

## SOME APPLICATIONS OF OPTIMAL CONTROL THEORY OF DISTRIBUTED SYSTEMS\*

ALFREDO BERMUDEZ<sup>1</sup>

**Abstract.** In this paper we present some applications of the J.-L. Lions' optimal control theory to real life problems in engineering and environmental sciences. More precisely, we deal with the following three problems: sterilization of canned foods, optimal management of waste-water treatment plants and noise control

**Mathematics Subject Classification.** 49J20, 49K20, 49M37, 49B22.

Received January 31, 2001.

### INTRODUCTION

Optimal control theory of distributed parameter systems is a fundamental tool in applied mathematics. Since the pioneer book by J.-L. Lions [24] published in 1968 many papers have been devoted to both its theoretical aspects and its practical applications. The present article belongs to the latter set: we review some work related to real-life applications of optimal control theory to some engineering and environmental problems that have been done by Galician researchers.

The first one concerns food technology (see [5] for details). The problem is to determine the time dependent optimal temperature inside a sterilizing chamber for canned foods. The industrial sterilization of canned foods is a process in which there usually occurs a degradation of nutrients and a deterioration of qualitative properties due to the temperature to which the food is overexposed in order to assure destroying pathogenic microorganisms. We face this problem using optimal control methods. Linear constraints on the control and nonlinear ones on the state will appear as a consequence of the technological restrictions and the request for quality and economy. Besides an existence result for the optimal control and an optimality system we give numerical results for some real industrial examples. They show that present industrial sterilization processes can be improved in terms of saving energy and nutrient retention.

The second problem falls into environmental engineering. We deal with the design and management of waste-water treatment systems discharging polluting effluent into an aquatic medium. This problem is formulated as a pointwise optimal control problem with state and control constraints. The main difficulties arise from the lack of regularity of the source term in the state system and from the pointwise constraints on the state variables. We develop theoretical analysis of the problem, propose an algorithm for its numerical resolution and give results for a realistic problem in the *ría* of Vigo (Spain).

---

*Keywords and phrases:* Optimal control, sterilization, canned food, water pollution, noise reduction.

\* *This work has been partially financed by research project PGIDT00PXI20701PR Xunta de Galicia (Spain).*

<sup>1</sup> Departamento de Matemática Aplicada, Universidade de Santiago de Compostela, 15706 Santiago de Compostela, Spain;  
e-mail: [mabermud@usc.es](mailto:mabermud@usc.es)

The third problem is related to acoustic engineering, more precisely, to noise reduction by active control methods. During the last two decades there has been an accelerating level of interest in the control of sound by active techniques. Many of the physical principles involved have long been established, but the technological means for the successful implementation of “active noise control”, namely the fast digital signal processors(DSP) have only recently become available. Noise reduction is based on the destructive interference of waves. Accordingly, the control consists of a (secondary) source of noise the properties of which have to be determined in an optimal way in order to cancel the original undesired (primary) noise.

## 1. OPTIMAL STERILIZATION OF CANNED FOODS

### 1.1. Introduction

In order to prevent sanitary problems, canned foods undergo a thermal sterilization process. The goal is to kill pathogenic microorganisms by keeping temperature at some level during some period of time. The problem is that an excessive heating may lead to degradation of nutrients and deterioration of qualitative properties and, additionally, to an unnecessary energy cost. In order to fulfill the request of quality and presentation for canned foods it is very important to know the optimal profile of temperature in the sterilization chamber during the process. Existing industrial sterilization retorts usually work with steam at constant temperature during a certain time after which cans are cooled using water. This procedure does not seem to be optimal. Apparently one could increase the temperature at the beginning of the process and after a certain time decrease it in a smooth way.

The objective of this work is to give a general methodology to determine the suitable profile of temperature by using optimal control techniques. Some theoretical and numerical difficulties arise as a consequence of the existence of constraints on the control (steam temperature in the sterilization chamber) and on the state (canned food temperature) which the process must satisfy. Detailed proofs of the results enounced below can be found in [5].

### 1.2. Mathematical modelling and optimal control problem

The aim of the design of a thermal process for canned food sterilization lies in finding the optimal steam temperature  $v(t)$  in the sterilization chamber, at each time  $t$ , in order to reach a given microorganism reduction within the can at the end of the process, in such a way that energy cost be minimized and nutrient retention maximized. We assume that cans are subjected to a variable steam temperature, which will be taken as the control of the system, during a time interval  $(0, T)$ . We suppose that heat transfer inside them occurs only by conduction. The optimization of sterilization processes including natural convection has been studied in [1].

Let us consider a can occupying a three-dimensional domain  $\Omega$  in the retort. We denote by  $\Gamma$  its boundary. Then heat transfer can be modelled by the following initial-boundary value problem,

$$\rho(\theta)c(\theta)\frac{\partial\theta}{\partial t} - \nabla \cdot (k(\theta)\nabla\theta) = 0, \text{ in } Q = \Omega \times (0, T) \quad (1.1)$$

$$k(\theta)\frac{\partial\theta}{\partial n} = \alpha(v - \theta), \text{ on } \Sigma = \Gamma \times (0, T) \quad (1.2)$$

$$\theta(x, 0) = \theta_0(x), \text{ in } \Omega \quad (1.3)$$

where

- $\theta$  is temperature;
- $\rho$  is density;
- $c$  is heat capacity;
- $k$  is thermal conductivity;
- $\alpha$  is a heat transfer coefficient to be measured depending, in particular, on the tin can;
- $v$  is the retort temperature;

- $n$  is an outward unit normal vector;
- $\theta_0$  is the initial temperature.

Parameters  $\rho$ ,  $c$ ,  $k$  and  $\alpha$  actually depend on temperature. However, as a first approximation, they will be considered constant in the present paper.

In order to express the requirements of sterilization in a mathematical way it is necessary to know several parameters relative to microbial resistance and their dependence on temperature (see, for instance, Olin *et al.* [29]). The microorganism destruction is assumed to be a first order process because it has been observed that the number of living cells decays exponentially in time during the exposure to a lethal temperature. This fact can be expressed mathematically by the equation

$$\frac{dc}{dt} = -\mathcal{K}c, \quad (1.4)$$

where  $c(x, t)$  represents the living organism concentration at point  $x$  at time  $t$ . Parameter  $\mathcal{K}$  is a function of temperature through Arrhenius' equation:

$$\mathcal{K} = \mathcal{K}_r \exp\left(-\frac{E}{R}\left(\frac{1}{\theta} - \frac{1}{\theta_r}\right)\right), \quad (1.5)$$

where  $\mathcal{K}_r$  is the value of  $\mathcal{K}$  at a reference temperature  $\theta_r$ ,  $E$  the activation energy and  $R$  the universal gas constant. We consider the following parameters

- $D_r$ : necessary time at a constant temperature  $\theta_r$  to reduce the microorganism or spore population in 90%;
- $z$ : necessary rise in temperature in order to reduce the value of  $D_r$  in one order of magnitude.

