AN OPTIMAL CONTROL PROBLEM FOR A KIRCHHOFF-TYPE EQUATION

M. Delgado[1], G.M. Figueiredo[2], I. Gayte[1] and C. Morales-Rodrigo[1]

Abstract. In this paper we study a control problem for a Kirchhoff-type equation. The method to obtain first order necessary optimality conditions is the Dubovitskii–Milyoutin formalism because the classical arguments do not work. We obtain a characterization of the optimal control by a partial differential system which is solved numerically.

1. Introduction

Nonlocal problems start out in the papers of Kirchhoff, [8], where he introduced his celebrated equation

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = \left(\dfrac{\tau_0}{\rho} + \dfrac{k}{2\rho l} \displaystyle\int_0^l \left(\dfrac{\partial u}{\partial x} \right)^2 \, \mathrm{d}x \right) \dfrac{\partial^2 u}{\partial x^2} \text{ for } 0 < x < l, \ t > 0, \\ \qquad\qquad u(0) = u(l) = 0, \end{cases} \tag{1.1}$$

which extended the D'Alembert equation of the vibration of a string with fixed ends subjected to transverse vibrations. The stationary version of the general problem with the Dirichlet condition can be written as

$$\begin{cases} -M \left(x, \displaystyle\int_\Omega |\nabla u|^2 \, \mathrm{d}x \right) \Delta u = f(x) \text{ in } \Omega, \\ u = 0 \qquad\qquad\qquad\qquad\quad \text{ on } \partial\Omega. \end{cases} \tag{1.2}$$

There are a relevant range of situations in Physics and Technology where problems with nonlocal terms appear, see for instance [1, 2, 4, 10, 14] and the survey [13].

In this paper we study an optimal control problem associated to (1.2) in the particular case $M(x, r) = a(x) + b(x)r$ where $a(x)$, $b(x)$ are positive continuous functions. Let us stress that since a and b are not constants the problem does not have a variational structure, as a consequence, the knowledge for this problem is very limited. In [15] it is proved that (1.2) has a solution in $H_0^1(\Omega)$, see also [4] for the uniqueness under the restriction $f \geq 0$.

Here we apply the Dubovitskii–Milyoutin formalism which enables us to obtain an optimality system which is verified for the optimal control, the associated state and the adjoint variable. We study its numerical

[1] Departement de Ecuaciones Diferenciales y Análisis Numérico, Faculdade de Matemáticas, Universidade de Sevilla Calle Tarfia s/n, 41012 Sevilla, Spain. madelgado@us.es; gayte@us.es; cristianm@us.es

[2] Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, 66.075-110 Belém-Pará, Brazil. giovany@ufpa.br

approximation. The simulations point out the uniqueness of solution of the optimality system, but we have not been able to prove it. This uniqueness would imply the uniqueness of the solution of the optimal control problem.

The numerical study of the Kirchhoff equation is not an easy task. Moreover, we are aware about very few papers in the topic, see [7,16]. In our case, the optimality system is composed by two coupled elliptic equations of the Kirchhoff-type. As far as we know, we have developed a new method that reduce the problem to the solution of several Poisson problems, which are obtained with the help of the Regula Falsi algorithm.

It is worth to quote that a number a recent papers of Lou and coworkers (see, for instance, [11]) consider optimal control problems for semilinear elliptic/parabolic partial differential equations with the leading term containing the control addressing both existence and necessary conditions. We cannot consider our problem in this framework: the weak convergence of the second member in H_0^1 is not enough to prove the existence of solution of the Kirchhoff problem and for this reason we address the control problem directly.

The paper is organized as follows. In Section 2, we study the existence of the optimal control. In Section 3, we applied the Dubovitskii-Milyoutin formalism to deduce the optimality system. In Section 4 we carry out the numerical approximation of the optimality system.

2. THE OPTIMAL CONTROL PROBLEM: THE EXISTENCE OF A SOLUTION

In this work we are interested in the study of the following optimal control problem:
To find a control a minimizing the functional $J : H_0^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$

$$J(u,a) = \frac{1}{2} \int_\Omega (u - u_d)^2 + \frac{B}{2} \int_\Omega a^2, \tag{2.1}$$

where (u,a) is a solution of the Kirchhoff equation

$$\begin{cases} -(a(x) + b(x) \int_\Omega |\nabla u|^2) \Delta u = f \text{ in } \Omega, \\ u = 0 \qquad\qquad\qquad\qquad \text{ on } \partial\Omega. \end{cases} \tag{2.2}$$

besides,

$$a \in \mathcal{U} = \{ v \in L^2(\Omega) : \quad v \geq a_0 > 0, \text{ a.e. } x \in \Omega \}, \tag{2.3}$$

the set $\Omega \subset \mathbb{R}^N$ is a bounded smooth open, $N \geq 1$, f is a given function in $L^\infty(\Omega)$, $f \geq f_0$, where f_0 is a positive real number, the functions a_0 and b are data and they verify $a_0 \in C(\overline{\Omega})$, $a_0 > 0$ and $b \in W^{1,\infty}(\Omega)$, $b \geq b_0 > 0$, u_d is also a known function, $u_d \in L^2(\Omega)$, and B is a positive parameter.

In the general topic of optimal control of nonlocal PDEs, it is not usual to study the control in a coefficient related with the nonlocal term. This question is mathematically relevant and moreover, this optimal control problem might give light to a parameter identification problem: to find a suitable function a, with L^2-norm small, such that u, the associated solution of a Kirchhoff problem, is close (in L^2-norm) to u_d.

For a given function $a \in L^2(\Omega)$, it is known the existence of $u \in H_0^1(\Omega)$, a solution of the Kirchhoff problem (see [15] for more general Kirchhoff equations), besides this solution is unique when $a \geq 0$. The ignorance about the existence and uniqueness of (2.2) when a is negative or changes its sign, is the reason why we cannot apply the classical theory (see [9]). We cannot write J like a functional j depending only on a and defined on an open set of $L^2(\Omega)$. We will explain this with more detail in Section 3.

We shall study the optimal control of J associated to the constraints (2.2) and (2.3) considering (u,a) as independent variables related through (2.2) and we will apply a generalization of the Lagrange multipliers theorem, called Dubovitskii–Milyutin formalism, which will provide us the Euler–Lagrange equation.
First of all, we define a weak solution of (2.2).

Definition 2.1. A function u is called a weak solution of (2.2) if $u \in H_0^1(\Omega)$ and verifies

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega \frac{f}{a + b \int_\Omega |\nabla u|^2} v \quad \forall v \in H_0^1(\Omega).$$

Our first step is to show that the admisible set is not empty.

Theorem 2.2. *For any $a \in \mathcal{U}$, there exists a unique $u \in H_0^1(\Omega)$ weak solution of the Kirchhoff problem (2.2). Besides, $u \in W^{2,p}(\Omega)$ for any $p < +\infty$, so it is a strong solution.*

The proof of this theorem is based on the following proposition:

Proposition 2.3. *Let $g : [0, +\infty) \to \mathbb{R}$ be the function defined by*

$$g(s) = s - \int_\Omega |\nabla u_s|^2,$$

where u_s is the solution of the Poisson problem

$$\begin{cases} -\Delta u_s = \dfrac{f}{a + bs} & in \ \Omega, \\ u_s = 0 & on \ \partial\Omega. \end{cases} \tag{2.4}$$

Then, g has got a unique root, r.

