quad \|v(\zeta_2)\|_{\mathcal{U}}^2 \leq C_2 \|f\|_{\mathcal{Z}}^2.$$

On en déduit que

$$\|\zeta^0\|_{\mathcal{Z}}^2 \leq C_1 \|v(\zeta_1)\|_{\mathcal{U}}^2 \leq 2C_1 \left( \|v(\zeta)\|_{\mathcal{U}}^2 + \|v(\zeta_2)\|_{\mathcal{U}}^2 \right) \leq 2C_1 \max(1, C_2) \left( \|v(\zeta)\|_{\mathcal{U}}^2 + \|f\|_{\mathcal{Z}}^2 \right).$$

□

**Définition 28** (Formulation mixte). On définit la fonctionnelle conjuguée

$$\tilde{\mathcal{J}}^*(\zeta) := \frac{1}{2}a(\zeta, \zeta) - \ell(\zeta), \quad \forall \zeta \in \mathfrak{Z},$$

ainsi que le lagrangien

$$\mathcal{L}(\zeta, \mu) := \frac{1}{2}a(\zeta, \zeta) + b(\zeta, \mu) - \ell(\zeta), \quad \forall (\zeta, \mu) \in \mathfrak{Z} \times \mathcal{Z},$$

où

$$\begin{aligned} a : \mathfrak{Z} \times \mathfrak{Z} &\rightarrow \mathbb{R}, & a(\zeta, \bar{\zeta}) &= \langle v(\zeta), v(\bar{\zeta}) \rangle_{\mathcal{U}}, \\ b : \mathfrak{Z} \times \mathcal{Z} &\rightarrow \mathbb{R}, & b(\zeta, \mu) &= \langle f, \mu \rangle_{\mathcal{Z}}, \quad \zeta = \zeta(\cdot; \zeta^0, f), \\ \ell : \mathfrak{Z} &\rightarrow \mathbb{R}, & \ell(\zeta) &= \langle z^0, \zeta^0 \rangle_{\mathcal{Z}}, \quad \zeta = \zeta(\cdot; \zeta^0, f). \end{aligned}$$

On considère alors la formulation mixte

$$\begin{cases} a(\zeta, \bar{\zeta}) + b(\bar{\zeta}, \mu) = \ell(\bar{\zeta}), & \forall \bar{\zeta} \in \mathfrak{Z}, \\ b(\zeta, \bar{\mu}) = 0, & \forall \bar{\mu} \in \mathcal{Z}. \end{cases} \quad \begin{array}{l} (1.44a) \\ (1.44b) \end{array}$$

*Remarque 35.* Au vu de la bijection entre la donnée initiale  $\zeta^0 \in \mathcal{Z}$  et la trajectoire  $\zeta = \zeta(\cdot; \zeta^0, 0) \in \mathfrak{Z}_0$ , il est clair que

$$\inf_{\zeta^0 \in \mathcal{Z}} \mathcal{J}^*(\zeta^0) = \inf_{\zeta \in \mathfrak{Z}_0} \tilde{\mathcal{J}}^*(\zeta).$$

Par ailleurs, la nouvelle variable  $\mu \in \mathcal{Z}$  joue ici le rôle de multiplicateur de Lagrange associé à la contrainte  $f = 0$  présente dans  $\mathfrak{Z}_0$ .

*Remarque 36.* En pratique, lors de la résolution numérique de la formulation mixte (1.44), il est souvent préférable d'augmenter le lagrangien. Ainsi, pour  $r > 0$  fixé, on considérera plutôt le lagrangien augmenté

$$\mathcal{L}_r(\zeta, \mu) := \frac{1}{2}a_r(\zeta, \zeta) + b(\zeta, \mu) - \ell(\zeta), \quad \forall (\zeta, \mu) \in \mathfrak{Z} \times \mathcal{Z},$$

avec

$$a_r(\zeta, \bar{\zeta}) := a(\zeta, \bar{\zeta}) + r \langle f, \bar{f} \rangle_{\mathcal{Z}}, \quad \forall \zeta, \bar{\zeta} \in \mathfrak{Z}, \quad \zeta = \zeta(\cdot; \zeta^0, f), \quad \bar{\zeta} = \bar{\zeta}(\cdot; \bar{\zeta}^0, \bar{f}).$$

**Proposition 29** (Formulation mixte). *Les propriétés suivantes sont vérifiées.*

- ▷ La formulation mixte (1.44) est bien posée.
- ▷ La solution  $(\hat{\zeta}, \hat{\mu}) \in \mathfrak{Z} \times \mathcal{Z}$  de (1.44) est l'unique point-selle du lagrangien  $\mathcal{L}$ .
- ▷ On a  $\hat{\zeta} \in \mathfrak{Z}_0$  et  $\hat{\zeta}$  est l'unique point de minimum de  $\tilde{\mathcal{J}}^*$  sur  $\mathfrak{Z}_0$ .
- ▷ Le contrôle de norme  $\mathcal{U}$  minimale pour (1.4) est donné par  $\hat{u} = v(\hat{\zeta})$ .
- ▷ La solution de (1.4) associée à  $z^0$  et  $\hat{u}$  est donnée par  $z(t; z^0, \hat{u}) = -\hat{\mu}(t)$ .

*Démonstration.* Pour montrer que la formulation mixte (1.44) est bien posée, on utilise les conditions suffisantes données dans [4, Th. 4.2.1]. La forme bilinéaire  $a$  est continue, symétrique et positive sur  $\mathfrak{Z} \times \mathfrak{Z}$ . La forme bilinéaire  $b$  est continue sur  $\mathfrak{Z} \times \mathcal{Z}$ . La forme linéaire  $\ell$  est continue sur  $\mathfrak{Z}$ . En effet, avec l'inégalité d'observabilité généralisée (1.43), on a

$$|\ell(\zeta)| \leq \|z^0\|_{\mathcal{Z}} \|\zeta^0\|_{\mathcal{Z}} \leq C \|z^0\|_{\mathcal{Z}} \|\zeta\|_{\mathfrak{Z}}, \quad \forall \zeta \in \mathfrak{Z}, \quad \zeta = \zeta(\cdot; \zeta^0, f).$$

La forme bilinéaire  $a$  est coercive sur le noyau

$$\mathcal{N}(b) = \left\{ \zeta \in \mathfrak{Z}; \quad b(\zeta, \mu) = 0, \forall \mu \in \mathcal{Z} \right\}$$

puisque  $\mathcal{N}(b) = \mathfrak{Z}_0$  et que  $a(\zeta, \zeta) = \|\zeta\|_{\mathfrak{Z}}^2$  pour tout  $\zeta \in \mathfrak{Z}_0$ . La forme bilinéaire  $b$  vérifie la condition *inf-sup* suivante. Il existe  $\delta > 0$  tel que

$$\inf_{\mu \in \mathcal{Z}} \sup_{\zeta \in \mathfrak{Z}} \frac{b(\zeta, \mu)}{\|\zeta\|_{\mathfrak{Z}} \|\mu\|_{\mathcal{Z}}} \geq \delta. \quad (1.45)$$

Soit  $\mu_0 \in \mathcal{Z}$ . On pose  $\zeta_0 \in \mathfrak{Z}$  tel que  $\zeta_0 = \zeta_0(\cdot; 0, \mu_0)$ . On a alors  $b(\zeta_0, \mu_0) = \|\mu_0\|_{\mathcal{Z}}^2$ . En outre, d'après le Lem. 5, il existe  $C > 0$  tel que

$$\|v(\zeta_0)\|_{\mathcal{U}} \leq C\|\mu_0\|_{\mathcal{Z}}.$$

Par suite, on a

$$\sup_{\zeta \in \mathfrak{Z}} \frac{b(\zeta, \mu_0)}{\|\zeta\|_{\mathfrak{Z}}\|\mu_0\|_{\mathcal{Z}}} \geq \frac{b(\zeta_0, \mu_0)}{\|\zeta_0\|_{\mathfrak{Z}}\|\mu_0\|_{\mathcal{Z}}} = \frac{\|\mu_0\|_{\mathcal{Z}}}{\|\zeta_0\|_{\mathfrak{Z}}} \geq \frac{1}{\sqrt{C^2 + \eta}}.$$

Le deuxième point est une conséquence de la symétrie et de la positivité de  $a$ .

Soit  $(\widehat{\zeta}, \widehat{\mu}) \in \mathfrak{Z} \times \mathcal{Z}$  la solution de la formulation mixte (1.44), avec  $\widehat{\zeta} = \widehat{\zeta}(\cdot; \widehat{\zeta}^0, \widehat{f})$ . D'après (1.44b), on a  $\widehat{f} = 0$  et donc  $\widehat{\zeta} \in \mathfrak{Z}_0$ . En outre, la condition (1.44a) implique  $a(\widehat{\zeta}, \zeta) = \ell(\zeta)$  pour tout  $\zeta \in \mathfrak{Z}_0$ . Par suite, on a

$$\widetilde{\mathcal{J}}^*(\widehat{\zeta}) = -\frac{1}{2}a(\widehat{\zeta}, \widehat{\zeta}) \leq \frac{1}{2}a(\zeta, \zeta) - a(\widehat{\zeta}, \zeta) = \widetilde{\mathcal{J}}^*(\zeta), \quad \forall \zeta \in \mathfrak{Z}_0.$$

En notant  $\widehat{u} = v(\widehat{\zeta})$ , d'après (1.44a), on a

$$-\langle \widehat{\mu}, f \rangle_{\mathcal{Z}} = -\langle z^0, \zeta^0 \rangle_{\mathcal{Z}} + \langle \widehat{u}, v(\zeta) \rangle_{\mathcal{U}}, \quad \forall \zeta \in \mathfrak{Z}, \quad \zeta = \zeta(\cdot; \zeta^0, f).$$

Avec le Lem. 6, on en déduit que  $-\widehat{\mu}$  est la solution de (1.4) avec  $z^0$  comme donnée initiale et  $\widehat{u}$  comme contrôle.  $\square$

**Lemme 6** (Solution par transposition). *Une fonction  $z \in \mathcal{Z}$  est la solution de (1.4) associée à  $z^0 \in \mathcal{Z}$  et  $u \in \mathcal{U}$  si et seulement si*

$$\langle z, f \rangle_{\mathcal{Z}} = -\langle z^0, \zeta(0) \rangle_{\mathcal{Z}} + \langle u, v(\zeta) \rangle_{\mathcal{U}}, \quad \forall f \in \mathcal{Z},$$

où  $\zeta$  est la solution du système rétrograde

$$\begin{cases} \frac{d}{dt}\zeta(t) = A\zeta(t) + f(t), & t \in (0, T), \\ \zeta(T) = 0. \end{cases}$$

*Démonstration.* Soit  $f \in \mathcal{Z}$ . La solution  $\zeta$  du système est donnée par

$$\zeta(s) = -\int_s^T \mathbb{T}(t-s)^* f(t) dt, \quad s \in [0, T].$$

Par ailleurs, si  $z$  est la solution de (1.4) associée à  $z^0 \in \mathcal{Z}$  et  $u \in \mathcal{U}$ , on a

$$\langle z(t), f(t) \rangle_{\mathcal{Z}} = \langle z^0, \mathbb{T}(t)^* f(t) \rangle_{\mathcal{Z}} + \int_0^t \langle u(s), B^* \mathbb{T}(t-s)^* f(t) \rangle_{\mathcal{U}} ds, \quad \forall t \in [0, T].$$

En intégrant sur  $t$  entre 0 et  $T$ , il vient

$$\begin{aligned} \langle z, f \rangle_{\mathcal{Z}} &= -\langle z^0, \zeta(0) \rangle_{\mathcal{Z}} + \int_0^T \int_0^t \langle u(s), B^* \mathbb{T}(t-s)^* f(t) \rangle_{\mathcal{U}} ds dt \\ &= -\langle z^0, \zeta(0) \rangle_{\mathcal{Z}} + \int_0^T \int_s^T \langle u(s), B^* \mathbb{T}(t-s)^* f(t) \rangle_{\mathcal{U}} dt ds \\ &= -\langle z^0, \zeta(0) \rangle_{\mathcal{Z}} + \langle u, v(\zeta) \rangle_{\mathcal{U}}. \end{aligned}$$

$\square$

Pour illustrer la formulation mixte (1.44), on la met en œuvre dans le cas de l'équation des ondes linéaire 1D avec contrôle interne. On renvoie à [8] pour plus de détails. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . Pour tout  $\varphi^0 = (\varphi^0, \varphi^1) \in V := H_0^1(\Omega) \times L^2(\Omega)$ , la fonctionnelle conjuguée s'écrit

$$\mathcal{J}^*(\varphi^0) := \frac{1}{2} \iint_Q |\partial_t \varphi|^2 - \langle y^0, \varphi^0 \rangle_{H_0^1} - \langle y^1, \varphi^1 \rangle_{L^2},$$

où  $\varphi$  est la solution du système adjoint

$$\begin{cases} \partial_t^2 \varphi - \Delta \varphi = 0 & \text{dans } Q, & (1.46a) \\ \varphi = 0 & \text{sur } \Sigma, & (1.46b) \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega. & (1.46c) \end{cases}$$

**Espace des trajectoires** On commence par introduire les espaces de trajectoires

$$\begin{aligned}\Phi &:= \left\{ \varphi \in L^2(Q); \quad (\varphi, \partial_t \varphi) \in C([0, T]; V), \quad \square \varphi \in L^2(Q) \right\}, \\ \Phi_0 &:= \left\{ \varphi \in L^2(Q); \quad (\varphi, \partial_t \varphi) \in C([0, T]; V), \quad \square \varphi = 0 \right\}.\end{aligned}$$

D'après la Prop. 28, si l'inégalité d'observabilité (1.21) est vérifiée, il existe alors  $C > 0$  tel que

$$\|(\varphi, \partial_t \varphi)(\cdot, 0)\|_V^2 \leq C \left( \|\partial_t \varphi\|_{L^2(Q)}^2 + \|\square \varphi\|_{L^2(Q)}^2 \right), \quad \forall \varphi \in \Phi.$$

Pour  $r > 0$  fixé, ceci implique que  $\Phi$  est un espace de Hilbert pour la norme  $\|\cdot\|_\Phi$  associée au produit scalaire

$$\langle \varphi, \bar{\varphi} \rangle_\Phi := \iint_q \partial_t \varphi \partial_t \bar{\varphi} + r \iint_Q \square \varphi \square \bar{\varphi}, \quad \forall \varphi, \bar{\varphi} \in \Phi.$$

**Formulation mixte** Pour tout  $\varphi \in \Phi$ , on définit à présent la nouvelle fonctionnelle conjuguée

$$\tilde{\mathcal{J}}^*(\varphi) := \frac{1}{2} \iint_q |\partial_t \varphi|^2 - \langle y^0, \varphi(\cdot, 0) \rangle_{H_0^1} - \langle y^1, \partial_t \varphi(\cdot, 0) \rangle_{L^2},$$

et on cherche à résoudre le problème de minimisation

$$\inf_{\varphi \in \Phi_0} \tilde{\mathcal{J}}^*(\varphi).$$

Pour ce faire, on traite la contrainte  $\square \varphi = 0$  présente dans  $\Phi_0$  via un multiplicateur de Lagrange  $\lambda \in \Lambda := L^2(Q)$ . On considère ainsi le lagrangien augmenté

$$\mathcal{L}_r(\varphi, \lambda) := \frac{1}{2} a_r(\varphi, \varphi) + b(\varphi, \lambda) - \ell(\varphi), \quad \forall (\varphi, \lambda) \in \Phi \times \Lambda,$$

ainsi que la formulation mixte

$$\begin{cases} a_r(\varphi, \bar{\varphi}) + b(\bar{\varphi}, \lambda) = \ell(\bar{\varphi}), & \forall \bar{\varphi} \in \Phi, \\ b(\varphi, \bar{\lambda}) = 0, & \forall \bar{\lambda} \in \Lambda, \end{cases} \quad \begin{array}{l} (1.47a) \\ (1.47b) \end{array}$$

où

$$\begin{aligned}a_r : \Phi \times \Phi &\rightarrow \mathbb{R}, & a_r(\varphi, \bar{\varphi}) &= \iint_q \partial_t \varphi \partial_t \bar{\varphi} + r \iint_Q \square \varphi \square \bar{\varphi}, \\ b : \Phi \times \Lambda &\rightarrow \mathbb{R}, & b(\varphi, \lambda) &= \iint_Q \square \varphi \lambda, \\ \ell : \Phi &\rightarrow \mathbb{R}, & \ell(\varphi) &= \langle y^0, \varphi(\cdot, 0) \rangle_{H_0^1} + \langle y^1, \partial_t \varphi(\cdot, 0) \rangle_{L^2}.\end{aligned}$$

En adaptant la Prop. 29, on peut alors montrer le résultat suivant.

**Proposition 30.** *Les propriétés suivantes sont vérifiées.*

- ▷ La formulation mixte (1.47) est bien posée.
- ▷ La solution  $(\hat{\varphi}, \hat{\lambda}) \in \Phi \times \Lambda$  de (1.47) est l'unique point-selle du lagrangien  $\mathcal{L}_r$ .
- ▷ On a  $\hat{\varphi} \in \Phi_0$  et  $\hat{\varphi}$  est l'unique point de minimum de  $\tilde{\mathcal{J}}^*$  sur  $\Phi_0$ .
- ▷ Le contrôle de norme  $L^2(q)$  minimale pour (1.19) est donné par  $\hat{u} = -\partial_t \hat{\varphi}|_q$ .
- ▷ On a  $\hat{\lambda} = -\partial_t y$ , où  $y$  est la solution de (1.19) associée à  $(y^0, y^1)$  et  $\hat{u}$ .

**Discrétisation** On passe maintenant à la discrétisation de la formulation mixte (1.47). On considère un maillage  $\mathcal{T}_h$  de  $Q = \Omega \times (0, T) \subset \mathbb{R}^2$ , c'est-à-dire que  $Q = \cup_{K \in \mathcal{T}_h} K$ . Le paramètre  $h > 0$  correspond à la taille de la plus grande maille de  $\mathcal{T}_h$ . Il est à noter que le maillage  $\mathcal{T}_h$  peut ici être choisi adapté à l'ouvert de contrôle  $q \subset Q$ . Pour ce faire, on construit deux sous-maillages  $\mathcal{T}_h^1$  et  $\mathcal{T}_h^2$  tels que

$$\bar{q} = \bigcup_{K \in \mathcal{T}_h^1} K \quad \text{et} \quad \overline{Q \setminus q} = \bigcup_{K \in \mathcal{T}_h^2} K,$$



de sorte que  $\mathcal{T}_h = \mathcal{T}_h^1 \sqcup \mathcal{T}_h^2$ . On introduit ensuite deux espaces de dimension finie  $\Phi_h \subset \Phi$  et  $\Lambda_h \subset \Lambda$ . La formulation mixte discrétisée s'écrit alors

$$\begin{cases} a_r(\varphi_h, \bar{\varphi}_h) + b(\bar{\varphi}_h, \lambda_h) = \ell(\bar{\varphi}_h), & \forall \bar{\varphi}_h \in \Phi_h, \\ b(\varphi_h, \bar{\lambda}_h) = 0, & \forall \bar{\lambda}_h \in \Lambda_h. \end{cases} \quad (1.48a)$$

$$(1.48b)$$

On note  $n_h := \dim(\Phi_h)$  et  $m_h := \dim(\Lambda_h)$ . La formulation mixte (1.48) est équivalente au système matriciel

$$\begin{pmatrix} A_{r,h} & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \{\varphi_h\} \\ \{\lambda_h\} \end{pmatrix} = \begin{pmatrix} L_h \\ 0 \end{pmatrix}, \quad (1.49)$$

où les matrices  $A_{r,h} \in \mathbb{R}^{n_h, n_h}$ ,  $B_h \in \mathbb{R}^{m_h, n_h}$  et  $L_h \in \mathbb{R}^{n_h}$  sont données par

$$\begin{aligned} \langle A_{r,h} \{\varphi_h\}, \{\bar{\varphi}_h\} \rangle_{n_h} &= a_r(\varphi_h, \bar{\varphi}_h), & \forall \varphi_h, \bar{\varphi}_h \in \Phi_h, \\ \langle B_h \{\varphi_h\}, \{\lambda_h\} \rangle_{m_h} &= b(\varphi_h, \lambda_h), & \forall \varphi_h \in \Phi_h, \forall \lambda_h \in \Lambda_h, \\ \langle L_h, \{\varphi_h\} \rangle_{n_h} &= \ell(\varphi_h), & \forall \varphi_h \in \Phi_h. \end{aligned}$$

En pratique, on inverse le système (1.49) par décomposition LU.

Le caractère bien posé de la formulation mixte (1.48) est une conséquence des deux propriétés suivantes.

▷ La forme bilinéaire  $a_r$  est coercive sur le noyau

$$\mathcal{N}_h(b) := \left\{ \varphi_h \in \Phi_h; \quad b(\varphi_h, \lambda_h) = 0, \forall \lambda_h \in \Lambda_h \right\}.$$

▷ La forme bilinéaire  $b$  vérifie la condition *inf-sup* discrète suivante. Il existe  $\delta_h > 0$  tel que

$$\inf_{\lambda_h \in \Lambda_h} \sup_{\varphi_h \in \Phi_h} \frac{b(\varphi_h, \lambda_h)}{\|\varphi_h\|_{\Phi} \|\lambda_h\|_{\Lambda}} \geq \delta_h. \quad (1.50)$$

La coercivité de  $a_r$  sur  $\mathcal{N}_h(b)$  provient de la présence du paramètre d'augmentation  $r > 0$ . En effet, comme  $a_r(\varphi, \varphi) = \|\varphi\|_{\Phi}^2$  pour tout  $\varphi \in \Phi$ , la forme  $a_r$  est coercive sur  $\Phi$  tout entier, et donc a fortiori sur  $\mathcal{N}_h(b) \subset \Phi_h \subset \Phi$ . La vérification de la condition (1.50) dépend quant à elle des espaces  $\Phi_h$  et  $\Lambda_h$  choisis. Dans cet exemple, on considère

$$\begin{aligned} \Phi_h &:= \left\{ \varphi_h \in C^1(\bar{Q}); \quad \varphi_h|_K \in \mathbb{H}(K), \forall K \in \mathcal{T}_h, \quad \varphi_h = 0 \text{ sur } \Sigma \right\}, \\ \Lambda_h &:= \left\{ \lambda_h \in C(\bar{Q}); \quad \lambda_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h, \quad \lambda_h = 0 \text{ sur } \Sigma \right\}. \end{aligned}$$

Ici,  $\mathbb{H}(K)$  désigne l'élément fini de *Hsieh-Clough-Tocher* réduit, tandis que  $\mathbb{P}_1(K)$  désigne l'élément fini de Lagrange de degré un. On considère également la matrice  $M_h \in \mathbb{R}^{m_h, m_h}$  donnée par

$$\langle M_h \{\lambda_h\}, \{\bar{\lambda}_h\} \rangle_{m_h} = \langle \lambda_h, \bar{\lambda}_h \rangle_{\Lambda}, \quad \forall \lambda_h, \bar{\lambda}_h \in \Lambda_h.$$

Avec les espaces choisis, les matrices  $A_{r,h}$  et  $M_h$  sont symétriques définies positives. En outre, on remarque que  $\ker(B_h^T) = \{0_{m_h}\}$ . On peut alors montrer que la constante *inf-sup* optimale vérifie

$$\delta_h^2 = \inf_{\lambda_h \in \Lambda_h} \frac{\langle B_h A_{r,h}^{-1} B_h^T \{\lambda_h\}, \{\lambda_h\} \rangle_{m_h}}{\langle M_h \{\lambda_h\}, \{\lambda_h\} \rangle_{m_h}} > 0.$$

En effet, pour  $\varphi_h \in \Phi_h$  et  $\lambda_h \in \Lambda_h$ , on a

$$\|\varphi_h\|_{\Phi}^2 = \langle A_{r,h}^{1/2} \{\varphi_h\}, A_{r,h}^{1/2} \{\varphi_h\} \rangle_{n_h} \quad \text{et} \quad b(\varphi_h, \lambda_h) = \langle A_{r,h}^{1/2} \{\varphi_h\}, A_{r,h}^{-1/2} B_h^T \{\lambda_h\} \rangle_{n_h}.$$

Par suite, il vient

$$\sup_{\varphi_h \in \Phi_h} \frac{b(\varphi_h, \lambda_h)}{\|\varphi_h\|_{\Phi}} = \sup_{\psi_h \in \Phi_h} \frac{\langle \{\psi_h\}, A_{r,h}^{-1/2} B_h^T \{\lambda_h\} \rangle_{n_h}}{\|\{\psi_h\}\|_{n_h}} = \|A_{r,h}^{-1/2} B_h^T \{\lambda_h\}\|_{n_h},$$

puis

$$\inf_{\lambda_h \in \Lambda_h} \sup_{\varphi_h \in \Phi_h} \frac{b(\varphi_h, \lambda_h)}{\|\varphi_h\|_{\Phi} \|\lambda_h\|_{\Lambda}} = \inf_{\lambda_h \in \Lambda_h} \left( \frac{\langle B_h A_{r,h}^{-1} B_h^T \{\lambda_h\}, \{\lambda_h\} \rangle_{m_h}}{\langle M_h \{\lambda_h\}, \{\lambda_h\} \rangle_{m_h}} \right)^{1/2}.$$

**Convergence** Afin de garantir la convergence de la solution discrétisée  $(\widehat{\varphi}_h, \widehat{\lambda}_h) \in \Phi_h \times \Lambda_h$  de (1.48) vers la solution continue  $(\widehat{\varphi}, \widehat{\lambda}) \in \Phi \times \Lambda$  de (1.47) quand  $h$  tend vers 0, il est nécessaire de montrer que la constante *inf-sup*  $\delta_h$  dans (1.50) est bornée inférieurement par rapport à  $h$ , c'est-à-dire qu'il existe  $\delta > 0$  tel que

$$\inf_{h>0} \delta_h \geq \delta. \quad (1.51)$$

Cependant, la condition (1.51) est généralement délicate à démontrer. Ainsi, pour les espaces  $\Phi_h$  et  $\Lambda_h$  utilisés ici, on dispose uniquement d'une justification numérique (voir [8, Sect. 4.5]). Par ailleurs, dans le cas où  $\delta_h \rightarrow 0$  quand  $h \rightarrow 0$ , on peut parfois prendre le paramètre d'augmentation  $r > 0$  comme une fonction de  $h$  pour rétablir la condition (1.51).

### Problème bidual

Dans cette section, on énonce un problème de minimisation, dual de la formulation mixte (1.44), et mettant en jeu uniquement le multiplicateur  $\mu \in \mathcal{Z}$ . On reprend ici les notations de la sous-section précédente. Pour  $r > 0$  fixé, on définit l'opérateur linéaire

$$\mathcal{A}_r : \begin{array}{ccc} \mathcal{Z} & \rightarrow & \mathcal{Z} \\ \mu & \mapsto & f \end{array},$$

où  $f$  est le second membre associé à la solution  $\zeta \in \mathfrak{Z}$ ,  $\zeta = \zeta(\cdot; \zeta^0, f)$ , du problème variationnel

$$a_r(\zeta, \bar{\zeta}) = b(\bar{\zeta}, \mu), \quad \forall \bar{\zeta} \in \mathfrak{Z}. \quad (1.52)$$

*Remarque 37.* Pour que le problème (1.52) soit bien posé, il est nécessaire de prendre le paramètre d'augmentation  $r > 0$ .

**Lemme 7.** *Soit  $r > 0$  fixé. L'opérateur  $\mathcal{A}_r$  est un isomorphisme symétrique et elliptique de  $\mathcal{Z}$  dans  $\mathcal{Z}$ .*

*Démonstration.* L'opérateur  $\mathcal{A}_r$  est continu. En effet, comme la forme bilinéaire  $a_r$  vérifie

$$a_r(\zeta, \zeta) \geq \min(1, \frac{r}{\eta}) \|\zeta\|_{\mathfrak{Z}}^2 \quad \text{et} \quad a_r(\zeta, \zeta) \geq r \|f\|_{\mathcal{Z}}^2, \quad \forall \zeta \in \mathfrak{Z}, \quad \zeta = \zeta(\cdot; \zeta^0, f),$$

d'après le théorème de Lax-Milgram (cf. [5, Cor. V.8]), la solution  $\zeta = \zeta(\cdot; \zeta^0, f)$  de (1.52) vérifie

$$\|f\|_{\mathcal{Z}} \leq \frac{\|\mu\|_{\mathcal{Z}}}{r}.$$

Soient  $\mu, \bar{\mu} \in \mathcal{Z}$ . On note  $\zeta = \zeta(\cdot; \zeta^0, f)$  et  $\bar{\zeta} = \bar{\zeta}(\cdot; \bar{\zeta}^0, \bar{f})$  les solutions correspondantes de (1.52). On peut alors écrire

$$\langle \mathcal{A}_r \mu, \bar{\mu} \rangle_{\mathcal{Z}} = b(\bar{\zeta}, \bar{\mu}) = a_r(\zeta, \bar{\zeta}),$$

ce qui implique que  $\mathcal{A}_r$  est symétrique défini positif. Enfin, pour montrer que  $\mathcal{A}_r$  est elliptique, on raisonne par l'absurde. On suppose qu'il existe une suite  $(\mu_n)_{n \geq 0} \subset \mathcal{Z}$  telle que

$$\|\mu_n\|_{\mathcal{Z}} = 1, \quad \forall n \geq 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} \langle \mathcal{A}_r \mu_n, \mu_n \rangle_{\mathcal{Z}} = 0.$$

Pour tout  $n \geq 0$ , on note  $\zeta_n = \zeta_n(\cdot; \zeta_n^0, f_n)$  la solution de (1.52) associée à  $\mu_n$ . Comme  $\langle \mathcal{A}_r \mu_n, \mu_n \rangle_{\mathcal{Z}} = a_r(\zeta_n, \zeta_n)$ , on en déduit que

$$\lim_{n \rightarrow \infty} \|v(\zeta_n)\|_{\mathcal{U}} = 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} \|f_n\|_{\mathcal{Z}} = 0.$$

Pour tout  $n \geq 0$ , on note  $\bar{\zeta}_n = \bar{\zeta}_n(\cdot; 0, \bar{f}_n)$  la solution de (1.42) avec donnée initiale nulle et  $\bar{f}_n = \mu_n - r f_n$  comme second membre. D'après le Lem. 5, en écrivant (1.52) avec  $\zeta = \zeta_n$ ,  $\bar{\zeta} = \bar{\zeta}_n$  et  $\mu = \mu_n$ , on trouve

$$\|\mu_n - r f_n\|_{\mathcal{Z}}^2 = \langle v(\zeta_n), v(\bar{\zeta}_n) \rangle_{\mathcal{U}} \leq \|v(\zeta_n)\|_{\mathcal{U}} \|v(\bar{\zeta}_n)\|_{\mathcal{U}} \leq C \|v(\zeta_n)\|_{\mathcal{U}} \|\mu_n - r f_n\|_{\mathcal{Z}},$$

d'où  $\|\mu_n - r f_n\|_{\mathcal{Z}} \leq C \|v(\zeta_n)\|_{\mathcal{U}}$ . On en conclut que  $\lim_{n \rightarrow \infty} \|\mu_n\|_{\mathcal{Z}} = 0$ , ce qui est une contradiction.  $\square$

**Proposition 31** (Fonctionnelle biduale). *Soit  $r > 0$  fixé. On note  $\zeta_0 \in \mathfrak{Z}$  la solution de*

$$a_r(\zeta_0, \bar{\zeta}) = \ell(\bar{\zeta}), \quad \forall \bar{\zeta} \in \mathfrak{Z}. \quad (1.53)$$

On définit la fonctionnelle

$$\tilde{\mathcal{J}}^{**}(\mu) := \frac{1}{2} \langle \mathcal{A}_r \mu, \mu \rangle_{\mathcal{Z}} - b(\zeta_0, \mu), \quad \forall \mu \in \mathcal{Z}.$$

On a alors

$$\sup_{\mu \in \mathcal{Z}} \inf_{\zeta \in \mathfrak{Z}} \mathcal{L}_r(\zeta, \mu) = - \inf_{\mu \in \mathcal{Z}} \tilde{\mathcal{J}}^{**}(\mu) + \mathcal{L}_r(\zeta_0, 0).$$

*Démonstration.* Soit  $\mu \in \mathcal{Z}$ . On note  $\hat{\zeta} \in \mathfrak{Z}$  le point de minimum de  $\zeta \mapsto \mathcal{L}_r(\zeta, \mu)$ . La fonction  $\hat{\zeta}$  vérifie

$$a_r(\hat{\zeta}, \bar{\zeta}) + b(\bar{\zeta}, \mu) = \ell(\bar{\zeta}), \quad \forall \bar{\zeta} \in \mathfrak{Z},$$

et peut être décomposée en  $\hat{\zeta} = \zeta_0 - \tilde{\zeta}$ , où  $\tilde{\zeta}$  vérifie

$$a_r(\tilde{\zeta}, \bar{\zeta}) = b(\bar{\zeta}, \mu), \quad \forall \bar{\zeta} \in \mathfrak{Z}.$$

Par suite, on a

$$\inf_{\zeta \in \mathfrak{Z}} \mathcal{L}_r(\zeta, \mu) = \mathcal{L}_r(\hat{\zeta}, \mu) = \mathcal{L}_r(\zeta_0 - \tilde{\zeta}, \mu) = X_1 + X_2 + X_3,$$

avec

$$\begin{cases} X_1 = \frac{1}{2} a_r(\tilde{\zeta}, \tilde{\zeta}) - b(\tilde{\zeta}, \mu) + b(\zeta_0, \mu), \\ X_2 = \frac{1}{2} a_r(\zeta_0, \zeta_0) - \ell(\zeta_0), \\ X_3 = \ell(\tilde{\zeta}) - a_r(\zeta_0, \tilde{\zeta}). \end{cases}$$

On remarque que  $X_3 = 0$  par définition de  $\zeta_0$  et que  $X_2 = \mathcal{L}_r(\zeta_0, 0)$ . Finalement, par définition de  $\tilde{\zeta}$ , il vient

$$X_1 = -\frac{1}{2} b(\tilde{\zeta}, \mu) + b(\zeta_0, \mu) = -\frac{1}{2} \langle \mathcal{A}_r \mu, \mu \rangle_{\mathcal{Z}} + b(\zeta_0, \mu).$$

□

Le problème de minimisation de  $\tilde{\mathcal{J}}^{**}$  est sans contrainte. En outre, l'opérateur  $\mathcal{A}_r$  est symétrique et elliptique. On peut donc approcher le point de minimum  $\hat{\mu}$  de  $\tilde{\mathcal{J}}^{**}$  de manière directe par la méthode du gradient conjugué (cf. Algo. 2). Une fois que l'on a approché le multiplicateur  $\hat{\mu}$ , on obtient le contrôle  $\hat{u} = v(\hat{\zeta})$  de norme  $\mathcal{U}$  minimale en résolvant

$$a_r(\hat{\zeta}, \bar{\zeta}) + b(\bar{\zeta}, \hat{\mu}) = \ell(\bar{\zeta}), \quad \forall \bar{\zeta} \in \mathfrak{Z}.$$

## Approximation par série de Fourier

Dans cette section, on expose une méthode permettant d'approcher le contrôle de norme minimale pour le système hyperbolique du second ordre (1.14) à l'aide de développements en série de Fourier. On reprend ici les notations des Sect. 1.1.4 et 1.1.3.

Les valeurs propres de l'opérateur  $A_0$  sont strictement positives et forment une suite croissante  $(\lambda_k^2)_{k \in \mathbb{N}^*}$  de limite  $+\infty$ . En outre, il existe une base orthonormée  $(\varphi_k)_{k \in \mathbb{N}^*}$  de  $H$ , où  $\varphi_k$  est une fonction propre associée à la valeur propre  $\lambda_k^2$ . Ceci provient du fait que l'opérateur  $A_0^{-1}$  est compact, autoadjoint et strictement positif.

**Lemme 8.** *Les valeurs propres de l'opérateur*

$$A = \begin{pmatrix} 0 & \text{Id} \\ -A_0 & 0 \end{pmatrix}$$

sont de la forme

$$\mu_k = \text{sgn}(k) i \lambda_{|k|}, \quad \text{pour } k \in \mathbb{Z}^*.$$

De plus, pour  $k \in \mathbb{Z}^*$ , la fonction

$$\xi_k = \begin{pmatrix} 1/\mu_k \\ 1 \end{pmatrix} \frac{\varphi_{|k|}}{\sqrt{2}}$$

est une fonction propre associée à la valeur propre  $\mu_k$  et la famille  $(\xi_k)_{k \in \mathbb{Z}^*}$  forme une base orthonormée de  $Z$ .

---

**Algorithm 2** Minimisation de  $\tilde{\mathcal{J}}^{**}$  par la méthode du gradient conjugué

---

- 1) *Initialisation.* On fixe  $\varepsilon > 0$  et on se donne  $\mu_0 \in \mathcal{Z}$ .  
On calcule  $\zeta_0 \in \mathfrak{Z}$  solution de

$$a_r(\zeta_0, \bar{\zeta}) = b(\bar{\zeta}, \mu_0) - \ell(\bar{\zeta}), \quad \forall \bar{\zeta} \in \mathfrak{Z}.$$

On calcule  $g_0 \in \mathcal{Z}$  solution de

$$\langle g_0, \bar{\mu} \rangle_{\mathcal{Z}} = b(\zeta_0, \bar{\mu}), \quad \forall \bar{\mu} \in \mathcal{Z}.$$

On pose  $p_0 = g_0$ .

- 2) *Itération.* Pour  $n \geq 0$ , on calcule  $\tilde{\zeta}_n \in \mathfrak{Z}$  solution de

$$a_r(\tilde{\zeta}_n, \bar{\zeta}) = b(\bar{\zeta}, p_n), \quad \forall \bar{\zeta} \in \mathfrak{Z}.$$

On pose

$$\alpha_n = \frac{\|g_n\|_{\mathcal{Z}}^2}{b(\tilde{\zeta}_n, p_n)}.$$

On met à jour

$$\mu_{n+1} = \mu_n - \alpha_n p_n.$$

On calcule  $g_{n+1} \in \mathcal{Z}$  solution de

$$\langle g_{n+1}, \bar{\mu} \rangle_{\mathcal{Z}} = \langle g_n, \bar{\mu} \rangle_{\mathcal{Z}} - \alpha_n b(\tilde{\zeta}_n, \bar{\mu}), \quad \forall \bar{\mu} \in \mathcal{Z}.$$

- 3) *Test de convergence.* Si  $\|g_{n+1}\|_{\mathcal{Z}} \leq \varepsilon \|g_0\|_{\mathcal{Z}}$ , l'algorithme s'arrête et on prend  $\mu_{n+1}$  comme approximation du point de minimum de  $\tilde{\mathcal{J}}^{**}$ . Sinon, on pose

$$\beta_n = \frac{\|g_{n+1}\|_{\mathcal{Z}}^2}{\|g_n\|_{\mathcal{Z}}^2}.$$

On met à jour

$$p_{n+1} = g_{n+1} + \beta_n p_n.$$

Effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).

---

*Démonstration.* Soit  $\mu \in \mathbb{C}$  et  $\zeta \in Z \setminus \{0\}$  tels que  $A\zeta = \mu\zeta$ . On a alors  $\zeta_2 = \mu\zeta_1$  et  $-A_0\zeta_1 = \mu\zeta_2$ . On en déduit que  $A_0\zeta_1 = -\mu^2\zeta_1$ . Par suite, il existe  $k \in \mathbb{N}^*$  tel que  $-\mu^2 = \lambda_k^2$ , d'où  $\mu = \pm i\lambda_k$ .

Pour  $k \in \mathbb{Z}^*$ , on calcule

$$A\xi_k = \begin{pmatrix} 1 \\ -\lambda_{|k|}^2/\mu_k \end{pmatrix} \frac{\varphi_{|k|}}{\sqrt{2}} = \begin{pmatrix} 1 \\ \mu_k \end{pmatrix} \frac{\varphi_{|k|}}{\sqrt{2}} = \mu_k \xi_k.$$

Pour  $p, q \in \mathbb{Z}^*$ , on calcule

$$\langle \xi_p, \xi_q \rangle_Z = \frac{\lambda_{|p|}^2}{2\mu_p\mu_q} \langle \varphi_{|p|}, \varphi_{|q|} \rangle_H + \frac{1}{2} \langle \varphi_{|p|}, \varphi_{|q|} \rangle_H = \frac{1}{2} \left( 1 + \frac{\mu_p}{\mu_q} \right) \delta_{|p|, |q|} = \delta_{p, q}.$$

La famille  $(\xi_k)_{k \in \mathbb{Z}^*}$  forme alors une base orthonormée de  $Z$  puisque l'opérateur  $A^{-1}$  est compact et antiadjoint.  $\square$

**Lemme 9.** Soit  $(c_k)_{k \in \mathbb{Z}^*} \subset \mathbb{C}$  tel que  $\sum_{k \in \mathbb{Z}^*} |c_k|^2 < +\infty$ . La solution  $\zeta$  de (1.9) associée à la donnée initiale

$$\zeta^0 = \sum_{k \in \mathbb{Z}^*} c_k \xi_k \in Z$$

est donnée par

$$\zeta(t) = \sum_{k \in \mathbb{Z}^*} c_k e^{\mu_k t} \xi_k, \quad t \in [0, T].$$

**Lemme 10.** Soient  $(a_k)_{k \in \mathbb{N}^*}, (b_k)_{k \in \mathbb{N}^*} \subset \mathbb{R}$  tels que  $\sum_{k \in \mathbb{N}^*} |\lambda_k a_k|^2 < +\infty$  et  $\sum_{k \in \mathbb{N}^*} |b_k|^2 < +\infty$ . La solution  $\zeta$  de (1.9) associée à la donnée initiale

$$\zeta^0 = \sum_{k \in \mathbb{N}^*} \begin{pmatrix} a_k \\ b_k \end{pmatrix} \varphi_k \in Z$$

est donnée par

$$\zeta(t) = \sum_{k \in \mathbb{N}^*} \begin{pmatrix} a_k \cos(\lambda_k t) + \frac{b_k}{\lambda_k} \sin(\lambda_k t) \\ b_k \cos(\lambda_k t) - \lambda_k a_k \sin(\lambda_k t) \end{pmatrix} \varphi_k, \quad t \in [0, T].$$

Le contrôle à zéro de norme  $\mathcal{U}$  minimale pour le système d'évolution (1.4) est donné par

$$\widehat{u}(z^0) = v(\widehat{\zeta}^0),$$

où  $\widehat{\zeta}^0 \in Z$  est la solution de

$$a(\widehat{\zeta}^0, \zeta^0) = \ell(\zeta^0), \quad \forall \zeta^0 \in Z. \quad (1.54)$$

Pour  $\zeta^0 \in Z$ , on rappelle que

$$v(\zeta^0) = -B^* \zeta,$$

où  $\zeta$  est la solution de (1.9) associée à  $\zeta^0$ . On a de plus

$$a(\widehat{\zeta}^0, \zeta^0) = \langle v(\widehat{\zeta}^0), v(\zeta^0) \rangle_{\mathcal{U}}.$$

En décomposant

$$\widehat{\zeta}^0 = \sum_{p \in \mathbb{Z}^*} \widehat{c}_p \xi_p \quad \text{et} \quad \zeta^0 = \sum_{q \in \mathbb{Z}^*} c_q \xi_q,$$

on vérifie que  $\widehat{\zeta}^0$  est solution de (1.54) si et seulement si

$$\sum_{p, q \in \mathbb{Z}^*} \widehat{c}_p c_q a(\xi_p, \xi_q) = \sum_{q \in \mathbb{Z}^*} c_q \ell(\xi_q)$$

pour tout  $(c_q)_{q \in \mathbb{Z}^*} \subset \mathbb{C}$  tel que  $\sum_{q \in \mathbb{Z}^*} |c_q|^2 < +\infty$ . On peut alors obtenir une approximation de la solution  $\widehat{\zeta}^0$  de (1.54) en résolvant le système linéaire issu de la troncature des développements au-delà de la  $N$ -ième harmonique (cf. Algo. 3).

---

**Algorithm 3** Approximation du contrôle par décomposition en série de Fourier

---

On fixe  $N \in \mathbb{N}^*$  et on note  $\mathbb{Z}_N^* = \llbracket -N, N \rrbracket \setminus \{0\}$ .

On définit  $\mathbf{A} \in M_{2N}(\mathbb{C})$  par

$$\mathbf{A}_{p,q} = a(\xi_q, \xi_p), \quad \forall p, q \in \mathbb{Z}_N^*.$$

On définit  $\mathbf{L} \in \mathbb{C}^{2N}$  par

$$\mathbf{L}_p = \ell(\xi_p), \quad \forall p \in \mathbb{Z}_N^*.$$

On calcule  $\mathbf{c} = (c_p)_{p \in \mathbb{Z}_N^*}$  solution du système

$$\mathbf{A}\mathbf{c} = \mathbf{L}.$$

On prend

$$\widehat{\zeta}_N^0 = \sum_{p \in \mathbb{Z}_N^*} c_p \xi_p$$

comme approximation de la solution  $\widehat{\zeta}^0$  de (1.54).

---

Pour illustrer l'Algo. 3, on le met en œuvre dans le cas de l'équation des ondes linéaire 1D avec contrôle interne. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . On considère le cas particulier d'un ouvert de contrôle  $q_\gamma \subset Q$  de la forme

$$q_\gamma := \left\{ (x, t) \in Q; \quad |x - \gamma(t)| < \delta_0 \right\},$$

avec  $\delta_0 > 0$  et  $\gamma \in W^{1,\infty}(0, T)$  fixés. En outre, pour tout  $(x, t) \in \overline{Q}$ , on définit  $\chi_\gamma(x, t) := \chi(x - \gamma(t))$ , avec  $\chi : \mathbb{R} \rightarrow [0, 1]$  une fonction paire telle que  $\text{Supp } \chi = [-\delta_0, \delta_0]$ . On a ainsi  $\text{Supp } \chi_\gamma = \overline{q_\gamma}$ . On s'intéresse alors au contrôle  $\widehat{u}$  de norme  $L_\chi^2(q_\gamma)$  minimale, où

$$L_\chi^2(q_\gamma) := L^2(q_\gamma; \chi_\gamma) = \left\{ u : q_\gamma \rightarrow \mathbb{R}; \quad \iint_{q_\gamma} u^2 \chi_\gamma < +\infty \right\}.$$

On utilise la caractérisation  $\widehat{u} = -\widehat{\varphi} \chi_\gamma|_{q_\gamma}$ , avec  $\widehat{\varphi}$  l'unique solution de (1.20) vérifiant

$$\iint_{q_\gamma} \widehat{\varphi} \overline{\varphi} \chi_\gamma = \langle y^1, \overline{\varphi}^0 \rangle_{L^2} - \langle y^0, \overline{\varphi}^1 \rangle_{H_0^1, H^{-1}}, \quad \forall (\overline{\varphi}^0, \overline{\varphi}^1) \in L^2(\Omega) \times H^{-1}(\Omega), \quad (1.55)$$

où  $\overline{\varphi}$  est la solution de (1.20) associée à  $(\overline{\varphi}^0, \overline{\varphi}^1)$ .

Soit  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . On commence par développer  $y^0$  et  $y^1$  sous la forme

$$y^0(x) = \sum_{p \geq 1} c_p(y^0) \sin(p\pi x) \quad \text{et} \quad y^1(x) = \sum_{p \geq 1} c_p(y^1) \sin(p\pi x),$$

où

$$c_p(y^0) = 2 \int_\Omega y^0(x) \sin(p\pi x) dx \quad \text{et} \quad c_p(y^1) = 2 \int_\Omega y^1(x) \sin(p\pi x) dx.$$

On cherche ensuite les solutions  $\varphi$  de (1.20) sous la forme

$$\varphi(x, t) = \sum_{p \geq 1} \left( a_p \cos(p\pi t) + \frac{b_p}{p\pi} \sin(p\pi t) \right) \sin(p\pi x).$$

Avec cette décomposition, on obtient

$$\begin{aligned} \iint_{q_\gamma} \varphi \overline{\varphi} \chi_\gamma &= \sum_{p, q \geq 1} a_p \overline{a}_q \iint_{q_\gamma} \cos(p\pi t) \cos(q\pi t) \sin(p\pi x) \sin(q\pi x) \chi_\gamma(x, t) dx dt \\ &+ \sum_{p, q \geq 1} a_p \frac{\overline{b}_q}{q\pi} \iint_{q_\gamma} \cos(p\pi t) \sin(q\pi t) \sin(p\pi x) \sin(q\pi x) \chi_\gamma(x, t) dx dt \\ &+ \sum_{p, q \geq 1} \frac{b_p}{p\pi} \overline{a}_q \iint_{q_\gamma} \sin(p\pi t) \cos(q\pi t) \sin(p\pi x) \sin(q\pi x) \chi_\gamma(x, t) dx dt \\ &+ \sum_{p, q \geq 1} \frac{b_p}{p\pi} \frac{\overline{b}_q}{q\pi} \iint_{q_\gamma} \sin(p\pi t) \sin(q\pi t) \sin(p\pi x) \sin(q\pi x) \chi_\gamma(x, t) dx dt \end{aligned} \quad (1.56)$$

et

$$\langle y^1, \bar{\varphi}^0 \rangle_{L^2} - \langle y^0, \bar{\varphi}^1 \rangle_{H_0^1, H^{-1}} = \frac{1}{2} \sum_{q \geq 1} c_q(y^1) \bar{a}_q - \frac{1}{2} \sum_{q \geq 1} c_q(y^0) \bar{b}_q. \quad (1.57)$$

Enfin, en fixant  $N \in \mathbb{N}^*$  et en tronquant les sommes ci-dessus au-delà du rang  $N$ , la caractérisation (1.55) se ramène au système linéaire  $\mathbf{A}\mathbf{c} = \mathbf{L}$ , avec  $\mathbf{c} = \{a_p, b_p\}_{1 \leq p \leq N}$ , et où les matrices  $\mathbf{A} \in \mathbb{R}^{2N, 2N}$  et  $\mathbf{L} \in \mathbb{R}^{2N}$  sont calculées via (1.56) et (1.57) respectivement.

Concernant la solution contrôlée  $y$  de (1.19) associée à  $(y^0, y^1)$  et  $\hat{u}$ , on la développe sous la forme

$$y(x, t) = \sum_{p=1}^N c_p(t) \sin(p\pi x).$$

En injectant cette décomposition dans (1.19), il vient

$$\begin{cases} c_p''(t) + (p\pi)^2 c_p(t) = f_p(t), \\ c_p(0) = c_p(y^0), \quad c_p'(0) = c_p(y^1), \end{cases} \quad 1 \leq p \leq N,$$

où

$$f_p(t) = -2 \sum_{q=1}^N \left( \hat{a}_q \cos(q\pi t) + \frac{\hat{b}_q}{q\pi} \sin(q\pi t) \right) \mathcal{I}_{p,q}(t)$$

et

$$\mathcal{I}_{p,q}(t) = \int_{\gamma(t)-\delta_0}^{\gamma(t)+\delta_0} \sin(p\pi x) \sin(q\pi x) \chi_\gamma(x, t) dx.$$

Pour  $1 \leq p \leq N$ , on a donc

$$c_p(t) = c_p(y^0) \cos(p\pi t) + \frac{c_p(y^1)}{p\pi} \sin(p\pi t) + \frac{1}{p\pi} \int_0^t f_p(s) \sin(p\pi(t-s)) ds.$$

En pratique, on évalue d'abord les intégrales  $\mathcal{I}_{p,q}(t)$  pour tous  $1 \leq p, q \leq N$ , avant d'effectuer l'intégration en temps. Par ailleurs, dans le cas  $\chi_\gamma = \mathbb{1}_{q_\gamma}$ , on a

$$\mathcal{I}_{p,q}(t) = \delta_0 \left( \cos((p-q)\pi\gamma(t)) \operatorname{sinc}((p-q)\pi\delta_0) - \cos((p+q)\pi\gamma(t)) \operatorname{sinc}((p+q)\pi\delta_0) \right).$$

### Approximation de la constante d'observabilité

Dans cette section, on présente une méthode itérative permettant d'approcher numériquement la constante d'observabilité  $C_{\text{obs}}$ . On reprend ici les notations de la Sect. 1.1.3. L'idée est de s'appuyer sur la relation  $C_{\text{opt}} = \sqrt{C_{\text{obs}}}$  (cf. [10, Th. 2.44]), qui lie la constante d'observabilité  $C_{\text{obs}}$  au coût de contrôle  $C_{\text{opt}}$ . Par définition du coût de contrôle, on a

$$C_{\text{opt}} = \sup_{z^0 \in Z} \frac{\|\hat{u}(z^0)\|_{\mathcal{U}}}{\|z^0\|_Z},$$

où  $\hat{u}(z^0)$  est le contrôle de norme  $\mathcal{U}$  minimale associé à  $z^0$ . Or, on sait que  $\hat{u}(z^0) = v(\hat{\zeta}^0)$ , avec  $\hat{\zeta}^0 = \Lambda^{-1}z^0$ , et où  $\Lambda$  est l'opérateur HUM introduit à la Sect. 1.1.3. En outre, avec la caractérisation (1.13) de  $\hat{\zeta}^0$ , on trouve

$$\|\hat{u}(z^0)\|_{\mathcal{U}}^2 = \|v(\hat{\zeta}^0)\|_{\mathcal{U}}^2 = \langle z^0, \hat{\zeta}^0 \rangle_Z = \langle z^0, \Lambda^{-1}z^0 \rangle_Z,$$

puis

$$C_{\text{obs}} = C_{\text{opt}}^2 = \sup_{z^0 \in Z} \frac{\|\hat{u}(z^0)\|_{\mathcal{U}}^2}{\|z^0\|_Z^2} = \sup_{z^0 \in Z} \frac{\langle \Lambda^{-1}z^0, z^0 \rangle_Z}{\|z^0\|_Z^2} = \left( \inf_{\zeta^0 \in Z} \frac{\langle \Lambda \zeta^0, \zeta^0 \rangle_Z}{\|\zeta^0\|_Z^2} \right)^{-1}.$$

La constante d'observabilité  $C_{\text{obs}}$  est donc la plus grande valeur propre de l'opérateur  $\Lambda^{-1}$  ou, de manière équivalente, l'inverse de la plus petite valeur propre de  $\Lambda$ . Pour approcher numériquement cette valeur propre, on peut utiliser la méthode de la puissance itérée (cf. Algo. 4), ou encore la méthode d'itération du quotient de Rayleigh (cf. Algo. 5).

---

**Algorithm 4** Approximation de  $C_{\text{obs}}$  par la méthode de la puissance itérée

---

- 1) *Initialisation.* On fixe  $\varepsilon > 0$  et on se donne  $z_0 \in Z$  tel que  $\|z_0\|_Z = 1$ .
- 2) *Itération.* Pour  $n \geq 0$ , on calcule  $\zeta_n \in Z$  solution de

$$\Lambda \zeta_n = z_n.$$

On met à jour

$$z_{n+1} = \frac{\zeta_n}{\|\zeta_n\|_Z} \quad \text{et} \quad \kappa_{n+1} = \|\zeta_n\|_Z.$$

On pose

$$g_{n+1} = \|z_n - z_{n+1}\|_Z.$$

- 3) *Test de convergence.* Si  $g_{n+1} \leq \varepsilon$ , l'algorithme s'arrête et on prend  $\kappa_{n+1}$  comme approximation de  $C_{\text{obs}}$ .  
Sinon, effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).
- 

---

**Algorithm 5** Approximation de  $C_{\text{obs}}$  par la méthode d'itération du quotient de Rayleigh

---

- 1) *Initialisation.* On fixe  $\varepsilon > 0$  et on se donne  $(z_0, \kappa_0) \in Z \times \mathbb{R}$  tel que  $\|z_0\|_Z = 1$ .
- 2) *Itération.* Pour  $n \geq 0$ , on calcule  $\zeta_n \in Z$  solution de

$$\Lambda \zeta_n - \kappa_n \zeta_n = z_n.$$

On met à jour

$$z_{n+1} = \frac{\zeta_n}{\|\zeta_n\|_Z} \quad \text{et} \quad \kappa_{n+1} = \frac{\|v(\zeta_n)\|_{\mathcal{U}}^2}{\|\zeta_n\|_Z^2}.$$

On pose

$$g_{n+1} = \frac{\|z_n - \langle z_n, z_{n+1} \rangle_Z z_{n+1}\|_Z}{\kappa_{n+1} \|\zeta_n\|_Z}.$$

- 3) *Test de convergence.* Si  $g_{n+1} \leq \varepsilon$ , l'algorithme s'arrête et on prend  $\kappa_{n+1}^{-1}$  comme approximation de  $C_{\text{obs}}$ .  
Sinon, effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).
-



Pour illustrer les Algo. 4 et 5, on les met en œuvre dans le cas de l'équation des ondes linéaire 1D avec contrôle interne. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . On considère plus précisément des ouverts de contrôle cylindriques de la forme

$$q = (x_0 - \delta_0, x_0 + \delta_0) \times (0, T), \quad \text{avec } \delta_0 > 0 \text{ et } x_0 \in [\delta_0, 1 - \delta_0]. \quad (1.58)$$

En effet, pour de tels ouverts de contrôle, il est possible de calculer explicitement la valeur de la constante d'observabilité. Dans la suite, on note  $\text{sinc}$  la fonction sinus cardinal définie par  $\text{sinc}(x) = \frac{\sin(x)}{x}$  si  $x \neq 0$  et  $\text{sinc}(0) = 1$ .

**Proposition 32.** *Soient  $T \in 2\mathbb{N}^*$  et  $\delta_0 > 0$ . Pour  $x_0 \in [\delta_0, 1 - \delta_0]$ , la constante d'observabilité associée à l'ouvert  $q$  défini en (1.58) est donnée par*

$$C_{\text{obs}}(x_0) = \frac{1}{\delta_0 T (1 - \sigma(x_0))}, \quad \text{avec } \sigma(x_0) = \max_{p \geq 1} \cos(2p\pi x_0) \text{sinc}(2p\pi\delta_0).$$

En outre, pour tout  $x_0 \in [\delta_0, 1 - \delta_0]$ , on a  $C_{\text{obs}}(x_0) \leq \bar{C}_{\text{obs}}$ , où

$$\bar{C}_{\text{obs}} = \frac{1}{\delta_0 T (1 - \text{sinc}(2\pi\delta_0))}.$$

*Démonstration.* Comme  $T \in 2\mathbb{N}^*$ , pour tous  $p, q \geq 1$ , on a

$$\begin{cases} \int_0^T \cos(p\pi t) \cos(q\pi t) dt = \frac{T}{2} \delta_{p,q}, \\ \int_0^T \sin(p\pi t) \sin(q\pi t) dt = \frac{T}{2} \delta_{p,q}, \\ \int_0^T \cos(p\pi t) \sin(q\pi t) dt = 0, \end{cases} \quad \text{et} \quad \begin{cases} \int_0^1 \cos(p\pi x) \cos(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \\ \int_0^1 \sin(p\pi x) \sin(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \end{cases}$$

où  $\delta_{p,q} = 0$  si  $p \neq q$  et  $\delta_{p,p} = 1$ . Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . On décompose

$$\varphi^0(x) = \sum_{p=1}^{+\infty} a_p \sin(p\pi x) \quad \text{et} \quad \varphi^1(x) = \sum_{p=1}^{+\infty} b_p \sin(p\pi x).$$

La solution  $\varphi$  de (1.20) s'écrit alors

$$\varphi(x, t) = \sum_{p=1}^{+\infty} \left( a_p \cos(p\pi t) + \frac{b_p}{p\pi} \sin(p\pi t) \right) \sin(p\pi x).$$

On calcule

$$\|\partial_t \varphi\|_{L^2(q)}^2 = \frac{T}{2} \sum_{p=1}^{+\infty} \mathcal{I}_p(x_0) \left( (p\pi)^2 a_p^2 + b_p^2 \right), \quad \text{avec } \mathcal{I}_p(x_0) = \int_{x_0 - \delta_0}^{x_0 + \delta_0} \sin^2(p\pi x) dx.$$

Or, on a

$$\mathcal{I}_p(x_0) = \delta_0 \left( 1 - \cos(2p\pi x_0) \text{sinc}(2p\pi\delta_0) \right),$$

d'où

$$\|\partial_t \varphi\|_{L^2(q)}^2 \geq T \min_{p \geq 1} \mathcal{I}_p(x_0) \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

Enfin, comme  $|\text{sinc}(2p\pi\delta_0)| \leq \text{sinc}(2\pi\delta_0)$  pour tout  $p \geq 1$ , on obtient la constante d'observabilité uniforme en remarquant que

$$\sigma(x_0) \leq \max_{p \geq 1} |\text{sinc}(2p\pi\delta_0)| \leq \text{sinc}(2\pi\delta_0).$$

□

Par ailleurs, pour des ouverts de contrôle cylindriques de la forme

$$q = \omega \times (0, T), \quad \text{avec } \omega \subset \Omega \text{ ouvert non vide}, \quad (1.59)$$

il est possible de donner une majoration de la constante d'observabilité ne dépendant que de  $T$  et de la mesure de  $\omega$ .

**Proposition 33.** [31, Prop. 2.1] & [20, Th. 1] Soit  $T \in 2\mathbb{N}^*$ . Pour tout ouvert non vide  $\omega \subset \Omega$ , la constante d'observabilité associée à l'ouvert  $q$  défini en (1.59) est majorée par

$$\bar{C}_{\text{obs}} = \frac{2}{|\omega|T(1 - \text{sinc}(\pi|\omega|))}.$$

*Démonstration.* Comme  $T \in 2\mathbb{N}^*$ , pour tous  $p, q \geq 1$ , on a

$$\left\{ \begin{array}{l} \int_0^T \cos(p\pi t) \cos(q\pi t) dt = \frac{T}{2} \delta_{p,q}, \\ \int_0^T \sin(p\pi t) \sin(q\pi t) dt = \frac{T}{2} \delta_{p,q}, \\ \int_0^T \cos(p\pi t) \sin(q\pi t) dt = 0, \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} \int_0^1 \cos(p\pi x) \cos(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \\ \int_0^1 \sin(p\pi x) \sin(q\pi x) dx = \frac{1}{2} \delta_{p,q}, \end{array} \right.$$

où  $\delta_{p,q} = 0$  si  $p \neq q$  et  $\delta_{p,p} = 1$ . Soit  $(\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega)$ . On décompose

$$\varphi^0(x) = \sum_{p=1}^{+\infty} a_p \sin(p\pi x) \quad \text{et} \quad \varphi^1(x) = \sum_{p=1}^{+\infty} b_p \sin(p\pi x).$$

La solution  $\varphi$  de (1.20) s'écrit alors

$$\varphi(x, t) = \sum_{p=1}^{+\infty} \left( a_p \cos(p\pi t) + \frac{b_p}{p\pi} \sin(p\pi t) \right) \sin(p\pi x).$$

On calcule

$$\|\partial_t \varphi\|_{L^2(q)}^2 = \frac{T}{2} \sum_{p=1}^{+\infty} \mathcal{I}_p(\omega) \left( (p\pi)^2 a_p^2 + b_p^2 \right), \quad \text{avec } \mathcal{I}_p(\omega) = \int_{\omega} \sin^2(p\pi x) dx.$$

Or, d'après [31, Prop. 2.1] ou [20, Lem. 1], on a

$$\mathcal{I}_p(\omega) \geq \frac{|\omega|}{2} - \frac{\sin(\pi|\omega|)}{2\pi} = \frac{|\omega|}{2} \left( 1 - \text{sinc}(\pi|\omega|) \right),$$

d'où

$$\|\partial_t \varphi\|_{L^2(q)}^2 \geq \frac{|\omega|T}{2} \left( 1 - \text{sinc}(\pi|\omega|) \right) \|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2.$$

□

## 1.2 Contrôlabilité d'une équation des ondes semi-linéaire

### 1.2.1 Contrôlabilité d'une équation des ondes semi-linéaire avec contrôle interne

Soit  $\Omega$  un domaine borné non vide de  $\mathbb{R}^d$ , avec  $d \in \{1, 2, 3\}$ . Le bord  $\Gamma := \partial\Omega$  est supposé de classe  $C^2$ . Alternativement, on peut supposer  $\Omega$  convexe et  $\Gamma$  lipschitzien. Pour  $T > 0$ , on pose  $Q := \Omega \times (0, T)$  et  $\Sigma := \Gamma \times (0, T)$ . On se donne également  $\omega \subset \Omega$  un ouvert non vide et on note  $q := \omega \times (0, T)$ . Enfin, dans la suite,  $g : \mathbb{R} \rightarrow \mathbb{R}$  désigne une fonction de classe  $C^1$  telle que

$$|g(r)| \leq C(1 + |r|) \ln(2 + |r|), \quad \forall r \in \mathbb{R},$$

pour un certain  $C > 0$ . On considère le système d'état

$$\begin{cases} \partial_t^2 y - \Delta y + g(y) = u \mathbb{1}_{\omega} & \text{dans } Q, & (1.60a) \\ y = 0 & \text{sur } \Sigma, & (1.60b) \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega. & (1.60c) \end{cases}$$

La fonction  $\mathbb{1}_{\omega} : \Omega \rightarrow \{0, 1\}$  désigne la fonction caractéristique de  $\omega$ . Par souci de concision, on notera parfois  $\square := \partial_t^2 - \Delta$  et  $\mathbf{y}^0 = (y^0, y^1)$ . De même, on pourra noter  $y(x, t; \mathbf{y}^0, u)$  la solution de (1.60) avec  $\mathbf{y}^0$  comme donnée initiale et  $u$  comme contrôle.

*Remarque 38* (Régularité des solutions). On pose  $V := H_0^1(\Omega) \times L^2(\Omega)$ . Le système d'état (1.60) est bien posé. Pour  $(y^0, y^1) \in V$  et  $u \in L^2(q)$ , il existe une unique solution  $y$  de (1.60) vérifiant

$$(y, \partial_t y) \in C([0, T]; V).$$

**Définition 29** (Contrôlabilité exacte). Le système (1.60) est dit exactement contrôlable en temps  $T$  si quels que soient  $(y^0, y^1), (z^0, z^1) \in V$ , il existe  $u \in L^2(q)$  tel que la solution correspondante  $y$  de (1.60) vérifie

$$(y, \partial_t y)(\cdot, T) = (z^0, z^1) \quad \text{dans } \Omega.$$

Par ailleurs, on appelle paire état-contrôle tout couple  $(y, u)$  composé d'un contrôle exact  $u$  et de la solution contrôlée  $y$  de (1.60) associée.

**Théorème 3** (Contrôlabilité exacte). [35, Th. 1] & [15, Th. 2.2] Soient  $x_0 \in \mathbb{R}^d \setminus \bar{\Omega}$  et  $\varepsilon > 0$ . On note

$$\Gamma_0 := \left\{ x \in \Gamma; \quad (x - x_0) \cdot \nu(x) > 0 \right\} \quad \text{et} \quad \mathcal{O}_\varepsilon(\Gamma_0) := \left\{ x \in \mathbb{R}^d; \quad \text{dist}(x, \Gamma_0) \leq \varepsilon \right\}.$$

On suppose que

$$T > T^* := 2 \max_{x \in \bar{\Omega}} |x - x_0| \quad \text{et} \quad \omega \supseteq \mathcal{O}_\varepsilon(\Gamma_0) \cap \Omega. \quad (\mathbf{H}_0)$$

Si de plus la semi-linéarité  $g$  vérifie

$$\limsup_{|r| \rightarrow +\infty} \frac{|g(r)|}{|r| \ln^\alpha |r|} = 0, \quad \text{avec} \quad \begin{cases} \alpha = 2 & \text{si } d = 1, \\ \alpha = 1/2 & \text{si } d \in \{2, 3\}, \end{cases} \quad (\mathbf{H}_1)$$

alors le système (1.60) est exactement contrôlable en temps  $T$ .

*Démonstration.* On renvoie à [35, Th. 1] pour le cas  $d = 1$  et à [15, Th. 2.2] pour le cas  $d \in \{2, 3\}$ . Dans les deux cas, la preuve est basée sur un argument de point fixe de Schauder. On définit d'abord la fonction  $\hat{g} \in C(\mathbb{R})$  par

$$\hat{g}(r) := \begin{cases} \frac{g(r) - g(0)}{r} & \text{si } r \neq 0, \\ g'(0) & \text{si } r = 0. \end{cases}$$

Pour  $\xi \in L^\infty(0, T; L^d(\Omega))$ , on considère le système linéarisé

$$\begin{cases} \partial_t^2 y - \Delta y + \hat{g}(\xi)y = u \mathbb{1}_\omega - g(0) & \text{dans } Q, & (1.61a) \\ y = 0 & \text{sur } \Sigma, & (1.61b) \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega, & (1.61c) \end{cases}$$

dont on sait qu'il est exactement contrôlable en temps  $T$ . On introduit ensuite l'opérateur

$$K : \begin{array}{ccc} L^\infty(0, T; L^d(\Omega)) & \rightarrow & L^\infty(0, T; L^d(\Omega)) \\ \xi & \mapsto & y_\xi \end{array}, \quad (1.62)$$

où  $y_\xi$  est la solution contrôlée de (1.61) associée au contrôle  $u_\xi$  de norme  $L^2(q)$  minimale, i.e.  $(y_\xi, \partial_t y_\xi)(\cdot, T) = (z^0, z^1)$  dans  $\Omega$ . On remarque alors que tout point fixe  $y$  de  $K$  fournit un contrôle exact  $u$  pour le système (1.60). Or, on peut montrer que l'opérateur  $K$  est compact et qu'il admet un point fixe par le théorème du point fixe de Schauder.  $\square$

## 1.3 Contrôlabilité du système de l'élasticité linéaire

### 1.3.1 Contrôlabilité du système de l'élasticité linéaire avec contrôle au bord

Soit  $\Omega$  un domaine borné non vide de  $\mathbb{R}^d$ , avec  $d \in \{1, 2, 3\}$ . Le bord  $\Gamma := \partial\Omega$  est supposé de classe  $C^2$ . Alternativement, on peut supposer  $\Omega$  convexe et  $\Gamma$  lipschitzien. Pour  $T > 0$ , on pose  $Q := \Omega \times (0, T)$  et  $\Sigma := \Gamma \times (0, T)$ . On se donne également  $\gamma \subset \Gamma$  un ouvert non vide et on note  $\sigma := \gamma \times (0, T)$ . Pour  $\mu > 0$  et  $\lambda \geq -\mu$  fixés, on considère le système d'état

$$\begin{cases} \partial_t^2 y - \mu \Delta y - (\lambda + \mu) \nabla \text{div } y = 0 & \text{dans } Q, & (1.63a) \\ y = u \mathbb{1}_\gamma & \text{sur } \Sigma, & (1.63b) \\ (y, \partial_t y)(\cdot, 0) = (y^0, y^1) & \text{dans } \Omega, & (1.63c) \end{cases}$$

ainsi que le système adjoint

$$\begin{cases} \partial_t^2 \varphi - \mu \Delta \varphi - (\lambda + \mu) \nabla \operatorname{div} \varphi = 0 & \text{dans } Q, \\ \varphi = 0 & \text{sur } \Sigma, \\ (\varphi, \partial_t \varphi)(\cdot, 0) = (\varphi^0, \varphi^1) & \text{dans } \Omega. \end{cases} \quad \begin{array}{l} (1.64a) \\ (1.64b) \\ (1.64c) \end{array}$$

L'état s'écrit  $y = (y_1, \dots, y_d)^\top$ , tandis que le contrôle s'écrit  $u = (u_1, \dots, u_d)^\top$ . On suppose ici que le contrôle  $u$  peut agir sur la partie  $\gamma$  du bord. La fonction  $\mathbb{1}_\gamma : \Gamma \rightarrow \{0, 1\}$  désigne la fonction caractéristique de  $\gamma$ . Dans la suite, on utilise des caractères gras pour noter les espaces de fonctions à valeurs vectorielles. On écrira par exemple  $\mathbf{L}^2(\Omega) := L^2(\Omega; \mathbb{R}^d)$ .

*Remarque 39.* L'équation (1.63a), parfois nommée équation de Navier-Cauchy, modélise l'évolution du champ de déplacement  $y$  d'un matériau élastique occupant le domaine  $\Omega$ . Les constantes  $\mu$  et  $\lambda$  sont appelées les coefficients de Lamé et décrivent les propriétés élastiques d'un matériau homogène et isotrope.

*Remarque 40 (Régularité des solutions).* On pose  $\mathbf{Y} := \mathbf{L}^2(\Omega) \times \mathbf{H}^{-1}(\Omega)$  et  $\mathbf{Z} := \mathbf{H}_0^1(\Omega) \times \mathbf{L}^2(\Omega)$ . Le système d'état (1.63) et le système adjoint (1.64) sont bien posés.

▷ Pour  $(y^0, y^1) \in \mathbf{Y}$  et  $u \in \mathbf{L}^2(\sigma)$ , il existe une unique solution  $y$  de (1.63) vérifiant

$$(y, \partial_t y) \in C([0, T]; \mathbf{Y}).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|y\|_{L^\infty(\mathbf{L}^2)} + \|\partial_t y\|_{L^\infty(\mathbf{H}^{-1})} \leq C_{\Omega, T} \left( \|(y^0, y^1)\|_{\mathbf{L}^2 \times \mathbf{H}^{-1}} + \|u\|_{\mathbf{L}^2(\sigma)} \right).$$

▷ Pour  $(\varphi^0, \varphi^1) \in \mathbf{Z}$ , il existe une unique solution  $\varphi$  de (1.64) vérifiant

$$(\varphi, \partial_t \varphi) \in C([0, T]; \mathbf{Z}).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|\varphi\|_{L^\infty(\mathbf{H}_0^1)} + \|\partial_t \varphi\|_{L^\infty(\mathbf{L}^2)} \leq C_{\Omega, T} \|(\varphi^0, \varphi^1)\|_{\mathbf{H}_0^1 \times \mathbf{L}^2}.$$

**Définition 30** (Contrôlabilité à zéro). Le système (1.63) est dit contrôlable à zéro en temps  $T$  si quel que soit  $(y^0, y^1) \in \mathbf{Y}$ , il existe  $u \in \mathbf{L}^2(\sigma)$  tel que la solution correspondante  $y$  de (1.63) vérifie

$$(y, \partial_t y)(\cdot, T) = (0, 0) \quad \text{dans } \Omega.$$

**Proposition 34** (Inégalité directe). [25, Sect. IV.1] L'assertion suivante est vérifiée.

▷ Soient  $(\varphi^0, \varphi^1) \in \mathbf{Z}$  et  $\varphi$  la solution de (1.64) correspondante. Il existe  $C_{\Omega, T} > 0$  tel que

$$\int_\sigma \left( \mu |\partial_\nu \varphi|^2 + (\lambda + \mu) |\operatorname{div} \varphi|^2 \right) \leq C_{\Omega, T} \left( \mu \|\varphi^0\|_{\mathbf{H}_0^1}^2 + (\lambda + \mu) \|\operatorname{div} \varphi^0\|_{\mathbf{L}^2}^2 + \|\varphi^1\|_{\mathbf{L}^2}^2 \right).$$

**Proposition 35** (Inégalité d'observabilité). Par dualité entre contrôlabilité et observabilité, l'assertion suivante est vérifiée.

▷ Le système (1.63) est contrôlable à zéro en temps  $T$  si et seulement s'il existe  $C > 0$  tel que

$$\mu \|\varphi^0\|_{\mathbf{H}_0^1}^2 + (\lambda + \mu) \|\operatorname{div} \varphi^0\|_{\mathbf{L}^2}^2 + \|\varphi^1\|_{\mathbf{L}^2}^2 \leq C \int_\sigma \left( \mu |\partial_\nu \varphi|^2 + (\lambda + \mu) |\operatorname{div} \varphi|^2 \right), \quad \forall (\varphi^0, \varphi^1) \in \mathbf{Z}, \quad (1.65)$$

où  $\varphi$  est la solution de (1.64) associée à  $(\varphi^0, \varphi^1)$ .

**Proposition 36** (Contrôle de norme minimale). En appliquant la méthode HUM, l'assertion suivante est vérifiée.

▷ On suppose que l'inégalité d'observabilité (1.65) est vérifiée. Pour  $(y^0, y^1) \in \mathbf{Y}$ , le contrôle de norme  $\mathbf{L}^2(\sigma)$  minimale pour le système (1.63) est donné par

$$\hat{u} = -(\mu \partial_\nu \hat{\varphi} + (\lambda + \mu) \operatorname{div}(\hat{\varphi}) \nu)|_\sigma,$$

avec  $\hat{\varphi}$  la solution de (1.64) associée au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) := \frac{1}{2} \int_\sigma \left| \mu \partial_\nu \varphi + (\lambda + \mu) \operatorname{div}(\varphi) \nu \right|^2 - \langle y^0, \varphi^1 \rangle_{\mathbf{L}^2} + \langle y^1, \varphi^0 \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1}, \quad \forall (\varphi^0, \varphi^1) \in \mathbf{Z},$$

où  $\varphi$  est la solution de (1.64) associée à  $(\varphi^0, \varphi^1)$ .

À la proposition suivante, on établit l'inégalité d'observabilité (1.65) pour un bord de contrôle  $\gamma \subset \Gamma$  défini par une condition géométrique, ainsi que pour  $T$  assez grand. La démonstration repose sur la méthode des multiplicateurs. Plus précisément, si  $x_0$  ci-dessous est pris dans  $\mathbb{R}^d \setminus \overline{\Omega}$ , le bord de contrôle  $\gamma = \Gamma_0$  correspond alors au complémentaire dans  $\Gamma$  de la partie du bord éclairée depuis  $x_0$ .

**Proposition 37** (Inégalité d'observabilité). [25, Sect. IV.1] Pour  $x_0 \in \mathbb{R}^d$  fixé, on définit le champ de vecteurs  $\theta \in [C^1(\overline{\Omega})]^d$  par  $\theta(x) = x - x_0$ . On introduit également la partie du bord

$$\Gamma_0 := \left\{ x \in \Gamma; \quad \theta(x) \cdot \nu(x) > 0 \right\}.$$

On pose  $R_0 := \max_{x \in \overline{\Omega}} |\theta(x)|$  et  $T^* := \frac{2}{\sqrt{\mu}} R_0$ . Pour tout  $T > T^*$ , l'inégalité suivante est vérifiée,

$$\mu \|\varphi^0\|_{\mathbf{H}_0^1}^2 + (\lambda + \mu) \|\operatorname{div} \varphi^0\|_{L^2}^2 + \|\varphi^1\|_{L^2}^2 \leq \frac{R_0}{T - T^*} \int_{\Sigma_0} \left( \mu |\partial_\nu \varphi|^2 + (\lambda + \mu) |\operatorname{div} \varphi|^2 \right), \quad \forall (\varphi^0, \varphi^1) \in \mathbf{Z},$$

où  $\varphi$  est la solution de (1.64) associée à  $(\varphi^0, \varphi^1)$ .

### 1.3.2 Système de l'élasticité linéaire sous forme vitesse-contrainte

Le système adjoint (1.64) peut être vu comme un système d'ordre un portant sur le déplacement  $\varphi$  et la vitesse  $\partial_t \varphi$ . Avec le changement de variables

$$w := \partial_t \varphi \quad \text{et} \quad \mathbf{q} := \nabla \varphi,$$

où  $\nabla \varphi$  désigne la matrice jacobienne de  $\varphi$ , on peut réécrire le système de l'élasticité sous forme vitesse-contrainte. En effet, en notant  $\mathbf{I}$  la matrice identité et

$$\boldsymbol{\sigma}(\mathbf{q}) := \mu \mathbf{q} + (\lambda + \mu) \operatorname{tr}(\mathbf{q}) \mathbf{I}$$

le tenseur des contraintes, le système (1.64) devient

$$\begin{cases} \partial_t w - \operatorname{div} \boldsymbol{\sigma}(\mathbf{q}) = 0 & \text{dans } Q, & (1.66a) \\ \partial_t \mathbf{q} - \nabla w = \mathbf{0} & \text{dans } Q, & (1.66b) \\ w = 0 & \text{sur } \Sigma, & (1.66c) \\ (w, \mathbf{q})(\cdot, 0) = (w^0, \mathbf{q}^0) & \text{dans } \Omega. & (1.66d) \end{cases}$$

Dans la suite, on note

$$\mathcal{M}(w, \mathbf{q}) := \left( \partial_t w - \operatorname{div} \boldsymbol{\sigma}(\mathbf{q}), \partial_t \mathbf{q} - \nabla w \right).$$

On définit l'opérateur  $\mathcal{A}_0 : \mathbf{H}_0^1(\Omega) \rightarrow \mathbf{H}^{-1}(\Omega)$  par  $\mathcal{A}_0 := -\mu \Delta - (\lambda + \mu) \nabla \operatorname{div}$ . Avec le changement de variables

$$v := y \quad \text{et} \quad \mathbf{p} := -\nabla \mathcal{A}_0^{-1} \partial_t y,$$

le système d'état (1.63) devient quant à lui

$$\begin{cases} \partial_t v - \operatorname{div} \boldsymbol{\sigma}(\mathbf{p}) = 0 & \text{dans } Q, & (1.67a) \\ \partial_t \mathbf{p} - \nabla v = \mathbf{0} & \text{dans } Q, & (1.67b) \\ v = u \mathbb{1}_\gamma & \text{sur } \Sigma, & (1.67c) \\ (v, \mathbf{p})(\cdot, 0) = (v^0, \mathbf{p}^0) & \text{dans } \Omega, & (1.67d) \end{cases}$$

avec  $v^0 = y^0$  et  $\mathbf{p}^0 = -\nabla \mathcal{A}_0^{-1} y^1$ .

*Remarque 41* (Régularité des solutions). On pose  $\mathbf{W} := \mathbf{L}^2(\Omega) \times \nabla \mathbf{H}_0^1(\Omega)$ , avec  $\nabla \mathbf{H}_0^1(\Omega) := \{ \nabla \varphi; \varphi \in \mathbf{H}_0^1(\Omega) \}$ . Le système d'état (1.67) et le système adjoint (1.66) sont bien posés.

▷ Pour  $(v^0, \mathbf{p}^0) \in \mathbf{W}$  et  $u \in \mathbf{L}^2(\sigma)$ , il existe une unique solution  $(v, \mathbf{p})$  de (1.67) vérifiant

$$(v, \mathbf{p}) \in C([0, T]; \mathbf{W}).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|(v, \mathbf{p})\|_{L^\infty(\mathbf{W})} \leq C_{\Omega, T} \left( \|(v^0, \mathbf{p}^0)\|_{L^\infty(\mathbf{W})} + \|u\|_{L^2(\sigma)} \right).$$

▷ Pour  $(w^0, \mathbf{q}^0) \in \mathbf{W}$ , il existe une unique solution  $(w, \mathbf{q})$  de (1.66) vérifiant

$$(w, \mathbf{q}) \in C([0, T]; \mathbf{W}).$$

De plus, il existe  $C_{\Omega, T} > 0$  tel que

$$\|(w, \mathbf{q})\|_{L^\infty(\mathbf{W})} \leq C_{\Omega, T} \|(w^0, \mathbf{q}^0)\|_{L^\infty(\mathbf{W})}.$$

## 1.4 Compléments à l'Article A

### 1.4.1 Version alternative de la preuve du Théorème 2.1

Dans cette section, on donne une version alternative de la preuve du Th. 2.1 de l'Art. A, dont on rappelle l'énoncé au Th. 4 ci-dessous. On se place dans le cas de l'équation des ondes linéaire 1D avec contrôle interne. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . On considère le système d'état (1.19) et on suppose que le contrôle peut agir dans un ouvert  $q \subset Q = \Omega \times (0, T)$  non cylindrique.

Pour  $(y^0, y^1) \in H_0^1(\Omega) \times L^2(\Omega)$ , on sait que la contrôlabilité à zéro de (1.19) est équivalente à l'inégalité d'observabilité (1.21) pour le système adjoint (1.20). Dans ce cadre, au Th. 4, on démontre que la constante  $C > 0$  dans l'inégalité d'observabilité (1.21) peut être choisie uniformément par rapport à l'ouvert de contrôle  $q$  pris dans un ensemble admissible  $\mathcal{Q}_{\text{ad}}^\varepsilon$  défini via une condition géométrique dérivée du critère (GCC) usuel. On renvoie à la Déf. 23 pour la définition du critère (GCC) et à la Déf. 22 pour la définition des lignes caractéristiques  $C_{x_0}^\pm$ .

*Remarque 42.* La preuve ci-dessous utilise les mêmes idées que celles de la preuve de l'Art. A. À titre de comparaison, on rappelle brièvement les éléments de la preuve de l'Art. A.

- ▷ Pour  $N > 1/\varepsilon$  fixé, on introduit la subdivision régulière  $S_N$  de  $\bar{\Omega}$ .
- ▷ À partir de la subdivision  $S_N$ , on construit un sous-espace  $V_N \subset V$  de dimension finie.
- ▷ À partir de la subdivision  $S_N$ , pour tout  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , on construit une "approximation"  $q_N \subset q$  vérifiant le critère (GCC).
- ▷ Pour  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , on démontre l'inégalité (1.21) avec  $(\varphi^0, \varphi^1) \in V_N$ . Pour ce faire, on associe à  $q$  un graphe  $G(q_N)$  et on exprime  $C_{\text{obs}}(q)$  à l'aide de la connectivité algébrique de  $G(q_N)$ .
- ▷ Par un argument de finitude, on en déduit l'inégalité uniforme (1.68) avec  $(\varphi^0, \varphi^1) \in V_N$ .
- ▷ On répète les étapes précédentes pour la sous-subdivision  $S_{pN}$ , avec  $p \geq 1$ .
- ▷ En passant à la limite  $p \rightarrow +\infty$ , on en déduit l'inégalité uniforme (1.68) avec  $(\varphi^0, \varphi^1) \in V$ .

Pour  $\varepsilon > 0$ , on introduit l'ensemble admissible

$$\mathcal{Q}_{\text{ad}}^\varepsilon := \left\{ q \subset Q \text{ ouvert; } \text{Int}_\varepsilon(q) \text{ vérifie (GCC)} \right\},$$

où

$$\text{Int}_\varepsilon(q) = \left\{ (x, t) \in q; \text{dist}((x, t), \partial q) > \varepsilon \right\}$$

désigne le  $\varepsilon$ -intérieur de  $q$ .

*Remarque 43.* Pour  $\varepsilon > 0$ , l'introduction de l'ensemble admissible  $\mathcal{Q}_{\text{ad}}^\varepsilon$  est motivée par les remarques suivantes.

- ▷ Comme  $\text{Int}_\varepsilon(q) \subset q$ , tout ouvert de contrôle  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  vérifie (GCC).
- ▷ Pour  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , toute ligne caractéristique  $C_{x_0}^\pm$  rencontre  $q$  sur une longueur supérieure à  $\varepsilon$ , i.e.

$$\mathcal{H}^1(q \cap C_{x_0}^\pm) \geq \varepsilon, \quad \forall x_0 \in \bar{\Omega},$$

où  $\mathcal{H}^1$  est la mesure de Hausdorff de dimension 1.