Then, approximating  $\theta\theta_r$  by  $(\theta_r + z)\theta_r$ , equation (1.5) yields

$$\mathcal{K} = \mathcal{K}_r \exp\left(\frac{\ln 10}{z}(\theta - \theta_r)\right), \quad (1.6)$$

from which we deduce

$$\ln \frac{\mathcal{K}}{\mathcal{K}_r} = \frac{\ln 10}{z}(\theta - \theta_r). \quad (1.7)$$

Now equation (1.4) gives the concentration of microorganisms at final time  $T$  by

$$c(x, T) = c(x, 0) \exp\left(-\int_0^T \mathcal{K}_r 10^{\frac{\theta(x,t) - \theta_r}{z}} dt\right). \quad (1.8)$$

Therefore, if we want to reach a reduction in the order of  $10^\beta$ , the following inequality would be satisfied at each point  $x$  in the domain defined by the canned food,

$$\exp\left(-\int_0^T \mathcal{K}_r 10^{\frac{\theta(x,t) - \theta_r}{z}} dt\right) \leq 10^{-\beta}. \quad (1.9)$$

This state constraint can be written in a simpler way as follows

$$\int_0^T 10^{\frac{\theta(x,t) - \theta_r}{z}} dt \geq \frac{\beta \ln 10}{\mathcal{K}_r}. \quad (1.10)$$

Moreover we are interested in optimizing the global retention of some particular nutrient for this given reduction in microorganism concentration. If the nutrient retention is also assumed to be a first order process we are led to maximize the function

$$\int_{\Omega} \exp \left( - \int_0^T \tilde{\kappa}_r 10^{\frac{\theta(x,t) - \bar{\theta}_r}{z}} dt \right) dx \tag{1.11}$$

which represents the whole nutrient content of the canned food after sterilization. Notice that this function is determined by the control  $v(t)$  (the retort temperature at time  $t$ ) because  $\theta$  is a function of  $v$  through the partial differential equation (1.1, 1.2). Due to the fact that activation energies for the nutrients are generally much higher than for microorganisms, thermal sterilization of foods is possible without destroying all their nutrients. The optimal control problem we are interested in consists in finding the (variable in time) steam temperature minimizing the energy cost and maximizing the nutrient retention, while satisfying the constraint (1.10).

In what follows we study, from a mathematical point of view, an optimal control problem including the one defined in the previous section as a particular case. More precisely we show existence of solution and obtain optimality conditions which will allow us to compute the gradient of the cost function (1.11). This gradient will be used for numerical solution.

Let  $\Omega$  be an open bounded subset of  $R^N$  with smooth boundary  $\Gamma$ . As usual we denote by  $\bar{\Omega}$  the closure of  $\Omega$ . Let  $T > 0$  be the maximum time,  $Q = \Omega \times (0, T)$  and  $\Sigma = \Gamma \times (0, T)$ . Given strictly positive functions  $\rho, c, k \in C^\infty(\bar{\Omega})$ , a positive constant  $\alpha$  and a function  $\theta_0 \in C(\bar{\Omega})$ , the state of the system,  $\theta$ , is obtained from the control  $v \in L^\infty(0, T)$  by solving the following initial value problem,

$$\rho c \frac{\partial \theta}{\partial t} - \nabla \cdot (k \nabla \theta) = 0 \quad \text{in } Q \tag{1.12}$$

$$k \frac{\partial \theta}{\partial n} + \alpha \theta = \alpha v \quad \text{on } \Sigma \tag{1.13}$$

$$\theta(x, 0) = \theta_0(x) \quad \text{in } \Omega. \tag{1.14}$$

Let the cost function be defined by

$$J(v) = \nu \int_0^T \Phi(v) dt - \int_{\Omega} e^{-\int_0^T G(\theta) dt} dx. \tag{1.15}$$

We assume  $\Phi$  and  $G$  are continuously differentiable real functions and denote their respective derivatives by  $\Phi'$  and  $G'$ . Furthermore we suppose  $\Phi$  is convex and  $G$  positive. Parameter  $\nu$  is a non negative real number representing the relative weight of the two terms in the cost function  $J$ . In the sterilization problem  $\int_0^T \Phi(v) dt$  represents the energy cost. We impose some **constraints on control and state**. More precisely,  $v$  is taken in the closed convex subset of  $L^2(0, T)$  given by

$$U_{ad} = \{v \in L^2(0, T) : v_1 \leq v(t) \leq v_2 \text{ a.e. in } (0, T)\}, \tag{1.16}$$

where  $v_1$  and  $v_2$  are given constants such that  $v_1 < v_2$ . In the sterilization problem  $v_1$  and  $v_2$  are technological limits for the retort temperature. Furthermore we suppose the state  $\theta$  has to satisfy the following inequality,

$$a(x) \leq \int_0^T F(\theta(x, t)) dt \text{ for all } x \in \Omega, \tag{1.17}$$

where  $a \in C(\bar{\Omega})$  and  $F$  is a continuous function with continuous derivative denoted by  $F'$ . Then the optimal control problem consists of finding  $u$  minimizing  $J$  in the set of admissible controls:

$$C = \left\{ v \in U_{ad} : a(x) \leq \int_0^T F(\theta(x, t)) dt \text{ in } \Omega \right\}. \quad (1.18)$$

We are interested in

- **existence of optimal control;**
- **optimality conditions;**
- **numerical methods.**

Difficulties arise in obtaining optimality conditions due to pointwise constraint (1.17) and the fact that the set

$$\{z \in L^2(\Omega) : a(x) \leq z(x) \text{ a.e. in } \Omega\} \quad (1.19)$$

has an empty interior in  $L^2(\Omega)$ . Indeed, optimality conditions for constrained optimization problems are generally proved by using existence results of Lagrange multipliers (see, for instance, Ekeland and Temam [17]). These results cannot be directly applied to our case because we are in an infinite-dimensional space (the space  $L^2(\Omega)$ ) where the admissible set (1.19) has an empty interior. To overcome this difficulty, we replace this functional space by  $C(\bar{\Omega})$ . For this we need to prove that the state  $\theta \in C(\bar{Q})$ . Getting optimality conditions is very important from the numerical point of view because they allow us to compute the derivative of the cost with respect to the control. This derivative can be used in descent algorithms for numerical solution. Notice that, in the present case, we have to deal with a large non-convex constrained mathematical programming problem.

### 1.3. Mathematical analysis

The following existence result can be found in [5]. The proof uses classical techniques of minimizing sequences and a compactness result (see Lions [24] and [25]).

**Proposition 1.1.** *Assume there exists  $v \in U_{ad}$  such that the corresponding state satisfies the constraint (1.17). Then the optimal control problem has a solution.*

Next proposition yields an optimality system to be satisfied by a solution of the above defined optimal control problem. The proof is based on an abstract theorem of existence of Lagrange multipliers (see Casas [11] or Bonnans and Casas [9]). For any set  $C$ , let  $\chi_C$  be its indicator function defined by

$$\chi_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

We denote the sub-differential of a convex lower semi-continuous function  $f$  defined in a Banach space by  $\partial f$  (cf. Ekeland and Temam [17]). Recall that  $y \in V'$  belongs to  $\partial f(x)$  if and only if  $\langle y, z - x \rangle_{V', V} \leq f(z) - f(x)$  for every  $z$ . We denote by  $\theta(v)$  the solution of the state equation (1.12–1.14) corresponding to a control  $v$ . Then we define the nonlinear operator  $H$  and the set  $E$  by

$$H(\theta(v))(x) = \int_0^T F(\theta(x, t)) dt, \quad (1.20)$$

$$E = \{\omega \in C(\bar{\Omega}) : a(x) \leq \omega(x) \text{ in } \Omega\}. \quad (1.21)$$

We have the following:

**Proposition 1.2.** *Let  $u \in U_{ad}$  be a solution of the optimal control problem. Under the following assumptions*