Proof. It is clear that g is continuous in $[0, +\infty)$. We will show that g is strictly increasing and it changes its sign between $s = 0$ and some $s > 0$. On multiplying (2.4) by u_s and integrating by parts we obtain

$$\int_\Omega |\nabla u_s|^2 = \int_\Omega \frac{f}{a + bs} u_s. \tag{2.5}$$

Then, if $s_1 < s_2$

$$\frac{f}{a + bs_1} \geq \frac{f}{a + bs_2},$$

and so, $-\Delta u_{s_1} \geq -\Delta u_{s_2}$. By the strong maximum principle $u_{s_1} \geq u_{s_2} \geq 0$. Using these two inequalities and (2.5) it is easy to prove that

$$s_1 - \int_\Omega |\nabla u_{s_1}|^2 < s_2 - \int_\Omega |\nabla u_{s_2}|^2$$

and so, g is strictly increasing.
Besides, $g(0) = -\int_\Omega |\nabla u_0|^2$, with u_0 the solution of (2.4) for $s = 0$, so $g(0) < 0$. Moreover, since

$$\int_\Omega |\nabla u_s|^2 = \int_\Omega \frac{f}{a + bs} u_s \leq \int_\Omega \frac{f}{a_0} u_0,$$

then $\int_\Omega |\nabla u_s|^2$ is bounded for any $s > 0$. So, $\lim_{s \to +\infty} g(s) = +\infty$.
Thus, the equation $g(s) = 0$ has a unique solution. $\qquad\square$

Proof of Theorem 2.2. As a consequence of Proposition 2.3, there exists a unique positive number r such that $g(r) = 0$, *i.e.*

$$r = \int_\Omega |\nabla u_r|^2,$$

and so, the function $u = u_r$ is the solution of (2.2). Besides, the second member of the equation of (2.4) belongs to $L^\infty(\Omega)$, so the solution $u = u_r \in W^{2,p}(\Omega)$, $p < +\infty$. This means that u, the solution of the Kirchhoff problem (2.2), is a strong solution. $\qquad\square$

Remark 2.4. Theorem 2.2 is true for any $a > 0$, in particular for $a \in \mathcal{U}$.

By the previous theorem, it makes sense to study the existence of optimal control of the problem

$$\min J(u, a) \tag{2.6}$$

with

$$J(u, a) = \frac{1}{2} \int_\Omega (u - u_d)^2 + \frac{B}{2} \int_\Omega a^2$$

$a \in \mathcal{U}$,

$$\begin{cases} -(a(x) + b(x) \int_\Omega |\nabla u|^2)\Delta u = f \text{ in } \Omega, \\ u = 0 \qquad\qquad\qquad\qquad \text{on } \partial\Omega. \end{cases}$$

Theorem 2.5. *There exists a solution* $(\hat{u}, \hat{a}) \in H_0^1(\Omega) \times L^2(\Omega)$ *of the optimal control problem* (2.6).

Proof. We will deduce the existence by the standard method of the minimizing sequences with the particularity that the functional only provides the boundedness of the sequences in $L^2(\Omega)$. Without loosing generality, we can assume that $u_d = 0$.

The proof will be divided in four steps.

The first step: The convergence of the minimizing sequence to $(\hat{u}, \hat{a})$.

Let us consider the infimum of J subject to (2.2) and (2.3) and a minimizing sequence, $\{J(u_n, a_n)\}$. If we call β the infimum of J, we have that there exists the limit of the sequence $J(u_n, a_n)$ and so, the sequences $\{u_n\}$ and $\{a_n\}$ are bounded in $L^2(\Omega)$.

Then, there exist two subsequences, that they will be denoted with the same subscript, and two functions $\hat{u}$, $\hat{a} \in L^2(\Omega)$, such that

$$u_n \rightharpoonup \hat{u} \text{ in } L^2(\Omega) - weak,$$
$$a_n \rightharpoonup \hat{a} \text{ in } L^2(\Omega) - weak.$$

The function $\hat{a}$ belongs to $\mathcal{U}$ because this set is convex and closed in $L^2(\Omega)$.

The functional J is weakly lower continuous because it is convex and continuous, so

$$\beta = \lim J(u_n, a_n) \geq J(\hat{u}, \hat{a}).$$

The second step: The convergence of the sequence $\{u_n\}$.

Since

$$-\Delta u_n = \frac{f}{a_n + b \int_\Omega |\nabla u_n|^2}$$

we get that

$$-\Delta u_n \geq \frac{f_0}{a_n + b \int_\Omega |\nabla u_n|^2} > 0.$$

Therefore

$$a_n = \frac{f}{-\Delta u_n} - b \int_\Omega |\nabla u_n|^2.$$

Moreover, by the elliptic regularity

$$u_n \in W^{2,p}(\Omega) \text{ for any } p \in (1, +\infty)$$

and

$$\|u_n\|_{W^{2,p}(\Omega)} \leq C \left\| \frac{f}{a_n + b \int_\Omega |\nabla u_n|^2} \right\|_{L^p(\Omega)}.$$

In particular, $\{u_n\} \subset H^2(\Omega)$ and $\|u_n\|_{H^2(\Omega)} \leq c$. So,

$$u_n \rightharpoonup \hat{u} \text{ in } H^2(\Omega) - weak.$$

The compact embedding $H^2(\Omega) \hookrightarrow H^1(\Omega)$ implies that

$$u_n \to \hat{u} \text{ in } H^1(\Omega) - strong$$

and

$$\Delta u_n \rightharpoonup \Delta \hat{u} \text{ in } L^2(\Omega) - weak.$$

The third step: We can assume that the functions a_n are bounded and continuous.

Consider a_1. Since the space of bounded and continuous functions on Ω, $C_b(\Omega)$, is dense in $L^2(\Omega)$, there exists a sequence $\{\tilde{a}_{1,n}\} \subset C_b(\Omega)$ which strongly converges to a_1 in $L^2(\Omega)$. For any $\tilde{a}_{1,n}$ there exists a unique $\tilde{u}_{1,n}$ solution of the Kirchhoff problem

$$\begin{cases} -(\tilde{a}_{1,n} + b \int_\Omega |\nabla \tilde{u}_{1,n}|^2)\Delta \tilde{u}_{1,n} = f \text{ in } \Omega, \\ \tilde{u}_{1,n} = 0 \qquad\qquad\qquad \text{ on } \partial\Omega. \end{cases} \tag{2.7}$$

Analogously to the second step, we obtain

$$\tilde{u}_{1,n} \to w \text{ in } H^1(\Omega)\text{-strong}$$

$$\Delta \tilde{u}_{1,n} \rightharpoonup \Delta w \text{ in } L^2(\Omega)\text{-weak.}$$

Then, we can pass to the limit in (2.7) and we have that

$$\begin{cases} -(a_1 + b \int_\Omega |\nabla w|^2)\Delta w = f \text{ in } \Omega, \\ w = 0 \qquad\qquad\qquad \text{ on } \partial\Omega. \end{cases}$$

By the uniqueness of the Kirchhoff problem, $w = u_1$. Then, there exists n_1 such that

$$\|\tilde{a}_{1,n_1} - a_1\|_{L^2(\Omega)} < 1,$$

$$\|\tilde{u}_{1,n_1} - u_1\|_{H^1(\Omega)} < 1$$

and

$$J(\tilde{u}_{1,n_1}, \tilde{a}_{1,n_1}) = \frac{1}{2}\int_\Omega |\tilde{u}_{1,n_1}|^2 + \frac{B}{2}\int_\Omega |\tilde{a}_{1,n_1}|^2 \le \frac{1}{2}(\|u_1\|_{L^2(\Omega)} + 1)^2 + \frac{B}{2}(\|a_1\|_{L^2(\Omega)} + 1)^2$$

$$= J(u_1, a_1) + \frac{1}{2} + \frac{B}{2} + \|u_1\|_{L^2(\Omega)} + B\|a_1\|_{L^2(\Omega)} \le J(u_1, a_1) + C,$$

where $C > 0$ is a constant which does not depend on n because the sequences $\|a_n\|_{L^2(\Omega)}$ and $\|u_n\|_{L^2(\Omega)}$ are bounded.