**Théorème 4** (Inégalité d'observabilité uniforme). *Pour  $\varepsilon > 0$  fixé, il existe  $\bar{C}_{\text{obs}}^\varepsilon > 0$  tel que pour tout  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ ,*

$$\|(\varphi^0, \varphi^1)\|_V^2 \leq \bar{C}_{\text{obs}}^\varepsilon \|\partial_t \varphi\|_{L^2(q)}^2, \quad \forall (\varphi^0, \varphi^1) \in V := H_0^1(\Omega) \times L^2(\Omega), \quad (1.68)$$

où  $\varphi$  est la solution de (1.20) associée à  $(\varphi^0, \varphi^1)$ .

**Preuve alternative** Soit  $\varepsilon > 0$  et soit  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . On commence par reformuler le problème. On définit la forme bilinéaire symétrique positive

$$\mathfrak{F}(\varphi^0, \bar{\varphi}^0) := \iint_q \partial_t \varphi \partial_t \bar{\varphi}, \quad \forall \varphi^0, \bar{\varphi}^0 \in V,$$

où  $\varphi$  et  $\bar{\varphi}$  sont respectivement les solutions de (1.20) associées à  $\varphi^0$  et  $\bar{\varphi}^0$ . Ainsi, montrer l'inégalité d'observabilité uniforme (1.68) revient à montrer qu'il existe une constante  $\bar{\mathfrak{C}}_\varepsilon > 0$  indépendante de l'ouvert de contrôle  $q$  et telle que

$$\mathfrak{F}(\varphi^0, \varphi^0) \geq \bar{\mathfrak{C}}_\varepsilon \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V.$$

On procède ensuite en plusieurs étapes.

- 1) On fixe  $N \in \mathbb{N}$  tel que  $2^N > 1/\varepsilon$ . À l'aide d'une subdivision  $S_N$  de  $\bar{\Omega}$ , on construit un ouvert de contrôle  $q_N$  tel que

$$\text{Int}_\varepsilon(q) \subset q_N \subset q.$$

On définit alors une forme bilinéaire symétrique positive  $\mathfrak{F}_N$  qui minore  $\mathfrak{F}$ , i.e.

$$\mathfrak{F}(\varphi^0, \varphi^0) \geq \mathfrak{F}_N(\varphi^0, \varphi^0), \quad \forall \varphi^0 \in V.$$

- 2) On construit une base orthonormée  $\mathcal{B} = (\psi_n^0)_{n \in \mathbb{N}}$  de  $V$  qui est également orthogonale pour  $\mathfrak{F}_N$  à partir du rang  $N$ , i.e.

$$\mathfrak{F}_N(\psi_n^0, \psi_m^0) = 0, \quad \forall n \geq N, \forall m \in \mathbb{N}, m \neq n.$$

On pose  $\mathcal{B}_N = (\psi_n^0)_{0 \leq n < N}$  et  $V_N = \text{Vect}(\mathcal{B}_N)$ . On réduit alors le problème à l'existence d'une constante  $\bar{\mathfrak{C}}_N > 0$  indépendante de l'ouvert de contrôle  $q$  et telle que

$$\mathfrak{F}_N(\varphi^0, \varphi^0) \geq \bar{\mathfrak{C}}_N \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V_N.$$

- 3) On conclut en utilisant le fait que l'ouvert de contrôle  $q_N$  vérifie (GCC) et que l'ensemble  $\{q_N; q \in \mathcal{Q}_{\text{ad}}^\varepsilon\}$  est fini.

**Étape 1** Soit  $N \in \mathbb{N}$  tel que  $2^N > 1/\varepsilon$ . On note  $h := 1/2^N$  et  $S_N := (x_k)_{0 \leq k \leq 2^N}$  la subdivision régulière de  $\bar{\Omega}$  en  $2^N$  intervalles, i.e.  $x_k = kh$ . En étendant la définition de  $x_k$  à  $k \in \mathbb{Z}$ , on pose

$$I_k := \begin{cases} [x_{k-1}, x_k] & \text{si } k > 0, \\ [x_k, x_{k+1}] & \text{si } k < 0, \end{cases} \quad \forall k \in \mathbb{Z}^*.$$

Pour  $k \in \mathbb{Z}^*$ , le milieu  $x_k^c$  du segment  $I_k$  est donné par

$$x_k^c := \begin{cases} \frac{x_{k-1} + x_k}{2} & \text{si } k > 0, \\ \frac{x_k + x_{k+1}}{2} & \text{si } k < 0. \end{cases}$$

Dans la suite, on note  $\mathcal{I}_N := \llbracket -2^N, 2^N \rrbracket \setminus \{0\}$  et on définit l'application  $\mathbf{n} : \mathbb{Z}^* \rightarrow \mathcal{I}_N$  par

$$\mathbf{n}(i) = \mathbf{i}, \quad \forall i \in \mathbb{Z}^*, \quad \text{où } \mathbf{i} \in \mathcal{I}_N \text{ vérifie } \mathbf{m}(x_i^c) = x_i^c.$$

**Définition 31.** Pour  $i, j \in \mathbb{Z}^*$ , on appelle carré élémentaire d'indice  $(i, j)$  l'ensemble  $C_{i,j}$  défini par

$$C_{i,j} := \left\{ (x, t) \in \mathbb{R}^2; \quad x - t \in I_i, \quad x + t \in I_j \right\}.$$

On note  $\mathcal{C}_N := \{C_{i,j}; i, j \in \mathbb{Z}^*\}$  l'ensemble des carrés élémentaires associés à la subdivision  $S_N$ . On note également  $\mathcal{C}_N(q)$  l'ensemble des carrés élémentaires dont l'intérieur est inclus dans  $q$ , i.e.

$$\mathcal{C}_N(q) := \left\{ C_{i,j} \in \mathcal{C}_N; \quad \overset{\circ}{C}_{i,j} \subset q \right\}.$$

Enfin, on définit l'ouvert de contrôle

$$q_N := \overbrace{\bigcup_{C_{i,j} \in \mathcal{C}_N(q)} C_{i,j}}^{\circ}.$$

**Lemme 11.** La condition  $2^N > 1/\varepsilon$  et la définition de  $q_N$  impliquent les inclusions

$$\text{Int}_\varepsilon(q) \subset q_N \subset q.$$

On en déduit que l'ouvert de contrôle  $q_N$  vérifie (GCC).

*Démonstration.* Soit  $z = (x, t) \in \text{Int}_\varepsilon(q)$ . Par définition, on a  $z \in q$  et  $\text{dist}(z, \partial q) > \varepsilon$ . Puisque les carrés élémentaires introduits à la Déf. 31 recouvrent  $\mathbb{R}^2$ , il existe  $i, j \in \mathbb{Z}^*$  tels que  $z \in C_{i,j}$ . En outre, comme  $\text{diam}(C_{i,j}) = h$ , on a  $C_{i,j} \subset \overline{B}(z, h)$  la boule fermée de centre  $z$  et de rayon  $h$ . Soit à présent  $\tilde{z} = (\tilde{x}, \tilde{t}) \in \overline{B}(z, h)$ . Pour tout  $\bar{z} = (\bar{x}, \bar{t}) \in \mathbb{R}^2 \setminus q$ , il vient

$$\text{dist}(\tilde{z}, \bar{z}) \geq |\text{dist}(\tilde{z}, z) - \text{dist}(z, \bar{z})| > \varepsilon - h > 0.$$

On obtient alors  $\text{dist}(\tilde{z}, \mathbb{R}^2 \setminus q) > 0$ , d'où  $\tilde{z} \in q$ . Par suite, on a  $C_{i,j} \subset \overline{B}(z, h) \subset q$  et enfin  $C_{i,j} \in \mathcal{C}_N(q)$ .  $\square$

**Définition 32.** On définit la forme bilinéaire symétrique positive

$$\mathfrak{F}_N(\varphi^0, \bar{\varphi}^0) := \iint_{q_N} \partial_t \varphi \partial_t \bar{\varphi} = \sum_{C_{i,j} \in \mathcal{C}_N(q)} \iint_{C_{i,j}} \partial_t \varphi \partial_t \bar{\varphi}, \quad \forall \varphi^0, \bar{\varphi}^0 \in V,$$

où  $\varphi$  et  $\bar{\varphi}$  sont respectivement les solutions de (1.20) associées à  $\varphi^0$  et  $\bar{\varphi}^0$ . Comme  $q_N \subset q$ , on a clairement

$$\mathfrak{F}(\varphi^0, \varphi^0) \geq \mathfrak{F}_N(\varphi^0, \varphi^0), \quad \forall \varphi^0 \in V.$$

**Étape 2** On introduit ici la base  $\mathcal{B}_{\text{Haar}}$  des ondelettes de Haar. Il s'agit d'une base hilbertienne de  $L^2(\Omega)$ . On définit la fonction mère  $\psi = \mathbb{1}_{[0,1/2]} - \mathbb{1}_{[1/2,1]}$ , ainsi que la fonction d'échelle  $\psi_0 = \mathbb{1}_{[0,1]}$ . Pour  $n \in \mathbb{N}$  et  $k \in \llbracket 1, 2^n \rrbracket$ , on pose

$$\psi_{n,k}(x) = 2^{n/2} \psi(2^n x - k + 1).$$

La base des ondelettes de Haar est alors donnée par

$$\mathcal{B}_{\text{Haar}} := \left\{ \psi_0, \psi_{n,k}, n \in \mathbb{N}, k \in \llbracket 1, 2^n \rrbracket \right\}.$$

**Définition 33.** On définit le couple  $\boldsymbol{\psi}_0^0 = (\psi_0^0, \psi_0^1) \in V$  par  $\psi_0^0 \equiv 0$  et  $\psi_0^1 = \psi_0$ . De même, pour  $n \in \mathbb{N}$  et  $k \in \llbracket 1, 2^n \rrbracket$ , on définit les couples  $\boldsymbol{\psi}_{n,k}^{0,+} = (\psi_{n,k}^{0,+}, \psi_{n,k}^{1,+}) \in V$  et  $\boldsymbol{\psi}_{n,k}^{0,-} = (\psi_{n,k}^{0,-}, \psi_{n,k}^{1,-}) \in V$  par

$$\psi_{n,k}^{0,\pm}(x) = \frac{1}{\sqrt{2}} \int_0^x \psi_{n,k}(y) dy \quad \text{et} \quad \psi_{n,k}^{1,\pm}(x) = \mp \frac{1}{\sqrt{2}} \psi_{n,k}(x).$$

On pose alors

$$\mathcal{B} := \left\{ \boldsymbol{\psi}_0^0, \boldsymbol{\psi}_{n,k}^{0,\pm}, n \in \mathbb{N}, k \in \llbracket 1, 2^n \rrbracket \right\}.$$

On vérifie aisément que  $\mathcal{B}$  est une base hilbertienne de  $V$ .

**Lemme 12.** La base  $\mathcal{B}$  est orthogonale pour  $\mathfrak{F}_N$  à partir du rang  $N$ . Plus précisément, pour  $n \geq N$  et  $k \in \llbracket 1, 2^n \rrbracket$ , on a

$$\mathfrak{F}_N(\boldsymbol{\psi}_{n,k}^{0,\pm}, \bar{\boldsymbol{\psi}}^0) = 0, \quad \forall \bar{\boldsymbol{\psi}}^0 \in \mathcal{B}, \bar{\boldsymbol{\psi}}^0 \neq \boldsymbol{\psi}_{n,k}^{0,\pm}.$$

*Démonstration.* Soient  $n_1 \geq N$  et  $k_1 \in \llbracket 1, 2^{n_1} \rrbracket$ . Il existe alors  $i_0 \in \llbracket 1, 2^N \rrbracket$  tel que  $\text{Supp } \boldsymbol{\psi}_{n_1, k_1}^{0,\pm} \subset I_{i_0}$ . Soient également  $n_2 \in \mathbb{N}$  et  $k_2 \in \llbracket 1, 2^{n_2} \rrbracket$  tels que  $(n_2, k_2) \neq (n_1, k_1)$ . On note respectivement  $\psi_1^\pm$  et  $\psi_2^\pm$  les solutions de (1.20) associées à  $\boldsymbol{\psi}_{n_1, k_1}^{0,\pm}$  et  $\boldsymbol{\psi}_{n_2, k_2}^{0,\pm}$ . On considère  $i, j \in \mathbb{Z}^*$  tels que  $C_{i,j} \in \mathcal{C}_N(q)$ . On note  $\mathbf{i} = \mathbf{n}(i) \in \mathcal{J}_N$  et  $\mathbf{j} = \mathbf{n}(j) \in \mathcal{J}_N$ , ainsi que  $s_i = \text{sgn}(i)$  et  $s_j = \text{sgn}(j)$ .

On montre ici que  $\iint_{C_{i,j}} \partial_t \psi_1^+ \partial_t \psi_2^+ = 0$ . Les trois autres cas (pour  $\psi_1^\pm$  et  $\psi_2^\pm$ ) se traitent de façon similaire.

Pour  $(\varphi^0, \varphi^1) \in V$ , on note  $\varphi$  la solution correspondante de (1.20). D'après la formule de d'Alembert (voir Prop. 15), pour tout  $(x, t) \in C_{i,j}$ , on a

$$\partial_t \varphi(x, t) = \frac{1}{2} \left( (\varphi^0)'(|r_j|) + s_j \varphi^1(|r_j|) \right) - \frac{1}{2} \left( (\varphi^0)'(|r_i|) - s_i \varphi^1(|r_i|) \right), \quad \text{où} \quad \begin{cases} r_i = \mathbf{m}(x - t) \in I_i, \\ r_j = \mathbf{m}(x + t) \in I_j. \end{cases}$$



Par suite, pour tout  $(x, t) \in C_{i,j}$ , on obtient

$$\partial_t \psi_1^+(x, t) = \frac{1-s_j}{2\sqrt{2}} \psi_{n_1, k_1}(|r_j|) - \frac{1+s_i}{2\sqrt{2}} \psi_{n_1, k_1}(|r_i|), \quad \text{où} \quad \begin{cases} r_i = \mathbf{m}(x-t) \in I_i, \\ r_j = \mathbf{m}(x+t) \in I_j. \end{cases}$$

On remarque alors que les seules configurations dans lesquelles  $\partial_t \psi_1^+ \neq 0$  sur  $C_{i,j}$  sont

$$+i = i_0 \quad \text{ou} \quad -j = i_0.$$

On traite ici le cas  $+i = i_0$  et  $-j \neq i_0$ ,  $j < 0$ . Les autres cas se traitent de façon similaire. Dans le cas présent, pour tout  $(x, t) \in C_{i,j}$ , on a

$$\begin{cases} \partial_t \psi_1^+(x, t) = -\frac{1}{\sqrt{2}} \psi_{n_1, k_1}(r_i), \\ \partial_t \psi_2^+(x, t) = \frac{1}{\sqrt{2}} \psi_{n_2, k_2}(-r_j) - \frac{1}{\sqrt{2}} \psi_{n_2, k_2}(r_i), \end{cases} \quad \text{où} \quad \begin{cases} r_i = \mathbf{m}(x-t) \in I_i, \\ r_j = \mathbf{m}(x+t) \in I_j. \end{cases}$$

À l'aide d'un changement de variables, on en déduit finalement que

$$\begin{aligned} \iint_{C_{i,j}} \partial_t \psi_1^+ \partial_t \psi_2^+ &= \frac{1}{4} \int_{I_i} \int_{I_j} \psi_{n_1, k_1}(r_i) (\psi_{n_2, k_2}(r_i) - \psi_{n_2, k_2}(-r_j)) \, dr_j \, dr_i \\ &= \frac{h}{4} \int_{I_{|i|}} \psi_{n_1, k_1} \psi_{n_2, k_2} - \frac{1}{4} \int_{I_{|i|}} \psi_{n_1, k_1} \int_{I_{|j|}} \psi_{n_2, k_2} = 0. \end{aligned}$$

En effet, comme  $\text{Supp } \psi_{n_1, k_1} \subset I_{i_0} = I_{|i|}$ , on a

$$\int_{I_{|i|}} \psi_{n_1, k_1} \psi_{n_2, k_2} = \int_{\Omega} \psi_{n_1, k_1} \psi_{n_2, k_2} = 0 \quad \text{et} \quad \int_{I_{|i|}} \psi_{n_1, k_1} = \int_{\Omega} \psi_{n_1, k_1} = 0.$$

□

**Lemme 13.** *Pour  $n \geq N$  et  $k \in \llbracket 1, 2^n \rrbracket$ , on a*

$$\mathfrak{F}_N(\psi_{n,k}^{0,\pm}, \psi_{n,k}^{0,\pm}) \geq \frac{h}{4} \|\psi_{n,k}^{0,\pm}\|_V^2.$$

*Démonstration.* Soient  $n \geq N$  et  $k \in \llbracket 1, 2^n \rrbracket$ . Il existe alors  $i_0 \in \llbracket 1, 2^n \rrbracket$  tel que  $\text{Supp } \psi_{n,k}^{0,\pm} \subset I_{i_0}$ . On note  $\psi^\pm$  la solution de (1.20) associée à  $\psi_{n,k}^{0,\pm}$ . On montre ici que  $\mathfrak{F}_N(\psi_{n,k}^{0,+}, \psi_{n,k}^{0,+}) \geq \frac{h}{4} \|\psi_{n,k}^{0,+}\|_V^2 = h/4$ . L'autre cas (pour  $\psi_{n,k}^{0,-}$ ) se traite de façon similaire. Comme  $q_N$  vérifie (GCC), il existe  $C_{i,j} \in \mathcal{C}_N(q)$  tel que

$$+\mathbf{n}(i) = i_0 \quad \text{ou} \quad -\mathbf{n}(j) = i_0.$$

On note  $\mathbf{i} = \mathbf{n}(i) \in \mathfrak{I}_N$  et  $\mathbf{j} = \mathbf{n}(j) \in \mathfrak{I}_N$ , ainsi que  $s_i = \text{sgn}(i)$  et  $s_j = \text{sgn}(j)$ . On traite ici le cas  $+i = i_0$  et  $-j \neq i_0$ . Les autres cas se traitent de façon similaire. En raisonnant comme au lemme précédent, pour tout  $(x, t) \in C_{i,j}$ , on a

$$\partial_t \psi^+(x, t) = -\frac{1}{\sqrt{2}} \psi_{n,k}(r_i), \quad \text{où} \quad \begin{cases} r_i = \mathbf{m}(x-t) \in I_i, \\ r_j = \mathbf{m}(x+t) \in I_j. \end{cases}$$

À l'aide d'un changement de variables, on en déduit finalement que

$$\mathfrak{F}_N(\psi_{n,k}^{0,+}, \psi_{n,k}^{0,+}) \geq \iint_{C_{i,j}} |\partial_t \psi^+|^2 = \frac{1}{4} \int_{I_i} \int_{I_j} |\psi_{n,k}(r_i)|^2 \, dr_j \, dr_i = \frac{h}{4} \int_{I_{|i|}} |\psi_{n,k}|^2 = \frac{h}{4}.$$

En effet, comme  $\text{Supp } \psi_{n,k} \subset I_{i_0} = I_{|i|}$ , on a

$$\int_{I_{|i|}} |\psi_{n,k}|^2 = \int_{\Omega} |\psi_{n,k}|^2 = 1.$$

□

*Remarque 44.* Dans le Lem. 13, la constante  $h/4$  est indépendante de l'ouvert de contrôle  $q$ . Par ailleurs, pour  $n \geq N$  et  $k \in \llbracket 1, 2^n \rrbracket$ , l'inégalité optimale s'écrit

$$\mathfrak{F}_N(\psi_{n,k}^{0,\pm}, \psi_{n,k}^{0,\pm}) \geq d_{i_0}^\pm(q) \frac{h}{4} \|\psi_{n,k}^{0,\pm}\|_V^2,$$

où  $i_0 \in \llbracket 1, 2^N \rrbracket$  vérifie  $\text{Supp } \psi_{n,k}^{0,\pm} \subset I_{i_0}$ , et où

$$d_{i_0}^\pm(q) := \text{Card} \left\{ C_{i,j} \in \mathcal{C}_N(q); \quad \pm \mathbf{n}(i) = i_0 \text{ ou } \mp \mathbf{n}(j) = i_0 \right\}.$$

On considère ensuite la sous-famille

$$\mathcal{B}_N := \left\{ \psi_0^0, \psi_{n,k}^{0,\pm}, n < N, k \in \llbracket 1, 2^n \rrbracket \right\} \subset \mathcal{B},$$

ainsi que le sous-espace vectoriel de dimension finie  $V_N := \text{Vect}(\mathcal{B}_N) \subset V$ . Pour tout  $\varphi^0 \in V$ , on peut alors décomposer

$$\varphi^0 = \varphi_N^0 + \sum_{n=N}^{+\infty} \sum_{k=1}^{2^n} \left( \alpha_{n,k}^+ \psi_{n,k}^{0,+} + \alpha_{n,k}^- \psi_{n,k}^{0,-} \right), \quad \text{avec } \varphi_N^0 \in V_N.$$

En utilisant le Lem. 12, il vient

$$\mathfrak{F}_N(\varphi^0, \varphi^0) = \mathfrak{F}_N(\varphi_N^0, \varphi_N^0) + \sum_{n=N}^{+\infty} \sum_{k=1}^{2^n} \left( |\alpha_{n,k}^+|^2 \mathfrak{F}_N(\psi_{n,k}^{0,+}, \psi_{n,k}^{0,+}) + |\alpha_{n,k}^-|^2 \mathfrak{F}_N(\psi_{n,k}^{0,-}, \psi_{n,k}^{0,-}) \right).$$

En outre, en utilisant le Lem. 13, on a la minoration

$$\sum_{n=N}^{+\infty} \sum_{k=1}^{2^n} \left( |\alpha_{n,k}^+|^2 \mathfrak{F}_N(\psi_{n,k}^{0,+}, \psi_{n,k}^{0,+}) + |\alpha_{n,k}^-|^2 \mathfrak{F}_N(\psi_{n,k}^{0,-}, \psi_{n,k}^{0,-}) \right) \geq \frac{h}{4} \sum_{n=N}^{+\infty} \sum_{k=1}^{2^n} \left( |\alpha_{n,k}^+|^2 + |\alpha_{n,k}^-|^2 \right).$$

On constate ainsi que notre problème de départ se réduit à montrer l'existence d'une constante  $\bar{\mathfrak{C}}_N > 0$  indépendante de l'ouvert de contrôle  $q$  et telle que

$$\mathfrak{F}_N(\varphi_N^0, \varphi_N^0) \geq \bar{\mathfrak{C}}_N \|\varphi_N^0\|_V^2, \quad \forall \varphi_N^0 \in V_N.$$

En effet, on aurait alors

$$\mathfrak{F}(\varphi^0, \varphi^0) \geq \mathfrak{F}_N(\varphi^0, \varphi^0) \geq \min(\bar{\mathfrak{C}}_N, h/4) \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V.$$

**Étape 3** On montre d'abord que pour  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  fixé, il existe  $\mathfrak{C}_N(q_N) > 0$  tel que

$$\mathfrak{F}_N(\varphi^0, \varphi^0) \geq \mathfrak{C}_N(q_N) \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V_N.$$

On justifie ensuite qu'il existe  $\bar{\mathfrak{C}}_N > 0$  tel que

$$\mathfrak{C}_N(q_N) \geq \bar{\mathfrak{C}}_N, \quad \forall q \in \mathcal{Q}_{\text{ad}}^\varepsilon.$$

**Lemme 14.** Soit  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . La forme bilinéaire  $\mathfrak{F}_N$  est symétrique définie positive sur  $V_N$ . Par suite, comme  $V_N$  est de dimension finie, on a

$$\mathfrak{F}_N(\varphi^0, \varphi^0) \geq \mathfrak{C}_N(q_N) \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V_N,$$

où  $\mathfrak{C}_N(q_N) > 0$  est la plus petite valeur propre de  $\mathfrak{F}_N$  restreint à  $V_N$ .

*Démonstration.* Soit  $\varphi^0 = (\varphi^0, \varphi^1) \in V_N$  tel que  $\mathfrak{F}_N(\varphi^0, \varphi^0) = 0$ . On note  $\varphi$  la solution de (1.20) associée à  $\varphi^0$ . Comme  $(\varphi^0, \varphi^1) \in V_N$ , il existe des réels  $(\alpha_p)_{1 \leq p \leq 2^N}$  et  $(\beta_p)_{1 \leq p \leq 2^N}$  tels que

$$(\varphi^0)'(x) = \sum_{p=1}^{2^N} \alpha_p \mathbb{1}_{I_p}(x) \quad \text{et} \quad \varphi^1(x) = \sum_{p=1}^{2^N} \beta_p \mathbb{1}_{I_p}(x), \quad \text{avec} \quad \sum_{p=1}^{2^N} \alpha_p = 0.$$

Pour  $\mathbf{p} \in \mathfrak{J}_N$ , on pose  $s_{\mathbf{p}} := \text{sgn}(\mathbf{p})$  et  $\gamma_{\mathbf{p}} := \alpha_{|\mathbf{p}|} + s_{\mathbf{p}}\beta_{|\mathbf{p}|}$ .

Pour  $i, j \in \mathbb{Z}^*$  tels que  $C_{i,j} \in \mathcal{C}_N(q)$ , on note  $\mathbf{i} = \mathbf{n}(i) \in \mathfrak{J}_N$  et  $\mathbf{j} = \mathbf{n}(j) \in \mathfrak{J}_N$ . D'après la formule de d'Alembert (voir Prop. 15), pour tout  $(x, t) \in C_{i,j}$ , on a

$$\partial_t \varphi(x, t) = \frac{1}{2} \left( (\varphi^0)'(|r_j|) + s_j \varphi^1(|r_j|) \right) - \frac{1}{2} \left( (\varphi^0)'(|r_i|) - s_i \varphi^1(|r_i|) \right), \quad \text{où} \quad \begin{cases} r_i = \mathbf{m}(x - t) \in I_i, \\ r_j = \mathbf{m}(x + t) \in I_j. \end{cases}$$

Par suite, pour tout  $(x, t) \in C_{i,j}$ , on obtient

$$\partial_t \varphi(x, t) = \frac{1}{2} \left( \alpha_{|j|} + s_j \beta_{|j|} \right) - \frac{1}{2} \left( \alpha_{|i|} - s_i \beta_{|i|} \right) = \frac{1}{2} \left( \gamma_j - \gamma_{-i} \right).$$

On en déduit que

$$\mathfrak{F}_N(\varphi^0, \varphi^0) = \frac{h^2}{8} \sum_{C_{i,j} \in \mathcal{C}_N(q)} \left( \gamma_j - \gamma_{-i} \right)^2 = 0.$$

On montre à présent que tous les réels  $(\gamma_{\mathbf{p}})_{\mathbf{p} \in \mathfrak{J}_N}$  sont égaux. Pour ce faire, on s'appuie sur le fait que  $q_N$  vérifie (GCC).

- i. On montre que  $\gamma_1 = \dots = \gamma_{2^N}$ . Soit  $\mathbf{p} \in \llbracket 1, 2^N - 1 \rrbracket$ . Comme  $q_N$  vérifie (GCC), on a  $q_N \cap C_{x_{\mathbf{p}}}^- \neq \emptyset$ . Il existe alors  $i, j \in \mathbb{Z}^*$  tels que

$$\left( C_{i,j}, C_{i,j+1} \in \mathcal{C}_N(q) \text{ avec } +\mathbf{n}(j) = \mathbf{p} \right) \quad \text{ou} \quad \left( C_{i,j}, C_{i-1,j} \in \mathcal{C}_N(q) \text{ avec } -\mathbf{n}(i) = \mathbf{p} \right).$$

Dans les deux cas, on trouve  $\gamma_{\mathbf{p}} = \gamma_{\mathbf{p}+1}$ .

- ii. On montre que  $\gamma_{-1} = \dots = \gamma_{-2^N}$ . Soit  $\mathbf{p} \in \llbracket 1, 2^N - 1 \rrbracket$ . Comme  $q_N$  vérifie (GCC), on a  $q_N \cap C_{x_{\mathbf{p}}}^+ \neq \emptyset$ . Il existe alors  $i, j \in \mathbb{Z}^*$  tels que

$$\left( C_{i,j}, C_{i+1,j} \in \mathcal{C}_N(q) \text{ avec } +\mathbf{n}(i) = \mathbf{p} \right) \quad \text{ou} \quad \left( C_{i,j}, C_{i,j-1} \in \mathcal{C}_N(q) \text{ avec } -\mathbf{n}(j) = \mathbf{p} \right).$$

Dans les deux cas, on trouve  $\gamma_{-\mathbf{p}} = \gamma_{-\mathbf{p}-1}$ .

- iii. On montre que  $\gamma_1 = \gamma_{-1}$ . Comme  $q_N$  vérifie (GCC), on a  $q_N \cap C_{x_0}^+ \neq \emptyset$ . Il existe alors  $i, j \in \mathbb{Z}^*$  tels que

$$\left( C_{i,j}, C_{i+1,j} \in \mathcal{C}_N(q) \text{ avec } +\mathbf{n}(i) = -1 \right) \quad \text{ou} \quad \left( C_{i,j}, C_{i,j-1} \in \mathcal{C}_N(q) \text{ avec } -\mathbf{n}(j) = -1 \right).$$

Dans les deux cas, on trouve  $\gamma_1 = \gamma_{-1}$ .

Il existe donc  $\gamma \in \mathbb{R}$  tel que  $\gamma_{\mathbf{p}} = \gamma$ , pour tout  $\mathbf{p} \in \mathfrak{J}_N$ . On en déduit d'une part que  $\varphi^1 \equiv 0$ . En effet, pour  $p \in \llbracket 1, 2^N \rrbracket$ , comme  $\gamma_p = \gamma_{-p}$ , on a  $\alpha_p + \beta_p = \alpha_p - \beta_p$ , d'où  $\beta_p = 0$ . On en déduit d'autre part que  $\varphi^0 \equiv 0$ . En effet, comme

$$2^{N+1} \gamma = \sum_{\mathbf{p} \in \mathfrak{J}_N} \gamma_{\mathbf{p}} = 2 \sum_{p=1}^{2^N} \alpha_p = 0,$$

on a  $\gamma = 0$ , d'où  $\alpha_p = 0$ , pour tout  $p \in \llbracket 1, 2^N \rrbracket$ . On en conclut que  $\varphi^0 \equiv (0, 0)$ .  $\square$

Comme l'ensemble  $\mathcal{C}_N(Q) := \left\{ C_{i,j} \in \mathcal{C}_N; \overset{\circ}{C}_{i,j} \subset Q \right\}$  est fini, on remarque que l'ensemble  $\mathcal{Q}_{\text{ad}}^N := \left\{ q_N; q \in \mathcal{Q}_{\text{ad}}^{\varepsilon} \right\}$  est lui aussi fini. On pose alors

$$\bar{\mathfrak{C}}_N := \min_{q_N \in \mathcal{Q}_{\text{ad}}^N} \mathfrak{C}_N(q_N) > 0,$$

où  $\mathfrak{C}_N(q_N) > 0$  est la valeur propre introduite au Lem. 14. Ceci conclut la preuve puisque pour tout  $q \in \mathcal{Q}_{\text{ad}}^{\varepsilon}$ , on a

$$\mathfrak{F}_N(\varphi^0, \varphi^0) \geq \bar{\mathfrak{C}}_N \|\varphi^0\|_V^2, \quad \forall \varphi^0 \in V_N.$$

### 1.4.2 Compléments à la Section 3

Dans cette section, on revient sur la Sect. 3 de l'Art. A, consacrée à l'optimisation du support de contrôle. Contrairement à l'Art. A où l'on utilise une régularisation en espace du support de contrôle, on utilise ici une régularisation en temps du support de contrôle. On se place dans le cas de l'équation des ondes linéaire 1D avec contrôle interne. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . Dans la suite, on emploie la notation (peu canonique)  $H_0^2(\Omega) := H^2(\Omega) \cap H_0^1(\Omega)$ .

On fixe  $T > 2$  et on choisit  $\tau > 0$  tel que  $T - 2\tau \geq 2$ . On se donne  $\eta \in C^1(\mathbb{R})$  tel que

$$\eta : \mathbb{R} \rightarrow [0, 1], \quad \eta(t) = \begin{cases} 1 & \text{si } t \in [\tau, T - \tau], \\ 0 & \text{si } t \notin (0, T). \end{cases}$$

On fixe  $\delta_0 > 0$ . Pour  $\gamma \in W^{1,\infty}(0, T)$ , on note

$$q_\gamma := \left\{ (x, t) \in Q; \quad |x - \gamma(t)| < \delta_0 \right\}.$$

On fixe  $M > 0$  et on introduit l'ensemble admissible

$$\mathcal{G}_{\text{ad}} := \left\{ \gamma \in W^{1,\infty}(0, T); \quad \|\gamma'\|_{L^\infty(0, T)} \leq M, \quad \delta_0 \leq \gamma \leq 1 - \delta_0 \right\}.$$

Pour  $\gamma \in \mathcal{G}_{\text{ad}}$ , on définit les espaces  $L^2$  à poids

$$\mathcal{U}_\gamma := \left\{ u : q_\gamma \rightarrow \mathbb{R}; \quad \iint_{q_\gamma} u^2 \eta^{-1} < +\infty \right\} \quad \text{et} \quad \mathcal{U}'_\gamma := \left\{ v : q_\gamma \rightarrow \mathbb{R}; \quad \iint_{q_\gamma} v^2 \eta < +\infty \right\},$$

que l'on munit des produits scalaires

$$\langle u, \bar{u} \rangle_{\mathcal{U}_\gamma} := \iint_{q_\gamma} u \bar{u} \eta^{-1}, \quad \forall u, \bar{u} \in \mathcal{U}_\gamma \quad \text{et} \quad \langle v, \bar{v} \rangle_{\mathcal{U}'_\gamma} := \iint_{q_\gamma} v \bar{v} \eta, \quad \forall v, \bar{v} \in \mathcal{U}'_\gamma.$$

Pour  $(y^0, y^1) \in H_0^2(\Omega) \times H_0^1(\Omega)$ , on s'intéresse au contrôle à zéro de norme  $\mathcal{U}_\gamma$  minimale pour le système (1.19). Avec la méthode d'unicité de Hilbert, on peut montrer que ce contrôle est donné par  $\hat{u}_\gamma = -\hat{\varphi}_\gamma \eta|_{q_\gamma}$ , où  $\hat{\varphi}_\gamma$  est la solution de (1.20) associée au point de minimum  $(\hat{\varphi}_\gamma^0, \hat{\varphi}_\gamma^1)$  de la fonctionnelle

$$\mathcal{J}_\gamma^*(\varphi^0, \varphi^1) := \frac{1}{2} \iint_{q_\gamma} |\varphi|^2 \eta + \langle y^0, \varphi^1 \rangle_{H_0^1, H^{-1}} - \langle y^1, \varphi^0 \rangle_{L^2}, \quad \forall (\varphi^0, \varphi^1) \in L^2(\Omega) \times H^{-1}(\Omega).$$

Il est à noter que, d'après [12], la régularité de  $\eta$  et de  $(y^0, y^1)$  implique que  $(\hat{\varphi}_\gamma^0, \hat{\varphi}_\gamma^1) \in H_0^1(\Omega) \times L^2(\Omega)$  et qu'il existe  $C_\gamma > 0$  tel que

$$\|(\hat{\varphi}_\gamma^0, \hat{\varphi}_\gamma^1)\|_{H_0^1 \times L^2} \leq C_\gamma \|(y^0, y^1)\|_{H_0^2 \times H_0^1}.$$

En outre, on peut trouver dans l'inégalité ci-dessus une constante  $\bar{C} > 0$  uniforme sur  $\mathcal{G}_{\text{ad}}$ .

On fixe  $(y^0, y^1) \in H_0^2(\Omega) \times H_0^1(\Omega)$  et on cherche à minimiser la norme du contrôle  $\hat{u}_\gamma$  par rapport à  $\gamma \in \mathcal{G}_{\text{ad}}$ . On considère ainsi le problème

$$\inf_{\gamma \in \mathcal{G}_{\text{ad}}} J(\gamma), \quad \text{où} \quad J(\gamma) := \|\hat{u}_\gamma\|_{\mathcal{U}_\gamma}^2 = \iint_{q_\gamma} |\hat{\varphi}_\gamma|^2 \eta.$$

**Proposition 38.** *La fonctionnelle  $J$  est continue sur  $\mathcal{G}_{\text{ad}}$  pour la norme  $L^\infty(0, T)$ .*

*Démonstration.* Soient  $(\gamma_n)_{n \geq 0} \subset \mathcal{G}_{\text{ad}}$  et  $\gamma \in \mathcal{G}_{\text{ad}}$  tels que  $\gamma_n \rightarrow \gamma$  dans  $L^\infty(0, T)$ . Pour alléger les notations, on pose  $q_n = q_{\gamma_n}$  et  $q = q_\gamma$ . Pour tout  $n \in \mathbb{N}$ , on note  $(\hat{\varphi}_n^0, \hat{\varphi}_n^1) \in H_0^1(\Omega) \times L^2(\Omega)$  le point de minimum de  $\mathcal{J}_{\gamma_n}^*$  et  $\hat{\varphi}_n$  la solution correspondante de (1.20). Il existe une constante  $\bar{C} > 0$ , indépendante de  $n$ , telle que

$$\|(\hat{\varphi}_n^0, \hat{\varphi}_n^1)\|_{H_0^1 \times L^2} \leq \bar{C} \|(y^0, y^1)\|_{H_0^2 \times H_0^1}.$$

Par suite, il existe  $(\hat{\varphi}^0, \hat{\varphi}^1) \in H_0^1(\Omega) \times L^2(\Omega)$  tel qu'à une extraction près,

$$(\hat{\varphi}_n^0, \hat{\varphi}_n^1) \rightharpoonup (\hat{\varphi}^0, \hat{\varphi}^1) \quad \text{dans} \quad H_0^1(\Omega) \times L^2(\Omega).$$

Sachant que la solution de (1.20) dépend continûment de la donnée initiale, il vient

$$(\widehat{\varphi}_n, \partial_t \widehat{\varphi}_n) \rightharpoonup^* (\widehat{\varphi}, \partial_t \widehat{\varphi}) \quad \text{dans } L^\infty(0, T; H_0^1(\Omega) \times L^2(\Omega)),$$

où  $\widehat{\varphi}$  est la solution de (1.20) associée à  $(\widehat{\varphi}^0, \widehat{\varphi}^1)$ . Comme l'injection de  $H_0^1(\Omega)$  dans  $C(\overline{\Omega})$  est compacte et que l'injection de  $C(\overline{\Omega})$  dans  $L^2(\Omega)$  est continue, d'après le lemme d'Aubin-Lions-Simon, à une extraction près, on a

$$\widehat{\varphi}_n \rightarrow \widehat{\varphi} \quad \text{dans } C(\overline{Q}).$$

Pour tout  $\overline{\varphi} \in L^\infty(0, T; L^2(\Omega))$ , on a alors

$$\iint_{q_n} \widehat{\varphi}_n \overline{\varphi} \eta = \iint_{q_n} \widehat{\varphi} \overline{\varphi} \eta + \iint_{q_n} (\widehat{\varphi}_n - \widehat{\varphi}) \overline{\varphi} \eta \rightarrow \iint_q \widehat{\varphi} \overline{\varphi} \eta.$$

En effet, la limite du premier terme s'obtient par convergence dominée, tandis que le second terme tend vers 0 puisque

$$\left| \iint_{q_n} (\widehat{\varphi}_n - \widehat{\varphi}) \overline{\varphi} \eta \right| \leq \|\overline{\varphi}\|_{L^1(Q)} \|\widehat{\varphi}_n - \widehat{\varphi}\|_{L^\infty(Q)}.$$

Le résultat de convergence ci-dessus permet de passer à la limite dans la condition d'optimalité de  $(\widehat{\varphi}_n^0, \widehat{\varphi}_n^1)$  et d'obtenir

$$\iint_q \widehat{\varphi} \overline{\varphi} \eta = -\langle y^0, \overline{\varphi}^1 \rangle_{H_0^1, H^{-1}} + \langle y^1, \overline{\varphi}^0 \rangle_{L^2}, \quad \forall (\overline{\varphi}^0, \overline{\varphi}^1) \in L^2(\Omega) \times H^{-1}(\Omega),$$

où  $\overline{\varphi}$  est la solution de (1.20) associée à  $(\overline{\varphi}^0, \overline{\varphi}^1)$ . Cette relation signifie que  $(\widehat{\varphi}^0, \widehat{\varphi}^1)$  est le point de minimum de  $\mathcal{J}_\gamma^*$ . En outre, ceci caractérise de manière unique la limite faible de toute sous-suite de  $(\widehat{\varphi}_n^0, \widehat{\varphi}_n^1)$  et implique que la suite entière converge faiblement vers  $(\widehat{\varphi}^0, \widehat{\varphi}^1)$ . Finalement, la continuité de  $J$  provient du passage à la limite

$$\iint_{q_n} |\widehat{\varphi}_n|^2 \eta = -\langle y^0, \widehat{\varphi}_n^1 \rangle_{H_0^1, H^{-1}} + \langle y^1, \widehat{\varphi}_n^0 \rangle_{L^2} \rightarrow -\langle y^0, \widehat{\varphi}^1 \rangle_{H_0^1, H^{-1}} + \langle y^1, \widehat{\varphi}^0 \rangle_{L^2} = \iint_q |\widehat{\varphi}|^2 \eta.$$

□

**Proposition 39.** *La fonctionnelle  $J$  atteint son minimum sur  $\mathcal{G}_{\text{ad}}$ .*

*Démonstration.* La fonctionnelle  $J$  étant bornée inférieurement, il existe une suite minimisante  $(\gamma_n)_{n \geq 0} \subset \mathcal{G}_{\text{ad}}$ . Par définition de  $\mathcal{G}_{\text{ad}}$ , cette suite est bornée dans  $W^{1,\infty}(0, T)$ . De plus, comme l'injection de  $W^{1,\infty}(0, T)$  dans  $L^\infty(0, T)$  est compacte, il existe  $\gamma \in L^\infty(0, T)$  tel qu'à une extraction près,  $\gamma_n \rightarrow \gamma$  dans  $L^\infty(0, T)$ . Toujours par définition de  $\mathcal{G}_{\text{ad}}$ , chaque  $\gamma_n$  est  $M$ -lipschitzienne, avec  $M$  indépendant de  $n$ . Ainsi, en passant à la limite simple dans les expressions

$$|\gamma_n(t) - \gamma_n(s)| \leq M|t - s|, \quad \delta_0 \leq \gamma_n(t) \leq 1 - \delta_0, \quad t, s \in [0, T],$$

on en déduit que  $\gamma \in \mathcal{G}_{\text{ad}}$ . Finalement, la continuité de  $J$  fournit  $J(\gamma_n) \rightarrow J(\gamma) = \inf_{\mathcal{G}_{\text{ad}}} J$ . □

**Proposition 40** (Dérivée au sens de Gateaux). *La dérivée directionnelle de  $J$  en  $\gamma \in \mathcal{G}_{\text{ad}}$  dans la direction admissible  $\overline{\gamma}$  s'écrit*

$$dJ(\gamma; \overline{\gamma}) = \lim_{\varepsilon \rightarrow 0} \frac{J(\gamma + \varepsilon \overline{\gamma}) - J(\gamma)}{\varepsilon} = \int_0^T \phi_\gamma(t) \overline{\gamma}(t) dt,$$

avec

$$\phi_\gamma(t) = \left( |\widehat{\varphi}_\gamma(\gamma(t) - \delta_0, t)|^2 - |\widehat{\varphi}_\gamma(\gamma(t) + \delta_0, t)|^2 \right) \eta(t), \quad t \in [0, T],$$

où  $\widehat{\varphi}_\gamma$  est la solution de (1.20) associée au point de minimum  $(\widehat{\varphi}_\gamma^0, \widehat{\varphi}_\gamma^1)$  de  $\mathcal{J}_\gamma^*$ .

*Démonstration.* Soit  $\varepsilon > 0$  assez petit. Pour alléger les notations, on pose  $q_\varepsilon = q_{\gamma + \varepsilon \overline{\gamma}}$  et  $q = q_\gamma$ . On note  $(\widehat{\varphi}_\varepsilon^0, \widehat{\varphi}_\varepsilon^1) \in H_0^1(\Omega) \times L^2(\Omega)$  le point de minimum de  $\mathcal{J}_{\gamma + \varepsilon \overline{\gamma}}^*$  et  $\widehat{\varphi}_\varepsilon$  la solution correspondante de (1.20). De même, on note  $(\widehat{\varphi}^0, \widehat{\varphi}^1) \in H_0^1(\Omega) \times L^2(\Omega)$  le point de minimum de  $\mathcal{J}_\gamma^*$  et  $\widehat{\varphi}$  la solution correspondante de (1.20). En exploitant les conditions d'optimalité pour  $\mathcal{J}_{\gamma + \varepsilon \overline{\gamma}}^*$  et  $\mathcal{J}_\gamma^*$ , il vient

$$J(\gamma + \varepsilon \overline{\gamma}) - J(\gamma) = \iint_{q_\varepsilon} |\widehat{\varphi}_\varepsilon|^2 \eta - \iint_q |\widehat{\varphi}|^2 \eta = \iint_q \widehat{\varphi}_\varepsilon \widehat{\varphi} \eta - \iint_{q_\varepsilon} \widehat{\varphi}_\varepsilon \widehat{\varphi} \eta.$$

On peut montrer que  $\widehat{\varphi}_\varepsilon \rightarrow \widehat{\varphi}$  dans  $C(\overline{Q})$  quand  $\varepsilon \rightarrow 0$ . Par suite, il vient

$$\frac{J(\gamma + \varepsilon\overline{\gamma}) - J(\gamma)}{\varepsilon} = \iint_Q \frac{\mathbb{1}_q - \mathbb{1}_{q_\varepsilon}}{\varepsilon} \widehat{\varphi}_\varepsilon \widehat{\varphi} \eta = \iint_Q \frac{\mathbb{1}_q - \mathbb{1}_{q_\varepsilon}}{\varepsilon} |\widehat{\varphi}|^2 \eta + \iint_Q \frac{\mathbb{1}_q - \mathbb{1}_{q_\varepsilon}}{\varepsilon} (\widehat{\varphi}_\varepsilon - \widehat{\varphi}) \widehat{\varphi} \eta \rightarrow \int_0^T \phi_\gamma \overline{\gamma}$$

quand  $\varepsilon \rightarrow 0$ . En effet, la limite du premier terme s'obtient par convergence dominée après avoir remarqué que

$$\iint_Q \frac{\mathbb{1}_q - \mathbb{1}_{q_\varepsilon}}{\varepsilon} |\widehat{\varphi}|^2 \eta = \int_0^T \left( \frac{1}{\varepsilon} \int_{a(t)}^{a(t)+\varepsilon\overline{\gamma}(t)} |\widehat{\varphi}(x,t)|^2 \eta(t) dx - \frac{1}{\varepsilon} \int_{b(t)}^{b(t)+\varepsilon\overline{\gamma}(t)} |\widehat{\varphi}(x,t)|^2 \eta(t) dx \right) dt,$$

où  $a(t) = \gamma(t) - \delta_0$  et  $b(t) = \gamma(t) + \delta_0$ . Le second terme tend quant à lui vers 0 puisque

$$\left| \iint_Q \frac{\mathbb{1}_q - \mathbb{1}_{q_\varepsilon}}{\varepsilon} (\widehat{\varphi}_\varepsilon - \widehat{\varphi}) \widehat{\varphi} \eta \right| \leq 2 \|\overline{\gamma}\|_{L^1(0,T)} \|\widehat{\varphi}\|_{L^\infty(Q)} \|\widehat{\varphi}_\varepsilon - \widehat{\varphi}\|_{L^\infty(Q)}.$$

□

Du point de vue numérique, pour  $\varepsilon > 0$  fixé, on cherche à résoudre le problème

$$\inf_{\gamma \in W^{1,\infty}} J_\varepsilon(\gamma), \quad \text{où } J_\varepsilon(\gamma) := J(\gamma) + \frac{\varepsilon}{2} \|\gamma'\|_{L^2(0,T)}^2.$$

Pour ce faire, on utilise la méthode du gradient à pas fixe (cf. Algo. 6).

---

**Algorithm 6** Minimisation de  $J_\varepsilon$  par la méthode du gradient à pas fixe

---

- 1) *Initialisation.* On fixe  $\varepsilon > 0$ ,  $\rho > 0$  et on se donne  $\gamma_0 \in W^{1,\infty}(0,T)$  tel que  $\delta_0 \leq \gamma_0 \leq 1 - \delta_0$ .
- 2) *Itération.* Pour  $n \geq 0$ , on calcule  $\phi_n^\varepsilon \in H^1(0,T)$  solution de

$$\langle \phi_n^\varepsilon, \overline{\gamma} \rangle_{L^2(0,T)} + \varepsilon \langle \phi_n^\varepsilon, \overline{\gamma}' \rangle_{L^2(0,T)} = \langle \phi_{\gamma_n}, \overline{\gamma} \rangle_{L^2(0,T)} + \varepsilon \langle \gamma_n', \overline{\gamma}' \rangle_{L^2(0,T)}, \quad \forall \overline{\gamma} \in H^1(0,T).$$

On met à jour

$$\gamma_{n+1} = \gamma_n - \rho \phi_n^\varepsilon.$$

On pose

$$g_{n+1} = \frac{|J_\varepsilon(\gamma_{n+1}) - J_\varepsilon(\gamma_n)|}{J_\varepsilon(\gamma_0)}.$$

- 3) *Test de convergence.* Si  $g_{n+1} \leq \varepsilon$ , l'algorithme s'arrête et on prend  $\gamma_{n+1}$  comme approximation du point de minimum de  $J_\varepsilon$ . Sinon, effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).
- 

On conclut cette section par l'énoncé d'un algorithme d'optimisation du support (cf. Algo. 7), dans le cas où l'ouvert de contrôle  $q$  est défini comme la partie de  $Q$  délimitée par deux courbes.

Pour  $\gamma = (a, b)$ , avec  $a, b \in W^{1,\infty}(0,T)$ , on note

$$q_\gamma := \left\{ (x, t) \in Q; \quad a(t) < x < b(t) \right\}.$$

Pour  $0 < \varsigma \leq 1$  et  $\varepsilon > 0$  fixés, on cherche à résoudre le problème

$$\inf_{\substack{\gamma = (a,b) \\ a,b \in W^{1,\infty}}} \sup_{\lambda \geq 0} J_{\lambda,\varepsilon}(\gamma), \quad \text{où } J_{\lambda,\varepsilon}(\gamma) := J(\gamma) + \lambda \left( \|\mathbb{1}_{q_\gamma}\|_{L^1(Q)} - \varsigma |Q| \right) + \frac{\varepsilon}{2} \left( \|a'\|_{L^2(0,T)}^2 + \|b'\|_{L^2(0,T)}^2 \right).$$

Il est à noter que  $\|\mathbb{1}_{q_\gamma}\|_{L^1(Q)} = \int_0^T (b-a)$ .

**Proposition 41** (Dérivée au sens de Gateaux). *La dérivée directionnelle de  $J$  en  $\gamma = (a, b)$  dans la direction admissible  $\overline{\gamma} = (\overline{a}, \overline{b})$  s'écrit*

$$dJ(\gamma; \overline{\gamma}) = \lim_{\varepsilon \rightarrow 0} \frac{J(\gamma + \varepsilon\overline{\gamma}) - J(\gamma)}{\varepsilon} = \int_0^T \left( \phi_a(t) \overline{a}(t) + \phi_b(t) \overline{b}(t) \right) dt,$$

avec

$$\phi_a(t) = |\widehat{\varphi}_\gamma(a(t), t)|^2 \eta(t) \quad \text{et} \quad \phi_b(t) = -|\widehat{\varphi}_\gamma(b(t), t)|^2 \eta(t), \quad t \in [0, T],$$

où  $\widehat{\varphi}_\gamma$  est la solution de (1.20) associée au point de minimum  $(\widehat{\varphi}_\gamma^0, \widehat{\varphi}_\gamma^1)$  de  $\mathcal{J}_\gamma^*$ .

---

**Algorithm 7** Minimisation de  $J_{\lambda,\epsilon}$  par la méthode du gradient à pas fixe

---

- 1) *Initialisation.* On fixe  $\epsilon > 0$ ,  $\rho > 0$  et on se donne  $\gamma_0 = (a_0, b_0)$ , avec  $a_0, b_0 \in W^{1,\infty}(0, T)$ , tel que  $0 \leq a_0 < b_0 \leq 1$  et  $|q_{\gamma_0}| = \varsigma|\Omega|$ .
- 2) *Itération.* Pour  $n \geq 0$ , on calcule  $\phi_{a_n}^\epsilon, \phi_{b_n}^\epsilon \in H^1(0, T)$  solutions respectives de

$$\begin{aligned} \langle \phi_{a_n}^\epsilon, \bar{a} \rangle_{L^2(0,T)} + \epsilon \langle \phi_{a_n}^{\epsilon'}, \bar{a}' \rangle_{L^2(0,T)} &= \langle \phi_{a_n}, \bar{a} \rangle_{L^2(0,T)} + \epsilon \langle a_n', \bar{a}' \rangle_{L^2(0,T)}, \quad \forall \bar{a} \in H^1(0, T), \\ \langle \phi_{b_n}^\epsilon, \bar{b} \rangle_{L^2(0,T)} + \epsilon \langle \phi_{b_n}^{\epsilon'}, \bar{b}' \rangle_{L^2(0,T)} &= \langle \phi_{b_n}, \bar{b} \rangle_{L^2(0,T)} + \epsilon \langle b_n', \bar{b}' \rangle_{L^2(0,T)}, \quad \forall \bar{b} \in H^1(0, T). \end{aligned}$$

On pose

$$\lambda_n = \frac{1}{2T} \int_0^T (\phi_{a_n}^\epsilon - \phi_{b_n}^\epsilon).$$

On met à jour

$$a_{n+1} = a_n - \rho(\phi_{a_n}^\epsilon - \lambda_n) \quad \text{et} \quad b_{n+1} = b_n - \rho(\phi_{b_n}^\epsilon + \lambda_n).$$

On pose  $\gamma_{n+1} = (a_{n+1}, b_{n+1})$  et

$$g_{n+1} = \frac{|J_{0,\epsilon}(\gamma_{n+1}) - J_{0,\epsilon}(\gamma_n)|}{J_{0,\epsilon}(\gamma_0)}.$$

- 3) *Test de convergence.* Si  $g_{n+1} \leq \epsilon$ , l'algorithme s'arrête et on prend  $\gamma_{n+1}$  comme approximation du point de minimum de  $J_{\lambda,\epsilon}$ . Sinon, effectuer  $n \leftarrow n + 1$  et revenir à l'étape 2).
-

# Bibliographie

- [1] M. ASCH AND G. LEBEAU, *Geometrical aspects of exact boundary controllability for the wave equation—a numerical study*, ESAIM Control Optim. Calc. Var., 3 (1998), pp. 163–212.
- [2] M. ASCH AND A. MÜNCH, *Uniformly controllable schemes for the wave equation on the unit square*, J. Optim. Theory Appl., 143 (2009), pp. 417–438.
- [3] C. BARDOS, G. LEBEAU, AND J. RAUCH, *Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary*, SIAM J. Control Optim., 30 (1992), pp. 1024–1065.
- [4] D. BOFFI, F. BREZZI, AND M. FORTIN, *Mixed finite element methods and applications*, vol. 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013.
- [5] 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].
- [6] E. BURMAN, A. FEIZMOHAMMADI, A. MÜNCH, AND L. OKSANEN, *Space time stabilized finite element methods for a unique continuation problem subject to the wave equation*, ESAIM Math. Model. Numer. Anal., 55 (2021), pp. S969–S991.
- [7] C. CASTRO, *Exact controllability of the 1-D wave equation from a moving interior point*, ESAIM Control Optim. Calc. Var., 19 (2013), pp. 301–316.
- [8] C. CASTRO, N. CÎNDEA, AND A. MÜNCH, *Controllability of the linear one-dimensional wave equation with inner moving forces*, SIAM J. Control Optim., 52 (2014), pp. 4027–4056.
- [9] C. CASTRO, S. MICU, AND A. MÜNCH, *Numerical approximation of the boundary control for the wave equation with mixed finite elements in a square*, IMA J. Numer. Anal., 28 (2008), pp. 186–214.
- [10] J.-M. CORON, *Control and nonlinearity*, vol. 136 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2007.
- [11] T. DUPONT, R. GLOWINSKI, W. KINTON, AND M. F. WHEELER, *Mixed finite-element methods for time-dependent problems : application to control*, in Finite elements in fluids, Vol. 8 (Huntsville, AL, 1989), Ser. Comput. Phys. Process. Mech. Thermal Sci., Hemisphere, Washington, DC, 1992, pp. 119–136.
- [12] S. ERVEDOZA AND E. ZUAZUA, *A systematic method for building smooth controls for smooth data*, Discrete Contin. Dyn. Syst. Ser. B, 14 (2010), pp. 1375–1401.
- [13] ———, *Numerical approximation of exact controls for waves*, SpringerBriefs in Mathematics, Springer, New York, 2013.
- [14] L. C. EVANS, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.
- [15] X. FU, J. YONG, AND X. ZHANG, *Exact controllability for multidimensional semilinear hyperbolic equations*, SIAM J. Control Optim., 46 (2007), pp. 1578–1614.
- [16] R. GLOWINSKI, *Ensuring well-posedness by analogy : Stokes problem and boundary control for the wave equation*, J. Comput. Phys., 103 (1992), pp. 189–221.
- [17] R. GLOWINSKI, W. KINTON, AND M. F. WHEELER, *A mixed finite element formulation for the boundary controllability of the wave equation*, Internat. J. Numer. Methods Engrg., 27 (1989), pp. 623–635.
- [18] 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 (1990), pp. 1–76.



- [19] R. GLOWINSKI, J.-L. LIONS, AND J. HE, *Exact and approximate controllability for distributed parameter systems*, vol. 117 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2008. A numerical approach.
- [20] A. HARAUX, T. LIARD, AND Y. PRIVAT, *How to estimate observability constants of one-dimensional wave equations? Propagation versus spectral methods*, J. Evol. Equ., 16 (2016), pp. 825–856.
- [21] L. F. HO, *Exact controllability of the one-dimensional wave equation with locally distributed control*, SIAM J. Control Optim., 28 (1990), pp. 733–748.
- [22] 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 (1999), pp. 407–438.
- [23] A. E. INGHAM, *Some trigonometrical inequalities with applications to the theory of series*, Math. Z., 41 (1936), pp. 367–379.
- [24] J. LE ROUSSEAU, G. LEBEAU, P. TERPOLILLI, AND E. TRÉLAT, *Geometric control condition for the wave equation with a time-dependent observation domain*, Anal. PDE, 10 (2017), pp. 983–1015.
- [25] J.-L. LIONS, *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1*, vol. 8 of Recherches en Mathématiques Appliquées [Research in Applied Mathematics], Masson, Paris, 1988. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos, G. Lebeau and J. Rauch.
- [26] S. MONTANER AND A. MÜNCH, *Approximation of controls for linear wave equations : a first order mixed formulation*, Math. Control Relat. Fields, 9 (2019), pp. 729–758.
- [27] A. MÜNCH, *A uniformly controllable and implicit scheme for the 1-D wave equation*, M2AN Math. Model. Numer. Anal., 39 (2005), pp. 377–418.
- [28] ———, *Optimal design of the support of the control for the 2-D wave equation : a numerical method*, Int. J. Numer. Anal. Model., 5 (2008), pp. 331–351.
- [29] ———, *Optimal location of the support of the control for the 1-D wave equation : numerical investigations*, Comput. Optim. Appl., 42 (2009), pp. 443–470.
- [30] A. PAZY, *Semi-groups of linear operators and applications to partial differential equations*, Department of Mathematics, University of Maryland, College Park, Md., 1974. Department of Mathematics, University of Maryland, Lecture Note, No. 10.
- [31] F. PERIAGO, *Optimal shape and position of the support for the internal exact control of a string*, Systems Control Lett., 58 (2009), pp. 136–140.
- [32] R. T. ROCKAFELLAR, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.
- [33] W. RUDIN, *Functional analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, second ed., 1991.
- [34] M. TUCSNAK AND G. WEISS, *Observation and control for operator semigroups*, Birkhäuser Advanced Texts : Basler Lehrbücher. [Birkhäuser Advanced Texts : Basel Textbooks], Birkhäuser Verlag, Basel, 2009.
- [35] E. ZUAZUA, *Exact controllability for semilinear wave equations in one space dimension*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 10 (1993), pp. 109–129.
- [36] E. ZUAZUA, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Rev., 47 (2005), pp. 197–243.

# Chapitre 2

## Résultats de la thèse

Les travaux réalisés durant la thèse de l’auteur sont essentiellement rassemblés dans quatre articles. Dans ce chapitre, on compile à la suite ces quatre articles par ordre de soumission. Chaque article est précédé d’un résumé succinct. On commence par situer l’article dans le contexte posé au premier chapitre. Puis, après avoir motivé le cadre choisi, on énonce les principaux résultats.

### A Optimisation du support de contrôle non cylindrique pour la contrôlabilité à zéro de l’équation des ondes 1D

Titre : *Optimization of non-cylindrical domains for the exact null controllability of the 1D wave equation.*

Auteurs : ARTHUR BOTTOIS, NICOLAE CÎNDEA et ARNAUD MÜNCH.

Publié dans *ESAIM: Control, Optimisation and Calculus of Variations, Volume 27, 2021.*

<https://doi.org/10.1051/cocv/2021010>

Aussi disponible à <https://arxiv.org/abs/1911.01284>

Contributions :

- La démonstration d’une inégalité d’observabilité uniforme par rapport au support de contrôle.
- La proposition et la mise en œuvre numérique d’un algorithme d’optimisation du support.

#### Résumé

Dans cet article, on traite le problème de la contrôlabilité à zéro pour l’équation des ondes linéaire unidimensionnelle avec contrôle interne. On reprend ici les notations de la Sect. 1.1.6 avec  $d = 1$  et  $\Omega = (0, 1)$ . On considère le système d’état (1.19) où le contrôle  $u$  peut agir dans l’ouvert non cylindrique  $q \subset Q$ . On sait que la contrôlabilité de (1.19) est équivalente à l’inégalité d’observabilité (1.22) (ou (1.21)). Dans ce cadre unidimensionnel, l’inégalité (1.22) a été établie dans [8] pour des ouverts de contrôle  $q$  vérifiant le critère (GCC) usuel (voir Déf. 23) portant sur les lignes caractéristiques. Ce résultat a par la suite été étendu dans [24] au cas de l’équation des ondes  $d$ -dimensionnelle pour des ouverts  $q$  vérifiant une condition de type (GCC) portant sur les bicaractéristiques généralisées. Cependant, ces travaux ne permettent pas d’obtenir la valeur de la constante d’observabilité  $C_{\text{obs}}(q)$ , ou d’explicitier la dépendance de cette constante vis-à-vis de telle ou telle caractéristique géométrique de  $q$ . Pour des ouverts cylindriques  $q = \omega \times (0, T)$ , avec  $\omega \subset \Omega$ , des contributions de cette nature peuvent être trouvées dans [31, Prop. 2.1] ou dans [20, Th. 1]. Dans ces travaux, les auteurs identifient une constante  $C_{\text{obs}}(|\omega|, T)$  ne dépendant que de la mesure de  $\omega$  et du temps de contrôle.

La première partie de l’article est consacrée à l’établissement d’une inégalité d’observabilité uniforme par rapport à l’ouvert non cylindrique  $q$ . Plus précisément, pour  $\varepsilon > 0$  fixé, on définit la classe

$$\mathcal{Q}_{\text{ad}}^\varepsilon = \left\{ q \subset Q \text{ ouvert; } \text{Int}_\varepsilon(q) \text{ vérifie (GCC)} \right\},$$

où  $\text{Int}_\varepsilon(q) = \{(x, t) \in q; \text{dist}((x, t), \partial q) > \varepsilon\}$  désigne le  $\varepsilon$ -intérieur de  $q$ . L’idée est ici de regrouper les ouverts de contrôle  $q$  vérifiant (GCC) et tels que toute ligne caractéristique intersecte  $q$  sur une longueur supérieure à  $\varepsilon$ . Le paramètre  $\varepsilon$  mesure donc à quel point le critère (GCC) est “confortablement” vérifié. On démontre alors le théorème suivant.

**Théorème** (Inégalité d’observabilité uniforme). *Pour  $\varepsilon > 0$  fixé, il existe une constante  $\overline{C}_{\text{obs}}^\varepsilon > 0$  ne dépendant que de  $\varepsilon$  et telle que pour tout  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ ,*

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq \overline{C}_{\text{obs}}^\varepsilon \|\partial_t \varphi\|_{L^2(q)}^2, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution du système adjoint (1.20) associée à  $(\varphi^0, \varphi^1)$ .

Dans la démonstration, on utilise une méthode dite propagative. Une telle méthode s’appuie sur la formule de d’Alembert et sur les propriétés de propagation des ondes le long des lignes caractéristiques. La philosophie de la preuve est la suivante. Pour tout ouvert de contrôle  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , on construit un ouvert approché  $q_\varepsilon$  vérifiant  $\text{Int}_\varepsilon(q) \subset q_\varepsilon \subset q$  et pour lequel on peut calculer explicitement la constante d’observabilité  $C_{\text{obs}}(q_\varepsilon)$ . La constante  $C_{\text{obs}}(q_\varepsilon)$  est exprimée à l’aide de la plus petite valeur propre d’une forme quadratique en dimension finie ou, de façon équivalente, de la deuxième valeur propre de la matrice laplacienne d’un certain graphe. De ce fait, le calcul de  $C_{\text{obs}}(q)$  peut être rapproché de la mesure de la connectivité d’un graphe. On conclut la preuve en justifiant que l’on peut choisir  $\overline{C}_{\text{obs}}^\varepsilon = \min_{q_\varepsilon} C_{\text{obs}}(q_\varepsilon)$  dans l’inégalité d’observabilité uniforme. Par ailleurs, on signale qu’une preuve alternative de l’inégalité d’observabilité uniforme peut être consultée à la Sect. 1.4.1.

Dans la deuxième partie de l’article, pour une donnée initiale  $(y^0, y^1)$  fixée, on cherche à identifier l’ouvert  $q$  permettant de minimiser l’énergie à dépenser pour contrôler l’onde. On fait deux hypothèses simplificatrices majeures. D’une part, on restreint notre étude à des supports de contrôle  $q_\gamma$  définis comme le déplacement rigide d’un intervalle de largeur fixe le long d’une trajectoire  $\gamma$ . Plus précisément, pour  $\delta > 0$  fixé et  $\gamma \in W^{1,\infty}(0, T)$ , on pose

$$q_\gamma = \left\{ (x, t) \in Q; \quad |x - \gamma(t)| < \delta \right\}.$$

D’autre part, plutôt que de considérer le contrôle  $\widehat{u}_\gamma$  de norme  $L^2(q_\gamma)$  minimale, on lui préfère ici le contrôle  $\widehat{u}_\gamma$  de norme

$$\mathcal{U}_\gamma := L^2(q_\gamma; \chi_\gamma) = \left\{ u : q_\gamma \rightarrow \mathbb{R}; \quad \iint_{q_\gamma} u^2 \chi_\gamma < +\infty \right\}$$

minimale, où la fonction poids  $\chi_\gamma \in C^2(\overline{Q})$  vérifie  $\text{Supp } \chi_\gamma = \overline{q}_\gamma$ . Ce choix de “régulariser” l’action du contrôle permet d’obtenir un contrôle plus régulier que simplement  $L^2$ . Dans ce cadre, le problème à résoudre s’écrit

$$\min_{\gamma \in \mathcal{G}_{\text{ad}}} J(\gamma), \quad \text{avec} \quad J(\gamma) = \|\widehat{u}_\gamma\|_{\mathcal{U}_\gamma}^2,$$

et où l’ensemble admissible  $\mathcal{G}_{\text{ad}}$  regroupe des trajectoires  $\gamma$  uniformément  $M$ -lipschitziennes, de sorte que  $\{q_\gamma; \gamma \in \mathcal{G}_{\text{ad}}\} \subset \mathcal{Q}_{\text{ad}}^\varepsilon$  pour un certain  $\varepsilon = \varepsilon(\delta, M)$ . On montre que la fonctionnelle  $J$  atteint son minimum sur  $\mathcal{G}_{\text{ad}}$ . Cependant, on voit facilement que  $J$  n’est pas convexe. Il existe donc potentiellement plusieurs points de minimum globaux, ainsi que des minima locaux. Par ailleurs, on signale que l’étude d’une variante du problème d’optimisation du support peut être consultée à la Sect. 1.4.2. Dans cette variante, plutôt qu’une régularisation en espace du support de contrôle, on considère une régularisation en temps.

La troisième partie de l’article se concentre sur la résolution numérique du problème d’optimisation du support de contrôle. Pour des raisons pratiques, la contrainte de Lipschitz uniforme présente dans  $\mathcal{G}_{\text{ad}}$  est remplacée par un terme de pénalisation  $L^2$ . Pour  $\varepsilon > 0$  fixé, le problème devient alors

$$\min_{\gamma \in W^{1,\infty}} J_\varepsilon(\gamma), \quad \text{avec} \quad J_\varepsilon(\gamma) = J(\gamma) + \frac{\varepsilon}{2} \|\gamma'\|_{L^2(0, T)}^2.$$

On résout numériquement par un algorithme de gradient à pas fixe. À partir d’une trajectoire initiale  $\gamma_0$ , on construit une suite minimisante  $\gamma_n \rightarrow \gamma^*$  définie par  $\gamma_{n+1} = \gamma_n - \rho \phi_n^\varepsilon$ , avec  $\rho > 0$  fixé. À chaque itération, la direction de descente  $\phi_n^\varepsilon$  est évaluée à l’aide de la dérivée directionnelle de  $J$  en  $\gamma_n$ . En outre, le contrôle  $\widehat{u}_{\gamma_n}$  de norme  $\mathcal{U}_{\gamma_n}$  minimale est approché par éléments finis via la formulation mixte espace-temps exposée dans [8]. On présente quelques simulations pour différentes données initiales  $(y^0, y^1)$ . Comme attendu, la trajectoire numérique optimale  $\gamma^*$  dépend fortement de l’initialisation  $\gamma_0$ , du fait de la présence de minima locaux. L’influence du paramètre de régularisation  $\varepsilon$  est également discutée.

## Article

## OPTIMIZATION OF NON-CYLINDRICAL DOMAINS FOR THE EXACT NULL CONTROLLABILITY OF THE 1D WAVE EQUATION

ARTHUR BOTTOIS\*, NICOLAE CÎNDEA AND ARNAUD MÜNCH

**Abstract.** This work is concerned with the null controllability of the one-dimensional wave equation over non-cylindrical distributed domains. The controllability in that case has been obtained by Castro *et al.* [SIAM J. Control Optim. **52** (2014)] for domains satisfying the usual geometric optic condition. We analyze the problem of optimizing the non-cylindrical support  $q$  of the control of minimal  $L^2(q)$ -norm. In this respect, we prove a uniform observability inequality for a class of domains  $q$  satisfying the geometric optic condition. The proof based on the d'Alembert formula relies on arguments from graph theory. Numerical experiments are discussed and highlight the influence of the initial condition on the optimal domains.

**Mathematics Subject Classification.** 49Q10, 93C20.

Received July 16, 2020. Accepted January 19, 2021.

### 1. INTRODUCTION AND MAIN RESULTS

In the last decades, an important literature was devoted to the study of vibrating structures controlled or stabilized using piezoelectric actuators [1, 12, 27, 28]. In most of the situations, the actuators occupy a fixed position in space. Nevertheless, some studies consider the controllability of elastic structures (plates or beams) by the action of mobile piezoelectric actuators [33] which might be available in the future. Actually, a mobile actuator can be emulated by an array of fixed actuators distributed along the structure under the constraint that only one actuator can be activated at a given time [20]. The present work aims to find the optimal trajectory of a piezoelectric actuator in order to control, with a minimal cost, the vibrations of a simplified one-dimensional structure corresponding to a given initial datum. Therefore, observing the state of the considered system, we can provide in real time a trajectory for the controller in order to minimize the cost of control. More precisely, we consider the null controllability of a one-dimensional wave equation by means of distributed controls acting in non-cylindrical domains. The main contribution of this work is the proof of a uniform observability inequality with respect to control domains in a precise class. A direct consequence of this uniform observability inequality is the existence of an optimal trajectory for a piezoelectric actuator of a given size in order to control to zero the vibrations of the considered structure with a control of minimal  $L^2$ -norm.

---

*Keywords and phrases:* Wave equation, Uniform observability, Optimal shape.

Université Clermont Auvergne, Laboratoire de Mathématiques Blaise Pascal CNRS-UMR 6620, Campus des Cézeaux, 63178 Aubière cedex, France.

\* Corresponding author: [Arnaud.Munch@uca.fr](mailto:Arnaud.Munch@uca.fr)

© The authors. Published by EDP Sciences, SMAI 2021

Let  $T$  be a positive real,  $\Omega = (0, 1)$ ,  $Q_T$  the domain  $\Omega \times (0, T)$ ,  $q$  a non-empty subset of  $Q_T$  and  $\Sigma_T = \partial\Omega \times (0, T)$ . We are concerned with the null distributed controllability of the 1D wave equation:

$$\begin{cases} y_{tt} - y_{xx} = v\mathbb{1}_q & \text{in } Q_T, \\ y = 0 & \text{on } \Sigma_T, \\ (y, y_t)(\cdot, 0) = (y_0, y_1) & \text{in } \Omega. \end{cases} \quad (1.1)$$

where  $y$  represents the transverse deflection of a simplified elastic structure of length 1 and  $v$  is the voltage applied to an actuator occupying the position  $q_t = q \cap \{(x, t); x \in \Omega\}$  at time  $t$ .

We assume that  $(y_0, y_1) \in \mathbf{V} := H_0^1(\Omega) \times L^2(\Omega)$ ;  $v$  is the control (a function in  $L^2(q)$ ) and  $y = y(x, t)$  is the associated state.  $\mathbb{1}_q$  from  $Q_T$  to  $\{0, 1\}$  denotes the indicator function of  $q$ . We also use the notation

$$Ly := y_{tt} - y_{xx}.$$

For any  $(y_0, y_1) \in \mathbf{V}$  and any  $v \in L^2(q)$ , there exists exactly one solution  $y$  to (1.1), with the following regularity  $y \in C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$  (see [18]). The null controllability problem for (1.1) at time  $T$  is the following: for each  $(y_0, y_1) \in \mathbf{V}$ , find a control  $v \in L^2(q)$  such that the corresponding solution  $y$  of (1.1) satisfies

$$(y, y_t)(\cdot, T) = (0, 0) \quad \text{in } \Omega.$$

As a consequence of the *Hilbert uniqueness method* introduced by J.-L. Lions [18], the null controllability of (1.1) is equivalent to an observability inequality for the associated adjoint problem: there exists a constant  $C_{\text{obs}}(q) > 0$  such that

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{W}}^2 \leq C_{\text{obs}}(q) \|\varphi\|_{L^2(q)}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W} := L^2(\Omega) \times H^{-1}(\Omega), \quad (1.2)$$

where  $(\varphi, \varphi_0, \varphi_1)$  solves

$$L\varphi = 0 \text{ in } Q_T, \quad \varphi = 0 \text{ on } \Sigma_T, \quad (\varphi, \varphi_t)(\cdot, 0) = (\varphi_0, \varphi_1) \text{ in } \Omega. \quad (1.3)$$

$C_{\text{obs}}(q)$  is the observability constant associated with the adjoint solution  $\varphi$  and depends on the control region  $q$ .

We address in this work the following extremal problem:

$$\inf_{q \in \mathcal{Q}_{\text{ad}}} \|v\|_{L^2(q)} \quad (1.4)$$

where  $v$  is the control of minimal  $L^2(q)$ -norm. The set  $\mathcal{Q}_{\text{ad}}$  is defined below and denotes an admissible class of non-cylindrical domains  $q$ , that is domains which evolve with respect to the time variable.

In the cylindrical case for which the control region  $q$  takes the form  $q := \omega \times (0, T)$ , where  $\omega$  denotes a subset of  $\Omega$ , the null controllability of (1.1) at any large  $T > T^*$  is well-known (for instance, see [2, 18]). The critical time  $T^*$  is related to the measure of the set  $\Omega \setminus \bar{\omega}$ . The controllability is also true in the multi-dimensional case assuming the celebrated geometric optic condition on the triplet  $(\Omega, \omega, T)$  introduced by Bardos, Lebeau and Rauch in [2]. The optimal problem (1.4) has been studied by Periago in [29] with  $\mathcal{Q}_{\text{ad}} = \mathcal{Q}_{\text{ad}, \tau} := \{q = \omega \times (0, T); \omega \subset \Omega, |\omega| = \tau|\Omega|, T \geq 2\}$  parameterized by  $\tau$  in  $(0, 1)$ . In particular, the analysis is based on the following uniform property proved by using Fourier series: there exists a constant  $C > 0$  which depends only on  $Q_T$  and  $\tau$  such that

$$\sup_{q \in \mathcal{Q}_{\text{ad}, \tau}} C_{\text{obs}}(q) < C. \quad (1.5)$$

We also mention that the problem of the optimal shape and position of the support has been numerically investigated in [23, 24] for the one and two dimensional wave equation. This problem is also considered by Privat, Trélat and Zuazua in [30] using spectral arguments. We also mention [31] which analyze the dependence of the observability constant with respect to the control region.

The non-cylindrical case is more involved on a mathematical viewpoint. It concerns however an increasing number of studies as it allows in many situations to obtain positive controllability results, in contrast to the cylindrical case. Focusing on the wave equation, one of the first contributions is due to Khapalov [16] providing observability results for a moving point sensor in the interior of the domain. This time-dependent observation allows to avoid the usual difficulties related to strategic or non-strategic points. In particular, in the 1D setting, for any  $T > 0$ , the existence of controls continuous almost everywhere in  $(0, T)$  and supported over curves continuous almost everywhere is obtained for data in  $H^2(\Omega) \cap H_0^1(\Omega) \times H_0^1(\Omega)$ . More recently, for initial data in  $\mathbf{W}$ , Castro analyzes in [5] the controllability from a moving interior point. By the way of the d'Alembert formula, a uniform observability inequality is proved for a precise set of curves  $\{(\gamma(t), t)\}_{t \in [0, T]}$  leading to moving controls in  $H^{-1}(\cup_{t \in (0, T)} \gamma(t) \times t)$ . Still in the 1D setting, Ciu, Liu and Gao [11] and Haak and Hoang [13] analyze the boundary and moving interior point observability for wave equations posed on a sufficiently regular time-dependent domain. In the  $N$ -dimensional case, Liu and Yong [19] employ the multiplier method to prove that the wave equation is controllable under the hypothesis that the distributed control region  $q$  covers the whole space domain  $\Omega$  before the time  $T$ . Under a similar hypothesis, we also mention the work [21] where the controllability of the damped wave equation  $y_{tt} - y_{xx} - \varepsilon y_{txx} = 0$  defined on the 1D torus is established. Because of the presence of an essential spectrum, such property does not hold true in the cylindrical case.

In the 1D setting, the geometric assumption has been relaxed in [6]. Precisely, the observability inequality (1.2) is obtained assuming that the distributed control domain  $q \subset Q_T$  is a finite union of connected open sets and satisfies the following hypothesis: any characteristic line starting from a point of  $\bar{\Omega} \times \{0\}$  and following the laws of geometric optics when reflected on the boundary  $\Sigma_T$  must meet the domain  $q$  (we refer to [6], Prop. 2.1). This geometric condition is the natural extension to non-cylindrical domains of the condition introduced in [2]. Following [5], the proof of this result is obtained by the way of the d'Alembert formula. It has been extended in [17] to the multi-dimensional case using microlocal analysis. We also mention the obtention of Carleman-type inequalities for general hyperbolic equations in [32].

With the aim to extend [29] to a non-cylindrical setting, we prove in a first part that the observability constant  $C_{\text{obs}}(q)$  in (1.2) is uniformly bounded from above for a class of domains satisfying the geometric optic condition. More precisely, for any parameter  $\varepsilon > 0$  small enough, we introduce the admissible set of control domains by

$$\mathcal{Q}_{\text{ad}}^\varepsilon = \{q \subset Q_T; \quad q \text{ open and } q^\varepsilon \text{ verifies the geometric optic condition}\}. \quad (1.6)$$

Here,  $q^\varepsilon = q^\varepsilon(q)$  denotes the  $\varepsilon$ -interior of any  $q \subset Q_T$  and is defined by

$$q^\varepsilon = \{(x, t) \in q; \quad d((x, t), \partial q) > \varepsilon\}. \quad (1.7)$$

We obtain that, for any  $\varepsilon > 0$ , there exists a constant  $C > 0$  which depends only on  $Q_T$  and  $\varepsilon$  such that

$$\sup_{q \in \mathcal{Q}_{\text{ad}}^\varepsilon} C_{\text{obs}}(q) < C. \quad (1.8)$$

As in [29], this uniform property allows, in a second part, to analyze the problem (1.4) of the optimal distribution of the control domain  $q$ . Preliminarily, for numerical purposes, we restrict the analysis to control domains of the form

$$q_\gamma = \{(x, t) \in Q_T; \quad |x - \gamma(t)| < \delta_0\} \quad (1.9)$$

for a given  $\delta_0 > 0$  and curves  $\gamma : (0, T) \rightarrow \Omega$  in the following set

$$\mathcal{G}_{\text{ad}} = \{\gamma \in W^{1,\infty}(0, T); \quad \|\gamma'\|_{L^\infty(0, T)} \leq M, \quad \delta_0 \leq \gamma \leq 1 - \delta_0\} \quad (1.10)$$

consisting of uniformly Lipschitz functions of fixed constant  $M > 0$ . For  $T \geq 2$  and  $\varepsilon > 0$  small enough, the class  $\{q_\gamma; \gamma \in \mathcal{G}_{\text{ad}}\}$  is a subset of  $\mathcal{Q}_{\text{ad}}^\varepsilon$ . In this setting, the optimal problem (1.4) reads as follows: for  $\delta_0 > 0$ ,  $M > 0$  and a given initial datum  $(y_0, y_1) \in \mathbf{V}$ , solve

$$\inf_{\gamma \in \mathcal{G}_{\text{ad}}} \|v\|_{L^2(q_\gamma)}^2 \quad (1.11)$$

where  $v$  is the control of minimal  $L^2(q_\gamma)$ -norm distributed over  $q_\gamma \subset Q_T$ .

This paper is organized as follows. In Section 2, we prove the uniform observability property (1.8) on  $\mathcal{Q}_{\text{ad}}^\varepsilon$  and its variant on the subset  $\{q_\gamma; \gamma \in \mathcal{G}_{\text{ad}}\}$  (see Thm. 2.1). This is achieved by defining an appropriate decomposition of the observation domains in  $\mathcal{Q}_{\text{ad}}^\varepsilon$ , and by using the d'Alembert formula. The proof also relies on arguments from graph theory. Then, in Section 3, following arguments from [15, 29], we analyze a variant of the extremal problem (1.4). Introducing a  $C^1$ -regularization of the support  $q_\gamma$ , we prove that the underlying cost is continuous over  $\mathcal{G}_{\text{ad}}$  for the  $L^\infty(0, T)$ -norm, and admits at least one local minimum (see Prop. 2). Section 4 illustrates the result with numerical experiments for the regularized minimization problem introduced in Section 3.3. Minimization sequences for the regularized cost are constructed using a gradient method: each iteration requires the computation of a null control, performed using the space-time formulation developed in [9] and used in [6], well-suited to the description of the non-cylindrical domains where the control acts. As expected, the optimal domain we obtain are closely related to the travelling waves generated by the initial condition. Section 5 concludes the work with some perspectives.

## 2. UNIFORM OBSERVABILITY WITH RESPECT TO THE DOMAIN OF OBSERVATION

We prove in this section the uniform observability property (1.8) with respect to the domain of observation. Precisely, we prove the following equivalent result for regular data in  $\mathbf{V}$ .

**Theorem 2.1.** *Let  $T > 0$  and let  $\varepsilon > 0$  be a small enough fixed parameter such that the set  $\mathcal{Q}_{\text{ad}}^\varepsilon$  defined by (1.6) is non-empty. There exists a constant  $C_{\text{obs}}^\varepsilon > 0$  such that for every  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , the following inequality holds*

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{V}}^2 \leq C_{\text{obs}}^\varepsilon \|\varphi_t\|_{L^2(q)}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{V}, \quad (2.1)$$

where  $\varphi$  is the solution of the wave equation (1.3) associated with the initial datum  $(\varphi_0, \varphi_1)$ .

We emphasize that the constant  $C_{\text{obs}}^\varepsilon$  does not depend on  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . We also emphasize that the parameter  $\varepsilon$  is introduced here to quantify to which extent a domain  $q$  satisfies the geometric optic condition used in [6]. As  $\varepsilon$  goes to zero, the observability constant  $C_{\text{obs}}^\varepsilon$  blows up (see Rem. 2.18). The same is true when  $\tau$  goes to zero in the study [29], see (1.5). Eventually, we emphasize that any domain  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  for some  $\varepsilon$  satisfies the geometric optic condition. Conversely, for any domain  $q$  satisfying the geometric optic condition, there exists  $\varepsilon > 0$  such that  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ .

In the remaining part of this section, we assume that the hypotheses of Theorem 2.1 are satisfied.

### 2.1. Proof of Theorem 2.1

The idea of the proof is to decompose the domain  $Q_T$  as the sum of “elementary squares” on which, in view of the d'Alembert formula, the terms in the inequality (2.1) can be simply computed and compared. The proof based on arguments from graph theory allows notably to relate the value of the observability constant to the spectrum of a Laplacian matrix, defined in terms of the graph associated to any domain  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ .

The proof requires some notations and technical lemmas. We decompose it into seven steps.