1.  $F$  is strictly increasing;
2. there exists an interval  $(t_1, t_2) \subset (0, T)$  and  $\varepsilon > 0$  such that  $u(t) \leq v_2(t) - \varepsilon$  a.e. in  $(t_1, t_2)$ ;

then there exist functions  $\theta, p, g$  and  $q$  such that

$$\theta \in C(\bar{Q}) \quad (1.22)$$

$$p \in L^2(0, T; H^2(\Omega)), \quad \frac{\partial p}{\partial t} \in L^2(0, T; L^2(\Omega)), \quad p|_{\Sigma} \in H^{\frac{1}{2}}(\Sigma) \quad (1.23)$$

$$g \in L^1(Q), \quad g|_{\Sigma} \in L^1(\Sigma) \quad (1.24)$$

$$q \in C(\bar{\Omega})' \quad (1.25)$$

and satisfy

$$\rho c \frac{\partial \theta}{\partial t} - \nabla \cdot (k \nabla \theta) = 0 \text{ in } Q, \quad (1.26)$$

$$k \frac{\partial \theta}{\partial n} + \alpha \theta = \alpha u \text{ on } \Sigma, \quad (1.27)$$

$$\theta(x, 0) = \theta_0(x) \text{ in } \Omega \quad (1.28)$$

$$-\rho c \frac{\partial p}{\partial t} - \nabla \cdot (k \nabla p) = G'(\theta) e^{-\int_0^T G(\theta) dt} \text{ in } Q, \quad (1.29)$$

$$k \frac{\partial p}{\partial n} + \alpha p = 0 \text{ on } \Sigma, \quad (1.30)$$

$$p(x, T) = 0 \text{ in } \Omega \quad (1.31)$$

$$\left\langle g, \rho c \frac{\partial \varphi}{\partial t} - \nabla \cdot (k \nabla \varphi) \right\rangle_{Y'Y} = \langle (DH(\theta(u)))^* q, \varphi \rangle_{C(\bar{Q})'C(\bar{Q})}, \quad \forall \varphi \in \mathfrak{R}, \quad (1.32)$$

where  $Y = \{w \in C(\bar{Q}) : w|_{\Sigma} = 0\}$  and

$$\mathfrak{R} = \left\{ \psi \in H^{2,1}(Q); \rho c \frac{\partial \psi}{\partial t} - \nabla \cdot (k \nabla \psi) \in Y, k \frac{\partial \psi}{\partial n} + \alpha \psi = 0, \psi(0) = 0 \right\}, \quad (1.33)$$

$$q \in \partial \chi_E(H(\theta(u))), \quad (1.34)$$

$$\left\{ \int_0^T \nu \Phi'(u)(v - u) dt + \int_0^T \left( \int_{\Gamma} \alpha p d\sigma \right) (v - u) dt \right\} + \quad (1.35)$$

$$+ \int_0^T \left( \int_{\Gamma} \alpha g d\sigma \right) (v - u) dt \geq 0, \quad \forall v \in U_{ad}. \quad (1.36)$$

**Remark 1.1.** Function  $p + g$  is called *adjoint state* in optimal control theory. As a by-product of the previous Proposition we can deduce that

$$\nu \Phi'(u) + \int_{\Gamma} \alpha(p + g) d\sigma,$$

which is a function of time, is a descent direction for the “state constrained” cost function. This fact will be used for numerical solution in the next paragraph.

#### 1.4. Numerical methods and results

For numerical solution, the state equations (1.12–1.14) are discretized using an implicit finite difference scheme in time (backward Euler) and continuous piecewise linear finite elements on a triangular mesh of  $\Omega$ . The solution of the discretized optimal control problem has been done by using a Newton method with an improvement simplifying the computation of the Hessian due to Bonnans [13]. Further details about computer implementation can be found in [5].

As a test example, let us consider a cylindrical can of tuna with 6.52 cm diameter and 3 cm height. The “energy cost” function has been defined by  $\Phi(v) = v^2$  and the following values of the parameters have been taken from Perez *et al.* [30],

$$\frac{k}{\rho c} = 1.2910^{-3} \frac{\text{cm}^2}{\text{s}}, \quad (1.37)$$

$$z = 10, \quad D_r = 0.21 \text{ min}, \quad \theta_r = 121.1 \text{ }^\circ\text{C}, \quad (1.38)$$

$$\tilde{z} = 25.56, \quad \tilde{D}_r = 178.6 \text{ min}, \quad \tilde{\theta}_r = 121.1 \text{ }^\circ\text{C}, \quad (1.39)$$

$$v_1 = 20 \text{ }^\circ\text{C}, \quad v_2 = 115 \text{ }^\circ\text{C} \quad (1.40)$$

$$\theta_0 = 15 \text{ }^\circ\text{C}, \quad \alpha = 90 \frac{\text{W}}{\text{cm}^2 \text{ }^\circ\text{C}}. \quad (1.41)$$

Due to symmetry, cylindrical coordinates are used. In Table 1 we show results corresponding to several numerical tests associated with different values of the process time (Col. 1), the upper bound of the steam temperature (Col. 2) and the  $F_c$  parameter (Col. 3) given by

$$F_c = \frac{\beta \ln 10}{\mathcal{K}_r}.$$

The fourth and fifth columns give respectively the “energy cost”  $\int_0^T \Phi(v)$  and the nutrient retention rate as given by (1.11). Finally Figures 1 and 2 show the optimal control (see the rightmost column of Tab. 1). These tests show how the industrial process, which usually works with steam at constant temperature during a certain time, is not the most adequate in terms of quality factor and energy cost. Indeed, optimal profiles of steam temperature, calculated for the same values of  $F_c$  and control bounds, give lower cost values and greater nutrient retention. In any case the values of  $F_c$  used in the industrial sterilization seem to be greater than necessary to ensure sterilization. In particular, the European Community (EC) rules require a value for  $F_c$  equal to 180 seconds for  $\theta_r = 121.1 \text{ }^\circ\text{C}$  (250  $^\circ\text{F}$ ), notably lower than that corresponding to industrial sterilization. Indeed, two very common industrial processes are,

1. 115  $^\circ\text{C}$  during 3600 s;
2. 110  $^\circ\text{C}$  during 7200 s.

If we compute the corresponding temperature inside a can by solving the state equation we obtain for the function in the left-hand side of (1.10) the values 268 s and 330 s, respectively, which are much greater than those prescribed by the EC rules. On the other hand the corresponding values of nutrient retention are 72.93% and 62.10%, respectively, which are lower than the optimal ones in Table 1.

## 2. AN OPTIMAL CONTROL PROBLEM RELATED TO MANAGEMENT OF WASTE-WATER TREATMENT SYSTEMS

### 2.1. Description of the problem

Let us consider a domain  $\Omega$  with boundary  $\Gamma$  occupied by shallow waters where polluting waste-waters are discharged through  $N_E$  outfalls each corresponding to a particular depuration plant. We also assume that inside

TABLE 1. Test results.

Process time	$v_2$	$F_c$	$(\sum u_i^2)^{\frac{1}{2}}$	Nutrient retention	Figures
3600 s	115 °C	180 s	338.8614	77.35%	–
3600 s	119 °C	268 s	329.0677	74.06%	1
7200 s	115 °C	180 s	256.1834	75.89%	–
7200 s	115 °C	330 s	280.5771	69.41%	2

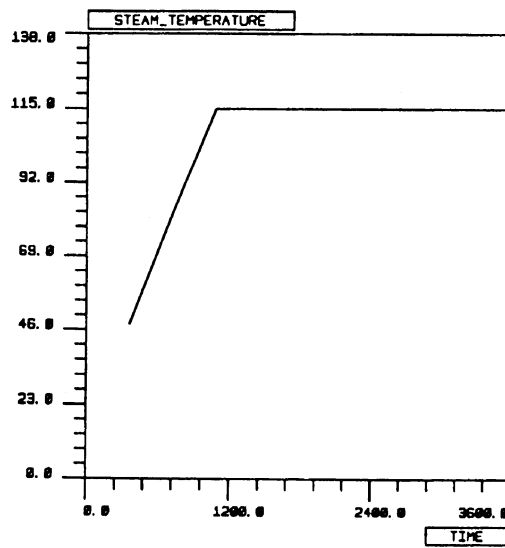


FIGURE 1. Optimal steam temperature in the sterilization chamber.