By the same procedure, we obtain a pair $(\tilde{u}_{2,n_2}, \tilde{a}_{2,n_2})$, with $n_2 > n_1$, such that

$$\|\tilde{a}_{2,n_2} - a_2\|_{L^2(\Omega)} < \frac{1}{2},$$

$$\|\tilde{u}_{2,n_2} - u_2\|_{H^1(\Omega)} < \frac{1}{2}$$

and

$$J(\tilde{u}_{2,n_2}, \tilde{a}_{2,n_2}) \le J(u_2, a_2) + C\frac{1}{2}.$$

Therefore, for any k we have a pair $(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k})$, with $n_k > n_{k-1}$, such that

$$\|\tilde{a}_{k,n_k} - a_k\|_{L^2(\Omega)} < \frac{1}{k},$$

$$\|\tilde{u}_{k,n_k} - u_k\|_{H^1(\Omega)} < \frac{1}{k}$$

and

$$J(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k}) \le J(u_k, a_k) + C\frac{1}{k}.$$

Then

$$\inf J(u,a) = \beta \le J(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k}) \le J(u_k, a_k) + C\frac{1}{k},$$

and taking the lower limit in the inequalities we obtain

$$\beta \le \liminf J(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k}) \le \lim J(u_k, a_k) = \beta.$$

So, we have that

$$\exists \lim J(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k}) = \beta.$$

This proves that the sequence $\{(\tilde{u}_{k,n_k}, \tilde{a}_{k,n_k})\}$ is a minimizing sequence, where $\tilde{a}_{k,n_k}$ is continuous and bounded, $\tilde{a}_{k,n_k} \in \mathcal{U}$ and $\tilde{u}_{k,n_k}$ is the solution of the Kirchhoff problem with $\tilde{a}_{k,n_k}$ and we claim

$$\tilde{a}_{k,n_k} \rightharpoonup \hat{a} \text{ in } L^2(\Omega)\text{-weak}$$

$$\tilde{u}_{k,n_k} \rightharpoonup \hat{u} \text{ in } L^2(\Omega)\text{-weak}.$$

In fact, for any $\varphi \in L^2(\Omega)$,

$$\int_\Omega (\tilde{a}_{k,n_k} - \hat{a})\varphi = \int_\Omega (\tilde{a}_{k,n_k} - a_k)\varphi + \int_\Omega (a_k - \hat{a})\varphi.$$

Given $\epsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that for any $k \ge k_0$ $\|\tilde{a}_{k,n_k} - a_k\|_{L^2(\Omega)} < \frac{1}{k} < \epsilon$. Then

$$\left| \int_\Omega (\tilde{a}_{k,n_k} - a_k)\varphi \right| \le \|\tilde{a}_{k,n_k} - a_k\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)} < \epsilon \|\varphi\|_{L^2(\Omega)}.$$

So, we have

$$\lim \int_\Omega (\tilde{a}_{k,n_k} - a_k)\varphi = 0$$

and, on the other hand

$$\lim \int_\Omega (a_k - \hat{a})\varphi = 0,$$

by the weak convergence of $\{a_n\}$.

A similar reasoning proves the convergence for $\{\tilde{u}_{k,n_k}\}$.

From now on we rename $\tilde{a}_{k,n_k}$ as a_k and $\tilde{u}_{k,n_k}$ as u_k.

The fourth step: The pair $(\hat{u}, \hat{a})$ is an admissible element, *i.e.*, it satisfies (2.2) and $\hat{a} \in \mathcal{U}$ (this last one has been already proved in the first step).

Since $\{a_n\}$ and $\{\Delta u_n\}$ weakly converge to $\hat{a}$ and $\Delta \hat{u}$, by Mazur's Lemma there exist two sequences, one for each sequence, constituted by convex linear combinations, such that they strongly converge (see [17]), *i.e.*

$$\sum_{i \in I_n} \lambda_i a_i \to \hat{a} \text{ in } L^2(\Omega)\text{-strong},$$

$$\sum_{j \in J_n} \mu_j \Delta u_j \to \Delta \hat{u} \text{ in } L^2(\Omega)\text{-strong}.$$

By Egoroff's theorem (see [5]), for any $\varepsilon > 0$ there exists a set $A_\varepsilon \subset \Omega$, with $|A_\varepsilon| < \varepsilon$, such that

$$\sum_{i \in I_n} \lambda_i a_i \to \hat{a} \text{ uniformly in } \Omega \setminus A_\varepsilon$$

$$\sum_{j \in J_n} \mu_j \Delta u_j \to \Delta \hat{u} \text{ uniformly in } \Omega \setminus A_\varepsilon. \tag{2.8}$$

Besides, the functions $b, f, \hat{a}, \Delta \hat{u}$ have continuous representatives in $\Omega \setminus A_\varepsilon$, equally denoted.

Let $D = \{x_n\}_n$ be a countable dense subset of Ω. We consider $D_\varepsilon = \{x_n \in D : x_n \in \Omega \setminus A_\varepsilon\}$.
For any $x_m \in D_\varepsilon$ and $\delta > 0$, the set

$$\{n \in \mathbb{N} : |a_n(x_m) - \hat{a}(x_m)| < \delta \text{ and } |\Delta u_n(x_m) - \Delta \hat{u}(x_m)| < \delta\}$$

is countably infinity. In fact, if we suppose that there is a finite number of n in this set, then there are infinite n such that

$$|a_n(x_m) - \hat{a}(x_m)| \geq \delta$$

or

$$|\Delta u_n(x_m) - \Delta \hat{u}(x_m)| \geq \delta.$$

If the set of n that verify $a_n(x_m) - \hat{a}(x_m) \geq \delta$ is not finite, the sequence of the convex linear combinations of a_n satisfies

$$\sum_{i \in I_n} \lambda_i a_i(x_m) - \hat{a}(x_m) \geq \delta,$$

but this is impossible by (2.8). Analogously if the inequality is $a_n(x_m) - \hat{a}(x_m) \leq -\delta$, and it happens the same if there are infinite numbers n which verify $|\Delta u_n(x_m) - \Delta \hat{u}(x_m)| \geq \delta$.

We define the set

$$N_1^{(1)} = \{n \in \mathbb{N} : |a_n(x_1) - \hat{a}(x_1)| < 1, |\Delta u_n(x_1) - \Delta \hat{u}(x_1)| < 1\}$$

which is infinite. Let be $n_1^{(1)}$ a chosen element in this set. In the step i we chose $n_i^{(1)} \in N_i^{(1)}$, where

$$N_i^{(1)} = \{n \in N_{i-1}^{(1)} : n > n_{i-1}^{(1)}, |a_n(x_i) - \hat{a}(x_i)| < 1, |\Delta u_n(x_i) - \Delta \hat{u}(x_i)| < 1\}.$$

We have just defined a subsequence $n^{(1)} = \{n_1^{(1)}, n_2^{(1)}, \ldots\}$ such that

$$|a_{n_i^{(1)}}(x_i) - \hat{a}(x_i)| < 1, |\Delta u_{n_i^{(1)}}(x_i) - \Delta \hat{u}(x_i)| < 1.$$

We repeat the reasoning decreasing the estimations:

In the step l, let be

$$N_1^{(l)} = \left\{n \in n^{(l-1)} : |a_n(x_1) - \hat{a}(x_1)| < \frac{1}{l}, |\Delta u_n(x_1) - \Delta \hat{u}(x_1)| < \frac{1}{l}\right\},$$

and $n_1^{(l)}$ an element of this set. Again, taking each element in the set D_ε we obtain a subsequence of $n^{(l-1)}$, named $n^{(l)}$ such that

$$n^{(l)} = \left\{n_1^{(l)}, n_2^{(l)}, \ldots\right\}$$

and

$$|a_{n_i^{(l)}}(x_i) - \hat{a}(x_i)| < \frac{1}{l}, |\Delta u_{n_i^{(l)}}(x_i) - \Delta \hat{u}(x_i)| < \frac{1}{l}. \tag{2.9}$$

We know that

$$-\left(a_{n_i^{(l)}}(x_i) + b(x_i) \int_\Omega |\nabla u_{n_i^{(l)}}|^2\right) \Delta u_{n_i^{(l)}}(x_i) = f(x_i).$$

By (2.9) and knowing that

$$\int_\Omega |\nabla u_n|^2 \rightarrow \int_\Omega |\nabla \hat{u}|^2$$

we pass to the limit when l tends to $+\infty$ and we obtain

$$-\left(\hat{a}(x_i) + b(x_i) \int_\Omega |\nabla \hat{u}|^2\right) \Delta \hat{u}(x_i) = f(x_i) \quad \forall x_i \in D_\varepsilon.$$

Since D_ε is a dense set in $\Omega \setminus A_\varepsilon$ and the functions in the previous equation are continuous in $\Omega \setminus A_\varepsilon$, we can extend the equation to every point $x \in \Omega \setminus A_\varepsilon$, so

$$-\left(\hat{a} + b\int_\Omega |\nabla\hat{u}|^2\right)\Delta\hat{u} = f \text{ in } \Omega \setminus A_\varepsilon.$$