**Step 1** – Solution of the adjoint problem (1.3) associated with piecewise affine initial conditions.

Let  $N \in \mathbb{N}^*$  and  $\kappa_N = 1/N$ . We denote by  $S_N = (x_i^N)_{0 \leq i \leq N}$  the regular subdivision of  $\bar{\Omega}$  in  $N$  intervals such that  $x_i^N = i/N$ . For any  $(\varphi_0, \varphi_1) \in \mathbf{V}$ , we then associate the functions  $\varphi_0^N, \varphi_1^N$  as follows: for all  $x \in \Omega$ , we set

$$\varphi_0^N(x) = \sum_{i=1}^N \left( \varphi_0(x_i^N) \frac{x - x_{i-1}^N}{\kappa_N} + \varphi_0(x_{i-1}^N) \frac{x_i^N - x}{\kappa_N} \right) \mathbb{1}_{[x_{i-1}^N, x_i^N]}(x), \quad (2.2)$$

$$\varphi_1^N(x) = \sum_{i=1}^N \beta_i^N \mathbb{1}_{[x_{i-1}^N, x_i^N]}(x), \quad \text{with } \beta_i^N = \frac{1}{\kappa_N} \int_{x_{i-1}^N}^{x_i^N} \varphi_1, \quad (2.3)$$

which are respectively affine and constant on each interval  $[x_{i-1}^N, x_i^N]$ . We also denote by  $(\varphi_0^N)' \in L^2(\Omega)$  the “derivative” of  $\varphi_0^N$ :

$$(\varphi_0^N)'(x) = \sum_{i=1}^N \alpha_i^N \mathbb{1}_{[x_{i-1}^N, x_i^N]}(x), \quad \text{with } \alpha_i^N = \frac{\varphi_0(x_i^N) - \varphi_0(x_{i-1}^N)}{\kappa_N}.$$

Using that  $\varphi_0 \in H_0^1(\Omega)$ , we easily check that

$$\|\varphi_0^N\|_{H_0^1(\Omega)}^2 = \frac{1}{N} \sum_{i=1}^N (\alpha_i^N)^2, \quad \|\varphi_1^N\|_{L^2(\Omega)}^2 = \frac{1}{N} \sum_{i=1}^N (\beta_i^N)^2 \quad \text{and} \quad \sum_{i=1}^N \alpha_i^N = 0.$$

In order to use the d'Alembert formula for the solution of (1.3) associated with the initial datum  $(\varphi_0^N, \varphi_1^N)$ , we extend these functions as odd functions to  $[-1, 1]$  and then by 2-periodicity to  $\mathbb{R}$ . In this respect, we first extend the definition of  $x_i^N$  to  $i \in \mathbb{Z}$  by putting  $x_i^N = i/N$  for every  $i \in \mathbb{Z}$ , and then denote by  $I_i^N$  for every  $i \in \mathbb{Z}^*$  the following interval:

$$I_i^N = \begin{cases} [x_{i-1}^N, x_i^N] & \text{if } i > 0, \\ [x_i^N, x_{i+1}^N] & \text{if } i < 0. \end{cases} \quad (2.4)$$

**Definition 2.2.** For any  $N \in \mathbb{N}^*$  and  $i \in \mathbb{Z}^*$ , we define the integer  $j_N(i)$  as follows:

$$\text{if } i > 0, \quad j_N(i) = \begin{cases} (i-1) \bmod (2N) + 1 & \text{if } (i-1) \bmod (2N) < N, \\ (i-1) \bmod (2N) - 2N & \text{if } (i-1) \bmod (2N) \geq N, \end{cases} \quad (2.5)$$

and if  $i < 0$ ,  $j_N(i) = -j_N(-i)$ . Remark that for every  $i \in \mathbb{Z}^*$ ,  $j_N(i) \in \mathbb{I}_N$  with

$$\mathbb{I}_N = \{-N, \dots, -1, 1, \dots, N\}. \quad (2.6)$$

Similarly, we extend  $\alpha_i^N$  and  $\beta_i^N$  for every  $i \in \mathbb{Z}^*$  as follows: if  $i \in \{-N, \dots, -1\}$ , we set  $\alpha_i^N = \alpha_{-i}^N$  and  $\beta_i^N = -\beta_{-i}^N$ ; if  $|i| > N$ , we set  $\alpha_i^N = \alpha_{j_N(i)}^N$  and  $\beta_i^N = \beta_{j_N(i)}^N$ .

**Definition 2.3.** For any  $N \in \mathbb{N}^*$  and  $i \in \mathbb{I}_N$ , let

$$\gamma_i^N = \alpha_i^N + \beta_i^N. \quad (2.7)$$

We extend the functions  $\varphi_0^N$  and  $\varphi_1^N$  as odd functions to  $[-1, 1]$  and by 2-periodicity to  $\mathbb{R}$ .



**Lemma 2.4.** For any  $N \in \mathbb{N}^*$ , the solution  $\varphi^N$  of (1.3) associated with the initial datum  $(\varphi_0^N, \varphi_1^N)$  satisfies:

$$\varphi_t^N(x, t) = \frac{1}{2} \sum_{i \in \mathbb{Z}^*} \sum_{j \in \mathbb{Z}^*} (\gamma_{j_N(i)}^N - \gamma_{-j_N(j)}^N) \mathbb{1}_{I_i^N}(x+t) \mathbb{1}_{I_j^N}(x-t), \quad \forall (x, t) \in Q_T. \quad (2.8)$$

*Proof.* Using the notations above, we obtain

$$(\varphi_0^N)'(x) = \sum_{i \in \mathbb{Z}^*} \alpha_i^N \mathbb{1}_{I_i^N}(x), \quad \varphi_1^N(x) = \sum_{i \in \mathbb{Z}^*} \beta_i^N \mathbb{1}_{I_i^N}(x), \quad \forall x \in \mathbb{R}. \quad (2.9)$$

Moreover, from the d'Alembert formula, the solution  $\varphi^N$  of (1.3) associated with the initial datum  $(\varphi_0^N, \varphi_1^N)$  is given as follows:

$$\varphi^N(x, t) = \frac{1}{2} \left( \varphi_0^N(x+t) + \varphi_0^N(x-t) \right) + \frac{1}{2} \int_{x-t}^{x+t} \varphi_1^N, \quad \forall (x, t) \in Q_T. \quad (2.10)$$

Taking the derivative with respect to  $t$  and replacing the expressions (2.9) in the above equation, we deduce that for all  $(x, t) \in Q_T$ , we have

$$\begin{aligned} \varphi_t^N(x, t) &= \frac{1}{2} \left( (\varphi_0^N)'(x+t) - (\varphi_0^N)'(x-t) + \varphi_1^N(x+t) + \varphi_1^N(x-t) \right) \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}^*} \left( (\alpha_i^N + \beta_i^N) \mathbb{1}_{I_i^N}(x+t) - (\alpha_i^N - \beta_i^N) \mathbb{1}_{I_i^N}(x-t) \right) \\ &= \frac{1}{2} \sum_{i \in \mathbb{Z}^*} \sum_{j \in \mathbb{Z}^*} (\alpha_i^N + \beta_i^N - \alpha_j^N + \beta_j^N) \mathbb{1}_{I_i^N}(x+t) \mathbb{1}_{I_j^N}(x-t). \end{aligned}$$

Using the properties of the function  $j_N$  defined in (2.5), we deduce that for  $i, j \in \mathbb{Z}^*$ ,

$$\begin{aligned} \alpha_i^N + \beta_i^N - \alpha_j^N + \beta_j^N &= \alpha_{j_N(i)}^N + \beta_{j_N(i)}^N - \alpha_{j_N(j)}^N + \beta_{j_N(j)}^N \\ &= (\alpha_{j_N(i)}^N + \beta_{j_N(i)}^N) - (\alpha_{-j_N(j)}^N + \beta_{-j_N(j)}^N) \\ &= \gamma_{j_N(i)}^N - \gamma_{-j_N(j)}^N \end{aligned}$$

and the result.  $\square$

**Step 2** – Representation of  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  using “elementary squares”.

In view of the expression (2.8), we introduce the following definition.

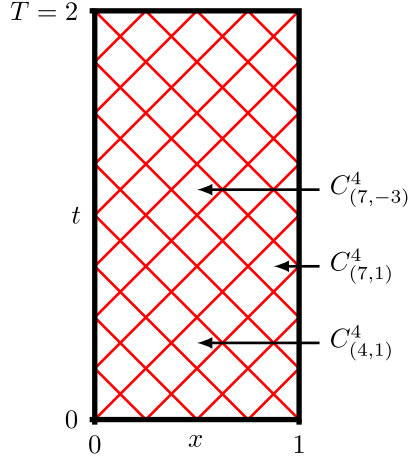
**Definition 2.5.** For any  $N \in \mathbb{N}^*$  and  $i, j \in \mathbb{Z}^*$ , we define the *elementary square* of indices  $(i, j)$  associated with the subdivision  $S_N$  as the following closed set of  $\mathbb{R}^2$ :

$$C_{(i,j)}^N = \{(x, t) \in \mathbb{R}^2 \text{ such that } x+t \in I_i^N \text{ and } x-t \in I_j^N\},$$

where the interval  $I_i^N$  is given by (2.4).

We also denote by  $\mathcal{C}_N = \{C_{(i,j)}^N; i, j \in \mathbb{Z}^*\}$  the set of all the elementary squares associated with the subdivision  $S_N$ . We easily check that  $\mathbb{R}^2 = \bigcup_{i,j \in \mathbb{Z}^*} C_{(i,j)}^N$ .

Figure 1 illustrates the way the elementary squares are indexed, with the example of the subdivision  $S_4$ .

FIGURE 1. Some elementary squares in  $\mathcal{C}_4$ .

**Remark 2.6.** For every  $i, j \in \mathbb{Z}^*$ , the coordinates of the center of the elementary square  $C_{(i,j)}^N$  associated with the subdivision  $S_N$  are given by

$$\begin{cases} x_{(i,j)}^N = \frac{m_i^N + m_j^N}{2}, \\ t_{(i,j)}^N = \frac{m_i^N - m_j^N}{2}, \end{cases} \quad \text{with} \quad m_i^N = \begin{cases} \frac{x_{i-1}^N + x_i^N}{2} & \text{if } i > 0, \\ \frac{x_i^N + x_{i+1}^N}{2} & \text{if } i < 0. \end{cases}$$

The area of every elementary square  $C_{(i,j)}^N \in \mathcal{C}_N$  is given by  $|C_{(i,j)}^N| = \frac{1}{2N^2}$ . Notice that for every  $i, j \in \mathbb{Z}^*$  with  $|i|, |j| > 1$ , the elementary squares having one side in common with the elementary square  $C_{(i,j)}^N$  are  $C_{(i\pm 1, j)}^N$  and  $C_{(i, j\pm 1)}^N$ .

**Definition 2.7.** For every  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , we denote by  $\mathcal{C}_N(q)$  and  $\mathcal{C}_N(Q_T)$  the sets of the elementary squares in  $\mathcal{C}_N$  with their interior included in  $q$  and  $Q_T$  respectively:

$$\mathcal{C}_N(q) = \{C_{(i,j)}^N \in \mathcal{C}_N; \overset{\circ}{C}_{(i,j)}^N \subset q\}, \quad \mathcal{C}_N(Q_T) = \{C_{(i,j)}^N \in \mathcal{C}_N; \overset{\circ}{C}_{(i,j)}^N \subset Q_T\}.$$

If  $N$  is large enough, the sets  $\mathcal{C}_N(q)$  and  $\mathcal{C}_N(Q_T)$  are non-empty.

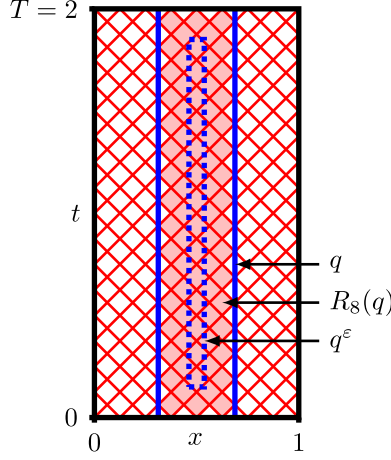
We also define  $R_N(q)$  as the union of the elementary squares in  $\mathcal{C}_N(q)$ :

$$R_N(q) = \overbrace{\bigcup_{C_{(i,j)}^N \in \mathcal{C}_N(q)} C_{(i,j)}^N}^{\circ}. \quad (2.11)$$

With these notations, we can prove the following lemma.

**Lemma 2.8.** Let  $N > 1/\varepsilon$  be a fixed integer. For every  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , the set  $\bigcup_{C_{(i,j)}^N \in \mathcal{C}_N(q)} C_{(i,j)}^N$  is a cover of the  $\varepsilon$ -interior  $q^\varepsilon$  of  $q$  (see (1.7)). Moreover, the set  $R_N(q)$  defined by (2.11) satisfies  $q^\varepsilon \subset R_N(q) \subset q$ .

*Proof.* Let  $X \in q^\varepsilon$ . By definition of  $q^\varepsilon$ , we have  $X \in q$  and  $d(X, \partial q) > \varepsilon$ . Since  $\mathbb{R}^2$  is covered by squares in  $\mathcal{C}_N$ , there exists  $C_{(i,j)}^N \in \mathcal{C}_N$  such that  $X \in C_{(i,j)}^N$ . Moreover, since  $\text{diam}(C_{(i,j)}^N) = \kappa_N$ , we have  $C_{(i,j)}^N \subset \overline{B}(X, \kappa_N)$ .

FIGURE 2. Cover  $R_8(q)$  of  $q^\varepsilon$ , for  $\varepsilon = 0.15$ .

Let  $Y \in \overline{B}(X, \kappa_N)$ . Then, for every  $Z \in \mathbb{R}^2 \setminus q$ , it follows that

$$d(Y, Z) \geq |d(Y, X) - d(X, Z)| > \varepsilon - \kappa_N > 0.$$

Consequently,  $d(Y, \mathbb{R}^2 \setminus q) > 0$  which implies  $Y \in q$ . Therefore,  $C_{(i,j)}^N \subset \overline{B}(X, \kappa_N) \subset q$  and, finally,  $C_{(i,j)}^N \in \mathcal{C}_N(q)$ .  $\square$

Figure 2 illustrates Lemma 2.8 in the case of the cylindrical observation domain  $q = (\frac{5}{16}, \frac{11}{16}) \times (0, 2)$ , for  $\varepsilon = 0.15$  and  $N = 8$ .

**Step 3** – Weighted graph associated with  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ .

In order to write several expressions in a simpler form, we use elements from graph theory.

More precisely, the introduction of a weighted graph in Definition 2.9 allows to reformulate the observability inequality for the solution  $\varphi^N$  in terms of a spectral inequality involving the Laplacian matrix of the graph (see Def. 2.11). In this way, the observability constant associated with  $\varphi^N$  is linked to the first strictly positive eigenvalue of the Laplacian matrix.

**Definition 2.9.** Let  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  an observation domain. We define the weighted graph  $G_N(q)$  as follows:

- $\mathbb{I}_N$  given by (2.6) is the set of vertices;
- for every  $i \in \mathbb{I}_N$ , the degree of the vertex  $i$  is given by:

$$d_i^N = \text{Card} \left( \{C_{(k,-l)}^N \in \mathcal{C}_N(q); \quad i \in \{j_N(k), j_N(l)\}\} \right);$$

- for every  $i, j \in \mathbb{I}_N$ , the weight of the edge linking the vertices  $i$  and  $j$  is

$$w_{i,j}^N = w_{j,i}^N = \text{Card} \left( \{C_{(k,-l)}^N \in \mathcal{C}_N(q); \quad \{i, j\} = \{j_N(k), j_N(l)\}\} \right).$$

**Definition 2.10.** Let  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  and let  $i, j \in \mathbb{I}_N$  be two vertices of the graph  $G_N(q)$ . We say that there is a path from  $i$  to  $j$  in  $G_N(q)$  and we denote  $i \sim j$  if the vertices  $i$  and  $j$  are in the same connected component of  $G_N(q)$ . In particular, if  $w_{i,j}^N \neq 0$ , then  $i \sim j$ .

We then recall the definition of the Laplacian matrix associated with a graph (see [4, 8]).

**Definition 2.11.** Let  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . The *Laplacian matrix* associated with the graph  $G_N(q)$  (see Def. 2.9) is the symmetric positive matrix  $A_N(q) \in \mathcal{M}_{2N}(\mathbb{R})$  defined by

$$A_N(q) = \begin{pmatrix} d_{-N}^N & \cdots & -w_{-N,-1}^N & -w_{-N,1}^N & \cdots & -w_{-N,N}^N \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -w_{-1,-N}^N & \cdots & d_{-1}^N & -w_{-1,1}^N & \cdots & -w_{-1,N}^N \\ -w_{1,-N}^N & \cdots & -w_{1,-1}^N & d_1^N & \cdots & -w_{1,N}^N \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -w_{N,-N}^N & \cdots & -w_{N,-1}^N & -w_{N,1}^N & \cdots & d_N^N \end{pmatrix}_{2N \times 2N}. \quad (2.12)$$

**Remark 2.12.** Remark that for every  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ , the graph  $G_N(q)$  has no loop, *i.e.*  $w_{i,i}^N = 0$  for every  $i \in \mathbb{I}_N$ . Indeed, the elementary squares  $C_{(k,-l)}^N$  such that  $j_N(k) = j_N(l) = i$  have their center  $x_{(k,-l)}^N \in \mathbb{Z}$  and, consequently, cannot be in  $Q_T$ .

Remark also that the Laplacian matrix  $A_N(q)$  of the graph  $G_N(q)$  verifies the following property (see [4, 8]): for every  $\eta = (\eta_{-N}, \dots, \eta_{-1}, \eta_1, \dots, \eta_N) \in \mathbb{R}^{2N}$ ,

$$\eta^T A_N(q) \eta = \sum_{i \in \mathbb{I}_N} d_i^N \eta_i^2 - \sum_{i,j \in \mathbb{I}_N} w_{i,j}^N \eta_i \eta_j = \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} (\eta_{j_N(i)} - \eta_{-j_N(j)})^2. \quad (2.13)$$

From now on, we consider that the assumption of Lemma 2.8 holds true, *i.e.* we take  $N > 1/\varepsilon$ . More precisely, we fix  $N$  the smallest integer strictly greater than  $1/\varepsilon$ .

**Lemma 2.13.** *Let  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . Then the associated graph  $G_N(q)$  is connected.*

*Proof.* Let  $i \in \{1, \dots, N-1\}$ . We denote by  $D_i^+$  the support of the characteristic line " $x+t = x_i^N$ ", starting from  $x_i^N$  in the direction of decreasing  $x$  and following the laws of geometric optics for its reflection on  $\Sigma_T$ . Since  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ ,  $q^\varepsilon$  satisfies the geometric optic condition; consequently, there exists  $(x^*, t^*) \in q^\varepsilon \cap D_i^+$ . From Lemma 2.8, we have  $q^\varepsilon \subset R_N(q)$ , so  $(x^*, t^*)$  belongs to the common side of two elementary squares in  $\mathcal{C}_N(q)$ :

$$\left( C_{(k,l)}^N \text{ and } C_{(k+1,l)}^N \text{ with } j_N(k) = i \right) \quad \text{or} \quad \left( C_{(l,k)}^N \text{ and } C_{(l,k-1)}^N \text{ with } j_N(k) = -i \right).$$

Therefore  $i \stackrel{\sim}{\sim} i+1$  and, so, the vertices  $\{1, \dots, N\}$  are in the same connected component.

Similarly, we denote by  $D_i^-$  the support of the characteristic line " $x-t = x_i^N$ ", starting from  $x_i^N$  in the direction of increasing  $x$  and following the laws of geometric optics for its reflection on  $\Sigma_T$ . Since  $q^\varepsilon$  satisfies the geometric optic condition, there exists  $(x^*, t^*) \in q^\varepsilon \cap D_i^-$ . From Lemma 2.8, we have  $q^\varepsilon \subset R_N(q)$ , so  $(x^*, t^*)$  belongs to the common side of two elementary squares in  $\mathcal{C}_N(q)$ :

$$\left( C_{(l,k)}^N \text{ and } C_{(l,k+1)}^N \text{ with } j_N(k) = i \right) \quad \text{or} \quad \left( C_{(k,l)}^N \text{ and } C_{(k-1,l)}^N \text{ with } j_N(k) = -i \right).$$

Therefore  $-i \stackrel{\sim}{\sim} -i-1$  and, so, the vertices  $\{-N, \dots, -1\}$  are in the same connected component.

It remains to show that the vertices  $N$  and  $-N$  belong to the same connected component. We denote by  $D_N^+$  the support of the characteristic line " $x+t = x_N^N$ ", starting from  $x_N^N$  in the direction of decreasing  $x$  and following the laws of geometric optics for its reflection on  $\Sigma_T$ . Since  $q^\varepsilon$  satisfies the geometric optic condition, there exists  $(x^*, t^*) \in q^\varepsilon \cap D_N^+$ . From Lemma 2.8, we have  $q^\varepsilon \subset R_N(q)$ , so  $(x^*, t^*)$  belongs to the common side of two elementary squares in  $\mathcal{C}_N(q)$ :

$$\left( C_{(k,l)}^N \text{ and } C_{(k+1,l)}^N \text{ with } j_N(k) = N \right) \quad \text{or} \quad \left( C_{(l,k)}^N \text{ and } C_{(l,k-1)}^N \text{ with } j_N(k) = -N \right).$$

Hence,  $N \simeq -N$ . □

**Remark 2.14.** A well-known graph theory result (see, for instance, [4], Prop. 1.3.7) states that the graph  $G_N(q)$  is connected if and only if  $\dim(\ker(A_N(q))) = 1$ . Moreover, if  $G_N(q)$  is connected, then  $\ker(A_N(q)) = \text{Vect}(\mathbf{1}_{2N})$ , where  $\mathbf{1}_{2N}$  is the vector in  $\mathbb{R}^{2N}$  with all its component equal to 1.

Let us denote  $\lambda_N(q) > 0$  the smallest non-zero eigenvalue of the matrix  $A_N(q)$ . This eigenvalue is known in graph theory as the *algebraic connectivity* of the graph. We also define  $\lambda_N$  by

$$\lambda_N = \min_{q \in \mathcal{Q}_{\text{ad}}^\varepsilon} \lambda_N(q) > 0. \quad (2.14)$$

Note that since the set  $\{G_N(q); q \in \mathcal{Q}_{\text{ad}}^\varepsilon\}$  has a finite number of elements,  $\lambda_N$  is well defined.

**Step 4** –  $p$ -refinement of the graph and connection with the spectrum of the Laplacian matrix.

**Definition 2.15.** For every  $p \in \mathbb{N}^*$ , we denote by  $\mathcal{C}_N^p(q)$  the set formed by the elementary squares associated with the subdivision  $S_{pN}$  and having their interior in  $R_N(q)$ :

$$\mathcal{C}_N^p(q) = \{C_{(i,j)}^{pN} \in \mathcal{C}_{pN}; \quad \mathring{C}_{(i,j)}^{pN} \subset R_N(q)\}.$$

We then define the graph  $G_N^p(q)$  following Definition 2.9, substituting  $N$  by  $pN$  and substituting  $\mathcal{C}_N(q)$  by  $\mathcal{C}_N^p(q)$  in the definitions of the vertex degrees and the edge weights. Finally, we denote  $A_N^p(q) \in \mathcal{M}_{2pN}(\mathbb{R})$  the Laplacian matrix associated with the graph  $G_N^p(q)$ . This matrix has the following block form:

$$A_N^p(q) = \begin{pmatrix} d_{-N}^N p I_p & \cdots & -w_{-N,-1}^N J_p & -w_{-N,1}^N J_p & \cdots & -w_{-N,N}^N J_p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -w_{-1,-N}^N J_p & \cdots & d_{-1}^N p I_p & -w_{-1,1}^N J_p & \cdots & -w_{-1,N}^N J_p \\ -w_{1,-N}^N J_p & \cdots & -w_{1,-1}^N J_p & d_1^N p I_p & \cdots & -w_{1,N}^N J_p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -w_{N,-N}^N J_p & \cdots & -w_{N,-1}^N J_p & -w_{N,1}^N J_p & \cdots & d_N^N p I_p \end{pmatrix}_{2pN \times 2pN}, \quad (2.15)$$

where  $I_p, J_p \in \mathcal{M}_p(\mathbb{R})$  are respectively the identity matrix and the matrix with all its elements equal to 1.

**Definition 2.16.** For any  $i \in \mathbb{Z}^*$ , we define the set  $\mathbb{J}_i^p$  by

$$\mathbb{J}_i^p = \begin{cases} \{p(i-1)+1, \dots, pi\} & \text{if } i > 0, \\ \{pi, \dots, p(i+1)-1\} & \text{if } i < 0. \end{cases} \quad (2.16)$$

For every  $\eta = (\eta_{-pN}, \dots, \eta_{-1}, \eta_1, \dots, \eta_{pN}) \in \mathbb{R}^{2pN}$ , we check that

$$\eta^T A_N^p(q) \eta = \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} \sum_{i' \in \mathbb{J}_i^p} \sum_{j' \in \mathbb{J}_j^p} (\eta_{i_{pN}(i')} - \eta_{j_{pN}(j')})^2.$$

For any  $p \in \mathbb{N}^*$ , the following lemma makes the link between the spectrum of the Laplacian matrix  $A_N^p(q)$  (see Def. 2.15) and the spectrum of the Laplacian matrix  $A_N(q)$  (see Def. 2.11).

**Lemma 2.17.** *Let  $p \in \mathbb{N}^*$ . The spectrum of the Laplacian matrix  $\frac{1}{p} A_N^p(q)$  (see (2.15)) is composed of the spectrum of the Laplacian matrix  $A_N(q)$  (see (2.12)), and the diagonal elements of  $A_N(q)$  repeated  $p-1$  times. Moreover,  $\dim(\ker(A_N^p(q))) = 1$  and  $\ker(A_N^p(q)) = \text{Vect}(\mathbf{1}_{2pN})$ .*

*Proof.* Let  $\eta = (\eta_{-pN}, \dots, \eta_{-1}, \eta_1, \dots, \eta_{pN}) \in \mathbb{R}^{2pN}$ . For every  $i \in \mathbb{I}_N$ , we denote by  $\Gamma_i = (\eta_{i'})_{i' \in \mathbb{J}_i^p} \in \mathbb{R}^p$  and we group these vectors in the matrix

$$\Gamma = (\Gamma_{-N} | \dots | \Gamma_{-1} | \Gamma_1 | \dots | \Gamma_N) \in \mathcal{M}_{p, 2N}(\mathbb{R}).$$

In view of (2.15), it follows that

$$\eta^T A_N^p(q) \eta = p \sum_{i \in \mathbb{I}_N} d_i^N \Gamma_i^T \Gamma_i - \sum_{i, j \in \mathbb{I}_N} w_{i, j}^N \Gamma_i^T J_p \Gamma_j.$$

Since  $J_p$  is a real symmetric matrix, there exists an orthonormal basis  $(b_k)_{1 \leq k \leq p}$  of  $\mathbb{R}^p$  diagonalizing  $J_p$ . Let us denote  $b_1 = \frac{1}{\sqrt{p}} \mathbf{1}_p$ . Then, there exists a diagonal matrix  $D \in \mathcal{M}_p(\mathbb{R})$  and a unitary matrix  $Q \in \mathcal{M}_p(\mathbb{R})$  such that  $J_p = Q D Q^T$ . These matrices have the following form:

$$D = \begin{pmatrix} p & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}_{p \times p} \quad \text{and} \quad Q = (b_1 | \dots | b_p)_{p \times p}.$$

We also define the matrix  $U = Q^T \Gamma \in \mathcal{M}_{p, 2N}(\mathbb{R})$ , and denote respectively

$$U_{k, \cdot} = (b_k^T \Gamma_i)_{i \in \mathbb{I}_N} \in \mathbb{R}^{2N} \quad \text{and} \quad U_{\cdot, i} = (b_k^T \Gamma_i)_{1 \leq k \leq p} \in \mathbb{R}^p$$

the rows and the columns of  $U$ . Then, for  $i, j \in \mathbb{I}_N$ , we have

$$\Gamma_i^T \Gamma_i = U_{\cdot, i}^T U_{\cdot, i} = \sum_{k=1}^p U_{k, i}^2 \quad \text{and} \quad \Gamma_i^T J_p \Gamma_j = U_{\cdot, i}^T D U_{\cdot, j} = p U_{1, i} U_{1, j}.$$

The spectrum of the matrix  $\frac{1}{p} A_N^p(q)$  can now be computed from

$$\begin{aligned} \frac{1}{p} \eta^T A_N^p(q) \eta &= \sum_{i \in \mathbb{I}_N} d_i^N U_{1, i}^2 - \sum_{i, j \in \mathbb{I}_N} w_{i, j}^N U_{1, i} U_{1, j} + \sum_{k=2}^p \sum_{i \in \mathbb{I}_N} d_i^N U_{k, i}^2 \\ &= U_{1, \cdot}^T A_N(q) U_{1, \cdot} + \sum_{k=2}^p U_{k, \cdot}^T \text{Diag}(A_N(q)) U_{k, \cdot}. \end{aligned}$$

Indeed, the expression above shows that  $A_N^p(q)$  is unitarily similar to the block diagonal matrix

$$\begin{pmatrix} A_N(q) & & & \\ & \text{Diag}(A_N(q)) & & \\ & & \ddots & \\ & & & \text{Diag}(A_N(q)) \end{pmatrix}_{2pN \times 2pN}.$$

□

**Step 5** – Uniform observability for the solution  $\varphi^N$  associated with  $(\varphi_0^N, \varphi_1^N)$  w.r.t.  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$ .

Let  $N$  be the smallest integer strictly greater than  $\varepsilon^{-1}$ . We prove an observability inequality for the function  $\varphi^N$  given by (2.10) and associated with the initial datum  $(\varphi_0^N, \varphi_1^N)$  defined in (2.2)–(2.3).

In view of Lemma 2.4, the function  $(\varphi_t^N)^2$  is constant on each elementary square  $C_{(i,j)}^N$  in  $\mathcal{C}_N$ :

$$(\varphi_t^N)^2|_{C_{(i,j)}^N} = \frac{1}{4}(\gamma_{j_N(i)}^N - \gamma_{-j_N(j)}^N)^2,$$

where  $\gamma_{j_N(i)}^N$  is given by (2.7) and  $j_N$  by (2.5). In view of the definition of the set  $R_N(q)$  (see (2.11)), we can estimate from below the  $L^2(q)$ -norm of  $\varphi_t^N$  as follows:

$$\begin{aligned} \iint_q (\varphi_t^N)^2 &\geq \iint_{R_N(q)} (\varphi_t^N)^2 = \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} \iint_{C_{(i,j)}^N} (\varphi_t^N)^2 \\ &= \frac{1}{8N^2} \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} (\gamma_{j_N(i)}^N - \gamma_{-j_N(j)}^N)^2. \end{aligned} \quad (2.17)$$

In the previous equality, we used that the area of every elementary square in  $\mathcal{C}_N$  is equal to  $\frac{1}{2N^2}$ . Combining (2.17) with the relation (2.13), we obtain

$$\iint_q (\varphi_t^N)^2 \geq \frac{1}{8N^2} (\gamma^N)^T A_N(q) \gamma^N, \quad (2.18)$$

with  $\gamma^N = (\gamma_{-N}^N, \dots, \gamma_{-1}^N, \gamma_1^N, \dots, \gamma_N^N) \in \mathbb{R}^{2N}$  and  $\gamma_i^N$  given by (2.7).

It is easy to see that  $\gamma^N \in \ker(A_N(q))^\perp$ . Indeed, applying Lemma 2.13, the graph  $G_N(q)$  is connected. Then, from Remark 2.14, we have  $\ker(A_N(q)) = \text{Vect}(\mathbf{1}_{2N})$  and, since  $\alpha_i^N = \alpha_{-i}^N$  and  $\beta_i^N = -\beta_{-i}^N$ , the vector  $\gamma^N$  verifies

$$(\gamma^N)^T \mathbf{1}_{2N} = \sum_{i \in \mathbb{I}_N} \gamma_i^N = 2 \sum_{i=1}^N \alpha_i^N = 0.$$

Then, using  $\lambda_N$  defined in (2.14), it follows that

$$\begin{aligned} (\gamma^N)^T A_N(q) \gamma^N &\geq \lambda_N \sum_{i \in \mathbb{I}_N} (\gamma_i^N)^2 = 2\lambda_N \sum_{i=1}^N \left( (\alpha_i^N)^2 + (\beta_i^N)^2 \right) \\ &= 2N\lambda_N \left( \|\varphi_0^N\|_{H_0^1(\Omega)}^2 + \|\varphi_1^N\|_{L^2(\Omega)}^2 \right). \end{aligned} \quad (2.19)$$

From (2.18) and (2.19), we deduce the following observability inequality:

$$\|(\varphi_0^N, \varphi_1^N)\|_{\mathbf{V}}^2 \leq \frac{4N}{\lambda_N} \|\varphi_t^N\|_{L^2(q)}^2,$$

where the constant  $\frac{4N}{\lambda_N}$  is independent of the domain  $q$  and the initial datum  $(\varphi_0, \varphi_1)$ .

**Step 6** – Uniform observability for the solution  $\varphi^{pN}$  associated with  $(\varphi_0^{pN}, \varphi_1^{pN})$  w.r.t.  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  and  $p \in \mathbb{N}^*$ .

The next step consists in obtaining a uniform observability inequality for an initial datum  $(\varphi_0^{pN}, \varphi_1^{pN})$  of the form (2.2)–(2.3) with respect to  $q \in \mathcal{Q}_{\text{ad}}^\varepsilon$  and  $p \in \mathbb{N}^*$ . As in the previous step, we get that  $(\varphi_t^{pN})^2$  is constant

on every elementary square  $C_{(i',j')}^{pN} \in \mathcal{C}_{pN}$ :

$$(\varphi_t^{pN})^2|_{C_{(i',j')}^{pN}} = \frac{1}{4}(\gamma_{j_{pN}(i')}^{pN} - \gamma_{-j_{pN}(j')}^{pN})^2,$$

where  $\varphi^{pN}$  is the solution of (1.3) associated with the initial datum  $(\varphi_0^{pN}, \varphi_1^{pN})$  given by (2.2)–(2.3), and  $\gamma_{j_{pN}(i')}^{pN}$  is defined by (2.7).

Then, remark that every elementary square  $C_{(i,j)}^N \in \mathcal{C}_N$  is the union of  $p^2$  elementary squares in  $\mathcal{C}_{pN}$ , or more precisely that

$$C_{(i,j)}^N = \bigcup_{i' \in \mathbb{J}_i^p} \bigcup_{j' \in \mathbb{J}_j^p} C_{(i',j')}^{pN}, \quad \forall i, j \in \mathbb{Z}^*.$$

Using the above expression in the evaluation of the  $L^2(q)$ -norm of  $\varphi_t^{pN}$ , we have

$$\begin{aligned} \iint_q (\varphi_t^{pN})^2 &\geq \iint_{R_N(q)} (\varphi_t^{pN})^2 = \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} \iint_{C_{(i,j)}^N} (\varphi_t^{pN})^2 \\ &= \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} \sum_{i' \in \mathbb{J}_i^p} \sum_{j' \in \mathbb{J}_j^p} \iint_{C_{(i',j')}^{pN}} (\varphi_t^{pN})^2 \\ &= \frac{1}{8p^2N^2} \sum_{C_{(i,j)}^N \in \mathcal{C}_N(q)} \sum_{i' \in \mathbb{J}_i^p} \sum_{j' \in \mathbb{J}_j^p} (\gamma_{j_{pN}(i')}^{pN} - \gamma_{-j_{pN}(j')}^{pN})^2 \\ &= \frac{1}{8p^2N^2} (\gamma^{pN})^T A_N^p(q) \gamma^{pN}. \end{aligned} \tag{2.20}$$

Since  $G_N(q)$  is a connected graph, the degree  $d_i^N$  of every vertex  $i \in \mathbb{I}_N$  verifies  $d_i^N \geq 1$ . Applying Lemma 2.17, the smallest non-zero eigenvalue  $\lambda_N^p(q)$  of  $\frac{1}{p}A_N^p(q)$  verifies

$$\lambda_N^p(q) = \min(\lambda_N(q), \min_{i \in \mathbb{I}_N} d_i^N) \geq \min(\lambda_N, 1) > 0.$$

The vector  $\gamma^{pN} = (\gamma_{i'}^{pN})_{i' \in \mathbb{I}_{pN}}$  belongs to  $\ker(A_N^p(q))^\perp$ . Indeed,

$$(\gamma^{pN})^T \mathbf{1}_{2pN} = \sum_{i' \in \mathbb{I}_{pN}} \gamma_{i'}^{pN} = 2 \sum_{i'=1}^{pN} \alpha_{i'}^{pN} = 0.$$

Setting  $\widehat{\lambda}_N = \min(\lambda_N, 1)$ , it follows that

$$\begin{aligned} \frac{1}{p} (\gamma^{pN})^T A_N^p(q) \gamma^{pN} &\geq \widehat{\lambda}_N \sum_{i' \in \mathbb{I}_{pN}} (\gamma_{i'}^{pN})^2 = 2\widehat{\lambda}_N \sum_{i'=1}^{pN} \left( (\alpha_{i'}^{pN})^2 + (\beta_{i'}^{pN})^2 \right) \\ &= 2pN\widehat{\lambda}_N \left( \|\varphi_0^{pN}\|_{H_0^1(\Omega)}^2 + \|\varphi_1^{pN}\|_{L^2(\Omega)}^2 \right). \end{aligned}$$



Consequently, combining the above relation with (2.20), we obtain the following observability inequality:

$$\|(\varphi_0^{pN}, \varphi_1^{pN})\|_{\mathbf{V}}^2 \leq \frac{4N}{\lambda_N} \|\varphi_t^{pN}\|_{L^2(q)}^2, \quad (2.21)$$

where the constant  $\frac{4N}{\lambda_N}$  is independent of the domain  $q$ , the initial datum  $(\varphi_0, \varphi_1)$  and the integer  $p$ .

**Step 7** – Passing to the limit  $p \rightarrow \infty$  and conclusion.

In order to finish the proof, we pass to the limit when  $p \rightarrow \infty$  in the observability inequality (2.21). We easily check that as  $p$  tends to  $\infty$ ,

$$\varphi_0^{pN} \rightarrow \varphi_0 \quad \text{in } H_0^1(\Omega) \quad \text{and} \quad \varphi_1^{pN} \rightarrow \varphi_1 \quad \text{in } L^2(\Omega).$$

Moreover, since the solution  $\varphi$  of the wave equation (1.3) depends continuously on its initial condition  $(\varphi_0, \varphi_1) \in \mathbf{V}$ , we can write

$$\varphi_t^{pN} \rightarrow \varphi_t \quad \text{in } L^2(0, T; L^2(\Omega)).$$

Eventually, passing to the limit in (2.21), we get

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{V}}^2 \leq \max \left\{ 4N, \frac{4N}{\lambda_N} \right\} \|\varphi_t\|_{L^2(q)}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{V},$$

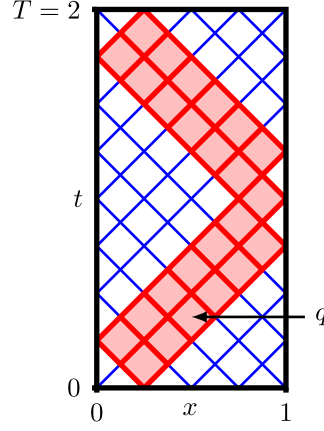
which concludes the proof with  $C_{\text{obs}}^\varepsilon = \max \left\{ 4N, \frac{4N}{\lambda_N} \right\}$ . We recall that  $N$  depends on  $\varepsilon$  by the condition  $N > 1/\varepsilon$ .

**Remark 2.18.** Let  $q \subset Q_T$  be a finite union of open sets. If  $q$  verifies the usual geometric optic condition, there exists  $\varepsilon > 0$  small enough such that  $q^\varepsilon$  still verifies the geometric optic condition. We then set  $N = \lfloor 1/\varepsilon \rfloor + 1$ . The associated graph  $G_N(q)$  being connected, there exists a relation (see, for instance, [22]) between the algebraic connectivity  $\lambda_N(q)$ , the number of vertices  $N_V$  and the diameter  $D_G$  of the graph. More precisely, we have  $\lambda_N(q) \geq \frac{4}{N_V D_G}$ . Since in our case  $N_V = 2N$  and  $D_G \leq 2N$ , we deduce that  $\lambda_N(q) \geq \frac{1}{N^2}$  and therefore that  $C_{\text{obs}}(q) \leq 4N^3$ . Thus, in the worst situation, we could have an observability constant of order  $1/\varepsilon^3$ . If we see  $\varepsilon$  as a measure of the “thickness” of the observation domain  $q$ , we get the same estimate of the observability constant as the one given in Proposition 2.1 of [29].

## 2.2. One explicit example

We illustrate the proof of Theorem 2.1 on a simple example for which the observation domain  $q$  depicted in Figure 3 (colored in red) is well adapted to the subdivision  $S_4$ . The study of this example is also the opportunity to develop a method for the computation of the observability constant for observation domains which are exactly the union of elementary squares associated with a given subdivision  $S_N$ , for a fixed integer  $N > 0$ .

We start by enumerating the elementary squares composing the observation domain  $q$ . In Table 1, we list, for  $i, j \in \mathbb{Z}^*$ , the elementary squares  $C_{(i,j)}^4$  included in  $q$  and the values of the indices  $j_4(i)$  and  $-j_4(j)$ , allowing to compute the Laplacian matrix  $A_4(q)$  associated with the corresponding graph  $G_4(q)$ .

FIGURE 3. Observation domain  $q$  adapted to  $S_4$ .TABLE 1. Elementary squares associated with  $S_4$  and belonging to  $\mathcal{C}_4(q)$ .

$C_{(i,j)}^4$	$j_4(i)$	$-j_4(j)$	$C_{(i,j)}^4$	$j_4(i)$	$-j_4(j)$	$C_{(i,j)}^4$	$j_4(i)$	$-j_4(j)$
$C_{(2,1)}^4$	2	-1	$C_{(6,-1)}^4$	-3	1	$C_{(9,-4)}^4$	1	4
$C_{(2,-1)}^4$	2	1	$C_{(7,1)}^4$	-2	-1	$C_{(8,-5)}^4$	-1	-4
$C_{(3,1)}^4$	3	-1	$C_{(7,-1)}^4$	-2	1	$C_{(9,-5)}^4$	1	-4
$C_{(3,-1)}^4$	3	1	$C_{(8,-1)}^4$	-1	1	$C_{(8,-6)}^4$	-1	-3
$C_{(4,1)}^4$	4	-1	$C_{(8,-2)}^4$	-1	2	$C_{(9,-6)}^4$	1	-3
$C_{(4,-1)}^4$	4	1	$C_{(9,-2)}^4$	1	2	$C_{(8,-7)}^4$	-1	-2
$C_{(5,1)}^4$	-4	-1	$C_{(8,-3)}^4$	-1	3	$C_{(9,-7)}^4$	1	-2
$C_{(5,-1)}^4$	-4	1	$C_{(9,-3)}^4$	1	3			
$C_{(6,1)}^4$	-3	-1	$C_{(8,-4)}^4$	-1	4			

The Laplacian matrix associated with the graph  $G_4(q)$  is given by

$$A_4(q) = \begin{pmatrix} 4 & 0 & 0 & -2 & -2 & 0 & 0 & 0 \\ 0 & 4 & 0 & -2 & -2 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 & -2 & 0 & 0 & 0 \\ -2 & -2 & -2 & 13 & -1 & -2 & -2 & -2 \\ -2 & -2 & -2 & -1 & 13 & -2 & -2 & -2 \\ 0 & 0 & 0 & -2 & -2 & 4 & 0 & 0 \\ 0 & 0 & 0 & -2 & -2 & 0 & 4 & 0 \\ 0 & 0 & 0 & -2 & -2 & 0 & 0 & 4 \end{pmatrix}_{8 \times 8}.$$

The spectrum of  $A_4(q)$  can be explicitly computed:

$$\text{Sp}(A_4(q)) = \{0, 4, 4, 4, 4, 4, 14, 16\}.$$

It confirms that the kernel of  $A_4(q)$  is one-dimensional – therefore  $G_4(q)$  is connected – and implies that the smallest non-zero eigenvalue of  $A_4(q)$  is  $\lambda_4(q) = 4$ . If we replace the subdivision  $S_4$  by the subdivision  $S_{4p}$  for

any  $p \in \mathbb{N}^*$ , then the Laplacian matrix associated with the graph  $G_4^p(q)$  is the following one:

$$A_4^p(q) = \begin{pmatrix} 4pI_p & 0_p & 0_p & -2J_p & -2J_p & 0_p & 0_p & 0_p \\ 0_p & 4pI_p & 0_p & -2J_p & -2J_p & 0_p & 0_p & 0_p \\ 0_p & 0_p & 4pI_p & -2J_p & -2J_p & 0_p & 0_p & 0_p \\ -2J_p & -2J_p & -2J_p & 13pI_p & -J_p & -2J_p & -2J_p & -2J_p \\ -2J_p & -2J_p & -2J_p & -J_p & 13pI_p & -2J_p & -2J_p & -2J_p \\ 0_p & 0_p & 0_p & -2J_p & -2J_p & 4pI_p & 0_p & 0_p \\ 0_p & 0_p & 0_p & -2J_p & -2J_p & 0_p & 4pI_p & 0_p \\ 0_p & 0_p & 0_p & -2J_p & -2J_p & 0_p & 0_p & 4pI_p \end{pmatrix}_{8p \times 8p}.$$

According to Lemma 2.17, the smallest non-zero eigenvalue of  $\frac{1}{p}A_4^p(q)$  is given by

$$\lambda_4^p(q) = \min(\lambda_4(q), \min_{i \in \mathbb{I}_4} d_i^4) = \min(4, 4) = 4.$$

Consequently, the observability constant associated with the observation domain  $q$  depicted in Figure 3 is given by

$$C_{\text{obs}}(q) = \frac{4 \cdot 4}{\lambda_4^p(q)} = 4.$$

### 2.3. A corollary

We prove a uniform observability inequality for the observation domains  $q_\gamma$  defined in (1.9), with  $\gamma \in \mathcal{G}_{\text{ad}}$  (see (1.10)), which will be used in the next section.

**Corollary 2.19.** *Let  $T \geq 2$ . There exists a constant  $C_{\text{obs}} > 0$  such that for every  $\gamma \in \mathcal{G}_{\text{ad}}$ ,*

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{W}}^2 \leq C_{\text{obs}} \|\varphi\|_{L^2(q_\gamma)}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W},$$

where  $\varphi$  is the solution of the wave equation (1.3) associated with the initial condition  $(\varphi_0, \varphi_1)$ .

*Proof.* We show that for any  $\varepsilon > 0$  small enough,  $\{q_\gamma; \gamma \in \mathcal{G}_{\text{ad}}\} \subset \mathcal{Q}_{\text{ad}}^\varepsilon$ . Let  $\gamma \in \mathcal{G}_{\text{ad}}$ . We introduce the sets  $\Gamma_\pm = \{(\gamma(t) \pm \delta_0, t); t \in [0, T]\}$ ,  $\tilde{\Gamma}_\pm = \{(\gamma(t) \pm \frac{\delta_0}{2}, t); t \in [0, T]\}$  and  $Q_T^\varepsilon = \Omega \times (\varepsilon, T - \varepsilon)$ .  $\gamma$  being a  $M$ -Lipschitz curve, we can show that

$$d(\tilde{\Gamma}_\pm, \Gamma_\pm) \geq \frac{\delta_0}{2\sqrt{M^2 + 1}}.$$

Then, for  $\varepsilon < \frac{\delta_0}{2\sqrt{M^2 + 1}}$ , we have  $\tilde{q}_\gamma \cap Q_T^\varepsilon \subset q_\gamma^\varepsilon$ , with the observation domain  $\tilde{q}_\gamma$  defined as in (1.9) with a half-width of  $\delta_0/2$ . The domain  $\tilde{q}_\gamma \cap Q_T^\varepsilon$  verifies the geometric optic condition because  $\delta_0 \leq \gamma \leq 1 - \delta_0$  and  $T - 2\varepsilon \geq 2(1 - \delta_0)$ . Consequently,  $q_\gamma^\varepsilon$  also verifies the geometric optic condition and  $q_\gamma \in \mathcal{Q}_{\text{ad}}^\varepsilon$ . We conclude the proof by noticing that the constant  $\frac{\delta_0}{2\sqrt{M^2 + 1}}$  is independent of the choice of  $\gamma$ .

Finally, the observability inequality follows from Theorem 2.1 with  $C_{\text{obs}} = C_{\text{obs}}^\varepsilon$ , for any  $\varepsilon < \frac{\delta_0}{2\sqrt{M^2 + 1}}$ . We conclude the proof by noticing that the constant  $\frac{\delta_0}{2\sqrt{M^2 + 1}}$  is independent of the choice of  $\gamma$ .  $\square$

We end this section by emphasizing the role of the bound  $M$  on the derivative of the curve  $\gamma$  introduced in the set  $\mathcal{G}_{\text{ad}}$ , namely  $\|\gamma'\|_{L^\infty(0, T)} \leq M$ . As can be seen in the previous proof, if  $M$  goes to infinity, then  $\varepsilon$  goes to zero, and consequently, in view of Remark 2.18, the constant  $C_{\text{obs}}^\varepsilon$  blows up. This is illustrated in Figure 4 where a ray starting from the point  $(\frac{1}{2}, 0)$  only intersects the control region on a narrow part corresponding to

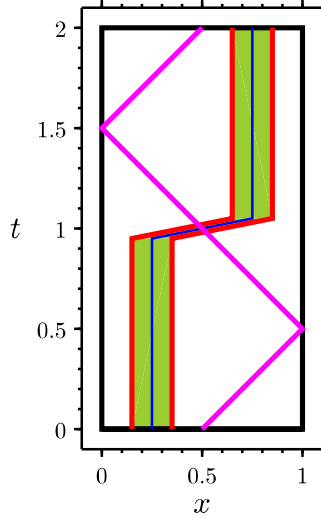


FIGURE 4. Ray intersecting a narrow part (in the neighborhood of the point  $(1/2, 1)$ ) of the control region corresponding to a large slope of  $\gamma$ .

a large slope of  $\gamma$ . Any initial condition generating a solution localized around such ray leads to an arbitrarily large cost of control. The constraint on the slope of  $\gamma$  has been introduced to avoid this phenomenon.

### 3. OPTIMIZATION OF THE SHAPE OF THE CONTROL DOMAIN

In this section, we study the problem of finding the optimal shape and position of the control domain, for a given initial condition  $(y_0, y_1) \in \mathbf{V}$ .

#### 3.1. Existence of an optimal domain

In order to obtain a well-posedness result, we consider a variant of the optimal problem (1.11) and replace the characteristic function  $\mathbb{1}_q$  in (1.1) by a regular function in space. More precisely, we fix  $\delta \in (0, \delta_0)$  and, for every  $\gamma \in \mathcal{G}_{\text{ad}}$ , we define  $\chi_\gamma(x, t) = \chi(x - \gamma(t))$ , with  $\chi : \mathbb{R} \rightarrow [0, 1]$  a  $C^1$  even function such that

$$\chi(x) = \begin{cases} 1 & \text{if } x \in (-\delta_0 + \delta, \delta_0 - \delta), \\ 0 & \text{if } x \notin (-\delta_0, \delta_0), \\ \in (0, 1) & \text{otherwise.} \end{cases} \quad (3.1)$$

In the sequel, we also use the function  $\chi'_\gamma$  defined by  $\chi'_\gamma(x, t) = \chi'(x - \gamma(t))$ . In this new setting, the HUM control now lives in the weighted space

$$L^2_\chi(q_\gamma) := L^2(q_\gamma; \chi_\gamma) = \left\{ v : q_\gamma \rightarrow \mathbb{R}; \quad \iint_{q_\gamma} v^2 \chi_\gamma < +\infty \right\}.$$

Moreover, we can adapt the uniform observability inequality given in Corollary 2.19. For  $T \geq 2$ , there exists a constant  $C_{\text{obs}} > 0$  such that for every  $\gamma \in \mathcal{G}_{\text{ad}}$ ,

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{W}}^2 \leq C_{\text{obs}} \|\varphi\|_{L^2_\chi(q_\gamma)}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W}, \quad (3.2)$$

where  $\varphi$  is the solution of (1.3) associated with the initial condition  $(\varphi_0, \varphi_1)$ .

Then, our optimization problem reads as follows: for a given initial datum  $(y_0, y_1) \in \mathbf{V}$ , solve

$$\inf_{\gamma \in \mathcal{G}_{\text{ad}}} J(\gamma) = \|v\|_{L^2_\chi(q_\gamma)}^2 = \iint_{q_\gamma} \varphi^2 \chi_\gamma, \quad (3.3)$$

where  $v$  is the control of minimal  $L^2_\chi$ -norm distributed over  $q_\gamma \subset Q_T$ , and  $\varphi$  is the associated adjoint state such that  $v = \varphi|_{q_\gamma}$ . This adjoint state can be characterized using the HUM method, it is the solution of (1.3) associated with the minimum  $(\varphi_0, \varphi_1)$  of the conjugate functional

$$\mathcal{J}_\gamma^*(\varphi_0, \varphi_1) = \frac{1}{2} \iint_{q_\gamma} \varphi^2 \chi_\gamma - \langle \varphi_1, y_0 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \langle \varphi_0, y_1 \rangle_{L^2(\Omega)}, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W}. \quad (3.4)$$

To show the well-posedness of (3.3), we follow the steps of Theorem 2.1 in [29]. We start with a convergence result on the function  $\chi_\gamma$ .

**Lemma 3.1.** *Let  $(\gamma_n)_{n \geq 0} \subset \mathcal{G}_{\text{ad}}$  and  $\gamma \in \mathcal{G}_{\text{ad}}$ . If  $\gamma_n \rightarrow \gamma$  in  $L^\infty(0, T)$ , then  $\chi_{\gamma_n} \rightarrow \chi_\gamma$  in  $L^\infty(Q_T)$ .*

*Proof.* It is a direct consequence of the inequality  $\|\chi_{\gamma_n} - \chi_\gamma\|_{L^\infty(Q_T)} \leq \|\chi'\|_{L^\infty(\mathbb{R})} \|\gamma_n - \gamma\|_{L^\infty(0, T)}$ .  $\square$

We then have that the following continuity result.

**Proposition 3.2.** *The cost  $J$  is continuous over  $\mathcal{G}_{\text{ad}}$  for the norm  $L^\infty(0, T)$ .*

*Proof.* Let  $(\gamma_n)_{n \geq 0} \subset \mathcal{G}_{\text{ad}}$  and  $\gamma \in \mathcal{G}_{\text{ad}}$  such that  $\gamma_n \rightarrow \gamma$  in  $L^\infty(0, T)$  as  $n \rightarrow \infty$ .

For any  $n \in \mathbb{N}$ , we denote  $(\varphi_0^n, \varphi_1^n) \in \mathbf{W}$  the minimum of  $\mathcal{J}_{\gamma_n}^*$ , and  $\varphi^n$  the corresponding solution of (1.3). Using the uniform observability inequality (3.2) and the optimality condition of  $\mathcal{J}_{\gamma_n}^*$ , it follows that

$$\begin{aligned} \|(\varphi_0^n, \varphi_1^n)\|_{\mathbf{W}}^2 &\leq C_{\text{obs}} \iint_{q_{\gamma_n}} (\varphi^n)^2 \chi_{\gamma_n} = C_{\text{obs}} \left( \langle \varphi_1^n, y_0 \rangle_{H^{-1}, H_0^1} - \langle \varphi_0^n, y_1 \rangle_{L^2} \right) \\ &\leq C_{\text{obs}} \|(\varphi_0^n, \varphi_1^n)\|_{\mathbf{W}} \|(y_0, y_1)\|_{\mathbf{V}} \end{aligned}$$

leading to the uniform bound  $\|(\varphi_0^n, \varphi_1^n)\|_{\mathbf{W}} \leq C_{\text{obs}} \|(y_0, y_1)\|_{\mathbf{V}}$ . Consequently, there exist two functions  $\varphi_0 \in L^2(\Omega)$  and  $\varphi_1 \in H^{-1}(\Omega)$  such that, up to a subsequence, as  $n \rightarrow \infty$

$$\varphi_0^n \rightharpoonup \varphi_0 \text{ weakly in } L^2(\Omega) \quad \text{and} \quad \varphi_1^n \rightharpoonup \varphi_1 \text{ weakly in } H^{-1}(\Omega).$$

From the continuous dependence of the solution of the wave equation with respect to the initial condition, it follows

$$\varphi^n \rightharpoonup \varphi \text{ weakly in } L^2(0, T; L^2(\Omega)),$$

where  $\varphi$  is the solution of (1.3) associated with  $(\varphi_0, \varphi_1)$ .

Let  $\psi \in L^2(0, T; L^2(\Omega))$ . We then have

$$\iint_{Q_T} \psi \varphi^n \chi_{\gamma_n} = \iint_{Q_T} \psi \varphi^n \chi_\gamma + \iint_{Q_T} \psi \varphi^n (\chi_{\gamma_n} - \chi_\gamma) \rightarrow \iint_{Q_T} \psi \varphi \chi_\gamma.$$

Indeed, we can take the weak limit in the first term because  $\psi \chi_\gamma \in L^2(0, T; L^2(\Omega))$ . Using Lemma 3.1 and the boundedness of  $(\varphi^n)_{n \geq 0}$  in  $L^2(0, T; L^2(\Omega))$ , the second term converges to 0 in view of the estimate

$$\left| \iint_{Q_T} \psi \varphi^n (\chi_{\gamma_n} - \chi_\gamma) \right| \leq \|\psi\|_{L^2(L^2)} \|\varphi^n\|_{L^2(L^2)} \|\chi_{\gamma_n} - \chi_\gamma\|_{L^\infty(Q_T)}.$$

Consequently, we obtain the convergence

$$\varphi^n \chi_{\gamma_n} \rightharpoonup \varphi \chi_\gamma \quad \text{weakly in } L^2(0, T; L^2(\Omega)).$$

Let now  $(\psi_0, \psi_1) \in \mathbf{W}$  and  $\psi$  the corresponding solution of (1.3). Taking the weak limit in the optimality condition

$$\iint_{q_{\gamma_n}} \psi \varphi^n \chi_{\gamma_n} = \langle \psi_1, y_0 \rangle_{H^{-1}, H_0^1} - \int_{\Omega} \psi_0 y_1,$$

we find

$$\iint_{q_\gamma} \psi \varphi \chi_\gamma = \langle \psi_1, y_0 \rangle_{H^{-1}, H_0^1} - \int_{\Omega} \psi_0 y_1.$$

This means that  $(\varphi_0, \varphi_1)$  is the minimum of  $\mathcal{J}_\gamma^*$ . Besides, we remark that this property uniquely characterizes the weak limit of any subsequence of  $(\varphi_0^n, \varphi_1^n)$ . This implies that the whole sequence  $(\varphi_0^n, \varphi_1^n)$  weakly converges. The continuity of  $J$  is finally obtained by taking the weak limit in the optimality condition

$$\iint_{q_{\gamma_n}} (\varphi^n)^2 \chi_{\gamma_n} = \langle \varphi_1^n, y_0 \rangle_{H^{-1}, H_0^1} - \int_{\Omega} \varphi_0^n y_1 \rightarrow \langle \varphi_1, y_0 \rangle_{H^{-1}, H_0^1} - \int_{\Omega} \varphi_0 y_1 = \iint_{q_\gamma} \varphi^2 \chi_\gamma.$$

□

The continuity of  $J$  then allows to show that the extremal problem (3.3) is well-posed.

**Proposition 3.3.** *The cost  $J$  reaches its minimum over  $\mathcal{G}_{\text{ad}}$ .*

*Proof.* The cost  $J$  being bounded from below, there exists a minimizing sequence  $(\gamma_n)_{n \geq 0} \subset \mathcal{G}_{\text{ad}}$ . By definition of  $\mathcal{G}_{\text{ad}}$ , this sequence is bounded in  $W^{1,\infty}(0, T)$ . Moreover, since  $W^{1,\infty}(0, T)$  is compactly embedded in  $L^\infty(0, T)$ , there exists a curve  $\gamma \in L^\infty(0, T)$  such that, up to a subsequence,  $\gamma_n \rightarrow \gamma$  in  $L^\infty(0, T)$ . From the definition of  $\mathcal{G}_{\text{ad}}$ , all the curves  $\gamma_n$  are  $M$ -Lipschitz, with  $M$  independent of  $n$ . So, taking the pointwise limit in the expressions

$$\begin{aligned} |\gamma_n(t) - \gamma_n(s)| &\leq M|t - s|, \quad \forall t, s \in [0, T], \\ \delta_0 &\leq \gamma_n(t) \leq 1 - \delta_0, \quad \forall t \in [0, T], \end{aligned}$$

we notice that  $\gamma \in \mathcal{G}_{\text{ad}}$ . Finally, using Proposition 3.2, we obtain  $J(\gamma_n) \rightarrow J(\gamma) = \inf_{\mathcal{G}_{\text{ad}}} J$  which means that  $\gamma$  is a minimum of  $J$  over  $\mathcal{G}_{\text{ad}}$ . □

It follows in particular from Proposition 3.2 that the optimal cost  $J(\gamma^*) = \inf_{\gamma \in \mathcal{G}_{\text{ad}}} J(\gamma)$  is stable with respect to the initial datum  $(y_0, y_1) \in \mathbf{V}$  in the following sense:  $J(\gamma^*) \leq C_{\text{obs}} \|(y_0, y_1)\|_{\mathbf{V}}^2$ , where  $C_{\text{obs}}$  is the observability constant appearing in Corollary 2.19. However, we highlight that the optimal curve  $\gamma^*$  does not vary continuously with respect to  $(y_0, y_1)$  (see Sect. 4.1).

**Remark 3.4.** We emphasize that the functional  $J$  may have several global minima. Indeed, if the initial datum  $(y_0, y_1) \in \mathbf{V}$  is such that  $y_0(x) = -y_0(1-x)$ ,  $y_1(x) = 0$  for  $x \in \Omega$ , there are examples (see [29] and example (EX1) in Sect. 4.1) where the optimal domain is not the cylinder centered at  $x = \frac{1}{2}$ . Hence, if  $\gamma^*$  is an optimal curve for  $J$  corresponding to this initial datum, from symmetry arguments, the curve  $1 - \gamma^*$  is also optimal.

### 3.2. First directional derivative of the cost $J$

We now give the expression of the directional derivative of the cost  $J$ .

**Definition 3.5.** Let  $\gamma, \bar{\gamma} \in W^{1,\infty}(0, T)$ , with  $\delta_0 \leq \gamma \leq 1 - \delta_0$ . The perturbation  $\bar{\gamma}$  is said admissible if and only if for any  $\eta > 0$  small enough, the perturbed curve  $\gamma_\eta = \gamma + \eta\bar{\gamma}$  verifies  $\delta_0 \leq \gamma_\eta \leq 1 - \delta_0$ .

**Lemma 3.6.** Let  $\chi \in C^2(\mathbb{R})$  and  $\gamma, \bar{\gamma} \in W^{1,\infty}(0, T)$ , with  $\delta_0 \leq \gamma \leq 1 - \delta_0$ . For any  $\eta > 0$ , we define the perturbed curve  $\gamma_\eta = \gamma + \eta\bar{\gamma}$ . Taking  $\eta \rightarrow 0$ , we then have

$$\frac{\chi_{\gamma_\eta} - \chi_\gamma}{\eta} \rightarrow -\bar{\gamma}\chi'_\gamma \quad \text{in } L^\infty(Q_T).$$

*Proof.* It is a direct consequence the inequality

$$\|\chi_{\gamma_\eta} - \chi_\gamma + \eta\bar{\gamma}\chi'_\gamma\|_{L^\infty(Q_T)} \leq \frac{\eta^2}{2} \|\chi''\|_{L^\infty(\mathbb{R})} \|\bar{\gamma}\|_{L^\infty(0, T)}^2.$$

□

**Proposition 3.7.** Let  $\chi \in C^2(\mathbb{R})$  and  $\gamma, \bar{\gamma} \in W^{1,\infty}(0, T)$ , with  $\delta_0 \leq \gamma \leq 1 - \delta_0$ . For any  $\eta > 0$ , we define the perturbed curve  $\gamma_\eta = \gamma + \eta\bar{\gamma}$ . If  $\bar{\gamma}$  is an admissible perturbation, then the directional derivative of  $J$  at  $\gamma$  in the direction  $\bar{\gamma}$ , denoted by  $dJ(\gamma; \bar{\gamma})$ , reads

$$dJ(\gamma; \bar{\gamma}) = \lim_{\eta \rightarrow 0} \frac{J(\gamma_\eta) - J(\gamma)}{\eta} = \int_0^T \bar{\gamma} \int_\Omega \varphi^2 \chi'_\gamma,$$

where  $\varphi$  is the solution of (1.3) associated with the minimum  $(\varphi_0, \varphi_1)$  of  $\mathcal{J}_\gamma^*$ .

*Proof.* For  $\eta > 0$  small enough, we denote by  $(\varphi_0^\eta, \varphi_1^\eta)$  the minimum of  $\mathcal{J}_{\gamma_\eta}^*$ , and  $\varphi^\eta$  the corresponding solution of (1.3). Similarly, we denote by  $(\varphi_0, \varphi_1)$  the minimum of  $\mathcal{J}_\gamma^*$ , and  $\varphi$  the corresponding solution of (1.3). Using the optimality conditions of  $\mathcal{J}_{\gamma_\eta}^*$  and  $\mathcal{J}_\gamma^*$ , we can write

$$\begin{aligned} J(\gamma_\eta) - J(\gamma) &= \iint_{q_{\gamma_\eta}} (\varphi^\eta)^2 \chi_{\gamma_\eta} - \iint_{q_\gamma} \varphi^2 \chi_\gamma \\ &= \left( \langle \varphi_1^\eta, y_0 \rangle_{H^{-1}, H_0^1} - \int_\Omega \varphi_0^\eta y_1 \right) - \left( \langle \varphi_1, y_0 \rangle_{H^{-1}, H_0^1} - \int_\Omega \varphi_0 y_1 \right) \\ &= \iint_{q_\gamma} \varphi^\eta \varphi \chi_\gamma - \iint_{q_{\gamma_\eta}} \varphi \varphi^\eta \chi_{\gamma_\eta} = - \iint_{Q_T} \varphi^\eta \varphi (\chi_{\gamma_\eta} - \chi_\gamma). \end{aligned}$$

Arguing as in the proof of Proposition 3.2, we can show that  $\varphi^\eta \rightharpoonup \varphi$  weakly in  $L^2(0, T; L^2(\Omega))$ . As a result, we have

$$\begin{aligned} \frac{J(\gamma_\eta) - J(\gamma)}{\eta} &= \iint_{Q_T} \varphi^\eta \bar{\gamma} \chi'_\gamma - \iint_{Q_T} \varphi \bar{\gamma} \left( \frac{\chi_{\gamma_\eta} - \chi_\gamma}{\eta} + \bar{\gamma} \chi'_\gamma \right) \\ &\rightarrow \iint_{Q_T} \varphi^2 \bar{\gamma} \chi'_\gamma = \int_0^T \bar{\gamma} \int_\Omega \varphi^2 \chi'_\gamma. \end{aligned}$$

Indeed, we can take the weak limit in the first term since  $\varphi\bar{\gamma}\chi'_\gamma \in L^2(0, T; L^2(\Omega))$ . Moreover, using Lemma 3.6 and the boundedness of  $(\varphi^\eta)_{\eta>0}$  in  $L^2(0, T; L^2(\Omega))$ , the second term converges to 0 in view of the estimate

$$\left| \iint_{Q_T} \varphi^\eta \varphi \left( \frac{\chi_{\gamma_\eta} - \chi_\gamma}{\eta} + \bar{\gamma}\chi'_\gamma \right) \right| \leq \|\varphi^\eta\|_{L^2(L^2)} \|\varphi\|_{L^2(L^2)} \left\| \frac{\chi_{\gamma_\eta} - \chi_\gamma}{\eta} + \bar{\gamma}\chi'_\gamma \right\|_{L^\infty(Q_T)}.$$

□

**Remark 3.8.** We emphasize that the expression of the directional derivative  $dJ(\gamma; \bar{\gamma})$  does not involve the solution of an adjoint boundary value problem. This is due to the fact that the curve  $\gamma$ , argument of the cost  $J$ , is associated with the control of minimal  $L^2(q_\gamma)$ -norm. We refer to the proof of Theorem 2.3 in [23] for more details.

### 3.3. Regularization and Gradient algorithm

At the practical level, in order to solve the optimal problem (3.3) numerically, we need to handle the Lipschitz constraint included in  $\mathcal{G}_{\text{ad}}$ . In this respect, we add a regularizing term to the cost  $J$  in order to keep the derivative of  $\gamma$  uniformly bounded. The optimization problem is now the following one: for  $\epsilon > 0$  fixed, solve

$$\min_{\substack{\gamma \in W^{1,\infty}(0,T) \\ \delta_0 \leq \gamma \leq 1 - \delta_0}} J_\epsilon(\gamma) = J(\gamma) + \frac{\epsilon}{2} \|\gamma'\|_{L^2(0,T)}^2.$$

The regularization parameter  $\epsilon$ , which can be compared to the Lipschitz constant  $M$  in the definition of  $\mathcal{G}_{\text{ad}}$ , controls the speed of variation of the curves  $\gamma \in W^{1,\infty}(0, T)$ .

We fix  $\gamma \in W^{1,\infty}(0, T)$  such that  $\delta_0 \leq \gamma \leq 1 - \delta_0$ . Using Proposition 3.7, for any admissible perturbation  $\bar{\gamma} \in W^{1,\infty}(0, T)$ , a direct calculation provides the expression of the directional derivative of  $J_\epsilon$

$$dJ_\epsilon(\gamma; \bar{\gamma}) = dJ(\gamma; \bar{\gamma}) + \epsilon \int_0^T \gamma' \bar{\gamma}' = \langle j_\gamma, \bar{\gamma} \rangle_{L^2(0,T)} + \epsilon \langle \gamma', \bar{\gamma}' \rangle_{L^2(0,T)},$$

with

$$j_\gamma(t) = \int_\Omega \varphi^2(x, t) \chi'_\gamma(x, t) dx, \quad \forall t \in [0, T].$$

In the expression of  $j_\gamma$ , the function  $\varphi$  is the solution of (1.3) associated with the minimum  $(\varphi_0, \varphi_1)$  of  $\mathcal{J}_\gamma^*$ . Consequently, a minimizing sequence  $(\gamma_n)_{n \in \mathbb{N}}$  for  $J_\epsilon$  is defined as follows:

$$\begin{cases} \gamma_0 \text{ given in } H^1(0, T), \\ \gamma_{n+1} = P_{[\delta_0, 1-\delta_0]}(\gamma_n - \rho j_{\gamma_n}^\epsilon), \quad \text{for } n \geq 0, \end{cases}$$

where  $P_{[\delta_0, 1-\delta_0]}$  is the pointwise projection in the interval  $[\delta_0, 1 - \delta_0]$ ,  $\rho > 0$  a descent step-size and  $j_{\gamma_n}^\epsilon \in H^1(0, T)$  is the solution of the variational formulation

$$\langle j_{\gamma_n}^\epsilon, \tilde{\gamma} \rangle_{L^2(0,T)} + \epsilon \langle j_{\gamma_n}^{\epsilon'}, \tilde{\gamma}' \rangle_{L^2(0,T)} = \langle j_{\gamma_n}, \tilde{\gamma} \rangle_{L^2(0,T)} + \epsilon \langle \gamma_n', \tilde{\gamma}' \rangle_{L^2(0,T)}, \quad \forall \tilde{\gamma} \in H^1(0, T). \quad (3.5)$$

This implies that  $dJ_\epsilon(\gamma_n; j_{\gamma_n}^\epsilon) = \|j_{\gamma_n}^\epsilon\|_{L^2(0,T)}^2 + \epsilon \|j_{\gamma_n}^{\epsilon'}\|_{L^2(0,T)}^2 \geq 0$ .

**Remark 3.9.** For simplicity, we have relaxed here the constraint  $\|\gamma'\|_{L^\infty(0,T)} \leq M$  appearing in the set  $\mathcal{G}_{\text{ad}}$  (see (1.10)) through the regularization term  $\frac{\epsilon}{2} \|\gamma'\|_{L^2(0,T)}^2$ , with  $\epsilon > 0$ . We may instead consider the term  $\frac{\epsilon}{2} \|\gamma''\|_{L^2(0,T)}^2$



ensuring that  $\|\gamma'\|_{L^\infty(0,T)} \leq M_\epsilon$ , with  $M_\epsilon = \mathcal{O}(\epsilon^{-1})$ . Our numerical experiments – some of them discussed in the next section – suggest that this later possibility is not necessary.

#### 4. NUMERICAL EXPERIMENTS

Before presenting some numerical experiments, let us briefly mention some aspects of the resolution of the underlying discretized problem.

– The discretization of the curve  $\gamma$  is performed as follows. For any fixed integer  $\mathcal{N} > 0$ , we denote  $\delta t = T/\mathcal{N}$  and define the uniform subdivision  $\{t_i\}_{i=0,\dots,\mathcal{N}}$  of  $[0, T]$  with  $t_i = i\delta t$ . We then approximate the curve  $\gamma$  in the space of dimension  $\mathcal{N} + 1$

$$\mathbb{P}_1^{\delta t} = \left\{ \gamma \in C([0, T]); \gamma|_{[t_{i-1}, t_i]} \text{ affine, } \forall i \in \{1, \dots, \mathcal{N}\} \right\}.$$

For any  $\gamma \in \mathbb{P}_1^{\delta t}$ ,  $\gamma = \sum_{i=0}^{\mathcal{N}} \gamma^i L_i^{\delta t}$  with  $(\gamma^i)_{0 \leq i \leq \mathcal{N}} \in \Omega^{\mathcal{N}+1}$  where  $(L_i^{\delta t})_{0 \leq i \leq \mathcal{N}}$  is the usual Lagrange basis. Consequently,  $\gamma$  is defined by the  $\mathcal{N} + 1$  points  $(\gamma^i, t_i) \in \Omega \times [0, T]$ . The knowledge of the initial curve  $\gamma_0 \in \mathbb{P}_1^{\delta t}$  such that  $\delta_0 \leq \gamma_0 \leq 1 - \delta_0$  determines such points and then a triangular mesh of  $Q_T$ . At each iteration  $n \geq 0$ , these points are updated along the  $x$ -axis according to the pointwise time descent direction  $j_{\gamma_n}^\epsilon \in H^1(0, T)$  (see (3.5)) as follows:

$$x_i^{n+1} = P_{[\delta_0, 1-\delta_0]} \left( x_i^n - \rho j_{\gamma_n}^\epsilon(t_i^n) \right), \quad t_i^{n+1} = t_i^n, \quad \forall i = 0, \dots, \mathcal{N} + 1.$$

We emphasize that a re-meshing of  $Q_T$  is performed at each iteration  $n$  according to the set of points  $(x_i^n, t_i^n)_{i=0,\dots,\mathcal{N}+1}$ .

Remark that, for a fixed value of  $\mathcal{N}$ , the curves  $\gamma \in \mathbb{P}_1^{\delta t}$  are uniformly Lipschitz with constant  $M = \frac{1}{\delta t}$ , as required in  $\mathcal{G}_{\text{ad}}$ .

– Each iteration of the algorithm requires the numerical approximation of the control of minimal  $L^2(q_{\gamma_n})$ -norm for the initial datum  $(y_0, y_1) \in \mathbf{V}$ . We use the space-time method described in Sections 3–4 of [6], which is very well-adapted to the description of  $\gamma$  embedded in a space-time mesh of  $Q_T$ . The minimization of the conjugate functional  $\mathcal{J}_{\gamma_n}^*$  (see (3.4)) with respect to  $(\varphi_0, \varphi_1) \in \mathbf{W}$  is replaced by the search of the unique saddle-point of the Lagrangian  $\mathcal{L} : \mathbf{Z} \times L^2(0, T; H_0^1(\Omega)) \rightarrow \mathbb{R}$  defined by

$$\mathcal{L}(\varphi, y) = \frac{1}{2} \|\varphi\|_{L^2(q_{\gamma_n})}^2 - \int_0^T \langle L\varphi, y \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \langle \varphi_t(\cdot, 0), y_0 \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} - \langle \varphi(\cdot, 0), y_1 \rangle_{L^2(\Omega)},$$

with  $\mathbf{Z} = C([0, T]; L^2(\Omega)) \cap C^1([0, T]; H^{-1}(\Omega))$ . The corresponding mixed formulation is solved with a conformal space-time finite element method, while a direct method is used to invert the discrete matrix. The interesting feature of the method, for which the adaptation of the mesh is very simple to handle, is that only a small part of the matrix – corresponding to the term  $\|\varphi\|_{L^2(q_{\gamma_n})}^2$  – is modified between two consecutive iterations  $n$  and  $n + 1$ . Moreover, this space-time framework leads to strong convergent approximations of the control as the fineness of the mesh goes to zero. We refer to [6] for more details.

##### 4.1. Numerical illustrations

We discuss several experiments performed with FreeFEM++ (see [14]) for various initial data and control domains. We notably use an UMFPACK type solver.

We denote in this section by  $J_\epsilon^n$  and  $dJ_\epsilon^n$  a numerical approximation of  $J_\epsilon(\gamma_n)$  and  $dJ_\epsilon(\gamma_n)$  respectively.

TABLE 2. **(EX1)** – Number of iterations, optimal value of the functional  $J_\epsilon$  and performance index, for the initial curves  $(\gamma_0^i)_{i \in \{1,2,3\}}$  given by (4.1).

Initial curve	$\gamma_0^1$	$\gamma_0^2$	$\gamma_0^3$
Number of iterations	33	33	84
$J_\epsilon(\gamma_{\text{opt}})$	47.09	47.09	47.93
$\Pi(\gamma_{\text{opt}})$	-0.32%	-0.32%	-2.11%

We fix  $\delta_0 = 0.15$  and  $\delta = \delta_0/4$ . Moreover, according to (3.1), we define the function  $\chi \in C^2(\mathbb{R})$  in  $[\delta_0 - \delta, \delta_0]$  as the unique polynomial of degree 5 such that  $\chi(\delta_0 - \delta) = 1$ ,  $\chi(\delta_0) = \chi'(\delta_0 - \delta) = \chi'(\delta_0) = \chi''(\delta_0 - \delta) = \chi''(\delta_0) = 0$  and vanishing on  $\mathbb{R} \setminus [\delta_0 - \delta, \delta_0]$ .

Concerning the stopping criterion for the descent algorithm, we observed that the usual one based on the relative quantity  $|J_\epsilon^n - J_\epsilon^{n-1}|/J_\epsilon^0$  is inefficient because too noisy. This is due to the uncertainty on the numerical computation of the adjoint state  $\varphi^n$  and the perturbation  $j_{\gamma_n}^\epsilon$ . Consequently, in order to better capture the variations of the sequence  $(J_\epsilon^n)_{n \in \mathbb{N}}$ , we replace  $J_\epsilon^n$  and  $J_\epsilon^{n-1}$  by the right and left  $p$ -point average respectively leading to the stopping criterion

$$\Delta J_\epsilon^n := \left| \frac{\frac{1}{p} \sum_{i=n}^{n+p-1} J_\epsilon^i - \frac{1}{p} \sum_{i=n-p}^{n-1} J_\epsilon^i}{J_\epsilon^0} \right| < \eta, \quad \text{for } p \in \mathbb{N}^* \text{ fixed.}$$

In the sequel, we fix  $p = 10$  and  $\eta = 10^{-3}$ .

Furthermore, in order to measure the gain obtained by using non-cylindrical control domains rather than cylindrical ones, we introduce a performance index associated with each optimal curves  $\gamma_{\text{opt}}$ : identifying any constant curve  $\gamma \equiv x_0$  with its value  $x_0 \in [\delta_0, 1 - \delta_0]$ , we compute the minimal cost  $\min_{x_0} J_\epsilon(x_0)$  for cylindrical domains; the performance index of  $\gamma_{\text{opt}}$  is then defined as follows:

$$\Pi(\gamma_{\text{opt}}) = 100 \left( 1 - \frac{J_\epsilon(\gamma_{\text{opt}})}{\min_{x_0} J_\epsilon(x_0)} \right).$$

In the sequel, in practice, the minimum of  $J_\epsilon$  with respect to  $x_0$  is searched among 13 distinct values equi-distributed between 0.2 and 0.8.

– We first consider the regular initial datum  $(y_0, y_1)$  given by

$$y_0(x) = \sin(2\pi x), \quad y_1(x) = 0, \quad \text{for } x \in (0, 1) \quad \text{(EX1)}$$

and  $T = 2$ ,  $\epsilon = 10^{-2}$ ,  $\rho = 10^{-4}$ . We initialize the descent algorithm with the following three initial curves:

$$\gamma_0^1(t) = \frac{2}{5}, \quad \gamma_0^2(t) = \frac{3}{5}, \quad \gamma_0^3(t) = \frac{1}{2} + \frac{1}{10} \cos\left(\pi \frac{t}{T}\right), \quad \text{for } t \in (0, T). \quad (4.1)$$

The corresponding initial and optimal domains are depicted in Figure 5 together with typical space-time meshes. The numbers of iterations until convergence, the values of the functional  $J_\epsilon$  evaluated at the optimal curve  $\gamma_{\text{opt}}$  and the performance indices of  $\gamma_{\text{opt}}$  are listed in Table 2.

In Figure 5, we observe that the optimal domain computed by the algorithm depends on the chosen initial domain. This indicates that the functional  $J_\epsilon$  have several local minima. Moreover, one can show that, among the cylindrical domains, there are two optimal values,  $x_0 = 1/4$  and  $x_0 = 3/4$ , leading to  $J_\epsilon(x_0) \approx 46.94$ . These values correspond to the extrema of the function  $x \rightarrow \sin(2\pi x)$ ,  $x \in [0, 1]$ . The simulations associated with

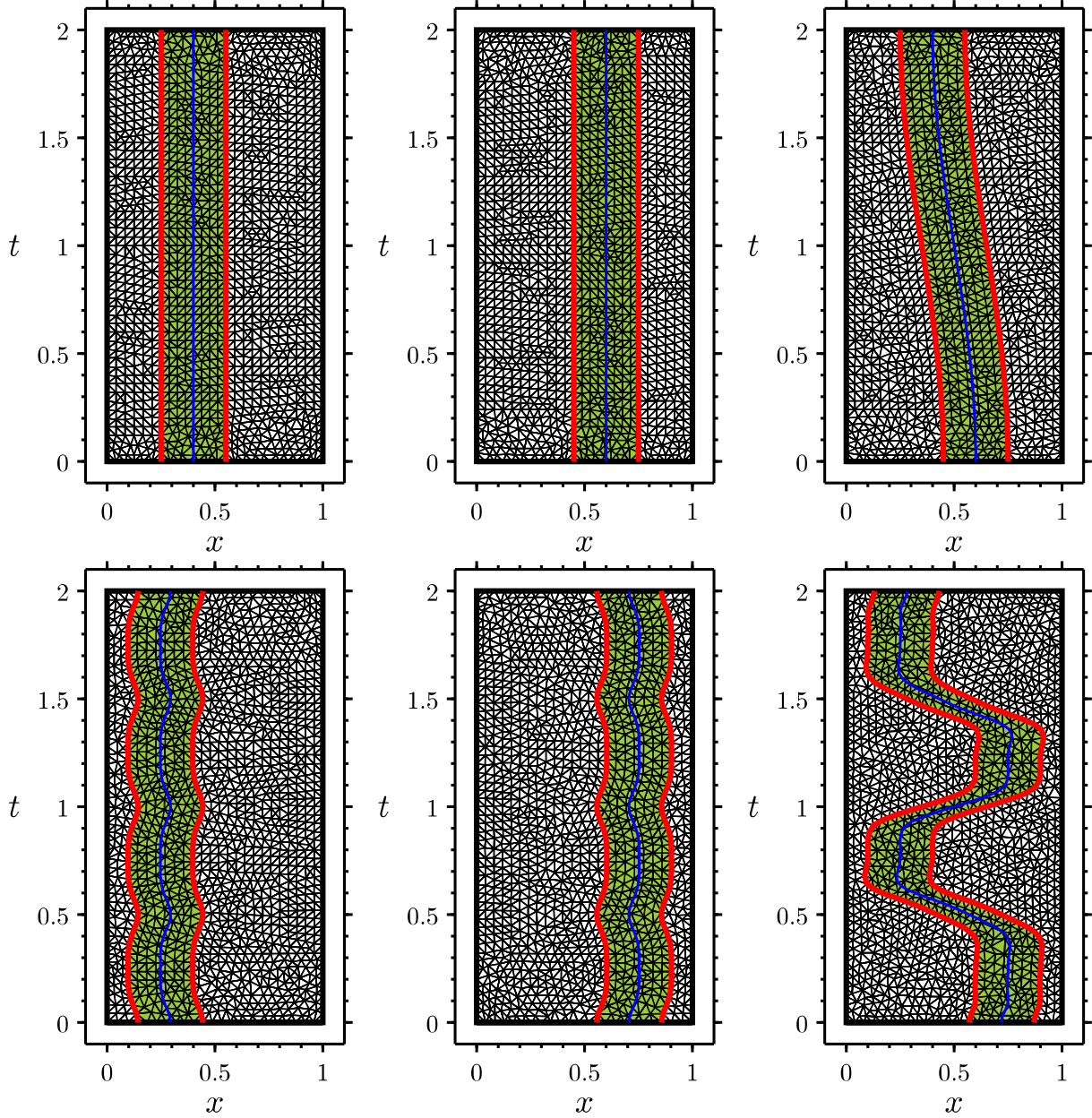


FIGURE 5. **(EX1)** – Initial *top* and optimal (*bottom*) control domains for the initial curves  $(\gamma_0^i)_{i \in \{1,2,3\}}$  given by (4.1) (from *left to right*).

the initial curves  $\gamma_0^1$  and  $\gamma_0^2$  are in agreement with this fact. On the other hand, the worst cylindrical domain corresponds to  $x_0 = 1/2$  (see Fig. 7-Left).

Eventually, the adjoint states  $\varphi$  (from which we obtain the control  $v = \varphi|_{q_0^n}$ ) computed for the optimal domains in Figure 5-Bottom, are displayed in Figure 6.

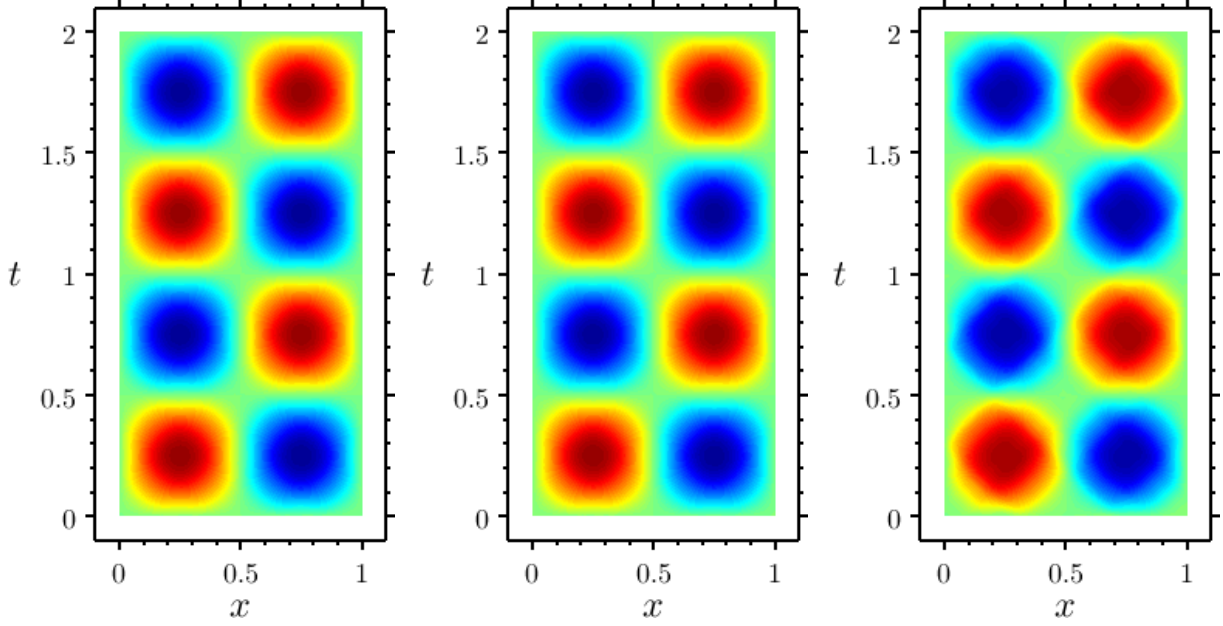


FIGURE 6. **(EX1)** – Isovalues of the adjoint states  $\varphi$  computed for the optimal domains obtained for the initial curves  $(\gamma_0^i)_{i \in \{1,2,3\}}$  given by (4.1) (from *left to right*) and for the initial datum **(EX1)**.