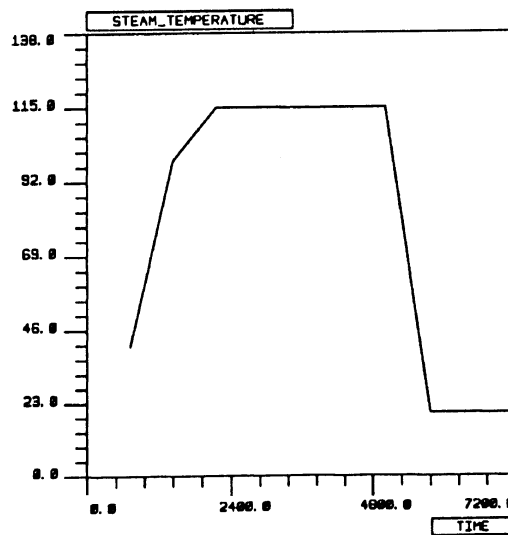


FIGURE 2. Optimal steam temperature in the sterilization chamber.



the domain  $\Omega$  there exist several protected areas  $A_i \subset \Omega$ ,  $i = 1, \dots, N_Z$  (bath, marine cultures, etc.) where it is necessary to assure water quality, *i.e.* pollution concentration must be lower than a given maximum level.

We recall that, in order to control the marine pollution, some parameters are used which indicate water quality and its capacity to keep the aquatic life. Among these indicators we mention dissolved oxygen, temperature, pH, etc. Oxygen is used by bacteria to decompose organic matter. This process can be measured in terms of the need of oxygen to do this task, the so called *Biological Oxygen Demand* (BOD). If the pollution level is not too high this need can be satisfied by the *Dissolved Oxygen* (DO). However, if the organic matter increases beyond a maximum value the DO is not enough for its decomposition, leading to important modifications (anaerobic processes) in the ecosystem. Hence, in each protected area  $A_i$ ,  $i = 1, \dots, N_Z$ , a threshold value  $\sigma_i$  of BOD may not be exceeded and a minimum level of DO  $\zeta_i$  must be guaranteed. This can be done by some chemical or biological treatments before discharging waste-waters into the sea which can be viewed as control mechanisms.

The cost of depuration at each plant is assumed to depend on the BOD of arriving waste-waters and some other features as technology, age, etc. Then, for a particular plant, the cost is a decreasing function of the BOD after treatment because getting a lower level of BOD leads to a more intensive depuration and, consequently, to a higher cost. Keeping in mind that absolute depuration is not feasible and that there exists a minimum fixed cost (even in the case where no treatment is needed), the cost function of the  $j$ -th treatment plant,  $f_j$ , takes a form similar to the one shown in Figure 5.

Then the problem is to determine, at each plant and for each time, the BOD level of discharges after treatment minimizing the global depuration cost while respecting the above mentioned constraints on the water quality in protected areas. From the mathematical point of view, this is an optimal control problem for a parabolic partial differential equation with pointwise state constraints and pointwise control. Indeed, discharges from each plant are made through outfalls located at some particular given points of the domain. Optimization of this location can be considered as a geometric optimal control problem. This has been done in [6].

### 2.2. Mathematical modelling. Optimal control problem

Let us suppose that the outfalls are located at points  $P_j \in \Omega$ ,  $j = 1, \dots, N_E$  and denote by  $m_j(t)$ ,  $j = 1, \dots, N_E$ , the rate of BOD discharged at point  $P_j$  at time  $t$ . According to the distributed parameter version of the Streeter–Phelps model, the evolution of the BOD and the DO in the domain  $\Omega \subset \mathbb{R}^2$  is governed by a system of partial differential equations (*cf.* [4,27,33]). Let us denote by  $\rho_1(x, t)$  and  $\rho_2(x, t)$  the concentrations of BOD and DO at a point  $x \in \Omega$  and a time  $t \in (0, T)$ , respectively. Then they are solution of the following boundary value problem:

$$\left. \begin{aligned} \frac{\partial \rho_1}{\partial t} + \vec{u} \nabla \rho_1 - \beta_1 \Delta \rho_1 &= -\kappa_1 \rho_1 + \frac{1}{h} \sum_{j=1}^{N_E} m_j \delta(x - P_j) && \text{in } \Omega \times (0, T) \\ \frac{\partial \rho_1}{\partial n} &= 0 && \text{on } \Gamma \times (0, T) \\ \rho_1(x, 0) &= 0 && \text{in } \Omega \end{aligned} \right\} \quad (2.42)$$

$$\left. \begin{aligned} \frac{\partial \rho_2}{\partial t} + \vec{u} \nabla \rho_2 - \beta_2 \Delta \rho_2 &= -\kappa_1 \rho_1 + \frac{1}{h} \kappa_2 (d_s - \rho_2) && \text{in } \Omega \times (0, T) \\ \frac{\partial \rho_2}{\partial n} &= 0 && \text{on } \Gamma \times (0, T) \\ \rho_2(x, 0) &= \rho_{20}(x) && \text{in } \Omega \end{aligned} \right\}$$

where  $h(x, t)$  and  $\vec{u}(x, t)$  are, respectively, the height and the mean horizontal velocity of the fluid layer, obtained by solving the Saint Venant (or shallow water) equations (see for instance [7]),  $\delta(x - P_j)$  represents the Dirac measure at point  $P_j$  and parameters  $\beta_1 > 0$  and  $\beta_2 > 0$  are horizontal diffusion coefficients involving dispersion and turbulence effects. Parameter  $\kappa_1 > 0$  is a kinetic coefficient, parameter  $\kappa_2 > 0$  is an oxygen transfer coefficient through the sea surface and  $d_s$  is the oxygen saturation density.

Let us assume now that there are  $N_Z$  protected areas,  $A_i$ ,  $i = 1, \dots, N_Z$ , where a maximum level of BOD and a minimum level of DO must be assured, namely,

$$\rho_1|_{A_i \times (0, T)} \leq \sigma_i, \quad i = 1, \dots, N_Z, \quad (2.43)$$

$$\rho_2|_{A_i \times (0, T)} \geq \zeta_i, \quad i = 1, \dots, N_Z. \quad (2.44)$$

Let  $f_j$ , the treatment cost at the  $j$ -th plant,  $j = 1, \dots, N_E$ , be given in  $C^2(0, \infty)$ . Then the total cost of the whole depuration system during a time interval  $[0, T]$  is given by

$$J(m) = \sum_{j=1}^{N_E} \int_0^T f_j(m_j(t)) dt. \quad (2.45)$$

The problem ( $\mathcal{P}$ ) of optimal management of the depuration system consists in finding the values of BOD  $m_j(t) > 0$ ,  $j = 1, \dots, N_E$ , along the time interval  $[0, T]$ , minimizing the cost function (2.45) under the state constraints (2.43) and (2.44).