Then,

$$\left\|-(\hat{a} + b\int_\Omega |\nabla\hat{u}|^2)\Delta\hat{u} - f\right\|_{L^2(\Omega)} = \left\|-(\hat{a} + b\int_\Omega |\nabla\hat{u}|^2)\Delta\hat{u} - f\right\|_{L^2(A_\varepsilon)},$$

and the measure of the set A_ε tends to zero when ε tends to zero. Taking the limit in this equality when ε tends to zero we obtain that

$$\left\|-(\hat{a} + b\int_\Omega |\nabla\hat{u}|^2)\Delta\hat{u} - f\right\|_{L^2(\Omega)} = 0,$$

i.e.

$$-\left(\hat{a} + b\int_\Omega |\nabla\hat{u}|^2\right)\Delta\hat{u} = f \text{ in } \Omega.$$

Since $(\hat{u}, \hat{a})$ satisfies the constraints and $J(\hat{u}, \hat{a}) = \beta$ we get that $(\hat{u}, \hat{a})$ is an optimal solution. $\qquad\square$

3. The optimality system

In this section we study a characterization of an optimal solution. We denote $(\hat{u}, \hat{a})$ an optimal solution, so $\hat{a} \in \mathcal{U}$ and $\hat{u}$ is its unique associated state, *i.e.* the solution of (2.2) associated to $\hat{a}$. For this purpose we will use the Dubovitskii–Milyutin theorem (D-M formalism) (see [6]). We could think that the minimization of the functional $j(a) = J(G(a), a)$, in $a \in \mathcal{U}$, where G is defined as follows:

$$G : a \mapsto u \in H_0^1(\Omega),$$

and u is the solution of Kirchhoff problem (2.2) for a could work. We know that G is well-defined on $\mathcal{U}$, by Theorem 2.2, and continuous, by the second step of the proof of Theorem 2.5, but, to differentiate G we need to define it on a open set of $L^2(\Omega)$ topology and this is not possible because we do not know if Kirchhoff problem (2.2) has a unique solution u associated to a when $a \not\geq 0$ for a.e. $x \in \Omega$, in general. So, we cannot define $G(a)$ in an open set of $L^2(\Omega)$ as it occurs in the classical control theory. This fact leads us to use D-M formalism.

Our goal is to write the optimal control $\hat{a}$ as a solution of a partial differential system and a variational inequality provided by the constraint $a \in \mathcal{U}$. We refer to [6] for the following theorem and definitions. The Dubovitskii–Milyutin theorem is:

Theorem 3.1. *Let X be a normed space. Assume that the functional $J: X \to \mathbb{R}$ has a local minimum with constraints $Z = \bigcap_{i=1}^{n+1} Z_i \subset X$ at a point $x_0 \in Z$. Assume that*

(a) *J is regularly decreasing at x_0, with decreasing (and convex) cone DC_0;*
(b) *the inequality constraints Z_i, $1 \leq i \leq n$, are regular at x_0, with feasible (and convex) cones FC_i, $1 \leq i \leq n$;*
(c) *the equality constraint Z_{n+1} is also regular at x_0, with tangent (and convex) cone TC_{n+1}.*

Then, there exist continuous linear functionals $f_0 \in DC_0^*$, $f_i \in FC_i^*$ for $1 \leq i \leq n$ and $f_{n+1} \in TC_{n+1}^*$ (we denote by * the corresponding dual cone), not all identically zero, such that they satisfy the Euler–Lagrange equation:

$$f_0 + \sum_{i=1}^n f_i + f_{n+1} = 0 \quad \text{in } X'.$$

In our case, since J is Fréchet differentiable, the dual decreasing cone at $(\hat{u}, \hat{a})$, DC_0, is given by [6]

$$f_0 = -\lambda J'(\hat{u}, \hat{a}), \quad \lambda \geq 0.$$

The constraint $a \in \mathcal{U}$, apparently an inequality constraint, has to be considered like an equality constraint because the interior of $\mathcal{U}$ in $L^2(\Omega)$ is empty. Then, the constraints (2.2) and (2.3) generate an only dual tangent cone at $(\hat{u}, \hat{a})$, called TC^*. Applying the Dubovitskii–Milyutin theorem, there exist $f_0 \in DC_0^*$ and $\tilde{f} \in TC^*$, not all indentically zero, such that they satisfy the Euler–Lagrange equation:

$$f_0 + \tilde{f} = 0 \quad \text{in } H^{-1}(\Omega) \times (L^2(\Omega))'. \tag{3.1}$$

Clearly $\lambda \neq 0$, because if $\lambda = 0$ both functionals f_0 and f would be indentically zero so, we can choose it equal to 1.

We do not know to characterize TC^*. To overcome this difficulty, we write f as a sum of two functionals, $\tilde{f} = f_1 + f_2$, related with the constraints (2.2) and (2.3). In fact, f_1 is related to (2.2) and it belongs to TC_1^*, the dual tangent cone at $(\hat{u}, \hat{a})$; the functional f_2 has the form $f_2 = (0, \tilde{f}_2)$, it belongs to the dual tangent cone related to (2.3), TC_2^*, and it verifies $\langle \tilde{f}_2, a - \hat{a} \rangle \geq 0$ for every $a \in \mathcal{U}$. The particular form of f_2 is due to the fact that (2.3) is independent of u.

When both dual cones are a system of cones of the same sense (see definition in [18]), it is true that

$$TC^* = TC_1^* + TC_2^*,$$

and the characterization of TC^* is reached. We have also sufficient conditions to know if two dual cones, one of which has the form of TC_2^*, are of the same sense.

Theorem 3.2 ((Walczak) Thm. 3.3 of [16]). *Let $E = X \times Y$, where X and Y are normed spaces. Denote by $K_1 \subset E$ a cone of the form*

$$K_1 = \{v_1 = (x_1, y_1) \in E, \ x_1 = Ay_1\},$$

where $x_1 \in X$, $y_1 \in Y$ and $A : Y \to X$ is a linear operator. Let $K_2 = X \times \tilde{K}_2$, where $\tilde{K}_2$ is a cone in Y. If the operator A is linear and continuous, then

$$K_1^* = \{(x_1^*, y_1^*) \in E^*, \ y_1^* = -A^* x_1^*\},$$
$$K_2^* = \{(0, y_2^*) \in E^*, \ y_2^* \in \tilde{K}_2^{\,*}\},$$

and the cones K_1^*, K_2^* are the same sense. Here, A^* stands for an operator dual to A.

Now we prove that TC_1^* has the form of K_1 in the Walczak theorem.

3.1. The characterization of the tangent cone at $(\hat{u}, \hat{a})$ associated to (2.2)

We call P the operator defined by (2.2):

$$P : H_0^1(\Omega) \times L^2(\Omega) \to H^{-1}(\Omega)$$
$$P(u, a) = -\left(a + b \int_\Omega |\nabla u|^2\right) \Delta u$$

Let be P' the derivative operator. We have

Proposition 3.3. *The operator $P'(\hat{u}, \hat{a})$ is surjective as a linear map from $H_0^1(\Omega) \times L^2(\Omega)$ to $H^{-1}(\Omega)$.*

Proof. Given any $g \in H^{-1}(\Omega)$, we will prove the existence of $(u, a) \in H_0^1(\Omega) \times L^2(\Omega)$ verifying $P'(\hat{u}, \hat{a})(u, a) = g$ i.e.

$$-\left(\hat{a} + b \int_\Omega |\nabla \hat{u}|^2\right) \Delta u - \left(a + 2b \int_\Omega \nabla \hat{u} \cdot \nabla u\right) \Delta \hat{u} = g. \tag{3.2}$$

Since u, a are independent, we are going to define u and we will obtain a from the equation.