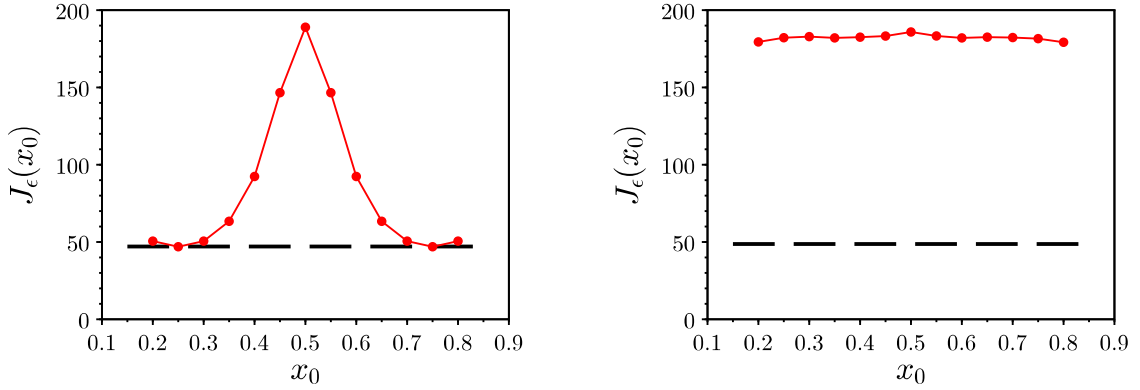


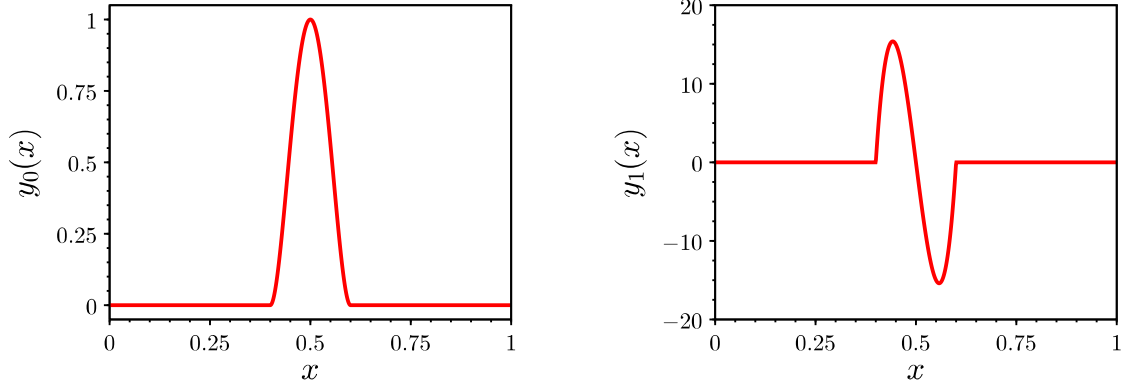
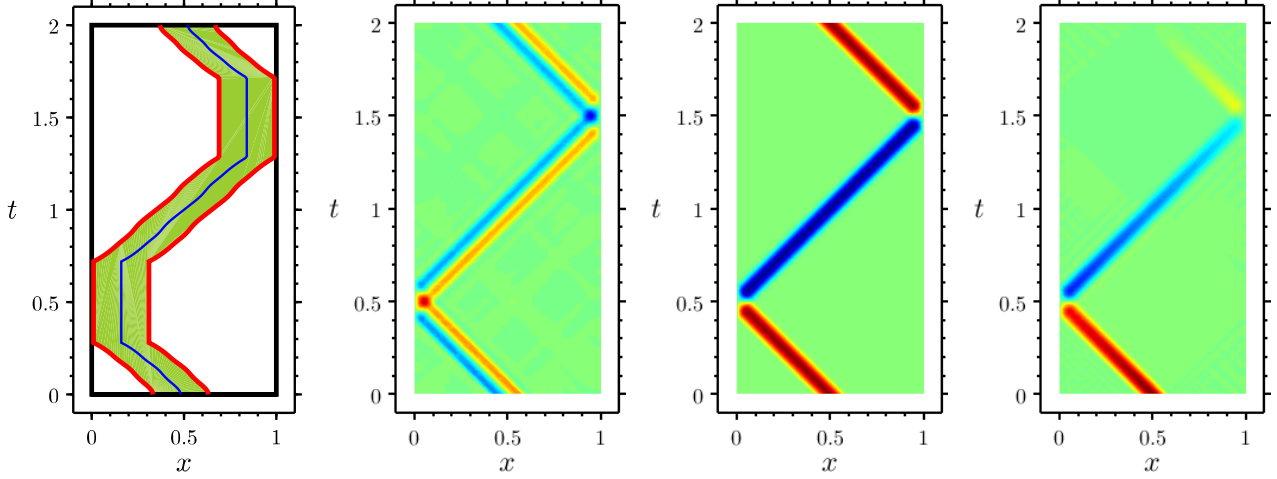
FIGURE 7. Values of  $J_\epsilon$  for constant curves  $\gamma \equiv x_0$  ( $\bullet$ ), for the initial datum **(EX1)** (*left*) and **(EX2)** (*right*). The dashed line (- -) represents the value of  $J_\epsilon(\gamma_{\text{opt}})$ , for the initial curves  $\gamma_0 \equiv 2/5$  (*left*) and  $\gamma_0 \equiv 1/2$  (*right*).

– We now consider the initial datum  $(y_0, y_1)$  given by

$$y_0(x) = (10x - 4)^2(10x - 6)^2 \mathbb{1}_{[0.4, 0.6]}(x), \quad y_1(x) = y_0'(x), \quad \text{for } x \in (0, 1). \quad \text{(EX2)}$$

As can be seen in Figure 9–3, this initial condition, plotted in Figure 8, generates a travelling wave.

For  $T = 2$ ,  $\epsilon = 10^{-2}$  and  $\rho = 10^{-4}$ , we initialize the descent algorithm with the curve  $\gamma_0 \equiv 1/2$ . The convergence is reached after 68 iterations and the optimal cost is  $J_\epsilon(\gamma_{\text{opt}}) \approx 48.70$ . Moreover, the minimal cost

FIGURE 8. Initial datum  $(y_0, y_1)$  defined in (EX2).FIGURE 9. (EX2) – From left to right, optimal control domain, isovalues of the corresponding adjoint state  $\varphi$ , isovalues of the uncontrolled and controlled wave over the optimal domain, for the initial curve  $\gamma_0 \equiv 1/2$ .

for cylindrical domains is  $\min_{x_0} J_\epsilon(x_0) \approx 179.22$  leading to a performance index  $\Pi(\gamma_{\text{opt}}) \approx 72.83\%$ . The non-cylindrical setup is in that case much more efficient than the cylindrical one. This is due to the fact that the domains we consider can follow very closely the propagation of the travelling wave. This can be noticed in Figure 9, where we display the optimal control domain, the corresponding adjoint state  $\varphi$ , the uncontrolled and controlled solutions over the optimal domain.

The evolution of the cost  $J_\epsilon^n$  and the derivative  $dJ_\epsilon^n$  with respect to  $n$  are displayed in Figure 10. Figure 7-Right depicts the values of the functional  $J_\epsilon$  for the constant curves  $\gamma \equiv x_0$  used to determine the best cylindrical domain and highlights the low variation of the cost with respect to the position of such domains.

This second example allows to emphasize the sensitivity of the optimal curve with respect to the initial condition. For any  $\delta$ , consider the initial condition  $(y_0^\delta, y_1^\delta) = (y_0, y_1) + \delta(w_0, w_1)$ , where  $(y_0, y_1)$  is given by (EX2) and supported in the interval  $[0.4, 0.6]$ , and where  $(w_0, w_1)$  is any initial condition supported for instance in  $[0.8, 0.9]$ . For  $\delta = 0$  and any  $T > 0$ , the optimal curve  $\gamma_{\text{opt}}$  follows the travelling wave starting in the interval  $[0.4, 0.6]$ . On the other hand, for any  $\delta \neq 0$ , the optimal curve  $\gamma_{\text{opt}}^\delta$  will also have to cross the travelling waves generated by the initial condition  $(w_0, w_1)$  starting in  $[0.8, 0.9]$ . If the controllability time  $T$  is small enough,

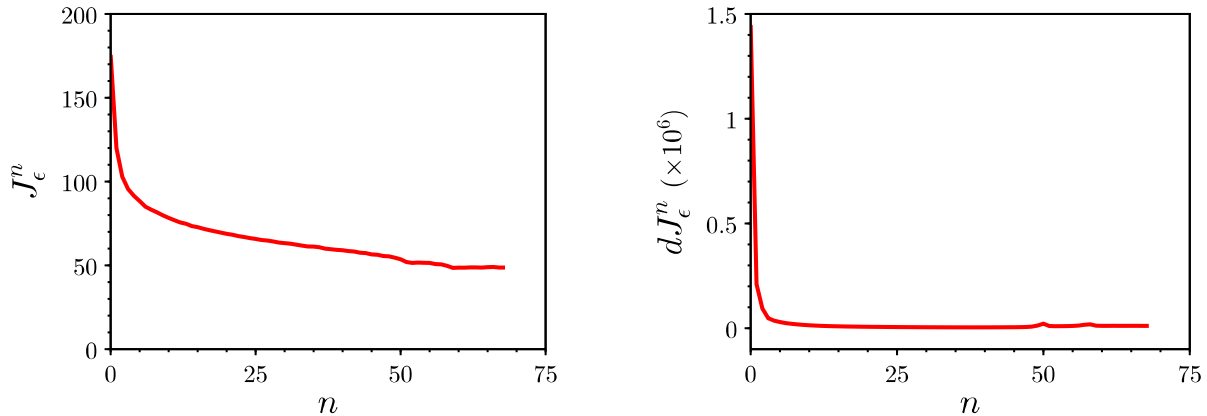


FIGURE 10. **(EX2)** – Evolution of the cost  $J_\epsilon^n$  (left) and the derivative  $dJ_\epsilon^n$  (right) for the initial curve  $\gamma_0 \equiv 1/2$ .

TABLE 3. **(EX3)** – Evolution of the optimal cost  $J_\epsilon(\gamma_{\text{opt}})$  w.r.t. the controllability time  $T$ .

$T$	2	1.5	1	0.5
$J_\epsilon(\gamma_{\text{opt}})$	41.02	58.02	94.78	239

these additional travelling waves remain far from the initial wave. It follows that, for any  $\delta \neq 0$ , the resulting optimal curve  $\gamma_{\text{opt}}^\delta$  will be not arbitrarily close to  $\gamma_{\text{opt}}$ .

– We now consider the initial datum  $(y_0, y_1)$  given by

$$y_0(x) = (10x - 4)^2(10x - 6)^2 \mathbb{1}_{[0.4, 0.6]}(x), \quad y_1(x) = 0, \quad \text{for } x \in (0, 1). \quad \text{(EX3)}$$

This initial condition generates two waves travelling waves in opposite directions, as can be seen in Figure 11.3. For  $T = 2$ ,  $\epsilon = 10^{-2}$  and  $\rho = 10^{-4}$ , we initialize the algorithm with the initial curve  $\gamma_0 \equiv 1/2$ . The convergence is observed after 111 iterations leading to  $J_\epsilon(\gamma_{\text{opt}}) \approx 41.02$ . Moreover, the minimal cost for cylindrical domains is  $\min_{x_0} J_\epsilon(x_0) \approx 85.08$  leading to a performance index  $\Pi(\gamma_{\text{opt}}) \approx 51.79\%$ . Once again, our non-cylindrical setup is much more efficient than the cylindrical one. It is still due to the fact that the domains we consider can follow the propagation of the travelling waves. This can be noticed in Figure 11, where we display the optimal control domain, the corresponding adjoint state  $\varphi$ , the uncontrolled and the controlled wave over the optimal domain.

In order to emphasize the influence of the controllability time on the optimal domain, for  $\epsilon = 10^{-2}$  and  $\gamma_0 \equiv 1/2$ , we use the descent algorithm with  $T = 1$  and  $\rho = 2.5 \times 10^{-5}$ , initialized with the curve  $\gamma_0 \equiv 1/2$ . Remark that the corresponding domain satisfies the geometric optic condition. The convergence is observed after 213 iterations and the optimal cost is  $J_\epsilon(\gamma_{\text{opt}}) \approx 94.78$ . Moreover, the minimal cost for cylindrical domains is  $\min_{x_0} J_\epsilon(x_0) \approx 183.98$  and the performance index is  $\Pi(\gamma_{\text{opt}}) \approx 48.48\%$ . Observe that the cylindrical domains associated with  $x_0 \notin (0.25, 0.75)$  do not verify the geometric optic condition. This highlights the necessity to use non-cylindrical domains. Compared to the simulation for  $T = 2$ , the optimal cost increases by a factor around 2.3. Figure 12 displays the optimal control domain, the corresponding adjoint state  $\varphi$ , the uncontrolled and controlled wave over the optimal domain. We remark that the projection of the optimal domain on the  $x$ -axis covers the whole domain  $\Omega$ , in contrast with the domain associated with  $T = 2$ .

In Table 3, we illustrate the influence of the controllability time  $T$  on the optimal cost  $J_\epsilon(\gamma_{\text{opt}})$ . More exactly, we remark that the cost is of order  $\mathcal{O}(T^{-1.27})$ .

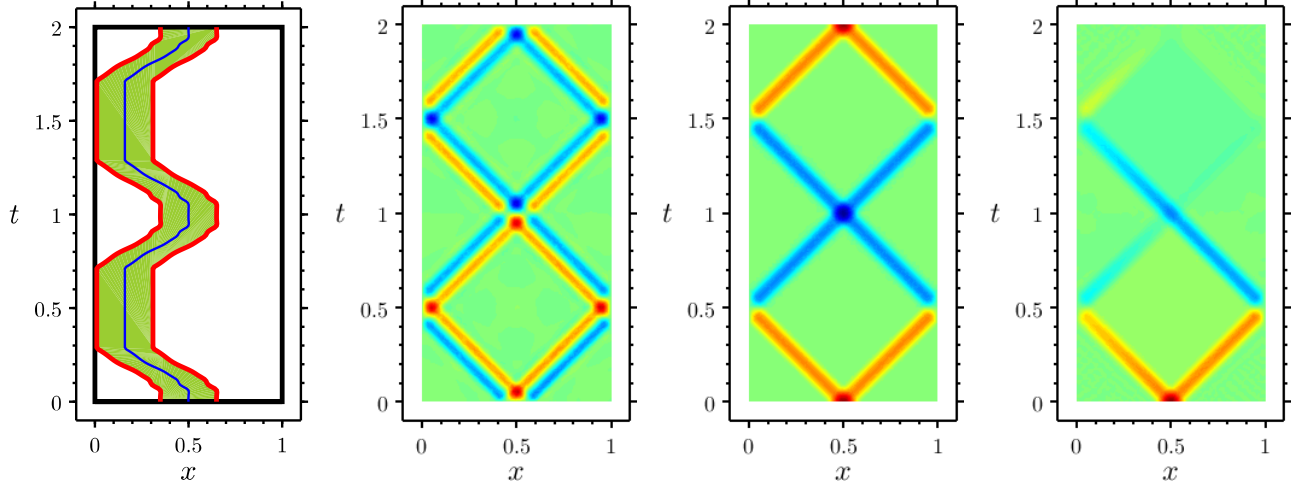


FIGURE 11. **(EX3)** – From left to right, optimal control domain, isovalues of the corresponding adjoint state  $\varphi$ , isovalues of the uncontrolled and controlled wave over the optimal domain, for  $T = 2$ , for the initial curve  $\gamma_0 \equiv 1/2$ .

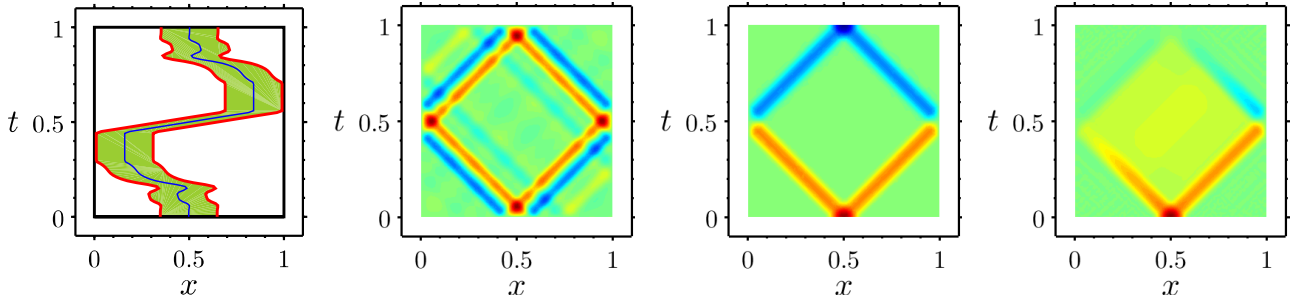


FIGURE 12. **(EX3)** – From left to right, optimal control domain, isovalues of the corresponding adjoint state  $\varphi$ , isovalues of the uncontrolled and controlled wave over the optimal domain, for  $T = 1$ , for the initial curve  $\gamma_0 \equiv 1/2$ .

– Eventually, in order to highlight the influence of the regularization parameter  $\epsilon$  on the optimal domain, we now consider the initial datum  $(y_0, y_1)$  given by

$$y_0(x) = \begin{cases} 3x & \text{if } 0 \leq x \leq 1/3, \\ 3(1-2x) & \text{if } 1/3 \leq x \leq 2/3, \\ -3(1-x) & \text{if } 2/3 \leq x \leq 1, \end{cases} \quad y_1(x) = 0, \quad \text{for } x \in (0, 1). \quad (\text{EX4})$$

For  $T = 2$  and  $\rho = 10^{-5}$ , we initialize the descent algorithm with the curve  $\gamma \equiv 1/2$  and consider  $\epsilon = 10^{-2}$  and  $\epsilon = 0$ . The numbers of iterations until convergence, the values of the functional  $J_\epsilon$  evaluated at the optimal curve  $\gamma_{\text{opt}}$  and the performance indices of  $\gamma_{\text{opt}}$  are listed in Table 4. For the initial datum **(EX4)**, the minimal cost for cylindrical domains is  $\min_{x_0} J_\epsilon(x_0) \approx 47.71$ .

In Figure 13, we clearly see the regularizing effect of  $\epsilon$  and the need of regularization in this case, as the optimal domain obtained when  $\epsilon = 0$  is very oscillating. Actually, since the initial position  $y_0$  is supported on the whole domain, the corresponding uncontrolled solution is not localized on a small part of  $Q_T$  (in contrast for instance to the second example, see **(EX2)**). The optimal domain is very likely composed of many parts



TABLE 4. (EX4) – Number of iterations, optimal value of the functional  $J_\epsilon$  and performance index, for  $\epsilon \in \{0, 10^{-2}\}$ , for the initial curve  $\gamma_0 \equiv 1/2$ .

$\epsilon$	0	$10^{-2}$
Number of iterations	247	389
$J_\epsilon(\gamma_{\text{opt}})$	60.35	43.23
$\Pi(\gamma_{\text{opt}})$	-26.51%	9.38%

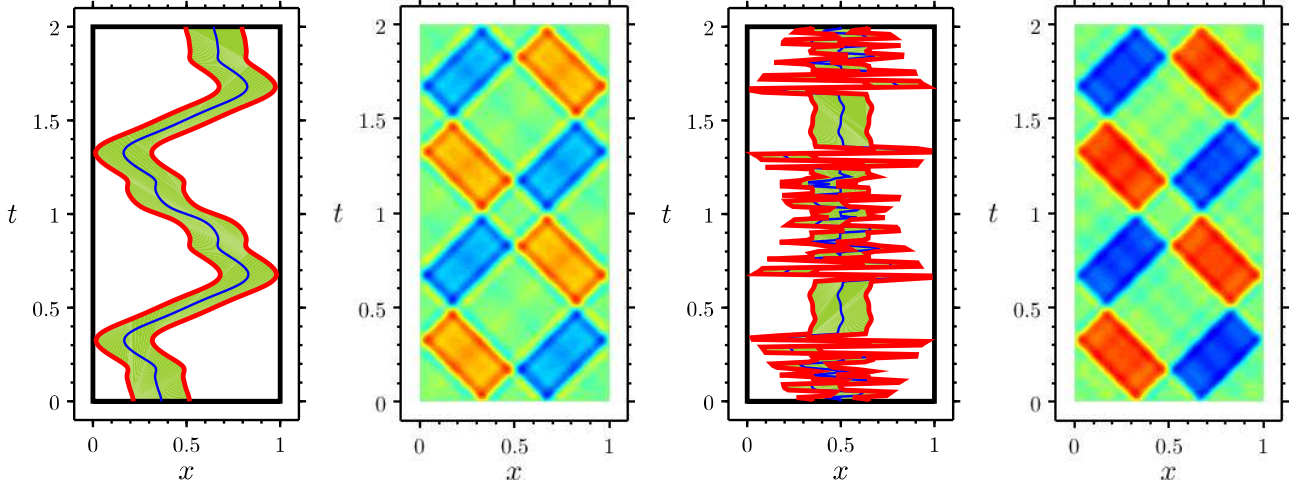


FIGURE 13. (EX4) – Optimal control domain and isovalues of the adjoint state for  $\epsilon = 10^{-2}$  (left),  $\epsilon = 0$  (right), for the initial curve  $\gamma_0 \equiv 1/2$ .

distributed in  $Q_T$  leading an homogenization phenomenon. Consequently, when  $\epsilon = 0$ , the algorithm tends to reach these parts leads to large oscillations of the curve  $\gamma$ . The  $\epsilon$ -term aims to reduce these oscillations by imposing a regularity property of the curves  $\gamma$  describing the domains  $q_\gamma$ . In this topic, we refer to the seminal work of Chenais [7]. Let us also mention that a well-posedness result is achieved in [29] by introducing a relaxation of the problem 1.4, *i.e.* by replacing characteristic functions by densities (see also [26] in the context of stabilization of the wave equation).

#### 4.2. Approximation of the observability constant by an iterative method

In this last part, we formally describe and use an algorithm allowing to approximate the observability constant appearing in (1.2), associated with any domain  $q \subset Q_T$ . The algorithm is based on the following characterization:

$$C_{\text{obs}}(q) = \sup_{\mathbf{y}_0 \in \mathbf{V}} \frac{\langle R\Lambda_q \mathbf{y}_0, \mathbf{y}_0 \rangle_{\mathbf{V}}}{\|\mathbf{y}_0\|_{\mathbf{V}}^2} \quad (4.2)$$

where  $\Lambda_q$  and  $R$  are respectively the control operator associated with the domain  $q$  and the duality operator between the space  $\mathbf{W}$  and  $\mathbf{V}$ :

$$\Lambda_q : \begin{cases} \mathbf{V} & \rightarrow \mathbf{W} \\ \mathbf{y}_0 & \mapsto \widehat{\varphi}_0 \end{cases}, \quad R : \begin{cases} \mathbf{W} & \rightarrow \mathbf{V} \\ (\varphi_0, \varphi_1) & \mapsto ((-\partial_x^2)^{-1}\varphi_1, -\varphi_0) \end{cases}.$$



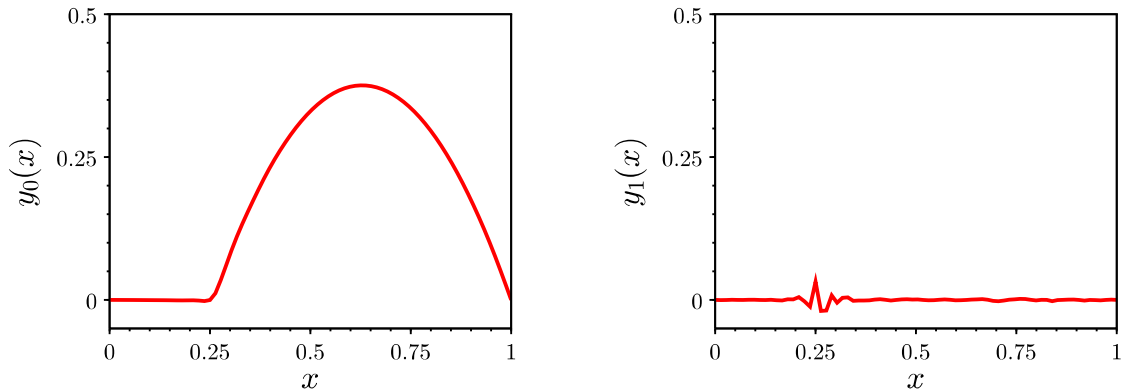


FIGURE 14. Most expensive initial data  $y_0$  (left) and  $y_1$  (right) to be controlled.

In the definition of  $\Lambda_q$ ,  $\widehat{\varphi}_0 \in \mathbf{W}$  is the minimum of the functional  $\mathcal{J}^*$  (see (3.4)) associated with  $\mathbf{y}_0 \in \mathbf{V}$ . The characterization (4.2) can be obtained by following the steps of Section 2 in [25] and Remark 2.98 in [10]. The main consequence of this characterization is that  $C_{\text{obs}}(q)$  can be viewed as the largest eigenvalue of the operator  $R\Lambda_q$  in  $\mathbf{V}$ . Consequently, we can formally adapt the power iteration method to our infinite-dimensional setting. The algorithm reads as follows. Let  $\mathbf{y}_0^0 \in \mathbf{V}$  be given such that  $\|\mathbf{y}_0^0\|_{\mathbf{V}} = 1$ . For  $n \geq 0$ , using the space-time finite element method described in Section 3–4 of [6], we compute  $\widehat{\varphi}_0^n = \Lambda_q \mathbf{y}_0^n$  then set  $\mathbf{z}_0^n = R\widehat{\varphi}_0^n$  and  $\mathbf{y}_0^{n+1} = \mathbf{z}_0^n / \|\mathbf{z}_0^n\|_{\mathbf{V}}$ . We finally have  $C_{\text{obs}}(q) = \lim_{n \rightarrow \infty} \|\mathbf{z}_0^n\|_{\mathbf{V}}$  while  $\mathbf{y}_0^n$  converges in  $\mathbf{V}$  to the most expensive initial datum to control. For the control domain of Figure 3, this algorithm initialized with  $\mathbf{y}_0^0 = K(x(1-x), 0)$ ,  $K$  such that  $\|\mathbf{y}_0^0\|_{\mathbf{V}} = 1$ , produces the following sequence  $\{\|\mathbf{z}_0^n\|_{\mathbf{V}}\}_{n \geq 0} = \{2.689, 3.829, 3.981, 3.994, 3.997, \dots\}$  converging toward the value 4, in agreement with the result of Section 2.2 based on a graph argument. The most expensive initial datum to be controlled is displayed in Figure 14. Remark that the initial datum solution of (4.2) is not unique.

## 5. CONCLUSION AND PERSPECTIVES

Making use of the d'Alembert formula for the solutions of the one dimensional wave equation, we have shown a uniform observability inequality with respect to a class of non cylindrical observation domains satisfying the geometric optic condition. The proof based on arguments from graph theory allows notably to relate the value of the observability constant to the spectrum of the Laplacian matrix, defined in term of the graph associated to any domain  $q \subset Q_T$ . The uniform observability property then allows to consider and analyze the problem of the control's optimal support associated to fixed initial conditions. For simplicity, the optimization is made over connected domains defined by regular curves. As expected, the optimal domains (approximated within a space-time finite element method) are closely related to the travelling waves generated by the initial conditions.

This work may be extended to several directions. First, the characterization of the observability constant in term of a computable eigenvalue problem in Section 4.2 may allow to consider the optimization of such constant with respect to the domain of observation, *i.e.*  $\inf_{q \in \mathcal{Q}_{\text{ad}}^{\varepsilon}} C_{\text{obs}}(q)$  (as done in [31]). The domain providing the smallest observability constant is nothing else than the trajectory minimizing the cost of control, *i.e.* the norm of the control corresponding to the *worst controllable* initial datum. Moreover, from an approximation point of view, we may also consider more general domains (than connected ones) and use, for instance, a level-set method to describe the geometry (as done in [23]). Eventually, this work may be adapted to the case of controls supported on single curves of  $Q_T$ , using the uniform observability property given in [5]. We refer to [3] for a numerical study in this case. The extension of this work to the multi-dimensional case studied in [17, 32] is also a challenge.

## REFERENCES

- [1] H. Banks, W. Fang, R. Silcox and R. Smith, Approximation methods for control of structural acoustics models with piezoceramic actuators. *J. Intell. Mater. Syst. Struct.* **4** (1993) 98–116.
- [2] C. Bardos, G. Lebeau and J. Rauch, Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.* **30** (1992) 1024–1065.
- [3] A. Bottois, Pointwise moving control for the 1-D wave equation – Numerical approximation and optimization of the support. To appear in *Radon Series on Computational and Applied Mathematics*. De Gruyter, In press. <https://hal.archives-ouvertes.fr/hal-02972968> (2021).
- [4] A.E. Brouwer and W.H. Haemers, Spectra of graphs, Universitext. Springer, New York (2012).
- [5] C. Castro, Exact controllability of the 1-D wave equation from a moving interior point. *ESAIM: COCV* **19** (2013) 301–316.
- [6] C. Castro, N. Cîndea and A. Münch, Controllability of the linear one-dimensional wave equation with inner moving forces. *SIAM J. Control Optim.* **52** (2014) 4027–4056.
- [7] D. Chenais, On the existence of a solution in a domain identification problem. *J. Math. Anal. Appl.* **52** (1975) 189–219.
- [8] F.R.K. Chung, Spectral graph theory. Vol. 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI (1997).
- [9] N. Cîndea and A. Münch, A mixed formulation for the direct approximation of the control of minimal  $L^2$ -norm for linear type wave equations. *Calcolo* **52** (2015) 245–288.
- [10] J.-M. Coron, Control and nonlinearity. Vol. 136 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2007).
- [11] L. Cui, X. Liu and H. Gao, Exact controllability for a one-dimensional wave equation in non-cylindrical domains. *J. Math. Anal. Appl.* **402** (2013) 612–625.
- [12] P. Destuynder, I. Legrain, L. Castet and N. Richard, Theoretical, numerical and experimental discussion on the use of piezoelectric devices for control–structure interaction. *Eur. J. Mech. A. Solids* **11** (1992) 181–213.
- [13] B.H. Haak and D.-T. Hoang, Exact observability of a 1-dimensional wave equation on a noncylindrical domain. *SIAM J. Control Optim.* **57** (2019) 570–589.
- [14] F. Hecht, New development in Freefem++. *J. Numer. Math.* **20** (2012) 251–265.
- [15] A. Henrot and M. Pierre, Variation et optimisation de formes. Une analyse géométrique. [A geometric analysis]. Vol. 48 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Berlin (2005).
- [16] A.Y. Khapalov, Controllability of the wave equation with moving point control. *Appl. Math. Optim.* **31** (1995) 155–175.
- [17] J. Le Rousseau, G. Lebeau, P. Terpolilli and E. Trélat, Geometric control condition for the wave equation with a time-dependent observation domain. *Anal. Partial Differ. Equ.* **10** (2017) 983–1015.
- [18] J.-L. Lions, Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos, G. Lebeau and J. Rauch. Tome 1. Vol. 8 of *Recherches en Mathématiques Appliquées [Research in Applied Mathematics]*. Masson, Paris (1988).
- [19] K. Liu and J. Yong, Rapid exact controllability of the wave equation by controls distributed on a time-variant subdomain. *Chin. Ann. Math. Ser. B* **20** (1999) 65–76. A Chinese summary appears in *Chin. Ann. Math. Ser. A* **20** (1999) 142.
- [20] K.A. Lurie, An introduction to the mathematical theory of dynamic materials. Vol. 15 of *Advances in Mechanics and Mathematics*. Second edition, Springer, Cham (2017) MR2305885.
- [21] P. Martin, L. Rosier and P. Rouchon, Null controllability of the structurally damped wave equation with moving control. *SIAM J. Control Optim.* **51** (2013) 660–684.
- [22] B. Mohar, The Laplacian spectrum of graphs. Vol. 2 of Graph theory, combinatorics, and applications. (Kalamazoo, MI, 1988). Wiley-Intersci. Publ., Wiley, New York (1991) 871–898.
- [23] A. Münch, Optimal design of the support of the control for the 2-D wave equation: a numerical method. *Int. J. Numer. Anal. Model.* **5** (2008) 331–351.
- [24] A. Münch, Optimal location of the support of the control for the 1-D wave equation: numerical investigations. *Comput. Optim. Appl.* **42** (2009) 443–470.
- [25] A. Münch, Numerical estimations of the cost of boundary controls for the equation  $y_t - \varepsilon y_{xx} + My_x = 0$  with respect to  $\varepsilon$ , in Recent advances in PDEs: analysis, numerics and control. Vol. 17 of *SEMA SIMAI Springer Ser.* Springer, Cham (2018) 159–191.
- [26] A. Münch, P. Pedregal and F. Periago, Optimal design of the damping set for the stabilization of the wave equation. *J. Differ. Equ.* **231** (2006) 331–358.
- [27] A.O. Özer, Potential formulation for charge or current-controlled piezoelectric smart composites and stabilization results: electrostatic versus quasi-static versus fully-dynamic approaches. *IEEE Trans. Automat. Control* **64** (2019) 989–1002.
- [28] A.O. Özer and K.A. Morris, Modeling and stabilization of current-controlled piezo-electric beams with dynamic electromagnetic field. *ESAIM:COCV* **26** (2020) 24.
- [29] F. Periago, Optimal shape and position of the support for the internal exact control of a string. *Systems Control Lett.* **58** (2009) 136–140.
- [30] Y. Privat, E. Trélat and E. Zuazua, Optimal location of controllers for the one-dimensional wave equation. *Ann. Inst. Henri Poincaré Anal. Non Linéaire* **30** (2013) 1097–1126.

- [31] Y. Privat, E. Trélat and E. Zuazua, Optimal observation of the one-dimensional wave equation. *J. Fourier Anal. Appl.* **19** (2013) 514–544.
- [32] A. Shao, On Carleman and observability estimates for wave equations on time-dependent domains. *Proc. Lond. Math. Soc.* **119** (2019) 998–1064.
- [33] M. Tucsnak, Control of plate vibrations by means of piezoelectric actuators. *Discrete Contin. Dynam. Syst.* **2** (1996) 281–293.

## B Contrôle ponctuel pour l'équation des ondes 1D – Approximation numérique et optimisation du support

Titre : *Pointwise moving control for the 1-D wave equation – Numerical approximation and optimization of the support.*

Auteur : ARTHUR BOTTOIS.

Publié dans *Radon Series on Computational and Applied Mathematics, Volume 29, 2022.*

<https://doi.org/10.1515/9783110695984-002>

Aussi disponible à <https://hal.archives-ouvertes.fr/hal-02972968>

### Résumé

Dans cet article, on traite le problème de la contrôlabilité à zéro pour l'équation des ondes linéaire unidimensionnelle avec contrôle ponctuel. On reprend ici les notations de la Sect. 1.1.8 avec  $d = 1$  et  $\Omega = (0, 1)$ , à la différence notable suivante. Dans la suite, l'action du contrôle  $u$  le long de la trajectoire  $x = \gamma(t)$  n'est pas régularisée en temps. Ceci revient à prendre  $\eta = \mathbb{1}_{[0,T]}$  dans le système d'état (1.33) et implique que  $u$  appartient naturellement à l'espace  $\mathcal{U} = (H^1(0, T))'$ , dual de  $H^1(0, T)$ . On peut donc s'attendre à observer des distributions de Dirac dans le comportement de  $u$  en  $t = 0$  et  $t = T$ . Dans ce cadre unidimensionnel, on sait que la contrôlabilité de (1.33) est équivalente à l'inégalité d'observabilité

$$\|(\varphi^0, \varphi^1)\|_{H_0^1 \times L^2}^2 \leq C_{\text{obs}}(\gamma) \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi$  est la solution du système adjoint (1.34) associée à  $(\varphi^0, \varphi^1)$ . Cette inégalité a été établie dans [8] pour des trajectoires de contrôle  $\gamma$  vérifiant certaines hypothèses géométriques analogues au critère (GCC).

Dans la première partie de l'article, on commence par appliquer la méthode HUM (cf. [25]) afin de montrer que le contrôle  $\hat{u}$  de norme  $\mathcal{U}$  minimale s'exprime via l'état adjoint  $\hat{\varphi}$  associé au point de minimum  $(\hat{\varphi}^0, \hat{\varphi}^1)$  de la fonctionnelle

$$\mathcal{J}^*(\varphi^0, \varphi^1) = \frac{1}{2} \|\varphi(\gamma, \cdot)\|_{H^1(0, T)}^2 - \langle y^0, \varphi^1 \rangle_{L^2} + \langle y^1, \varphi^0 \rangle_{H^{-1}, H_0^1}, \quad \forall (\varphi^0, \varphi^1) \in H_0^1(\Omega) \times L^2(\Omega),$$

où  $\varphi(\gamma, \cdot)$  désigne la fonction  $t \mapsto \varphi(\gamma(t), t)$ . Dans un deuxième temps, en vue de calculer une approximation numérique du contrôle de norme minimale, on adapte l'approche espace-temps exposée dans [8] pour l'équation des ondes avec contrôle interne. En introduisant un multiplicateur de Lagrange  $\lambda$  pour la contrainte  $\partial_t^2 \varphi - \Delta \varphi = 0$ , on montre que  $\hat{\varphi}$  est solution d'une formulation mixte de type

$$\begin{cases} a_r(\varphi, \bar{\varphi}) + b(\bar{\varphi}, \lambda) = \ell(\bar{\varphi}), & \forall \bar{\varphi} \in \Phi, \\ b(\varphi, \bar{\lambda}) = 0, & \forall \bar{\lambda} \in \Lambda. \end{cases}$$

On montre que la formulation mixte est bien posée et qu'en outre, le multiplicateur  $\hat{\lambda}$  associé à  $\hat{\varphi}$  n'est autre que la solution contrôlée  $\hat{y}$  de (1.33) associée à  $\hat{u}$ . Dans un troisième temps, on propose et on discute une discrétisation type éléments finis de la formulation mixte. La résolution de cette formulation mixte discrétisée permet alors d'approcher numériquement le contrôle  $\hat{u}$  et la solution contrôlée  $\hat{y}$ . Avec une simulation, on illustre la convergence des approximations  $\hat{u}_h$  et  $\hat{y}_h$  quand le paramètre de discrétisation  $h$  tend vers 0. On illustre également l'absence de contrôlabilité pour une trajectoire stationnaire agissant en un point non stratégique.

Dans la deuxième partie de l'article, pour une donnée initiale  $(y^0, y^1)$  fixée, on cherche à identifier la trajectoire  $\gamma$  permettant de minimiser l'énergie à dépenser pour contrôler l'onde. Le problème à résoudre s'écrit

$$\inf_{\gamma \in \mathcal{G}_{\text{ad}}} J(\gamma), \quad \text{avec} \quad J(\gamma) = \frac{1}{2} \|\hat{u}_\gamma\|_{\mathcal{U}}^2,$$

et où l'ensemble admissible  $\mathcal{G}_{\text{ad}}$  regroupe des trajectoires  $\gamma$  régulières, typiquement de classe  $C^2([0, T])$ , pour lesquelles il existe une constante d'observabilité  $\bar{C}_{\text{obs}} > 0$  uniforme. En l'état et en l'absence de modifications qui apporterait plus de régularité sur le contrôle, on ne sait pas si le problème d'optimisation est bien posé. On se concentre donc sur sa résolution numérique. D'une part, pour donner plus de sens au problème, on pénalise la longueur  $L(\gamma)$  de la trajectoire. D'autre part, afin d'éviter que les

trajectoires ne présentent de brusques variations, on pénalise aussi la “courbure”  $\gamma''$ . Ainsi, pour  $\varepsilon > 0$ ,  $\eta > 0$  et  $\bar{L} \geq T$  fixés, on réécrit le problème

$$\inf_{\gamma \in \mathcal{G}_{\text{ad}}} J_{\varepsilon, \eta}(\gamma), \quad \text{avec} \quad J_{\varepsilon, \eta}(\gamma) = J(\gamma) + \frac{\varepsilon}{2} \|\gamma''\|_{L^2(0, T)}^2 + \frac{\eta}{2} |(L(\gamma) - \bar{L})^+|^2,$$

et où  $(\cdot)^+$  désigne la partie positive. On considère ici  $\mathcal{G}_{\text{ad}} = \mathcal{S}_5$ , l'espace des splines de degré 5 associées à une subdivision régulière de  $[0, T]$  fixée. On résout numériquement le problème par un algorithme de gradient à pas fixe. À partir d'une trajectoire initiale  $\gamma_0$ , on construit une suite minimisante  $\gamma_n \rightarrow \gamma^*$  définie par  $\gamma_{n+1} = \gamma_n - \rho \phi_n^{\varepsilon, \eta}$ , avec  $\rho > 0$  fixé. À chaque itération, la direction de descente  $\phi_n^{\varepsilon, \eta}$  est évaluée à l'aide de la dérivée de Gateaux de  $J$  en  $\gamma_n$ . En outre, le contrôle  $\hat{u}_{\gamma_n}$  de norme  $\mathcal{U}$  minimale est approché via la formulation mixte introduite précédemment. On présente quelques simulations pour différentes données initiales  $(y^0, y^1)$ . Comparativement au cas du contrôle interne traité à la Sect. A, le problème paraît ici plus difficile à résoudre, dans le sens où la fonctionnelle  $J_{\varepsilon, \eta}$  semble être assez “plate” au voisinage de ses points de minimum.

## Article



---

Arthur Bottois\*

# Pointwise moving control for the 1-D wave equation

Numerical approximation and optimization of the support

**Abstract:** We consider the exact null controllability of the 1-D wave equation with an interior pointwise control acting on a moving point  $(\gamma(t))_{t \in (0, T)}$ . We approximate a control of minimal norm through a mixed formulation solved by using a conformal space-time finite element method. We then introduce a gradient-type approach in order to optimize the trajectory  $\gamma$  of the control point. Several experiments are discussed.

**Keywords:** exact controllability, wave equation, pointwise control, mixed formulation, finite element approximation

## 1 Introduction

Let  $T > 0$ . We consider the linear one-dimensional wave equation in the interval  $\Omega = (0, 1)$ , with a pointwise control  $v$  acting on a moving point  $x = \gamma(t)$ ,  $t \in [0, T]$ . The state equation reads

$$\begin{cases} y_{tt} - y_{xx} = v(t)\delta_{\gamma(t)}(x) & \text{in } Q_T = \Omega \times (0, T), \\ y = 0 & \text{on } \Sigma_T = \partial\Omega \times (0, T), \\ (y, y_t)(\cdot, 0) = (y_0, y_1) & \text{in } \Omega. \end{cases} \quad (1)$$

Here,  $\delta_{\gamma(t)}$  is the Dirac measure at  $x = \gamma(t)$  and  $\gamma$  represents the trajectory in time of the control point. The curve  $\gamma : [0, T] \rightarrow \Omega$  is assumed to be piecewise  $C^1$ . We also denote by  $\mathbf{H}'$  the dual space of  $\mathbf{H} := H^1(0, T)$ . For  $v \in \mathbf{H}'$ , we refer to Section 2.1 for the well-posedness of (1).

The exact null controllability problem for (1) at time  $T > 0$  is the following. *Given a trajectory  $\gamma : [0, T] \rightarrow \Omega$ , for any initial datum  $(y_0, y_1) \in \mathbf{V} := L^2(\Omega) \times H^{-1}(\Omega)$ , find a control  $v \in \mathbf{H}'$  such that the corresponding solution  $y$  of (1)*

---

\*Corresponding author: Arthur Bottois, Université Clermont Auvergne, Laboratoire de Mathématiques Blaise Pascal CNRS-UMR 6620, Campus des Cézeaux, F-63178 Aubière cedex, France, e-mail: arthur.bottois@uca.fr

satisfies

$$(y, y_t)(\cdot, T) = (0, 0) \quad \text{in } \Omega.$$

As a consequence of the *Hilbert uniqueness method* (HUM) introduced by J.-L. Lions [25], the controllability of (1) is equivalent to an observability inequality for the associated adjoint problem. Indeed, the state equation (1) is controllable if and only if there exists a constant  $C_{\text{obs}}(\gamma) > 0$  such that

$$\|(\varphi_0, \varphi_1)\|_{\mathbf{W}}^2 \leq C_{\text{obs}}(\gamma) \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W} := H_0^1(\Omega) \times L^2(\Omega), \quad (2)$$

where  $\varphi \in C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$  solves

$$L\varphi = 0 \text{ in } Q_T, \quad \varphi = 0 \text{ on } \Sigma_T, \quad (\varphi, \varphi_t)(\cdot, 0) = (\varphi_0, \varphi_1) \text{ in } \Omega. \quad (3)$$

Here, the notation  $\varphi(\gamma, \cdot)$  stands for the function  $\varphi(\gamma(t), t)$  with  $t \in (0, T)$ , while  $L$  denotes the wave operator

$$L = \partial_t^2 - \partial_x^2.$$

Under additional assumptions on  $\gamma$ , a proof of (2) can be found in [8]. We emphasize that the observability constant  $C_{\text{obs}}(\gamma)$  depends on the control trajectory  $\gamma$ . In what follows, we say that  $\gamma$  is an *admissible trajectory* if the observability inequality (2) holds true.

In this work, we investigate the issue of the numerical approximation of the control  $\widehat{v}_\gamma$  of minimal  $\mathbf{H}'$ -norm and the associated controlled state. We also tackle the problem of optimizing the support of control, which is done numerically by minimizing the norm  $\|\widehat{v}_\gamma\|_{\mathbf{H}'}$  with respect to the trajectory  $\gamma$ .

Let us now mention some references related to pointwise control. This problem arises naturally in practical situations when the size of the control domain is very small compared to the size of the physical system. For a stationary control point  $\gamma \equiv x_0 \in \Omega$ , the controllability of (1) depends strongly on the location of  $x_0$  [24, 26, 14]. Indeed, one can show that controllability holds if and only if the controllability time  $T$  is large enough, i.e.  $T \geq 2|\Omega|$ , and if there is no eigenfunction of the Dirichlet Laplacian vanishing at  $x = x_0$ . The constraint on  $T$  is due to the finite speed of propagation of the solution of the wave equation (1). A point  $x_0$  satisfying the previous spectral property is referred to as a *strategic point*. Furthermore,  $x_0$  is a strategic point if and only if it is irrational with respect to the length of  $\Omega$ , making controllability very unstable. Consequently, controls acting on stationary points are usually difficult to implement in practice. It is often more convenient to control along curves for which the strategic point property holds a.e. in  $[0, T]$ .

For a moving control point  $x = \gamma(t)$ , several sufficient conditions to ensure controllability have been studied [20, 22, 8, 1]. In [22], the author proves the

existence of controls in  $L^2(0, T)$  acting on a point rapidly bouncing between two positions. In [8, Proposition 4.1], the author shows, using the d'Alembert formula, that the observability inequality (2) holds under some geometric restrictions on the trajectory  $\gamma$ . By duality, this implies the existence of controls in  $\mathbf{H}'$  for initial data in  $\mathbf{V}$ . The geometric requirements are related to the usual *geometric control condition* (GCC) introduced for controls acting over domains  $\omega \subset \Omega$  [3, 23]. Among the constraints given to guarantee that  $\gamma$  is admissible, there must exist two constants  $c_1, c_2 > 0$  and a finite number of subintervals  $(I_j)_{0 \leq j \leq J} \subset [0, T]$  such that, for each subinterval  $I_j$ ,  $\gamma \in C^1(I_j)$ ,  $1 - |\gamma'|$  does not change sign in  $I_j$  and  $c_1 \leq |\gamma'| \leq c_2$  in  $I_j$ . The constants appearing in the proof of the observability inequality (2) depend only on  $c_1$  and  $c_2$  (see [8, Remark 4.2]). Thus, it is possible to write a uniform observability inequality for trajectories in a suitable class, i.e. there exists  $C > 0$  such that  $C_{\text{obs}}(\gamma) \leq C$  for every  $\gamma$  in that class.

In the context of feedback stabilization, we mention [2]. For parabolic equations, we also mention [26, 21]. Finally, for the computation of pointwise controls for the Burgers equation, we refer to [4, 31].

The main contributions of this paper are the following. First, we use the HUM method to characterize the control  $\hat{v}$  of minimal  $\mathbf{H}'$ -norm, also known as the HUM control. We then turn our attention to the numerical approximation of this control and the associated controlled state. Usually (see [16, 29]), such an approximation is computed by minimizing the so-called conjugate functional  $\mathcal{J}_\gamma^* : \mathbf{W} \rightarrow \mathbb{R}$  defined by

$$\mathcal{J}_\gamma^*(\varphi_0, \varphi_1) = \frac{1}{2} \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2 - \int_{\Omega} y_0 \varphi_1 + \langle y_1, \varphi_0 \rangle_{-1,1}, \quad (4)$$

where  $\varphi$  is the solution of (3) associated with  $(\varphi_0, \varphi_1)$ , and  $\langle \cdot, \cdot \rangle_{-1,1}$  stands for the duality product in  $H_0^1(\Omega)$ . Here, instead, we notice that the unconstrained minimization of  $\mathcal{J}_\gamma^*(\varphi_0, \varphi_1)$  is equivalent to the minimization of another functional  $\tilde{\mathcal{J}}_\gamma^*(\varphi)$  (cf. (17)) over  $\varphi$  satisfying the constraint  $L\varphi = 0$ . This constraint is taken into account using a Lagrange multiplier which leads to a mixed formulation where the space and time variables are embedded. We follow the steps of [13, 9], where a similar formulation is used for controls distributed over non-cylindrical domains  $q \subset Q_T$ . It is worth mentioning that this space-time approach is well-adapted to our moving point situation, since we can achieve a good description of the trajectory  $\gamma$  embedded in a space-time mesh of  $Q_T$ . From a numerical point of view, we build a Galerkin approximation of the mixed formulation using conformal space-time finite elements. This allows to compute the optimal adjoint state  $\hat{\varphi}$ , linked to the HUM control  $\hat{v}$  by the relation (9). This also gives an approximation of the Lagrange multiplier, which turns out to be the controlled state associated with  $\hat{v}$ .



Another aspect of this work is the numerical optimization of the support of control. For a given initial datum  $(y_0, y_1) \in \mathbf{V}$ , we want to minimize the norm  $\|\widehat{v}_\gamma\|_{\mathbf{H}'}$  of the HUM control  $\widehat{v}_\gamma$  with respect to the trajectory  $\gamma$ . To do so, we consider the functional

$$J(\gamma) = \frac{1}{2} \|\widehat{v}_\gamma\|_{\mathbf{H}'}^2 \quad (5)$$

and we implement a gradient-type algorithm. In order to find a descent direction at each iteration, we establish a formula for the directional derivative of  $J$ . The values of  $J$  are computed using the approximate control arising from the mixed formulation mentioned previously. We perform several numerical experiments and compare our results with those obtained in [6] for controls distributed over non-cylindrical domains  $q \subset Q_T$ . In the simulations, the admissible set of trajectories  $\gamma$  is discretized using splines functions of degree 5.

The rest of the paper is organized in three sections. First, in Section 2, we briefly give some theoretical results. Namely, we justify the existence of weak solutions for the state equation (1), and we characterize the control of minimal  $\mathbf{H}'$ -norm using the HUM method. We also analyse the extremal problem  $\min_\gamma J(\gamma)$  (cf. (5)) and compute the directional derivative of  $J$  with respect to  $\gamma$ . In a second step, in Section 3, we present the space-time mixed formulation used to approximate the control and the controlled state. We also discuss some issues related to the discretization of that formulation. Finally, in Section 4, we give several numerical experiments. We illustrate the convergence of the approximated control as the discretization parameter goes to zero. For stationary control points  $\gamma \equiv x_0 \in \Omega$ , we illustrate the lack of controllability at non-strategic points. We also describe the gradient-type algorithm designed to optimize the support of control and discuss some results.

## 2 Some theoretical results

### 2.1 Existence of weak solutions for the state equation

The weak solution of (1) is defined by transposition (see [27]). For any  $\psi \in L^1(0, T; L^2(\Omega))$ , let  $\varphi \in C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$  be the solution of the backward adjoint equation

$$L\varphi = \psi \text{ in } Q_T, \quad \varphi = 0 \text{ on } \Sigma_T, \quad (\varphi, \varphi_t)(\cdot, T) = (0, 0) \text{ in } \Omega.$$

Multiplying (1) by  $\varphi$  and integrating by parts, we formally obtain

$$\iint_{Q_T} y\psi = \langle v, \varphi(\gamma, \cdot) \rangle_{\mathbf{H}', \mathbf{H}} - \int_{\Omega} y_0 \varphi_t(\cdot, 0) + \langle y_1, \varphi(\cdot, 0) \rangle_{-1, 1}, \quad \forall \psi \in L^1(0, T; L^2(\Omega)), \quad (6)$$

where  $\langle \cdot, \cdot \rangle_{-1, 1}$  and  $\langle \cdot, \cdot \rangle_{\mathbf{H}', \mathbf{H}}$  denote respectively the duality products in  $H_0^1(\Omega)$  and  $\mathbf{H}$ . We adopt identity (6) as the definition of the solution of (1) in the sense of transposition. One can then prove the following result (see [8, Theorem 2.1]).

**Lemma 1.** *Let  $\gamma : [0, T] \rightarrow \Omega$  be piecewise  $C^1$ . If there exists a subdivision  $(t_i)_{0 \leq i \leq m}$  of  $[0, T]$  such that, on each subinterval  $[t_{i-1}, t_i]$ ,  $\gamma$  is  $C^1$  and  $1 - |\gamma'|$  does not change sign, there exists a unique solution  $y$  to (1) in the sense of transposition. This solution has the regularity  $y \in C([0, T]; L^2(\Omega))$  and  $y_t \in L^2([0, T]; H^{-1}(\Omega))$ .*

## 2.2 Characterization of the HUM control

In order to give a characterization of the controls for (1), for any  $(\varphi_0, \varphi_1) \in \mathbf{W}$ , let  $\varphi$  be the solution of the adjoint equation (3). Multiplying (1) by  $\varphi$  and integrating by parts, we get that  $v \in \mathbf{H}'$  is a control if and only if

$$\langle v, \varphi(\gamma, \cdot) \rangle_{\mathbf{H}', \mathbf{H}} = \int_{\Omega} y_0 \varphi_1 - \langle y_1, \varphi_0 \rangle_{-1, 1}, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W}. \quad (7)$$

Then, by a straightforward application of the HUM method (see [8, Section 6]), we can readily characterize the control of minimal  $\mathbf{H}'$ -norm for (1). Let us consider the conjugate functional  $\mathcal{J}_\gamma^*$  defined in (4). If  $\gamma$  is an admissible trajectory, that is if the observability inequality (2) holds, we can see that  $\mathcal{J}_\gamma^*$  is continuous, strictly convex and coercive. Thus,  $\mathcal{J}_\gamma^*$  has a unique minimum point  $(\widehat{\varphi}_0, \widehat{\varphi}_1) \in \mathbf{W}$ , which satisfies the optimality condition

$$\langle \widehat{\varphi}(\gamma, \cdot), \varphi(\gamma, \cdot) \rangle_{\mathbf{H}} = \int_{\Omega} y_0 \varphi_1 - \langle y_1, \varphi_0 \rangle_{-1, 1}, \quad \forall (\varphi_0, \varphi_1) \in \mathbf{W}, \quad (8)$$

where  $\widehat{\varphi}$  and  $\varphi$  are the solutions of (3) associated with  $(\widehat{\varphi}_0, \widehat{\varphi}_1)$  and  $(\varphi_0, \varphi_1)$  respectively. For sufficient conditions guaranteeing that a trajectory  $\gamma$  is admissible, we refer to [8, Theorem 2.4]. Examples of such admissible trajectories can be found in Figure 3 and [8, Section 3]. In view of (7), one can then see that the control  $\widehat{v}$  of minimal  $\mathbf{H}'$ -norm for (1) has the following form.

**Lemma 2** (HUM control). *Let  $\gamma \in C^1([0, T])$  piecewise. If  $\gamma$  is an admissible trajectory, the control  $\widehat{v}$  of minimal  $\mathbf{H}'$ -norm for (1) is given by*

$$\begin{aligned} \widehat{v}(t) = & -\frac{d^2}{dt^2}\widehat{\varphi}(\gamma(t), t) + \widehat{\varphi}(\gamma(t), t) \\ & + \frac{d}{dt}\widehat{\varphi}(\gamma(t), t)\delta_T(t) - \frac{d}{dt}\widehat{\varphi}(\gamma(t), t)\delta_0(t), \quad \forall t \in (0, T), \end{aligned} \quad (9)$$

where  $\widehat{\varphi}$  is the solution of (3) associated with the minimum point  $(\widehat{\varphi}_0, \widehat{\varphi}_1)$  of  $\mathcal{J}_\gamma^*$ , and  $\delta_0, \delta_T$  denote respectively the Dirac measures at  $t = 0$  and  $t = T$ . Moreover, the norm of  $\widehat{v}$  can be computed by

$$\|\widehat{v}\|_{\mathbf{H}'}^2 = \|\widehat{\varphi}(\gamma, \cdot)\|_{\mathbf{H}}^2 = \int_0^T \varphi^2(\gamma(t), t) dt + \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt. \quad (10)$$

## 2.3 Optimization of the support of control

We focus here on the optimization of the control trajectory. More precisely, for  $(y_0, y_1) \in \mathbf{V}$  fixed, we want to minimize the norm  $\|\widehat{v}\|_{\mathbf{H}'}$  (cf. (10)) of the HUM control with respect to the curve  $\gamma$ , i.e. solve

$$\min_{\gamma \in \mathcal{G}} J(\gamma), \quad \text{where } J(\gamma) = \frac{1}{2} \int_0^T \varphi^2(\gamma(t), t) dt + \frac{1}{2} \int_0^T \left| \frac{d}{dt} \varphi(\gamma(t), t) \right|^2 dt, \quad (11)$$

and where  $\varphi$  is the solution of (3) associated with the minimum point  $(\varphi_0, \varphi_1)$  of  $\mathcal{J}_\gamma^*$ . The admissible set  $\mathcal{G}$  is composed of smooth trajectories, typically of class  $C^2([0, T])$ . We also require that the observability inequality (2) holds uniformly on  $\mathcal{G}$ , meaning that there exists  $C > 0$  such that  $C_{\text{obs}}(\gamma) \leq C$  for every  $\gamma \in \mathcal{G}$ . This property can be achieved with the hypotheses of [8, Theorem 2.4]. In Section 4, we discretize  $\mathcal{G}$  using the space  $\mathcal{S}_5$  of degree 5 splines, adapted to a fixed regular subdivision of  $[0, T]$ .

As it stands, we do not know if the extremal problem (11) is well-posed. To establish the lower semi-continuity of  $J$ , it could be possible to exploit the works [18, 19] where, in the context of the heat equation, the authors consider a shape optimization problem with respect to a curve. In the process, it might be necessary to have a more regular control, which would probably require more regular initial data  $(y_0, y_1)$  (see [15]).

Moreover, a longer trajectory  $\gamma$  allows intuitively a smaller cost of control. Consequently, to give more sense to the problem, we penalize the length  $L(\gamma)$  of the curve  $\gamma$ . Similarly, in order to avoid too fast variations of the trajectory, we

also regularize the “curvature”  $\gamma''$ . A similar strategy has been introduced and discussed in [6]. Thus, for  $\varepsilon > 0$  small enough,  $\eta > 0$  large enough and  $\bar{L} \geq T$  fixed, we consider the following regularized-penalized extremal problem

$$\min_{\gamma \in \mathcal{G}} J_{\varepsilon, \eta}(\gamma), \quad \text{where } J_{\varepsilon, \eta}(\gamma) = J(\gamma) + \frac{\varepsilon}{2} \|\gamma''\|_{L^2(0, T)}^2 + \frac{\eta}{2} \left( (L(\gamma) - \bar{L})^+ \right)^2, \quad (12)$$

and where  $(\cdot)^+$  stands for the positive part.

We solve this problem numerically in Section 4, using a gradient-type algorithm. In order to evaluate a descent direction for  $J_{\varepsilon, \eta}$  at each iteration of the algorithm, we compute the derivatives of  $J$  and  $J_{\varepsilon, \eta}$  with respect to  $\gamma$ .

**Lemma 3.** *Let  $\gamma \in C^2([0, T])$  be an admissible trajectory and let  $\bar{\gamma} \in C^2([0, T])$  be a perturbation. The directional derivative of  $J$  at  $\gamma$  in the direction  $\bar{\gamma}$ , defined by  $dJ(\gamma; \bar{\gamma}) := \lim_{\nu \rightarrow 0} \frac{J(\gamma + \nu \bar{\gamma}) - J(\gamma)}{\nu}$ , reads as follows*

$$\begin{aligned} dJ(\gamma; \bar{\gamma}) = & - \int_0^T \varphi(\gamma(t), t) \varphi_x(\gamma(t), t) \bar{\gamma}(t) dt \\ & - \int_0^T \frac{d}{dt} \varphi(\gamma(t), t) \frac{d}{dt} \left( \varphi_x(\gamma(t), t) \bar{\gamma}(t) \right) dt, \end{aligned}$$

where  $\varphi$  is the solution of (3) associated with the minimum point  $(\varphi_0, \varphi_1)$  of  $\mathcal{J}_\gamma^*$ . Similarly, the directional derivative of  $J_{\varepsilon, \eta}$  at  $\gamma$  in the direction  $\bar{\gamma}$  is given by

$$dJ_{\varepsilon, \eta}(\gamma; \bar{\gamma}) = dJ(\gamma; \bar{\gamma}) + \varepsilon \langle \gamma'', \bar{\gamma}'' \rangle_{L^2(0, T)} + \eta (L(\gamma) - \bar{L})^+ dL(\gamma; \bar{\gamma}),$$

where

$$L(\gamma) = \int_0^T \sqrt{1 + \gamma'^2} \quad \text{and} \quad dL(\gamma; \bar{\gamma}) = \int_0^T \frac{\gamma'}{\sqrt{1 + \gamma'^2}} \bar{\gamma}' dt.$$

*Proof.* We provide only a formal proof. Rigorous demonstrations of similar lemmas can be found in [30, 6], for controls distributed over domains  $q \subset Q_T$ . For any admissible trajectory  $\gamma \in C^2([0, T])$  and any perturbation  $\bar{\gamma} \in C^2([0, T])$ , we get

$$\begin{aligned} dJ(\gamma; \bar{\gamma}) = & \int_0^T \varphi(\gamma(t), t) \left( \varphi'(\gamma(t), t) + \varphi_x(\gamma(t), t) \bar{\gamma}(t) \right) dt \\ & + \int_0^T \frac{d}{dt} \varphi(\gamma(t), t) \frac{d}{dt} \left( \varphi'(\gamma(t), t) + \varphi_x(\gamma(t), t) \bar{\gamma}(t) \right) dt. \end{aligned} \quad (13)$$

Here,  $\varphi'$  denotes the derivative of  $\varphi$  with respect to  $\gamma$ . To simplify (13), we differentiate the optimality condition (8) with respect to  $\gamma$ . It gives

$$\begin{aligned} & \int_0^T \left( \varphi'(\gamma(t), t) + \varphi_x(\gamma(t), t)\bar{\gamma}(t) \right) \psi(\gamma(t), t) dt + \int_0^T \varphi(\gamma(t), t) \psi_x(\gamma(t), t) \bar{\gamma}(t) dt \\ & + \int_0^T \frac{d}{dt} \left( \varphi'(\gamma(t), t) + \varphi_x(\gamma(t), t)\bar{\gamma}(t) \right) \frac{d}{dt} \psi(\gamma(t), t) dt \\ & + \int_0^T \frac{d}{dt} \varphi(\gamma(t), t) \frac{d}{dt} \left( \psi_x(\gamma(t), t) \bar{\gamma}(t) \right) dt = 0, \quad \forall (\psi_0, \psi_1) \in \mathbf{W}, \end{aligned}$$

where  $\psi$  is the solution of (3) associated with  $(\psi_0, \psi_1)$ . Evaluating the previous expression for  $(\psi_0, \psi_1) = (\varphi_0, \varphi_1)$ , we can eliminate the derivative  $\varphi'$  from (13) and obtain the announced result.  $\square$

### 3 Mixed formulation

In this section, in order to approximate the HUM control for (1) and the associated controlled state, we present a space-time mixed formulation based on the optimality condition (8). We follow the steps of [9, Section 3.1], where a similar formulation is built for controls distributed over domains  $q \subset Q_T$ . From a numerical point of view, this space-time formulation is very appropriate for the moving point situation considered in this work. Indeed, after the discretization step, we solve the formulation using a space-time triangular mesh, which is constructed from boundary vertices placed on the border of  $Q_T$  and on the curve  $\gamma$ .

#### 3.1 Mixed formulation

We start by a lemma extending the observability inequality (2). For this, we first need to introduce the functional space

$$\Phi := \left\{ \varphi \in C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega)); \quad L\varphi \in L^2(0, T; L^2(\Omega)) \right\}.$$

**Lemma 4** (Generalized observability inequality). *Let  $\gamma \in C^1([0, T])$  piecewise. If  $\gamma$  is an admissible trajectory, there exists a constant  $\tilde{C}_{\text{obs}}(\gamma) > 0$  such that*

$$\|(\varphi, \varphi_t)(\cdot, 0)\|_{\mathbf{W}}^2 \leq \tilde{C}_{\text{obs}}(\gamma) \left( \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2 + \|L\varphi\|_{L^2(0, T; L^2(\Omega))}^2 \right), \quad \forall \varphi \in \Phi. \quad (14)$$

*Proof.* Let  $\varphi \in \Phi$ . We can decompose  $\varphi = \psi_1 + \psi_2$ , where  $\psi_1, \psi_2 \in \Phi$  solve

$$\begin{cases} L\psi_1 = 0 & \text{in } Q_T, & \psi_1 = 0 & \text{on } \Sigma_T, & (\psi_1, \psi_{1,t})(\cdot, 0) = (\varphi, \varphi_t)(\cdot, 0) & \text{in } \Omega, \\ L\psi_2 = L\varphi & \text{in } Q_T, & \psi_2 = 0 & \text{on } \Sigma_T, & (\psi_2, \psi_{2,t})(\cdot, 0) = (0, 0) & \text{in } \Omega. \end{cases}$$

From Duhamel's principle and the conservation of energy, one can show (see [8, Section 5]) the following so-called hidden regularity property for  $\psi_2$ , there exists a constant  $c(\gamma) > 0$  such that

$$\|\psi_2(\gamma, \cdot)\|_{\mathbf{H}}^2 \leq c(\gamma) \|L\varphi\|_{L^2(0,T;L^2(\Omega))}^2. \quad (15)$$

Combining (2) for  $\psi_1$  and (15) for  $\psi_2$ , we obtain

$$\begin{aligned} \|(\varphi, \varphi_t)(\cdot, 0)\|_{\mathbf{W}}^2 &= \|(\psi_1, \psi_{1,t})(\cdot, 0)\|_{\mathbf{W}}^2 \leq C_{\text{obs}}(\gamma) \|\psi_1(\gamma, \cdot)\|_{\mathbf{H}}^2 \\ &\leq 2C_{\text{obs}}(\gamma) \left( \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2 + \|\psi_2(\gamma, \cdot)\|_{\mathbf{H}}^2 \right) \\ &\leq \tilde{C}_{\text{obs}}(\gamma) \left( \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2 + \|L\varphi\|_{L^2(0,T;L^2(\Omega))}^2 \right). \end{aligned}$$

□

As for (2), it is possible to find a class of admissible trajectories  $\gamma$  such that the generalized observability inequality (14) holds uniformly (see [8, Theorem 2.4]), i.e. there exists  $\tilde{C} > 0$  such that  $\tilde{C}_{\text{obs}}(\gamma) \leq \tilde{C}$  for every  $\gamma$  in that class. In addition, the inequality (14) implies the following property on the space  $\Phi$ .

**Lemma 5.** *Let  $\gamma \in C^1([0, T])$  piecewise. If  $\gamma$  is an admissible trajectory, the space  $\Phi$  is a Hilbert space with the inner product*

$$\langle \varphi, \bar{\varphi} \rangle_{\Phi} = \langle \varphi(\gamma, \cdot), \bar{\varphi}(\gamma, \cdot) \rangle_{\mathbf{H}} + \tau \langle L\varphi, L\bar{\varphi} \rangle_{L^2(0,T;L^2(\Omega))}, \quad \forall \varphi, \bar{\varphi} \in \Phi, \quad (16)$$

for  $\tau > 0$  fixed.

*Proof.* The semi-norm  $\|\cdot\|_{\Phi}$  associated with the inner product is trivially a norm in view of the generalized observability inequality (14). It remains to prove that  $\Phi$  is complete with respect to this norm. Let  $(\varphi_k)_{k \geq 1} \subset \Phi$  be a Cauchy sequence for the norm  $\|\cdot\|_{\Phi}$ . So, there exists  $f \in L^2(0, T; L^2(\Omega))$  such that  $L\varphi_k \rightarrow f$  in  $L^2(0, T; L^2(\Omega))$ . As a consequence of (14), there also exists  $(\varphi_0, \varphi_1) \in \mathbf{W}$  such that  $(\varphi_k, \varphi_{k,t})(\cdot, 0) \rightarrow (\varphi_0, \varphi_1)$  in  $\mathbf{W}$ . Therefore,  $(\varphi_k)_{k \geq 1}$  can be considered as a sequence of solutions of the wave equation with convergent initial data and convergent right-hand sides. By the continuous dependence of the solution of the wave equation on the data,  $\varphi_k \rightarrow \varphi$  in  $C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$ , where  $\varphi$  is the solution of the wave equation with initial datum  $(\varphi_0, \varphi_1) \in \mathbf{W}$  and right-hand side  $f \in L^2(0, T; L^2(\Omega))$ . Thus,  $\varphi \in \Phi$ . □

We can now turn to the set-up of the mixed formulation. In order to avoid the minimization of the conjugate functional  $\mathcal{J}_\gamma^*$  (cf. (4)) with respect to  $(\varphi_0, \varphi_1)$ , we remark that the solution  $\varphi$  of (3) is completely and uniquely determined by the initial datum  $(\varphi_0, \varphi_1)$ . Then, the main idea of the reformulation is to keep  $\varphi$  as main variable and consider instead the minimization of

$$\tilde{\mathcal{J}}_\gamma^*(\varphi) = \frac{1}{2} \|\varphi(\gamma, \cdot)\|_{\mathbf{H}}^2 - \int_{\Omega} y_0 \varphi_t(\cdot, 0) + \langle y_1, \varphi(\cdot, 0) \rangle_{-1,1} \quad (17)$$

over

$$\Phi_0 := \left\{ \varphi \in \Phi; \quad L\varphi = 0 \in L^2(0, T; L^2(\Omega)) \right\}.$$

Indeed, we clearly have

$$\min_{(\varphi_0, \varphi_1) \in \mathbf{W}} \mathcal{J}_\gamma^*(\varphi_0, \varphi_1) = \mathcal{J}_\gamma^*(\hat{\varphi}_0, \hat{\varphi}_1) = \tilde{\mathcal{J}}_\gamma^*(\hat{\varphi}) = \min_{\varphi \in \Phi_0} \tilde{\mathcal{J}}_\gamma^*(\varphi),$$

where  $\hat{\varphi}$  is the solution of (3) associated with the minimum point  $(\hat{\varphi}_0, \hat{\varphi}_1)$  of  $\mathcal{J}_\gamma^*$ . Besides, the minimum point  $\hat{\varphi}$  of  $\tilde{\mathcal{J}}_\gamma^*$  is unique. So, the new variable is the function  $\varphi$  with the constraint  $L\varphi = 0$  in  $L^2(0, T; L^2(\Omega))$ . To deal with this constraint, we introduce a Lagrange multiplier  $\lambda \in \Lambda := L^2(0, T; L^2(\Omega))$ . We thus consider the following problem: find  $(\varphi, \lambda) \in \Phi \times \Lambda$  solution of

$$\begin{cases} a(\varphi, \bar{\varphi}) - b(\bar{\varphi}, \lambda) = \ell(\bar{\varphi}), & \forall \bar{\varphi} \in \Phi, \\ b(\varphi, \bar{\lambda}) = 0, & \forall \bar{\lambda} \in \Lambda, \end{cases} \quad (18)$$

where we have set

$$\begin{aligned} a : \Phi \times \Phi &\rightarrow \mathbb{R}, & a(\varphi, \bar{\varphi}) &= \langle \varphi(\gamma, \cdot), \bar{\varphi}(\gamma, \cdot) \rangle_{\mathbf{H}}, \\ b : \Phi \times \Lambda &\rightarrow \mathbb{R}, & b(\varphi, \lambda) &= \langle L\varphi, \lambda \rangle_{L^2(0, T; L^2(\Omega))}, \\ \ell : \Phi &\rightarrow \mathbb{R}, & \ell(\varphi) &= \int_{\Omega} y_0 \varphi_t(\cdot, 0) - \langle y_1, \varphi(\cdot, 0) \rangle_{-1,1}. \end{aligned}$$

The introduction of this problem is justified by the result below.

**Theorem 1** (Mixed formulation). *Let  $\gamma \in C^1([0, T])$  piecewise. If  $\gamma$  is an admissible trajectory, we have the following properties;*

- *The mixed formulation (18) is well-posed.*
- *The unique solution  $(\varphi, \lambda) \in \Phi \times \Lambda$  is the unique saddle point of the Lagrangian  $\mathcal{L} : \Phi \times \Lambda \rightarrow \mathbb{R}$  defined by*

$$\mathcal{L}(\varphi, \lambda) = \frac{1}{2} a(\varphi, \varphi) - b(\varphi, \lambda) - \ell(\varphi).$$

- *The optimal function  $\varphi$  is the minimum point of  $\tilde{\mathcal{J}}_\gamma^*$  over  $\Phi_0$ . Besides, the optimal function  $\lambda \in \Lambda$  is the solution of the controlled wave equation (1), with the control  $v$  associated with  $\varphi$  (cf. (9)).*

*Proof.* We easily check that the bilinear form  $a$  is continuous over  $\Phi \times \Phi$ , symmetric and positive. Similarly, we check that the bilinear form  $b$  is continuous over  $\Phi \times \Lambda$ . Furthermore, the continuity of the linear form  $\ell$  over  $\Phi$  is a direct consequence of the generalized observability inequality (14),

$$|\ell(\varphi)| \leq \|(y_0, y_1)\|_{\mathbf{V}} \sqrt{2\tilde{C}_{\text{obs}}(\gamma) \max(1, \tau^{-1})} \|\varphi\|_{\Phi}, \quad \forall \varphi \in \Phi.$$

Therefore, to prove the well-posedness of the mixed formulation (18), we only need to check the following two properties (see [7]).

- The form  $a$  is coercive on the kernel  $\mathcal{N}(b) := \{\varphi \in \Phi; b(\varphi, \lambda) = 0, \forall \lambda \in \Lambda\}$ .
- The form  $b$  satisfies the usual “inf-sup” condition over  $\Phi \times \Lambda$ , i.e. there exists a constant  $\delta > 0$  such that

$$\inf_{\lambda \in \Lambda} \sup_{\varphi \in \Phi} \frac{b(\varphi, \lambda)}{\|\varphi\|_{\Phi} \|\lambda\|_{\Lambda}} \geq \delta. \quad (19)$$

From the definition of  $a$ , the first point is clear. Indeed, for any  $\varphi \in \mathcal{N}(b) = \Phi_0$ ,  $a(\varphi, \varphi) = \|\varphi\|_{\Phi}^2$ . We now check the inf-sup condition (19). For any  $\lambda_0 \in \Lambda$ , we define the unique element  $\varphi_0 \in \Phi$  such that

$$L\varphi_0 = \lambda_0 \text{ in } Q_T, \quad \varphi_0 = 0 \text{ on } \Sigma_T, \quad (\varphi_0, \varphi_{0,t})(\cdot, 0) = (0, 0) \text{ in } \Omega.$$

It implies  $b(\varphi_0, \lambda_0) = \|\lambda_0\|_{\Lambda}^2$  and

$$\sup_{\varphi \in \Phi} \frac{b(\varphi, \lambda_0)}{\|\varphi\|_{\Phi} \|\lambda_0\|_{\Lambda}} \geq \frac{b(\varphi_0, \lambda_0)}{\|\varphi_0\|_{\Phi} \|\lambda_0\|_{\Lambda}} = \frac{\|\lambda_0\|_{\Lambda}}{\sqrt{\|\varphi_0(\gamma, \cdot)\|_{\mathbf{H}}^2 + \tau \|\lambda_0\|_{\Lambda}^2}}.$$

We then use the following estimate (see [8, Section 5]), there exists a constant  $c(\gamma) > 0$  such that

$$\|\varphi_0(\gamma, \cdot)\|_{\mathbf{H}}^2 \leq c(\gamma) \|\lambda_0\|_{\Lambda}^2.$$

Combining the two previous inequalities, we obtain

$$\sup_{\varphi \in \Phi} \frac{b(\varphi, \lambda_0)}{\|\varphi\|_{\Phi} \|\lambda_0\|_{\Lambda}} \geq \frac{1}{\sqrt{c(\gamma) + \tau}}, \quad \forall \lambda_0 \in \Lambda.$$

Hence, the inequality (19) holds with  $\delta = (c(\gamma) + \tau)^{-\frac{1}{2}}$ .

The second point of the theorem is due to the symmetry and positivity of the bilinear form  $a$ . Regarding the third point, the equality  $b(\varphi, \bar{\lambda}) = 0$  for all  $\bar{\lambda} \in \Lambda$  implies that  $L\varphi = 0$  in  $L^2(0, T; L^2(\Omega))$ . Besides, for  $\bar{\varphi} \in \Phi_0$ , the first equation of (18) gives  $a(\varphi, \bar{\varphi}) = \ell(\bar{\varphi})$ . So, if  $(\varphi, \lambda) \in \Phi \times \Lambda$  solves the mixed formulation, then  $\varphi \in \Phi_0$  and  $\mathcal{L}(\varphi, \lambda) = \tilde{\mathcal{J}}_{\gamma}^*(\varphi)$ . Moreover, again due to the symmetry and positivity



of  $a$ , the function  $\varphi$  is the minimum point of  $\tilde{\mathcal{J}}_\gamma^*$  over  $\Phi_0$ . Indeed, for any  $\bar{\varphi} \in \Phi_0$ , we have

$$\tilde{\mathcal{J}}_\gamma^*(\varphi) = -\frac{1}{2}a(\varphi, \varphi) \leq \frac{1}{2}a(\bar{\varphi}, \bar{\varphi}) - a(\varphi, \bar{\varphi}) = \frac{1}{2}a(\bar{\varphi}, \bar{\varphi}) - \ell(\bar{\varphi}) = \tilde{\mathcal{J}}_\gamma^*(\bar{\varphi}).$$

Finally, the first equation of (18) reads

$$\langle \varphi(\gamma, \cdot), \bar{\varphi}(\gamma, \cdot) \rangle_{\mathbf{H}} - \langle L\bar{\varphi}, \lambda \rangle_{\Lambda} = \int_{\Omega} y_0 \bar{\varphi}_t(\cdot, 0) - \langle y_1, \bar{\varphi}(\cdot, 0) \rangle_{-1,1}, \quad \forall \bar{\varphi} \in \Phi.$$

Since the control  $v$  of minimal  $\mathbf{H}'$ -norm is given by (9), we get

$$\iint_{Q_T} \lambda L\bar{\varphi} = \langle v, \bar{\varphi}(\gamma, \cdot) \rangle_{\mathbf{H}', \mathbf{H}} - \int_{\Omega} y_0 \bar{\varphi}_t(\cdot, 0) + \langle y_1, \bar{\varphi}(\cdot, 0) \rangle_{-1,1}, \quad \forall \bar{\varphi} \in \Phi.$$

But this means that  $\lambda$  is solution in a weak sense of the wave equation (1) associated with the initial datum  $(y_0, y_1) \in \mathbf{V}$  and the control  $v \in \mathbf{H}'$ .  $\square$

Consequently, the search of the HUM control for (1) is reduced to the resolution of the mixed formulation (18), or equivalently to the search of the saddle point of  $\mathcal{L}$ . Moreover, for numerical purposes, it is convenient to “augment” the Lagrangian  $\mathcal{L}$  and to consider instead the Lagrangian  $\mathcal{L}_r$  defined, for any  $r > 0$ , by

$$\begin{cases} \mathcal{L}_r(\varphi, \lambda) = \frac{1}{2}a_r(\varphi, \varphi) - b(\varphi, \lambda) - \ell(\varphi), \\ a_r(\varphi, \bar{\varphi}) = a(\varphi, \bar{\varphi}) + r \langle L\varphi, L\bar{\varphi} \rangle_{L^2(0,T;L^2(\Omega))}. \end{cases}$$

Since  $a(\varphi, \varphi) = a_r(\varphi, \varphi)$  for  $\varphi \in \Phi_0$ , the Lagrangians  $\mathcal{L}$  and  $\mathcal{L}_r$  share the same saddle point.

## 3.2 Discretization

We now turn to the discretization of the mixed formulation (18). Let  $(\Phi_h)_{h>0} \subset \Phi$  and  $(\Lambda_h)_{h>0} \subset \Lambda$  be two families of finite-dimensional spaces. For any  $h > 0$ , we introduce the following approximated problem: find  $(\varphi_h, \lambda_h) \in \Phi_h \times \Lambda_h$  solution of

$$\begin{cases} a_r(\varphi_h, \bar{\varphi}_h) - b(\bar{\varphi}_h, \lambda_h) = \ell(\bar{\varphi}_h), & \forall \bar{\varphi}_h \in \Phi_h, \\ b(\varphi_h, \bar{\lambda}_h) = 0, & \forall \bar{\lambda}_h \in \Lambda_h. \end{cases} \quad (20)$$

To prove the well-posedness of this mixed formulation, we again have to check the following two properties. First, the bilinear form  $a_r$  is coercive on the kernel  $\mathcal{N}_h(b) := \{\varphi_h \in \Phi_h; b(\varphi_h, \lambda_h) = 0, \forall \lambda_h \in \Lambda_h\}$ . Actually, from the relation

$$a_r(\varphi, \varphi) \geq \min(1, r/\tau) \|\varphi\|_{\Phi}^2, \quad \forall \varphi \in \Phi,$$

the form  $a_r$  is coercive on the full space  $\Phi$ , and so a fortiori on  $\mathcal{N}_h(b) \subset \Phi_h \subset \Phi$ . The second property is a discrete inf-sup condition, there exists a constant  $\delta_h > 0$  such that

$$\inf_{\lambda_h \in \Lambda_h} \sup_{\varphi_h \in \Phi_h} \frac{b(\varphi_h, \lambda_h)}{\|\varphi_h\|_{\Phi_h} \|\lambda_h\|_{\Lambda_h}} \geq \delta_h. \quad (21)$$

The spaces  $\Phi_h$  and  $\Lambda_h$  are finite-dimensional, so the infimum and the supremum in (21) are reached. Moreover, from the properties of  $a_r$  and with the finite element spaces  $\Phi_h, \Lambda_h$  chosen below, it is standard to prove that  $\delta_h$  is strictly positive. Consequently, for any  $h > 0$ , there exists a unique couple  $(\varphi_h, \lambda_h) \in \Phi_h \times \Lambda_h$  solution of the discrete mixed formulation (21).

On the other hand, if we could show that  $\inf_{h>0} \delta_h > 0$ , it would ensure the convergence of the solution  $(\varphi_h, \lambda_h)$  of the discrete formulation (20) towards the solution  $(\varphi, \lambda)$  of the continuous formulation (18). However, this property is usually difficult to prove and depends strongly on the choice made for the spaces  $\Phi_h, \Lambda_h$ . We analyse numerically this property in Section 3.3.

Let us consider a triangulation  $\mathcal{T}_h$  of  $Q_T$ , i.e.  $\cup_{K \in \mathcal{T}_h} K = \overline{Q_T}$ . We denote  $h := \max\{\text{diam}(K); K \in \mathcal{T}_h\}$ , where  $\text{diam}(K)$  is the diameter of the triangle  $K$ . In what follows, the space-time mesh  $\mathcal{T}_h$  is built from a discretization of the border of  $Q_T$  and the curve  $\gamma$  (see Figure 5). Thus, the fineness of  $\mathcal{T}_h$  will be given either by  $h$  or by the number  $N_{\mathcal{T}}$  of vertices per unit of length. This also means that some vertices are supported on  $\gamma$ , making the mesh well-adapted to the control trajectory. The mesh is generated using the software FreeFEM++ (see [17]).

The finite-dimensional space  $\Phi_h$  must be chosen such that  $L\varphi_h$  belongs to  $L^2(0, T; L^2(\Omega))$ , for any  $\varphi_h \in \Phi_h$ . Therefore, any space of functions continuously differentiable with respect to both  $x$  and  $t$  is a conformal approximation of  $\Phi$ . We define the space  $\Phi_h$  as follows

$$\Phi_h := \left\{ \varphi_h \in C^1(\overline{Q_T}); \quad \varphi_h|_K \in \mathbb{P}(K), \quad \forall K \in \mathcal{T}_h, \quad \varphi_h = 0 \text{ on } \Sigma_T \right\} \subset \Phi,$$

where  $\mathbb{P}(K)$  stands for the complete *Hsieh-Clough-Tocher* finite element (HCT for short) of class  $C^1$ . It is a so-called composite finite element. It involves 12 degrees of freedom which are, for each triangle  $K$ , the values of  $\varphi_h, \varphi_{h,x}, \varphi_{h,t}$  on the three vertices, and the values of the normal derivative of  $\varphi$  in the middle of the three edges. We refer to [12] and [5, 28] for the precise definition and the implementation of such finite element. We also introduce the finite-dimensional space

$$\Lambda_h := \left\{ \lambda_h \in C^0(\overline{Q_T}); \quad \lambda_h|_K \in \mathbb{Q}(K), \quad \forall K \in \mathcal{T}_h, \quad \lambda_h = 0 \text{ on } \Sigma_T \right\} \subset \Lambda,$$

where  $\mathbb{Q}(K)$  is the space of affine functions in both  $x$  and  $t$  on the element  $K$ .