### 2.3. Existence of solution of the state equation

**Definition 2.1.** Given  $r, s \in [1, 2)$ ,  $\frac{2}{r} + \frac{2}{s} > 3$ , we say that  $\rho = (\rho_1, \rho_2) \in [L^r(0, T; W^{1,s}(\Omega))]^2$ , is a solution of the system (2.42) if for all  $\Phi = (\Phi_1, \Phi_2) \in [L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))]^2 \cap [C^1(\bar{\Omega} \times [0, T])]^2$  such that  $\Phi(\cdot, T) = 0$ , the following equality holds,

$$\begin{aligned} & \int_0^T \int_{\Omega} \left\{ -\frac{\partial \Phi_1}{\partial t} \rho_1 - \frac{\partial \Phi_2}{\partial t} \rho_2 + \beta_1 \nabla \Phi_1 \nabla \rho_1 + \beta_2 \nabla \Phi_2 \nabla \rho_2 + \bar{u} \Phi_1 \nabla \rho_1 \right. \\ & \quad \left. + \bar{u} \Phi_2 \nabla \rho_2 + \kappa_1 \Phi_1 \rho_1 + \kappa_1 \Phi_2 \rho_1 + \frac{1}{h(x, t)} \kappa_2 \Phi_2 \rho_2 \right\} dx dt \\ & = \sum_{j=1}^{N_E} \int_0^T \frac{1}{h(P_j, t)} \Phi_1(P_j, t) m_j(t) dt \\ & \quad + \int_0^T \int_{\Omega} \frac{1}{h(x, t)} \kappa_2 d_s \Phi_2(x, t) dx dt + \int_{\Omega} \Phi_2(x, 0) \rho_{20}(x) dx. \end{aligned} \quad (2.46)$$

Let  $\mathcal{A}$  be the operator defined by

$$\begin{aligned} \langle \mathcal{A}(w_1, w_2), (z_1, z_2) \rangle & = \int_{\Omega} \left( -\beta_1 \Delta w_1 z_1 - \beta_2 \Delta w_2 z_2 \right. \\ & \quad \left. + \bar{u} \nabla w_1 z_1 + \bar{u} \nabla w_2 z_2 + \kappa_1 w_1 z_1 + \kappa_1 w_1 z_2 + \frac{1}{h} \kappa_2 w_2 z_2 \right) dx, \end{aligned}$$

for  $(w_1, w_2), (z_1, z_2)$  such that the previous expression makes sense. Then we have the following result:

**Theorem 2.1.** *There exists a unique pair*

$$\rho = (\rho_1, \rho_2) \in [L^r(0, T; W^{1,s}(\Omega))]^2 \cap [L^2(0, T; L^2(\Omega))]^2,$$

with

$$\frac{\partial \rho}{\partial t} = \left( \frac{\partial \rho_1}{\partial t}, \frac{\partial \rho_2}{\partial t} \right) \in [L^r(0, T; (W^{1,s'}(\Omega))')]^2$$

for all  $r, s \in [1, 2)$ ,  $\frac{2}{r} + \frac{2}{s} > 3$ , such that  $\rho$  is solution of (2.42) and verifies

$$\int_0^T \left\langle -\frac{\partial \Phi}{\partial t} + \mathcal{A}^*(\Phi), \rho \right\rangle dt = \sum_{j=1}^{N_E} \int_0^T \frac{1}{h(P_j, t)} \Phi_1(P_j, t) m_j(t) dt + \int_0^T \int_{\Omega} \frac{1}{h(x, t)} \kappa_2 d_s \Phi_2(x, t) dx dt + \int_{\Omega} \Phi_2(x, 0) \rho_{20}(x) dx, \tag{2.47}$$

for all  $\Phi = (\Phi_1, \Phi_2) \in \mathcal{B}$ , where:

$$\mathcal{B} = \left\{ \Phi = (\Phi_1, \Phi_2) \in [L^2(0, T; H^2(\Omega)) \cap H^1(0, T; L^2(\Omega))]^2 : -\frac{\partial \Phi}{\partial t} + \mathcal{A}^*(\Phi) \in [L^2(0, T; L^2(\Omega))]^2, \frac{\partial \Phi}{\partial n_{\mathcal{A}^*}}|_{\Gamma \times (0, T)} = 0, \Phi(\cdot, T) = 0 \right\}.$$

Furthermore, there exist constants  $C_k$ ,  $k = 1, \dots, 6$ , such that

$$\|\rho\|_{[L^r(0, T; W^{1, s}(\Omega))]^2} \leq C_1 \sum_{i=1}^{N_E} \|m_i\|_{L^\infty(0, T)} + C_2 \|\rho_{20}\|_{C(\bar{\Omega})} + C_3 d_s. \tag{2.48}$$

and

$$\|\rho\|_{[L^2(0, T; L^2(\Omega))]^2} \leq C_4 \sum_{i=1}^{N_E} \|m_i\|_{L^\infty(0, T)} + C_5 \|\rho_{20}\|_{C(\bar{\Omega})} + C_6 d_s. \tag{2.49}$$

The proof of this theorem can be seen in [28] as well as the following regularity and continuity results:

**Lemma 2.1.** Functions  $\rho_1$  and  $\rho_2$  are continuous in  $\bar{A}_i \times [0, T]$ ,  $\forall i = 1, \dots, N_Z$ .

**Lemma 2.2.** There exist constants  $\hat{C}_1, \hat{C}_2, \hat{C}_3$  such that:

$$\|\rho\|_{[C(\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T])]^2} \leq \hat{C}_1 \sum_{i=1}^{N_E} \|m_i\|_{L^\infty(0, T)} + \hat{C}_2 \|\rho_{20}\|_{C(\bar{\Omega})} + \hat{C}_3 d_s.$$

**2.4. Existence of control. Optimality system.**

**Theorem 2.2.** If there exists a feasible control  $\tilde{m} \in U_{ad}$  such that:

$$\begin{aligned} \tilde{\rho}_1|_{A_i \times (0, T)} &\leq \sigma_i, & i = 1, \dots, N_Z \\ \tilde{\rho}_2|_{A_i \times (0, T)} &\geq \zeta_i, & i = 1, \dots, N_Z \end{aligned}$$

then the optimal control problem has, at least, a solution.

*Proof.* See [28]. □

Now we are interested in writing an optimality system. Firstly, let us consider the mappings

$$\begin{aligned} F_1 : m \in (L^\infty(0, T))^{N_E} &\longrightarrow F_1(m) = \rho_1|_{\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]} \in C(\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]), \\ F_2 : m \in (L^\infty(0, T))^{N_E} &\longrightarrow F_2(m) = \rho_2|_{\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]} \in C(\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]). \end{aligned}$$

We have the

**Lemma 2.3.** *The mappings  $F_1$  y  $F_2$  are Gâteaux derivable. Moreover*

$$DF_1(m)(n) = \omega_1|_{\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]},$$

$$DF_2(m)(n) = \omega_2|_{\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]},$$

where  $\omega_1$  and  $\omega_2$  are the solution of the linearized system

$$\left. \begin{aligned} \frac{\partial \omega_1}{\partial t} + \bar{u} \nabla \omega_1 - \beta_1 \Delta \omega_1 &= -\kappa_1 \omega_1 + \frac{1}{h} \sum_{j=1}^{N_E} n_j \delta(x - P_j) && \text{in } \Omega \times (0, T) \\ \frac{\partial \omega_1}{\partial n} &= 0 && \text{on } \Gamma \times (0, T) \\ \omega_1(x, 0) &= 0 && \text{in } \Omega \\ \frac{\partial \omega_2}{\partial t} + \bar{u} \nabla \omega_2 - \beta_2 \Delta \omega_2 &= -\kappa_1 \omega_1 - \frac{1}{h} \kappa_2 \omega_2 && \text{in } \Omega \times (0, T) \\ \frac{\partial \omega_2}{\partial n} &= 0 && \text{on } \Gamma \times (0, T) \\ \omega_2(x, 0) &= 0 && \text{in } \Omega \end{aligned} \right\}. \quad (2.50)$$