Let u be the solution of the Poisson problem

$$
\begin{cases}
-\Delta u = \dfrac{1}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} g \text{ in } \Omega, \\
u = 0 \qquad\qquad\qquad \text{on } \partial\Omega.
\end{cases}
$$

Then, replacing Δu in (3.2) we obtain

$$
-\left(a + 2b \int_\Omega \nabla \hat{u} \cdot \nabla u \right) \Delta \hat{u} = 0,
$$

and so,

$$
a = -2b \int_\Omega \nabla \hat{u} \cdot \nabla u. \qquad\qquad \square
$$

By the Lyusternik theorem see [6], the tangent cone TC_1 is given by

$$
TC_1 = \{(u, a) \in H_0^1(\Omega) \times L^2(\Omega) : \ P'(\hat{u}, \hat{a})(u, a) = 0\},
$$

i.e., $(u, a) \in H_0^1(\Omega) \times L^2(\Omega)$ such that

$$
\begin{cases}
-\left(\hat{a} + b \int_\Omega |\nabla \hat{u}|^2 \right) \Delta u - \left(a + 2b \int_\Omega \nabla \hat{u} \cdot \nabla u \right) \Delta \hat{u} = 0 \text{ in } \Omega, \\
u = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{on } \partial\Omega.
\end{cases} \tag{3.3}
$$

In order to apply the Walczak theorem we need to prove that the operator which provides u in (3.3) given any $a \in L^2(\Omega)$ is well-defined, linear and continuous. These assertions are formulated in the following theorem.

Theorem 3.4. *The operator $L : L^2(\Omega) \to H_0^1(\Omega)$ defined by*

$$
a \in L^2(\Omega) \mapsto u = La,
$$

being u the solution of (3.3), *is well defined, linear and continuous.*

Proof. By the equality $\int_\Omega \nabla \hat{u} \cdot \nabla u = \int_\Omega (-\Delta \hat{u}) u$, we can write (3.3) as

$$
\begin{cases}
-\Delta u = \dfrac{2b \Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} \left(\int_\Omega (-\Delta \hat{u}) u \right) + \dfrac{\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} a \text{ in } \Omega, \\
u = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial\Omega.
\end{cases} \tag{3.4}
$$

for $\hat{a} \in \mathcal{U}$ an optimal control, $\hat{u}$ its unique associated state, and for each $a \in L^2(\Omega)$. In what follows, we will prove that there is a unique solution to the linear nonlocal problem (3.4). To this end, we define the function $h : [0, +\infty) \to \mathbb{R}$ as

$$
h(r) = r - \int_\Omega (-\Delta \hat{u}) u_r,
$$

where u_r is the solution of the Poisson problem

$$
\begin{cases}
-\Delta u_r = \dfrac{2b \Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} r + \dfrac{\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} a \text{ in } \Omega, \\
u_r = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial\Omega.
\end{cases} \tag{3.5}
$$

We will show that h possesses a unique root. We know that h is continuous because u_r depends continuously on r. We claim that h is strictly increasing function. In fact, let $r_1 < r_2$; then the respective solutions of (3.5) verify

$$
-\Delta(u_{r_1} - u_{r_2}) = \dfrac{2b \Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} (r_1 - r_2) > 0,
$$

because $\Delta \hat{u} < 0$. From the maximum principle it follows that $u_{r_1} > u_{r_2}$ and, therefore, $h(r_1) < h(r_2)$.

To prove that h changes sign, we first consider the case $a > 0$ in Ω. In this case, u_0 is the solution of

$$\begin{cases} -\Delta u_0 = \dfrac{\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} a \text{ in } \Omega, \\ u_0 = 0 \qquad\qquad\qquad \text{ on } \partial\Omega. \end{cases} \tag{3.6}$$

So, $u_0 < 0$ in Ω and $h(0) > 0$. On the other hand, if $r < 0$, then $u_r > u_0$ in Ω and

$$h(r) = r + \int_\Omega (-\Delta \hat{u}) u_r < r + \int_\Omega (-\Delta \hat{u}) u_0.$$

Consequently,

$$\lim_{r \to -\infty} h(r) = -\infty,$$

and there exists a change of sign in $(-\infty, 0)$ and, for the monotonicity, there exists a unique $r^+ < 0$ such that $h(r^+) = 0$. Analogously, in the case $a < 0$ the change of sign appears in $(0, +\infty)$ and there exists a unique $r^- > 0$ such that $h(r^-) = 0$.

If $a \in L^2(\Omega)$, we can write $a = a^+ - a^-$, being

$$a^+ = \max\{a, 0\}, \qquad \text{and} \qquad a^- = \min\{-a, 0\}.$$

Denoting by r^+ and r^- the respective roots of h for a^+ and a^-, it is easy to check that $u_{r^+ + r^-} = u_{r^+} + u_{r^-}$, and

$$h(r^+ + r^-) = r^+ + r^- - \int_\Omega (-\Delta \hat{u}) u_{r^+ + r^-} = r^+ - \int_\Omega (-\Delta \hat{u}) u_{r^+} + r^- - \int_\Omega (-\Delta \hat{u}) u_{r^-} = 0.$$

The uniqueness follows considering that (3.4) has only one solution for each $a \in L^2(\Omega)$. In fact if u_1, u_2 are two solutions for the same a, then the difference $w = u_1 - u_2$, verifies

$$-\Delta w = \frac{2b\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} \int_\Omega (-\Delta \hat{u}) w;$$

if $\int_\Omega (-\Delta \hat{u}) w > 0$, the maximum principle assures that $w < 0$ and this is not possible because $\Delta \hat{u} < 0$. We may argue in a similar way if $\int_\Omega (-\Delta \hat{u}) w < 0$. Therefore, $\int_\Omega (-\Delta \hat{u}) w = 0$, $-\Delta w = 0$ and $w = 0$.

Therefore (3.4) has a unique solution which implies that L is well defined. It is obviously linear. The continuity follows if we write L as the composition of two continuous maps: the first one is $a \in L^2(\Omega) \mapsto r^* \in \mathbb{R} : h(r^*) = 0$ and the second one is $r \in \mathbb{R} \mapsto u_r \in H_0^1(\Omega)$. The continuity of the second map has been already mentioned. With respect to the first map, $a \in L^2(\Omega) \mapsto r^* \in \mathbb{R} : h(r^*) = 0$, we are going to obtain an easy formulation of this functional.

In fact,

$$h(r) = r + \int_\Omega (\Delta \hat{u}\, u_r).$$

Using that

$$\int_\Omega (\Delta \hat{u}\, u_r) = \int_\Omega (\hat{u} \Delta u_r)$$

and

$$-\Delta u_r = \frac{2b\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} r + \frac{\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} a,$$

we obtain that

$$h(r) = r \left(1 - \int_\Omega \frac{\hat{u} 2b\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} \right) - \int_\Omega \left(\frac{\hat{u}\Delta \hat{u}}{\hat{a} + b \int_\Omega |\nabla \hat{u}|^2} a \right).$$

Then

$$r^* = \frac{\int_\Omega \left(\frac{\hat{u}\Delta\hat{u}}{\hat{a}+b\int_\Omega |\nabla\hat{u}|^2} a \right)}{1 - \int_\Omega \frac{\hat{u}2b\Delta\hat{u}}{\hat{a}+b\int_\Omega |\nabla\hat{u}|^2}}$$

and the mapping $a \mapsto r^*$ is continuous. $\qquad\square$

We can apply the Theorem 3.3 in [18] to the characterization of the functional f in (3.1) and we obtain the Euler–Lagrange equation

$$-J'(\hat{u}, \hat{a}) + f_1 + f_2 = 0 \quad \text{in } H^{-1}(\Omega) \times (L^2(\Omega))'. \tag{3.7}$$

The functional f_1 belongs to the dual tangent cone associated to the constraint (2.2) at the point $(\hat{u}, \hat{a})$. It is known that f_1 satisfies

$$\langle f_1, (u, a) \rangle = 0 \quad \forall (u, a) \text{ satisfying } (3.3).$$

The functional $\tilde{f}_2 \in L^2(\Omega)'$ satisfies

$$\langle \tilde{f}_2, a - \hat{a} \rangle \geq 0 \quad \forall a \in \mathcal{U}. \tag{3.8}$$

If we choose any $(u, a) \in TC_1$ and we apply (3.7) to (u, a) it results

$$\int_\Omega (\hat{u} - u_d) u + B \int_\Omega \hat{a} a = \langle \tilde{f}_2, a \rangle \quad \forall (u, a) \in TC_1. \tag{3.9}$$

3.2. The adjoint problem

In order to write the necessary condition of the optimality as a optimality system, *i.e.*, as a system of partial differential equations together with a variational inequality (due to (2.3)), we will use the standard technique of getting the adjoint problem. The statement of the adjoint problem can be deduced as in the following way: Let us multiply (3.3) by a function $p \in H_0^1(\Omega)$, which will be the adjoint function and we will determine later,

$$\left\langle -\left(\hat{a} + b \int_\Omega |\nabla\hat{u}|^2\right) \Delta u, p \right\rangle - \left\langle \left(a + 2b \int_\Omega \nabla\hat{u} \cdot \nabla u\right) \Delta\hat{u}, p \right\rangle = 0.$$

These dual products are scalar products in $L^2(\Omega)$ because p is taken in $H_0^1(\Omega)$ and u, $\hat{u}$ are in $H_0^1(\Omega)$ too, besides we know by Theorem 2.2 that $\hat{u} \in W^{2,p}(\Omega)$ for any $p < +\infty$.