Let  $n_h = \dim(\mathbf{\Phi}_h)$  and  $m_h = \dim(\mathbf{\Lambda}_h)$ . We define the matrices  $A_{r,h} \in \mathbb{R}^{n_h, n_h}$ ,  $B_h \in \mathbb{R}^{m_h, n_h}$ ,  $M_h \in \mathbb{R}^{m_h, m_h}$  and the vector  $L_h \in \mathbb{R}^{n_h}$  by

$$\begin{aligned} \langle A_{r,h}\{\varphi_h\}, \{\bar{\varphi}_h\} \rangle &= a_r(\varphi_h, \bar{\varphi}_h), & \forall \varphi_h, \bar{\varphi}_h \in \mathbf{\Phi}_h, \\ \langle B_h\{\varphi_h\}, \{\lambda_h\} \rangle &= b(\varphi_h, \lambda_h), & \forall \varphi_h \in \mathbf{\Phi}_h, \forall \lambda_h \in \mathbf{\Lambda}_h, \\ \langle M_h\{\lambda_h\}, \{\bar{\lambda}_h\} \rangle &= \langle \lambda_h, \bar{\lambda}_h \rangle_{\mathbf{\Lambda}}, & \forall \lambda_h, \bar{\lambda}_h \in \mathbf{\Lambda}_h, \\ \langle L_h, \{\varphi_h\} \rangle &= \ell(\varphi_h), & \forall \varphi_h \in \mathbf{\Phi}_h, \end{aligned}$$

where  $\{\varphi_h\} \in \mathbb{R}^{n_h}$  and  $\{\lambda_h\} \in \mathbb{R}^{m_h}$  denote the vectors associated with  $\varphi_h \in \mathbf{\Phi}_h$  and  $\lambda_h \in \mathbf{\Lambda}_h$  respectively. With these notations, the discrete mixed formulation (20) reads as follows: find  $\{\varphi_h\} \in \mathbb{R}^{n_h}$  and  $\{\lambda_h\} \in \mathbb{R}^{m_h}$  such that

$$\begin{pmatrix} A_{r,h} & -B_h^T \\ -B_h & 0 \end{pmatrix} \begin{pmatrix} \{\varphi_h\} \\ \{\lambda_h\} \end{pmatrix} = \begin{pmatrix} L_h \\ 0 \end{pmatrix}. \quad (22)$$

For any  $r > 0$ , the matrix  $A_{r,h}$  is symmetric and positive definite. However, the matrix in (22) is symmetric but not positive definite. The system (22) is solved by the LU method with FreeFEM++ (see [17]).

### 3.3 Discrete inf-sup test

Here, we test numerically the discrete inf-sup condition (21), and more precisely the property  $\inf_{h>0} \delta_h > 0$ . For simplicity, we take  $\tau = r > 0$  in (16), so that  $a_{r,h}(\varphi, \bar{\varphi}) = \langle \varphi, \bar{\varphi} \rangle_{\mathbf{\Phi}}$  for all  $\varphi, \bar{\varphi} \in \mathbf{\Phi}$ . It is readily seen (see [10]) that the discrete inf-sup constant satisfies

$$\delta_h = \inf \left\{ \sqrt{\mu}; \quad B_h A_{r,h}^{-1} B_h^T \{\lambda_h\} = \mu M_h \{\lambda_h\}, \quad \forall \{\lambda_h\} \in \mathbb{R}^{m_h} \setminus \{0\} \right\}. \quad (23)$$

For any  $h > 0$ , the matrix  $B_h A_{r,h}^{-1} B_h^T$  is symmetric and positive definite, so the constant  $\delta_h$  is strictly positive. The generalized eigenvalue problem (23) is solved by the inverse power method (see [11]). Given  $\{u_h^0\} \in \mathbb{R}^{m_h}$  such that  $\|\{u_h^0\}\|_2 = 1$ , for any  $n \in \mathbb{N}$ , compute iteratively  $(\{\varphi_h^n\}, \{\lambda_h^n\}) \in \mathbb{R}^{n_h} \times \mathbb{R}^{m_h}$  and  $\{u_h^{n+1}\} \in \mathbb{R}^{m_h}$  as follows

$$\begin{cases} A_{r,h}\{\varphi_h^n\} - B_h^T\{\lambda_h^n\} = 0, \\ B_h\{\varphi_h^n\} = M_h\{u_h^n\}, \end{cases} \quad \{u_h^{n+1}\} = \frac{\{\lambda_h^n\}}{\|\{\lambda_h^n\}\|_2}.$$

The discrete inf-sup constant  $\delta_h$  is then given by  $\delta_h = \lim_{n \rightarrow \infty} \|\{\lambda_h^n\}\|_2^{-\frac{1}{2}}$ .

We now compute  $\delta_h$  for decreasing values of the fineness  $h$ , and for different values of the parameter  $r$ , namely  $r = 10^{-2}$ ,  $r = h$  and  $r = h^2$ . We use the control

trajectory  $\gamma$  defined in (Ex1- $\gamma$ ). The values that we obtain are collected in Table 1. In view of the results for  $r = 10^{-2}$ , the constant  $\delta_h$  does not seem to be uniformly bounded by below as  $h \rightarrow 0$ . Thus, we may conclude that the finite elements used here do not “pass” the discrete inf-sup test. As we shall see in the next section, this fact does not prevent the convergence of the sequences  $(\varphi_h)_{h>0}$  and  $(\lambda_h)_{h>0}$ , at least for the cases we have considered. Interestingly, we also observe that  $\delta_h$  remains bounded by below with respect to  $h$  when  $r$  depends appropriately on  $h$ , as for instance in the case  $r = h^2$ .

$h (\times 10^{-2})$	6.46	3.51	2.66	2.17	1.37	1.21
$r = 10^{-2}$	1.8230	1.7947	1.7845	1.6749	1.6060	1.5008
$r = h$	1.4575	1.3806	1.3269	1.2402	1.4188	1.3851
$r = h^2$	1.8873	1.8885	1.8783	1.8697	1.8982	1.8920

**Tab. 1:** Discrete inf-sup constant  $\delta_h$  w.r.t.  $h$  and  $r$ , for  $\gamma$  defined in (Ex1- $\gamma$ ).

## 4 Numerical simulations

In this section, we solve on various examples the discrete mixed formulation (20) to compute the HUM control for (1) and the associated controlled state. First, we determine the rate of convergence of the approximated control/controlled state, as the discretization parameter  $h$  goes to zero. Second, for stationary control points  $\gamma \equiv x_0$ , we illustrate the blow-up of the cost of control at non-strategic points. Finally, we introduce a gradient-type algorithm to solve the problem (12) of optimizing the support of control. The algorithm is then tested on two different initial data. From now on, we set  $T = 2$  and  $r = 10^{-2}$ .

### 4.1 Convergence of the approximated control

In order to measure the rate of convergence of the approximated control with respect to the mesh fineness  $h$ , we use the initial datum

$$y_0(x) = \sin(\pi x), \quad y_1(x) = 0, \quad \forall x \in \Omega, \quad (\text{Ex1-}y_0)$$

and the control trajectory

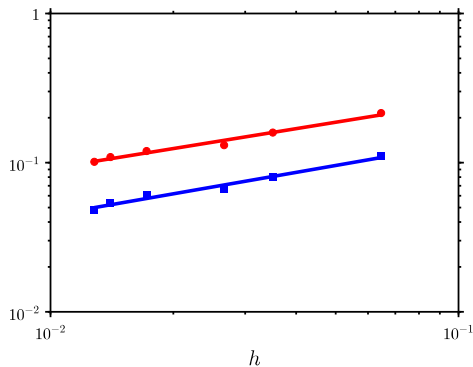
$$\gamma(t) = \frac{1}{5} + \frac{3}{5} \frac{t}{T}, \quad \forall t \in [0, T]. \quad (\text{Ex1-}\gamma)$$

This curve  $\gamma$  is an admissible trajectory (see [8, Example 3.2]), i.e. the system (1) is controllable. To compare with the approximated solution  $(\varphi_h, \lambda_h)$  of (20),

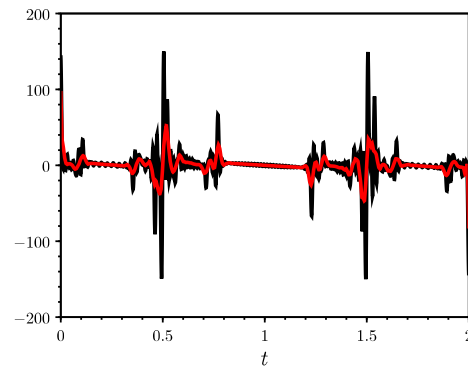
using the optimality condition (8), we compute another approximation  $(\varphi, \lambda)$  by Fourier expansion (see Appendix A), with  $N_F = 100$  harmonics. We then evaluate the errors  $\|\varphi(\gamma, \cdot) - \varphi_h(\gamma, \cdot)\|_{\mathbf{H}}$  and  $\|\lambda - \lambda_h\|_{\mathbf{\Lambda}}$  for the six levels of fineness  $N_{\mathcal{T}} = 25, 50, 75, 100, 125, 150$ . We gather the results in Table 2 and display them in Figure 1. By linear regression, we find a convergence rate in  $h^{0.44}$  for  $\varphi_h$  and in  $h^{0.48}$  for  $\lambda_h$ . In Figure 3, we represent the adjoint state  $\varphi_h$  and the controlled state  $\lambda_h$ , for  $N_{\mathcal{T}} = 150$ . The HUM control  $v_h$  computed from  $\varphi_h$  by (9) is shown in Figure 2, together with the “exact” control  $v$  obtained by Fourier expansion.

$N_{\mathcal{T}}$	25	50	75	100	125	150
$h (\times 10^{-2})$	6.46	3.51	2.66	1.72	1.40	1.28
$\ \varphi(\gamma, \cdot) - \varphi_h(\gamma, \cdot)\ _{\mathbf{H}} (\times 10^{-1})$	2.15	1.59	1.31	1.20	1.09	1.01
$\ \lambda - \lambda_h\ _{\mathbf{\Lambda}} (\times 10^{-2})$	11.0	8.06	6.69	6.05	5.38	4.81

**Tab. 2: (Ex1)** – Error on the approximated solution  $(\varphi_h, \lambda_h)$  of (20) w.r.t.  $h$ .



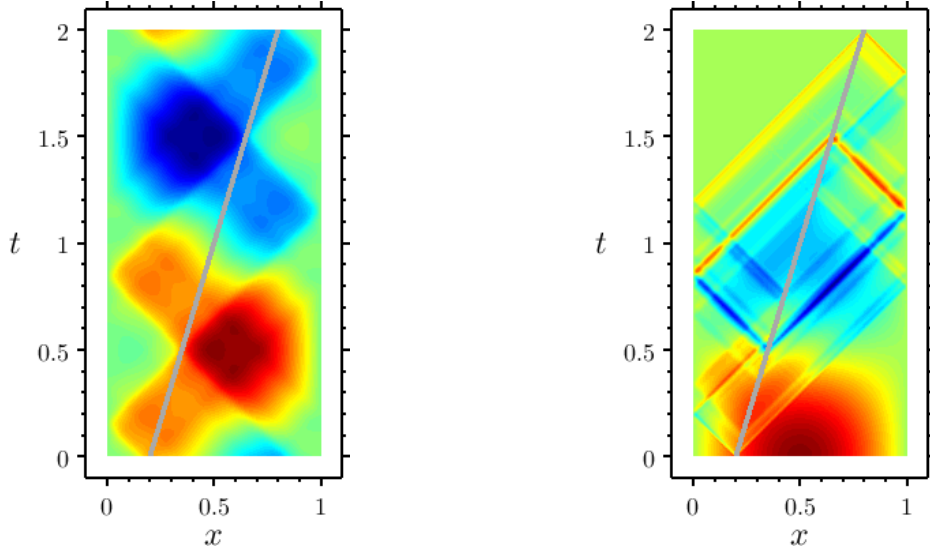
**Fig. 1: (Ex1)** – Error on the approximated solution  $(\varphi_h, \lambda_h)$  of (20) vs.  $h$  –  $\|\varphi(\gamma, \cdot) - \varphi_h(\gamma, \cdot)\|_{\mathbf{H}}$  (•),  $\|\lambda - \lambda_h\|_{\mathbf{\Lambda}}$  (■).



**Fig. 2: (Ex1)** – Controls  $v_h$  (–) and  $v$  (–), for  $N_{\mathcal{T}} = 150$ .

## 4.2 Blow-up at non-strategic points

In the case of a stationary control point  $\gamma \equiv x_0 \in \Omega$ , it is well-known that one has to choose a so-called strategic point (see [26]) to ensure the controllability of (1). A point  $x_0$  is strategic if and only if  $\sin(p\pi x_0) \neq 0$  for every  $p \geq 1$ . Moreover, a given initial datum  $(y_0, y_1) \in \mathbf{V}$  can be controlled if and only if  $\sin(p\pi x_0) \neq 0$  for every  $p \geq 1$  such that one of the Fourier coefficients  $c_p(y_0), c_p(y_1)$  are non-zero.



**Fig. 3: (Ex1)** – Iso-values of the adjoint state  $\varphi_h$  (left) and controlled state  $\lambda_h$  (right), for  $N_{\mathcal{T}} = 150$ .

Therefore, for  $(y_0, y_1) \in \mathbf{V}$  fixed, we expect the cost of control to blow up as  $x_0$  gets closer to a non-strategic location. To illustrate this property, we use the initial datum

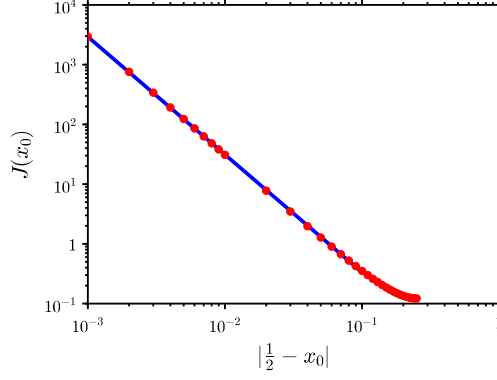
$$y_0(x) = \sin(2\pi x), \quad y_1(x) = 0, \quad \forall x \in \Omega, \quad (\mathbf{Ex2}\text{-}y_0)$$

and we evaluate the functional  $J(x_0)$  (cf. (11)) for several control locations  $x_0$  spread in the interval  $(\frac{1}{4}, \frac{1}{2})$ . With the initial datum considered,  $x^* = \frac{1}{2}$  is the unique non-strategic point. In Figure 4, we display  $J(x_0)$  w.r.t. the distance  $|x^* - x_0|$ . As expected, we note that the cost of control blows up when  $x_0 \rightarrow x^*$ . More precisely, we have  $J(x_0) \sim_{x^*} C_0 |x^* - x_0|^{-1.97}$ .

### 4.3 Optimization of the support using splines

We now focus on solving numerically the problem (12) with a gradient-type algorithm. To do so, the control trajectories  $\gamma$  considered are degree 5 splines adapted to a fixed subdivision of  $[0, T]$ . For any integer  $N \geq 1$ , we denote  $S_N = (t_i)_{0 \leq i \leq N}$  the regular subdivision of  $[0, T]$  in  $N$  intervals. With  $\kappa = T/N$ , the subdivision points are  $t_i = i\kappa$ . In the simulations below, we use  $N = 20$ . We then define the set  $\mathcal{S}_5$  of degree 5 splines adapted to the subdivision  $S_N$ . Such a spline  $\gamma \in \mathcal{S}_5$  is of class  $C^2([0, T])$  and is uniquely determined by the  $3(N + 1)$  conditions

$$\gamma(t_i) = x_i, \quad \gamma'(t_i) = p_i, \quad \gamma''(t_i) = c_i, \quad 0 \leq i \leq N,$$



**Fig. 4: (Ex2)** –  $J(x_0)$  vs.  $|x^* - x_0|$ , for stationary control points  $x_0$ .

where  $\mathbf{x} = (x_i)_{0 \leq i \leq N}$ ,  $\mathbf{p} = (p_i)_{0 \leq i \leq N}$  and  $\mathbf{c} = (c_i)_{0 \leq i \leq N}$  represent the spline parameters. We also introduce the degree 5 polynomial basis  $(P_{k,l})_{\substack{k=0,1,2 \\ l=0,1}}$  on  $[0, 1]$  characterized by

$$P_{k,l}^{(k')}(l') = \delta_{k,k'} \delta_{l,l'}, \quad \text{for } k, k' \in \{0, 1, 2\}, \quad l, l' \in \{0, 1\}.$$

Here,  $P_{k,l}^{(k')}$  stands for the  $k'$ -th derivative of  $P_{k,l}$  and  $\delta_{k,k'}$  is the Kronecker delta, i.e.  $\delta_{k,k'} = 1$  if  $k = k'$  and  $\delta_{k,k'} = 0$  otherwise. For the sake of presentation, we briefly rename the parameters  $(\mathbf{x}, \mathbf{p}, \mathbf{c}) = (\mathbf{s}^0, \mathbf{s}^1, \mathbf{s}^2)$ . It allows to decompose  $\gamma$  into

$$\gamma(t) = \sum_{i=1}^N \sum_{k=0}^2 \left( \mathbf{s}_{i-1}^k P_{k,0}^i(t) + \mathbf{s}_i^k P_{k,1}^i(t) \right) \mathbb{1}_{[t_{i-1}, t_i]}(t), \quad \forall t \in [0, T],$$

where we have set  $P_{k,l}^i(t) = \kappa^k P_{k,l} \left( \frac{t-t_{i-1}}{\kappa} \right)$ . With this decomposition, the optimization problem (12) is reduced to a finite-dimensional problem in the space of parameters, i.e.

$$\min_{\gamma \in \mathcal{S}_5} J_{\varepsilon, \eta}(\gamma) = \min_{\mathbf{s}} \tilde{J}_{\varepsilon, \eta}(\mathbf{s}), \quad \text{where } \mathbf{s} = (\mathbf{x}, \mathbf{p}, \mathbf{c}) \in \mathbb{R}^{3(N+1)}.$$

In order to get a descent direction for  $J_{\varepsilon, \eta}$  at  $\gamma \in \mathcal{S}_5$ , we consider the following variational problem: find  $j_\gamma \in \mathcal{S}_5$  solution of

$$\begin{aligned} \langle j_\gamma, \bar{\gamma} \rangle_{\mathbf{H}} + \varepsilon \langle j_\gamma'', \bar{\gamma}'' \rangle_{L^2(0, T)} &= \mathrm{d}J(\gamma; \bar{\gamma}) + \varepsilon \langle \gamma'', \bar{\gamma}'' \rangle_{L^2(0, T)} \\ &+ \eta (L(\gamma) - \bar{L})^+ \mathrm{d}L(\gamma; \bar{\gamma}), \quad \forall \bar{\gamma} \in \mathcal{S}_5. \end{aligned} \quad (24)$$

Indeed, using Lemma 3, we can see that  $\mathrm{d}J_{\varepsilon, \eta}(\gamma; j_\gamma) = \|j_\gamma\|_{\mathbf{H}}^2 + \varepsilon \|j_\gamma''\|_{L^2(0, T)}^2 \geq 0$ . The problem (24) is solved by the finite element method using FreeFEM++. We denote by  $P_\Omega$  the projection in  $\Omega$ . Then, the gradient algorithm for solving (12) is

given by Algorithm 1. We point out that a re-meshing of  $Q_T$  is performed at each iteration, in order to be conform with the current trajectory  $\gamma_n$ . We illustrate the algorithm on two examples.

---

**Algo. 1:** Gradient descent
 

---

**Initialization** Choose a trajectory  $\gamma_0 \in \mathcal{S}_5$  such that  $0 < \gamma_0 < 1$ .

**For each**  $n \geq 0$  **do**

- ▷ Compute the solution  $\varphi_h$  of (20) associated with  $\gamma_n$ .
- ▷ Evaluate the costs  $J(\gamma_n)$  and  $J_{\varepsilon, \eta}(\gamma_n)$ .
- ▷ Compute the solution  $j_{\gamma_n}$  of (24).
- ▷ Update the trajectory  $\gamma_n$  by setting

$$\gamma_{n+1} = P_{\Omega}(\gamma_n - \rho j_{\gamma_n}), \quad \text{with } \rho > 0 \text{ fixed.}$$

**End**

---

**Example 1 – Sine function**

To test Algorithm 1, we first use the initial datum

$$y_0(x) = 10 \sin(\pi x), \quad y_1(x) = 0, \quad \forall x \in \Omega. \quad (\mathbf{Ex3-y_0})$$

We initialize the algorithm with the trajectory  $\gamma_0 \in \mathcal{S}_5$  associated with the parameters

$$x_i = \frac{3}{20} + \frac{1}{5} \frac{t_i}{T}, \quad p_i = \frac{1}{5T}, \quad c_i = 0, \quad 0 \leq i \leq N. \quad (\mathbf{Ex3-\gamma_0})$$

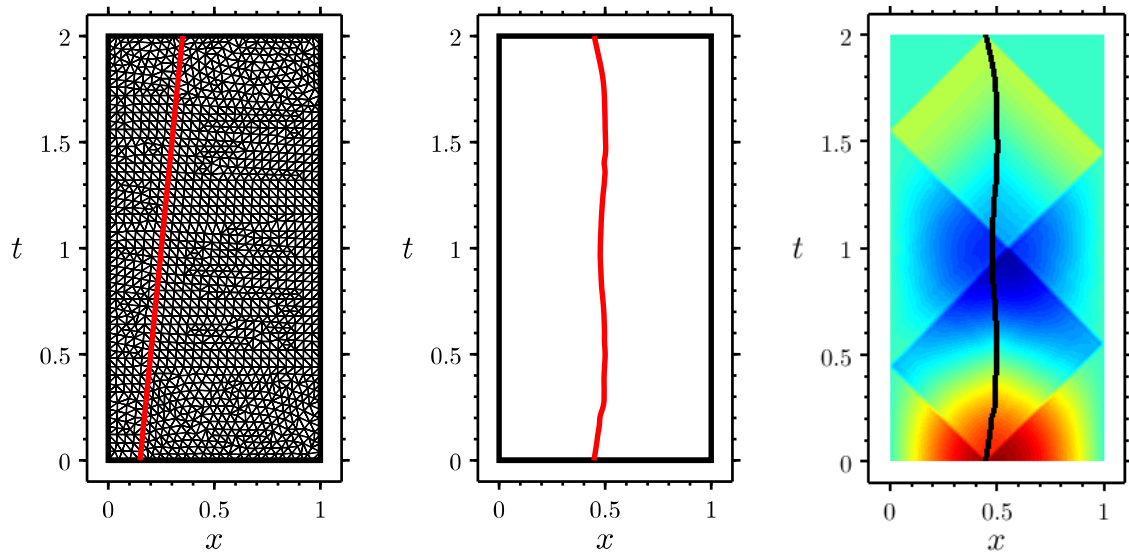
We set  $\varepsilon = 10^{-4}$ ,  $\eta = 10^3$ ,  $\bar{L} = 2.01$  and  $\rho = 10^{-2}$ . The initial trajectory  $\gamma_0$ , the optimal trajectory  $\gamma^*$  and the optimal controlled state  $\lambda^*$  are displayed in Figure 5. We observe that the optimal trajectory we get is close to a stationary control point located in  $x_0 = \frac{1}{2}$ , the maximum point of  $\sin(\pi x)$ . This is coherent with the case of controls distributed over domains  $q \subset Q_T$  (see [6, Example **EX1**]).

**Example 2 – Travelling wave**

To test again the similarities between the pointwise control case and the distributed control case, we now use the initial datum

$$y_0(x) = (10x - 3)^2(10x - 7)^2 \mathbb{1}_{[0.3, 0.7]}(x), \quad y_1(x) = y_0'(x), \quad \forall x \in \Omega. \quad (\mathbf{Ex4-y_0})$$



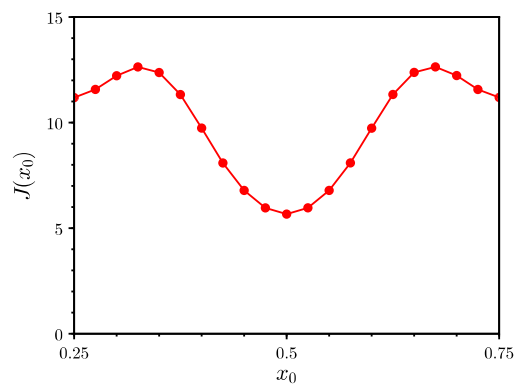


**Fig. 5: (Ex3)** – Initial trajectory  $\gamma_0$ , optimal trajectory  $\gamma^*$  and optimal controlled state  $\lambda^*$  (from left to right). The left figure also illustrates the type of mesh used to solve (20).

To see whether the control trajectory is likely to “follow” the wave associated with **(Ex4–y<sub>0</sub>)** as it is the case in [6, Example **EX2**]), we define the trajectories

$$g_{x_0}(t) = f_{x_0}(t) + 0.15 \cos(5\pi(t - x_0)), \quad \text{for any } x_0 \in \Omega.$$

Here,  $f_{x_0}$  is the characteristic line “ $x + t = x_0$ ” of the wave equation. The trajectory  $g_{\frac{1}{2}}$  is displayed in Figure 8-left. Then, for several values of  $x_0$  in  $\Omega$ , we evaluate the functional  $J(g_{x_0})$  associated with the initial datum **(Ex4–y<sub>0</sub>)**. The results are displayed in Figure 6, and we can see that  $J$  reaches its minimum for  $x_0 = \frac{1}{2}$ .



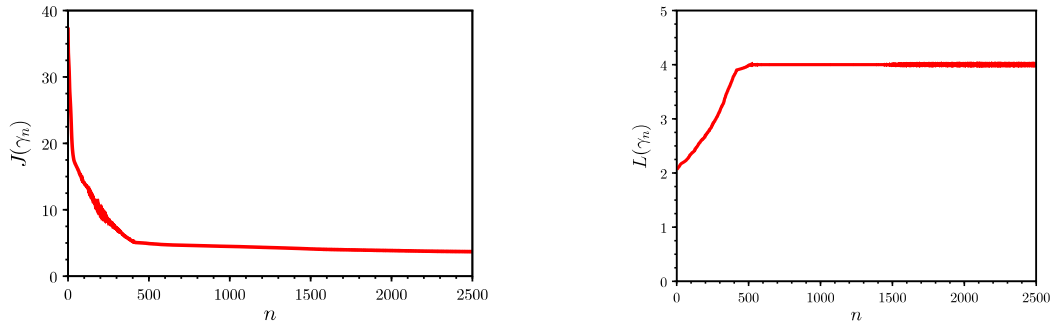
**Fig. 6: (Ex4)** –  $J(g_{x_0})$  vs.  $x_0$ .

We then employ Algorithm 1 for two different initial trajectories  $\gamma_0 \in \mathcal{S}_5$ , respectively defined by

$$x_i = g_{\frac{1}{2}}(t_i), \quad p_i = g'_{\frac{1}{2}}(t_i), \quad c_i = g''_{\frac{1}{2}}(t_i), \quad 0 \leq i \leq N, \quad (\mathbf{Ex4.1}-\gamma_0)$$

$$x_i = \frac{1}{4} + \frac{1}{2} \frac{t_i}{T}, \quad p_i = \frac{1}{2T}, \quad c_i = 0, \quad 0 \leq i \leq N. \quad (\mathbf{Ex4.2}-\gamma_0)$$

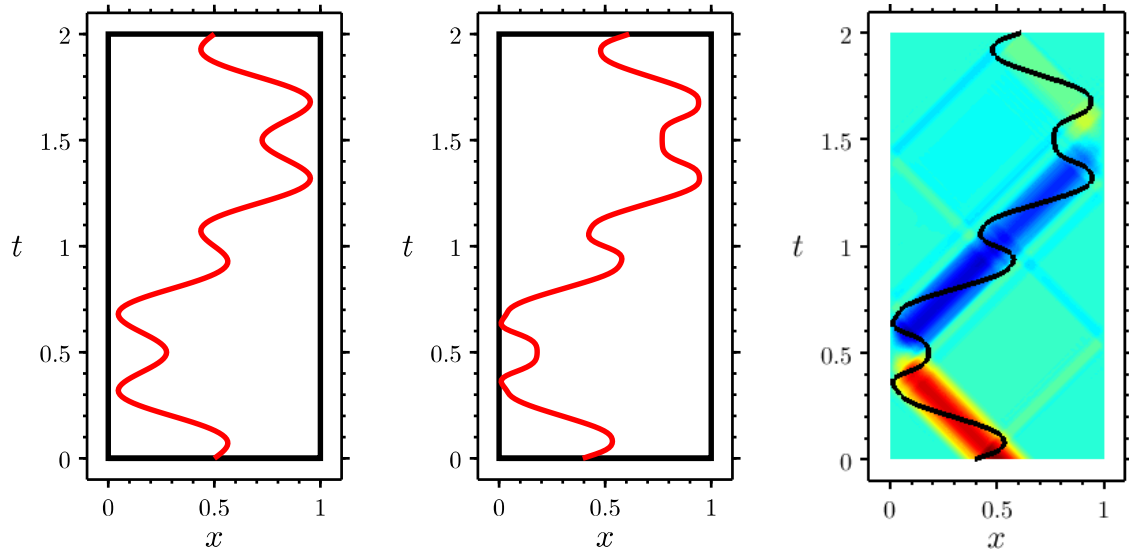
We set  $\varepsilon = 10^{-4}$ ,  $\eta = 10^3$ ,  $\bar{L} = 4$  and  $\rho = 10^{-2}$ . For the two examples (**Ex4.1**) and (**Ex4.2**), we display the initial trajectory  $\gamma_0$ , the optimal trajectory  $\gamma^*$  and the optimal controlled state  $\lambda^*$  in Figures 8-9 respectively. In the first setup, we observe that the optimal trajectory remains close to the wave support, which is coherent with the distributed control case. In the second setup, the optimal trajectory also seems to get closer to the wave support, but the convergence is very slow. This can be seen in Figure 7, where the evolution of the functional  $J(\gamma_n)$  and the curve length  $L(\gamma_n)$  are shown. The optimal costs are respectively  $J(\gamma^*) = 3.92$  for (**Ex4.1**) and  $J(\gamma^*) = 3.69$  for (**Ex4.2**). The difference is negligible compared to the initial cost  $J(\gamma_0) = 37.45$  for the example (**Ex4.2**).



**Fig. 7:** (**Ex4.2**) – Functional  $J(\gamma_n)$  (left) and curve length  $L(\gamma_n)$  (right).

## 5 Conclusion

On the basis of [9] that deals with controls distributed over non-cylindrical domains, we have built a mixed formulation characterizing the HUM control acting on a moving point. The formulation involves the adjoint state and a Lagrange multiplier which turns out to coincide with the controlled state. This approach leads to a variational formulation over a Hilbert space without distinction between the space

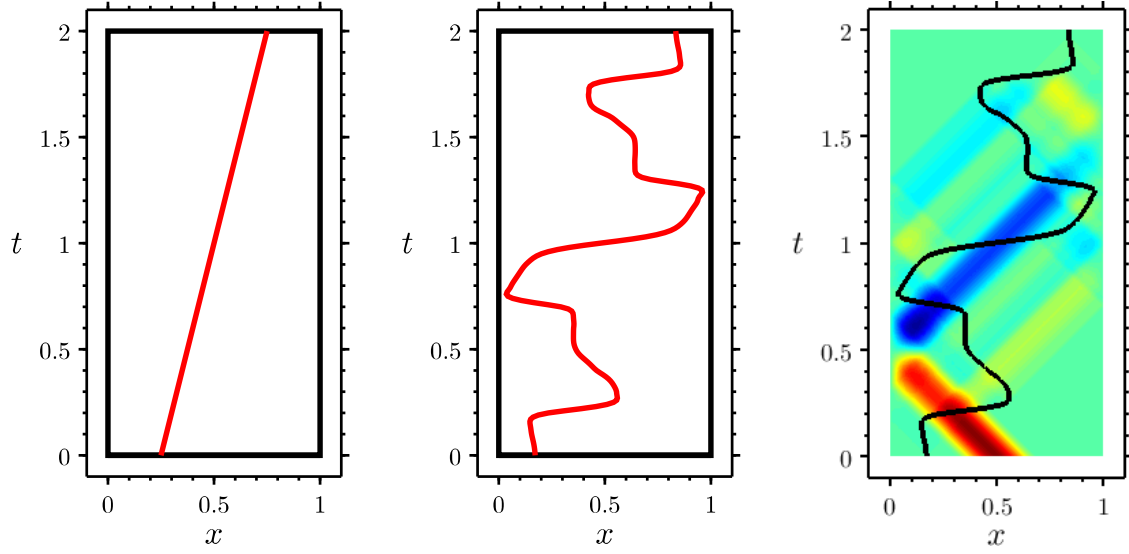


**Fig. 8: (Ex4.1)** – Initial trajectory  $\gamma_0$ , optimal trajectory  $\gamma^*$  and optimal controlled state  $\lambda^*$  (from left to right).

and time variables, making it very appropriate to our moving point situation. We have shown the well-posedness of the formulation using the observability inequality proved in [8]. At a practical level, the mixed formulation is discretized and solved in the finite element framework. The resolution amounts to solve a sparse symmetric system. From a numerical point of view, we have provided evidence of the convergence of the approximated control for regular initial data.

Still from a numerical perspective, for a fixed initial datum, we have considered the natural problem of optimizing the support of control. We have solved this problem with a simple gradient algorithm. For simplicity, the optimization is made over very regular trajectories. The results we get are similar with those obtained in [6], where the same problem is studied for controls distributed over non-cylindrical domains. Although, the convergence towards the optimal trajectory seems to be generally much slower.

This work may be extended to several directions. First, as it is done in [30] for distributed controls, one could try to justify rigorously the well-posedness of the support optimization problem. In that context, it could be interesting to find the minimal regularity necessary for the control trajectories. Besides, one could try to implement other types of algorithm for solving the problem, as for instance an algorithm based on the level-set method. Another challenge is the extension of the observability inequality to the multidimensional case, where we cannot make use of the d'Alembert formula.



**Fig. 9: (Ex4.2)** – Initial trajectory  $\gamma_0$ , optimal trajectory  $\gamma^*$  and optimal controlled state  $\lambda^*$  (from left to right).

## A Fourier expansion of the HUM control

In this appendix, we expand in terms of Fourier series the adjoint state  $\varphi$  linked to the HUM control  $v$  by the relation (9), as well as the associated controlled state  $y$ . These expansions are used to evaluate the errors  $\|v - v_h\|_{\mathbf{H}'}$  and  $\|y - y_h\|_{\mathbf{\Lambda}}$  in Section 4. One can show that  $\varphi$  and  $y$  take the form

$$\varphi(x, t) = \sum_{p \geq 1} \left( a_p \cos(p\pi t) + \frac{b_p}{p\pi} \sin(p\pi t) \right) \sin(p\pi x), \quad (25)$$

$$y(x, t) = \sum_{p \geq 1} c_p(t) \sin(p\pi x). \quad (26)$$

We set

$$\begin{cases} \xi_p^a(t) = \cos(p\pi t) \sin(p\pi \gamma(t)), \\ \xi_p^b(t) = \frac{1}{p\pi} \sin(p\pi t) \sin(p\pi \gamma(t)), \end{cases} \quad \forall p \geq 1.$$

Injecting (25) in the terms appearing in the optimality condition (8), we get

$$\begin{aligned} \int_0^T \varphi(\gamma(t), t) \bar{\varphi}(\gamma(t), t) dt &= \sum_{p, q \geq 1} a_p \bar{a}_q \int_0^T \xi_p^a \xi_q^a + \sum_{p, q \geq 1} b_p \bar{b}_q \int_0^T \xi_p^b \xi_q^b \\ &+ \sum_{p, q \geq 1} a_p \bar{b}_q \int_0^T \xi_p^a \xi_q^b + \sum_{p, q \geq 1} b_p \bar{a}_q \int_0^T \xi_p^b \xi_q^a, \end{aligned} \quad (27)$$

$$\int_0^T \frac{d}{dt} \varphi(\gamma(t), t) \frac{d}{dt} \bar{\varphi}(\gamma(t), t) dt = \sum_{p, q \geq 1} a_p \bar{a}_q \int_0^T \xi_p^{a'} \xi_q^{a'} + \sum_{p, q \geq 1} b_p \bar{b}_q \int_0^T \xi_p^{b'} \xi_q^{b'} \quad (28)$$

$$+ \sum_{p, q \geq 1} a_p \bar{b}_q \int_0^T \xi_p^{a'} \xi_q^{b'} + \sum_{p, q \geq 1} b_p \bar{a}_q \int_0^T \xi_p^{b'} \xi_q^{a'},$$

$$\int_{\Omega} y_0 \bar{\varphi}_1 = \frac{1}{2} \sum_{p \geq 1} c_p(y_0) \bar{b}_p \quad \text{and} \quad \langle y_1, \bar{\varphi}_0 \rangle_{-1,1} = \frac{1}{2} \sum_{p \geq 1} c_p(y_1) \bar{a}_p. \quad (29)$$

Here,  $c_p(y_0)$  and  $c_p(y_1)$  are the Fourier coefficients of  $y_0$  and  $y_1$ . Thus, optimality condition (8) can be rewritten

$$\left\langle \mathcal{M}_{\gamma} \begin{pmatrix} \{a_p\}_{p \geq 1} \\ \{b_p\}_{p \geq 1} \end{pmatrix}, \begin{pmatrix} \{\bar{a}_q\}_{q \geq 1} \\ \{\bar{b}_q\}_{q \geq 1} \end{pmatrix} \right\rangle = \left\langle \mathcal{F}_{y_0}, \begin{pmatrix} \{\bar{a}_q\}_{q \geq 1} \\ \{\bar{b}_q\}_{q \geq 1} \end{pmatrix} \right\rangle, \quad \forall (\bar{a}_q, \bar{b}_q)_{q \geq 1}, \quad (30)$$

where the positive definite matrix  $\mathcal{M}_{\gamma}$  and the vector  $\mathcal{F}_{y_0}$  are obtained from (27-28) and (29) respectively. The resolution of the infinite-dimensional system (30) (reduced to a finite-dimensional one by truncation) provides an approximation of the adjoint state  $\varphi$  linked to the HUM control  $v$  by (9).

Injecting (26) in the wave equation (1), we find that  $c_p(t)$  satisfies

$$\begin{cases} c_p''(t) + (p\pi)^2 c_p(t) = 2v(t) \sin(p\pi\gamma(t)), & \forall t > 0, \\ c_p(0) = c_p(y_0), \quad c_p'(0) = c_p(y_1). \end{cases}$$

We then have

$$c_p(t) = c_p(y_0) \cos(p\pi t) + \frac{c_p(y_1)}{p\pi} \sin(p\pi t) + \frac{2}{p\pi} \int_0^t v(s) \sin(p\pi\gamma(s)) \sin(p\pi(t-s)) ds.$$

Finally, by integration by parts, we deduce

$$\begin{aligned} c_p(t) &= c_p(y_0) \cos(p\pi t) + \frac{c_p(y_1)}{p\pi} \sin(p\pi t) \\ &+ \frac{2}{p\pi} \int_0^t \varphi(\gamma(s), s) \sin(p\pi\gamma(s)) \sin(p\pi(t-s)) ds \\ &- 2 \int_0^t \frac{d}{ds} \varphi(\gamma(s), s) \sin(p\pi\gamma(s)) \cos(p\pi(t-s)) ds \\ &+ 2 \int_0^t \frac{d}{ds} \varphi(\gamma(s), s) \cos(p\pi\gamma(s)) \gamma'(s) \sin(p\pi(t-s)) ds. \end{aligned}$$

## References

- [1] A. Agresti, D. Andreucci, and P. Loreti. Observability for the wave equation with variable support in the Dirichlet and Neumann cases. In *International Conference on Informatics in Control, Automation and Robotics*, pages 51–75. Springer, 2018.
- [2] A. Bamberger, J. Jaffre, and J.-P. Yvon. Punctual control of a vibrating string: numerical analysis. *Comput. Math. Appl.*, 4(2):113–138, 1978.
- [3] C. Bardos, G. Lebeau, and J. Rauch. Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.*, 30(5):1024–1065, 1992.
- [4] M. Berggren and R. Glowinski. Controllability issues for flow-related models: a computational approach. Technical report, 1994.
- [5] M. Bernadou and K. Hassan. Basis functions for general Hsieh-Clough-Tocher triangles, complete or reduced. *Internat. J. Numer. Methods Engrg.*, 17(5):784–789, 1981.
- [6] A. Bottois, N. Cîndea, and A. Münch. Optimization of non-cylindrical domains for the exact null controllability of the 1D wave equation. *To be published*, 2019.
- [7] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.
- [8] C. Castro. Exact controllability of the 1-D wave equation from a moving interior point. *ESAIM Control Optim. Calc. Var.*, 19(1):301–316, 2013.
- [9] C. Castro, N. Cîndea, and A. Münch. Controllability of the linear one-dimensional wave equation with inner moving forces. *SIAM J. Control Optim.*, 52(6):4027–4056, 2014.
- [10] D. Chapelle and K.-J. Bathe. The inf-sup test. *Comput. & Structures*, 47(4-5):537–545, 1993.
- [11] F. Chatelin. *Eigenvalues of matrices*, volume 71 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2012. With exercises by Mario Ahués and the author, Translated with additional material by Walter Ledermann, Revised reprint of the 1993 edition [MR1232655].
- [12] P. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

- [13] N. Cîndea and A. Münch. A mixed formulation for the direct approximation of the control of minimal  $L^2$ -norm for linear type wave equations. *Calcolo*, 52(3):245–288, 2015.
- [14] R. Dáger and E. Zuazua. *Wave propagation, observation and control in 1-d flexible multi-structures*, volume 50 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006.
- [15] S. Ervedoza and E. Zuazua. A systematic method for building smooth controls for smooth data. *Discrete Contin. Dyn. Syst. Ser. B*, 14(4):1375–1401, 2010.
- [16] R. Glowinski, J.-L. Lions, and J. He. *Exact and approximate controllability for distributed parameter systems*, volume 117 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2008. A numerical approach.
- [17] F. Hecht. New development in freefem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.
- [18] A. Henrot, W. Horn, and J. Sokołowski. Domain optimization problem for stationary heat equation. *Appl. Math. Comput. Sci.*, 6(2):353–374, 1996. Shape optimization and scientific computations (Warsaw, 1994).
- [19] A. Henrot and J. Sokołowski. A shape optimization problem for the heat equation. In *Optimal control (Gainesville, FL, 1997)*, volume 15 of *Appl. Optim.*, pages 204–223. Kluwer Acad. Publ., Dordrecht, 1998.
- [20] A. Khapalov. Controllability of the wave equation with moving point control. *Appl. Math. Optim.*, 31(2):155–175, 1995.
- [21] A. Khapalov. Mobile point controls versus locally distributed ones for the controllability of the semilinear parabolic equation. *SIAM J. Control Optim.*, 40(1):231–252, 2001.
- [22] A. Khapalov. Observability and stabilization of the vibrating string equipped with bouncing point sensors and actuators. *Math. Methods Appl. Sci.*, 24(14):1055–1072, 2001.
- [23] J. Le Rousseau, G. Lebeau, P. Terpolilli, and E. Trélat. Geometric control condition for the wave equation with a time-dependent observation domain. *Anal. PDE*, 10(4):983–1015, 2017.
- [24] J.-L. Lions. *Some methods in the mathematical analysis of systems and their control*. Kexue Chubanshe (Science Press), Beijing; Gordon & Breach Science Publishers, New York, 1981.
- [25] 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. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos, G. Lebeau and J. Rauch.

- [26] J.-L. Lions. Pointwise control for distributed systems. In *Control and estimation in distributed parameter systems*, pages 1–39. SIAM, 1992.
- [27] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
- [28] A. Meyer. A simplified calculation of reduced HCT-basis functions in a finite element context. *Comput. Methods Appl. Math.*, 12(4):486–499, 2012.
- [29] A. Münch. A uniformly controllable and implicit scheme for the 1-D wave equation. *M2AN Math. Model. Numer. Anal.*, 39(2):377–418, 2005.
- [30] F. Periago. Optimal shape and position of the support for the internal exact control of a string. *Systems Control Lett.*, 58(2):136–140, 2009.
- [31] Á. M. Ramos, R. Glowinski, and J. Périaux. Pointwise control of the Burgers equation and related Nash equilibrium problems: computational approach. *J. Optim. Theory Appl.*, 112(3):499–516, 2002.



# C Construction de contrôles exacts pour une équation des ondes semi-linéaire

Titre : *Constructive exact controls for semi-linear wave equations.*

Auteurs : ARTHUR BOTTOIS, JÉRÔME LEMOINE et ARNAUD MÜNCH.

Soumis en juin 2021.

<https://hal.archives-ouvertes.fr/hal-03112309>

## Résumé

Dans cet article, on traite le problème de la contrôlabilité exacte pour une équation des ondes semi-linéaire avec contrôle interne. On reprend ici les notations de la Sect. 1.2.1 avec  $d \in \{1, 2, 3\}$ . On considère le système d'état (1.60) où le contrôle  $u$  peut agir dans l'ouvert cylindrique  $q = \omega \times (0, T)$ . La contrôlabilité exacte pour le système (1.60) a été établie dans [15, Th. 2.2] (voir aussi Th. 3) sous des hypothèses classiques de type (GCC) pour l'ouvert  $\omega$  et sous l'hypothèse de croissance asymptotique

$$\limsup_{|r| \rightarrow +\infty} \frac{|g(r)|}{|r| \ln^{1/2} |r|} = 0 \quad (\mathbf{H}_1)$$

pour la semi-linéarité  $g$ . La démonstration s'appuie sur le théorème du point fixe de Schauder afin de justifier l'existence d'un point fixe pour l'opérateur  $K$  défini en (1.62).

L'objectif du présent article est de construire une suite  $(u_k)_{k \in \mathbb{N}} \subset L^2(q)$  convergeant fortement vers un contrôle exact  $u$  pour le système (1.60). Tout d'abord, on remarque que la stratégie consistant à considérer les itérés de Picard  $y_{k+1} = K(y_k)$ ,  $k \geq 0$ , échoue en général. En effet, la plupart du temps, l'opérateur  $K$  n'est pas contractant, et ce même si  $g$  est globalement lipschitzien. On utilise ici une approche moindres carrés. On définit l'espace de Hilbert

$$\mathcal{H} = \left\{ (y, u) \in L^2(Q) \times L^2(q); \quad (y, \partial_t y) \in C([0, T]; \mathbf{V}) \text{ et } \square y \in L^2(Q) \right\}.$$

Pour  $(y^0, y^1), (z^0, z^1) \in \mathbf{V}$  fixés, on introduit également les sous-espaces fermés

$$\begin{aligned} \mathcal{A} &= \left\{ (y, u) \in \mathcal{H}; \quad (y, \partial_t y)(\cdot, 0) = (y^0, y^1) \text{ et } (y, \partial_t y)(\cdot, T) = (z^0, z^1) \right\}, \\ \mathcal{A}_0 &= \left\{ (y, u) \in \mathcal{H}; \quad (y, \partial_t y)(\cdot, 0) = (0, 0) \text{ et } (y, \partial_t y)(\cdot, T) = (0, 0) \right\}. \end{aligned}$$

On considère alors le problème d'optimisation (non convexe)

$$\min_{(y, u) \in \mathcal{A}} E(y, u), \quad \text{avec} \quad E(y, u) = \frac{1}{2} \|\square y + g(y) - u \mathbb{1}_\omega\|_{L^2(Q)}^2.$$

La fonctionnelle moindres carrés  $E$  mesure la proximité d'un couple  $(y, u) \in \mathcal{A}$  avec l'ensemble des paires état-contrôle de (1.60). De surcroît, on observe que tout zéro  $(y, u) \in \mathcal{A}$  de  $E$  est une paire état-contrôle de (1.60). On cherche donc à construire une suite minimisante  $(y_k, u_k)_{k \in \mathbb{N}} \subset \mathcal{A}$  qui converge vers un zéro de  $E$ . Pour ce faire, on ajoute deux hypothèses sur  $g$  légèrement plus fortes que  $(\mathbf{H}_1)$ . Le résultat principal de l'article est reproduit ci-dessous.

**Théorème.** *Soit  $s \in (0, 1]$ . On pose  $\beta^*(s) = \sqrt{\frac{s}{2C_{\Omega, T}(2s+1)}}$ . On suppose que*

$$[g']_s = \sup_{\substack{a, b \in \mathbb{R} \\ a \neq b}} \frac{|g'(a) - g'(b)|}{|a - b|^s} < +\infty, \quad (\overline{\mathbf{H}}_s)$$

$$\exists \alpha \geq 0, \exists \beta \in [0, \beta^*(s)), \quad |g'(r)| \leq \alpha + \beta \ln^{1/2}(1 + |r|), \quad \forall r \in \mathbb{R}. \quad (\overline{\mathbf{H}}_2)$$

Pour  $(y_0, u_0) \in \mathcal{A}$ , on définit la suite  $(y_k, u_k)_{k \in \mathbb{N}}$  par

$$(y_{k+1}, u_{k+1}) = (y_k, u_k) - \lambda_k (Y_k, U_k), \quad \text{avec} \quad \lambda_k = \arg \min_{\lambda \in [0, 1]} E((y_k, u_k) - \lambda(Y_k, U_k)),$$

et où  $(Y_k, U_k) \in \mathcal{A}_0$  est la paire état-contrôle HUM du système linéarisé

$$\begin{cases} \partial_t^2 Y - \Delta Y + g'(y_k)Y = U \mathbb{1}_\omega + (\square y_k + g(y_k) - u_k \mathbb{1}_\omega) & \text{dans } Q, \\ Y = 0 & \text{sur } \Sigma, \\ (Y, \partial_t Y)(\cdot, 0) = (0, 0) & \text{dans } \Omega. \end{cases}$$

Plus précisément,  $U_k$  est ici le contrôle à zéro de norme  $L^2(q)$  minimale du système ci-dessus. La suite  $(y_k, u_k)_{k \in \mathbb{N}}$  converge fortement vers une paire état-contrôle  $(y^*, u^*)$  de (1.60). En outre, la convergence est au moins linéaire, et même d'ordre  $1 + s$  après un nombre fini d'itérations.

Il est à noter que l'algorithme de descente pour  $E$  s'apparente à un algorithme de Newton amorti pour la fonction  $F(y, u) = \square y + g(y) - u \mathbb{1}_\omega$ . Pour illustrer la convergence de la suite minimisante, on complète l'article par quelques simulations numériques en une et deux dimensions. On initialise la suite avec la paire état-contrôle HUM du système (1.60) "linéaire", i.e. pour  $g \equiv 0$ . On considère une semi-linéarité "critique" de la forme

$$g(r) = -c_g r \ln^{1/2}(2 + |r|), \quad \forall r \in \mathbb{R},$$

avec  $c_g \in \mathbb{R}$ . Cette semi-linéarité  $g$  vérifie  $(\overline{\mathbf{H}}_s)$  pour  $s = 1$  et  $(\overline{\mathbf{H}}_2)$  pour  $|c_g|$  assez petit. Pour des petites valeurs de  $c_g$ , on met en évidence la convergence quadratique de la suite vers une paire état-contrôle de (1.60). Par ailleurs, on trouve un pas de descente  $\lambda_k$  systématiquement proche de 1, ce qui signifie que l'algorithme de descente pour  $E$  coïncide avec l'algorithme de Newton classique pour  $F$ . Pour des valeurs de  $c_g$  plus élevées, on met en évidence une convergence linéaire, puis quadratique à partir d'un certain rang. On observe également un pas de descente  $\lambda_k$  initialement proche de 0, puis croissant jusqu'à atteindre 1.

## Article

# Constructive exact controls for semi-linear wave equations

ARTHUR BOTTOIS      JÉRÔME LEMOINE      ARNAUD MÜNCH\*

September 5, 2022

## Abstract

The exact distributed controllability of the semi-linear wave equation  $\partial_{tt}y - \Delta y + g(y) = f1_\omega$  posed over multi-dimensional and bounded domains, assuming that  $g \in \mathcal{C}^1(\mathbb{R})$  satisfies the growth condition  $\limsup_{|r| \rightarrow \infty} g(r)/(|r| \ln^{1/2}|r|) = 0$  has been obtained by Fu, Yong and Zhang in 2007. The proof based on a non constructive Leray-Schauder fixed point theorem makes use of precise estimates of the observability constant for a linearized wave equation. Assuming that  $g'$  does not grow faster than  $\beta \ln^{1/2}|r|$  at infinity for  $\beta > 0$  small enough and that  $g'$  is uniformly Hölder continuous on  $\mathbb{R}$  with exponent  $s \in (0, 1]$ , we design a constructive proof yielding an explicit sequence converging to a controlled solution for the semi-linear equation, at least with order  $1 + s$  after a finite number of iterations. Numerical experiments in the two-dimensional case illustrate the results. This work extends to a multi-dimensional case and completes with some numerical experiments the study by Münch and Trélat devoted to the one-dimensional situation.

**AMS Classifications:** 35L71, 49M15, 93E24.

**Keywords:** Semilinear wave equation, exact controllability, least-squares approach.

## 1 Introduction

Let  $\Omega$  be a bounded domain of  $\mathbb{R}^d$ ,  $d \in \{2, 3\}$  with  $\mathcal{C}^{1,1}$  boundary and  $\omega \subset\subset \Omega$  be a non-empty open set. Let  $T > 0$  and denote  $Q_T := \Omega \times (0, T)$ ,  $q_T := \omega \times (0, T)$  and  $\Sigma_T := \partial\Omega \times (0, T)$ . We consider the semi-linear wave equation

$$\begin{cases} \partial_{tt}y - \Delta y + g(y) = f1_\omega, & \text{in } Q_T, \\ y = 0, & \text{on } \Sigma_T, \\ (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1), & \text{in } \Omega, \end{cases} \quad (1)$$

where  $(u_0, u_1) \in \mathbf{V} := H_0^1(\Omega) \times L^2(\Omega)$  is the initial state of  $y$  and  $f \in L^2(q_T)$  is a *control* function. Here and throughout the paper,  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a function of class  $\mathcal{C}^1$  such that  $|g(r)| \leq C(1 + |r|) \ln(2 + |r|)$  for every  $r \in \mathbb{R}$  and some  $C > 0$ . Then, (1) has a unique global weak solution in  $\mathcal{C}([0, T]; H_0^1(\Omega)) \cap \mathcal{C}^1([0, T]; L^2(\Omega))$  (see [6, 8]).

The exact controllability for (1) in time  $T$  is formulated as follows: for any  $(u_0, u_1), (z_0, z_1) \in \mathbf{V}$ , find a control function  $f \in L^2(q_T)$  such that the weak solution of (1) satisfies  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$ . Assuming a growth condition on the non-linearity  $g$  at infinity, this problem has been solved in [18].

---

\*Université Clermont Auvergne, CNRS, LMBP, F-63000 Clermont-Ferrand, France; arthur.bottois@uca.fr, jerome.lemoine@uca.fr, arnaud.munch@uca.fr (Corresponding author).

**Theorem 1.** [18, Theorem 2.2] For any  $x_0 \in \mathbb{R}^d \setminus \overline{\Omega}$ , let  $\Gamma_0 = \{x \in \partial\Omega, (x - x_0) \cdot \nu(x) > 0\}$  and, for any  $\epsilon > 0$ ,  $\mathcal{O}_\epsilon(\Gamma_0) = \{y \in \mathbb{R}^d \mid |y - x| \leq \epsilon \text{ for } x \in \Gamma_0\}$ . Assume

$$(\mathbf{H}_0) \quad T > 2 \max_{x \in \overline{\Omega}} |x - x_0| \text{ and } \omega \supseteq \mathcal{O}_\epsilon(\Gamma_0) \cap \Omega \text{ for some } \epsilon > 0.$$

If  $g$  satisfies

$$(\mathbf{H}_1) \quad \limsup_{|r| \rightarrow \infty} \frac{|g(r)|}{|r| \ln^{1/2} |r|} = 0$$

then (1) is exactly controllable in time  $T$ .

This result improves [29] where a stronger condition of the support  $\omega$  is made, namely that  $\omega$  is a neighborhood of  $\partial\Omega$  and that  $T > \text{diam}(\Omega \setminus \omega)$ . In Theorem 1,  $\Gamma_0$  is the usual star-shaped part of the whole boundary of  $\Omega$  introduced in [30] and  $\nu(x)$  denotes the outward normal derivative at any point  $x \in \partial\Omega$ .

A special case of Theorem 1 is when  $g$  is globally Lipschitz continuous, which gives the main result of [40], later generalized to an abstract setting in [23] using a global version of the inverse function theorem and improved in [38] for control domains  $\omega$  satisfying the classical multiplier method of Lions [30]. Theorem 1 extends to the multi-dimensional case the result of [41] under the condition  $\limsup_{|r| \rightarrow \infty} \frac{|g(r)|}{|r| \ln^2 |r|} = 0$ , relaxed later on in [6], following [14], and in [31]. The exact controllability for subcritical non-linearities is obtained in [12] assuming the sign condition  $rg(r) \geq 0$  for every  $r \in \mathbb{R}$ . This latter assumption has been weakened in [22] to an asymptotic sign condition leading to a semi-global controllability result in the sense that the final data  $(z_0, z_1)$  is prescribed in a precise subset of  $\mathbf{V}$ .

The proof given in [18, 29] is based on a fixed-point argument introduced in [39, 41] that reduces the exact controllability problem to the obtention of suitable *a priori* estimates for the linearized wave equation with a potential (see Proposition 6 in Appendix A). More precisely, it is shown that the operator  $\Lambda : L^\infty(0, T; L^d(\Omega)) \rightarrow L^\infty(0, T; L^d(\Omega))$  where  $y := \Lambda(z)$  is a controlled solution through the control function  $f$  of the linear boundary value problem

$$\begin{cases} \partial_{tt}y - \Delta y + \widehat{g}(z)y = -g(0) + f1_\omega, & \text{in } Q_T, \\ y = 0, & \text{on } \Sigma_T, \\ (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1), & \text{in } \Omega, \end{cases} \quad \widehat{g}(r) := \begin{cases} \frac{g(r) - g(0)}{r} & \text{if } r \neq 0, \\ g'(0) & \text{if } r = 0, \end{cases} \quad (2)$$

satisfying  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$  has a fixed point. The control  $f$  is chosen in [29] as the one of minimal  $L^2(Q_T)$ -norm. The existence of a fixed point for the compact operator  $\Lambda$  is obtained by using the Leray-Schauder's degree theorem. In particular, under the growth assumption  $(\mathbf{H}_1)$ , it is shown a stability property of the operator  $\Lambda$ , i.e. the existence of a constant  $M = M(\|(u_0, u_1)\|_{\mathbf{V}}, \|(z_0, z_1)\|_{\mathbf{V}})$  such that  $\Lambda(B_{L^\infty(0, T; L^d(\Omega))}(0, M)) \subset B_{L^\infty(0, T; L^d(\Omega))}(0, M)$ .

This article is concerned with the determination of strongly convergent sequences  $(y_k, f_k)_{k \in \mathbb{N}}$  toward a state-control pair for the nonlinear system (1). The controllability of nonlinear partial differential equations has attracted a large number of works in the last decades (see the monograph [11] and references therein). However, very few are concerned with the approximation of exact controls for nonlinear partial differential equations, and the construction of convergent control approximations for controllable nonlinear equations remains in general an open question. This is mainly due to the fact that the available controllability results are based on non constructive fixed arguments. Thus, the Picard iterates  $(y_k)_{k \in \mathbb{N}}$  associated with the operator  $\Lambda$ , defined for any  $y_0 \in L^\infty(0, T; L^d(\Omega))$  by  $y_{k+1} = \Lambda(y_k)$ ,  $k \geq 0$ , remains bounded in  $L^\infty(0, T; L^d(\Omega))$  but has no reason to converge (we refer to [17] where divergence is observed numerically in a parabolic case).

Recently, two constructions of convergent sequences have been proposed in the one-dimensional case with  $\Omega = (0, 1)$ : the first one in [36] is based on a least-squares approach: the extremal problem

$$\min_{(y, f) \in \mathcal{A}} E(y, f), \quad E(y, f) := \|\partial_{tt}y - \partial_{xx}y + g(y) - f1_\omega\|_{L^2(Q_T)}^2 \quad (3)$$

is considered where  $\mathcal{A}$  is a closed subset of  $L^2(Q_T) \times L^2(Q_T)$  containing the initial condition and controllability requirement at the initial and final time respectively. Assuming notably that  $g \in C^1(\mathbb{R})$  satisfies for some  $\beta > 0$  small enough the asymptotic condition

$$(\mathbf{H}'_1) \quad \limsup_{|r| \rightarrow \infty} \frac{|g'(r)|}{\ln^{1/2}|r|} \leq \beta$$

a minimizing sequence for  $E$  is constructed and proved to converge strongly to a state-control pair for (1). We refer to [36, Theorem 2.3] for a precise statement. The least-squares approach turns out to be related to the operator  $\Lambda_N : L^\infty(Q_T) \rightarrow L^\infty(Q_T)$ , where for any  $z \in L^\infty(Q_T)$ ,  $y = \Lambda_N(z)$  is a controlled solution through the control  $f$  of

$$\begin{cases} \partial_{tt}y - \partial_{xx}y + g'(z)y = f 1_\omega + g'(z)z - g(z), & \text{in } Q_T, \\ y = 0, & \text{on } \Sigma_T, \\ (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1) & \text{in } \Omega \end{cases} \quad (4)$$

satisfying  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$ . Again, for each  $z$ , the control of minimal  $L^2(Q_T)$  norm is considered. A similar strategy has been successfully applied in [27] for a semi-linear 1D heat equation. The second one in [2] focuses on the boundary controllability and considers the operator  $\Lambda_F : L^\infty(Q_T) \rightarrow L^\infty(Q_T)$ , where for any  $z \in L^\infty(Q_T)$ ,  $y = \Lambda_F(z)$  is a controlled solution through the control  $f$  of

$$\begin{cases} \partial_{tt}y - \partial_{xx}y = -g(z), & \text{in } Q_T, \\ y(0, \cdot) = 0, \quad y(1, \cdot) = f, & \text{on } (0, T), \\ (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1) & \text{in } \Omega \end{cases} \quad (5)$$

satisfying  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$ . For each  $z$ , the state-control pair  $(y, f)$  is chosen as the minimizer of an  $L^2$  functional involving parametrized Carleman weights. Under the asymptotic condition  $(\mathbf{H}'_1)$ , it is shown that the operator  $\Lambda_F$  is contracting for a large enough parameter. We refer to [2, Theorem 8] for a precise statement. This provides a convergent sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  to a state-control pair for the nonlinear equation. Remark that a similar operator has also been used recently for a semi-linear heat equation in [37].

The objective of the present paper is two-fold: first, we extend the least-squares approach introduced in [36] to a multi-dimensional case. With respect to the one-dimensional case studied in [36], the controlled solution is not anymore in  $L^\infty(Q_T)$  but in  $L^\infty(0, T; L^p(\Omega))$  for some  $p$  related to the dimension of  $\Omega$ . This requires a finer analysis in order to estimate the observability constant. Second, we give some numerical illustrations of the method (not given in [36]) both in the one and two dimensional cases. This requires the approximation of exact controls for linear wave equations, which is known to be a delicate issue, since the works of Glowinski (see the monograph [20] and the recent review [34]).

The paper is organized as follows. In Sections 2 and 3, we adapt [36] to the higher dimension without reproducing all the arguments. More precisely, we define the non-convex optimization problem (6) involving the least-squares functional  $E$ . We show that any critical point  $(y, f)$  for  $E$  such that  $g'(y) \in L^\infty(0, T; L^d(\Omega))$  is also a zero of  $E$ . This is done by introducing a descent direction  $(Y^1, F^1)$  for  $E$  at any  $(y, f)$  for which  $E'(y, f) \cdot (Y^1, F^1)$  is proportional to  $\sqrt{E(y, f)}$ . A minimizing sequence based on  $(Y^1, F^1)$  is then proved to converge to a controlled pair for the semi-linear wave equation (1) under assumptions on  $g$  similar to  $(\mathbf{H}'_1)$ . We refer to Theorem 2 for a precise statement of our result. Section 4 provides several comments and Section 5 then illustrates the result with some numerical experiments in one and two dimensions. Section 6 concludes. In Appendix A, we recall some *a priori* estimates for the linearized wave equation with potential in  $L^\infty(0, T; L^d(\Omega))$  and source term in  $L^2(Q_T)$ .

As far as we know, the method introduced and analyzed in this work is the first one providing an explicit, algorithmic construction of exact controls for semi-linear wave equations with non-Lipschitz non-linearity and defined over multi-dimensional bounded domains.

This work extends the one-dimensional study addressed in [36], for which the solution is uniformly bounded with respect to both the time and space variable. In contrast, the multi-dimensional case for

1 which the solution does not belong to  $L^\infty(Q_T)$ , requires finer analysis : we refer for instance to Lemma  
2 5. With respect to [36], some proof very closed to the one dimensional case are omitted; on the contrary,  
3 the proof of Proposition 5, left to the reader in [36], is given in the present work. With respect to [36],  
4 we also provide some numerical experiments (including for  $d = 1$ ).

5 For parabolic equations with Lipschitz non-linearity, we mention [24]. Actually, this work devoted to  
6 controllability problems takes their roots in earlier works, namely [25, 26], concerned with the approxi-  
7 mation of solution of Navier-Stokes type problem, through least-square methods: they refine the analysis  
8 performed in [28, 33] inspired from the seminal contribution [3].

9 Throughout, we denote by  $\|\cdot\|_\infty$  the usual norm in  $L^\infty(\mathbb{R})$ , by  $(\cdot, \cdot)_X$  the scalar product of  $X$  (if  $X$   
10 is a Hilbert space) and by  $\langle \cdot, \cdot \rangle_{X,Y}$  the duality product between  $X$  and  $Y$ . The notation  $\|\cdot\|_{2,q_T}$  stands  
11 for  $\|\cdot\|_{L^2(Q_T)}$  and  $\|\cdot\|_p$  for  $\|\cdot\|_{L^p(Q_T)}$ ,  $p \in \mathbb{N}^*$ . We also denote by  $C$  a positive constant depending only  
12 on  $\Omega$  and  $T$  that may vary from line to line.

13 In the rest of the paper, we assume that the open set  $\omega$  and the time  $T$  satisfy  $(\mathbf{H}_0)$ .

## 14 2 The least-squares functional and its properties

### 15 2.1 The least-squares problem

16 We define the Hilbert space

$$\mathcal{H} = \left\{ (y, f) \in L^2(Q_T) \times L^2(Q_T), y \in \mathcal{C}([0, T]; H_0^1(\Omega)) \cap \mathcal{C}^1([0, T]; L^2(\Omega)) \mid \partial_{tt}y - \Delta y \in L^2(Q_T) \right\}$$

17 endowed with the scalar product

$$\begin{aligned} ((y, f), (\bar{y}, \bar{f}))_{\mathcal{H}} &= (y, \bar{y})_2 + ((y(\cdot, 0), \partial_t y(\cdot, 0)), (\bar{y}(\cdot, 0), \partial_t \bar{y}(\cdot, 0)))_{\mathbf{V}} \\ &\quad + (\partial_{tt}y - \Delta y, \partial_{tt}\bar{y} - \Delta \bar{y})_2 + (f, \bar{f})_{2,q_T} \end{aligned}$$

18 and the norm  $\|(y, f)\|_{\mathcal{H}} := \sqrt{((y, f), (y, f))_{\mathcal{H}}}$ .

**Remark 1.** *Endowed with the norm  $\|(y, \partial_t y)\|_{L^\infty(0,T;\mathbf{V})} := \|y\|_{L^\infty(0,T;H_0^1(\Omega))} + \|\partial_t y\|_{L^\infty(0,T;L^2(\Omega))}$ , the  
space  $\mathcal{C}([0, T]; H_0^1(\Omega)) \cap \mathcal{C}^1([0, T]; L^2(\Omega))$  is a Banach space and  $\mathcal{H} \hookrightarrow (\mathcal{C}([0, T]; H_0^1(\Omega)) \cap \mathcal{C}^1([0, T]; L^2(\Omega))) \times$   
 $L^2(Q_T)$  continuously. Indeed, if  $(y, f) \in \mathcal{H}$ , we get from [30, Lemme 3.6, p.39] that*

$$\|(y, \partial_t y)\|_{L^\infty(0,T;\mathbf{V})} \leq C \left( \|\partial_{tt}y - \Delta y\|_{L^2(Q_T)} + \|(y(\cdot, 0), \partial_t y(\cdot, 0))\|_{\mathbf{V}} \right)$$

19 from which we deduce that  $\|(y, \partial_t y)\|_{L^\infty(0,T;\mathbf{V})} + \|f\|_{L^2(Q_T)} \leq C \|(y, f)\|_{\mathcal{H}}$ .

20 Let  $(u_0, u_1), (z_0, z_1) \in \mathbf{V}$ . We define the non-empty subspaces of  $\mathcal{H}$

$$\begin{aligned} \mathcal{A} &= \left\{ (y, f) \in \mathcal{H} \mid (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1), (y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1) \right\}, \\ \mathcal{A}_0 &= \left\{ (y, f) \in \mathcal{H} \mid (y(\cdot, 0), \partial_t y(\cdot, 0)) = (0, 0), (y(\cdot, T), \partial_t y(\cdot, T)) = (0, 0) \right\}. \end{aligned}$$

21 Remark that  $(0, 0) \in \mathcal{A}_0$  while  $\mathcal{A}$  contains the controlled pairs for the linear wave equation.

22 We consider the following non convex extremal problem:

$$\inf_{(y,f) \in \mathcal{A}} E(y, f), \quad E(y, f) := \frac{1}{2} \|\partial_{tt}y - \Delta y + g(y) - f1_\omega\|_2^2. \quad (6)$$

23 The functional  $E$  is well-defined in  $\mathcal{A}$ . Precisely,

24 **Lemma 1.** *There exists a positive constant  $C > 0$  such that  $E(y, f) \leq C(1 + \|(y, f)\|_{\mathcal{H}}^3)$  for any  $(y, f) \in \mathcal{A}$ .*

*Proof.* A *a priori* estimate for the linear wave equation reads as

$$\|(y, \partial_t y)\|_{L^\infty(0,T;\mathbf{V})}^2 \leq C \left( \|\partial_{tt} y - \Delta y\|_2^2 + \|(u_0, u_1)\|_{\mathbf{V}}^2 \right)$$

for any  $y$  such that  $(y, f) \in \mathcal{A}$ . Using that  $|g(r)| \leq C(1+|r|) \log(2+|r|)$  for every  $r \in \mathbb{R}$  and some  $C > 0$ , we infer that

$$\begin{aligned} \|g(y)\|_2^2 &\leq C^2 \int_{Q_T} \left( (1+|y|) \log(2+|y|) \right)^2 \\ &\leq C^2 \int_{Q_T} (1+|y|)^3 \leq C^2 (|Q_T|^3 + \|y\|_{L^3(Q_T)}^3) \\ &\leq C^2 (|Q_T|^3 + \|y\|_{L^\infty(0,T;H_0^1(\Omega))}^3) \end{aligned}$$

1 for which we get  $E(y, f) \leq C(\|\partial_{tt} y - \Delta y\|_2^2 + \|f\|_{2,q_T}^2 + |Q_T|^3 + \|y\|_{L^\infty(0,T;H_0^1(\Omega))}^3)$  leading to the result.  $\square$

2 Within the hypotheses of Theorem 1, the infimum of the functional of  $E$  is zero and is reached by  
3 at least one pair  $(y, f) \in \mathcal{A}$ , solution of (1) and satisfying  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$ . Conversely, any  
4 pair  $(y, f) \in \mathcal{A}$  for which  $E(y, f)$  vanishes is solution of (1). In spite of the lack of convexity of the  
5 functional  $E$ , we are going to construct a minimizing sequence which always converges to a zero of  $E$ . In  
6 this respect, we formally look, for any  $(y, f) \in \mathcal{A}$ , for a pair  $(Y^1, F^1) \in \mathcal{A}_0$  solution of

$$\begin{cases} \partial_{tt} Y^1 - \Delta Y^1 + g'(y) \cdot Y^1 = F^1 1_\omega + (\partial_{tt} y - \Delta y + g(y) - f 1_\omega), & \text{in } Q_T, \\ Y^1 = 0, & \text{on } \Sigma_T, \\ (Y^1(\cdot, 0), \partial_t Y^1(\cdot, 0)) = (0, 0), & \text{in } \Omega. \end{cases} \quad (7)$$

7 Since  $(Y^1, F^1)$  belongs to  $\mathcal{A}_0$ ,  $F^1$  is a null control for  $Y^1$ . Among the controls of this linear equation,  
8 we select the control of minimal  $L^2(q_T)$  norm. In the sequel, we shall call the corresponding solution  
9  $(Y^1, F^1) \in \mathcal{A}_0$  the solution of *minimal control norm*. We have the following property.

**Proposition 1.** For any  $(y, f) \in \mathcal{A}$ , there exists a pair  $(Y^1, F^1) \in \mathcal{A}_0$  solution of (7). Moreover, the pair  $(Y^1, F^1)$  of minimal control norm satisfies the following estimates :

$$\|(Y^1, \partial_t Y^1)\|_{L^\infty(0,T;\mathbf{V})} + \|F^1\|_{2,q_T} \leq C e^{C\|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \sqrt{E(y, f)}, \quad (8)$$

10 and

$$\|(Y^1, F^1)\|_{\mathcal{H}} \leq C(1 + \|g'(y)\|_{L^\infty(0,T;L^3(\Omega))}) e^{C\|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \sqrt{E(y, f)} \quad (9)$$

for some positive constant  $C > 0$ .

11 *Proof.* The first estimate follows Proposition 7 and the equality  $\|\partial_{tt} y - \Delta y + g(y) - f 1_\omega\|_2 = \sqrt{2E(y, f)}$ .

12 The second one follows from

$$\begin{aligned} \|(Y^1, F^1)\|_{\mathcal{H}} &\leq \|\partial_{tt} Y^1 - \Delta Y^1\|_2 + \|Y^1\|_2 + \|F^1\|_{2,q_T} + \|Y^1(\cdot, 0), \partial_t Y^1(\cdot, 0)\|_{\mathbf{V}} \\ &\leq \|Y^1\|_2 + \|g'(y)Y^1\|_2 + 2\|F^1\|_{2,q_T} + \sqrt{2}\sqrt{E(y, f)} \\ &\leq C(1 + \|g'(y)\|_{L^\infty(0,T;L^3(\Omega))}) e^{C\|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \sqrt{E(y, f)} \end{aligned}$$

13 using that  $\|g'(y)Y^1\|_2^2 \leq \int_0^T \|g'(y)\|_{L^3(\Omega)}^2 \|Y^1\|_{L^6(\Omega)}^2 \leq C\|g'(y)\|_{L^\infty(0,T;L^3(\Omega))}^2 \|Y^1\|_{L^\infty(0,T;H_0^1(\Omega))}^2$ .  $\square$

## 14 2.2 Main properties of the functional $E$

15 Given any  $s \in (0, 1]$ , we introduce for any  $g \in C^1(\mathbb{R})$  the following hypothesis:

$$1 \quad (\overline{\mathbf{H}}_s) \quad [g']_s := \sup_{\substack{a, b \in \mathbb{R} \\ a \neq b}} \frac{|g'(a) - g'(b)|}{|a - b|^s} < +\infty$$

2 meaning that  $g'$  is uniformly Hölder continuous with exponent  $s$ . In particular,  $g$  satisfies  $(\overline{\mathbf{H}}_1)$  if and  
3 only if  $g'$  is Lipschitz continuous (in this case,  $g'$  is almost everywhere differentiable and  $g'' \in L^\infty(\mathbb{R})$ ,  
4 and we have  $[g']_s \leq \|g''\|_\infty$ ).

5 The interest of the pair  $(Y^1, F^1) \in \mathcal{A}_0$  lies in the following result.

**Proposition 2.** *Assume that  $g$  satisfies  $(\overline{\mathbf{H}}_s)$  for some  $s \in (0, 1]$ . Let  $(y, f) \in \mathcal{A}$  and let  $(Y^1, F^1) \in \mathcal{A}_0$   
6 be a solution of (7). Then the derivative of  $E$  at the point  $(y, f) \in \mathcal{A}$  along the direction  $(Y^1, F^1)$  satisfies*

$$E'(y, f) \cdot (Y^1, F^1) = 2E(y, f). \quad (10)$$

*Proof.* We check that for all  $(Y, F) \in \mathcal{A}_0$  the functional  $E$  is differentiable at the point  $(y, f) \in \mathcal{A}$  along  
the direction  $(Y, F) \in \mathcal{A}_0$ . For any  $\lambda \in \mathbb{R}$ , simple computations lead to the equality

$$E(y + \lambda Y, f + \lambda F) = E(y, f) + \lambda E'(y, f) \cdot (Y, F) + h((y, f), \lambda(Y, F))$$

7 with

$$E'(y, f) \cdot (Y, F) := (\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega, \partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega)_2 \quad (11)$$

and

$$\begin{aligned} h((y, f), \lambda(Y, F)) &:= \frac{\lambda^2}{2} (\partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega, \partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega)_2 \\ &\quad + \lambda (\partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega, l(y, \lambda Y))_2 \\ &\quad + (\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega, l(y, \lambda Y)) + \frac{1}{2} (l(y, \lambda Y), l(y, \lambda Y)) \end{aligned}$$

8 where  $l(y, \lambda Y) := g(y + \lambda Y) - g(y) - \lambda g'(y)Y$ . The application  $(Y, F) \rightarrow E'(y, f) \cdot (Y, F)$  is linear and  
9 continuous from  $\mathcal{A}_0$  to  $\mathbb{R}$  as it satisfies

$$\begin{aligned} |E'(y, f) \cdot (Y, F)| &\leq \|\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega\|_2 \|\partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega\|_2 \\ &\leq \sqrt{2E(y, f)} \left( \|\partial_{tt}Y - \Delta Y\|_2 + \|g'(y)\|_{L^\infty(0, T; L^3(\Omega))} \|Y^1\|_{L^\infty(0, T; H_0^1(\Omega))} + \|F\|_{2, q_T} \right) \\ &\leq \sqrt{2E(y, f)} \max(1, \|g'(y)\|_{L^\infty(0, T; L^3(\Omega))}) \|(Y, F)\|_{\mathcal{H}}. \end{aligned} \quad (12)$$

Similarly, for all  $\lambda \in \mathbb{R}^*$ ,

$$\begin{aligned} \left| \frac{1}{\lambda} h((y, f), \lambda(Y, F)) \right| &\leq \frac{|\lambda|}{2} \|\partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega\|_2^2 \\ &\quad + \left( |\lambda| \|\partial_{tt}Y - \Delta Y + g'(y)Y - F \mathbf{1}_\omega\|_2 + \sqrt{2E(y, f)} + \frac{1}{2} \|l(y, \lambda Y)\|_2 \right) \frac{1}{|\lambda|} \|l(y, \lambda Y)\|_2. \end{aligned}$$

For any  $(x, y) \in \mathbb{R}^2$  and  $\lambda \in \mathbb{R}$ , we then write  $g(x + \lambda y) - g(x) = \int_0^\lambda y g'(x + \xi y) d\xi$  leading to

$$\begin{aligned} |g(x + \lambda y) - g(x) - \lambda g'(x)y| &\leq \left| \int_0^\lambda |y| |g'(x + \xi y) - g'(x)| d\xi \right| \\ &\leq \left| \int_0^\lambda |y|^{1+s} |\xi|^s \frac{|g'(x + \xi y) - g'(x)|}{|\xi y|^s} d\xi \right| \leq [g']_s |y|^{1+s} \frac{|\lambda|^{1+s}}{1+s}. \end{aligned}$$

10 It follows that  $|l(y, \lambda Y)| = |g(y + \lambda Y) - g(y) - \lambda g'(y)Y| \leq [g']_s \frac{|\lambda|^{1+s}}{1+s} |Y|^{1+s}$  and

$$\frac{1}{|\lambda|} \|l(y, \lambda Y)\|_2 \leq [g']_s \frac{|\lambda|^s}{1+s} \| |Y|^{1+s} \|_2. \quad (13)$$

11 But  $\| |Y|^{1+s} \|_2^2 = \|Y\|_{2^{(s+1)}}^{2(s+1)} \leq C \|Y\|_{L^\infty(0, T; L^4(\Omega))}^{2(s+1)}$ ; consequently, for  $s > 0$ ,  $\frac{1}{|\lambda|} \|l(y, \lambda Y)\|_2 \rightarrow 0$  as  $\lambda \rightarrow 0$   
12 and  $|h((y, f), \lambda(Y, F))| = o(\lambda)$ . Eventually, the equality (10) follows from the definition of the pair  
13  $(Y^1, F^1)$  given in (7).  $\square$



1 Remark that from the equality (11), the derivative  $E'(y, f)$  is independent of  $(Y, F)$ . We can then  
 2 define the norm  $\|E'(y, f)\|_{\mathcal{A}'_0} := \sup_{(Y, F) \in \mathcal{A}_0 \setminus \{0\}} \frac{E'(y, f) \cdot (Y, F)}{\|(Y, F)\|_{\mathcal{H}}}$  associated with  $\mathcal{A}'_0$ , the topological dual of  
 3  $\mathcal{A}_0$ .

4 Combining the equality (10) and the inequality (8), we deduce the following estimate of  $E(y, f)$  in  
 5 term of the norm of  $E'(y, f)$ .

**Proposition 3.** *For any  $(y, f) \in \mathcal{A}$ , the following inequalities hold true:*

$$\begin{aligned} \frac{1}{\sqrt{2} \max(1, \|g'(y)\|_{L^\infty(0, T; L^3(\Omega))})} \|E'(y, f)\|_{\mathcal{A}'_0} &\leq \sqrt{E(y, f)} \\ &\leq \frac{1}{\sqrt{2}} C \left(1 + \|g'(y)\|_{L^\infty(0, T; L^3(\Omega))}\right) e^{C\|g'(y)\|_{L^\infty(0, T; L^d(\Omega))}^2} \|E'(y, f)\|_{\mathcal{A}'_0} \end{aligned} \quad (14)$$

where  $C$  is the positive constant from Proposition 1.

*Proof.* (10) rewrites  $E(y, f) = \frac{1}{2} E'(y, f) \cdot (Y^1, F^1)$  where  $(Y^1, F^1) \in \mathcal{A}_0$  is solution of (7) and therefore, with (9)

$$\begin{aligned} E(y, f) &\leq \frac{1}{2} \|E'(y, f)\|_{\mathcal{A}'_0} \|(Y^1, F^1)\|_{\mathcal{A}_0} \\ &\leq \frac{1}{2} C (1 + \|g'(y)\|_{L^\infty(0, T; L^3(\Omega))}) e^{C\|g'(y)\|_{L^\infty(0, T; L^d(\Omega))}^2} \|E'(y, f)\|_{\mathcal{A}'_0} \sqrt{E(y, f)}. \end{aligned}$$

7 On the other hand, the left inequality follows from (12). □

8 Consequently, any *critical point*  $(y, f) \in \mathcal{A}$  of  $E$  (i.e.,  $E'(y, f)$  vanishes) such that  $\|g'(y)\|_{L^\infty(0, T; L^3(\Omega))}$   
 9 is finite is a zero for  $E$ , a pair solution of the controllability problem. In other words, any sequence  
 10  $(y_k, f_k)_{k>0}$  satisfying  $\|E'(y_k, f_k)\|_{\mathcal{A}'_0} \rightarrow 0$  as  $k \rightarrow \infty$  and for which  $\|g'(y_k)\|_{L^\infty(0, T; L^3(\Omega))}$  is uniformly  
 11 bounded is such that  $E(y_k, f_k) \rightarrow 0$  as  $k \rightarrow \infty$ . We insist that this property does not imply the convexity  
 12 of the functional  $E$  (and *a fortiori* the strict convexity of  $E$ , which actually does not hold here in view of  
 13 the multiple zeros for  $E$ ) but show that a minimizing sequence for  $E$  can not be stuck in a local minimum.

14 On the other hand, the left inequality indicates the functional  $E$  is flat around its zero set. As a  
 15 consequence, gradient-based minimizing sequences may achieve a low speed of convergence (we refer to  
 16 [35] and also [28] devoted to the Navier-Stokes equation where this phenomenon is observed).

17

18 We end this section with the following fundamental estimate.

**Lemma 2.** *Assume that  $g$  satisfies  $(\overline{\mathbf{H}}_s)$  for some  $s \in [0, 1]$ . For any  $(y, f) \in \mathcal{A}$ , let  $(Y^1, F^1) \in \mathcal{A}_0$  be  
 defined by (7). For any  $\lambda \in \mathbb{R}$  the following estimate holds*

$$E((y, f) - \lambda(Y^1, F^1)) \leq E(y, f) \left( |1 - \lambda| + |\lambda|^{1+s} c(y) E(y, f)^{s/2} \right)^2 \quad (15)$$

with  $c(y) := \frac{C}{(1+s)\sqrt{2}} [g']_s d(y)^{1+s}$  and  $d(y) := C e^{C\|g'(y)\|_{L^\infty(0, T; L^d(\Omega))}^2}$ .