From this lemma one can obtain a first order optimality system satisfied by any solution of the optimal control problem. In order to express them in a simple way we introduce functions  $p_1, p_2$  solutions (in the sense of Def. 2 below) of the following boundary value problems,

$$\left. \begin{aligned} -\frac{\partial p_1}{\partial t} - \nabla \cdot (\bar{u} p_1) - \beta_1 \Delta p_1 + \kappa_1 (p_1 + p_2) &= \mu_1|_{\Omega \times (0, T)} && \text{in } \Omega \times (0, T) \\ \beta_1 \frac{\partial p_1}{\partial n} + \bar{u} \cdot \bar{n} p_1 &= 0 && \text{on } \Gamma \times (0, T) \\ p_1(x, T) &= \mu_1|_{\Omega \times \{T\}} && \text{in } \Omega \\ -\frac{\partial p_2}{\partial t} - \nabla \cdot (\bar{u} p_2) - \beta_2 \Delta p_2 + \frac{1}{h} \kappa_2 p_2 &= \mu_2|_{\Omega \times (0, T)} && \text{in } \Omega \times (0, T) \\ \beta_2 \frac{\partial p_2}{\partial n} + \bar{u} \cdot \bar{n} p_2 &= 0 && \text{on } \Gamma \times (0, T) \\ p_2(x, T) &= \mu_2|_{\Omega \times \{T\}} && \text{in } \Omega \end{aligned} \right\} \quad (2.51)$$

where  $\mu_1, \mu_2$  are regular Borel measures in  $\bar{\Omega} \times [0, T]$ . The weak solution of the system (2.51) can be defined by transposition techniques (see Casas [12]):

**Definition 2.** Given  $r, s \in [1, 2)$ ,  $\frac{2}{r} + \frac{2}{s} > 3$ , we say that  $p = (p_1, p_2) \in [L^r(0, T; W^{1,s}(\Omega))]^2$  is a solution of the system (2.51) if, for all  $z = (z_1, z_2) \in [L^2(0, T; H^1(\Omega)) \cap C^1(\bar{\Omega} \times [0, T])]^2$  such that  $z(\cdot, 0) = 0$ , the following equality holds,

$$\begin{aligned} \int_0^T \int_{\Omega} \left\{ \frac{\partial z_1}{\partial t} p_1 + \frac{\partial z_2}{\partial t} p_2 + \rho_1 \nabla z_1 \nabla p_1 + \rho_2 \nabla z_2 \nabla p_2 + \bar{u} \nabla z_1 p_1 + \bar{u} \nabla z_2 p_2 \right. \\ \left. + \kappa_1 z_1 p_1 + \kappa_1 z_1 p_2 + \frac{1}{h(x, t)} \kappa_2 z_2 p_2 \right\} dx dt = \int_0^T \int_{\Omega} z_1 d\mu_1(x, t) \\ + \int_0^T \int_{\Omega} z_1 d\mu_1(x, t) + \int_{\Omega} z_1(x, T) d\mu_{1T}(x) + \int_{\Omega} z_2(x, T) d\mu_{2T}(x). \end{aligned}$$

Let us define the sets  $S_1$  and  $S_2$  by

$$S_1 = \{s \in C(\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]) : s(x, t) \leq \sigma_j, \forall (x, t) \in \bar{A}_j \times [0, T], j = 1, \dots, N_Z\},$$

$$S_2 = \{\omega \in C(\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]) : \omega(x, t) \geq \zeta_j, \forall (x, t) \in \bar{A}_j \times [0, T], j = 1, \dots, N_Z\}.$$

We have the following result:

**Theorem 2.3.** *Let  $m \in U_{ad}$  be an optimal control. Then, there exist two functions  $\rho_1, \rho_2 \in L^r(0, T; W^{1,s}(\Omega)) \cap L^2(0, T; L^2(\Omega))$ , for all  $r, s \in [1, 2)$ ,  $\frac{2}{r} + \frac{2}{s} > 3$ , solution of (2.42) and two functions  $p_1, p_2 \in L^r(0, T; W^{1,s}(\Omega))$ , solution of (2.51), where  $\mu_1, \mu_2$  are two Borel measures, with support in  $\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]$ , such that,*

$$\mu_i|_{\cup_{i=1}^{N_Z} \bar{A}_i \times [0, T]} \in \partial \chi_{S_i}(F_i(m)), \quad i = 1, 2, \tag{2.52}$$

and furthermore

$$\sum_{j=1}^{N_E} \left\{ \int_0^T f_j'(m_j(t))(n_j(t) - m_j(t)) dt + \int_0^T \frac{1}{h(P_j, t)} p_1(P_j, t)(n_j(t) - m_j(t)) dt \right\} \geq 0, \quad \forall n \in U_{ad}. \tag{2.53}$$

### 2.5. Numerical solution

The first step in the numerical resolution of the control problem is solving the state system. For this purpose we carry out a time discretization, treating the convective term with a method of characteristics. For the time interval  $[0, T]$ , we choose a number  $N \in \mathbb{N}$  and define  $\Delta t = \frac{T}{N} > 0$  and  $t_n = n\Delta t, \forall n = 0, \dots, N$ . Then, for the semi-discretized problem, we consider a variational formulation and approximate it by a finite element discretization. We introduce a polygonal set  $\Omega_h$  approximating the domain  $\Omega$  and choose an admissible triangulation  $\tau_h$  of it (see [15]) with triangles of diameter  $\leq h$  and such that the vertices in the boundary of  $\Omega_h$  also lie in the boundary of  $\Omega$ . We define the following finite element space,

$$V_h = \{v_h \in C^0(\bar{\Omega}), v_h|_K \in P_1, K \in \tau_h\}.$$

Then the discretized problem is solved making use of the velocity and height of water column provided by Saint Venant equations. Moreover the discretized cost  $\hat{J}$  and constraints  $g$  are given by

$$\begin{aligned} \hat{J} : \mathbb{R}^{N \times N_E} &\longrightarrow \mathbb{R} \\ m &\longrightarrow \hat{J}(m) = \Delta t \sum_{j=1}^{N_E} \sum_{n=0}^{N-1} C_{jn} f_j(m_{jn}), \\ g : \mathbb{R}^{N \times N_E} &\longrightarrow \mathbb{R}^{N \times N_{VZ}} \times \mathbb{R}^{N \times N_{VZ}} \times \mathbb{R}^{N \times N_E} \\ m &\longrightarrow g(m) = (\underbrace{\tilde{\rho}_{1h} - \sigma, \zeta - \tilde{\rho}_{2h}}_{= g_1(m)}, \underbrace{\underline{m} - m, m - \bar{m}}_{= g_2(m)})^t, \end{aligned}$$

where

- $m$  is the vector consisting of all of the discharges at all times;
- $m_{jn}$  is the amount of BOD discharged in  $P_j$  at time  $t_n$ ;
- $C_{jn}$  are the weights of the quadrature formula;

- $N_{VZ}$  is the number of vertices in the protected areas;
- $\tilde{\rho}_{ih}$  is a vector of values of  $\rho_{ih}$  at each vertex included in the protected areas and for all times.

We remark that function  $g$  can be decomposed into  $g_1$ , which is putting together the constraints about the water quality, and  $g_2$  collecting the bound constraints. Then, the optimal control problem is approximated by the following discretized problem,

$$(\mathcal{P}_{\mathcal{F}}) \left\{ \begin{array}{l} \min_{m \in \mathbb{R}^{N \times N_E}} \hat{J}(m) \\ \text{such that } g(m) \leq 0. \end{array} \right.$$

This problem has solution because the cost function is convex and the constraints set is compact. It has been solved by using a feasible points method which is based on a globally convergent algorithm introduced by Herskovits [19] and Panier *et al.* [31]. Let us introduce the following notation,

- $p$  is the dimension of the control and  $q$  is the number of the constraints on the state;
- $(\lambda, \theta)$  is the vector of the dual variables;
- $L(m, \lambda, \theta)$  is the Lagrangian;
- $H(m, \lambda, \theta)$  is the Hessian.