We integrate by parts to get

$$\left\langle u, -\Delta\left(\left(\hat{a} + b \int_\Omega |\nabla\hat{u}|^2\right) p\right) \right\rangle - \langle \Delta\hat{u}, ap \rangle - \langle \Delta\hat{u}, 2bp \rangle \int_\Omega \nabla\hat{u} \cdot \nabla u = 0.$$

We define the adjoint problem as

$$\begin{cases} -\Delta\left(\left(\hat{a} + b \int_\Omega |\nabla\hat{u}|^2\right) p\right) + \Delta\hat{u} \int_\Omega 2bp(\Delta\hat{u}) = \hat{u} - u_d \text{ in } \Omega, \\ p = 0 \hspace{7cm} \text{on } \partial\Omega. \end{cases} \tag{3.10}$$

This is a linear problem with a nonlocal term $\Delta\hat{u} \int_\Omega 2bp(\Delta\hat{u})$.

Proposition 3.5. *The adjoint problem* (3.10) *has a unique solution* p.

Proof. Proposition 3.5 is similar to Theorem 3.4. We consider the problem

$$\begin{cases} -\Delta\left(\left(\hat{a} + b \int_\Omega |\nabla\hat{u}|^2\right) p_r\right) = (-\Delta\hat{u})r + \hat{u} - u_d \text{ in } \Omega, \\ p_r = 0 \hspace{6cm} \text{on } \partial\Omega, \end{cases} \tag{3.11}$$

the function $h(r) = r - \int_\Omega 2b(\Delta\hat{u})p_r$ and we argue in the same fashion. $\qquad\square$

3.3. The optimality system

Taking into account that the adjoint problem we obtain that the first term on the left hand side of (3.9) is equal to $\langle \Delta \hat{u}, ap \rangle$, so the Euler–Lagrange equation is written by

$$\int_{\Omega} ap\Delta\hat{u} \,+\, B \int_{\Omega} \hat{a}a \;=\; \langle \tilde{f}_2, a \rangle \quad \forall (u,a) \in TC_1. \tag{3.12}$$

This equation defines $\tilde{f}_2$ on $L^2(\Omega)$ because by Theorem 3.4, for any $a \in L^2(\Omega)$ there exists a unique u such that $(u,a) \in TC_1$. So, we get a characterization of $\tilde{f}_2$, that is, $\tilde{f}_2 = p\Delta\hat{u} + B\hat{a}$.

By the condition (3.8), we have that

$$\langle p\Delta\hat{u} + B\hat{a}, a - \hat{a} \rangle \geq 0 \quad \forall a \in \mathcal{U},$$

but this is the scalar product in $L^2(\Omega)$, since $\Delta\hat{u} \in L^{\infty}(\Omega)$, so

$$(p\Delta\hat{u} + B\hat{a}, a - \hat{a}) \geq 0 \quad \forall a \in \mathcal{U},$$

which is equivalent to say that

$$\left(-\frac{p\Delta\hat{u}}{B} - \hat{a}, a - \hat{a} \right) \leq 0 \quad \forall a \in \mathcal{U},$$

i.e.

$$\hat{a} = P_{\mathcal{U}} \left(-\frac{p\Delta\hat{u}}{B} \right),$$

where P is the projection operator on $\mathcal{U}$ in $L^2(\Omega)$. The optimality system is given by

$$\begin{cases} -(\hat{a} + b\int_{\Omega}|\nabla\hat{u}|^2)\Delta\hat{u} \,=\, f \text{ in } \Omega, \quad \hat{u} = 0 \text{ on } \partial\Omega, \\ -\Delta\left(\left(\hat{a} + b\int_{\Omega}|\nabla\hat{u}|^2\right)p \right) + \Delta\hat{u}\int_{\Omega}2bp(\Delta\hat{u}) \,=\, \hat{u} - u_d \text{ in } \Omega, \quad p = 0 \text{ on } \partial\Omega, \\ \hat{a} = P_{\mathcal{U}}\left(-\frac{p\Delta\hat{u}}{B} \right). \end{cases} \tag{3.13}$$

The projection operator for the convex $\mathcal{U}$ is well characterized. It is

$$P_{\mathcal{U}}(f) = \max\{f, a_0\} \quad \forall f \in L^2(\Omega).$$

So,

$$\hat{a} = \begin{cases} a_0 & \text{if } -\frac{p\Delta\hat{u}}{B} < a_0 \\ -\frac{p\Delta\hat{u}}{B} & \text{in other case,} \end{cases}$$

i.e.

$$\hat{a} = \begin{cases} a_0 & \text{if } -p\Delta\hat{u} < Ba_0, \\ \frac{-p}{B}\Delta\hat{u} & \text{if } -p\Delta\hat{u} \geq Ba_0. \end{cases}$$

In other words,

$$a = \max\left\{ a_0, \frac{-p}{B}\Delta\hat{u} \right\}. \tag{3.14}$$

There is another formula for a which is deduced by (3.14) and which will fulfill a stopping criterion in next section. We replace $\Delta\hat{u}$ by $\frac{f}{\hat{a}+br}$, with $r = \int_{\Omega}|\nabla\hat{u}|^2$. Then,

$$a = \max\left\{ a_0, \frac{pf}{a + br} \right\};$$

so, we take

$$a = \max\{a_0(x), \beta(x)\},$$

being

$$\beta(x) = \frac{-b(x)r + \sqrt{b(x)^2 r^2 + \frac{4p(x)f(x)}{B}}}{2}, \tag{3.15}$$

the positive root of the equation $z^2 + b(x)rz - \frac{p(x)f(x)}{B} = 0$.

The optimality system is written by these formulae:

$$\begin{cases} -(\hat{a} + b \int_\Omega |\nabla \hat{u}|^2)\Delta \hat{u} \;=\; f \text{ in } \Omega, \;\; \hat{u} = 0 \text{ on } \partial\Omega, \\[2mm] -\Delta\left((\hat{a} + b \int_\Omega |\nabla \hat{u}|^2)\, p\right) + \Delta\hat{u} \int_\Omega 2bp(\Delta\hat{u}) \;=\; \hat{u} - \mathbf{u_d} \text{ in } \Omega, \;\; p = 0 \text{ on } \partial\Omega, \\[2mm] \hat{a} = \max\{a_0(x), \beta(x)\}, \;\; \beta(x) = \frac{-b(x)r + \sqrt{b(x)^2 r^2 + \frac{4p(x)f(x)}{B}}}{2}, \;\; \beta \geq 0. \end{cases} \tag{3.16}$$

Remark 3.6. We do not know if the solution of this system is unique.