20 *Proof.* Estimate (13) applied with  $Y = Y^1$  reads

$$\|l(y, \lambda Y^1)\|_2 \leq [g']_s \frac{|\lambda|^{1+s}}{1+s} \||Y^1|^{1+s}\|_2. \quad (16)$$

21 But  $\||Y^1|^{1+s}\|_2^2 = \|Y^1\|_{\frac{2(s+1)}{2(s+1)}}^{2(s+1)} \leq C \|Y^1\|_{L^\infty(0, T; H_0^1(\Omega))}^{2(s+1)}$  which together with (8) lead to

$$\||Y^1|^{1+s}\|_2 \leq C \left( C e^{C\|g'(y)\|_{L^\infty(0, T; L^d(\Omega))}^2} \right)^{1+s} E(y, f)^{\frac{1+s}{2}}. \quad (17)$$

1 Eventually, we write

$$\begin{aligned}
& 2E((y, f) - \lambda(Y^1, F^1)) \\
&= \left\| (\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega) - \lambda(\partial_{tt}Y^1 - \Delta Y^1 + g'(y)Y^1 - F \mathbf{1}_\omega) + l(y, -\lambda Y^1) \right\|_2^2 \\
&= \left\| (1 - \lambda)(\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega) + l(y, -\lambda Y^1) \right\|_2^2 \\
&\leq \left( \left\| (1 - \lambda)(\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega) \right\|_2 + \left\| l(y, -\lambda Y^1) \right\|_2 \right)^2 \\
&\leq 2 \left( |1 - \lambda| \sqrt{E(y, f)} + [g']_s \frac{|\lambda|^{1+s}}{1+s} \left\| |Y^1|^{1+s} \right\|_2 \right)^2 \\
&\leq 2 \left( |1 - \lambda| \sqrt{E(y, f)} + [g']_s \frac{|\lambda|^{1+s}}{1+s} C \left( C e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}} \right)^{1+s} E(y, f)^{\frac{1+s}{2}} \right)^2
\end{aligned} \tag{18}$$

2 and we get the result.  $\square$

### 3 Convergence of a minimizing sequence for $E$

4 Equality (10) shows that  $-(Y^1, F^1)$  given by the solution of (7) is a descent direction for  $E$ . Therefore,  
5 we can define, for any fixed  $m \geq 1$ , a minimizing sequence  $(y_k, f_k)_{k>0} \in \mathcal{A}$  as follows:

$$\begin{cases} (y_0, f_0) \in \mathcal{A}, \\ (y_{k+1}, f_{k+1}) = (y_k, f_k) - \lambda_k(Y_k^1, F_k^1), \quad k \in \mathbb{N}, \\ \lambda_k = \operatorname{argmin}_{\lambda \in [0, m]} E((y_k, f_k) - \lambda(Y_k^1, F_k^1)), \end{cases} \tag{19}$$

6 where  $(Y_k^1, F_k^1) \in \mathcal{A}_0$  is the solution of minimal control norm of

$$\begin{cases} \partial_{tt}Y_k^1 - \Delta Y_k^1 + g'(y_k) \cdot Y_k^1 = F_k^1 \mathbf{1}_\omega + (\partial_{tt}y_k - \Delta y_k + g(y_k) - f_k \mathbf{1}_\omega), & \text{in } Q_T, \\ Y_k^1 = 0, & \text{on } \Sigma_T, \\ (Y_k^1(\cdot, 0), \partial_t Y_k^1(\cdot, 0)) = (0, 0), & \text{in } \Omega. \end{cases} \tag{20}$$

7 The real  $m \geq 1$  is arbitrarily fixed and is introduced in order to keep the sequence  $(\lambda_k)_{k \in \mathbb{N}}$  bounded.

8 Given any  $s \in (0, 1]$ , we set

$$\beta^*(s) := \sqrt{\frac{s}{2C(2s+1)}} \tag{21}$$

9 where  $C > 0$ , only depending on  $\Omega$  and  $T$ , is the constant appearing in Proposition 7. In this section, we  
10 prove our main result.

**Theorem 2.** *Assume that  $g'$  satisfies  $(\overline{\mathbf{H}}_s)$  for some  $s \in (0, 1]$  and*

**(H<sub>2</sub>)** *There exists  $\alpha \geq 0$  and  $\beta \in [0, \beta^*(s))$  such that  $|g'(r)| \leq \alpha + \beta \ln^{1/2}(1 + |r|)$  for every  $r \in \mathbb{R}$ .*

11 *Then, for any  $(y_0, f_0) \in \mathcal{A}$ , the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  defined by (19) strongly converges to a pair  $(\bar{y}, \bar{f}) \in \mathcal{A}$  satisfying (1) and the condition  $(y(\cdot, T), y_t(\cdot, T)) = (z_0, z_1)$ , for all  $(u_0, u_1), (z_0, z_1) \in \mathbf{V}$ . Moreover, the convergence is at least linear and is at least of order  $1 + s$  after a finite number of iterations*

12 This result remains true for  $s = 0$  (remark that  $[g']_0 < \infty$  is equivalent to  $g' \in L^\infty(\mathbb{R})$ ) under the  
13 additional smallness assumption on  $\|g'\|_{L^\infty(\mathbb{R})}$ :

14 **(H<sub>3</sub>)**  $\sqrt{2C} \|g'\|_\infty e^{C \|g'\|_\infty^2 |\Omega|^{2/d}} < 1$

with  $C$  the constant appearing in Proposition 7. We refer to [36] for the details devoted in the case  $d = 1$ .

The proof of 2 consists in showing that the decreasing sequence  $(E(y_k, f_k))_{k \in \mathbb{N}}$  converges to zero. In view of (14), this property is related to the uniform property of the observability constant  $e^{C\|g'(y_k)\|_{L^\infty(0,T;L^d(\Omega))}^2}$  with respect to  $k$ . In order to fix some notations and the main ideas of the proof of Theorem 2, we first prove in Section 3.1 the convergence of the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  under the stronger condition that  $g' \in L^\infty(\mathbb{R})$ , sufficient to ensure the boundedness of the sequence  $(e^{C\|g'(y_k)\|_{L^\infty(0,T;L^d(\Omega))}^2})_{k \in \mathbb{N}}$ . Then, in Section 3.2, we prove Theorem 2 by showing that under the assumption  $(\mathbf{H}_2)$ , the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  is actually bounded in  $\mathcal{A}$ . This implies the same property for the real sequence  $e^{C\|g'(y_k)\|_{L^\infty(0,T;L^d(\Omega))}^2}$ , and then the announced convergence.

### 3.1 Proof of the convergence under the additional assumption $g' \in L^\infty(\mathbb{R})$

We establish the following preliminary result which coincides with Theorem 2 in the simpler case  $\beta = 0$ .

**Proposition 4.** *Assume that  $g'$  satisfies  $(\overline{\mathbf{H}}_s)$  for some  $s \in (0, 1]$  and that  $g' \in L^\infty(\mathbb{R})$ . For any  $(y_0, f_0) \in \mathcal{A}$ , the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  defined by (19) strongly converges to a pair  $(\overline{y}, \overline{f}) \in \mathcal{A}$  satisfying (1) and the condition  $(y(\cdot, T), y_t(\cdot, T)) = (z_0, z_1)$ . Moreover, the convergence is at least linear and is at least of order  $1 + s$  after a finite number of iterations.*

Proceeding as in [26, 36], Proposition 4 follows from the following lemma.

**Lemma 3.** *Under the hypotheses of Proposition 4, for any  $(y_0, f_0) \in \mathcal{A}$ , there exists a  $k_0 \in \mathbb{N}$  such that the sequence  $(E(y_k, f_k))_{k \geq k_0}$  tends to 0 as  $k \rightarrow \infty$  with at least a rate  $s + 1$ .*

*Proof.* Since  $g' \in L^\infty(\mathbb{R})$ , the nonnegative constant  $c(y_k)$  in (15) is uniformly bounded w.r.t.  $k$ : we introduce the real  $c > 0$  as follows

$$c(y_k) \leq c := \frac{C}{(1+s)\sqrt{2}} [g']_s \left( C e^{C\|g'\|_\infty^2 |\Omega|^{2/d}} \right)^{1+s}, \quad \forall k \in \mathbb{N}. \quad (22)$$

$|\Omega|$  denotes the measure of the domain  $\Omega$ . For any  $(y_k, f_k) \in \mathcal{A}$ , let us then denote the real function  $p_k$  by

$$p_k(\lambda) := |1 - \lambda| + \lambda^{1+s} c E(y_k, f_k)^{s/2}, \quad \forall \lambda \in [0, m].$$

Lemma 2 with  $(y, f) = (y_k, f_k)$  then allows to write that

$$\sqrt{E(y_{k+1}, f_{k+1})} = \min_{\lambda \in [0, m]} \sqrt{E((y_k, f_k) - \lambda(Y_k^1, F_k^1))} \leq p_k(\widetilde{\lambda}_k) \sqrt{E(y_k, f_k)} \quad (23)$$

with  $p_k(\widetilde{\lambda}_k) := \min_{\lambda \in [0, m]} p_k(\lambda)$ . Assume first that  $s > 0$ . The optimal  $\widetilde{\lambda}_k$  is given by

$$\widetilde{\lambda}_k := \begin{cases} \frac{1}{(1+s)^{1/s} c^{1/s} \sqrt{E(y_k, f_k)}}, & \text{if } (1+s)^{1/s} c^{1/s} \sqrt{E(y_k, f_k)} \geq 1, \\ 1, & \text{if } (1+s)^{1/s} c^{1/s} \sqrt{E(y_k, f_k)} < 1 \end{cases}$$

leading to

$$p_k(\widetilde{\lambda}_k) := \begin{cases} 1 - \frac{s}{(1+s)^{\frac{1}{s}+1} c^{1/s} \sqrt{E(y_k, f_k)}}, & \text{if } (1+s)^{1/s} c^{1/s} \sqrt{E(y_k, f_k)} \geq 1, \\ c E(y_k, f_k)^{s/2}, & \text{if } (1+s)^{1/s} c^{1/s} \sqrt{E(y_k, f_k)} < 1. \end{cases} \quad (24)$$

Accordingly, we may distinguish two cases :

• If  $(1+s)^{1/s}c^{1/s}\sqrt{E(y_0, f_0)} < 1$ , then  $c^{1/s}\sqrt{E(y_0, f_0)} < 1$ , and thus  $c^{1/s}\sqrt{E(y_k, f_k)} < 1$  for all  $k \in \mathbb{N}$  since the sequence  $(E(y_k, f_k))_{k \in \mathbb{N}}$  is decreasing. Hence (23) implies that

$$c^{1/s}\sqrt{E(y_{k+1}, f_{k+1})} \leq (c^{1/s}\sqrt{E(y_k, f_k)})^{1+s} \quad \forall k \in \mathbb{N}.$$

1 It follows that  $c^{1/s}\sqrt{E(y_k, f_k)} \rightarrow 0$  as  $k \rightarrow \infty$  with a rate equal to  $1+s$ .

2 • If  $(1+s)^{1/s}c^{1/s}\sqrt{E(y_0, f_0)} \geq 1$  then we check that the set  $I := \{k \in \mathbb{N}, (1+s)^{1/s}c^{1/s}\sqrt{E(y_k, f_k)} \geq 1\}$   
3 is a finite subset of  $\mathbb{N}$ ; indeed, for all  $k \in I$ , (23) implies that

$$c^{1/s}\sqrt{E(y_{k+1}, f_{k+1})} \leq \left(1 - \frac{s}{(1+s)^{\frac{1}{s}+1}} \frac{1}{c^{1/s}\sqrt{E(y_k, f_k)}}\right) c^{1/s}\sqrt{E(y_k, f_k)} = c^{1/s}\sqrt{E(y_k, f_k)} - \frac{s}{(1+s)^{\frac{1}{s}+1}} \quad (25)$$

4 and the strict decrease of the sequence  $(c^{1/s}\sqrt{E(y_k, f_k)})_{k \in I}$ . Thus there exists  $k_0 \in \mathbb{N}$  such that for all  
5  $k \geq k_0$ ,  $(1+s)^{1/s}c^{1/s}\sqrt{E(y_k, f_k)} < 1$ , that is  $I$  is a finite subset of  $\mathbb{N}$ . Arguing as in the first case, it  
6 follows that  $\sqrt{E(y_k, f_k)} \rightarrow 0$  as  $k \rightarrow \infty$ .

7 It follows in particular from (24) that the sequence  $(p_k(\tilde{\lambda}_k))_{k \in \mathbb{N}}$  decreases as well.

8 □

*Proof.* (of Proposition 4) In view of (9), we write

$$(1 + \|g'(y)\|_{L^\infty(0,T;L^3(\Omega))}) e^{C\|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \leq (1 + \|g'\|_\infty |\Omega|^{1/3}) e^{C\|g'\|_\infty^2 |\Omega|^{2/d}} \leq e^{2C\|g'\|_\infty^2 |\Omega|^{2/d}}$$

9 using that  $(1+u)e^{u^2} \leq e^{2u^2}$  for all  $u \in \mathbb{R}^+$ . It follows that

$$\sum_{n=0}^k |\lambda_n| \| (Y_n^1, F_n^1) \|_{\mathcal{H}} \leq m C e^{C\|g'\|_\infty^2 |\Omega|^{2/d}} \sum_{n=0}^k \sqrt{E(y_n, f_n)}. \quad (26)$$

10 Using that  $p_n(\tilde{\lambda}_n) \leq p_0(\tilde{\lambda}_0)$  for all  $n \geq 0$ , we can write for  $n > 0$ ,

$$\sqrt{E(y_n, f_n)} \leq p_{n-1}(\tilde{\lambda}_{n-1}) \sqrt{E(y_{n-1}, f_{n-1})} \leq p_0(\tilde{\lambda}_0) \sqrt{E(y_{n-1}, f_{n-1})} \leq (p_0(\tilde{\lambda}_0))^n \sqrt{E(y_0, f_0)}. \quad (27)$$

Then, using that  $p_0(\tilde{\lambda}_0) = \min_{\lambda \in [0, m]} p_0(\lambda) < 1$  (since  $p_0(0) = 1$  and  $p'_0(0) < 0$ ), we finally obtain the uniform estimate

$$\sum_{n=0}^k |\lambda_n| \| (Y_n^1, F_n^1) \|_{\mathcal{H}} \leq m C e^{C\|g'\|_\infty^2 |\Omega|^{2/d}} \frac{\sqrt{E(y_0, f_0)}}{1 - p_0(\tilde{\lambda}_0)}$$

11 for which we deduce (since  $\mathcal{H}$  is a complete space) that the serie  $\sum_{n \geq 0} \lambda_n (Y_n^1, F_n^1)$  converges in  $\mathcal{A}_0$ .

12 Writing from (19) that  $(y_{k+1}, f_{k+1}) = (y_0, f_0) - \sum_{n=0}^k \lambda_n (Y_n^1, F_n^1)$ , we conclude that  $(y_k, f_k)$  strongly  
13 converges in  $\mathcal{A}$  to  $(\bar{y}, \bar{f}) := (y_0, f_0) + \sum_{n \geq 0} \lambda_n (Y_n^1, F_n^1)$ .

14 Let us now pass to the limit in (20). We write that  $\|g(y_k) - g(\bar{y})\|_{L^2(Q_T)} \leq \|g'\|_\infty \|y_k - \bar{y}\|_{L^2(Q_T)}$  and  
15 thus  $g(y_k) \rightarrow g(\bar{y})$  in  $L^2(Q_T)$ . Moreover,  $(g'(y_k))_{k \in \mathbb{N}}$  is a bounded sequence of  $L^2(Q_T)$  since  $g' \in L^\infty$ .  
16 Then, using that  $(Y_k^1, F_k^1)$  goes to zero as  $k \rightarrow \infty$  in  $\mathcal{A}_0$ , we pass to the limit in (20) and get that  
17  $(\bar{y}, \bar{f}) \in \mathcal{A}$  solves

$$\begin{cases} \partial_{tt}\bar{y} - \Delta\bar{y} + g(\bar{y}) = \bar{f}1_\omega, & \text{in } Q_T, \\ \bar{y} = 0, & \text{on } \Sigma_T, \\ (\bar{y}(\cdot, 0), \partial_t\bar{y}(\cdot, 0)) = (y_0, y_1), & \text{in } \Omega. \end{cases} \quad (28)$$

1 Since the limit  $(\bar{y}, \bar{f})$  belongs to  $\mathcal{A}$ , we have that  $(\bar{y}(\cdot, T), \bar{y}_t(\cdot, T)) = (z_0, z_1)$  in  $\Omega$ . Moreover, for all  $k > 0$

$$\begin{aligned} \|(\bar{y}, \bar{f}) - (y_k, f_k)\|_{\mathcal{H}} &= \left\| \sum_{p=k+1}^{\infty} \lambda_p (Y_p^1, F_p^1) \right\|_{\mathcal{H}} \leq m \sum_{p=k+1}^{\infty} \|(Y_p^1, F_p^1)\|_{\mathcal{H}} \\ &\leq m C \sum_{p=k+1}^{\infty} \sqrt{E(y_p, f_p)} \leq m C \sum_{p=k+1}^{\infty} p_0 (\tilde{\lambda}_0)^{p-k} \sqrt{E(y_k, f_k)} \\ &\leq m C \frac{p_0 (\tilde{\lambda}_0)}{1 - p_0 (\tilde{\lambda}_0)} \sqrt{E(y_k, f_k)} \end{aligned} \quad (29)$$

2 and conclude from Lemma 3 the convergence of order at least  $1 + s$  after a finite number of iterates.  $\square$

3 **Remark 2.** In particular, along the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  defined by (19), (29) is a kind of coercivity  
4 property for the functional  $E$ . We emphasize, in view of the non uniqueness of the zeros of  $E$ , that an  
5 estimate (similar to (29)) of the form  $\|(\bar{y}, \bar{f}) - (y, f)\|_{\mathcal{H}} \leq C \sqrt{E(y, f)}$  does not hold for all  $(y, f) \in \mathcal{A}$ .  
6 We also emphasize that the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  and its limits  $(\bar{y}, \bar{f})$  are uniquely determined from the  
7 initialization  $(y_0, f_0) \in \mathcal{A}$  and from our selection criterion for the control  $F^1$ .

**Remark 3.** Estimate (26) implies the uniform estimate on the sequence  $(\|(y_k, f_k)\|_{\mathcal{H}})_{k \in \mathbb{N}}$ :

$$\begin{aligned} \|(y_k, f_k)\|_{\mathcal{H}} &\leq \|(y_0, f_0)\|_{\mathcal{H}} + m C e^{C \|g'\|_{\infty}^2 |\Omega|^{2/d}} \sum_{n=0}^{k-1} \sqrt{E(y_n, f_n)} \\ &\leq \|(y_0, f_0)\|_{\mathcal{H}} + m C e^{C \|g'\|_{\infty}^2 |\Omega|^{2/d}} \frac{\sqrt{E(y_0, f_0)}}{1 - p_0 (\tilde{\lambda}_0)}. \end{aligned}$$

In particular, for the less favorable case for which  $(1 + s)^{1/s} c^{1/s} \sqrt{E(y_0, f_0)} \geq 1$ , we get  $\frac{\sqrt{E(y_0, f_0)}}{1 - p_0 (\tilde{\lambda}_0)} =$   
 $\frac{(1+s)^{\frac{1}{s}+1}}{s} c^{1/s} E(y_0, f_0)$ , (see (24)) leading to

$$\|(y_k, f_k)\|_{\mathcal{H}} \leq \|(y_0, f_0)\|_{\mathcal{H}} + m C e^{C \|g'\|_{\infty}^2 |\Omega|^{2/d}} \frac{(1 + s)^{\frac{1}{s}+1}}{s} c^{1/s} E(y_0, f_0),$$

8 and then, in view of (22), to the explicit estimate in term of the data

$$\|(y_k, f_k)\|_{\mathcal{H}} \leq \|(y_0, f_0)\|_{\mathcal{H}} + m \frac{(1 + s)}{s} \left( \frac{C \|g'\|_{\infty}}{\sqrt{2}} \right)^{1/s} \left( C e^{C \|g'\|_{\infty}^2 |\Omega|^{2/d}} \right)^{\frac{2s+1}{s}} E(y_0, f_0).$$

9 **Remark 4.** For  $s > 0$ , recalling that the constant  $c$  is defined in (22), if  $(1 + s)^{1/s} c^{1/s} \sqrt{E(y_0, f_0)} \geq 1$ ,  
10 inequality (25) implies that

$$c^{1/s} \sqrt{E(y_k, f_k)} \leq c^{1/s} \sqrt{E(y_0, f_0)} - k \frac{s}{(1 + s)^{\frac{1}{s}+1}}, \quad \forall k \in I.$$

Hence, the number of iteration  $k_0$  to achieve a rate  $1 + s$  is estimated as follows :

$$k_0 = \left\lfloor (1 + s) \left( c^{1/s} (1 + s)^{1/s} \sqrt{E(y_0, f_0)} \right) - \frac{1}{s} \right\rfloor + 1$$

11 where  $\lfloor \cdot \rfloor$  denotes the integer part. As expected, this number increases with  $\sqrt{E(y_0, f_0)}$  and  $\|g'\|_{\infty}$ . If  
12  $(1 + s)^{1/s} c^{1/s} \sqrt{E(y_0, f_0)} < 1$ , then  $k_0 = 0$ . In particular, as  $s \rightarrow 0^+$ ,  $k_0 \rightarrow \infty$  if  $c > 1$ , i.e. if  $(\mathbf{H}_3)$  does  
13 not hold.

14 The following convergence also holds, independently of the dimension of  $\Omega$ . We refer to [36, Section  
15 3, step 2] for the proof.

16 **Lemma 4.** Assume that  $g'$  satisfies  $(\bar{\mathbf{H}}_s)$  for some  $s \in (0, 1]$  and that  $g' \in L^\infty(\mathbb{R})$ . The sequence  $(\lambda_k)_{k > k_0}$   
defined in (19) converges to 1 as  $k \rightarrow \infty$  at least with order  $1 + s$ .

## 3.2 Proof of Theorem 2

In this section, we relax the condition  $g' \in L^\infty(\mathbb{R})$  and prove Theorem 3.2, for  $s > 0$  under the assumption  $(\mathbf{H}_2)$ . This assumption implies notably that  $|g(r)| \leq C(1 + |r|) \ln(2 + |r|)$  for every  $r \in \mathbb{R}$ , mentioned in the introduction to state the well-posedness of (1). The case  $\beta = 0$  corresponds to the case developed in the previous section, i.e.  $g' \in L^\infty(\mathbb{R})$ .

Within this more general framework, the difficulty is to have a uniform control with respect to  $k$  of the observability constant  $C e^{C \|g'(y_k)\|_{L^\infty(0,T;L^d(\Omega))}^2}$  appearing in the estimates for  $(Y_k^1, F_k^1)$ , see Proposition 1. In other terms, we have to show that the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  uniquely defined in (19) is uniformly bounded in  $\mathcal{A}$ , for any  $(y_0, f_0) \in \mathcal{A}$ .

The following intermediate result is crucial as it gives an estimate of the observability constant in term of an  $L^\infty(0, T, L^p(\Omega))$  norm of the state.

**Lemma 5.** *Let  $C > 0$ , only depending on  $\Omega$  and  $T$  be the constant appearing in Proposition 7. Assume that  $g$  satisfies the growth condition  $(\mathbf{H}_2)$  and  $2C\beta^2 \leq 1$ . Then for any  $(y, f) \in \mathcal{A}$ ,*

$$e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \leq 2C \max(1, e^{2C\alpha^2 |\Omega|^2}) \left( 1 + \frac{\|y\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\beta^2}$$

for any  $p^* \in \mathbb{N}^*$  with  $p^* < \infty$  if  $d = 2$  and  $p^* \leq 6$  if  $d = 3$ .

*Proof.* We use the following inequality (direct consequence of the inequality (3.8) in [29]):

$$e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \leq C \left( 1 + \sup_{t \in (0,T)} \int_{\Omega} e^{C|g'(y)|^2} \right), \quad \forall (y, f) \in \mathcal{A}. \quad (30)$$

Writing that  $|g'(y)|^2 \leq 2(\alpha^2 + \beta^2 \ln(1 + |y|))$ , we get that  $\int_{\Omega} e^{C|g'(y)|^2} \leq e^{2C\alpha^2} \int_{\Omega} (1 + |y|)^{2C\beta^2}$ . Assuming  $2C\beta^2 \leq p^*$ , Hölder inequality leads to

$$\begin{aligned} \int_{\Omega} e^{C|g'(y)|^2} &\leq e^{2C\alpha^2} \left( \int_{\Omega} (1 + |y|)^{p^*} \right)^{\frac{2C\beta^2}{p^*}} |\Omega|^{1 - \frac{2C\beta^2}{p^*}} \\ &\leq e^{2C\alpha^2} |\Omega| \left( 1 + \frac{\|y\|_{L^{p^*}(\Omega)}}{|\Omega|^{1/p^*}} \right)^{2C\beta^2}. \end{aligned}$$

It follows, by (30), that for every  $(y, f) \in \mathcal{A}$ ,

$$\begin{aligned} e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} &\leq C \left( 1 + e^{2C\alpha^2} |\Omega| \left( 1 + \frac{\|y\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\beta^2} \right) \\ &\leq C \max(1, e^{2C\alpha^2} |\Omega|) \left( 1 + \left( 1 + \frac{\|y\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\beta^2} \right) \\ &\leq 2^{2C\beta^2} C \max(1, e^{2C\alpha^2} |\Omega|) \left( 1 + \frac{\|y\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\beta^2} \end{aligned}$$

and the result.  $\square$

**Lemma 6.** *Assume that  $g$  satisfies the growth condition  $(\mathbf{H}_2)$  and  $2C\beta^2 \leq 1$ . For any  $(y, f) \in \mathcal{A}$ , the unique solution  $(Y^1, F^1) \in \mathcal{A}_0$  of (7) satisfies*

$$\|(Y^1, \partial_t Y^1)\|_{L^\infty(0,T;\mathbf{V})} + \|F^1\|_{2,q_T} \leq d(y) \sqrt{E(y, f)}$$

with  $d(y) := C_3(\alpha) \left( 1 + \frac{\|y\|_{L^\infty(0,T;L^1(\Omega))}}{|\Omega|} \right)^{2C\beta^2}$  and  $C_3(\alpha) := 2C \max(1, e^{2C\alpha^2} |\Omega|)$ .

1 *Proof.* Lemma 5 with  $p^* = 1$  and estimate (8) lead to the result.  $\square$

2 With these notations, the term  $c(y)$  in (15) rewrites as

$$c(y) = \frac{C}{(1+s)\sqrt{2}} [g']_s d(y)^{1+s}, \quad \forall (y, f) \in \mathcal{A}, \forall s \in (0, 1]. \quad (31)$$

3 *Proof.* (of Theorem 2) If the initialization  $(y_0, f_0) \in \mathcal{A}$  is such that  $E(y_0, f_0) = 0$ , then the sequence  
4  $(y_k, f_k)_{k \in \mathbb{N}}$  constant equal to  $(y_0, f_0)$  is convergent. We assume in the sequel that  $E(y_0, f_0) > 0$ .

5 We are going to prove that, for any  $\beta < \beta^*(s)$ , there exists a constant  $M > 0$  such that the sequence  
6  $(y_k)_{k \in \mathbb{N}}$  defined by (19) enjoys the uniform property

$$\|y_k\|_{L^\infty(0, T; L^1(\Omega))} \leq M, \quad \forall k \in \mathbb{N}. \quad (32)$$

7 The convergence of the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  in  $\mathcal{A}$  will then follow by proceeding as in Section 3.1. Remark  
8 preliminary that the assumption  $\beta < \beta^*(s)$  implies  $2C\beta^2 < \frac{s}{2s+1} \leq 1$  since  $s \in (0, 1]$ .

*Proof of the uniform property (32) for some  $M$  large enough-* As for  $n = 0$ , from any initialization  $(y_0, f_0)$   
chosen in  $\mathcal{A}$ , it suffices to take  $M$  larger than  $M_1 := \|y_0\|_{L^\infty(0, T; L^1(\Omega))}$ . We then proceed by induction  
and assume that, for some  $n \in \mathbb{N}$ ,  $\|y_k\|_{L^\infty(0, T; L^1(\Omega))} \leq M$  for all  $k \leq n$ . This implies in particular that,

$$d(y_k) \leq d_M(\beta) := C_3(\alpha) \left(1 + \frac{M}{|\Omega|}\right)^{2C\beta^2}, \quad \forall k \leq n$$

9 and then

$$c(y_k) \leq c_M(\beta) := \frac{C}{(1+s)\sqrt{2}} [g']_s d_M^{1+s}(\beta), \quad \forall k \leq n. \quad (33)$$

10 Then, we write that  $\|y_{n+1}\|_{L^\infty(0, T; L^1(\Omega))} \leq \|y_0\|_{L^\infty(0, T; L^1(\Omega))} + \sum_{k=0}^n \lambda_k \|Y_k^1\|_{L^\infty(0, T; L^1(\Omega))}$ . But, Lemma  
11 6 implies that  $\|Y_k^1\|_{L^\infty(0, T; L^1(\Omega))} \leq d_M(\beta) \sqrt{E(y_k, f_k)}$  for all  $k \leq n$  leading to

$$\|y_{n+1}\|_{L^\infty(0, T; L^1(\Omega))} \leq \|y_0\|_{L^\infty(0, T; L^1(\Omega))} + m d_M(\beta) \sum_{k=0}^n \sqrt{E(y_k, f_k)}. \quad (34)$$

12 Moreover, inequality (27) implies that  $\sum_{k=0}^n \sqrt{E(y_k, f_k)} \leq \frac{1}{1-p_0(\widetilde{\lambda}_0)} \sqrt{E(y_0, f_0)}$  where  $p_0(\widetilde{\lambda}_0)$  is given by  
13 (24) with  $c = c_M(\beta)$ .

14 Now, we take  $M$  large enough so that  $(1+s)^{1/s} c_M^{1/s}(\beta) \sqrt{E(y_0, f_0)} \geq 1$  i.e.

$$\left(\frac{C}{\sqrt{2}} [g']_s\right)^{1/s} C_3(\alpha)^{2/s} \left(1 + \frac{M}{|\Omega|}\right)^{\frac{4C\beta^2}{s}} \sqrt{E(y_0, f_0)} \geq 1. \quad (35)$$

15 Such  $M$  exists since  $\sqrt{E(y_0, f_0)} > 0$  is independent of  $M$  and since the left hand side is of order  
16  $\mathcal{O}(M^{\frac{4C\beta^2}{s}})$  with  $\frac{4C\beta^2}{s} > 0$ . We denote by  $M_2$  the smallest value of  $M$  such that (35) hold true.

Then, from (24), we get that  $p_0(\widetilde{\lambda}_0) = 1 - \frac{s}{(1+s)^{\frac{1}{s}+1}} \frac{1}{c_M^{1/s}(\beta) \sqrt{E(y_0, f_0)}}$  and therefore

$$\frac{1}{1-p_0(\widetilde{\lambda}_0)} = \frac{(1+s)^{\frac{1}{s}+1}}{s} c_M^{1/s}(\beta) \sqrt{E(y_0, f_0)}$$

so that  $\sum_{k=0}^n \sqrt{E(y_k, f_k)} \leq \frac{(1+s)^{\frac{1}{s}+1}}{s} c_M^{1/s}(\beta) E(y_0, f_0)$ . It follows from (34) that

$$\|y_{n+1}\|_{L^\infty(0, T; L^1(\Omega))} \leq \|y_0\|_{L^\infty(0, T; L^1(\Omega))} + m d_M(\beta) \frac{(1+s)^{\frac{1}{s}+1}}{s} c_M^{1/s}(\beta) E(y_0, f_0).$$

The definition of  $c_M(\beta)$  (see (33)) then gives

$$\begin{aligned} \|y_{n+1}\|_{L^\infty(0,T;L^1(\Omega))} &\leq \|y_0\|_{L^\infty(0,T;L^1(\Omega))} \\ &\quad + \frac{m(1+s)}{s} \left( \frac{C[g']_s}{\sqrt{2}} \right)^{1/s} \left( C_3(\alpha) \right)^{1+\frac{2}{s}} E(y_0, f_0) \left( 1 + \frac{M}{|\Omega|} \right)^{\frac{(2C\beta^2)(2s+1)}{s}}. \end{aligned}$$

1 Now, we take  $M > 0$  large enough so that the right hand side is bounded by  $M$ , i.e.

$$\|y_0\|_{L^\infty(0,T;L^1(\Omega))} + \frac{m(1+s)}{s} \left( \frac{C[g']_s}{\sqrt{2}} \right)^{1/s} \left( C_3(\alpha) \right)^{1+\frac{2}{s}} E(y_0, f_0) \left( 1 + \frac{M}{|\Omega|} \right)^{\frac{(2C\beta^2)(2s+1)}{s}} \leq M. \quad (36)$$

2 Such  $M$  exists under the assumption  $\beta < \beta^*(s)$  equivalent to  $\frac{(2C\beta^2)(2s+1)}{s} < 1$ . We denote by  $M_3$  the  
3 smallest value of  $M$  such that (36) holds true. Eventually, taking  $M := \max(M_1, M_2, M_3)$ , we get that  
4  $\|y_{n+1}\|_{L^\infty(0,T;L^1(\Omega))} \leq M$  as well. We have then proved by induction the uniform property (32) for some  
5  $M$  large enough.

*Proof of the convergence of the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$*  - In view of Lemma 5 with  $p^* = 1$ , the uniform property (32) implies that the observability constant  $C e^{C \|g'(y_k)\|_{L^\infty(0,T;L^d(\Omega))}^2}$  appearing in the estimates for  $(Y_k^1, F_k^1)$  (see Proposition 1) is uniformly bounded with respect to the parameter  $k$ . As a consequence, the constant  $c(y_k)$  appearing in the instrumental estimate (15) is bounded by  $c_M(\beta)$  given by (33). Consequently, the developments of Section 3.1 apply with  $c = c_M(\beta)$ . Theorem 2 then follows from the proof of Proposition 4 except for the limit in (20) with respect to  $k$  (since  $g'$  is not anymore in  $L^\infty(Q_T)$ ). Since  $g \in C^1(\mathbb{R})$ , a.e in  $Q_T$  there exists  $0 \leq \theta(x, t) \leq 1$  such that

$$\begin{aligned} |g(y_k(x, t)) - g(\bar{y}(x, t))| &= |g'(\bar{y}(x, t) + \theta(x, t)y_k(x, t))| |y_k(x, t) - \bar{y}(x, t)| \\ &\leq (\alpha + \beta \ln^{1/2} (1 + |\bar{y}(x, t) + \theta(x, t)y_k(x, t)|)) |y_k(x, t) - \bar{y}(x, t)| \\ &\leq (\alpha + \beta (|\bar{y}(x, t)|^{1/2} + |y_k(x, t)|^{1/2})) |y_k(x, t) - \bar{y}(x, t)| \end{aligned}$$

and thus

$$\|g(y_k) - g(\bar{y})\|_2 \leq (\alpha |Q_T|^{1/4} + \beta (\|\bar{y}\|_2^{1/2} + \|y_k\|_2^{1/2})) \|y_k - \bar{y}\|_4.$$

Since  $y_k \rightarrow \bar{y}$  in  $L^4(Q_T)$ , it follows that  $g(y_k) \rightarrow g(\bar{y})$  in  $L^2(Q_T)$ . Moreover, since  $(y_k)_{k \in \mathbb{N}}$  is a bounded sequence of  $L^4(Q_T)$ , the estimate

$$\|g'(y_k)\|_2 \leq C(\alpha + \beta \|y_k\|_2^{1/2}) \|y_k\|_4$$

6 implies that  $(g'(y_k))_{k \in \mathbb{N}}$  is a bounded sequence of  $L^2(Q_T)$ . Then, using that  $(Y_k^1, F_k^1)$  goes to zero as  
7  $k \rightarrow \infty$  in  $\mathcal{A}_0$ , we pass to the limit in (20) and get that  $(\bar{y}, \bar{f}) \in \mathcal{A}$  solves (28).  $\square$

**Remark 5.** Remark that  $M := \max(M_2, M_3)$  since  $M_3 \geq M_1$ . The constant  $M_2$  can be made explicit since the constraint (35) implies that

$$\left( \frac{C[g']_s}{\sqrt{2}} \right)^{1/s} C_3(\alpha)^{2/s} \left( 1 + \frac{M}{|\Omega|} \right)^{\frac{4C\beta^2}{s}} \sqrt{E(y_0, f_0)} \geq 1.$$

equivalent to

$$\left( 1 + \frac{M}{|\Omega|} \right)^{2C\beta^2} \geq C_3(\alpha)^{-1} \sqrt{E(y_0, f_0)}^{-s/2} \left( \frac{C}{\sqrt{2}} [g']_s \right)^{-1/2}.$$

8 In particular,  $M_2$  is large for small values of  $\sqrt{E(y_0, f_0)}$ , for any  $s > 0$ . On the other hand, the constant  
9  $M_3$  is no explicit, hence whether  $M_2 > M_3$  or  $M_3 > M_2$  depend on the values of  $\sqrt{E(y_0, f_0)}$  and  
10  $\|y_0\|_{L^\infty(0,T;L^1(\Omega))}$ . Remark that  $\sqrt{E(y_0, f_0)}$  can be large and  $\|y_0\|_{L^\infty(0,T;L^1(\Omega))}$  small, and vice versa.



## 4 Comments

**Asymptotic condition.** The asymptotic condition  $(\mathbf{H}_2)$  on  $g'$  is slightly stronger than the asymptotic condition  $(\mathbf{H}_1)$  made in [18]: this is due to our linearization of (1) which involves  $r \rightarrow g'(r)$  while the linearization (2) in [18] involves  $r \rightarrow (g(r) - g(0))/r$ . There exist cases covered by Theorem 1 in which exact controllability for (1) is true but that are not covered by Theorem 2. Note however that the example  $g(r) = a + br + cr \ln^{1/2}(1 + |r|)$ , for any  $a, b \in \mathbb{R}$  and for any  $c > 0$  small enough (which is somehow the limit case in Theorem 1) satisfies  $(\mathbf{H}_2)$  as well as  $(\overline{\mathbf{H}}_s)$  for any  $s \in (0, 1]$ .

**Link with Newton method.** Defining  $F : \mathcal{A} \rightarrow L^2(Q_T)$  by  $F(y, f) := (\partial_{tt}y - \Delta y + g(y) - f \mathbf{1}_\omega)$ , we have  $E(y, f) = \frac{1}{2} \|F(y, f)\|_2^2$  and we observe that, for  $\lambda_k = 1$ , the algorithm (19) coincides with the Newton algorithm associated to the mapping  $F$ . This explains the super-linear convergence property in Theorem 2, in particular the quadratic convergence when  $s = 1$ . The optimization of the parameter  $\lambda_k$  gives to a global convergence property of the algorithm and leads to the so-called damped Newton method applied to  $F$  (we refer to [13, chapter 8]).

**Initialization with the controlled pair of the linear equation.** The number of iterates to achieve convergence (notably to enter in a super-linear regime) depends on the size of the value  $E(y_0, f_0)$ . A natural example of an initialization  $(y_0, f_0) \in \mathcal{A}$  is the unique solution of minimal control norm of (1) with  $g = 0$  (i.e., in the linear case). Under the assumption  $(\mathbf{H}_2)$ , this leads to the estimate

$$E(y_0, f_0) = \frac{1}{2} \|g(y_0)\|_2^2 \leq |g(0)|^2 |Q_T| + 2 \int_{Q_T} |y_0|^2 (\alpha^2 + \beta^2 \ln(1 + |y_0|)).$$

**Local controllability when removing the growth condition  $(\mathbf{H}_2)$ .** If the real  $E(y_0, f_0)$  is small enough, we may remove the growth condition  $(\mathbf{H}_2)$  on  $g'$ .

**Proposition 5.** Assume  $g'$  satisfies  $(\overline{\mathbf{H}}_s)$  for some  $s \in (0, 1]$ . Let  $(y_k, f_k)_{k>0}$  be the sequence of  $\mathcal{A}$  defined in (19). There exists a constant  $C([g']_s)$  such that if  $E(y_0, f_0) \leq C([g']_s)$ , then  $(y_k, f_k)_{k \in \mathbb{N}} \rightarrow (\bar{y}, \bar{f})$  in  $\mathcal{A}$  where  $\bar{f}$  is a null control for  $\bar{y}$  solution of (1). Moreover, the convergence is at least linear and is at least of order  $1 + s$  after a finite number of iterations.

*Proof.* In this proof, the notation  $\|\cdot\|_{\infty, d}$  stands for  $\|\cdot\|_{L^\infty(0, T; L^d(\Omega))}$ . We note  $D := \frac{C}{(1+s)\sqrt{2}} [g']_s$  and  $e_k := c(y_k) E(y_k, f_k)^{s/2}$  with  $c(y) := Dd(y)^{1+s}$  and  $d(y) := C e^{C \|g'(y)\|_{\infty, d}^2}$ . (23) then reads

$$\sqrt{E(y_{k+1}, f_{k+1})} \leq \min_{\lambda \in [0, m]} (|1 - \lambda| + \lambda^{1+s} e_k) \sqrt{E(y_k, f_k)}. \quad (37)$$

We write  $|g'(y_k) - g'(y_k - \lambda_k Y_k^1)| \leq [g']_s |\lambda_k Y_k^1|^s$  so that

$$\|g'(y_{k+1})\|_{L^\infty(0, T; L^d(\Omega))}^2 \leq \|g'(y_k)\|_{\infty, d}^2 + ([g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s)^2 + 2 \|g'(y_k)\|_{\infty, d} [g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s$$

and

$$e^{C \|g'(y_{k+1})\|_{\infty, d}^2} \leq e^{C \|g'(y_k)\|_{\infty, d}^2} e^{C ([g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s)^2} e^{2C \|g'(y_k)\|_{\infty, d} ([g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s)}$$

leading to

$$\frac{c(y_{k+1})}{c(y_k)} \leq \left( e^{C ([g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s)^2} e^{2C \|g'(y_k)\|_{\infty, d} ([g']_s \lambda_k^s \|Y_k^1\|_{\infty, d}^s)} \right)^{1+s}.$$

We infer that  $\|(Y_k^1)^s\|_{\infty, d} = \|Y_k^1\|_{\infty, sd}^s$ . Moreover, estimate (8) leads to

$$\|Y_k^1\|_{\infty, sd}^s \leq d^s(y_k) E(y_k, f_k)^{s/2} = \frac{c(y_k)^{\frac{s}{1+s}}}{D^{\frac{s}{1+s}}} E(y_k, f_k)^{s/2} \leq D^{-\frac{s}{1+s}} c(y_k) E(y_k, f_k)^{s/2}$$

using that  $c(y_k) \geq 1$  (by increasing the constant  $C$  if necessary). Consequently,

$$e^{C([g']_s \lambda^s \|Y_k^1\|_{\infty,d})^2} \leq e^{C([g']_s \lambda^s D^{-\frac{s}{1+s}} e_k)^2} := e^{C_1 e_k^2}.$$

Similarly,

$$\begin{aligned} \|g'(y_k)\|_{\infty,d} \|Y_k^1\|_{\infty,d}^s &\leq \|g'(y_k)\|_{\infty,d} d^s E(y_k, f_k)^{s/2} \\ &\leq \|g'(y_k)\|_{\infty,d} \left( C e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \right)^s E(y_k, f_k)^{s/2} \\ &\leq \left( C e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2} \right)^{s+1} E(y_k, f_k)^{s/2} \leq \frac{c(y_k)}{D} E(y_k, f_k)^{s/2} = \frac{e_k}{D} \end{aligned}$$

using that  $a \leq C e^{C a^2}$  for all  $a \geq 0$  and  $C > 0$  large enough. It follows that

$$e^{2C \|g'(y_k)\|_{\infty,d} ([g']_s \lambda_k^s \|Y_k^1\|_{\infty,d})} \leq e^{2C [g']_s \lambda_k^s \frac{e_k}{D}} := e^{C_2 e_k}$$

1 and then  $\frac{c(y_{k+1})}{c(y_k)} \leq (e^{C_1 e_k^2 + C_2 e_k})^{1+s}$ . By multiplying (37) by  $c(y_{k+1})$ , we obtain the inequality

$$e_{k+1} \leq \min_{\lambda \in [0,m]} (|1 - \lambda| + e_k \lambda^{1+s}) (e^{C_1 e_k^2 + C_2 e_k})^{1+s} e_k.$$

2 If  $2e_k < 1$ , the minimum is reached for  $\lambda = 1$  leading  $\frac{e_{k+1}}{e_k} \leq e_k (e^{C_1 e_k^2 + C_2 e_k})^{1+s}$ . Consequently, if the ini-  
3 tial guess  $(y_0, f_0)$  belongs to the set  $\{(y_0, f_0) \in \mathcal{A}, e_0 < 1/2, e_0 (e^{C_1 e_0^2 + C_2 e_0})^{1+s} < 1\}$ , the sequence  $(e_k)_{k>0}$   
4 goes to zero as  $k \rightarrow \infty$ . Since  $c(y_k) \geq 1$  for all  $k \in \mathbb{N}$ , this implies that the sequence  $(E(y_k, f_k))_{k>0}$  goes  
5 to zero as well. Moreover, from (8), we get  $D \|Y_k^1, F_k^1\|_{\mathcal{H}} \leq e_k \sqrt{E(y_k, f_k)}$  and repeating the arguments  
6 of the proof of Proposition 4, we conclude that the sequence  $(y_k, f_k)_{k>0}$  converges to a controlled pair  
7 for (1).  $\square$

8 These computations does not use the assumption  $(\mathbf{H}_2)$  on the nonlinearity  $g$ . However, the smallness  
9 assumption on  $e_0$  requires a smallness assumption on  $E(y_0, f_0)$  (since  $c(y_0) > 1$ ). This is equivalent  
10 to assume the controllability of (1). Alternatively, in the case  $g(0) = 0$ , the smallness assumption on  
11  $E(y_0, f_0)$  is achieved as soon as  $\|(u_0, u_1)\|_{\mathbf{V}}$  is small enough. Therefore, the convergence result stated in  
12 Proposition 5 is equivalent to the local controllability property for (1). Proposition 5 can also be seen  
13 as a consequence of the usual convergence of the Newton method: when  $E(y_0, f_0)$  is small enough, i.e.,  
14 when the initialization is close enough to the solution, then  $\lambda_k = 1$  for every  $k \in \mathbb{N}$  and we recover the  
15 standard Newton method.

16 **Weakening of the condition  $(\overline{\mathbf{H}}_s)$ .** Given any  $s \in (0, 1]$ , we introduce for any  $g \in \mathcal{C}^1(\mathbb{R})$  the following  
17 hypothesis :

18  $(\overline{\mathbf{H}}'_s)$  There exist  $\overline{\alpha}, \overline{\beta}, \gamma \in \mathbb{R}^+$  such that  $|g'(a) - g'(b)| \leq |a - b|^s (\overline{\alpha} + \overline{\beta}(|a|^\gamma + |b|^\gamma))$ ,  $\forall a, b \in \mathbb{R}$

which coincides with  $(\overline{\mathbf{H}}_s)$  if  $\gamma = 0$  for  $\overline{\alpha} + \overline{\beta} = [g']_s$ . If  $\gamma \in (0, 1)$  is small enough and related to the  
constant  $\beta$  appearing in the growth condition  $(\mathbf{H}_2)$ , Theorem 2 still holds if  $(\overline{\mathbf{H}}_s)$  is replaced by the  
weaker hypothesis  $(\overline{\mathbf{H}}'_s)$ . Precisely, if  $g$  satisfies  $(\mathbf{H}_2)$  and  $(\overline{\mathbf{H}}'_s)$  for some  $s \in (0, 1]$ , then the sequence  
 $(y_k, f_k)_{k \in \mathbb{N}}$  defined by (19) fulfills the estimate

$$E(y_{k+1}, f_{k+1}) \leq E(y_k, f_k) \min_{\lambda \in [0,m]} (|1 - \lambda| + \lambda^{1+s} c(y_k) E(y_k, f_k)^{s/2})^2$$

19 with  $c(y) := \frac{1}{(1+s)\sqrt{2}} (\overline{\alpha} + 2\overline{\beta} \|y_k\|_{\infty,6\gamma}^\gamma + \overline{\beta} m^\gamma d(y)^\gamma E(y_0, f_0)^{\gamma/2}) d(y)^{1+s}$  and  $d(y) := C e^{C \|g'(y)\|_{L^\infty(0,T;L^d(\Omega))}^2}$ .  
20 Using Lemma 5 with  $p^* = 6\gamma \leq 6$  and proceeding as in the proof of Theorem 2, one may prove by induction  
21 that the sequence  $(\|y_k\|_{L^\infty(0,T;L^6(\Omega))})_{k \in \mathbb{N}}$  is uniformly bounded under the condition  $\frac{\gamma + 2C\beta^2(1+2s)}{s} < 1$  and  
22 then deduce the convergence of the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$ .

## 5 Numerical illustrations

In this section, we illustrate our results of convergence. We provide some practical details about the algorithm (19), then discuss some experiments in one and two dimensions performed with the software FreeFem++ [21].

### 5.1 Algorithm

We introduce a cut-off  $\chi$  of the form  $\chi(x, t) = \chi_1(x)\chi_2(t)$ , where  $\chi_1 \in \mathcal{C}_0^\infty(\omega)$  and  $\chi_2 \in \mathcal{C}_0^\infty(0, T)$  take values in  $[0, 1]$ . In the sequel, we consider controls of minimal  $L_\chi^2(q_T)$ -norm, with  $L_\chi^2(q_T) := \{f \mid \int_{q_T} f^2 \chi^{-1} < +\infty\}$ . Besides, for  $n \in \mathbb{N}$ , we denote  $e_n := \partial_{tt}y_n - \Delta y_n + g(y_n) - f_n 1_\omega$ . Then, algorithm (19) can be expanded as follows.

1. *Initialization* – We compute the state-control pair  $(y_0, f_0) \in \mathcal{A}$  solution of

$$\begin{cases} \partial_{tt}y_0 - \Delta y_0 = f_0 1_\omega, & \text{in } Q_T, \\ y_0 = 0, & \text{on } \Sigma_T, \\ (y_0(\cdot, 0), \partial_t y_0(\cdot, 0)) = (u_0, u_1), & \text{in } \Omega, \\ (y_0(\cdot, T), \partial_t y_0(\cdot, T)) = (z_0, z_1), & \text{in } \Omega, \end{cases} \quad (38)$$

where  $f_0$  is the control of minimal  $L_\chi^2(q_T)$ -norm. We then evaluate  $e_0 = g(y_0)$ .

Assume now that  $(y_n, f_n) \in \mathcal{A}$  and  $e_n \in L^2(Q_T)$  are computed for some  $n \geq 0$ .

2. *Evaluation of the least-squares functional* – We compute the error functional  $E(y_n, f_n) = \frac{1}{2}\|e_n\|_2^2$ . If  $\sqrt{2E(y_n, f_n)} \leq 10^{-5}$ , then the algorithm stops.

3. *Descent direction* – We compute the state-control pair  $(Y_n^1, F_n^1) \in \mathcal{A}_0$  solution of

$$\begin{cases} \partial_{tt}Y_n^1 - \Delta Y_n^1 + g'(y_n)Y_n^1 = F_n^1 1_\omega + e_n, & \text{in } Q_T, \\ Y_n^1 = 0, & \text{on } \Sigma_T, \\ (Y_n^1(\cdot, 0), \partial_t Y_n^1(\cdot, 0)) = (0, 0), & \text{in } \Omega, \\ (Y_n^1(\cdot, T), \partial_t Y_n^1(\cdot, T)) = (0, 0), & \text{in } \Omega, \end{cases} \quad (39)$$

where  $F_n^1$  is the control of minimal  $L_\chi^2(q_T)$ -norm.

4. *Optimal descent step* – We compute the optimal descent step  $\lambda_n$  as the minimizer in  $[0, 1]$  of  $\lambda \mapsto E((y_n, f_n) - \lambda(Y_n^1, F_n^1))$ , evaluated using the expression

$$E((y_n, f_n) - \lambda(Y_n^1, F_n^1)) = \frac{1}{2}\|(1 - \lambda)e_n + l_n(\lambda)\|_2^2,$$

where  $l_n(\lambda) := g(y_n - \lambda Y_n^1) - g(y_n) + \lambda g'(y_n)Y_n^1$ . This is done with 20 iterations of the trichotomy method on the interval  $[0, 1]$ .

5. *Update* – We set  $(y_{n+1}, f_{n+1}) = (y_n, f_n) - \lambda_n(Y_n^1, F_n^1)$ . We then evaluate  $e_{n+1} = (1 - \lambda_n)e_n + l_n(\lambda_n)$  and return to step 2.

In the sequel, we denote by  $k^* = \min\{k \geq 0 \mid \sqrt{2E(y_k, f_k)} \leq 10^{-5}\}$  and define the corresponding approximation of the solution

$$(y^*, f^*) = (y_0, f_0) - \sum_{k=0}^{k^*} \lambda_k (Y_k^1, F_k^1) \in \mathcal{A}.$$

Then, in order to measure *a posteriori* the quality of this approximation, we shall compute the relative term

$$\mathcal{E}_T := \frac{\|(y, \partial_t y)(\cdot, T; f^*)\|_{\mathbf{V}}}{\|(y, \partial_t y)(\cdot, T; 0)\|_{\mathbf{V}}},$$

1 where  $y(\cdot, \tau; f^*)$  (resp.  $y(\cdot, \tau; 0)$ ) is the solution at time  $\tau$  of (1) with control equal to  $f = f^*$  (resp.  $f = 0$ ).

2  
3 The introduction of the cut-off  $\chi$  together with regularity assumptions on the initial datum  $(u_0, u_1)$   
4 make the state-control pairs  $(y_0, f_0)$  and  $(Y_k^1, F_k^1)$  regular as well (we refer to [15]). This allows to  
5 give a meaning to  $e_k = \partial_{tt} y_k - \Delta y_k + g(y_k) - f_k 1_\omega$  as an  $L^2(Q_T)$  function. Moreover, this involves  
6 stability properties with respect to the discretization parameters for the standard finite-dimensional  
7 approximations of systems (38) to (41) below.

## 8 5.2 Experiments in 2D

We consider a two-dimensional case for which  $\Omega = (0, 1)^2$ . The controllability time is equal to  $T = 3$  and the control domain  $\omega$  is depicted on Figure 1. Moreover, for any real constant  $c_g$ , we consider the non-linear function  $g$  defined by

$$g(r) = -c_g r \ln^{1/2}(2 + |r|), \quad \forall r \in \mathbb{R}.$$

9 We check that  $g$  satisfies  $(\overline{\mathbf{H}}_s)$  for  $s = 1$  and  $(\mathbf{H}_2)$  for  $|c_g|$  small enough. Remark that the unfavorable  
10 situation in which the norm of the corresponding uncontrolled solution of (1) grows corresponds to  
11 positive values of  $c_g$ . As for the initial and final data, we consider  $(u_0, u_1) = (100 \sin(\pi x_1) \sin(\pi x_2), 0)$   
12 and  $(z_0, z_1) = (0, 0)$  respectively.

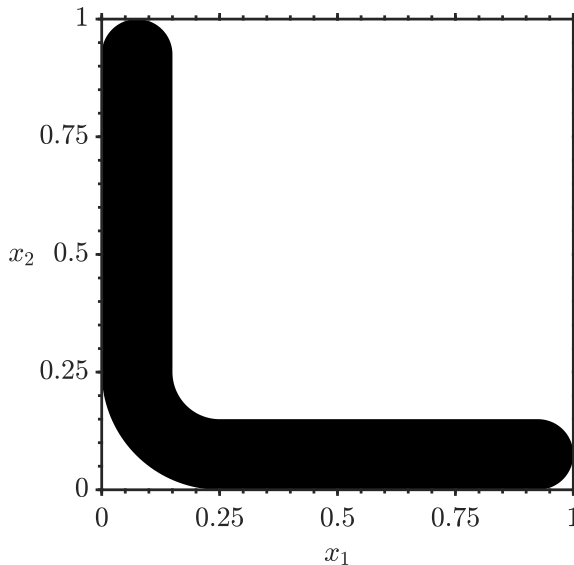


Figure 1: Control domain  $\omega \subset \Omega = (0, 1)^2$  (black part).

13 In order to determine the state-control pairs  $(y_0, f_0)$  and  $(Y_n^1, F_n^1)$  of (38) and (39) respectively, we  
14 employ the method introduced by Glowinski-Li-Lions in the seminal work [19], based on the unconstrained  
15 minimization of the conjugate functional.

16 Concerning problem (38), for any  $(w_0, w_1) \in \mathbf{H} := L^2(\Omega) \times H^{-1}(\Omega)$ , we consider the adjoint system

$$\begin{cases} \partial_{tt} \varphi - \Delta \varphi = 0, & \text{in } Q_T, \\ \varphi = 0, & \text{on } \Sigma_T, \\ (\varphi(\cdot, T), \partial_t \varphi(\cdot, T)) = (w_0, w_1), & \text{in } \Omega, \end{cases} \quad (40)$$

and the functional

$$J_0^*(w_0, w_1) := \frac{1}{2} \int_{q_T} |\varphi|^2 \chi - \langle u_0, \partial_t \varphi(\cdot, 0) \rangle_{H_0^1(\Omega), H^{-1}(\Omega)} + (u_1, \varphi(\cdot, 0))_{L^2(\Omega)} \\ + \langle z_0, w_1 \rangle_{H_0^1(\Omega), H^{-1}(\Omega)} - (z_1, w_0)_{L^2(\Omega)}.$$

1 Then, the control  $f_0$  of minimal  $L_\chi^2(q_T)$ -norm is given by  $f_0 = \varphi_0 \chi$ , where  $\varphi_0$  is the solution of (40)  
 2 associated with the minimizer  $(\widehat{w}_0, \widehat{w}_1)$  of  $J_0^*$  over  $\mathbf{H}$ . The resolution of this minimization problem is  
 3 done using the Fletcher-Reeves conjugate gradient algorithm, initialized with  $(w_0, w_1) = (0, 0)$ . The  
 4 stopping criterion is  $\|g_p\|_{\mathbf{H}} \leq 10^{-5} \|g_0\|_{\mathbf{H}}$ , where  $g_p$  denotes the gradient of  $J_0^*$  at iteration  $p$ .

5 Concerning problem (39), for any  $(w_0, w_1) \in \mathbf{H}$ , we consider the adjoint system

$$\begin{cases} \partial_{tt} \varphi - \Delta \varphi + g'(y_n) \varphi = 0, & \text{in } Q_T, \\ \varphi = 0, & \text{on } \Sigma_T, \\ (\varphi(\cdot, T), \partial_t \varphi(\cdot, T)) = (w_0, w_1), & \text{in } \Omega, \end{cases} \quad (41)$$

and the functional

$$J_n^*(w_0, w_1) := \frac{1}{2} \int_{q_T} |\varphi|^2 \chi + \int_{Q_T} e_n \varphi.$$

6 Then, the control  $F_n^1$  of minimal  $L_\chi^2(q_T)$ -norm is given by  $F_n^1 = \varphi_n \chi$ , where  $\varphi_n$  is the solution of (41)  
 7 associated with the minimizer  $(\widehat{w}_0, \widehat{w}_1)$  of  $J_n^*$  over  $\mathbf{H}$ . The resolution of this minimization problem is  
 8 done using the Fletcher-Reeves conjugate gradient algorithm, initialized with the minimizer of  $J_{n-1}^*$ . The  
 9 stopping criterion is  $\|g_p\|_{\mathbf{H}} \leq 10^{-5} \|g_0\|_{\mathbf{H}}$ , where  $g_p$  denotes the gradient of  $J_n^*$  at iteration  $p$ .

10 To compute the solution of the state systems (38)-(39) and the adjoint systems (40)-(41), we use a  
 11 time-marching method combining an explicit centered finite-difference scheme in time and a finite-element  
 12 approximation in space. We consider a uniform discretization  $(t_i)_{i=0, \dots, N}$  of the time interval  $[0, T]$  and  
 13 denote by  $\delta t = T/N$  the time discretization parameter. Besides, we consider a family  $\mathcal{T} = \{\mathcal{T}_h, h > 0\}$   
 14 of regular triangulations of  $\Omega$  such that  $\overline{\Omega} = \cup_{K \in \mathcal{T}_h} K$ . The family is indexed by  $h = \max_{K \in \mathcal{T}_h} |K|$ .  
 15 For every time  $t_i$ , the variables  $\varphi_0(\cdot, t_i)$ ,  $\varphi_n(\cdot, t_i)$ ,  $y_0(\cdot, t_i)$  and  $Y_n^1(\cdot, t_i)$  are approximated in the space  
 16  $P_h = \{p_h \in \mathcal{C}(\overline{Q_T}) \mid p_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}$  where  $\mathbb{P}_1(K)$  denotes the space of polynomials of degree  
 17 one. We refer to [5] for convergence results in this setting. We also refer to [1, 7, 16, 32]. In the sequel,  
 18 we mainly use a regular triangulation  $\mathcal{T}_h$  with fineness  $h = 1/64$  and a time step equal to  $\delta t = h/3$  in  
 19 order to satisfy the CFL condition arising from the explicit scheme with respect to the time variable.

20 We now present some simulations for several values of the constant  $c_g$ .

21 • Case  $c_g = 1$  – Table 1 collects some norms from the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  associated with the value  
 22  $c_g = 1$ . The convergence of the algorithm is observed after  $k^* = 3$  iterations. The optimal steps  $\lambda_k$  are  
 23 equal to one so that the algorithm (19) coincides with the Newton algorithm (see the second item of  
 24 Section 4). Figure 2 depicts with respect to the time variable the  $L^2(\Omega)$ -norm of the controlled solution  
 25  $y^* = y_{k=k^*}$  (red solid line) to be compared with the  $L^2(\Omega)$ -norm of the controlled solution  $y_{k=0}$  of  
 26 the linear equation (blue dash-dotted line) used to initialize the algorithm (equivalently, this controlled  
 27 solution corresponds to  $c_g = 0$ ). The effect of the non-linearity is reduced as the dynamics of the two  
 28 controlled solutions are similar. The figure also depicts the  $L^2(\Omega)$ -norm of the uncontrolled solution (blue  
 29 dashed line) and displays a periodic behavior. Similarly, Figure 3 depicts the  $L_\chi^2(q_T)$ -norm of the null  
 30 control  $f^* = f_{k=k^*}$  (red solid line) and  $f_{k=0}$  (blue dash-dotted line). By construction, these controls  
 31 vanish at the initial and final times. The corresponding value of the relative error  $\mathcal{E}_T = 2.53 \times 10^{-4}$   
 32 indicates a notable reduction of the solution at time  $T$  thought the action of the control.

33 • Case  $c_g = 5$  – Table 2 and Figures 4 and 5 collect the results obtained for the value  $c_g = 5$ . The  
 34 relative error takes the value  $\mathcal{E}_T = 2.36 \times 10^{-4}$ . The convergence is quadratic and is obtained after  $k^* = 4$   
 35 iterations.

36 • Case  $c_g = 10$  – Table 3 and Figures 6 and 7 collect the results obtained for the value  $c_g = 10$ . We  
 37 compute the relative error  $\mathcal{E}_T = 2.64 \times 10^{-5}$ . The convergence is observed after  $k^* = 4$  iterations. As

#iterate $k$	$\sqrt{2E(y_k, f_k)}$	$\lambda_k$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_\chi(q_T)}}{\ f_{k-1}\ _{L^2_\chi(q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_\chi(q_T)}$
0	$7.32 \times 10^1$	1	—	—	37.653	1339.39
1	$9.62 \times 10^{-1}$	1	$1.72 \times 10^{-1}$	$3.29 \times 10^{-1}$	37.113	1265.62
2	$1.03 \times 10^{-5}$	1	$3.83 \times 10^{-4}$	$1.06 \times 10^{-3}$	37.115	1265.77
3	$6.42 \times 10^{-15}$	—	$4.44 \times 10^{-9}$	$9.34 \times 10^{-9}$	37.115	1265.77

Table 1:  $c_g = 1$  – Norms of  $(y_k, f_k)$  w.r.t.  $k$  defined by the algorithm (19).

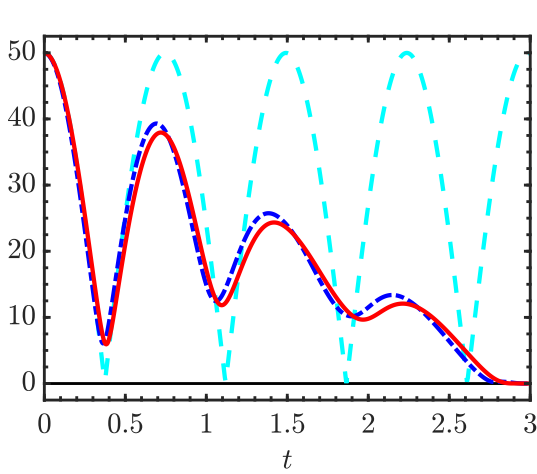


Figure 2:  $c_g = 1$  – (—)  $\|y^*(\cdot, t)\|_{L^2(\Omega)}$  ;  
(---)  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  ; (—)  $\|y(\cdot, t; 0)\|_{L^2(\Omega)}$ .

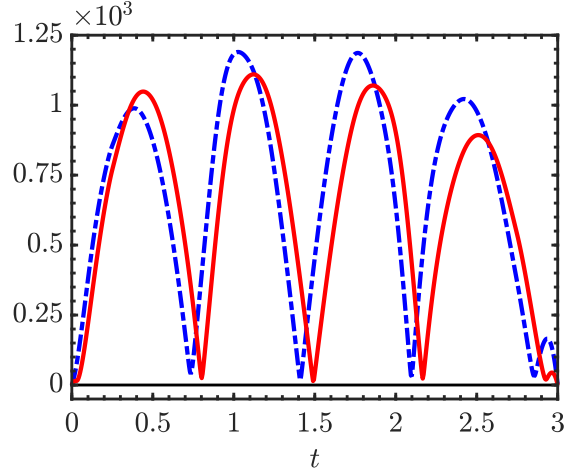


Figure 3:  $c_g = 1$  – (—)  $\|f^*(\cdot, t)\|_{L^2_\chi(\omega)}$  ;  
(---)  $\|f_0(\cdot, t)\|_{L^2_\chi(\omega)}$ .

#iterate $k$	$\sqrt{2E(y_k, f_k)}$	$\lambda_k$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_\chi(q_T)}}{\ f_{k-1}\ _{L^2_\chi(q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_\chi(q_T)}$
0	$3.66 \times 10^2$	1	—	—	37.653	1339.39
1	$4.87 \times 10^1$	0.996	$9.50 \times 10^{-1}$	$1.05 \times 10^0$	31.353	1223.44
2	$8.65 \times 10^{-1}$	1	$7.67 \times 10^{-2}$	$1.50 \times 10^{-1}$	32.101	1348.12
3	$5.82 \times 10^{-5}$	1	$3.91 \times 10^{-4}$	$7.37 \times 10^{-4}$	32.104	1348.09
4	$7.17 \times 10^{-14}$	—	$1.45 \times 10^{-8}$	$3.24 \times 10^{-8}$	32.104	1348.09

Table 2:  $c_g = 5$  – Norms of  $(y_k, f_k)$  w.r.t.  $k$  defined by the algorithm (19).

1 before, the optimal steps are very close to one. The main difference with the previous situations for which  
2  $c_g = 1$  and  $c_g = 5$  is the behavior of the uncontrolled solution which grows exponentially with respect to  
3 the time variable, as shown in Figure 6. As expected, this larger value of  $c_g$  induces a larger gap between  
4 the non-linear control and the linear one. We observe notably that the non-linear control  $f^*$  acts stronger  
5 from the beginning, precisely in order to balance the initial exponential growth of the solution outside  
6 the subset  $\omega$ . We also observe that the control reduces the oscillations of the corresponding controlled  
7 solution (in comparison with the solution of the linear equation). For larger values of  $c_g$ , we suspect a  
8 different dynamic yielding to the first values of the optimal step  $\lambda_k$  being far from one (as observed in [26]  
9 for the resolution of the Navier-Stokes system with large values of the Reynolds number). However, for  
10 larger values of  $c_g$  (for instance  $c_g = 20$ ), the exponential growth behavior leads to numerical instabilities  
11 and overflow in the computation of the controlled pair  $(Y_k^1, F_k^1)$  solution of (39), where the potential  
12  $g'(y_k)$  appears. This leads to the divergence of the conjugate gradient algorithm including for very fine  
13 discretizations and the non-convergence of the least-squares algorithm. In the next section, we shall

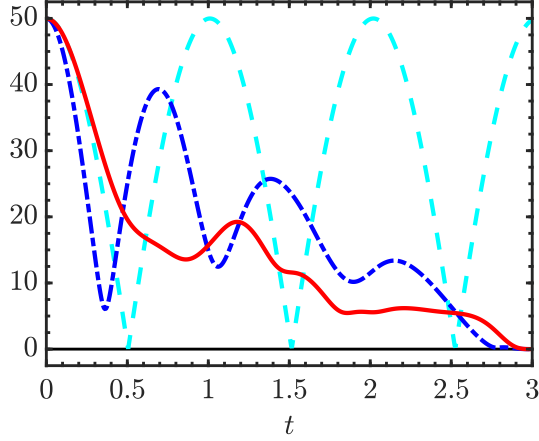


Figure 4:  $c_g = 5$  – (—)  $\|y^*(\cdot, t)\|_{L^2(\Omega)}$  ;  
 (---)  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  ; (---)  $\|y(\cdot, t; 0)\|_{L^2(\Omega)}$ .

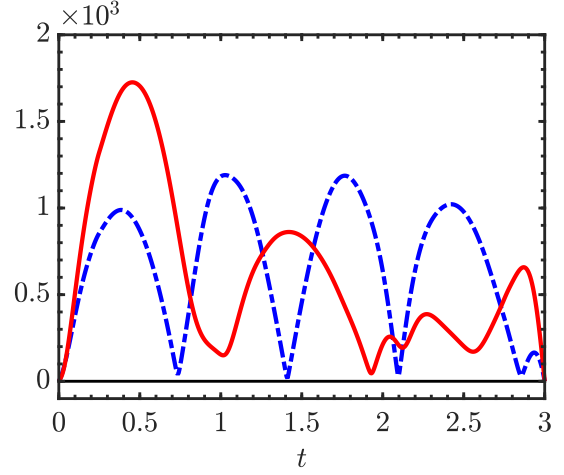


Figure 5:  $c_g = 5$  – (—)  $\|f^*(\cdot, t)\|_{L^2_x(\omega)}$  ;  
 (---)  $\|f_0(\cdot, t)\|_{L^2_x(\omega)}$ .

1 employ a different method of approximation allowing to consider larger values of  $c_g$ .

#iterate $k$	$\sqrt{2E}(y_k, f_k)$	$\lambda_k$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_x(Q_T)}}{\ f_{k-1}\ _{L^2_x(Q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_x(Q_T)}$
0	$7.32 \times 10^2$	1	—	—	37.653	1339.39
1	$1.59 \times 10^2$	0.998	$1.76 \times 10^0$	$9.40 \times 10^{-1}$	57.718	1164.11
2	$2.43 \times 10^0$	1	$8.82 \times 10^{-2}$	$1.30 \times 10^{-1}$	59.934	1138.75
3	$4.09 \times 10^{-3}$	1	$1.62 \times 10^{-3}$	$5.66 \times 10^{-3}$	59.892	1141.01
4	$1.80 \times 10^{-9}$	—	$1.10 \times 10^{-6}$	$2.96 \times 10^{-6}$	59.891	1141.01

Table 3:  $c_g = 10$  – Norms of  $(y_k, f_k)$  w.r.t.  $k$  defined by the algorithm (19).

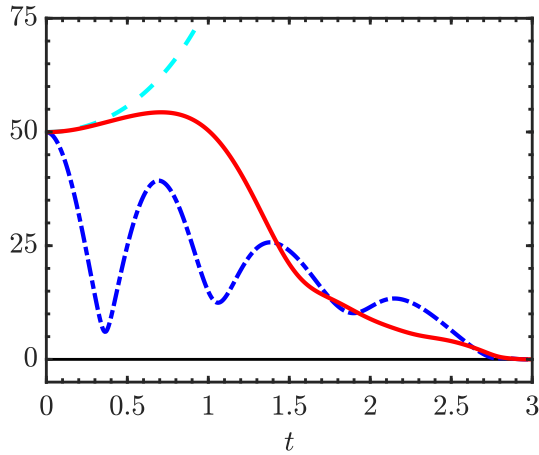


Figure 6:  $c_g = 10$  – (—)  $\|y^*(\cdot, t)\|_{L^2(\Omega)}$  ;  
 (---)  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  ; (---)  $\|y(\cdot, t; 0)\|_{L^2(\Omega)}$ .

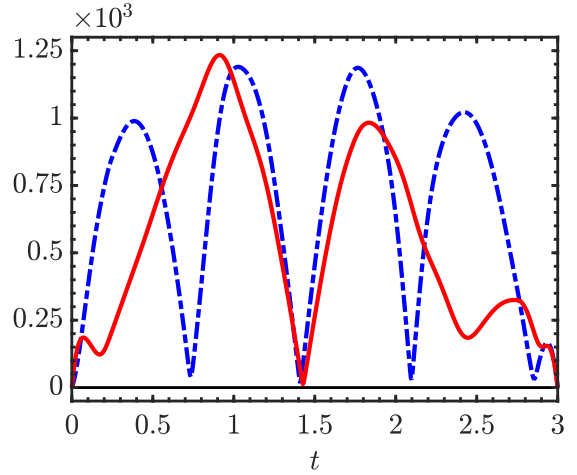


Figure 7:  $c_g = 10$  – (—)  $\|f^*(\cdot, t)\|_{L^2_x(\omega)}$  ;  
 (---)  $\|f_0(\cdot, t)\|_{L^2_x(\omega)}$ .

2 • Case  $c_g = -20$  – For negative values of  $c_g$  leading to  $rg(r) \geq 0$  for every  $r$ , the situation is more

1 favorable from a computational viewpoint. Table 4 and Figures 8 and 9 are concerned with the value  
2  $c_g = -20$ . The convergence is observed after  $k^* = 4$  iterations and leads to  $\mathcal{E}_T = 3.64 \times 10^{-4}$ . We observe  
3 that the uncontrolled solution oscillates faster as  $c_g$  decreases. This leads to an oscillatory dynamic of  
4 the optimal control pair  $(y^*, f^*)$ . We also observe that the norm of the control  $f^*$  is significantly greater  
5 than the norm of  $f_0$ , the initial control associated with the linear case.

#iterate $k$	$\sqrt{2E(y_k, f_k)}$	$\lambda_k$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_\chi(Q_T)}}{\ f_{k-1}\ _{L^2_\chi(Q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_\chi(Q_T)}$
0	$1.46 \times 10^3$	1	–	–	37.653	1339.39
1	$2.70 \times 10^2$	0.985	$1.38 \times 10^0$	$1.69 \times 10^0$	42.479	2601.16
2	$1.55 \times 10^1$	1	$1.57 \times 10^{-1}$	$1.43 \times 10^{-1}$	44.309	2696.17
3	$1.94 \times 10^{-2}$	1	$3.68 \times 10^{-3}$	$5.13 \times 10^{-3}$	44.34	2700.73
4	$9.66 \times 10^{-9}$	–	$2.88 \times 10^{-6}$	$4.80 \times 10^{-6}$	44.34	2700.73

Table 4:  $c_g = -20$  – Norms of  $(y_k, f_k)$  w.r.t.  $k$  defined by the algorithm (19).

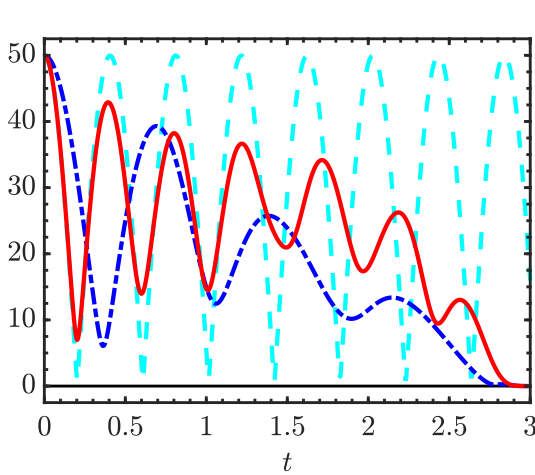


Figure 8:  $c_g = -20$  – (—)  $\|y^*(\cdot, t)\|_{L^2(\Omega)}$  ;  
(---)  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  ; (---)  $\|y(\cdot, t; 0)\|_{L^2(\Omega)}$ .

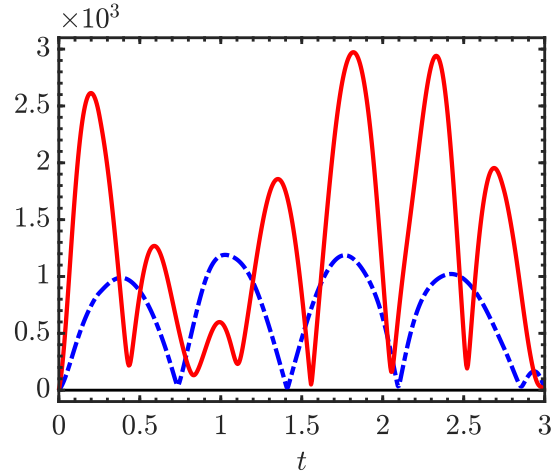


Figure 9:  $c_g = -20$  – (—)  $\|f^*(\cdot, t)\|_{L^2_\chi(\omega)}$  ;  
(---)  $\|f_0(\cdot, t)\|_{L^2_\chi(\omega)}$ .

Table 5 associated with the value  $c_g = 5$  provides a numerical evidence of the convergence of the approximation  $(y_h^*, f_h^*)$  with respect to the value of  $h$ . Actually, in view of the inequality

$$\|f - f_k^h\| \leq \|f - f_k\| + \|f_k - f_k^h\|, \quad \forall k \in \mathbb{N}, \forall h > 0,$$

6 the convergence result stated in Theorem 2 for the sequence  $(f_k)_{k \in \mathbb{N}}$  and the convergence, for any  $k$ , of  
7 the approximation  $(f_k^h)_{h > 0}$  of the linear control  $f_k$  implies that  $f_k^h$  is a finite dimensional approximation  
8 of  $f$ , a control for (1). We observe that the level of the discretization has no influence on the speed of  
9 convergence of the least-squares algorithm.

10 To end this section, we compare our least-squares approach with two fixed-point methods. We first  
11 consider the method associated with the operator  $\Lambda : L^\infty(0, T; L^d(\Omega)) \rightarrow L^\infty(0, T; L^d(\Omega))$  mentioned in  
12 the introduction where  $y = \Lambda(z)$  solves (2). This leads to the algorithm:

$$y_0 \in L^2(Q_T), \quad y_{k+1} = \Lambda(y_k), \quad k \geq 0. \quad (42)$$

13 With the same data and initialization, Table 6 collects some norms with respect to  $k$  for  $c_g = 5$ . The  
14  $L^2_\chi(Q_T)$ -norm of the control is smaller than the one from the least-squares algorithm (967.97 vs 1348.09)



$h$	$n_h^*$	$\ y_h^*\ _{L^2(Q_T)}$	$\ f_h^*\ _{L^2_\chi(q_T)}$	$\mathcal{E}_{T,h}$
1/10	4	27.278	753.111	$8.88 \times 10^{-2}$
1/20	4	31.431	1397.7	$1.00 \times 10^{-2}$
1/40	4	32.026	1353.28	$8.91 \times 10^{-4}$
1/80	4	32.123	1350.24	$1.37 \times 10^{-4}$
1/100	4	32.134	1350.11	$8.27 \times 10^{-5}$
1/120	4	32.139	1350.06	$5.71 \times 10^{-5}$

Table 5:  $c_g = 5$  – Norm of  $(y_h^*, f_h^*)$  w.r.t.  $h$ .

1 but leads to a larger  $L^2(Q_T)$ -norm of the controlled solution (36.901 vs 32.104). The convergence is linear  
2 and reached after  $k^* = 10$  iterations leading to  $\mathcal{E}_T = 2.15 \times 10^{-4}$ . Figure 10 displays the time evolution of  
3 the norms of  $y_{k^*}$  and  $f_{k^*}$  for the final iteration. We observe that the approximation obtained differs from  
4 those of Figures 4 and 5. For these data, the sequence converges for  $|c_g| < 15$  approximately. For larger  
5 values, we observe the non-convergence of the method suggesting that the operator  $\Lambda$  is not contracting  
6 in general.

#iterate $k$	$\sqrt{2E(y_k, f_k)}$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_\chi(q_T)}}{\ f_{k-1}\ _{L^2_\chi(q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_\chi(q_T)}$
0	$3.66 \times 10^2$	–	–	37.653	1339.39
1	$4.73 \times 10^1$	$1.04 \times 10^0$	$1.31 \times 10^0$	37.828	1031.1
2	$2.65 \times 10^0$	$5.83 \times 10^{-2}$	$1.71 \times 10^{-1}$	36.867	972.97
3	$1.70 \times 10^{-1}$	$3.74 \times 10^{-3}$	$1.39 \times 10^{-2}$	36.901	967.508
4	$1.05 \times 10^{-2}$	$2.54 \times 10^{-4}$	$1.02 \times 10^{-3}$	36.9	968.005
5	$2.42 \times 10^{-3}$	$5.27 \times 10^{-5}$	$1.25 \times 10^{-4}$	36.901	967.973
6	$5.20 \times 10^{-4}$	$1.20 \times 10^{-5}$	$2.63 \times 10^{-5}$	36.901	967.972
7	$1.62 \times 10^{-4}$	$3.57 \times 10^{-6}$	$7.13 \times 10^{-6}$	36.901	967.97
8	$4.39 \times 10^{-5}$	$9.84 \times 10^{-7}$	$1.96 \times 10^{-6}$	36.901	967.97
9	$1.28 \times 10^{-5}$	$2.83 \times 10^{-7}$	$5.56 \times 10^{-7}$	36.901	967.97
10	$3.59 \times 10^{-6}$	$7.99 \times 10^{-8}$	$1.57 \times 10^{-7}$	36.901	967.97

Table 6:  $c_g = 5$  – Norms for the sequence defined by the fixed-point algorithm (42).

7 The second fixed-point method is associated with the operator  $\Lambda_F : L^2(Q_T) \rightarrow L^2(Q_T)$  and defined  
8 by  $y = \Lambda_F(z)$ , where  $y$  is a controlled solution of

$$\begin{cases} \partial_{tt}y - \Delta y = f1_\omega - g(z), & \text{in } Q_T, \\ y = 0, & \text{on } \Sigma_T, \\ (y(\cdot, 0), \partial_t y(\cdot, 0)) = (u_0, u_1), & \text{in } \Omega, \end{cases} \quad (43)$$

9 satisfying  $(y(\cdot, T), \partial_t y(\cdot, T)) = (z_0, z_1)$ . The function  $f$  is selected as the control of minimal  $L^2_\chi(q_T)$ -norm.  
10 Any fixed point of  $\Lambda_F$  is a controlled solution for (1). Theorem 1 implies the existence of at least one  
11 fixed point for  $\Lambda_F$ . The controllability of the system (43) allows to define the sequence  $(y_k)_{k \in \mathbb{N}}$  as follows:

$$y_0 \in L^2(Q_T), \quad y_{k+1} = \Lambda_F(y_k), \quad k \geq 0. \quad (44)$$

12 With the same data and initialization, Table 7 collects some norms with respect to  $k$  for  $c_g = 5$ . The  
13 function  $f_k$  is the control of minimal  $L^2_\chi(q_T)$ -norm for  $y_k$  solution of (43). The  $L^2$ -norm of the controlled  
14 pair is greater than the one obtained from the least-squares algorithm. The convergence is significantly  
15 slower and reached after  $k^* = 50$  iterations leading to  $\mathcal{E}_T = 2.02 \times 10^{-4}$ . The convergence is again linear.  
16 Figures 11 depicts the time evolution of the norms of  $y_{k^*}$  and  $f_{k^*}$  for the final iteration. We check that

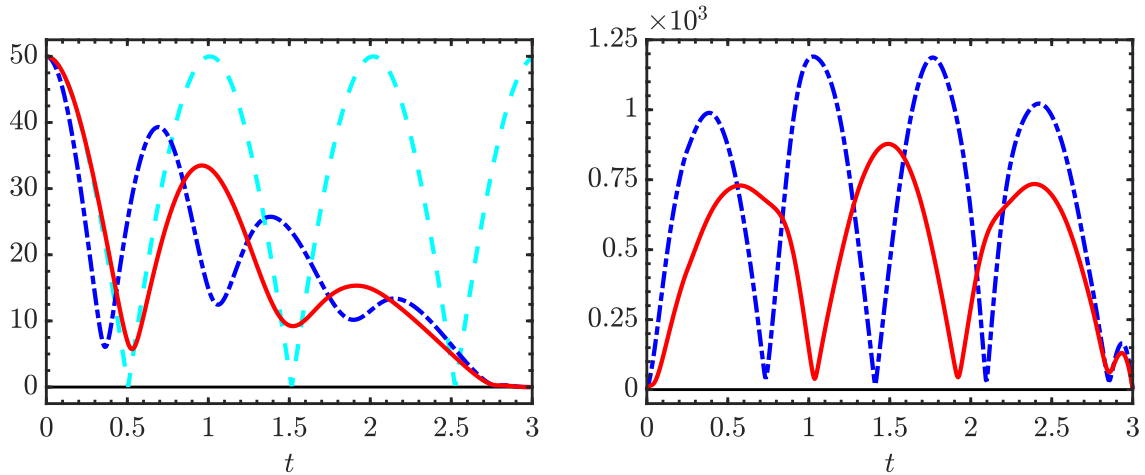


Figure 10: Fixed-point algorithm (42);  $c_g = 5$  – **Left:**  $\|y_{k^*}(\cdot, t)\|_{L^2(\Omega)}$  (—) and  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  (---) vs  $t$ ; **Right:**  $\|f_{k^*}(\cdot, t)\|_{L^2_\chi(\omega)}$  (—) and  $\|f_0(\cdot, t)\|_{L^2_\chi(\omega)}$  (---) vs  $t$ .

- 1 the approximation obtained differs from those of Figures 4 and 5. For these data, the sequence converges
- 2 for  $|c_g| < 7$  approximately. For larger values, we observe the non-convergence of the method suggesting
- 3 that the operator  $\Lambda_F$  is not contracting in general. We refer to [2] where, for  $d = 1$  and boundary
- 4 controllability, the operator  $\Lambda_F$  is proved to be contracting when associated to a different control cost.

#iterate $k$	$\sqrt{2E(y_k, f_k)}$	$\frac{\ y_k - y_{k-1}\ _{L^2(Q_T)}}{\ y_{k-1}\ _{L^2(Q_T)}}$	$\frac{\ f_k - f_{k-1}\ _{L^2_\chi(q_T)}}{\ f_{k-1}\ _{L^2_\chi(q_T)}}$	$\ y_k\ _{L^2(Q_T)}$	$\ f_k\ _{L^2_\chi(q_T)}$
0	$3.66 \times 10^2$	—	—	37.653	1339.39
1	$3.29 \times 10^2$	$8.82 \times 10^{-1}$	$1.64 \times 10^0$	50.427	2333.24
2	$1.97 \times 10^2$	$3.95 \times 10^{-1}$	$7.55 \times 10^{-1}$	51.157	2339.62
3	$7.41 \times 10^1$	$1.41 \times 10^{-1}$	$2.58 \times 10^{-1}$	51.492	2032.67
4	$2.94 \times 10^1$	$5.55 \times 10^{-2}$	$8.26 \times 10^{-2}$	52.085	1988.9
5	$1.58 \times 10^1$	$2.91 \times 10^{-2}$	$2.91 \times 10^{-2}$	52.667	1994.82
6	$1.05 \times 10^1$	$1.91 \times 10^{-2}$	$2.05 \times 10^{-2}$	53.102	2003.63
7	$7.31 \times 10^0$	$1.31 \times 10^{-2}$	$1.27 \times 10^{-2}$	53.43	2009.52
8	$5.19 \times 10^0$	$9.21 \times 10^{-3}$	$9.58 \times 10^{-3}$	53.671	2014.45
9	$3.71 \times 10^0$	$6.53 \times 10^{-3}$	$6.63 \times 10^{-3}$	53.848	2018.1
10	$2.66 \times 10^0$	$4.66 \times 10^{-3}$	$4.84 \times 10^{-3}$	53.978	2020.95
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
46	$2.87 \times 10^{-5}$	$4.87 \times 10^{-8}$	$5.37 \times 10^{-8}$	54.334	2029.72
47	$2.10 \times 10^{-5}$	$3.56 \times 10^{-8}$	$3.93 \times 10^{-8}$	54.334	2029.72
48	$1.54 \times 10^{-5}$	$2.60 \times 10^{-8}$	$2.87 \times 10^{-8}$	54.334	2029.72
49	$1.13 \times 10^{-5}$	$1.91 \times 10^{-8}$	$2.10 \times 10^{-8}$	54.334	2029.72
50	$8.23 \times 10^{-6}$	$1.39 \times 10^{-8}$	$1.54 \times 10^{-8}$	54.334	2029.72

Table 7:  $c_g = 5$  – Norms for the sequence defined by the fixed-point algorithm (44).