Then the first order Karush–Kuhn–Tucker optimality conditions for the discretized problem are,

$$\nabla \hat{J}(m) + \nabla g_1 \lambda - I \theta = 0, \quad (2.54)$$

$$G_1(m) \lambda = 0, \quad G_2(m) \theta = 0, \quad (2.55)$$

$$\lambda \geq 0, \quad \theta \geq 0, \quad (2.56)$$

$$g_1(m) \leq 0, \quad g_2(m) \leq 0, \quad (2.57)$$

where  $G_1(m)$  and  $G_2(m)$  are diagonal matrices of order  $q$  and  $p$ , respectively, with diagonal elements being the values of the corresponding functions  $g_i(m)$ . The basic idea of the feasible points algorithm consists of solving the system of equations (2.54, 2.55) in  $(m, \lambda, \theta)$  by using a fixed point method, in such a way that conditions (2.56, 2.57) hold at each iteration. We refer to [28] for further details.

## 2.6. Numerical results

In this section we present the numerical results obtained when solving the problem for a realistic situation. We have taken a two-dimensional mesh of the *ría* of Vigo (Spain) as domain  $\Omega$ , where we have considered two protected areas and two points of discharge (see Fig. 3). We have also supposed that it is necessary to guarantee lower level of pollution in zone 1 than in zone 2. The values of the physical parameters and data can be seen in Table 2.

The cost function is given in Figure 5. We assume that pollutant concentration of waste-waters arriving to the sewage farm is  $150 \text{ Kg/m}^3$  so the depuration cost above this value is constant. The velocity and the height of water column have been obtained by solving the Saint Venant equations in this domain. Figures 3 and 4 show the BOD concentration at high tide and at low tide, respectively. The constraints are satisfied everywhere in the protected areas; at high tide they saturate at one of the vertices in zone 1 but at low tide, after a tidal cycle, the saturation takes place at one of the vertices in zone 2.

The values of the optimal discharges can be seen in Figure 6. The BOD rate of discharge is greater during rising tide at point 2 than at point 1. However, during ebb tide (after  $t = 60$ ) the flow rate decreases at  $P_2$  and increases at  $P_1$ . This is an obvious consequence of the outfalls position: in fact, during rising tide,  $P_2$  is *better* than  $P_1$ , but during ebb tide the opposite holds.

TABLE 2. Parameters for solving the problem ( $\mathcal{P}_{\mathcal{F}}$ ) on the *ría* of Vigo.

Saint Venant parameters	Empiric Coefficients	non-dimensional constants
Tidal cycle: $T=12.4$ h Tidal run: 2.8 m Water density: $1000 \text{ kg/m}^3$ Air density: $1.28 \text{ kg/m}^3$ North latitude: $0.7326$ rad Wind direction: $3.9269$ rad Wind velocity: $10 \text{ km/h}$ Angular velocity of Earth: $7.92 \cdot 10^{-5} \text{ rad/s}$	$\beta_1 = \beta_2 = 2000 \text{ m}^2/\text{s}$ $\kappa_1 = 1.15 \cdot 10^{-5} \text{ s}^{-1}$ $\kappa_2 = 9 \cdot 10^{-12} \text{ s}^{-1}$ $d_s = 8.98 \cdot 10^{-3} \text{ Kg/m}^3$ $\rho_{20} = 8.082 \cdot 10^{-3} \text{ Kg/m}^3$ $\sigma_1 = 3.48398 \cdot 10^{-4} \text{ Kg/m}^3$ $\sigma_2 = 5 \cdot 10^{-4} \text{ Kg/m}^3$ $\zeta_1 = 8.05255 \cdot 10^{-3} \text{ Kg/m}^3$ $\zeta_2 = 8.03218 \cdot 10^{-3} \text{ Kg/m}^3$ $\underline{m} = 0 \text{ Kg/m}^3$ $\overline{m} = 150 \text{ Kg/m}^3$	$N = 120$ $\xi_1 = 10^{-8}$ $\xi_2 = 10^{-8}$ $\mu_1 = 10^{-5}$ $\mu_2 = 10^{-5}$ $\lambda^I = 1$ $\theta^I = 1$

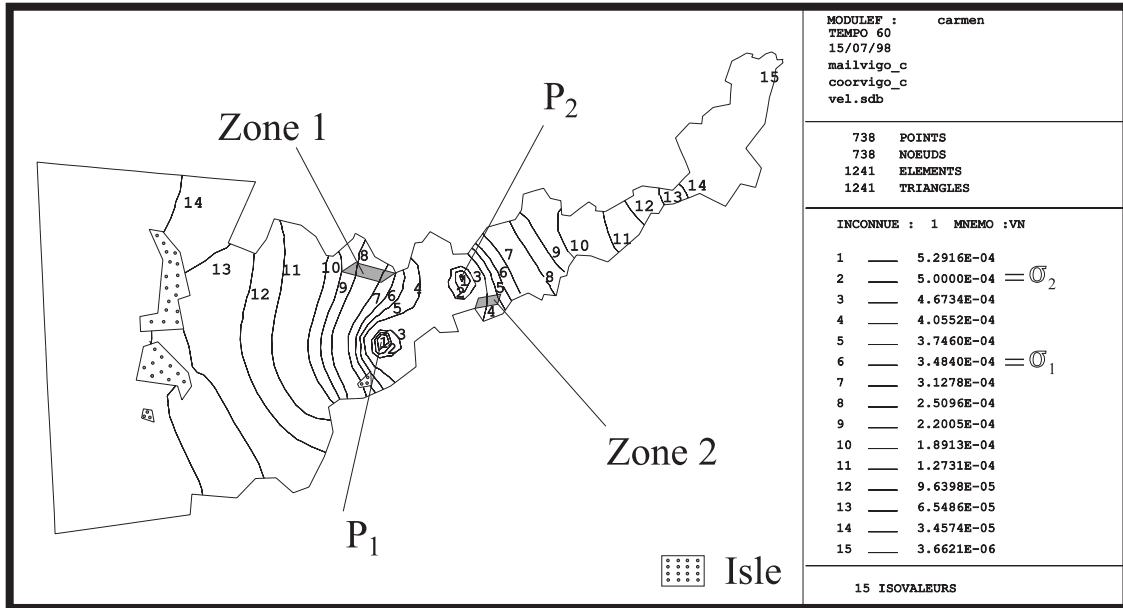


FIGURE 3. BOD concentration at high tide.

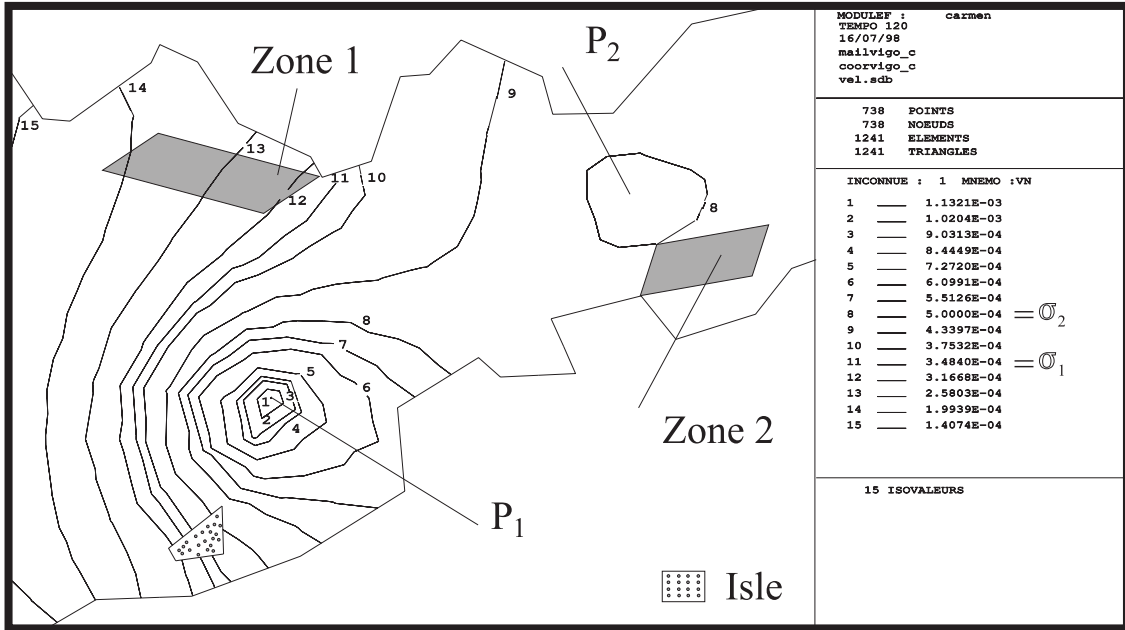


FIGURE 4. BOD concentration at low tide.