4. The numerical approximation

In this section we are going to solve the optimality system in any dimension. As we will see, the key point is to solve different Poisson problems. This is very interesting because, in one hand, the algorithm is independent of the dimension and, on the other hand, there are no ill conditioned problems. The algorithm consists of the following three steps:

First step:
We choose an initial control $a \in L^2(\Omega)$ and solve the Kirchhoff problem for this a. If we define $r = \int_\Omega |\nabla u|^2$, this problem is written as following

$$\begin{cases} -\Delta u = \dfrac{f}{a + br} \text{ in } \Omega, \;\; u = 0 \text{ on } \partial\Omega, \\[2mm] r = \int_\Omega |\nabla u|^2. \end{cases}$$

Proposition 2.3 asserts us that the nonlinear equation

$$s = \int_\Omega \frac{f}{a + bs} u_s$$

has a unique real solution, denoted by r, which can be calculated by an algorithm to find roots of functions (for example, it can be used the Regula Falsi method). Then, we obtain u_r as the solution of the Poisson problem (2.4) with $s = r$.

As far as we know, this is a new way to solve numerically the Kirchhoff problem. Other works, [12], apply the Newton method and solve a linear system whose solution is an approximation of the discretized function u and the nonlocal term in the Kirchhoff equation, both of them are solved together. The disadvantage of this method is that the coefficient matrix of the linear system is ill conditioned in many cases. Other numerical works solve the Kirchhoff problem in time using the separation of variables (see [16]).

Second step:

We solve the adjoint problem. For it, we call s the nonlocal term:

$$s = \int_\Omega 2b(\Delta u)p. \tag{4.1}$$

Then, we define

$$\psi = (a + br)p - us, \tag{4.2}$$

which is the solution of the Poisson problem

$$\begin{cases} -\Delta\psi = u - u_d \text{ in } \Omega, \\ \psi = 0 \qquad \text{on } \partial\Omega. \end{cases} \tag{4.3}$$

We obtain an expression of p by (4.2)

$$p = \frac{\psi + us}{a + br}, \tag{4.4}$$

which is replaced in (4.1). Using that

$$\Delta u = -\frac{f}{a + br}$$

we obtain a new formula for s

$$s = \frac{-\int_\Omega \dfrac{2bf}{(a + br)^2}\psi}{1 + \int_\Omega \dfrac{2bf}{(a + br)^2}u}. \tag{4.5}$$

And, now we can calculate p by (4.4).

Thus, we have solved the Poisson problem (4.3), we have obtained s by (4.5) and, finally, we have got p by the formula (4.4).

Third step:

We take a new a in order to repeat the first and second step. This a is obtained as

$$a = \max\{a_0(x), \beta(x)\},$$

and $\beta(x)$ is the positive root of the equation $z^2 + b(x)rz - \frac{p(x)f(x)}{B} = 0$.

The algorithm finishes when the difference between the new a, and the previous one achieves a stop condition. We write the algorithm:

- Initialize a
- Until convergence:
  - Compute the state u given by the solution of the problem (2.2) for this a, compute the adjoint state p and compute β.
  - Compute $a^+ = P_{\mathcal{U}}(\beta)$.
  - Update $a = a^+$.

In the following example we solve the optimality system numerically for the following data: $u_d = 0$, the functions are $a_0 = 1$, $b = 1$ and $f(x) = \frac{2+\pi}{2}\sin(x)$, $\Omega = (0, \pi)$ and the differential equations are solved by finite differences.

We show the graphs of the optimal control, $\hat{a}$, and the optimal state, $\hat{u}$, in two cases: the first is for the parameter $B = 0.4$, which gives $\hat{a} = a_0$, and the second case is for $B = 0.0004$ which gives $\hat{a} = -p\Delta u/B$ in $(0, \pi)$ and $\hat{a} = a_0$ in the boundary of Ω.
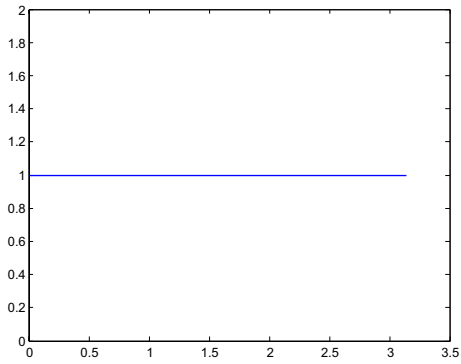
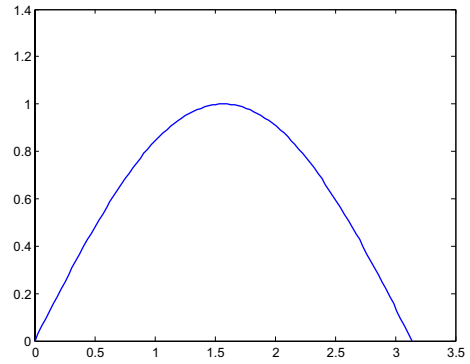FIGURE 1. Function $\hat{a}$ with $B = 0.4$.

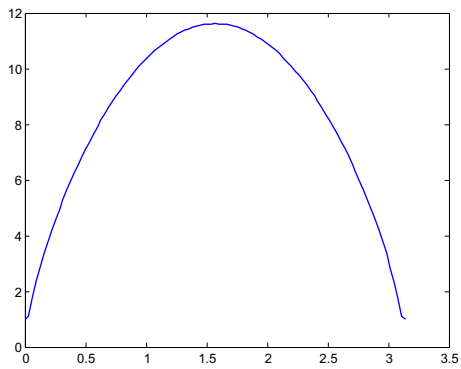FIGURE 2. Function $\hat{u}$ with $B = 0.4$.

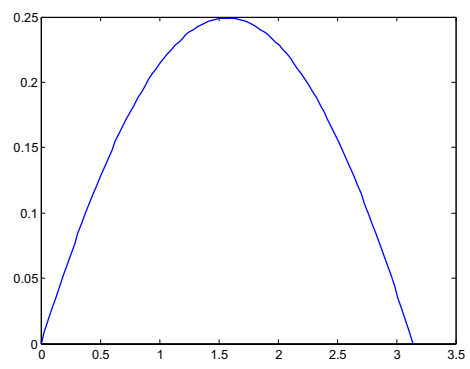

FIGURE 3. Function $\hat{a}$, $B = 0.0004$.

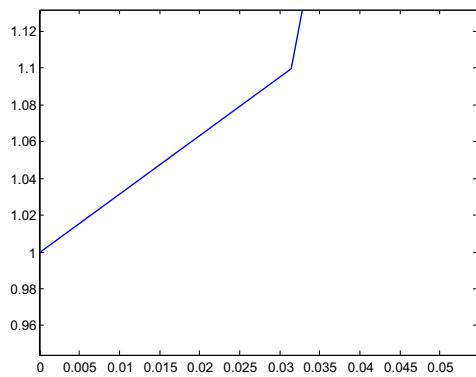FIGURE 4. Function $\hat{u}$, $B = 0.0004$.
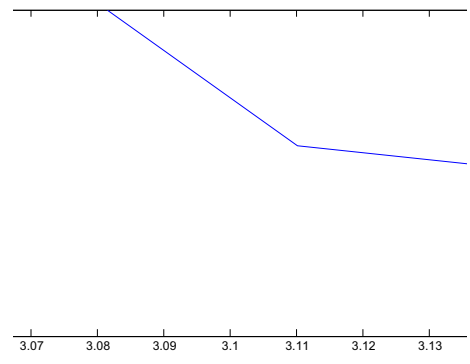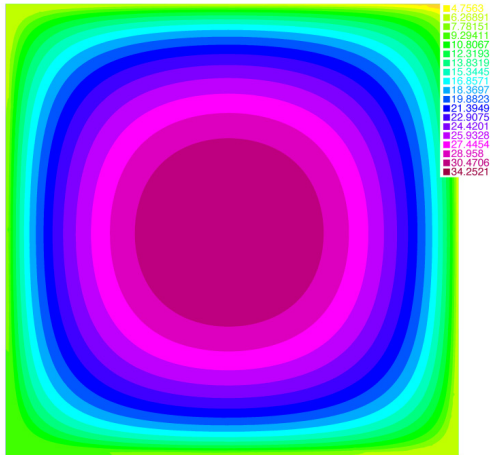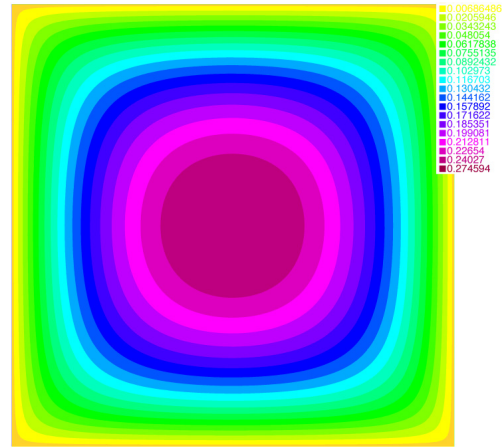

FIGURE 5. Function $\hat{a}$, $B = 0.0004$.