### 5.3 Experiments in 1D

- 6 In order to bypass the numerical instabilities observed for large values of  $c_g$  in Subsection 5.2, we employ,
- 7 in the one-dimensional setting, a different method, not based on the minimization of  $J^*$  but on the direct

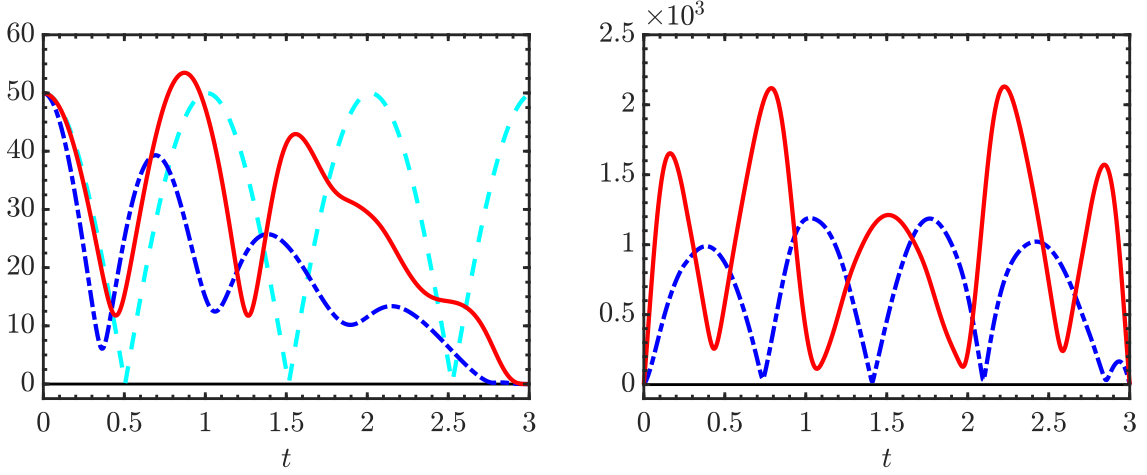


Figure 11: Fixed-point algorithm (44);  $c_g = 5$  – **Left:**  $\|y_{k^*}(\cdot, t)\|_{L^2(\Omega)}$  (—) and  $\|y_0(\cdot, t)\|_{L^2(\Omega)}$  (---) vs  $t$ ; **Right:**  $\|f_{k^*}(\cdot, t)\|_{L^2_\chi(\omega)}$  (—) and  $\|f_0(\cdot, t)\|_{L^2_\chi(\omega)}$  (---) vs  $t$ .

1 approximation of the optimality condition associated with the controllability. We refer to [9, 10] where  
 2 this method has been introduced and to [4] for a numerical analysis.

We consider  $\Omega = (0, 1)$ . The controllability time is equal to  $T = 2.5$  and the control domain is the interval  $\omega = (0.2, 0.4)$ . For any  $c_g \in \mathbb{R}$  and  $\alpha > 0$ , we consider the non-linear function  $g$  defined by

$$g(r) = -c_g r \ln^\alpha(2 + |r|), \quad \forall r \in \mathbb{R}.$$

3 For  $\alpha = 1/2$ , we check that  $g$  satisfies  $(\overline{\mathbf{H}}_s)$  for  $s = 1$  and  $(\mathbf{H}_2)$  for  $|c_g|$  small enough. Moreover,  
 4 in this one-dimensional setting, it is known (see [41]) that the semi-linear wave equation (1) is exactly  
 5 controllable up to  $\alpha = 2$ . As for the initial and final data, we consider  $(u_0, u_1) = (100 \sin(\pi x_1) \sin(\pi x_2), 0)$   
 6 and  $(z_0, z_1) = (0, 0)$  respectively.

7 In order to determine the state-control pairs  $(y_0, f_0)$  and  $(Y_n^1, F_n^1)$  of (38) and (39) respectively, we  
 8 employ the space-time mixed formulation method used in [10].

Concerning problem (38), we set  $\Lambda := L^2(0, T; H_0^1(\Omega))$ ,

$$\Phi_0 := \{\varphi \in L^2(Q_T); \partial_{tt}\varphi - \Delta\varphi \in L^2(0, T; H^{-1}(\Omega)) \text{ in } Q_T, \varphi = 0 \text{ on } \Sigma_T\},$$

and for  $(\varphi, \lambda) \in \Phi_0 \times \Lambda$ , we consider the Lagrangian

$$\begin{aligned} \mathcal{L}_0(\varphi, \lambda) := & \frac{1}{2} \int_{q_T} |\varphi|^2 \chi - \int_0^T \langle \partial_{tt}\varphi - \Delta\varphi, \lambda \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} - \langle u_0, \partial_t\varphi(\cdot, 0) \rangle_{H_0^1(\Omega), H^{-1}(\Omega)} + (u_1, \varphi(\cdot, 0))_{L^2(\Omega)} \\ & + \langle z_0, \partial_t\varphi(\cdot, T) \rangle_{H_0^1(\Omega), H^{-1}(\Omega)} - (z_1, \varphi(\cdot, T))_{L^2(\Omega)}. \end{aligned}$$

9 Then, the control  $f_0$  of minimal  $L^2_\chi(q_T)$ -norm is given by  $f_0 = \varphi_0 \chi$ , where  $(\varphi_0, \lambda_0) \in \Phi_0 \times \Lambda$  is the unique  
 10 saddle point of  $\mathcal{L}_0$ . Note also that it appears that  $\lambda_0$  is the controlled solution  $y_0$  of (38) associated with  
 11 the control  $f_0$ .

Concerning problem (39), we set

$$\Phi_n := \{\varphi \in L^2(Q_T); \partial_{tt}\varphi - \Delta\varphi + g'(y_n)\varphi \in L^2(0, T; H^{-1}(\Omega)) \text{ in } Q_T, \varphi = 0 \text{ on } \Sigma_T\},$$

and for  $(\varphi, \lambda) \in \Phi_n \times \Lambda$ , we consider the Lagrangian

$$\mathcal{L}_n(\varphi, \lambda) := \frac{1}{2} \int_{q_T} |\varphi|^2 \chi - \int_0^T \langle \partial_{tt}\varphi - \Delta\varphi + g'(y_n)\varphi, \lambda \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \int_{Q_T} e_n \varphi.$$

Then, the control  $F_n^1$  of minimal  $L_\chi^2(q_T)$ -norm is given by  $F_n^1 = \varphi_n \chi$ , where  $(\varphi_n, \lambda_n) \in \Phi_n \times \Lambda$  is the unique saddle point of  $\mathcal{L}_n$ . Note also that it appears that  $\lambda_n$  is the controlled solution  $Y_n^1$  of (39) associated with the control  $F_n^1$ .

To approximate the saddle point of  $\mathcal{L}_0$  and  $\mathcal{L}_n$ , we solve a finite-element discretization of the mixed formulation associated with these Lagrangians. We consider a family  $\mathcal{T} = \{\mathcal{T}_h, h > 0\}$  of regular triangulations of  $Q_T$  such that  $\overline{Q_T} = \cup_{K \in \mathcal{T}_h} K$ . The family is indexed by  $h = \max_{K \in \mathcal{T}_h} |K|$ . The functions  $\varphi_0$  and  $\varphi_n$  are approximated in the space  $\Phi_h = \{\varphi_h \in C^1(\overline{Q_T}) \mid \varphi_h|_K \in \mathbb{P}(K), \forall K \in \mathcal{T}_h\}$ , where  $\mathbb{P}(K)$  denotes the reduced Hsieh-Clough-Tocher  $C^1$ -element; while the functions  $\lambda_0$  and  $\lambda_n$  are approximated in the space  $\Lambda_h = \{\lambda_h \in C(\overline{Q_T}) \mid \lambda_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}$ , where  $\mathbb{P}_1(K)$  denotes the space of polynomials of degree 1. In the sequel, we mainly use a regular triangulation  $\mathcal{T}_h$  with fineness  $h = 1/64$ .

We now present some simulations for large values of  $c_g$  in the case  $\alpha = 1/2$  and  $\alpha = 2$ .

- Case  $\alpha = 1/2$ ,  $c_g = 50$  – We set  $\alpha = 1/2$  and we compute the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  associated with the value  $c_g = 50$ . The convergence of the algorithm is observed after  $k^* = 21$  iterations. The corresponding value of the relative error is  $\mathcal{E}_T = 6.24 \times 10^{-17}$ . Figure 12-left depicts the evolution of the error  $\sqrt{2E(y_k, f_k)}$  (red dots, left axis), as well as the evolution of the optimal steps  $\lambda_k$  (blue stars, right axis). Here, we can clearly see the two rates of convergence described in Theorem 2. At first, the values of  $\lambda_k$  are close to 0, while the error decreases linearly. Afterwards, around iteration  $k = 16$ , the values of  $\lambda_k$  reach 1, while the error decreases quadratically. Note that, due to numerical instabilities, the numerical method used in Subsection 5.2 does not converge for  $c_g$  greater than 7.

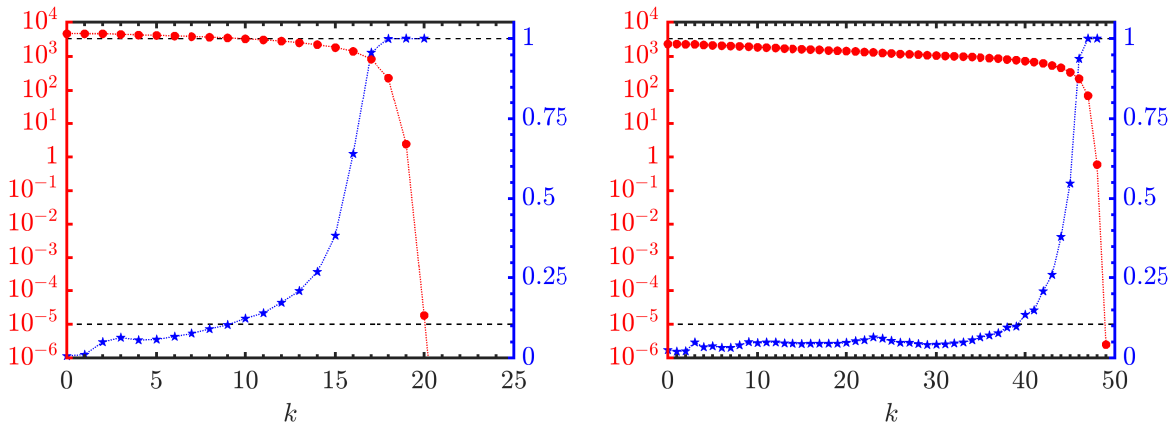


Figure 12: Evolution of  $\sqrt{2E(y_k, f_k)}$  (•, left axis) and  $\lambda_k$  (★, right axis) w.r.t  $k$ ; Left:  $\alpha = 1/2$ ,  $c_g = 50$ ; Right:  $\alpha = 2$ ,  $c_g = 3$ .

- Case  $\alpha = 2$ ,  $c_g = 3$  – We set  $\alpha = 2$  and we compute the sequence  $(y_k, f_k)_{k \in \mathbb{N}}$  associated with the value  $c_g = 3$ . The convergence of the algorithm is observed after  $k^* = 49$  iterations. The corresponding value of the relative error is  $\mathcal{E}_T = 1.18 \times 10^{-17}$ . Figure 12-right depicts the evolution of the error  $\sqrt{2E(y_k, f_k)}$  (red dots, left axis), as well as the evolution of the optimal step  $\lambda_k$  (blue stars, right axis). In this case, the switch between the linear convergence (corresponding to the damped Newton regime) and the quadratic convergence (corresponding to the classical Newton regime) occurs around iteration  $k = 45$ .

## 6 Conclusion

Exact controllability of (1) has been established in [18], under a growth condition on  $g$ , by means of a Leray-Schauder fixed point argument that is not constructive. In this paper, under a slightly stronger growth condition and under the additional assumption that  $g'$  is uniformly Hölder continuous

with exponent  $s \in (0, 1]$ , we have designed an explicit algorithm and proved its convergence of a controlled solution of (1). Moreover, the convergence is super-linear of order greater than or equal to  $1 + s$  after a finite number of iterations. Our approach gives a new and constructive proof of the exact controllability of (1). The method is general and may be applied to any other equations or systems - not necessarily of hyperbolic nature - for which a precise observability estimate for the linearized problem is available: we refer to [24, 27, 37] addressing the case of the heat equation. We also mention the recent extension [2] of this constructive method to address controllability problem (initially investigated in [40]).

Numerical experiments reported are in agreement with the theoretical convergence: in particular, the convergence is, after a finite number of iterations, super-linear to be compared with the linear rate observed with algorithms derived from simpler linearizations. The experiments also confirm that the numerical method developed in [9, 10] and based on the direct resolution of the optimality system (39)-(41) turns out to be very robust with respect to the size of the potential and allows to consider large amplitudes of the nonlinearity, in contrast with the standard minimization of the corresponding conjugate functional (introduced in [19]).

## A Appendix: controllability results for the linearized wave equation

We recall some *a priori* estimates for the linearized wave equation with potential in  $L^\infty(0, T; L^d(\Omega))$  and right hand side in  $L^2(Q_T)$ .

**Proposition 6.** [29, Theorem 2.1] *Assume that  $\omega$  and  $T$  satisfy the assumptions of Theorem 1. For any  $A \in L^\infty(0, T; L^d(\Omega))$ , and  $(\phi_0, \phi_1) \in \mathbf{H} := L^2(\Omega) \times H^{-1}(\Omega)$ , the weak solution  $\phi$  of*

$$\begin{cases} \partial_{tt}\phi - \Delta\phi + A\phi = 0, & \text{in } Q_T, \\ \phi = 0, & \text{on } \Sigma_T, \\ (\phi(\cdot, 0), \partial_t\phi(\cdot, 0)) = (\phi_0, \phi_1), & \text{in } \Omega, \end{cases} \quad (45)$$

satisfies the observability inequality  $\|\phi_0, \phi_1\|_{\mathbf{H}} \leq C e^{C\|A\|_{L^\infty(0, T; L^d(\Omega))}^2} \|\phi\|_{2, Q_T}$  for some  $C > 0$  only depending on  $\Omega$  and  $T$ .

Classical arguments then lead to following controllability result.

**Proposition 7.** *Let  $A \in L^\infty(0, T; L^d(\Omega))$ ,  $B \in L^2(Q_T)$  and  $(z_0, z_1) \in \mathbf{V}$ . Assume that  $\omega$  and  $T$  satisfy the assumptions of Theorem 1. There exists a control function  $u \in L^2(Q_T)$  such that the weak solution of*

$$\begin{cases} \partial_{tt}z - \Delta z + Az = u1_\omega + B, & \text{in } Q_T, \\ z = 0, & \text{on } \Sigma_T, \\ (z(\cdot, 0), \partial_t z(\cdot, 0)) = (z_0, z_1), & \text{in } \Omega, \end{cases} \quad (46)$$

satisfies  $(z(\cdot, T), z_t(\cdot, T)) = (0, 0)$  in  $\Omega$ . Moreover, the unique pair  $(u, z)$  of minimal control norm satisfies

$$\|u\|_{2, Q_T} + \|(z, \partial_t z)\|_{L^\infty(0, T; \mathbf{V})} \leq C \left( \|B\|_2 + \|z_0, z_1\|_{\mathbf{V}} \right) e^{C\|A\|_{L^\infty(0, T; L^d(\Omega))}^2} \quad (47)$$

for some constant  $C > 0$  only depending on  $\Omega$  and  $T$ .

Let  $p^* \in \mathbb{N}^*$  such that  $p^* < \infty$  if  $d = 2$  and  $p^* < 6$  if  $d = 3$ . We next discuss some properties of the operator  $\Lambda : L^\infty(0, T; L^{p^*}(\Omega)) \rightarrow L^\infty(0, T; L^{p^*}(\Omega))$  defined by  $\Lambda(\xi) = y_\xi$ , a null controlled solution of the linear boundary value problem (2) with the control  $f_\xi$  of minimal  $L^2(Q_T)$  norm. Proposition 7 with

1  $B = -g(0)$  gives

$$\|(y_\xi, \partial_t y_\xi)\|_{L^\infty(0,T;\mathbf{V})} \leq C \left( \|u_0, u_1\|_{\mathbf{V}} + \|g(0)\|_2 \right) e^{C\|\hat{g}(\xi)\|_{L^\infty(0,T;L^d(\Omega))}^2} \quad (48)$$

2 where the function  $\hat{g}$  is defined in (2). We assume that  $g \in \mathcal{C}^1(\mathbb{R})$  satisfies the following asymptotic  
3 condition (slightly weaker than  $(\mathbf{H}_1)$ ): there exists a  $\bar{\beta}$  small enough such that  $\limsup_{|r| \rightarrow \infty} \frac{|g(r)|}{|r| \ln^{1/2}|r|} \leq \bar{\beta}$ ,  
4 i.e.

5  $(\hat{\mathbf{H}}_1)$  There exist  $\bar{\alpha} \geq 0$  and  $\bar{\beta} \geq 0$  small enough such that  $|g(r)| \leq \bar{\alpha} + \bar{\beta}(1 + |r|) \ln^{1/2}(1 + |r|)$  for  
6 every  $r \in \mathbb{R}$ .

7 This implies that  $\hat{g}$  satisfies  $|\hat{g}(r)| \leq \bar{\alpha} + \bar{\beta} \ln^{1/2}(1 + |r|)$  for every  $r \in \mathbb{R}$  and some constant  $\bar{\alpha} > 0$ .  
8 This also implies that  $\hat{g}(\xi) \in L^\infty(0,T;L^d(\Omega))$  for any  $\xi \in L^\infty(0,T;L^{p^*}(\Omega))$ . Assuming  $2C\bar{\beta}^2 \leq 1$  and  
9 proceeding as in the proof of Lemma 5, we get

$$e^{C\|\hat{g}(\xi)\|_{L^\infty(0,T;L^d(\Omega))}^2} \leq C_1 \left( 1 + \frac{\|\xi\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\bar{\beta}^2}, \quad \forall \xi \in L^\infty(0,T;L^{p^*}(\Omega))$$

for some  $C_1 = C_1(\alpha)$ . Using (48), we then infer that

$$\|y_\xi\|_{L^\infty(0,T;L^{p^*}(\Omega))} \leq C \left( \|u_0, u_1\|_{\mathbf{V}} + \|g(0)\|_2 \right) C_1 \left( 1 + \frac{\|\xi\|_{L^\infty(0,T;L^{p^*}(\Omega))}}{|\Omega|^{1/p^*}} \right)^{2C\bar{\beta}^2}, \quad \forall \xi \in L^\infty(0,T;L^{p^*}(\Omega)).$$

10 Taking  $\bar{\beta}$  small enough so that  $2C\bar{\beta}^2 < 1$ , we conclude that there exists  $M > 0$  such that  $\|\xi\|_{L^\infty(0,T;L^{p^*}(\Omega))} \leq M$   
11  $M$  implies  $\|\Lambda(\xi)\|_{L^\infty(0,T;L^{p^*}(\Omega))} \leq M$ . This is the argument (introduced in [41] for the one-dimensional  
12 case and) used in [29] to prove the controllability of (1).

## 13 References

- 14 [1] M. Asch and A. Münch. Uniformly controllable schemes for the wave equation on the unit square.  
15 *J. Optim. Theory Appl.*, 143(3):417–438, 2009.
- 16 [2] Kuntal Bhandari, Jérôme Lemoine, and Arnaud Münch. Exact boundary controllability of 1d semi-  
17 linear wave equation through a constructive approach. *To appear in Mathematics of Controls, Signals*  
18 *and Systems*.
- 19 [3] M. O. Bristeau, O. Pironneau, R. Glowinski, J. Periaux, and P. Perrier. On the numerical solution  
20 of nonlinear problems in fluid dynamics by least squares and finite element methods. I. Least square  
21 formulations and conjugate gradient. *Comput. Methods Appl. Mech. Engrg.*, 17(18)(part, part 3):619–  
22 657, 1979.
- 23 [4] Erik Burman, Ali Feizmohammadi, Arnaud Münch, and Lauri Oksanen. Space time finite element  
24 methods for control problem subject to the wave equation. *arXiv 1912.00891*.
- 25 [5] Erik Burman, Ali Feizmohammadi, and Lauri Oksanen. A fully discrete numerical control method  
26 for the wave equation. *SIAM J. Control Optim.*, 58(3):1519–1546, 2020.
- 27 [6] Piermarco Cannarsa, Vilmos Komornik, and Paola Loreti. One-sided and internal controllability of  
28 semilinear wave equations with infinitely iterated logarithms. *Discrete Contin. Dyn. Syst.*, 8(3):745–  
29 756, 2002.
- 30 [7] Carlos Castro, Sorin Micu, and Arnaud Münch. Numerical approximation of the boundary control  
31 for the wave equation with mixed finite elements in a square. *IMA J. Numer. Anal.*, 28(1):186–214,  
32 2008.

- 1 [8] Thierry Cazenave and Alain Haraux. Équations d'évolution avec non linéarité logarithmique. *Ann.*  
2 *Fac. Sci. Toulouse Math. (5)*, 2(1):21–51, 1980.
- 3 [9] Nicolae Cîndea, Enrique Fernández-Cara, and Arnaud Münch. Numerical controllability of the wave  
4 equation through primal methods and Carleman estimates. *ESAIM Control Optim. Calc. Var.*,  
5 19(4):1076–1108, 2013.
- 6 [10] Nicolae Cîndea and Arnaud Münch. A mixed formulation for the direct approximation of the control  
7 of minimal  $L^2$ -norm for linear type wave equations. *Calcolo*, 52(3):245–288, 2015.
- 8 [11] Jean-Michel Coron. *Control and nonlinearity*, volume 136 of *Mathematical Surveys and Monographs*.  
9 American Mathematical Society, Providence, RI, 2007.
- 10 [12] B. Dehman and G. Lebeau. Analysis of the HUM control operator and exact controllability for  
11 semilinear waves in uniform time. *SIAM J. Control Optim.*, 48(2):521–550, 2009.
- 12 [13] Peter Deuffhard. *Newton methods for nonlinear problems*, volume 35 of *Springer Series in Compu-*  
13 *tational Mathematics*. Springer-Verlag, Berlin, 2004. Affine invariance and adaptive algorithms.
- 14 [14] O. Yu. Èmanuilov. Boundary controllability of semilinear evolution equations. *Uspekhi Mat. Nauk*,  
15 44(3(267)):185–186, 1989.
- 16 [15] Sylvain Ervedoza and Enrique Zuazua. A systematic method for building smooth controls for smooth  
17 data. *Discrete Contin. Dyn. Syst. Ser. B*, 14(4):1375–1401, 2010.
- 18 [16] Sylvain Ervedoza and Enrique Zuazua. *Numerical approximation of exact controls for waves*.  
19 SpringerBriefs in Mathematics. Springer, New York, 2013.
- 20 [17] Enrique Fernández-Cara and Arnaud Münch. Numerical null controllability of semi-linear 1-D heat  
21 equations: fixed point, least squares and Newton methods. *Math. Control Relat. Fields*, 2(3):217–246,  
22 2012.
- 23 [18] Xiaoyu Fu, Jiongmin Yong, and Xu Zhang. Exact controllability for multidimensional semilinear  
24 hyperbolic equations. *SIAM J. Control Optim.*, 46(5):1578–1614, 2007.
- 25 [19] Roland Glowinski, Chin Hsien Li, and Jacques-Louis Lions. A numerical approach to the exact  
26 boundary controllability of the wave equation. I. Dirichlet controls: description of the numerical  
27 methods. *Japan J. Appl. Math.*, 7(1):1–76, 1990.
- 28 [20] Roland Glowinski, Jacques-Louis Lions, and Jiwen He. *Exact and approximate controllability for*  
29 *distributed parameter systems*, volume 117 of *Encyclopedia of Mathematics and its Applications*.  
30 Cambridge University Press, Cambridge, 2008. A numerical approach.
- 31 [21] F. Hecht. New development in Freefem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.
- 32 [22] Romain Joly and Camille Laurent. A note on the semiglobal controllability of the semilinear wave  
33 equation. *SIAM J. Control Optim.*, 52(1):439–450, 2014.
- 34 [23] I. Lasiecka and R. Triggiani. Exact controllability of semilinear abstract systems with application  
35 to waves and plates boundary control problems. *Appl. Math. Optim.*, 23(2):109–154, 1991.
- 36 [24] Jérôme Lemoine, Irene Marín-Gayte, and Arnaud Münch. Approximation of null controls for semi-  
37 linear heat equations using a least-squares approach. *ESAIM Control Optim. Calc. Var.*, 27:Paper  
38 No. 63, 28, 2021.
- 39 [25] Jérôme Lemoine and Arnaud Münch. A Fully Space-Time Least-Squares Method for the Unsteady  
40 Navier–Stokes System. *J. Math. Fluid Mech.*, 23(4):Paper No. 102, 2021.

- 1 [26] Jérôme Lemoine and Arnaud Münch. Resolution of the implicit Euler scheme for the Navier-Stokes  
2 equation through a least-squares method. *Numer. Math.*, 147(2):349–391, 2021.
- 3 [27] Jérôme Lemoine and Arnaud Münch. Constructive exact control of semilinear 1d heat equations.  
4 *Mathematical Control and Related Fields*, 2022.
- 5 [28] Jérôme Lemoine, Arnaud Münch, and Pablo Pedregal. Analysis of continuous  $H^{-1}$ -least-squares  
6 methods for the steady Navier-Stokes system. *Appl. Math. Optim.*, 83(1):461–488, 2021.
- 7 [29] Liangyu Li and Xu Zhang. Exact controllability for semilinear wave equations. *J. Math. Anal. Appl.*,  
8 250(2):589–597, 2000.
- 9 [30] J.-L. Lions. *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1*,  
10 volume 8 of *Recherches en Mathématiques Appliquées [Research in Applied Mathematics]*. Masson,  
11 Paris, 1988. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos,  
12 G. Lebeau and J. Rauch.
- 13 [31] Patrick Martinez and Judith Vancostenoble. Exact controllability in “arbitrarily short time” of the  
14 semilinear wave equation. *Discrete Contin. Dyn. Syst.*, 9(4):901–924, 2003.
- 15 [32] Sorin Micu, Ionel Roventă, and Laurențiu Emanuel Temereancă. Approximation of the controls for  
16 the wave equation with a potential. *Numer. Math.*, 144(4):835–887, 2020.
- 17 [33] Arnaud Münch. A least-squares formulation for the approximation of controls for the Stokes system.  
18 *Math. Control Signals Systems*, 27(1):49–75, 2015.
- 19 [34] Arnaud Münch. Approximation of exact controls for semilinear wave and heat equations through  
20 space-time methods. *To appear in Handbook of Numerical analysis, volume 23, Numerical control*  
21 *part B*, 2023.
- 22 [35] Arnaud Münch and Pablo Pedregal. Numerical null controllability of the heat equation through a  
23 least squares and variational approach. *European J. Appl. Math.*, 25(3):277–306, 2014.
- 24 [36] Arnaud Münch and Emmanuel Trélat. Constructive exact control of semilinear 1D wave equations  
25 by a least-squares approach. *SIAM J. Control Optim.*, 60(2):652–673, 2022.
- 26 [37] Ervedoza Sylvain, Lemoine Jérôme, and Arnaud Münch. Exact controllability of semilinear heat  
27 equation through a constructive approach. *To appear in AIMS Evolution equations and control*  
28 *theory*.
- 29 [38] Xu Zhang. Explicit observability estimate for the wave equation with potential and its application.  
30 *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 456(1997):1101–1115, 2000.
- 31 [39] Enrique Zuazua. Exact controllability for the semilinear wave equation. *J. Math. Pures Appl. (9)*,  
32 69(1):1–31, 1990.
- 33 [40] Enrique Zuazua. Exact boundary controllability for the semilinear wave equation. In *Nonlinear*  
34 *partial differential equations and their applications. Collège de France Seminar, Vol. X (Paris, 1987–*  
35 *1988)*, volume 220 of *Pitman Res. Notes Math. Ser.*, pages 357–391. Longman Sci. Tech., Harlow,  
36 1991.
- 37 [41] Enrique Zuazua. Exact controllability for semilinear wave equations in one space dimension. *Ann.*  
38 *Inst. H. Poincaré Anal. Non Linéaire*, 10(1):109–129, 1993.



# D Contrôlabilité de l'élasticité linéaire vu comme un système d'ordre 1 et approximation numérique par une formulation mixte espace-temps stabilisée

Titre : *Controllability of the linear elasticity as a first-order system using a stabilized space-time mixed formulation.*

Auteurs : ARTHUR BOTTOIS et NICOLAE CÎNDEA.

Publié dans *Mathematical Control and Related Fields, Volume 12, 2022.*

<http://doi.org/10.3934/mcrf.2022028>

Aussi disponible à <https://hal.archives-ouvertes.fr/hal-03512656>

## Résumé

Dans cet article, on traite le problème de la contrôlabilité à zéro pour le système de l'élasticité linéaire avec contrôle au bord. On reprend ici les notations de la Sect. 1.3.1 avec  $d \in \{1, 2, 3\}$ . On considère le système d'état (1.63) où le contrôle  $u$  peut agir sur la partie  $\gamma$  du bord. On sait que la contrôlabilité de (1.63) est équivalente à l'inégalité d'observabilité (1.65). Cette inégalité a été établie dans [25, Ch. IV.1] pour des bords de contrôle  $\gamma$  vérifiant certaines hypothèses géométriques analogues au critère (GCC).

L'objectif du présent article est d'approcher numériquement le contrôle  $\hat{u}$  de norme  $L^2(\sigma)$  minimale pour le système (1.63). Pour ce faire, on adapte la méthode exposée dans [26] pour l'équation des ondes. À l'aide du changement de variables  $(w, \mathbf{q}) = (\partial_t \varphi, \nabla \varphi)$ , on réécrit le système adjoint (1.64) sous la forme vitesse-contrainte (1.66). Cette réécriture est intéressante à plusieurs titres. D'une part, le contrôle  $\hat{u}$  s'exprime naturellement en fonction de la déformation  $\nabla \varphi$ . D'autre part, le système (1.66) étant d'ordre 1 en espace et en temps, on peut utiliser des éléments finis  $C^0$  lors de l'approximation numérique de  $\hat{u}$ , en lieu et place d'éléments  $C^1$  pour le système d'ordre 2. On pose  $\mathbf{W} = L^2(\Omega) \times \nabla \mathbf{H}_0^1(\Omega)$  et on note  $\boldsymbol{\sigma}(\mathbf{q}) = \mu \mathbf{q} + (\lambda + \mu) \operatorname{tr}(\mathbf{q}) \mathbf{I}$  le tenseur des contraintes. On vérifie alors que  $\hat{u} = -\boldsymbol{\sigma}(\hat{\mathbf{q}})\nu|_\sigma$ , où  $(\hat{w}, \hat{\mathbf{q}})$  est la solution de (1.66) associée au point de minimum  $(\hat{w}^0, \hat{\mathbf{q}}^0)$  de la fonctionnelle

$$\hat{\mathcal{J}}^*(w^0, \mathbf{q}^0) = \frac{1}{2} \int_\sigma |\boldsymbol{\sigma}(\mathbf{q})\nu|^2 - \langle v^0, w^0 \rangle_{L^2} - \langle \mathbf{p}^0, \mathbf{q}^0 \rangle_{\nabla \mathbf{H}_0^1}, \quad \forall (w^0, \mathbf{q}^0) \in \mathbf{W},$$

où  $(w, \mathbf{q})$  est la solution de (1.66) associée à  $(w^0, \mathbf{q}^0)$ , et avec  $v^0 = y^0$ ,  $\mathbf{p}^0 = -\nabla \mathcal{A}_0^{-1} y^1$ . Pour minimiser la fonctionnelle  $\hat{\mathcal{J}}^*$ , on introduit un multiplicateur de Lagrange  $(v, \boldsymbol{\rho})$  pour les contraintes  $\partial_t w - \operatorname{div} \boldsymbol{\sigma}(\mathbf{q}) = 0$  et  $\partial_t \mathbf{q} - \nabla w = \mathbf{0}$ . On montre alors que  $(\hat{w}, \hat{\mathbf{q}})$  est solution d'une formulation mixte espace-temps de type

$$\begin{cases} a_r((w, \mathbf{q}), (\bar{w}, \bar{\mathbf{q}})) + b((\bar{w}, \bar{\mathbf{q}}), (v, \boldsymbol{\rho})) = \ell(\bar{w}, \bar{\mathbf{q}}), & \forall (\bar{w}, \bar{\mathbf{q}}) \in \mathbf{Z}, \\ b((w, \mathbf{q}), (\bar{v}, \bar{\boldsymbol{\rho}})) = 0, & \forall (\bar{v}, \bar{\boldsymbol{\rho}}) \in \boldsymbol{\Lambda}. \end{cases} \quad (\mathbf{FM}_r)$$

On montre que la formulation mixte est bien posée et qu'en outre, le multiplicateur  $\hat{v}$  associé à  $(\hat{w}, \hat{\mathbf{q}})$  n'est autre que la solution contrôlée  $\hat{y}$  de (1.63) associée à  $\hat{u}$ . Afin de s'affranchir de la condition *inf-sup* qui pèse sur le type d'éléments finis utilisé dans la discrétisation de la formulation mixte, on ajoute un terme de stabilisation au lagrangien  $\mathcal{L}_r$  associé à  $(\mathbf{FM}_r)$ . L'idée est ici d'exploiter le fait que  $(\hat{v}, \hat{\boldsymbol{\rho}})$  est solution du système d'état (1.67) et que  $\hat{v} = \boldsymbol{\sigma}(\hat{\mathbf{q}})\nu$  sur le bord de contrôle  $\sigma$ . On considère ainsi la formulation mixte  $(\mathbf{FM}_{r,\alpha})$  associée au lagrangien stabilisé

$$\mathcal{L}_{r,\alpha}((w, \mathbf{q}), (v, \boldsymbol{\rho})) = \mathcal{L}_r((w, \mathbf{q}), (v, \boldsymbol{\rho})) - \frac{\alpha}{2} \|\boldsymbol{\sigma}(\mathbf{q})\nu - v\|_{L^2(\sigma)}^2 - \frac{\alpha}{2} \|\mathcal{M}(v, \boldsymbol{\rho})\|_{L^2(0,T;\mathbf{W})}^2.$$

On vérifie que les solutions de  $(\mathbf{FM}_r)$  et  $(\mathbf{FM}_{r,\alpha})$  coïncident. On propose enfin une discrétisation de  $(\mathbf{FM}_{r,\alpha})$  et on conclut l'article par quelques simulations numériques en deux dimensions. En se plaçant dans une configuration où l'on connaît explicitement le contrôle  $\hat{u}$ , on observe une convergence de l'approximation  $\hat{u}_h$  vers  $\hat{u}$ . On constate cependant un phénomène de verrouillage numérique quand le paramètre de discrétisation  $h$  tend vers 0.

## Article



## CONTROLLABILITY OF THE LINEAR ELASTICITY AS A FIRST-ORDER SYSTEM USING A STABILIZED SPACE-TIME MIXED FORMULATION

ARTHUR BOTTOIS\* AND NICOLAE CÎNDEA

Université Clermont Auvergne, CNRS  
Laboratoire de Mathématiques Blaise Pascal  
F-63000 Clermont-Ferrand, France

(Communicated by Emmanuel Trélat)

**ABSTRACT.** The aim of this paper is to study the boundary controllability of the linear elasticity system as a first-order system in both space and time. Using the observability inequality known for the usual second-order elasticity system, we deduce an equivalent observability inequality for the associated first-order system. Then, the control of minimal  $L^2$ -norm can be found as the solution to a space-time mixed formulation. This first-order framework is particularly interesting from a numerical perspective since it is possible to solve the space-time mixed formulation using only piecewise linear  $C^0$ -finite elements. Numerical simulations illustrate the theoretical results.

**1. Introduction and notations.** This work is devoted to the study of the boundary null controllability for the linear elasticity system by means of a first-order equivalent formulation. The main motivation to consider this first order framework is a computational one. If there is a rich literature concerning the numerical approximation of controls for the wave equation (see, for instance, [14, 23, 1, 10, 12, 11] and references therein), there are only a few works studying the linear elasticity case (see, for instance, [9, 19]).

The present work is based on the reformulation of the boundary null controllability of the linear elasticity system as a space-time mixed formulation. This idea was already exploited for the controllability of the wave equation in [7, 8]. Notice that, in order to have a conformal finite element approximation of the solution of this mixed formulation for the wave equation,  $C^1$  finite elements are needed. To avoid this issue, in [16], a first-order formulation was introduced for the numerical approximation of the boundary control of the wave equation. Another possibility is to employ a non-conformal stabilized approximation (see, for instance, [4, 5]).

In the sequel we adapt the framework proposed in [16] to the linear elasticity system. This will allow us to numerical approach the boundary null control of minimal  $L^2$  norm for the elasticity system using a space-time stabilized mixed formulation and  $P_1$  finite elements.

---

2020 *Mathematics Subject Classification.* 35Q93, 49K20, 74B05.

*Key words and phrases.* Controllability, linear elasticity, first-order system, space-time mixed formulation, numerical simulation.

\* Corresponding author: Arthur Bottois.

Let  $n \in \mathbb{N}^*$  and let  $\Omega \subset \mathbb{R}^n$  be a non-empty bounded domain with a regular enough boundary  $\Gamma = \partial\Omega$ . For  $T > 0$ , we define  $Q = \Omega \times (0, T)$  and  $\Sigma = \Gamma \times (0, T)$ . We also set  $\Gamma_0 \subset \Gamma$  and  $\Sigma_0 = \Gamma_0 \times (0, T)$ . We then consider the following linear elasticity system:

$$\mathbf{u}'' - \mu \Delta \mathbf{u} - (\lambda + \mu) \nabla \operatorname{div} \mathbf{u} = \mathbf{0} \quad \text{in } Q, \quad (1a)$$

$$\mathbf{u} = \mathbf{h} \quad \text{on } \Sigma_0, \quad (1b)$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Sigma \setminus \Sigma_0, \quad (1c)$$

$$(\mathbf{u}, \mathbf{u}')(\cdot, 0) = (\mathbf{u}^0, \mathbf{u}^1) \quad \text{in } \Omega, \quad (1d)$$

where  $\mathbf{u} = (u_1, \dots, u_n)^\top$  is the state and  $\mathbf{h} = (h_1, \dots, h_n)^\top$  is a boundary control acting on  $\Sigma_0$ . Here and henceforth, we denote by  $\mathbf{u}'$  the derivative of  $\mathbf{u}$  with respect to the time variable  $t$ . Moreover, we use bold characters for the vector-valued functions and the associated spaces. We write for instance  $\mathbf{L}^2(\Omega) = L^2(\Omega; \mathbb{R}^n)$ . Equation (1a), often referred to as the Navier-Cauchy equation, models the evolution of the (small) displacement  $\mathbf{u}$  of an elastic body occupying the domain  $\Omega$ . The constants  $\mu > 0$  and  $\lambda$  satisfying  $\lambda + \mu \geq 0$  are called the Lamé's coefficients. They describe the elastic properties of an isotropic material. For more details related to the model, we refer to the textbook [21].

We first recall some facts about the null controllability of system (1) and the control of minimal  $L^2$ -norm. If the boundary  $\Gamma$  is regular,  $C^2$  being sufficient, for any  $(\mathbf{u}^0, \mathbf{u}^1) \in \mathbf{L}^2(\Omega) \times \mathbf{H}^{-1}(\Omega)$  and  $\mathbf{h} \in \mathbf{L}^2(\Sigma_0)$ , there exists (see, for instance, [15]) a unique transposition solution  $\mathbf{u}$  of (1) with regularity

$$\mathbf{u} \in C([0, T]; \mathbf{L}^2(\Omega)) \cap C^1([0, T]; \mathbf{H}^{-1}(\Omega)). \quad (2)$$

The null controllability problem for (1) in time  $T$  is the following. For each initial datum  $(\mathbf{u}^0, \mathbf{u}^1) \in \mathbf{L}^2(\Omega) \times \mathbf{H}^{-1}(\Omega)$ , find a control  $\mathbf{h} \in \mathbf{L}^2(\Sigma_0)$  such that the corresponding solution  $\mathbf{u}$  of (1) satisfies

$$(\mathbf{u}, \mathbf{u}')(\cdot, T) = (\mathbf{0}, \mathbf{0}) \quad \text{in } \Omega. \quad (3)$$

It is well-known (see, for instance, [15] or [17]) that, under some geometric conditions on  $\Omega$  and  $\Gamma_0$ , system (1) is null controllable at any large time  $T > T^*$ , where  $T^* = \frac{2}{\sqrt{\mu}} R(\mathbf{x}_0)$  and  $R(\mathbf{x}_0)$  depends on the geometry of the domain  $\Omega$  and the control boundary  $\Gamma_0$ . Moreover, as a consequence of the Hilbert Uniqueness Method (HUM) of J.-L. Lions [15], the controllability of (1) is equivalent to an observability inequality for the associated adjoint system

$$\varphi'' - \mu \Delta \varphi - (\lambda + \mu) \nabla \operatorname{div} \varphi = \mathbf{0} \quad \text{in } Q, \quad (4a)$$

$$\varphi = \mathbf{0} \quad \text{on } \Sigma, \quad (4b)$$

$$(\varphi, \varphi')(\cdot, 0) = (\varphi^0, \varphi^1) \quad \text{in } \Omega. \quad (4c)$$

More exactly, there exists a constant  $C_{\text{obs}} > 0$  such that, for any initial datum  $(\varphi^0, \varphi^1) \in \mathcal{W} = \mathbf{H}_0^1(\Omega) \times \mathbf{L}^2(\Omega)$ , the solution  $\varphi$  of (4) satisfies

$$\|(\varphi^0, \varphi^1)\|_{\mathcal{W}}^2 \leq C_{\text{obs}} \int_{\Sigma_0} \left( \mu \left| \frac{\partial \varphi}{\partial \nu} \right|^2 + (\lambda + \mu) |\operatorname{div} \varphi|^2 \right), \quad (5)$$

with

$$\|(\varphi^0, \varphi^1)\|_{\mathcal{W}}^2 = \mu \|\varphi^0\|_{\mathbf{H}_0^1(\Omega)}^2 + (\lambda + \mu) \|\operatorname{div} \varphi^0\|_{L^2(\Omega)}^2 + \|\varphi^1\|_{L^2(\Omega)}^2.$$

As we point out in Remark 1, the boundary term appearing in the right-hand side of inequality (5) is well-defined.

We recall that there is an infinite number of controls in time  $T$ . Thus, the problem of finding the control of minimal  $L^2$ -norm arises naturally. In fact, the control of minimal  $L^2$ -norm is unique and can be found (see [15]) as the solution to a minimization problem involving the conjugate functional

$$\mathcal{J}^*(\varphi^0, \varphi^1) = \frac{1}{2} \int_{\Sigma_0} |\tau(\varphi)|^2 + \langle \mathbf{u}^0, \varphi^1 \rangle_{L^2(\Omega)} - \langle \mathbf{u}^1, \varphi^0 \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1}, \quad (6)$$

with

$$\tau(\varphi) = \mu \frac{\partial \varphi}{\partial \boldsymbol{\nu}} + (\lambda + \mu) \operatorname{div} \varphi \boldsymbol{\nu},$$

and where  $\varphi$  is the solution of the adjoint system (4) associated with  $(\varphi^0, \varphi^1) \in \mathcal{W}$ . We denote by  $\langle \cdot, \cdot \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1}$  the duality pairing between  $\mathbf{H}^{-1}(\Omega)$  and  $\mathbf{H}_0^1(\Omega)$ . For any  $\mathbf{u}^1 \in \mathbf{H}^{-1}(\Omega)$  and  $\varphi^0 \in \mathbf{H}_0^1(\Omega)$ , the pairing can be expressed

$$\langle \mathbf{u}^1, \varphi^0 \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1} = \mu \int_{\Omega} \nabla(\mathcal{A}_0^{-1} \mathbf{u}^1) : \nabla \varphi^0 + (\lambda + \mu) \int_{\Omega} \operatorname{div}(\mathcal{A}_0^{-1} \mathbf{u}^1) \operatorname{div} \varphi^0, \quad (7)$$

where we denote by  $\nabla \varphi^0$  the Jacobian matrix of the vector-valued function  $\varphi^0$  and by  $\mathcal{A}_0^{-1} : \mathbf{H}^{-1}(\Omega) \rightarrow \mathbf{H}_0^1(\Omega)$  the inverse of the operator  $\mathcal{A}_0 = -\mu \Delta - (\lambda + \mu) \nabla \operatorname{div}$ . More precisely, we have  $\mathbf{g} = \mathcal{A}_0^{-1} \mathbf{f}$  if  $\mathbf{g}$  is solution to

$$\begin{aligned} -\mu \Delta \mathbf{g} - (\lambda + \mu) \nabla \operatorname{div} \mathbf{g} &= \mathbf{f} && \text{in } \Omega, \\ \mathbf{g} &= \mathbf{0} && \text{on } \Gamma. \end{aligned}$$

Besides, it is known that the control  $\widehat{\mathbf{h}}$  of minimal  $L^2(\Sigma_0)$ -norm for (1) can be written

$$\widehat{\mathbf{h}} = \tau(\widehat{\varphi})|_{\Sigma_0} = \left( \mu \frac{\partial \widehat{\varphi}}{\partial \boldsymbol{\nu}} + (\lambda + \mu) \operatorname{div} \widehat{\varphi} \boldsymbol{\nu} \right) \Big|_{\Sigma_0}, \quad (8)$$

where  $\widehat{\varphi}$  is the solution of (4) associated with the minimum point  $(\widehat{\varphi}^0, \widehat{\varphi}^1) \in \mathcal{W}$  of the functional  $\mathcal{J}^*$ . We mention that the boundedness by below and the coercivity of  $\mathcal{J}^*$  are consequences of the observability inequality (5).

**Remark 1.** We recall here the following existence, uniqueness and boundary regularity result for the adjoint system. For any  $(\varphi^0, \varphi^1) \in \mathcal{W}$  and  $\mathbf{f} \in L^1(0, T; L^2(\Omega))$ , there exists a unique weak solution

$$\varphi \in C([0, T]; \mathbf{H}_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega)) \quad (9)$$

to the system

$$\varphi'' - \mu \Delta \varphi - (\lambda + \mu) \nabla \operatorname{div} \varphi = \mathbf{f} \quad \text{in } Q, \quad (10a)$$

$$\varphi = \mathbf{0} \quad \text{on } \Sigma, \quad (10b)$$

$$(\varphi, \varphi')(\cdot, 0) = (\varphi^0, \varphi^1) \quad \text{in } \Omega. \quad (10c)$$

Moreover, we have  $\frac{\partial \varphi}{\partial \boldsymbol{\nu}} \in L^2(\Sigma)$ ,  $\operatorname{div} \varphi \in L^2(\Sigma)$  and there exists a positive constant  $C = C(\Omega, T)$  such that

$$\int_{\Sigma} \left( \mu \left| \frac{\partial \varphi}{\partial \boldsymbol{\nu}} \right|^2 + (\lambda + \mu) |\operatorname{div} \varphi|^2 \right) \leq C \left( \|(\varphi^0, \varphi^1)\|_{\mathcal{W}}^2 + \|\mathbf{f}\|_{L^1(0, T; L^2(\Omega))}^2 \right).$$

In view of this regularity result, the boundary term appearing in (5) makes sense.

**Remark 2.** Similarly to the wave equation, the minimal controllability time  $T^*$  blows up when the Lamé's coefficient  $\mu$  goes to 0, and is independent of the value of the second Lamé's coefficient  $\lambda$ . An interesting question which we aim to investigate in a further work is the controllability of the linear elasticity system (1) when the control  $\mathbf{h} = (h_1, \dots, h_n)^\top$  has some components equal to zero. In this situation, since the value of  $\lambda$  quantifies the coupling between the equations in (1), we presume that  $T^*$  will depend on the value of  $\lambda + \mu$ . Indeed, if  $\lambda = -\mu$ , system (1) consists of  $n$  independent wave equations. Hence, controllability cannot hold when the control acts only on some of these equations.

As it was pointed out in [8], we can reformulate the minimization of  $\mathcal{J}^*$  in terms of trajectories of the elasticity system. Thus, there is a bijective relation between the trajectories and the initial data. As a consequence, the function  $\widehat{\varphi}$  appearing in (8) can be found as the minimum point of the new conjugate functional

$$\widetilde{\mathcal{J}}^*(\varphi) = \frac{1}{2} \int_{\Sigma_0} |\tau(\varphi)|^2 + \langle \mathbf{u}^0, \varphi'(\cdot, 0) \rangle_{L^2(\Omega)} - \langle \mathbf{u}^1, \varphi(\cdot, 0) \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1}, \quad (11)$$

where  $\varphi \in \Phi$ , with

$$\Phi = \left\{ \varphi \in L^2(Q); \varphi \text{ satisfies (4a)–(4b), } \frac{\partial \varphi}{\partial \nu} \in L^2(\Sigma) \text{ and } \operatorname{div} \varphi \in L^2(\Sigma) \right\}. \quad (12)$$

The minimization of the functional  $\mathcal{J}^*$  is also exploited for the numerical approximation of the control of minimal  $L^2$ -norm (see, for instance, [12, 11, 23] for the wave equation and [9, 19] for the elasticity system).

Following the ideas employed in [16] for the controllability of the wave equation, we write (4) as a first-order hyperbolic system. More precisely, if  $\varphi$  is a solution of (4a), we define the new variables

$$\mathbf{w} = \varphi' \quad \text{and} \quad \mathbf{Q} = \nabla \varphi, \quad (13)$$

where  $\nabla \varphi$  is the Jacobian matrix of  $\varphi$ . Then, equation (4a) implies that  $(\mathbf{w}, \mathbf{Q})$  solves the following first-order system,

$$\mathbf{w}' - \mu \operatorname{div} \mathbf{Q} - (\lambda + \mu) \nabla \operatorname{tr} \mathbf{Q} = \mathbf{0} \quad \text{in } Q, \quad (14a)$$

$$\mathbf{Q}' - \nabla \mathbf{w} = \mathbf{0} \quad \text{in } Q. \quad (14b)$$

Reciprocally, if we are able to establish the well-posedness of (14) in some suitable functional space, by uniqueness, we may recover solutions of (4a). A slightly different reformulation of the elasticity system as a first order system was proposed in [2]. Note that the new variable  $\mathbf{Q} = (q_{ij})_{1 \leq i, j \leq n}$  is a matrix, that  $\operatorname{tr} \mathbf{Q} = \sum_{i=1}^n q_{ii}$  is the trace of  $\mathbf{Q}$  and that  $\operatorname{div} \mathbf{Q}$  is the vector defined by

$$\operatorname{div} \mathbf{Q} = \left( \sum_{j=1}^n \frac{\partial q_{ij}}{\partial x_j} \right)_{1 \leq i \leq n}. \quad (15)$$

At this point, it is possible to reformulate the minimization problem associated with  $\widetilde{\mathcal{J}}^*$  in terms of the solution to a first-order system. Before doing so, we introduce the initial-boundary value problem associated with (14). First, we must choose the adequate boundary conditions for  $(\mathbf{w}, \mathbf{Q})$ . Since the solution  $\varphi$  of (4) satisfies the boundary condition  $\varphi = \mathbf{0}$  on  $\Sigma$ , we also have  $\varphi' = \mathbf{0}$  on  $\Sigma$ . Thus, we shall impose  $\mathbf{w} = \mathbf{0}$  on  $\Sigma$  in order to be consistent with (13). In fact, this boundary condition is sufficient for our purposes and it is not necessary to impose a boundary condition on  $\mathbf{Q}$ . Regarding the initial condition, the natural choice is to

set  $(\mathbf{w}, \mathbf{Q})(\cdot, 0) = (\varphi^1, \nabla \varphi^0)$ , which means that the initial datum for  $(\mathbf{w}, \mathbf{Q})$  should belong to

$$\mathbf{W} = L^2(\Omega) \times \nabla H_0^1, \quad \text{with} \quad \nabla H_0^1 = \left\{ \nabla \varphi; \varphi \in H_0^1(\Omega) \right\}. \quad (16)$$

The space  $\nabla H_0^1$  is a Hilbert space with respect to the inner product

$$\langle \mathbf{Q}, \overline{\mathbf{Q}} \rangle_{\nabla H_0^1} = \mu \langle \mathbf{Q}, \overline{\mathbf{Q}} \rangle_{L^2(\Omega)} + (\lambda + \mu) \langle \text{tr } \mathbf{Q}, \text{tr } \overline{\mathbf{Q}} \rangle_{L^2(\Omega)}. \quad (17)$$

Moreover,  $\nabla H_0^1$  is a closed subspace of  $L^2(\Omega; M_n(\mathbb{R}))$ . The Hilbert space  $\mathbf{W}$  is endowed with the inner product

$$\langle (\mathbf{w}, \mathbf{Q}), (\overline{\mathbf{w}}, \overline{\mathbf{Q}}) \rangle_{\mathbf{W}} = \langle \mathbf{w}, \overline{\mathbf{w}} \rangle_{L^2(\Omega)} + \langle \mathbf{Q}, \overline{\mathbf{Q}} \rangle_{\nabla H_0^1}. \quad (18)$$

Moreover, for  $\mathbf{Q} \in \nabla H_0^1$ , we denote by  $\boldsymbol{\sigma}(\mathbf{Q})$  the following stress tensor

$$\boldsymbol{\sigma}(\mathbf{Q}) = \mu \mathbf{Q} + (\lambda + \mu) \text{tr } \mathbf{Q} \mathbf{I}, \quad (19)$$

where  $\mathbf{I}$  is the identity matrix. Remark that equation (14a) then writes

$$\mathbf{w}' - \text{div } \boldsymbol{\sigma}(\mathbf{Q}) = \mathbf{0} \quad \text{in } Q. \quad (20)$$

We also define

$$\mathbf{V} = H_0^1(\Omega) \times \nabla H_0^2, \quad \text{with} \quad \nabla H_0^2 = \left\{ \nabla \varphi; \varphi \in H^2(\Omega) \cap H_0^1(\Omega) \right\}. \quad (21)$$

The space  $\mathbf{W}$  is the closure of  $\mathbf{V}$  with respect to the norm  $\|\cdot\|_{\mathbf{W}}$ . Besides, for any  $\mathbf{Q} \in \nabla H_0^2$ , note that  $\text{div } \boldsymbol{\sigma}(\mathbf{Q}) \in L^2(\Omega)$ . With these notations, the first-order initial-boundary value problem associated with (4) is the following. For any  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$ , find  $(\mathbf{w}, \mathbf{Q})$  solution of (14) and satisfying the following boundary and initial value conditions

$$\mathbf{w} = \mathbf{0} \quad \text{on } \Sigma, \quad (22a)$$

$$(\mathbf{w}, \mathbf{Q})(\cdot, 0) = (\mathbf{w}^0, \mathbf{Q}^0) \quad \text{in } \Omega. \quad (22b)$$

In what follows, we denote

$$\mathcal{M}(\mathbf{w}, \mathbf{Q}) = \left( \mathbf{w}' - \text{div } \boldsymbol{\sigma}(\mathbf{Q}), \mathbf{Q}' - \nabla \mathbf{w} \right). \quad (23)$$

In order to state a suitable well-posedness result for the first-order system (14)–(22), we consider the closed and densely defined operator  $\mathcal{A} : \mathbf{V} \rightarrow \mathbf{W}$  defined by

$$\mathcal{A}(\mathbf{w}, \mathbf{Q}) = \left( \text{div } \boldsymbol{\sigma}(\mathbf{Q}), \nabla \mathbf{w} \right), \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{V}. \quad (24)$$

We show in Appendix A that  $\mathcal{A}$  generates a  $C_0$ -group of contractions which we denote  $(\mathbb{T}_t)_{t \in \mathbb{R}}$ . Using this semigroup approach (see, for instance, [18] or [22]), we can introduce the notion of mild solution.

**Definition 1.1.** Let  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^1(0, T; \mathbf{W})$ . We say that  $(\mathbf{w}, \mathbf{Q})$  is a mild solution of

$$\mathbf{w}' - \text{div } \boldsymbol{\sigma}(\mathbf{Q}) = \mathbf{f} \quad \text{in } Q, \quad (25a)$$

$$\mathbf{Q}' - \nabla \mathbf{w} = \mathbf{F} \quad \text{in } Q, \quad (25b)$$

$$\mathbf{w} = \mathbf{0} \quad \text{on } \Sigma, \quad (25c)$$

$$(\mathbf{w}, \mathbf{Q})(\cdot, 0) = (\mathbf{w}^0, \mathbf{Q}^0) \quad \text{in } \Omega, \quad (25d)$$

if

$$(\mathbf{w}, \mathbf{Q})(\cdot, t) = \mathbb{T}_t(\mathbf{w}^0, \mathbf{Q}^0) + \int_0^t \mathbb{T}_{t-s}(\mathbf{f}, \mathbf{F})(\cdot, s) ds, \quad \forall t \in (0, T). \quad (26)$$

As a consequence of well-known facts concerning the theory of semigroups and the definition of mild solutions, we have the following well-posedness statement for (25).

**Lemma 1.2.** *Let  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^1(0, T; \mathbf{W})$ .*

- i) The first-order system (25) has a unique mild solution  $(\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{W})$  and there exists a positive constant  $C = C(\Omega, T)$  such that*

$$\sup_{t \in (0, T)} \|(\mathbf{w}, \mathbf{Q})(\cdot, t)\|_{\mathbf{W}}^2 \leq C \left( \|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right). \quad (27)$$

- ii) If  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{V}$  and  $(\mathbf{f}, \mathbf{F}) \in C^1([0, T]; \mathbf{W})$ , then the mild solution satisfies  $(\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{V}) \cap C^1([0, T]; \mathbf{W})$  and the equation  $\mathcal{M}(\mathbf{w}, \mathbf{Q}) = (\mathbf{f}, \mathbf{F})$  holds in  $\mathbf{W}$ .*

Similarly to the boundary regularity result stated in Remark 1, we have a boundary regularity result for the solutions of (25). Indeed, we prove in Lemma 2.2 that if  $(\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{W})$  is a solution of (25), then  $\mathbf{Q}\nu \in L^2(\Sigma)$ ,  $\text{tr } \mathbf{Q} \in L^2(\Sigma)$  and the inequality

$$\int_{\Sigma} \left( \mu |\mathbf{Q}\nu|^2 + (\lambda + \mu) |\text{tr } \mathbf{Q}|^2 \right) \leq C \left( \|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right)$$

holds true for some positive constant  $C = C(\Omega, T)$ .

We are now in position to restate the minimization problem associated with  $\tilde{\mathcal{J}}^*$  (cf (11)) in terms of the solution to the first-order system (25). A natural way to proceed is to consider a minimization problem in which the quantities  $\varphi'$  and  $\nabla \varphi$  are replaced by  $\mathbf{w}$  and  $\mathbf{Q}$  respectively. We then define the following spaces of trajectories,

$$\mathbf{Z} = \left\{ (\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{W}) \text{ mild solution of (25) associated with } \right. \quad (28)$$

$$\left. (\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}, (\mathbf{f}, \mathbf{F}) \in L^2(0, T; \mathbf{W}) \right\},$$

$$\mathbf{Z}_0 = \left\{ (\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{W}) \text{ mild solution of (25) associated with } \right. \quad (29)$$

$$\left. (\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}, (\mathbf{f}, \mathbf{F}) = (\mathbf{0}, \mathbf{0}) \right\}.$$

With these definitions, the adequate minimization problem in terms of the first-order system is

$$\min_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0} \hat{\mathcal{J}}^*(\mathbf{w}, \mathbf{Q}), \quad (30)$$

where

$$\hat{\mathcal{J}}^*(\mathbf{w}, \mathbf{Q}) = \frac{1}{2} \int_{\Sigma_0} |\sigma(\mathbf{Q})\nu|^2 + \langle \mathbf{u}^0, \mathbf{w}(\cdot, 0) \rangle_{L^2(\Omega)} - \langle \nabla(\mathcal{A}_0^{-1} \mathbf{u}^1), \mathbf{Q}(\cdot, 0) \rangle_{\nabla H_0^1}. \quad (31)$$

As for the elasticity system, the boundedness by below and the coercivity of  $\hat{\mathcal{J}}^*$  can be obtained as a consequence of an observability inequality which takes the form

$$\|(\mathbf{w}, \mathbf{Q})(\cdot, 0)\|_{\mathbf{W}}^2 \leq \mu^{-1} C_{\text{obs}} \int_{\Sigma_0} |\sigma(\mathbf{Q})\nu|^2,$$

for any  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0$  and under suitable conditions on  $\Omega$ ,  $\Gamma_0$ ,  $(\mu, \lambda)$  and  $T$ .

The remaining part of this work is structured as follows. In Section 2, we prove an a priori estimate and a generalized observability inequality for system (25). The third section rewrites the minimization problem (30) as a mixed formulation. The

well-posedness of this mixed formulation is a direct consequence of the estimates proved in Section 2 and provides an easily computable expression for the control of minimal  $L^2$ -norm associated with (1). Sections 4 and 5 propose a stabilized version of the mixed formulation introduced in Section 3 and an equivalent reformulation of this stabilized mixed formulation as a dual extremal problem involving only the state variable. Such a reformulation is particularly interesting from the numerical perspective and is exploited in Section 6 to numerically illustrate the controllability for the wave equation and the elasticity system written as first-order systems in both space and time.

**2. Boundary regularity and observability.** We first derive an identity which is useful to obtain a priori estimates for (25).

**Lemma 2.1.** *With  $\Omega$  of class  $C^2$ , let  $\beta \in C^1(\bar{\Omega}; \mathbb{R}^n)$  be a vector field. For any solution  $(\mathbf{w}, \mathbf{Q})$  of (25) associated with  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{V}$  and  $(\mathbf{f}, \mathbf{F}) \in C^1([0, T]; \mathbf{W})$ , the following identity holds,*

$$\begin{aligned} & \frac{1}{2} \int_{\Sigma} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right) (\beta \cdot \boldsymbol{\nu}) = \\ & \left[ \int_{\Omega} \mathbf{w} \cdot (\mathbf{Q}\beta) \right]_0^T + \frac{1}{2} \iint_Q \left( |\mathbf{w}|^2 - \mu |\mathbf{Q}|^2 - (\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right) \operatorname{div} \beta \\ & + \mu \iint_Q \mathbf{Q} : (\mathbf{Q}\nabla\beta) + (\lambda + \mu) \iint_Q \operatorname{tr} \mathbf{Q} \operatorname{tr}(\mathbf{Q}\nabla\beta) \\ & - \iint_Q \mathbf{f} \cdot (\mathbf{Q}\beta) - \iint_Q \mathbf{w} \cdot (\mathbf{F}\beta). \end{aligned} \quad (32)$$

*Proof.* Let  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{V}$  and  $(\mathbf{f}, \mathbf{F}) \in C^1([0, T]; \mathbf{W})$ . There exists two functions  $\varphi^0 \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$  and  $\psi \in C^1([0, T]; \mathbf{H}_0^1(\Omega))$  such that  $\mathbf{Q}^0 = \nabla\varphi^0$  and  $\mathbf{F} = \nabla\psi$ . We start by some properties of  $(\mathbf{w}, \mathbf{Q})$  on the boundary  $\Sigma$ . The matrix  $\mathbf{Q}$  satisfies

$$\mathbf{Q}' - \nabla\mathbf{w} = \mathbf{F} \quad \text{in } Q \quad \text{and} \quad \mathbf{Q}(\cdot, 0) = \mathbf{Q}^0 \quad \text{in } \Omega,$$

therefore

$$\mathbf{Q}(\mathbf{x}, t) = \nabla\varphi^0(\mathbf{x}) + \int_0^t \nabla\mathbf{w}(\mathbf{x}, \tau) \, d\tau + \int_0^t \nabla\psi(\mathbf{x}, \tau) \, d\tau, \quad \forall (\mathbf{x}, t) \in Q. \quad (33)$$

Let  $1 \leq i \leq n$ . Since  $\varphi^0 \in \mathbf{H}_0^1(\Omega)$ , we have  $\nabla\varphi_i^0 = \frac{\partial\varphi_i^0}{\partial\nu}\boldsymbol{\nu}$  on  $\Gamma$ . Similarly, for any  $\tau \in [0, T]$ , since  $\mathbf{w}(\cdot, \tau) = \mathbf{0}$  on  $\Gamma$  and  $\psi(\cdot, \tau) \in \mathbf{H}_0^1(\Omega)$ , we also have  $\nabla w_i(\cdot, \tau) = \frac{\partial w_i}{\partial\nu}(\cdot, \tau)\boldsymbol{\nu}$  and  $\nabla\psi_i(\cdot, \tau) = \frac{\partial\psi_i}{\partial\nu}(\cdot, \tau)\boldsymbol{\nu}$  on  $\Gamma$ . Hence, if  $\mathbf{q}_i$  denotes the  $i$ -th row of  $\mathbf{Q}$ , it follows from (33) that  $\mathbf{q}_i = (\mathbf{q}_i \cdot \boldsymbol{\nu})\boldsymbol{\nu}$  on  $\Sigma$ . Thus, on  $\Sigma$ , we get

$$|\mathbf{Q}|^2 = |\mathbf{Q}\boldsymbol{\nu}|^2, \quad (\mathbf{Q}\boldsymbol{\nu}) \cdot \boldsymbol{\nu} = \operatorname{tr} \mathbf{Q}, \quad \mathbf{Q}\beta = (\beta \cdot \boldsymbol{\nu}) \mathbf{Q}\boldsymbol{\nu}, \quad (\mathbf{Q}\beta) \cdot \boldsymbol{\nu} = \operatorname{tr} \mathbf{Q} (\beta \cdot \boldsymbol{\nu}). \quad (34)$$

Multiplying equation (25a) by  $\mathbf{Q}\beta$  and integrating over  $Q$ , we have

$$\iint_Q \mathbf{w}' \cdot (\mathbf{Q}\beta) - \mu \iint_Q \operatorname{div} \mathbf{Q} \cdot (\mathbf{Q}\beta) - (\lambda + \mu) \iint_Q \nabla \operatorname{tr} \mathbf{Q} \cdot (\mathbf{Q}\beta) = \iint_Q \mathbf{f} \cdot (\mathbf{Q}\beta). \quad (35)$$

We treat each term one after the other. We recall that there exists a function  $\varphi \in C^1([0, T]; \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega))$  such that  $\mathbf{Q} = \nabla\varphi$ .



- For the first term, by integration by parts, we get

$$\begin{aligned} \iint_Q \mathbf{w}' \cdot (\mathbf{Q}\beta) &= \left[ \int_\Omega \mathbf{w} \cdot (\mathbf{Q}\beta) \right]_0^T - \iint_Q \mathbf{w} \cdot (\mathbf{Q}'\beta) \\ &= \left[ \int_\Omega \mathbf{w} \cdot (\mathbf{Q}\beta) \right]_0^T - \iint_Q \mathbf{w} \cdot (\nabla \mathbf{w} \beta) - \iint_Q \mathbf{w} \cdot (\mathbf{F}\beta). \end{aligned}$$

Moreover, since  $\mathbf{w} = \mathbf{0}$  on  $\Sigma$ , we have

$$\iint_Q \mathbf{w} \cdot (\nabla \mathbf{w} \beta) = \frac{1}{2} \iint_Q \nabla |\mathbf{w}|^2 \cdot \beta = -\frac{1}{2} \iint_Q |\mathbf{w}|^2 \operatorname{div} \beta.$$

Thus, we get

$$\iint_Q \mathbf{w}' \cdot (\mathbf{Q}\beta) = \left[ \int_\Omega \mathbf{w} \cdot (\mathbf{Q}\beta) \right]_0^T + \frac{1}{2} \iint_Q |\mathbf{w}|^2 \operatorname{div} \beta - \iint_Q \mathbf{w} \cdot (\mathbf{F}\beta). \quad (36)$$

- For the second term, by integration by parts and since  $\mathbf{Q} = \nabla \varphi$ , we get

$$\begin{aligned} \iint_Q \operatorname{div} \mathbf{Q} \cdot (\mathbf{Q}\beta) &= \int_\Sigma (\mathbf{Q}\nu) \cdot (\mathbf{Q}\beta) - \iint_Q \mathbf{Q} : \nabla (\mathbf{Q}\beta) \\ &= \int_\Sigma |\mathbf{Q}\nu|^2 (\beta \cdot \nu) - \iint_Q \mathbf{Q} : (\nabla \mathbf{Q}\beta) - \iint_Q \mathbf{Q} : (\mathbf{Q}\nabla \beta). \end{aligned}$$

Moreover, we have

$$\begin{aligned} \iint_Q \mathbf{Q} : (\nabla \mathbf{Q}\beta) &= \frac{1}{2} \iint_Q \nabla |\mathbf{Q}|^2 \cdot \beta = \frac{1}{2} \int_\Sigma |\mathbf{Q}|^2 (\beta \cdot \nu) - \frac{1}{2} \iint_Q |\mathbf{Q}|^2 \operatorname{div} \beta \\ &= \frac{1}{2} \int_\Sigma |\mathbf{Q}\nu|^2 (\beta \cdot \nu) - \frac{1}{2} \iint_Q |\mathbf{Q}|^2 \operatorname{div} \beta. \end{aligned}$$

Thus, we get

$$\iint_Q \operatorname{div} \mathbf{Q} \cdot (\mathbf{Q}\beta) = \frac{1}{2} \int_\Sigma |\mathbf{Q}\nu|^2 (\beta \cdot \nu) + \frac{1}{2} \iint_Q |\mathbf{Q}|^2 \operatorname{div} \beta - \iint_Q \mathbf{Q} : (\mathbf{Q}\nabla \beta). \quad (37)$$

- For the third term, by integration by parts and since  $\mathbf{Q} = \nabla \varphi$ , we get

$$\begin{aligned} \iint_Q \nabla \operatorname{tr} \mathbf{Q} \cdot (\mathbf{Q}\beta) &= \int_\Sigma \operatorname{tr} \mathbf{Q} (\mathbf{Q}\beta) \cdot \nu - \iint_Q \operatorname{tr} \mathbf{Q} \operatorname{div} (\mathbf{Q}\beta) \\ &= \int_\Sigma |\operatorname{tr} \mathbf{Q}|^2 (\beta \cdot \nu) - \iint_Q \operatorname{tr} \mathbf{Q} \nabla \operatorname{tr} \mathbf{Q} \cdot \beta - \iint_Q \operatorname{tr} \mathbf{Q} \operatorname{tr} (\mathbf{Q}\nabla \beta). \end{aligned}$$

Moreover, we have

$$\iint_Q \operatorname{tr} \mathbf{Q} \nabla \operatorname{tr} \mathbf{Q} \cdot \beta = \frac{1}{2} \iint_Q \nabla |\operatorname{tr} \mathbf{Q}|^2 \cdot \beta = \frac{1}{2} \int_\Sigma |\operatorname{tr} \mathbf{Q}|^2 (\beta \cdot \nu) - \frac{1}{2} \iint_Q |\operatorname{tr} \mathbf{Q}|^2 \operatorname{div} \beta.$$

Thus, we get

$$\iint_Q \nabla \operatorname{tr} \mathbf{Q} \cdot (\mathbf{Q}\beta) = \frac{1}{2} \int_\Sigma |\operatorname{tr} \mathbf{Q}|^2 (\beta \cdot \nu) + \frac{1}{2} \iint_Q |\operatorname{tr} \mathbf{Q}|^2 \operatorname{div} \beta - \iint_Q \operatorname{tr} \mathbf{Q} \operatorname{tr} (\mathbf{Q}\nabla \beta). \quad (38)$$

Replacing (36) to (38) in (35), we obtain identity (32).  $\square$

We recall that, if  $\mathbf{Q} \in \nabla \mathbf{H}_0^2$ , then  $\sigma(\mathbf{Q})\nu = \mu \mathbf{Q}\nu + (\lambda + \mu) \operatorname{tr} \mathbf{Q}\nu$  is well-defined as a distribution in  $\mathbf{H}^{-\frac{1}{2}}(\Gamma)$  and Green formula is satisfied,

$$\int_\Omega \operatorname{div} \sigma(\mathbf{Q}) \cdot \bar{\mathbf{w}} + \int_\Omega \sigma(\mathbf{Q}) : \nabla \bar{\mathbf{w}} = \langle \sigma(\mathbf{Q})\nu, \bar{\mathbf{w}} \rangle_{\mathbf{H}^{-\frac{1}{2}}, \mathbf{H}^{\frac{1}{2}}}, \quad \forall \bar{\mathbf{w}} \in \mathbf{H}^1(\Omega), \quad (39)$$

where  $\langle \cdot, \cdot \rangle_{\mathbf{H}^{-\frac{1}{2}}, \mathbf{H}^{\frac{1}{2}}}$  denotes the duality pairing between  $\mathbf{H}^{-\frac{1}{2}}(\Gamma)$  and  $\mathbf{H}^{\frac{1}{2}}(\Gamma)$ . The next lemma improves the regularity of  $\sigma(\mathbf{Q})\boldsymbol{\nu}$  when  $(\mathbf{w}, \mathbf{Q})$  is a solution of (25) belonging to  $C([0, T]; \mathbf{V}) \cap C^1([0, T]; \mathbf{W})$  and allows to define  $\sigma(\mathbf{Q})\boldsymbol{\nu}$  when  $(\mathbf{w}, \mathbf{Q})$  is a mild solution of (25).

**Lemma 2.2.** *With  $\Omega$  of class  $C^2$ , let  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{V}$ ,  $(\mathbf{f}, \mathbf{F}) \in C^1([0, T]; \mathbf{W})$  and let  $(\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{V}) \cap C^1([0, T]; \mathbf{W})$  be the corresponding solution of (25). Then,  $\mathbf{Q}\boldsymbol{\nu} \in \mathbf{L}^2(\Sigma)$ ,  $\text{tr } \mathbf{Q} \in L^2(\Sigma)$  and there exists a positive constant  $C = C(\Omega, T)$  such that*

$$\int_{\Sigma} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\text{tr } \mathbf{Q}|^2 \right) \leq C \left( \|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right). \quad (40)$$

Moreover, we have  $\sigma(\mathbf{Q})\boldsymbol{\nu} \in \mathbf{L}^2(\Sigma)$  and

$$\int_{\Sigma} |\sigma(\mathbf{Q})\boldsymbol{\nu}|^2 \leq (\lambda + 3\mu) C \left( \|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right). \quad (41)$$

Consequently, there is a unique bounded map

$$\Lambda : \mathbf{W} \times L^1(0, T; \mathbf{W}) \rightarrow \mathbf{L}^2(\Sigma)$$

such that  $\Lambda((\mathbf{w}^0, \mathbf{Q}^0), (\mathbf{f}, \mathbf{F})) = \sigma(\mathbf{Q})\boldsymbol{\nu}$  when  $(\mathbf{w}, \mathbf{Q}) \in C([0, T]; \mathbf{V}) \cap C^1([0, T]; \mathbf{W})$ . Accordingly, we denote  $\sigma(\mathbf{Q})\boldsymbol{\nu} = \Lambda((\mathbf{w}^0, \mathbf{Q}^0), (\mathbf{f}, \mathbf{F}))$  when  $(\mathbf{w}, \mathbf{Q})$  is the mild solution of (25) associated with  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^1(0, T; \mathbf{W})$ . In that case, estimate (41) still holds.

*Proof.* Due to the regularity of the boundary  $\Gamma$ , there exists a vector field  $\boldsymbol{\beta} \in C^1(\bar{\Omega}; \mathbb{R}^n)$  such that  $\boldsymbol{\beta} = \boldsymbol{\nu}$  on  $\Gamma$ . Thus, with this choice of  $\boldsymbol{\beta}$ , identity (32) implies

$$\int_{\Sigma} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\text{tr } \mathbf{Q}|^2 \right) \leq C \left( \sup_{t \in (0, T)} \|(\mathbf{w}, \mathbf{Q})(\cdot, t)\|_{\mathbf{W}}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right).$$

We then apply estimate (27) to obtain (40). Besides, from (34), we know that  $(\mathbf{Q}\boldsymbol{\nu}) \cdot \boldsymbol{\nu} = \text{tr } \mathbf{Q}$  on  $\Sigma$ . It follows that

$$\begin{aligned} \int_{\Sigma} |\sigma(\mathbf{Q})\boldsymbol{\nu}|^2 &= \int_{\Sigma} \left| \mu \mathbf{Q}\boldsymbol{\nu} + (\lambda + \mu) \text{tr } \mathbf{Q} \boldsymbol{\nu} \right|^2 \\ &= \int_{\Sigma} \left( \mu^2 |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu)^2 |\text{tr } \mathbf{Q}|^2 + 2\mu(\lambda + \mu) \text{tr } \mathbf{Q} (\mathbf{Q}\boldsymbol{\nu}) \cdot \boldsymbol{\nu} \right) \\ &= \int_{\Sigma} \left( \mu^2 |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + 3\mu)(\lambda + \mu) |\text{tr } \mathbf{Q}|^2 \right) \\ &\leq (\lambda + 3\mu) \int_{\Sigma} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\text{tr } \mathbf{Q}|^2 \right). \end{aligned}$$

□

**Remark 3.** The above lemma can also be proved using the following Rellich-Nečas identities. If  $\mathbf{q}, \boldsymbol{\beta} \in C^1(\bar{\Omega}; \mathbb{R}^n)$  are two vector fields, we can check that

$$\begin{aligned} 2 \text{div}((\mathbf{q} \cdot \boldsymbol{\beta})\mathbf{q}) - \text{div}(|\mathbf{q}|^2\boldsymbol{\beta}) &= 2(\mathbf{q} \cdot \boldsymbol{\beta}) \text{div } \mathbf{q} - |\mathbf{q}|^2 \text{div } \boldsymbol{\beta} + 2(\nabla \boldsymbol{\beta} \mathbf{q}) \cdot \mathbf{q} \\ &\quad + 2\left( (\nabla \mathbf{q} \mathbf{q}) \cdot \boldsymbol{\beta} - (\nabla \mathbf{q} \boldsymbol{\beta}) \cdot \mathbf{q} \right). \end{aligned}$$

Consequently, for  $\mathbf{Q} \in C^1(\bar{\Omega}; M_n(\mathbb{R}))$ , we can adapt the previous identity in order to get

$$\begin{aligned} 2 \operatorname{div}(\mathbf{Q}^\top (\mathbf{Q}\boldsymbol{\beta})) - \operatorname{div}(|\mathbf{Q}|^2 \boldsymbol{\beta}) &= 2(\mathbf{Q}\boldsymbol{\beta}) \cdot \operatorname{div} \mathbf{Q} - |\mathbf{Q}|^2 \operatorname{div} \boldsymbol{\beta} + 2(\mathbf{Q}\nabla \boldsymbol{\beta}) : \mathbf{Q} \\ &\quad + 2(\boldsymbol{\beta}^\top \nabla \mathbf{Q}^\top - \nabla \mathbf{Q} \boldsymbol{\beta}) : \mathbf{Q}. \end{aligned}$$

Similarly, we show that

$$\begin{aligned} 2 \operatorname{div}((\mathbf{Q}\boldsymbol{\beta}) \operatorname{tr} \mathbf{Q}) - \operatorname{div}(|\operatorname{tr} \mathbf{Q}|^2 \boldsymbol{\beta}) &= \\ 2(\mathbf{Q}\boldsymbol{\beta}) \cdot \nabla \operatorname{tr} \mathbf{Q} - |\operatorname{tr} \mathbf{Q}|^2 \operatorname{div} \boldsymbol{\beta} + 2 \operatorname{tr}(\mathbf{Q}\nabla \boldsymbol{\beta}) \operatorname{tr} \mathbf{Q} \\ &\quad + 2(\operatorname{tr}(\boldsymbol{\beta}^\top \nabla \mathbf{Q}^\top) - \operatorname{tr}(\nabla \mathbf{Q} \boldsymbol{\beta})) \operatorname{tr} \mathbf{Q}. \end{aligned}$$

Finally, note that if  $\mathbf{Q} \in \nabla \mathbf{H}_0^2$ , we have  $\boldsymbol{\beta}^\top \nabla \mathbf{Q}^\top = \nabla \mathbf{Q} \boldsymbol{\beta}$ .

**Lemma 2.3.** *We recall that  $\mathbf{Z}_0$  is the set of mild solutions  $(\mathbf{w}, \mathbf{Q})$  of (22), with  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$ . For any  $T > T^*$ , there exists a constant  $C_{\text{obs}} > 0$  such that*

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq C_{\text{obs}} \int_{\Sigma_0} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right), \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0. \quad (42)$$

The critical time  $T^* > 0$  depends on  $\Omega$ ,  $\Gamma_0$  and  $(\lambda, \mu)$ . Besides, we have

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq \mu^{-1} C_{\text{obs}} \int_{\Sigma_0} |\boldsymbol{\sigma}(\mathbf{Q})\boldsymbol{\nu}|^2, \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0. \quad (43)$$

*Proof.* To obtain (42), we use the observability inequality (5) associated with the adjoint system (4). Let  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$ . Then, there exists  $\varphi^0 \in \mathbf{H}_0^1(\Omega)$  such that  $\mathbf{Q}^0 = \nabla \varphi^0$ . We set  $\varphi^1 = \mathbf{w}^0$  and we define  $(\mathbf{w}, \mathbf{Q}) = (\varphi', \nabla \varphi)$ , where  $\varphi$  is the solution of (4) associated with  $(\varphi^0, \varphi^1)$ . If we apply the observability inequality (5) to  $\varphi$ , we get that for  $T > T^*$ , the following estimate holds,

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq C_{\text{obs}} \int_{\Sigma_0} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right).$$

Finally, it is easy to check that  $(\mathbf{w}, \mathbf{Q})$  is the mild solution of (22) associated with  $(\mathbf{w}^0, \mathbf{Q}^0)$ . Besides, it follows from (34) that

$$\begin{aligned} \int_{\Sigma_0} |\boldsymbol{\sigma}(\mathbf{Q})\boldsymbol{\nu}|^2 &= \int_{\Sigma_0} \left( \mu^2 |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + 3\mu)(\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right) \\ &\geq \mu \int_{\Sigma_0} \left( \mu |\mathbf{Q}\boldsymbol{\nu}|^2 + (\lambda + \mu) |\operatorname{tr} \mathbf{Q}|^2 \right). \end{aligned}$$

□

**Proposition 1.** *We recall that  $\mathbf{Z}$  is the set of mild solutions  $(\mathbf{w}, \mathbf{Q})$  of (25), with  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^2(0, T; \mathbf{W})$ . For any  $T > T^*$ , there exists a positive constant  $C = C(\Omega, T)$  such that*

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq C \left( \|\boldsymbol{\sigma}(\mathbf{Q})\boldsymbol{\nu}\|_{L^2(\Sigma_0)}^2 + \|(\mathbf{f}, \mathbf{F})\|_{L^1(0, T; \mathbf{W})}^2 \right), \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}. \quad (44)$$

*Proof.* We can split a solution  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  as  $(\mathbf{w}, \mathbf{Q}) = (\mathbf{w}_1, \mathbf{Q}_1) + (\mathbf{w}_2, \mathbf{Q}_2)$ , where  $(\mathbf{w}_1, \mathbf{Q}_1)$  solves (25) with  $(\mathbf{f}, \mathbf{F}) = (\mathbf{0}, \mathbf{0})$  and  $(\mathbf{w}_2, \mathbf{Q}_2)$  solves (25) with  $(\mathbf{w}^0, \mathbf{Q}^0) = (\mathbf{0}, \mathbf{0})$ . By Lemma 2.3 and Lemma 2.2, we have respectively

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq \mu^{-1} C_{\text{obs}} \|\boldsymbol{\sigma}(\mathbf{Q}_1)\boldsymbol{\nu}\|_{L^2(\Sigma_0)}^2$$

and

$$\|\sigma(\mathbf{Q}_2)\nu\|_{L^2(\Sigma_0)}^2 \leq (\lambda + 3\mu) C \|(\mathbf{f}, \mathbf{F})\|_{L^1(0,T;\mathbf{W})}^2.$$

We then get the result using

$$\|\sigma(\mathbf{Q}_1)\nu\|_{L^2(\Sigma_0)}^2 \leq 2\|\sigma(\mathbf{Q})\nu\|_{L^2(\Sigma_0)}^2 + 2\|\sigma(\mathbf{Q}_2)\nu\|_{L^2(\Sigma_0)}^2.$$

□

**Remark 4.** i) Since  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  is a mild solution of (25) for some  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^2(0, T; \mathbf{W})$ , by Lemma 2.2, the normal trace  $\sigma(\mathbf{Q})\nu \in L^2(\Sigma_0)$  is well-defined.

ii) For any  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$ , there exists two unique pairs  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$  and  $(\mathbf{f}, \mathbf{F}) \in L^2(0, T; \mathbf{W})$  such that  $(\mathbf{w}, \mathbf{Q})$  is the corresponding solution of (25). We then denote  $\mathcal{M}(\mathbf{w}, \mathbf{Q}) = (\mathbf{f}, \mathbf{F})$ . With this notation, we have

$$\mathbf{Z}_0 = \left\{ (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}; \quad \mathcal{M}(\mathbf{w}, \mathbf{Q}) = (\mathbf{0}, \mathbf{0}) \right\}.$$

iii) Note that the generalized observability inequality (44) implies

$$\|(\mathbf{w}^0, \mathbf{Q}^0)\|_{\mathbf{W}}^2 \leq C \left( \|\sigma(\mathbf{Q})\nu\|_{L^2(\Sigma_0)}^2 + \|\mathcal{M}(\mathbf{w}, \mathbf{Q})\|_{L^2(0,T;\mathbf{W})}^2 \right), \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}. \quad (45)$$

For any fixed  $\eta > 0$ , we endow  $\mathbf{Z}$  with the inner product

$$\langle (\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \rangle_{\mathbf{Z}} = \int_{\Sigma_0} (\sigma(\mathbf{Q})\nu) \cdot (\sigma(\bar{\mathbf{Q}})\nu) + \eta \int_0^T \langle \mathcal{M}(\mathbf{w}, \mathbf{Q}), \mathcal{M}(\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \rangle_{\mathbf{W}}. \quad (46)$$

We denote  $\|\cdot\|_{\mathbf{Z}}$  the associated norm.

**Lemma 2.4.** *The space  $\mathbf{Z}$  endowed with the inner product  $\langle \cdot, \cdot \rangle_{\mathbf{Z}}$  is a Hilbert space.*

*Proof.* We check that  $\mathbf{Z}$  is closed with respect to the norm  $\|\cdot\|_{\mathbf{Z}}$ . To do so, let  $\{(\mathbf{w}_k, \mathbf{Q}_k)\}_{k \geq 0} \subset \mathbf{Z}$  be a sequence converging to some  $(\mathbf{w}, \mathbf{Q})$  for the norm  $\|\cdot\|_{\mathbf{Z}}$ . Then,  $\mathcal{M}(\mathbf{w}_k, \mathbf{Q}_k) = (\mathbf{f}_k, \mathbf{F}_k)$  converges to some  $(\mathbf{f}, \mathbf{F}) \in L^2(0, T; \mathbf{W})$  and the generalized observability inequality (45) implies that  $(\mathbf{w}_k^0, \mathbf{Q}_k^0)$  converges to some  $(\mathbf{w}^0, \mathbf{Q}^0) \in \mathbf{W}$ . By the mapping properties of the semigroup which defines the mild solutions of (25), we get that  $(\mathbf{w}_k, \mathbf{Q}_k)$  converges to the mild solution of (25) associated with  $(\mathbf{w}^0, \mathbf{Q}^0)$  and  $(\mathbf{f}, \mathbf{F})$ . We therefore have  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$ . □

**3. Mixed formulation.** Problem (30) is a constrained minimization problem in  $\mathbf{Z}$  where the constraint is the equation solved by  $(\mathbf{w}, \mathbf{Q})$ , i.e.  $\mathcal{M}(\mathbf{w}, \mathbf{Q}) = (\mathbf{0}, \mathbf{0})$ . We take into account this constraint by introducing a Lagrange multiplier  $(\zeta, \Theta) \in \Lambda = L^2(0, T; \mathbf{W})$ . It leads to the following mixed formulation. Find  $((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) \in \mathbf{Z} \times \Lambda$  solution of

$$a((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) + b((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta, \Theta)) = \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}, \quad (47a)$$

$$b((\mathbf{w}, \mathbf{Q}), (\bar{\zeta}, \bar{\Theta})) = 0, \quad \forall (\bar{\zeta}, \bar{\Theta}) \in \Lambda, \quad (47b)$$

where

$$a : \mathbf{Z} \times \mathbf{Z} \rightarrow \mathbb{R}, \quad a((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = \langle \sigma(\mathbf{Q})\nu, \sigma(\bar{\mathbf{Q}})\nu \rangle_{L^2(\Sigma_0)}, \quad (48)$$

$$b : \mathbf{Z} \times \Lambda \rightarrow \mathbb{R}, \quad b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = \langle \mathcal{M}(\mathbf{w}, \mathbf{Q}), (\zeta, \Theta) \rangle_{L^2(0,T;\mathbf{W})}, \quad (49)$$

$$\ell : \mathbf{Z} \rightarrow \mathbb{R}, \quad \ell(\mathbf{w}, \mathbf{Q}) = \langle (-\mathbf{u}^0, \nabla(-\mathcal{A}_0)^{-1}\mathbf{u}^1), (\mathbf{w}, \mathbf{Q})(\cdot, 0) \rangle_{\mathbf{W}}. \quad (50)$$

**Theorem 3.1.** *We have the following.*

i) *The mixed formulation (47) is well-posed.*

ii) Let  $\mathcal{L} : \mathbf{Z} \times \mathbf{\Lambda} \rightarrow \mathbb{R}$  be the Lagrangian defined by

$$\mathcal{L}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = \frac{1}{2}a((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) + b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) - \ell(\mathbf{w}, \mathbf{Q}). \quad (51)$$

Then, the unique solution  $((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) \in \mathbf{Z} \times \mathbf{\Lambda}$  of (47) is solution to the saddle-point problem

$$\sup_{(\zeta, \Theta) \in \mathbf{\Lambda}} \inf_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \mathcal{L}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)).$$

iii) The solution  $(\mathbf{w}, \mathbf{Q})$  of (47) is the minimum point of  $\widehat{\mathcal{J}}^*$  over  $\mathbf{Z}_0$ . Hence, the control of minimal  $\mathbf{L}^2(\Sigma_0)$ -norm for (1) is given by

$$\mathbf{h} = \sigma(\mathbf{Q})\boldsymbol{\nu}|_{\Sigma_0}. \quad (52)$$

Besides, the Lagrange multiplier  $\zeta$  is the associated controlled state of (1). This means that  $\zeta$  is the solution of (1) in the sense of transposition associated with the initial datum  $(\mathbf{u}^0, \mathbf{u}^1)$  and the control  $\mathbf{h}$ , and that  $\zeta$  satisfies  $(\zeta, \zeta')(\cdot, T) = (\mathbf{0}, \mathbf{0})$  in  $\Omega$ .

**Remark 5.** We can also consider the augmented Lagrangian  $\mathcal{L}_r$ , which is defined for any  $r = (r_1, r_2)$ ,  $r_1, r_2 > 0$ , by

$$\mathcal{L}_r((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = \frac{1}{2}a_r((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) + b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) - \ell(\mathbf{w}, \mathbf{Q}), \quad (53)$$

$$a_r((\mathbf{w}, \mathbf{Q}), (\overline{\mathbf{w}}, \overline{\mathbf{Q}})) = a((\mathbf{w}, \mathbf{Q}), (\overline{\mathbf{w}}, \overline{\mathbf{Q}})) + r_1 \langle \mathcal{M}(\mathbf{w}, \mathbf{Q}), \mathcal{M}(\overline{\mathbf{w}}, \overline{\mathbf{Q}}) \rangle_{L^2(0, T; \mathbf{W})} + r_2 \langle \mathbf{Q}\boldsymbol{\nu}^\perp, \overline{\mathbf{Q}}\boldsymbol{\nu}^\perp \rangle_{L^2(\Sigma_0)}. \quad (54)$$

Since  $a_r((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) = a((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q}))$  for any  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0$ , the Lagrangians  $\mathcal{L}$  and  $\mathcal{L}_r$  share the same saddle point. Remark that the augmented term associated with  $r_2$  is zero. Nevertheless, we choose to add here this term which will be useful for the numerical implementation. Indeed, we need to impose  $\mathbf{Q} \in \nabla \mathbf{H}_0^1$  for the numerical spaces. In dimension two,  $\boldsymbol{\nu}^\perp$  is any unit vector orthogonal to  $\boldsymbol{\nu}$ , hence if the  $r_2$  is non zero, then this augmentation term impose that the tangent part of  $\mathbf{Q}$  is close to zero.

*Proof.* i) In virtue of [3, Theorem 4.2.1], we have to check the following properties.

- 1) The bilinear forms  $a$  and  $b$  are continuous. It is obvious from the definition of  $\|\cdot\|_{\mathbf{Z}}$ .
- 2) The linear form  $\ell$  is continuous. It is a direct consequence of the generalized observability inequality (45).
- 3) The form  $a$  is coercive on the kernel

$$\mathcal{N}(b) = \left\{ (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}; \quad b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = 0, \forall (\zeta, \Theta) \in \mathbf{\Lambda} \right\}.$$

It is clear from the definition of  $a$ . If  $(\mathbf{w}, \mathbf{Q}) \in \mathcal{N}(b) = \mathbf{Z}_0$ , we then have  $a((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) = \|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{Z}}^2$ .

- 4) The form  $b$  satisfies an *inf-sup* condition over  $\mathbf{Z} \times \mathbf{\Lambda}$ . There exists a constant  $\delta > 0$  such that

$$\inf_{(\zeta, \Theta) \in \mathbf{\Lambda}} \sup_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \frac{b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta))}{\|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{Z}} \|(\zeta, \Theta)\|_{\mathbf{\Lambda}}} \geq \delta. \quad (55)$$

We fix  $(\widehat{\zeta}, \widehat{\theta}) \in \mathbf{\Lambda}$  and consider the unique  $(\widehat{w}, \widehat{Q}) \in \mathbf{Z}$  such that  $\mathcal{M}(\widehat{w}, \widehat{Q}) = (\widehat{\zeta}, \widehat{\theta})$  and  $(\widehat{w}, \widehat{Q})(\cdot, 0) = (\mathbf{0}, \mathbf{0})$ . The trace inequality (41) implies

$$\|\sigma(\widehat{Q})\nu\|_{L^2(\Sigma_0)} \leq C\|(\widehat{\zeta}, \widehat{\theta})\|_{\mathbf{\Lambda}} = C\|\mathcal{M}(\widehat{w}, \widehat{Q})\|_{\mathbf{\Lambda}}. \quad (56)$$

We also have  $b((\widehat{w}, \widehat{Q}), (\widehat{\zeta}, \widehat{\theta})) = \|(\widehat{\zeta}, \widehat{\theta})\|_{\mathbf{\Lambda}}^2$ . Therefore,

$$\begin{aligned} \sup_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \frac{b((\mathbf{w}, \mathbf{Q}), (\widehat{\zeta}, \widehat{\theta}))}{\|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{Z}} \|(\widehat{\zeta}, \widehat{\theta})\|_{\mathbf{\Lambda}}} &\geq \frac{\|(\widehat{\zeta}, \widehat{\theta})\|_{\mathbf{\Lambda}}}{\|(\widehat{w}, \widehat{Q})\|_{\mathbf{Z}}} = \frac{\|\mathcal{M}(\widehat{w}, \widehat{Q})\|_{\mathbf{\Lambda}}}{\|(\widehat{w}, \widehat{Q})\|_{\mathbf{Z}}} \\ &= \frac{\|\mathcal{M}(\widehat{w}, \widehat{Q})\|_{\mathbf{\Lambda}}}{\left(\|\sigma(\widehat{Q})\nu\|_{L^2(\Sigma_0)}^2 + \eta\|\mathcal{M}(\widehat{w}, \widehat{Q})\|_{\mathbf{\Lambda}}^2\right)^{\frac{1}{2}}} \geq \frac{1}{\sqrt{C^2 + \eta}}, \end{aligned}$$

where the last inequality is a consequence of (56). Thus, inequality (55) holds with  $\delta = (C^2 + \eta)^{-\frac{1}{2}}$ .

ii) This property is due to the symmetry and positivity of  $a$ .

iii) If  $((\mathbf{w}, \mathbf{Q}), (\zeta, \theta)) \in \mathbf{Z} \times \mathbf{\Lambda}$  solves the mixed formulation (47), then equation (47b) implies  $\mathcal{M}(\mathbf{w}, \mathbf{Q}) = (\mathbf{0}, \mathbf{0})$ . Consequently, we have  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}_0$  and  $\mathcal{L}((\mathbf{w}, \mathbf{Q}), (\zeta, \theta)) = \widehat{\mathcal{J}}^*(\mathbf{w}, \mathbf{Q})$ .

Besides, equation (47a) holds for any  $(\overline{w}, \overline{Q}) \in \mathbf{Z}$ . In particular, if choosing  $\mathbf{h}$  given by (52), we have

$$\int_{\Sigma_0} \mathbf{h} \cdot (\sigma(\overline{Q})\nu) + \iint_Q \zeta \cdot \overline{\mathbf{f}} = \langle (-\mathbf{u}^0, \nabla(-\mathcal{A}_0)^{-1}\mathbf{u}^1), (\overline{w}, \overline{Q})(\cdot, 0) \rangle_{\mathbf{W}}, \quad (57)$$

for any  $(\overline{w}, \overline{Q}) \in \mathbf{Z}$  such that  $\mathcal{M}(\overline{w}, \overline{Q}) = (\overline{\mathbf{f}}, \mathbf{0})$ , with  $\overline{\mathbf{f}} \in L^2(Q)$ . Let  $(\overline{w}, \overline{Q}) \in \mathbf{Z}$  with  $\mathcal{M}(\overline{w}, \overline{Q}) = (\overline{\mathbf{f}}, \mathbf{0})$  for some  $\overline{\mathbf{f}} \in L^2(Q)$ . Then, it is easy to check that  $(\overline{w}, \overline{Q})$  can be realized as  $(\overline{w}, \overline{Q}) = (\overline{\varphi}', \nabla \overline{\varphi})$ , where  $\overline{\varphi} \in C([0, T]; \mathbf{H}_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$  is the unique weak solution of

$$\begin{aligned} \overline{\varphi}'' - \mu \Delta \overline{\varphi} - (\lambda + \mu) \nabla \operatorname{div} \overline{\varphi} &= \overline{\mathbf{f}} && \text{in } Q, \\ \overline{\varphi} &= \mathbf{0} && \text{on } \Sigma, \\ (\overline{\varphi}, \overline{\varphi}')(\cdot, T) &= (\mathbf{0}, \mathbf{0}) && \text{in } \Omega. \end{aligned} \quad (58)$$

Hence, the formulation (57) is equivalent to

$$\begin{aligned} \int_{\Sigma_0} \mathbf{h} \cdot \tau(\overline{\varphi}) + \iint_Q \zeta \cdot \left( \overline{\varphi}'' - \mu \Delta \overline{\varphi} - (\lambda + \mu) \nabla \operatorname{div} \overline{\varphi} \right) &= \\ &= - \langle \mathbf{u}^0, \overline{\varphi}'(\cdot, 0) \rangle_{L^2(\Omega)} + \langle \mathbf{u}^1, \overline{\varphi}(\cdot, 0) \rangle_{\mathbf{H}^{-1}, \mathbf{H}_0^1}, \end{aligned}$$

for any solution  $\overline{\varphi}$  of (58), with  $\overline{\mathbf{f}} \in L^2(Q)$ . It means that  $\zeta$  is a transposition solution of (1), so  $\zeta$  is the controlled state associated with the initial datum  $(\mathbf{u}^0, \mathbf{u}^1)$  and the control  $\mathbf{h}$ .  $\square$

**4. Stabilized mixed formulation.** In this section, we introduce a stabilized mixed formulation equivalent to (47). This stabilization adds some uniform coercivity with respect to the multiplier variable, which allows to bypass the Babuška-Brezzi *inf-sup* condition. We define the Hilbert space

$$\begin{aligned} \Xi &= \left\{ (\zeta, \theta) \in C([0, T]; \mathbf{W}) \text{ mild solution of (60) associated with} \right. \\ &\quad \left. (\zeta^0, \theta^0) \in \mathbf{W}, \mathbf{h} \in L^2(\Sigma_0), (\mathbf{g}, \mathbf{G}) \in L^2(0, T; \mathbf{W}) \right\}. \end{aligned} \quad (59)$$

For  $(\zeta^0, \Theta^0) \in \mathbf{W}$ ,  $\mathbf{h} \in L^2(\Sigma_0)$  and  $(\mathbf{g}, \mathbf{G}) \in L^2(0, T; \mathbf{W})$ , we say that  $(\zeta, \Theta) \in C([0, T]; \mathbf{W})$  is a mild solution of

$$\begin{aligned} \mathcal{M}(\zeta, \Theta) &= (\mathbf{g}, \mathbf{G}) && \text{in } Q, \\ \zeta &= \mathbf{h} && \text{on } \Sigma_0, \\ \zeta &= \mathbf{0} && \text{on } \Sigma \setminus \Sigma_0, \\ (\zeta, \Theta)(\cdot, 0) &= (\zeta^0, \Theta^0) && \text{in } \Omega, \end{aligned} \quad (60)$$

if

$$(\zeta, \Theta)(\cdot, t) = \mathbb{T}_t(\zeta^0, \Theta^0) + \int_0^t \mathbb{T}_{t-s} \mathcal{B} \mathbf{h}(\cdot, s) ds + \int_0^t \mathbb{T}_{t-s} (\mathbf{g}, \mathbf{G})(\cdot, s) ds, \quad (61)$$

for any  $t \in [0, T]$ , where  $\mathcal{B} : L^2(\Gamma_0) \rightarrow \mathbf{V}'$  is defined by transposition

$$\langle \mathcal{B} \mathbf{h}, (\mathbf{w}, \mathbf{Q}) \rangle_{\mathbf{V}', \mathbf{V}} = -\langle \mathbf{h}, \sigma(\mathbf{Q}) \boldsymbol{\nu} \rangle_{L^2(\Gamma_0)}, \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{V}.$$

The space  $\Xi$  is endowed with the norm given by the following inner product,

$$\begin{aligned} \langle (\zeta, \Theta), (\bar{\zeta}, \bar{\Theta}) \rangle_{\Xi} &= \langle \zeta, \bar{\zeta} \rangle_{L^2(\Sigma_0)} + \langle \mathcal{M}(\zeta, \Theta), \mathcal{M}(\bar{\zeta}, \bar{\Theta}) \rangle_{L^2(0, T; \mathbf{W})} \\ &+ \langle (\zeta, \Theta)(\cdot, 0), (\bar{\zeta}, \bar{\Theta})(\cdot, 0) \rangle_{\mathbf{W}} + \langle (\zeta, \Theta)(\cdot, T), (\bar{\zeta}, \bar{\Theta})(\cdot, T) \rangle_{\mathbf{W}}. \end{aligned} \quad (62)$$

Let  $r = (r_1, r_2)$ ,  $r_1, r_2 > 0$ , and  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ ,  $\alpha_1 \in (0, 1)$ ,  $\alpha_2, \alpha_3, \alpha_4 > 0$ . We want to find  $((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) \in \mathbf{Z} \times \Xi$  such that

$$a_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) + b_{\alpha}((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta, \Theta)) = \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}, \quad (63a)$$

$$b_{\alpha}((\mathbf{w}, \mathbf{Q}), (\bar{\zeta}, \bar{\Theta})) - c_{\alpha}((\zeta, \Theta), (\bar{\zeta}, \bar{\Theta})) = d_{\alpha}(\bar{\zeta}, \bar{\Theta}), \quad \forall (\bar{\zeta}, \bar{\Theta}) \in \Xi, \quad (63b)$$

where  $a_{r, \alpha} : \mathbf{Z} \times \mathbf{Z} \rightarrow \mathbb{R}$ ,  $b_{\alpha} : \mathbf{Z} \times \Xi \rightarrow \mathbb{R}$ ,  $c_{\alpha} : \Xi \times \Xi \rightarrow \mathbb{R}$  and  $d_{\alpha} : \Xi \rightarrow \mathbb{R}$  are defined by

$$a_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = a_r((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) - \alpha_1 \langle \sigma(\mathbf{Q}) \boldsymbol{\nu}, \sigma(\bar{\mathbf{Q}}) \boldsymbol{\nu} \rangle_{L^2(\Sigma_0)}, \quad (64)$$

$$b_{\alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = b((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) + \alpha_1 \langle \sigma(\mathbf{Q}) \boldsymbol{\nu}, \zeta \rangle_{L^2(\Sigma_0)}, \quad (65)$$

$$\begin{aligned} c_{\alpha}((\zeta, \Theta), (\bar{\zeta}, \bar{\Theta})) &= \alpha_1 \langle \zeta, \bar{\zeta} \rangle_{L^2(\Sigma_0)} + \alpha_2 \langle \mathcal{M}(\zeta, \Theta), \mathcal{M}(\bar{\zeta}, \bar{\Theta}) \rangle_{L^2(0, T; \mathbf{W})} \\ &+ \alpha_3 \langle (\zeta, \Theta)(\cdot, 0), (\bar{\zeta}, \bar{\Theta})(\cdot, 0) \rangle_{\mathbf{W}} + \alpha_4 \langle (\zeta, \Theta)(\cdot, T), (\bar{\zeta}, \bar{\Theta})(\cdot, T) \rangle_{\mathbf{W}}, \end{aligned} \quad (66)$$

$$d_{\alpha}(\zeta, \Theta) = \alpha_3 \langle (-\mathbf{u}^0, \nabla(-\mathcal{A}_0)^{-1} \mathbf{u}^1), (\zeta, \Theta)(\cdot, 0) \rangle_{\mathbf{W}}. \quad (67)$$

In view of [3, Remark 4.3.1], if  $((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) \in \mathbf{Z} \times \Xi$  solves (63), then it is solution to the saddle-point problem

$$\inf_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \sup_{(\zeta, \Theta) \in \Xi} \mathcal{L}_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)),$$

with the Lagrangian  $\mathcal{L}_{r, \alpha} : \mathbf{Z} \times \Xi \rightarrow \mathbb{R}$  given by

$$\begin{aligned} \mathcal{L}_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) &= \\ &\mathcal{L}_r((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) - \frac{\alpha_1}{2} \|\sigma(\mathbf{Q}) \boldsymbol{\nu} - \zeta\|_{L^2(\Sigma_0)}^2 - \frac{\alpha_2}{2} \|\mathcal{M}(\zeta, \Theta)\|_{L^2(0, T; \mathbf{W})}^2 \\ &- \frac{\alpha_3}{2} \|(-\mathbf{u}^0, -\nabla(-\mathcal{A}_0)^{-1} \mathbf{u}^1) - (\zeta, \Theta)(\cdot, 0)\|_{\mathbf{W}}^2 - \frac{\alpha_4}{2} \|(\zeta, \Theta)(\cdot, T)\|_{\mathbf{W}}^2 \\ &= \frac{1}{2} a_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) + b_{\alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) \\ &\quad - \frac{1}{2} c_{\alpha}((\zeta, \Theta), (\zeta, \Theta)) - \ell(\mathbf{w}, \mathbf{Q}) - d_{\alpha}(\zeta, \Theta), \end{aligned}$$

where  $\mathcal{L}_r$  is the augmented Lagrangian defined in (53).

**Remark 6.** Using integration by parts, we could find an alternative expression for  $b_\alpha$ . For  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  and  $(\zeta, \Theta) \in \Xi$ , we have

$$\begin{aligned} b_\alpha((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) &= -\langle (\mathbf{w}, \mathbf{Q}), \mathcal{M}(\zeta, \Theta) \rangle_{L^2(0, T; \mathbf{W})} - (1 - \alpha_1) \langle \sigma(\mathbf{Q})\nu, \zeta \rangle_{L^2(\Sigma_0)} \\ &\quad - \langle (\mathbf{w}, \mathbf{Q})(\cdot, 0), (\zeta, \Theta)(\cdot, 0) \rangle_{\mathbf{W}} + \langle (\mathbf{w}, \mathbf{Q})(\cdot, T), (\zeta, \Theta)(\cdot, T) \rangle_{\mathbf{W}}. \end{aligned}$$

This expression may be useful for the numerical simulations. This idea is similar to the partitioned finite elements methods (see, for instance, [6, 20]).

**Theorem 4.1.** For any  $r = (r_1, r_2)$ ,  $r_1, r_2 > 0$ , and any  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ ,  $\alpha_1 \in (0, 1)$ ,  $\alpha_2, \alpha_3, \alpha_4 > 0$ , the stabilized mixed formulation (63) is well-posed.

*Proof.* From [3, Proposition 4.2.1], it is a straightforward consequence of the boundedness of  $a_{r, \alpha}$ ,  $b_\alpha$ ,  $c_\alpha$ ,  $d_\alpha$  and  $\ell$ , as well as the coercivity of  $a_{r, \alpha}$  and  $c_\alpha$  on  $\mathbf{Z}$  and  $\Xi$  respectively.  $\square$

**Proposition 2.** The solution of the mixed formulation (47) and the solution of the stabilized mixed formulation (63) coincide.

*Proof.* For any  $r = (r_1, r_2)$ ,  $r_1, r_2 > 0$ , and any  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ ,  $\alpha_1 \in (0, 1)$ ,  $\alpha_2, \alpha_3, \alpha_4 > 0$ , we check that the saddle point  $((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r)) \in \mathbf{Z} \times \Lambda$  of  $\mathcal{L}_r$  is also a saddle point of  $\mathcal{L}_{r, \alpha}$ . We observe that, by Theorem 3.1,  $\zeta_r$  is a controlled solution of (1), in the sens of the transposition, with initial datum in  $\mathbf{W}$  and boundary control in  $L^2(\Sigma_0)$ . Thus,  $((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r)) \in \mathbf{Z} \times \Xi$ . Besides, since  $((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r))$  solves the augmented formulation associated with (47), for any  $(\zeta, \Theta) \in \Xi$ , we have

$$\begin{aligned} \mathcal{L}_{r, \alpha}((\mathbf{w}_r, \mathbf{Q}_r), (\zeta, \Theta)) &\leq \mathcal{L}_r((\mathbf{w}_r, \mathbf{Q}_r), (\zeta, \Theta)) \\ &\leq \mathcal{L}_r((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r)) = \mathcal{L}_{r, \alpha}((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r)). \end{aligned}$$

Therefore,  $(\zeta_r, \Theta_r)$  maximizes  $(\zeta, \Theta) \mapsto \mathcal{L}_{r, \alpha}((\mathbf{w}_r, \mathbf{Q}_r), (\zeta, \Theta))$ . Conversely, the functional  $F : \mathbf{Z} \rightarrow \mathbb{R}$  given by  $F(\mathbf{w}, \mathbf{Q}) = \mathcal{L}_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\zeta_r, \Theta_r))$  admits a unique extremal point due to the ellipticity of  $a_{r, \alpha}$ . Moreover, for any  $(\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}$ , we have

$$\begin{aligned} \frac{d}{d\varepsilon} F((\mathbf{w}, \mathbf{Q}) + \varepsilon(\bar{\mathbf{w}}, \bar{\mathbf{Q}})) \Big|_{\varepsilon=0} &= a_r((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) + b((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta_r, \Theta_r)) - \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \\ &\quad - \alpha_1 \langle \sigma(\mathbf{Q})\nu - \zeta_r, \sigma(\bar{\mathbf{Q}})\nu \rangle_{L^2(\Sigma_0)}. \end{aligned}$$

Therefore, from (47), we get

$$\frac{d}{d\varepsilon} F((\mathbf{w}_r, \mathbf{Q}_r) + \varepsilon(\bar{\mathbf{w}}, \bar{\mathbf{Q}})) \Big|_{\varepsilon=0} = 0, \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}.$$

Thus,  $(\mathbf{w}_r, \mathbf{Q}_r)$  minimizes  $(\mathbf{w}, \mathbf{Q}) \mapsto \mathcal{L}_{r, \alpha}((\mathbf{w}, \mathbf{Q}), (\zeta_r, \Theta_r))$ . Consequently, the pair  $((\mathbf{w}_r, \mathbf{Q}_r), (\zeta_r, \Theta_r))$  is also a saddle point of  $\mathcal{L}_{r, \alpha}$ . The result follows from the uniqueness of the saddle point.  $\square$

**5. Dual extremal problem.** Here, we derive an extremal problem which is dual to (30) and only involves the primal variable  $(\zeta, \Theta)$ . For  $r = (r_1, r_2)$  and  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ , we define the linear operator  $\mathcal{A}_{r, \alpha} : \Xi \rightarrow \Xi$  in the following way. For any  $(\zeta, \Theta) \in \Xi$ ,  $\mathcal{A}_{r, \alpha}(\zeta, \Theta) \in \Xi$  is the unique solution of

$$\langle \mathcal{A}_{r, \alpha}(\zeta, \Theta), (\bar{\zeta}, \bar{\Theta}) \rangle_{\Xi} = b_\alpha((\mathbf{w}, \mathbf{Q}), (\bar{\zeta}, \bar{\Theta})) + c_\alpha((\zeta, \Theta), (\bar{\zeta}, \bar{\Theta})), \quad \forall (\bar{\zeta}, \bar{\Theta}) \in \Xi, \quad (68)$$



where  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  is the unique solution of

$$a_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = b_\alpha((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta, \Theta)), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}. \quad (69)$$

The conditions  $r_1, r_2 > 0$  and  $\alpha_1 < 1$  imply that the augmented bilinear form  $a_{r,\alpha}$  is coercive on  $\mathbf{Z}$ , hence (69) has a unique solution.

**Lemma 5.1.** *The operator  $\mathcal{A}_{r,\alpha}$  is a strongly elliptic, symmetric isomorphism from  $\Xi$  onto  $\Xi$ .*

*Proof.* For simplicity, we take  $\eta = \frac{r_1}{1-\alpha_1}$  in (46). Since  $a_{r,\alpha}$  satisfies the coercivity estimate

$$a_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\mathbf{w}, \mathbf{Q})) \geq (1 - \alpha_1) \|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{Z}}^2, \quad \forall (\mathbf{w}, \mathbf{Q}) \in \mathbf{Z},$$

then the Lax-Milgram lemma implies that for  $(\zeta, \Theta) \in \Xi$ , the corresponding solution  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  of (69) satisfies

$$\|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{Z}} \leq \frac{\|b_\alpha\|}{1 - \alpha_1} \|(\zeta, \Theta)\|_{\Xi}.$$

Therefore, taking  $(\bar{\zeta}, \bar{\Theta}) = \mathcal{A}_{r,\alpha}(\zeta, \Theta)$  in (68) leads to

$$\|\mathcal{A}_{r,\alpha}(\zeta, \Theta)\|_{\Xi} \leq \left( \frac{\|b_\alpha\|^2}{1 - \alpha_1} + \|c_\alpha\| \right) \|(\zeta, \Theta)\|_{\Xi},$$

which gives the boundedness of  $\mathcal{A}_{r,\alpha}$ . Let  $(\zeta, \Theta), (\bar{\zeta}, \bar{\Theta}) \in \Xi$  and denote respectively  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}$  and  $(\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}$  the corresponding solutions of (69). Then, from (68), we have

$$\langle \mathcal{A}_{r,\alpha}(\zeta, \Theta), (\bar{\zeta}, \bar{\Theta}) \rangle_{\Xi} = a_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) + c_\alpha((\zeta, \Theta), (\bar{\zeta}, \bar{\Theta})), \quad (70)$$

which implies the positivity and symmetry of  $\mathcal{A}_{r,\alpha}$ . Besides, from the coercivity of  $c_\alpha$ , there exists  $C > 0$  such that, for any  $(\zeta, \Theta) \in \Xi$ ,

$$\langle \mathcal{A}_{r,\alpha}(\zeta, \Theta), (\zeta, \Theta) \rangle_{\Xi} \geq c_\alpha((\zeta, \Theta), (\zeta, \Theta)) \geq C \|(\zeta, \Theta)\|_{\Xi}^2. \quad \square$$

The ellipticity of  $\mathcal{A}_{r,\alpha}$  allows to introduce a coercive functional  $\hat{\mathcal{J}}^{**}$  in the next lemma.

**Lemma 5.2.** *For any  $r = (r_1, r_2)$ ,  $r_1, r_2 > 0$ , let  $(\mathbf{w}_0, \mathbf{Q}_0) \in \mathbf{Z}$  be the unique solution of*

$$a_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}. \quad (71)$$

Let also  $\hat{\mathcal{J}}^{**} : \Xi \rightarrow \mathbb{R}$  be the functional defined by

$$\hat{\mathcal{J}}^{**}(\zeta, \Theta) = \frac{1}{2} \langle \mathcal{A}_{r,\alpha}(\zeta, \Theta), (\zeta, \Theta) \rangle_{\Xi} - b_\alpha((\mathbf{w}_0, \mathbf{Q}_0), (\zeta, \Theta)) + d_\alpha(\zeta, \Theta).$$

Then, the following equality holds,

$$\sup_{(\zeta, \Theta) \in \Xi} \inf_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \mathcal{L}_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) = - \inf_{(\zeta, \Theta) \in \Xi} \hat{\mathcal{J}}^{**}(\zeta, \Theta) + \mathcal{L}_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\mathbf{0}, \mathbf{0})).$$

*Proof.* For any  $(\zeta, \Theta) \in \Xi$ , let  $(\mathbf{w}^*, \mathbf{Q}^*) \in \mathbf{Z}$  be the minimum point of  $(\mathbf{w}, \mathbf{Q}) \mapsto \mathcal{L}_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta))$ . Thus,  $(\mathbf{w}^*, \mathbf{Q}^*)$  satisfies the equation

$$a_{r,\alpha}((\mathbf{w}^*, \mathbf{Q}^*), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) + b_\alpha((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta, \Theta)) = \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z},$$

and can be decomposed as  $(\mathbf{w}^*, \mathbf{Q}^*) = (\mathbf{w}_0, \mathbf{Q}_0) - (\tilde{\mathbf{w}}, \tilde{\mathbf{Q}})$ , where  $(\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}) \in \mathbf{Z}$  solves

$$a_{r,\alpha}((\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = b_\alpha((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta, \Theta)), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}.$$

We then have

$$\begin{aligned} \inf_{(\mathbf{w}, \mathbf{Q}) \in \mathbf{Z}} \mathcal{L}_{r,\alpha}((\mathbf{w}, \mathbf{Q}), (\zeta, \Theta)) &= \mathcal{L}_{r,\alpha}((\mathbf{w}^*, \mathbf{Q}^*), (\zeta, \Theta)) \\ &= \mathcal{L}_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0) - (\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}), (\zeta, \Theta)) \\ &= X_1 + X_2 + X_3 + X_4, \end{aligned}$$

with

$$\begin{cases} X_1 = \frac{1}{2} a_{r,\alpha}((\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}), (\tilde{\mathbf{w}}, \tilde{\mathbf{Q}})) - b_\alpha((\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}), (\zeta, \Theta)) - \frac{1}{2} c_\alpha((\zeta, \Theta), (\zeta, \Theta)), \\ X_2 = b_\alpha((\mathbf{w}_0, \mathbf{Q}_0), (\zeta, \Theta)) - d_\alpha(\zeta, \Theta), \\ X_3 = \frac{1}{2} a_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\mathbf{w}_0, \mathbf{Q}_0)) - \ell(\mathbf{w}_0, \mathbf{Q}_0), \\ X_4 = -a_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\tilde{\mathbf{w}}, \tilde{\mathbf{Q}})) + \ell(\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}). \end{cases}$$

From the definition of  $(\mathbf{w}_0, \mathbf{Q}_0)$ , we have  $X_4 = 0$  and  $X_3 = \mathcal{L}_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\mathbf{0}, \mathbf{0}))$ . Finally, the definition of  $(\tilde{\mathbf{w}}, \tilde{\mathbf{Q}})$  implies

$$X_1 = -\frac{1}{2} b_\alpha((\tilde{\mathbf{w}}, \tilde{\mathbf{Q}}), (\zeta, \Theta)) - \frac{1}{2} c_\alpha((\zeta, \Theta), (\zeta, \Theta)) = -\frac{1}{2} \langle \mathcal{A}_{r,\alpha}(\zeta, \Theta), (\zeta, \Theta) \rangle_{\Xi}$$

and the result follows.  $\square$

From the ellipticity of the operator  $\mathcal{A}_{r,\alpha}$ , the minimization of  $\hat{\mathcal{J}}^{**}$  in  $\mathbf{Z}$  is well-posed. We observe that, in contrast with the initial problem of finding the control of minimal  $L^2$ -norm, the minimization of  $\hat{\mathcal{J}}^{**}$  in  $\mathbf{Z}$  does not entail any constraint.

Due to the symmetry and ellipticity of the operator  $\mathcal{A}_{r,\alpha}$ , the *conjugate gradient* method is well-suited for the numerical minimization of  $\hat{\mathcal{J}}^{**}$ . More exactly, the Polak-Ribière version of the conjugate gradient method reads as follows: for a given convergence criterion  $\epsilon > 0$ , we follow the steps in Algorithm 1.

**6. Numerical approximation and simulations.** In this section, we numerically illustrate the proposed method for the approximation of the control and the controlled solution for the two-dimensional wave equation and for the two-dimensional elasticity system. More precisely, we put in place the procedure described in Section 5 by considering a conformal finite elements approximation of the spaces  $\mathbf{Z}$ ,  $\Xi$  and of the operators  $\mathcal{A}_{r,\alpha}$ ,  $\mathcal{M}$ , together with the inner products of these spaces. In that purpose, we consider uniform triangulations  $\mathcal{T}_h$  (structured or not) of the domain  $Q$ . For such each triangulation, we denote by  $h > 0$  the largest diameter of tetrahedra composing  $\mathcal{T}_h$ . Two examples of domains  $Q$  and the associated meshes are displayed in Figure 1.

We define  $\Xi_h \subset \Xi$  as follows:

$$\begin{aligned} \Xi_h = \left\{ (\zeta_h, \Theta_h) \in \Xi; \quad \zeta = \begin{pmatrix} \zeta_{h,1} \\ \zeta_{h,2} \end{pmatrix}, \quad \Theta = \begin{pmatrix} \Theta_{h,11} & \Theta_{h,12} \\ \Theta_{h,21} & \Theta_{h,22} \end{pmatrix}, \right. \\ \left. \zeta_{h,i|K}, \Theta_{h,ij|K} \in \mathbb{P}_1 \text{ for every } K \in \mathcal{T}_h \text{ and } i, j \in \{1, 2\} \right\}, \quad (72) \end{aligned}$$

with  $\mathbb{P}_1$  being the space of polynomials of variables  $(x, y)$  and of degree less than 1. Algorithm 1 easily adapts to this finite dimensional framework considering finite elements approximations of operators  $\mathcal{A}_{r,\alpha}$ ,  $\mathcal{M}$  and of the inner product over  $\Xi$ . The implementation of this algorithm was done using FreeFEM [13]. If it is not specified otherwise, we choose the augmentation parameters  $r_1 = r_2 = 10^{-1}$  and the stabilization parameters  $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 10^{-3}$ . Remark that the control

**Algorithm 1** Conjugate algorithm for iteratively solving (70)

- 1) *Initialization.* We choose an initial candidate  $(\zeta_0, \Theta_0) \in \Xi$ .

We compute  $(\mathbf{w}_0, \mathbf{Q}_0) \in \mathbf{Z}$  solution of

$$a_{r,\alpha}((\mathbf{w}_0, \mathbf{Q}_0), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = b_\alpha((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\zeta_0, \Theta_0)) - \ell(\bar{\mathbf{w}}, \bar{\mathbf{Q}}), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}.$$

We compute  $(\mathbf{g}_0, \mathbf{G}_0) \in \Xi$  solution of

$$\begin{aligned} \langle (\mathbf{g}_0, \mathbf{G}_0), (\bar{\zeta}, \bar{\Theta}) \rangle_\Xi &= b_\alpha((\mathbf{w}_0, \mathbf{Q}_0), (\bar{\zeta}, \bar{\Theta})) \\ &\quad + c_\alpha((\zeta_0, \Theta_0), (\bar{\zeta}, \bar{\Theta})) + d_\alpha(\bar{\zeta}, \bar{\Theta}), \quad \forall (\bar{\zeta}, \bar{\Theta}) \in \Xi. \end{aligned}$$

We set  $(\mathbf{v}_0, \mathbf{P}_0) = (\mathbf{g}_0, \mathbf{G}_0)$ .

- 2) *Iteration.* For  $n \geq 0$ , we compute  $(\tilde{\mathbf{w}}_n, \tilde{\mathbf{Q}}_n) \in \mathbf{Z}$  solution of

$$a_{r,\alpha}((\tilde{\mathbf{w}}_n, \tilde{\mathbf{Q}}_n), (\bar{\mathbf{w}}, \bar{\mathbf{Q}})) = b_\alpha((\bar{\mathbf{w}}, \bar{\mathbf{Q}}), (\mathbf{v}_n, \mathbf{P}_n)), \quad \forall (\bar{\mathbf{w}}, \bar{\mathbf{Q}}) \in \mathbf{Z}.$$

We compute  $\mathcal{A}_{r,\alpha}(\mathbf{v}_n, \mathbf{P}_n) \in \Xi$  solution of

$$\begin{aligned} \langle \mathcal{A}_{r,\alpha}(\mathbf{v}_n, \mathbf{P}_n), (\bar{\zeta}, \bar{\Theta}) \rangle_\Xi &= b_\alpha((\tilde{\mathbf{w}}_n, \tilde{\mathbf{Q}}_n), (\bar{\zeta}, \bar{\Theta})) \\ &\quad + c_\alpha((\mathbf{v}_n, \mathbf{P}_n), (\bar{\zeta}, \bar{\Theta})), \quad \forall (\bar{\zeta}, \bar{\Theta}) \in \Xi. \end{aligned}$$

We compute

$$\alpha_n = \frac{\|(\mathbf{g}_n, \mathbf{G}_n)\|_\Xi^2}{\langle \mathcal{A}_{r,\alpha}(\mathbf{v}_n, \mathbf{P}_n), (\mathbf{v}_n, \mathbf{P}_n) \rangle_\Xi}$$

and we update

$$\begin{aligned} (\zeta_{n+1}, \Theta_{n+1}) &= (\zeta_n, \Theta_n) - \alpha_n(\mathbf{v}_n, \mathbf{P}_n), \\ (\mathbf{w}_{n+1}, \mathbf{Q}_{n+1}) &= (\mathbf{w}_n, \mathbf{Q}_n) - \alpha_n(\tilde{\mathbf{w}}_n, \tilde{\mathbf{Q}}_n), \\ (\mathbf{g}_{n+1}, \mathbf{G}_{n+1}) &= (\mathbf{g}_n, \mathbf{G}_n) - \alpha_n \mathcal{A}_{r,\alpha}(\mathbf{v}_n, \mathbf{P}_n). \end{aligned}$$

- 3) *Convergence test.* If  $\|(\mathbf{g}_{n+1}, \mathbf{G}_{n+1})\|_\Xi \leq \epsilon \|(\mathbf{g}_0, \mathbf{G}_0)\|_\Xi$ , we stop and set  $(\zeta_n, \Theta_n)$  as the approximation of the minimum point of  $\hat{\mathcal{J}}^{**}$ . Else, we compute

$$\beta_n = \frac{\langle (\mathbf{g}_{n+1}, \mathbf{G}_{n+1}), (\mathbf{g}_{n+1}, \mathbf{G}_{n+1}) - (\mathbf{g}_n, \mathbf{G}_n) \rangle_\Xi}{\|(\mathbf{g}_n, \mathbf{G}_n)\|_\Xi^2}$$

and we update

$$(\mathbf{v}_{n+1}, \mathbf{P}_{n+1}) = (\mathbf{g}_{n+1}, \mathbf{G}_{n+1}) + \beta_n(\mathbf{v}_n, \mathbf{P}_n).$$

Do  $n \leftarrow n + 1$  and return to step 2).

$\mathbf{h}$  can be obtained in two different manners. On the one hand, since  $\zeta$  correspond to the controlled state of the system (1), we have  $\mathbf{h} = \zeta|_{\Sigma_0}$ . On the other hand, we can compute  $\mathbf{h}$  using the dual variables  $(\mathbf{w}, \mathbf{Q})$  and the relation (52).

As a first step for the validation of the proposed method, we employ Algorithm 1 for the approximation of the boundary control and the corresponding controlled state for a two-dimensional wave equation. Indeed, choosing  $\mu = 1$  and  $\lambda = -\mu$ , system (1) is nothing else than two independent wave equations with constant coefficients in dimension two. The boundary controllability of two-dimensional wave equations was firstly considered from a numerical perspective in [11, 12]. An example of initial data and the corresponding boundary control of minimal  $L^2$ -norm are explicitly given in this paper. The numerical experiments corresponding to the

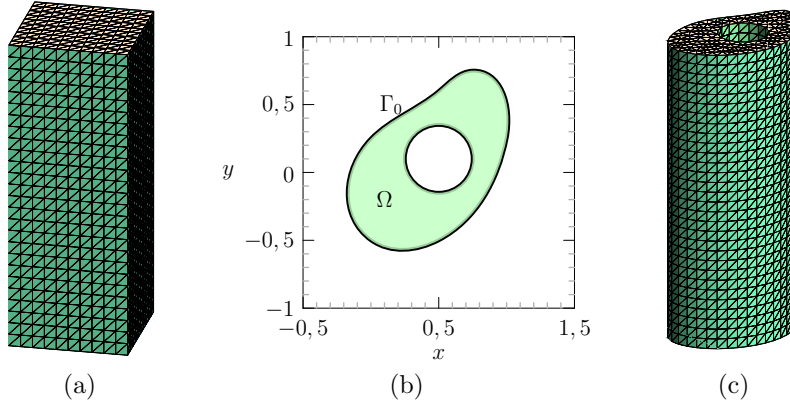


FIGURE 1. Domains  $Q$  and associated meshes. (a) A structured mesh of  $Q = (0, 1)^2 \times (0, T)$ . (b) An example of non-convex domain  $\Omega$  and (c) a mesh of  $Q = \Omega \times (0, T)$  associated to this domain.

domain  $\Omega = (0, 1)^2$  and to the initial data in [12, 10] for which the boundary control is explicitly known are presented in Subsection 6.1.

Once the method validated for the boundary controllability of the 2-D wave equation, numerical experiments were conducted for the elasticity system in dimension 2. More exactly, we set  $\mu = \lambda = 1$ . We numerically illustrate this situation in Subsection 6.2 for two different geometries of  $\Omega$ : the unit square and the non-convex domain represented in Figure 1 (b).

**6.1. Two dimensional wave equation.** In this subsection, we set  $\Omega = (0, 1)^2$  the unit square with the control acting on its whole boundary ( $\Gamma_0 = \Gamma$ ). We consider the initial datum constructed in [11, 12] using Fourier series. For this initial datum, displayed in Figure 2, and for  $T = \frac{1}{\sqrt{2}}(3 + \frac{3}{4})$ , the boundary control can be explicitly computed (see [12] for details).

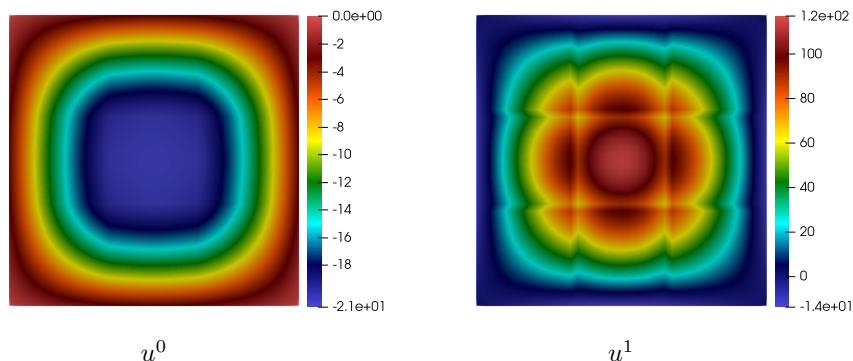


FIGURE 2. Initial datum  $(u^0, u^1)$  constructed in [12].

More exactly, for this choice of  $T$ , we have

$$\begin{aligned}
u^0(x, y) &= -\pi (\sin(\pi x) \cos(2\pi y) + \cos(2\pi x) \sin(\pi y)) \\
&\quad + \left( -2\sqrt{2}\pi T + \frac{28}{3\sqrt{2}} \sin(\sqrt{2}\pi T) \right) \sin(\pi x) \sin(\pi y) \\
&\quad + 4 \sin(\pi x) \sum_{p=3, p \text{ odd}}^{255} \frac{p}{p^2-1} \left( \frac{3}{p^2-4} - \frac{2}{\sqrt{p^2+1}} \sin(\sqrt{p^2+1}\pi T) \right) \sin(p\pi y) \\
&\quad + 4 \sin(\pi y) \sum_{p=3, p \text{ odd}}^{255} \frac{p}{p^2-1} \left( \frac{3}{p^2-4} - \frac{2}{\sqrt{p^2+1}} \sin(\sqrt{p^2+1}\pi T) \right) \sin(p\pi x), \\
u^1(x, y) &= -\sqrt{2}\pi^2 (\sin(\pi x) \cos(2\pi y) + \cos(2\pi x) \sin(\pi y)) \\
&\quad + \left( 2\sqrt{2}\pi + 4\pi^2 T - \frac{28\pi}{3} \cos(\sqrt{2}\pi T) \right) \sin(\pi x) \sin(\pi y) \\
&\quad + 4 \sin(\pi x) \sum_{p=3, p \text{ odd}}^{255} \frac{p}{p^2-1} \left( \frac{3\sqrt{2}\pi}{p^2-4} + 2\pi \cos(\sqrt{p^2+1}\pi T) \right) \sin(p\pi y) \\
&\quad + 4 \sin(\pi y) \sum_{p=3, p \text{ odd}}^{255} \frac{p}{p^2-1} \left( \frac{3\sqrt{2}\pi}{p^2-4} + 2\pi \cos(\sqrt{p^2+1}\pi T) \right) \sin(p\pi x).
\end{aligned}$$

As pointed out in Remark 2, for  $\lambda = -\mu$ , system (1) consists of two independent wave equations:

$$\begin{cases} u_1'' - \mu \Delta u_1 = 0 & \text{in } Q, \\ u_1 = h_1 & \text{on } \Sigma, \\ (u_1, u_1')(\cdot, 0) = (u^0, u^1) & \text{in } \Omega, \end{cases} \quad \begin{cases} u_2'' - \mu \Delta u_2 = 0 & \text{in } Q, \\ u_2 = h_2 & \text{on } \Sigma, \\ (u_2, u_2')(\cdot, 0) = (u^0, u^1) & \text{in } \Omega. \end{cases}$$

In the rest of this subsection, we only consider the equation satisfied by  $u_1$ .

We consider five levels of meshes described in Table 1. The coarsest of these meshes is depicted in Figure 1 (a).

Mesh number	1	2	3	4	5
Diameter $h$ of elements	$\frac{1}{10}$	$\frac{1}{20}$	$\frac{1}{30}$	$\frac{1}{40}$	$\frac{1}{50}$
Number of nodes	3 267	23 814	76 880	179 867	345 933
Number of tetrahedra	15 600	127 200	426 600	1 017 600	1 980 000

TABLE 1. Description of five meshes of the domain  $Q = (0, 1)^2 \times (0, T)$ .

In Figure 3 (a), we display the evolution of the norm of the residuals  $\|(g_n, \mathbf{G}_n)\|_{\Xi}$  appearing in Algorithm 1. Here and henceforth, we fixed the value of the stopping criterion parameter in the conjugate gradient algorithm Algorithm 1 to  $\epsilon = 10^{-10}$ . In Figure 3 (b), we display the norm of the boundary control obtained from  $(\zeta_n, \Theta_n)$  once the conjugate gradient algorithm has converged.

In Figure 4 we display the evolution of the norm of the  $L^2$  error between the exact control (explicitly known) and the numerical controls computed from dual and primal variables for two different choices of stabilization parameter  $\alpha_i$ . Analyzing these results, we observe a convergence of order  $h^{0.5}$  for the numerical controls to the explicit control, for the controls obtained from both the primal and dual variables when  $\alpha_i = 10^{-1}$ . For an augmentation parameter  $\alpha_i = 10^{-3}$  we observe a better convergence (of order  $h$ ) for the control computed from the primal variable. The same order of convergence is observed for large values of  $h$  when the control is

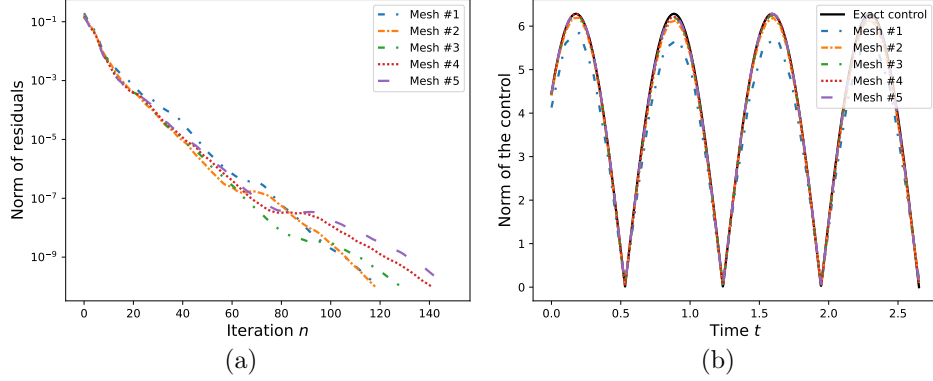


FIGURE 3. Results for initial datum  $(u^0, u^1)$  displayed in Figure 2. (a) Evolution of the norm residuals  $(g_n, G_n)$ . (b) Norm  $L^2$  of the control  $h(t)$ .

computed from the dual variable, but the approximation is affected by a numerical locking phenomenon at 0.02 which is probably due to the values of the stabilization parameters which are not large enough. A more involved study will be needed for finding the optimal values of these parameters.

We also remark that the number of iterations needed for the convergence of the conjugate gradient algorithm remains stable with respect to the discretization parameter  $h$ . Nevertheless, as expected, for smaller values of augmentation and stabilization parameters a higher number of iterations is need for converge of the conjugate gradient algorithm described in Algorithm 1. In Figure 5, we display the solution  $(\zeta_n, \Theta_n)$  obtained once the Algorithm 1 has converged.

Similar results are obtained for different choices of parameter  $r = (r_1, r_2)$  and  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ .

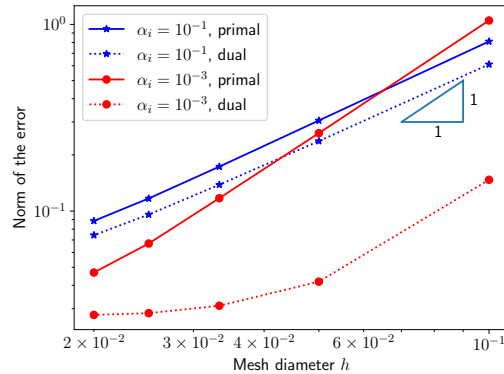


FIGURE 4. Norm of the error between the exact and numerical controls for different values of  $h$  and two different values of  $\alpha_1$ .

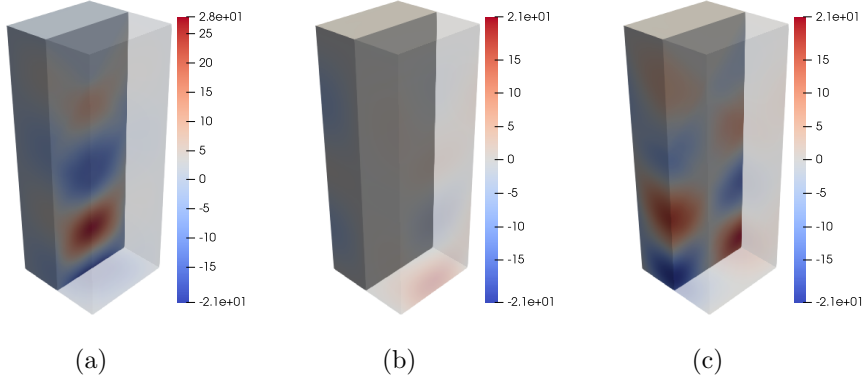


FIGURE 5. Solution  $(\zeta_n, \Theta_n)$  obtained once the conjugate algorithm converged for the mesh 5. (a)  $\zeta_n$ . (b)  $\Theta_{n,1}$ . (c)  $\Theta_{n,2}$ .

Another interesting illustration (see Figure 6 (a)) is the evolution with respect to the time of the norm of the primal and dual solutions  $(\zeta, \Theta)$  and  $(w, Q)$ , respectively. As expected, the energy norm of the primal variable goes to zero when the time is close to the control time  $T$  and the energy norm of the dual variable is conserved in time.

**6.2. Elasticity system.** In this subsection, we illustrate the approximation of the boundary control for the two-dimensional elasticity system ( $\mu = \lambda = 1$ ) for two different domain geometries. In the first case, we consider  $\Omega$  the unit square and in the second one  $\Omega$  is a non convex domain with regular boundary.

**6.2.1. Control acting on the whole boundary  $\Gamma$  of  $\Omega = (0, 1)^2$ .** For simplicity, we consider here the same geometric situation as in Section 6.1:  $\Omega = (0, 1)^2$ ,  $T = \frac{1}{\sqrt{2}}(3 + \frac{3}{4})$  and the initial datum  $(u^0, u^1) \in L^2(\Omega) \times H^{-1}(\Omega)$  given by

$$u^0 = (u^0, u^0)^\top, \quad u^1 = (u^1, u^1)^\top, \quad (73)$$

with  $u^0$  and  $u^1$  depicted in Figure 2. The conjugate gradient algorithm converges for every mesh in less than 200 iterations for a norm of the residuals smaller than  $\epsilon = 10^{-10}$ . The norm  $L^2(\Omega)$  of the controls  $h(t)$  is depicted in Figure 7 for all the five levels of mesh considered here. In the left column of this figure we display the evolution of the norm of the controls computed as the boundary trace of  $\zeta_n$  for  $\alpha_1 = 10^{-3}$  (top) and  $\alpha_1 = 0.9$  (bottom). The norm of these controls are very similar for both values of the parameter  $\alpha_1$ . On the right side of the figure we depict the norm of the control computed from  $Q_n$  using (52) for the same values of  $\alpha_1$  as before. In that case we observe that for  $\alpha_1 = 10^{-3}$  the control has some extra oscillations near  $t = 0$  and  $t = T$ . As expected, for values of  $\alpha_1$  close to 1, the controls obtained from the primal variable  $\zeta$  are very close to the controls obtained from the dual variable  $Q$ . Similarly to the wave equation, the controls seem to converge with respect to the discretization parameter  $h$ . In Figure 6 (b) we also observe a very similar comportment to the wave equation for the evolution in time of the norms of primal and dual variables.

The six components of the solution  $(\zeta_n, \Theta_n)$  obtained once the Algorithm 1 has converged for the finest mesh in Table 1 are graphically represented in Figure 8.

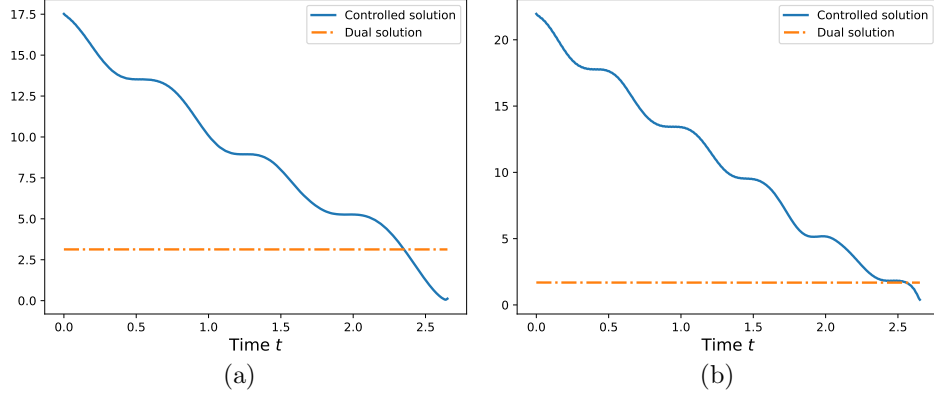


FIGURE 6. Evolution with respect to the time  $t$  of the norms of primal and dual solutions for: (a) the wave equation; (b) the elasticity system and the initial data in Figure 2.

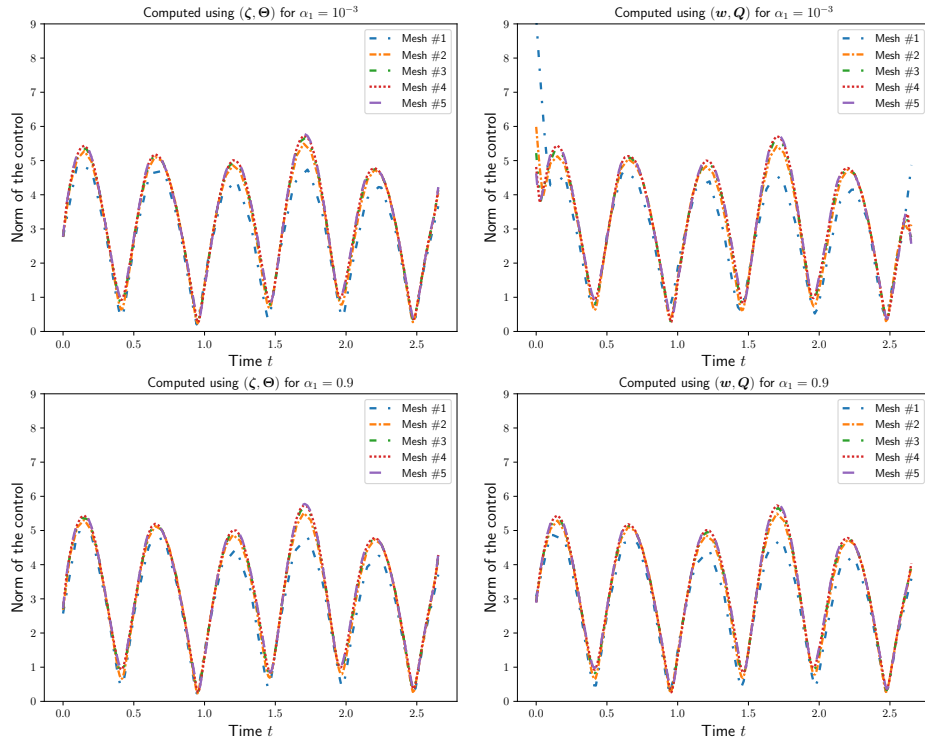


FIGURE 7. Norm of the control  $h$  for the five meshes described in Table 1 computed from  $(\zeta_n, \Theta_n)$  (left) and from  $(w_n, Q_n)$  (right), for  $\alpha_1 = 10^{-3}$  (up) and  $\alpha_1 = 9 \times 10^{-1}$  (bottom), respectively.

6.2.2. *Control acting on a part of the regular boundary of a non-convex  $\Omega$ .* In this subsection we consider the non-convex domain  $\Omega$  graphically represented in Figure 1 (b). The control acts only on the exterior boundary  $\Gamma_0$  of  $\Omega$ .



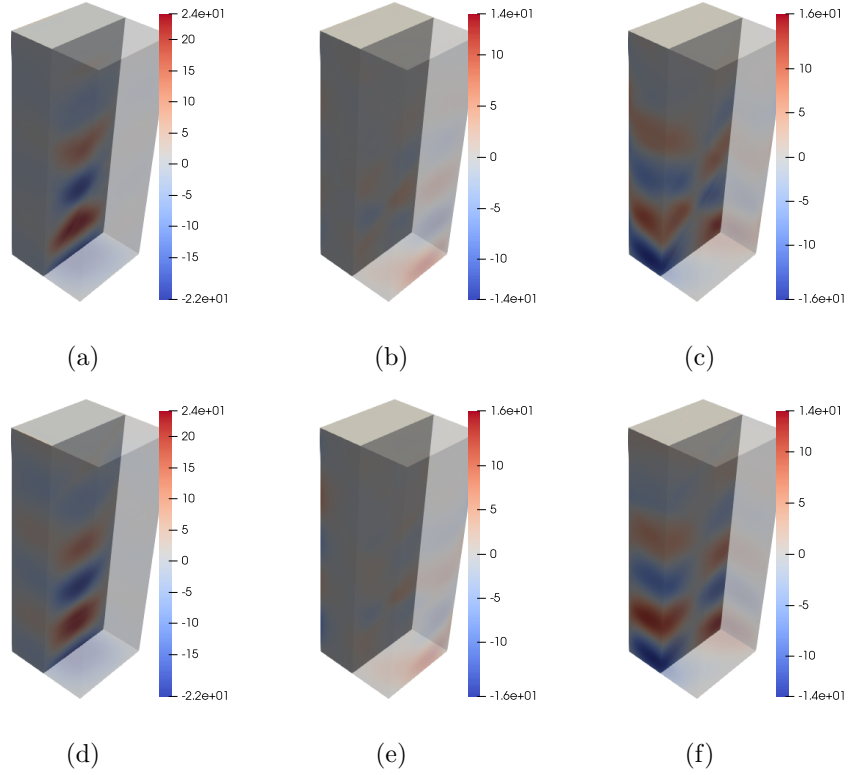


FIGURE 8. The six components of the solution for the initial datum in Figure 2 and the finest mesh in Table 1. (a)  $\zeta_{n,1}$ . (b)  $\Theta_{n,11}$ . (c)  $\Theta_{n,12}$ . (d)  $\zeta_{n,2}$ . (e)  $\Theta_{n,21}$ . (f)  $\Theta_{n,22}$ .

For this choice of  $\Omega$ , we consider a control time  $T = 3$  and the initial data  $(\mathbf{u}^0, \mathbf{u}^1) \in \mathbf{L}^2(\Omega) \times \mathbf{H}^{-1}(\Omega)$  given by

$$\mathbf{u}^0 = \mathbf{u}^1 \text{ solution to } \begin{cases} -\Delta \mathbf{u}^0 = 100 & \text{in } \Omega, \\ \mathbf{u}^0 = 0 & \text{on } \Gamma. \end{cases} \quad (74)$$

We consider five levels of meshes described in Table 2. The coarsest of these meshes is plotted in Figure 1 (c).

Mesh number	1	2	3	4	5
Diameter $h$ of elements	$\frac{1}{10}$	$\frac{1}{20}$	$\frac{1}{30}$	$\frac{1}{40}$	$\frac{1}{50}$
Number of nodes	4 557	29 707	99 094	212 234	406 945
Number of tetrahedra	21 510	155 700	515 080	1 185 760	2 303 550

TABLE 2. Description of five meshes of the domain  $Q = \Omega \times (0, T)$  for  $\Omega$  displayed in Figure 1 (b).

For the set of simulations corresponding to this non-convex domain and initial data given by (74) we fixed  $\alpha_1 = 0.9$  and we consider  $\alpha_2 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$ .

The parameter  $\alpha_2$  enforces that the primal variable verifies the controlled elasticity system in  $L^2(0, T; \mathbf{W})$ .

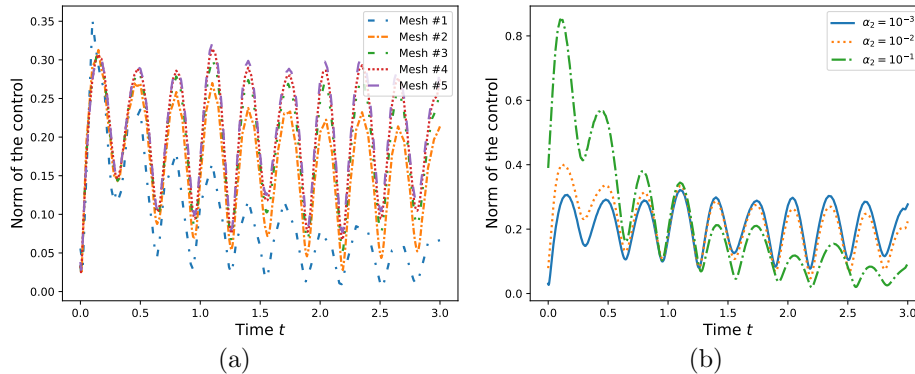


FIGURE 9. Norm of the control for initial data given by (74). (a)  $\alpha_2 = 10^{-3}$  and different meshes. (b) Computation on the mesh #5 and different values for  $\alpha_2$ .

The norm of the control for  $\alpha_2 = 10^{-3}$  and the five meshes described in Table 2 is displayed in Figure 9 (a). As expected we observe the convergence of the controls with respect to the discretization parameter  $h$ . In Figure 9 (b) we illustrate the norm of the control computed on the finest mesh in Table 2 for three different values of  $\alpha_2$ . As expected, for larger values of  $\alpha_2$  the primal solution  $\zeta$  is more regular and the corresponding control is closer to zero near the control time  $T$ .

**7. Conclusion and perspectives.** This paper studied the boundary null controllability of the elasticity system reformulated as a first order in time and space system. We also proposed a constructive method to compute the boundary control of minimal  $L^2$ -norm for the elasticity system as the solution of a stabilized mixed formulation. Although the numerical computation of the control remains extremely challenging for the elasticity system, the proposed method was numerically implemented and the obtained results are very promising. The facts that we considered a first order system in space and time combined to the very careful augmentation and stabilization of the mixed formulation were essential for the numerical implementation. More exactly, we are able to successfully solve the mixed formulation by approaching it using only  $P_1$  finite elements and by the means of a conjugate gradient algorithm. The proposed numerical method seems robust with respect to the values of the Lamé's coefficients. Nevertheless, several difficulties remain and in what follows we propose some perspectives to tackle them.

As for all space-time methods, the main numerical difficulty was related to the high dimension of discrete system to solve. Indeed, the time appears as a third dimension for a system being two-dimensional in space. Even using a parallel solver for the linear systems to solve in the gradient descent algorithm, this was not enough for very refined meshes. In a further work, we will investigate the possibility to iteratively solve the mixed formulation characterizing the control and the controlled solution by a domain decomposition technique combined to a conjugate gradient method. This will allow to fully parallelize the method and not only the resolution of the involved linear systems as it is actually the case.

Another delicate question is the choice of the values for the augmentation parameters  $(r_i)_{1 \leq i \leq 2}$  and for the stabilization parameters  $(\alpha_i)_{1 \leq i \leq 4}$ . For larger values of these parameters we need less iterations for the conjugate gradient algorithm convergence but the errors are bigger. For smaller values of parameters we need a larger number of iterations for the conjugate gradient algorithm and the convergence of control seems better. A more involved numerically study is needed in order to numerically investigate the dependence of the *inf-sup* constant associated to the mixed formulation to these parameters. We will consider this question in a further work.

**Appendix A. Operator  $\mathcal{A}$  is the generator of a group of contractions.** The following lemma states that the operator  $\mathcal{A}$  given by (24) generates a group of contractions.

**Lemma A.1.** *The operator  $\mathcal{A} : \mathbf{V} \rightarrow \mathbf{W}$  defined by (24) is the infinitesimal generator of a  $C_0$ -group of contractions  $(\mathbb{T}_t)_{t \in \mathbb{R}}$  on  $\mathbf{W}$ .*

*Proof.* It is sufficient to apply the Hille-Yosida theorem to the operators  $\mathcal{A}$  and  $-\mathcal{A}$ . We verify the hypotheses of this theorem for  $\mathcal{A}$ , the situation being very similar for  $-\mathcal{A}$ . More precisely, we should check that  $(0, +\infty) \subset \rho(\mathcal{A})$  and that, for every  $\xi > 0$ , the corresponding resolvent operator verifies  $\|(\xi \mathcal{I}_d - \mathcal{A})^{-1}\| \leq \xi^{-1}$ .

Let  $\xi > 0$  and  $(\mathbf{f}, \mathbf{F}) \in \mathbf{W}$ . In order to show that  $\xi \in \rho(\mathcal{A})$ , we should find  $(\mathbf{w}, \mathbf{Q}) \in \mathbf{V}$  such that  $(\xi \mathcal{I}_d - \mathcal{A})(\mathbf{w}, \mathbf{Q}) = (\mathbf{f}, \mathbf{F})$ . This equation is equivalent to

$$\xi \mathbf{w} - \mu \operatorname{div} \mathbf{Q} - (\lambda + \mu) \nabla \operatorname{tr} \mathbf{Q} = \mathbf{f}, \quad (75a)$$

$$\xi \mathbf{Q} - \nabla \mathbf{w} = \mathbf{F}. \quad (75b)$$

From (75b), we express  $\mathbf{Q}$  as a function of  $\mathbf{w}$ . Replacing in (75a), we obtain

$$\xi \mathbf{w} - \xi^{-1} \mu \Delta \mathbf{w} - \xi^{-1} (\lambda + \mu) \nabla \operatorname{div} \mathbf{w} = \mathbf{f} + \xi^{-1} \mu \operatorname{div} \mathbf{F} + \xi^{-1} (\lambda + \mu) \nabla \operatorname{tr} \mathbf{F},$$

which can be written in variational form as: find  $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$  such that, for every  $\bar{\mathbf{w}} \in \mathbf{H}_0^1(\Omega)$ ,

$$\begin{aligned} \xi \int_{\Omega} \mathbf{w} \cdot \bar{\mathbf{w}} + \xi^{-1} \mu \int_{\Omega} \nabla \mathbf{w} : \nabla \bar{\mathbf{w}} + \xi^{-1} (\lambda + \mu) \int_{\Omega} \operatorname{div} \mathbf{w} \operatorname{div} \bar{\mathbf{w}} = \\ \int_{\Omega} \mathbf{f} \cdot \bar{\mathbf{w}} - \xi^{-1} \mu \int_{\Omega} \mathbf{F} : \nabla \bar{\mathbf{w}} - \xi^{-1} (\lambda + \mu) \int_{\Omega} \operatorname{tr} \mathbf{F} \operatorname{div} \bar{\mathbf{w}}. \end{aligned} \quad (76)$$

Using the Lax-Milgram theorem, the above equation admits a unique weak solution  $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$ . Then, we recover  $\mathbf{Q} = \xi^{-1} (\nabla \mathbf{w} + \mathbf{F}) \in \nabla \mathbf{H}_0^1$ . From this expression for  $\mathbf{Q}$ , we deduce

$$\int_{\Omega} \mathbf{Q} : \nabla \bar{\mathbf{w}} = \xi^{-1} \int_{\Omega} (\nabla \mathbf{w} + \mathbf{F}) : \nabla \bar{\mathbf{w}} \quad \text{and} \quad \operatorname{tr} \mathbf{Q} = \xi^{-1} (\operatorname{div} \mathbf{w} + \operatorname{tr} \mathbf{F}).$$

Together with (76), it implies that, for every  $\bar{\mathbf{w}} \in \mathbf{H}_0^1(\Omega)$ ,

$$\mu \int_{\Omega} \mathbf{Q} : \nabla \bar{\mathbf{w}} + (\lambda + \mu) \int_{\Omega} \operatorname{tr} \mathbf{Q} \operatorname{div} \bar{\mathbf{w}} = \int_{\Omega} (\mathbf{f} - \xi \mathbf{w}) \cdot \bar{\mathbf{w}}.$$

Let  $\varphi \in \mathbf{H}_0^1(\Omega)$  such that  $\mathbf{Q} = \nabla \varphi$ , from the previous variational formulation, we see that

$$\varphi = \mathcal{A}_0^{-1} (\mathbf{f} - \xi \mathbf{w}).$$

Therefore, since  $\mathbf{f} - \xi \mathbf{w} \in \mathbf{L}^2(\Omega)$ , we have  $\varphi \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$  and  $\mathbf{Q} \in \nabla \mathbf{H}_0^2$ . Finally,  $\xi \in \rho(\mathcal{A})$  and we obtain  $(0, +\infty) \subset \rho(\mathcal{A})$ .

In order to prove the estimate for the norm of the resolvent operator  $(\xi \mathcal{I}_d - \mathcal{A})^{-1}$ , from (75), we easily get

$$\begin{aligned} \xi \int_{\Omega} |\mathbf{w}|^2 - \mu \int_{\Omega} \operatorname{div} \mathbf{Q} \cdot \mathbf{w} - (\lambda + \mu) \int_{\Omega} \nabla \operatorname{tr} \mathbf{Q} \cdot \mathbf{w} &= \int_{\Omega} \mathbf{f} \cdot \mathbf{w}, \\ \xi \int_{\Omega} |\mathbf{Q}|^2 - \int_{\Omega} \nabla \mathbf{w} : \mathbf{Q} &= \int_{\Omega} \mathbf{F} : \mathbf{Q}. \end{aligned}$$

Integrating by parts and summing the above two relations, we obtain

$$\begin{aligned} \xi \int_{\Omega} |\mathbf{w}|^2 + \xi \int_{\Omega} |\mathbf{Q}|^2 + (\mu - 1) \int_{\Omega} \mathbf{Q} : \nabla \mathbf{w} \\ + (\lambda + \mu) \int_{\Omega} \operatorname{tr} \mathbf{Q} \operatorname{div} \mathbf{w} = \int_{\Omega} \mathbf{f} \cdot \mathbf{w} + \int_{\Omega} \mathbf{F} : \mathbf{Q}. \end{aligned} \quad (77)$$

From (75b), we can equivalently write that

$$\nabla \mathbf{w} = \xi \mathbf{Q} - \mathbf{F} \quad \text{and} \quad \operatorname{div} \mathbf{w} = \xi \operatorname{tr} \mathbf{Q} - \operatorname{tr} \mathbf{F}.$$

Replacing in (77), it follows that

$$\begin{aligned} \xi \int_{\Omega} |\mathbf{w}|^2 + \xi \int_{\Omega} |\mathbf{Q}|^2 + (\mu - 1) \int_{\Omega} \mathbf{Q} : (\xi \mathbf{Q} - \mathbf{F}) \\ + (\lambda + \mu) \int_{\Omega} \operatorname{tr} \mathbf{Q} (\xi \operatorname{tr} \mathbf{Q} - \operatorname{tr} \mathbf{F}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{w} + \int_{\Omega} \mathbf{F} : \mathbf{Q}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \xi \int_{\Omega} |\mathbf{w}|^2 + \xi \mu \int_{\Omega} |\mathbf{Q}|^2 + \xi (\lambda + \mu) \int_{\Omega} |\operatorname{tr} \mathbf{Q}|^2 = \\ \int_{\Omega} \mathbf{f} \cdot \mathbf{w} + \mu \int_{\Omega} \mathbf{F} : \mathbf{Q} + (\lambda + \mu) \int_{\Omega} \operatorname{tr} \mathbf{F} \operatorname{tr} \mathbf{Q}. \end{aligned}$$

Finally, applying the Cauchy-Schwarz inequality, we get that

$$\xi \|(\mathbf{w}, \mathbf{Q})\|_{\mathbf{w}} \leq \|(\mathbf{f}, \mathbf{F})\|_{\mathbf{w}} \quad \text{and} \quad \|(\xi \mathcal{I}_d - \mathcal{A})^{-1}\|_{\mathcal{L}(\mathbf{w})} \leq \xi^{-1},$$

which concludes the proof of the lemma.  $\square$

## REFERENCES

- [1] M. Asch and G. Lebeau, [Geometrical aspects of exact boundary controllability for the wave equation—a numerical study](#), *ESAIM Control Optim. Calc. Var.*, **3** (1998), 163–212.
- [2] E. Bécache, P. Joly and C. Tsogka, [A new family of mixed finite elements for the linear elastodynamic problem](#), *SIAM J. Numer. Anal.*, **39** (2002), 2109–2132.
- [3] D. Boffi, F. Brezzi and M. Fortin, *Mixed Finite Element Methods and Applications*, volume 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013.
- [4] E. Burman, A. Feizmohammadi, A. Münch and L. Oksanen, [Space time stabilized finite element methods for a unique continuation problem subject to the wave equation](#), *ESAIM Math. Model. Numer. Anal.*, **55** (2021), S969–S991.
- [5] E. Burman, A. Feizmohammadi, A. Munch and L. Oksanen, [Spacetime finite element methods for control problems subject to the wave equation](#), working paper or preprint, September 2021.
- [6] F. L. Cardoso-Ribeiro, D. Matignon and L. Lefèvre, [A partitioned finite element method for power-preserving discretization of open systems of conservation laws](#), *IMA J. Math. Control Inform.*, **38** (2021), 493–533.
- [7] N. Cîndea, E. Fernández-Cara and A. Münch, [Numerical controllability of the wave equation through primal methods and Carleman estimates](#), *ESAIM Control Optim. Calc. Var.*, **19** (2013), 1076–1108.
- [8] N. Cîndea and A. Münch, [A mixed formulation for the direct approximation of the control of minimal  \$L^2\$ -norm for linear type wave equations](#), *Calcolo*, **52** (2015), 245–288.

- [9] R. Font and F. Periago, [Numerical simulation of the boundary exact control for the system of linear elasticity](#), *Appl. Math. Lett.*, **23** (2010), 1021–1026.
- [10] R. Glowinski, [Ensuring well-posedness by analogy: Stokes problem and boundary control for the wave equation](#), *J. Comput. Phys.*, **103** (1992), 189–221.
- [11] R. Glowinski, W. Kinton and M. F. Wheeler, [A mixed finite element formulation for the boundary controllability of the wave equation](#), *Internat. J. Numer. Methods Engrg.*, **27** (1989), 623–635.
- [12] 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** (1990), 1–76.
- [13] F. Hecht, [New development in freefem++](#), *J. Numer. Math.*, **20** (2012), 251–265.
- [14] L. I. Ignat and E. Zuazua, [Convergence of a two-grid algorithm for the control of the wave equation](#), *J. Eur. Math. Soc. (JEMS)*, **11** (2009), 351–391.
- [15] 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. Contrôlabilité exacte. [Exact controllability], With appendices by E. Zuazua, C. Bardos, G. Lebeau and J. Rauch.
- [16] S. Montaner and A. Münch, [Approximation of controls for linear wave equations: A first order mixed formulation](#), *Math. Control Relat. Fields*, **9** (2019), 729–758.
- [17] A. Osses, [A rotated multiplier applied to the controllability of waves, elasticity, and tangential Stokes control](#), *SIAM J. Control Optim.*, **40** (2001), 777–800.
- [18] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, volume 44 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 1983.
- [19] A. Rincon and I.-S. Liu, On numerical approximation of an optimal control problem in linear elasticity, *Divulg. Mat.*, **11** (2003), 91–107.
- [20] A. Serhani, D. Matignon and G. Haine, Partitioned finite element method for port-Hamiltonian systems with boundary damping: anisotropic heterogeneous 2D wave equations, *IFAC-PapersOnLine*, **52** (2019), 96–101.
- [21] R. Temam and A. Miranville, *Mathematical Modeling in Continuum Mechanics*, Cambridge University Press, Cambridge, second edition, 2005.
- [22] M. Tucsnak and G. Weiss, *Observation and Control for Operator Semigroups*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 2009.
- [23] E. Zuazua, [Propagation, observation, and control of waves approximated by finite difference methods](#), *SIAM Review*, **47** (2005), 197–243.

Received January 2022; revised May 2022; early access June 2022.

*E-mail address:* [arthur.bottois@uca.fr](mailto:arthur.bottois@uca.fr)

*E-mail address:* [nicolae.cindea@uca.fr](mailto:nicolae.cindea@uca.fr)

# Perspectives

On donne ici quelques perspectives et quelques pistes pour prolonger chacun des articles du Chapitre 2.

- A. Dans cet article, on a utilisé la formulation mixte espace-temps introduite dans [8] pour approcher numériquement le contrôle. Pour s’assurer de la convergence de l’approximation quand le paramètre de discrétisation  $h$  tend vers 0, on peut vérifier que les éléments finis choisis (HCT réduit et  $\mathbb{P}_1$ ) satisfont une condition *inf-sup* discrète, c’est-à-dire qu’ils satisfont une inégalité *inf-sup* uniformément par rapport à  $h$ . Or, il semble n’exister actuellement qu’une justification numérique de cette propriété. Ainsi, on pourrait chercher à montrer rigoureusement que les éléments finis choisis satisfont la condition *inf-sup* discrète, ou plus généralement à caractériser quels types d’éléments finis satisfont cette condition.

D’autre part, pour un état initial fixé, on a cherché à optimiser le support de contrôle. On a vu que le problème était bien posé pour des supports de contrôle  $q_\gamma$  définis comme le déplacement rigide d’un intervalle  $\omega$  le long de courbes  $\gamma$  uniformément lipschitziennes. On pourrait tenter d’étendre le problème à des supports de contrôle plus généraux, par exemple en prenant simplement  $q$  ouvert avec éventuellement un nombre fixé de composantes connexes. Dans le cas où le support de contrôle est pris indépendant du temps, un certain nombre de travaux (voir [29, 31, 28]) indiquent que le problème est mal posé dans le sens où l’optimum n’est pas atteint dans l’ensemble des fonctions caractéristiques  $\mathbb{1}_q$  mais dans l’ensemble des densités  $\theta \in L^\infty(Q; [0, 1])$ . Pour un support de contrôle dépendant du temps, il semble possible, au moins dans quelques cas simples comme le contrôle d’une onde sinusoïdale pure, de calculer explicitement la densité optimale.

- B. Dans cet article, on a adapté la formulation mixte espace-temps introduite dans [8] pour approcher numériquement le contrôle. Cependant, la convergence de l’approximation n’est pas garantie puisque les éléments finis choisis (HCT réduit et  $\mathbb{P}_1$ ) pour les simulations ne satisfont pas la condition *inf-sup* discrète, sauf à prendre le paramètre d’augmentation  $r$  proportionnel à  $h^2$ . On pourrait donc chercher à identifier des éléments finis qui satisfont (au moins “numériquement”) la condition *inf-sup*.

D’autre part, pour un état initial fixé, on a cherché à optimiser numériquement le support de contrôle. On n’a cependant pas montré que le problème était bien posé. Pour ce faire, il serait utile de disposer d’une inégalité d’observabilité uniforme par rapport au support de contrôle, à l’image de celle démontrée dans l’Art. A. Une telle inégalité peut être déduite de [7]. Ceci étant, le critère que doit vérifier le support de contrôle pour que l’inégalité d’observabilité soit vérifiée est un peu opaque. Un axe de recherche pourrait être de formuler un critère géométrique simple, à l’image du critère (GCC) utilisé pour un contrôle interne ou au bord (voir [24]).

- C. Dans cet article, le contrôle du problème linéarisé (lié au calcul de la direction de descente dans l’algorithme moindres carrés) est approché par minimisation directe de la fonctionnelle conjuguée (voir [18]). Cependant, cette méthode ne donne pas un résultat satisfaisant pour des semi-linéarités  $g$  vérifiant  $rg(r) \leq 0$  et telles que la solution non contrôlée du système semi-linéaire croît exponentiellement. Pour remédier à cela, on pourrait approcher (d’abord en 1D puis éventuellement en 2D) le contrôle du problème linéarisé via la formulation mixte espace-temps introduite dans [8]. Cette méthode semble prometteuse du fait de sa moindre sensibilité vis-à-vis de la taille du terme de potentiel présent dans le système linéarisé. En outre, la mise en œuvre de cette méthode pourrait permettre de vérifier numériquement l’optimalité des hypothèses faites sur la semi-linéarité  $g$ .

- D. Dans cet article, on a adapté la formulation mixte espace-temps stabilisée introduite dans [26] pour approcher numériquement le contrôle. Cependant, en l’état, on constate un phénomène de verrouillage numérique qui met en défaut la convergence de l’approximation quand le paramètre de discrétisation  $h$  tend vers 0. Il faudrait ainsi réussir à identifier l’origine de ce phénomène. Pour ce faire, on pourrait

chercher une estimation de l'erreur en fonction des différents paramètres de stabilisation.

Par ailleurs, la formulation mixte est gourmande en ressources mémoires. Il s'agit en effet de résoudre un système de taille  $12n_h \times 12n_h$ , où  $n_h$  est la dimension de l'espace des fonctions  $\mathbb{P}_1$  sur une triangulation  $\mathcal{T}_h$  du domaine espace-temps à trois dimensions (deux en espace et une en temps). Ceci reste vrai même si on résout la formulation mixte via la formulation biduale, ce qui revient à utiliser le complément de Schur d'un des blocs de la matrice du système et qui ramène le problème à la résolution itérée de deux systèmes de taille  $6n_h \times 6n_h$ . Pour diminuer davantage la taille des systèmes à résoudre, on pourrait envisager d'employer une méthode de décomposition de domaine de type Schwarz. Ceci permettrait de se ramener à la résolution itérée d'une multitude de systèmes de petite taille.

Une autre stratégie pour réduire les besoins en ressources mémoires pourrait être d'approcher le contrôle par minimisation directe de la fonctionnelle conjuguée (voir [17, 11]). Cette méthode passe par l'approximation de trajectoires du système de l'élasticité par différences finies en temps et par éléments finis en espace. Cependant, les expérimentations menées avec des éléments finis classiques (de type Raviart–Thomas par exemple) se sont révélées peu concluantes. On pourrait alors songer à utiliser des éléments finis non conformes pour obtenir une approximation plus robuste du contrôle.