FIGURE 6. Function $\hat{a}$, $B = 0.0004$.

Figure 1 shows $\hat{a}$ when B is big. By (3.14) $\hat{a} = a_0$, and a_0 is a constant function, $a_0 = 1$. The associated state is provided in Figure 2.

When B is small, $\hat{a} = \frac{-p\Delta u}{B}$, as we can see in Figure 3. Nevertheless, on $\partial\Omega$ the function $\hat{a}$ is equal to a_0 because $p \in H_0^1(\Omega)$ and so, its trace is zero. Figure 4 shows the associated state. We can observe that it is smaller than the state drawn in Figure 2.

FIGURE 7. Function $\hat{a}$, $B = 0.00004$.

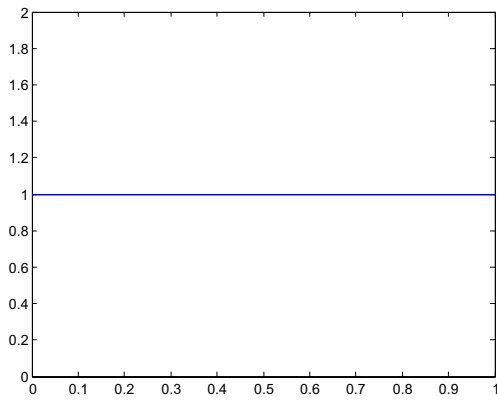FIGURE 8. Function $\hat{u}$, $B = 0.00004$.

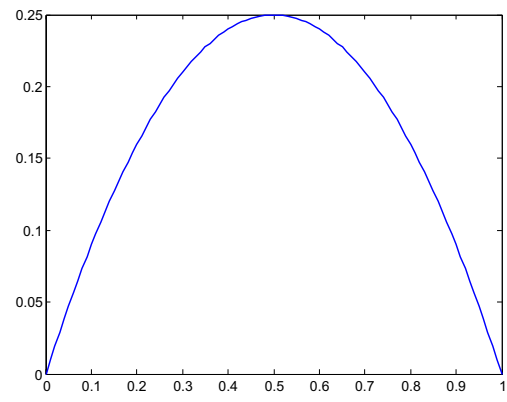FIGURE 9. Function $\hat{a}$, $B = 1$.
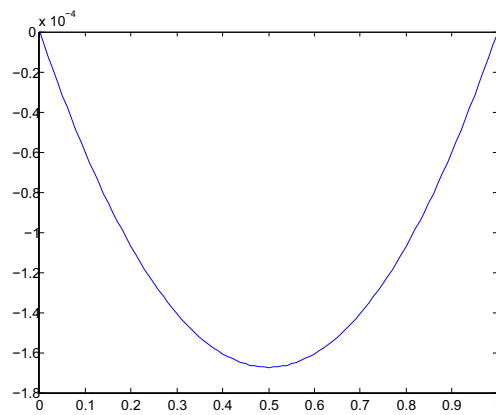

FIGURE 10. Function $\hat{u}$, $B = 1$.

FIGURE 11. Difference $\hat{u} - u_d$, $B = 1$.

We can see a detail of $\hat{a}$ with $B = 0.0004$:

Figures 5 and 6 provide a detail of $\hat{a}$ in both sides of $\partial\Omega$, $x = 0$, $x = \pi$. In this two points, $\hat{a}$ is equal to $a_0 = 1$, and for the rest of x, $\hat{a} = \frac{-p\Delta u}{B}$.

In a two dimensional case, we show the resolution with the data: $u_d = 0$, $a_0 = -3x - 3y + 10$, $b = x^2 + y^2$, $f(x) = 100$, $\Omega = (0, 1) \times (0, 1)$ and $B = 0.00004$. The partial differential equations are solved by P1 finite element (software FreeFem).

Figure 8 provides the values of $\hat{u}$. Since B is small we are prioritizing to minimize the term of the state in the functional J. In this case, we want $\hat{u}$ to be as near as it is possible to $u_d = 0$ (in the norm $L^2(\Omega)$). Nevertheless, Figure 7 shows that $\hat{a}$ is big.

Next, we show in an example how the exact solution of the optimal control problem given in (2.1)–(2.3) is well approximated by a solution of the optimality system (3.16). Let be $\Omega = (0, 1)$, $a_0 = 1$, $b = 1/2$, $B = 1$ and $u_d = -x(x-1)$. This function u_d verifies a Kirchhoff problem (2.2) with $f = 7/3$, so $u = u_d$ and $a = a_0$ minimize J and then, (u_d, a_0) is the solution of the optimal control problem. Although we do not know if the optimality system has a unique solution, the results we have got by the numerical resolution of the optimality system are those we should expect, (u_d, a_0). The following graphics are such solutions. Figure 9 provides $\hat{a}$, Figure 10 shows $\hat{u}$ and Figure 11 draws the difference between $\hat{u}$ and u_d. As we can see, this difference is about 1e-4.

References

[1] A. Arosio, On the nonlinear Timoshenko-Kirchhoff beam equation. *Chin. Annal. Math.* **20** (1999) 495–506.

[2] M. Chipot and J.F. Rodrigues, On a class of nonlocal nonlinear problems. *RAIRO Model. Math. Anal. Numer.* **26** (1992) 447–467.

[3] L.C. Evans, Partial Differential Equations. Vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Berkeley (1997).

[4] G.M. Figueiredo, C. Morales-Rodrigo, J.R. Santos Junior and A. Suárez, Study of a nonlinear Kirchhoff equation with non-homogeneous material. *J. Math. Anal. Appl.* **416** (2014) 597–608.

[5] G.B. Folland, Real Analysis. *Modern Techniques and Their Applications*. A Wiley-Interscience publication, New York (1984).

[6] I.V. Girsanov, Lectures on mathematical theory of extremum problems. Vol. 67 of *Lectures notes in Economics and mathematical systems*. Springer Verlag, Berlin (1972).

[7] T. Gudi, Finite element method for a nonlocal problem of Kirchhoff type. *SIAM J. Numer. Anal.* **50** (2012) 657–668.

[8] G. Kirchhoff, Mechanik. Teubner, Leipzig (1883).

[9] J.L. Lions, *Contrôle optimal de systèmes gouvernés par des Équations aux dérivées partielles*. Dunod, Paris (1968).

[10] J.L. Lions, On some questions in boundary value problems of Mathematical Physics, in *International Symposium on Continuum Mechanics and Partial Differential Equations, Rio de Janeiro 1977*. Vol. 30 of *Mathematical Studies* (1978) 284–346.

[11] H. Lou and J. Yong, Optimality conditions for semi linear elliptic equations with leading term containing control. *SIAM J. Control Optim.* **48** (2009) 2366–2387.

[12] T.F. Ma, Existence results and numerical solutions for a beam equation with nonlinear boundary conditions. *Appl. Numer. Math.* **47** (2003) 189–196.

[13] T.F. Ma, Remarks on a elliptic equation of Kirchhoff type. *Nonlin. Anal.* **63** (2005) 1967–1977.

[14] T.F. Ma and J.E. Muñoz Rivera, Positive solutions for a nonlinear nonlocal elliptic transmission problem. *Appl. Math. Lett.* **16** (2003) 243–248.

[15] R.G. Nascimento, *Problemas elípticos não locais do tipo p-Kirchhoff*. Doct. dissertation, Unicamp (2008).

[16] J. Peradze, A numerical algorithm for the nonlinear Kirchhoff string equation. *Numer. Math.* **102** (2005) 311–342.

[17] W. Rudin, Functional Analysis. Mc Graw-Hill, Inc. (1991).

[18] S. Walczak, Some properties of cones in normed spaces and their applications to investigating extremal problems. *J. Optim. Theory Appl.* **42** (1984) 561–582.