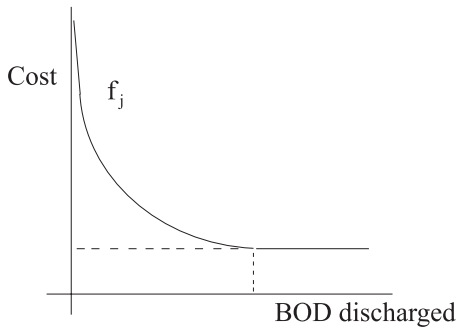


FIGURE 5. Cost function.

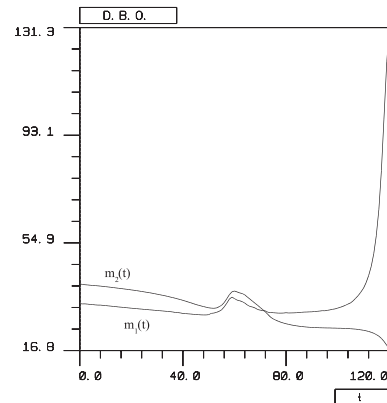


FIGURE 6. Optimal discharges during a tidal cycle.

### 3. ACTIVE CONTROL OF NOISE

#### 3.1. Introduction

Noise reduction is an important problem in acoustical and environmental engineering. While passive methods are good for mid and high frequencies they are not efficient for low frequencies. However, the latter can be efficiently reduced by active control. This is an old concept that has generated increased interest during the past decades due to the development of fast digital signal processors (DSP). It is based on the principle of destructive



interference of waves: an opposite pressure is generated by a secondary source to cancel an undesired noise. In order to achieve significant reduction, this source must produce, with great precision, an equal amplitude but inverted replica of the noise to be cancelled. Applications of these techniques are now successfully used to reduce noise in aircrafts or cars. Reference books on this subject are [3] and [26]. The general principles of active control of noise was described in a early patent by Leug in 1936. A microphone detects the undesired noise and provides an input signal to an electronic control system. The transfer from the microphone to the loudspeaker is adjusted so that the sound wave generated will destructively interfere the noise to be cancelled.

In this paper we state the problem of active control of noise in the framework of the optimal control theory of distributed systems. For the sake of simplicity we only consider that the noise to be cancelled has one single frequency but we can also control broad-band noise or non-periodic noise. Two problems are successively considered. In a first step complex amplitudes are taken as control and the objective is to minimize pressure at some particular points in the domain. In a second step, the locations of the loud-speakers are optimized with respect to the same objective function. A third step that is not included here would consist in choosing the locations of the microphones in view of minimizing the global noise, *i.e.* the norm of the pressure in the whole domain under consideration rather than at some finite number of points. We refer to [18] for further details.

### 3.2. Mathematical model

Let us consider a bounded domain  $\Omega_F \subset \mathbb{R}^n$  ( $n = 2$  ó  $3$ ) either convex polygonal or with smooth boundary enclosing an acoustic fluid, (*i.e.* inviscid compressible and barotropic). The propagation of acoustic waves is modelled by the well-know equation

$$\frac{1}{c^2} \frac{\partial^2 p(x, t)}{\partial t^2} - \Delta p(x, t) = f(x, t) \quad \text{in } \Omega_F,$$

where  $p$  is the pressure fluctuation,  $c$  is the sound speed and  $f$  is a source term. In our case  $f$  is going to be the secondary source of noise produced by loud-speakers, that is, the control. Moreover we have a primary source of noise acting on a part  $\Gamma_N$  of the boundary:

$$\frac{\partial p(x, t)}{\partial \mathbf{n}} = g(x, t) \quad \text{on } \Gamma_N \subset \partial\Omega_F,$$

where  $\mathbf{n}$  is an outward unit vector normal to the boundary. This means that the normal displacements are imposed on  $\Gamma_N$ . In practice it corresponds to vibrations of the enclosure. In this paper we assume that both  $g$  and  $f$  are harmonic sources, *i.e.*

$$g(x, t) = \text{Re} (G(x) e^{-i\omega t}), \quad f(x, t) = \text{Re} (F(x) e^{-i\omega t}),$$

where  $G(x)$  y  $F(x)$  are complex functions called complex amplitudes. Actually their modulus are the physical amplitudes while their arguments are called phase angles. Since the model is linear its solutions are also harmonic with the same frequency:

$$p(x, t) = \text{Re} (P(x) e^{-i\omega t}).$$

In this situation, the acoustic wave equation yields the following Helmholtz problem to be satisfied by the complex amplitude of pressure,

$$\left\{ \begin{array}{ll} -\Delta P - \left(\frac{\omega}{c}\right)^2 P = F & \text{in } \Omega_F, \\ \frac{\partial P}{\partial \mathbf{n}} = \frac{i\omega\rho}{Z(\omega)} P & \text{on } \Gamma_Z \subset \partial\Omega_F, \\ \frac{\partial P}{\partial \mathbf{n}} = G & \text{on } \Gamma_N \subset \partial\Omega_F, \end{array} \right. \quad (3.58)$$

where  $\rho$  is the density of the fluid and  $Z(\omega) \in \mathbb{C}$  is the so called *wall impedance*,

$$Z(\omega) := \beta(\omega) + \frac{\alpha(\omega)}{\omega} i.$$

The boundary condition on  $\Gamma_Z$  allows modelling the behaviour of absorbing viscoelastic materials on the walls of the enclosure. They are used as passive systems to reduce noise. Functions  $\beta(\omega)$  and  $\alpha(\omega)$  are related to the viscous and elastic responses, respectively. Both are strictly positive functions of the angular frequency  $\omega$ .

In our case  $F(x)$  is a linear combination of  $N_a$  Dirac's delta measures supported at some given points,  $\{x_1^a, \dots, x_{N_a}^a\}$ , with complex amplitudes,  $\{u_1, \dots, u_{N_a}\}$ , to be determined:

$$F(x) = \sum_{i=1}^{N_a} u_i \delta_{x_i^a}(x), \text{ with } u_i \in \mathbb{C}.$$

This amounts to consider loud-speakers as acoustic monopoles (see for instance [26]). In order to state the active control of noise as an optimal control problem we do the following choices,

- the *state of the system* is the pressure  $P(x)$  in the domain  $\Omega_F$ ;
- the *control variable*  $\mathbf{u}$  is the vector of complex amplitudes of loud-speakers (actuators),

$$\mathbf{u} = (u_1, \dots, u_{N_a}) \in \mathbb{C}^{N_a},$$

which define the source term  $F(x)$ . The set of admissible controls is a convex closed set  $U_{ad} \subseteq \mathbb{C}^{N_a}$ ;

- the *model of the system* relating the control to the state is the Helmholtz problem (3.58);
- the *observation*  $\mathbf{z}$  is the set of pressure values at  $N_s$  microphones (sensors) located at given points  $\{x_1^s, \dots, x_{N_s}^s\}$  contained in  $\Omega_F$ . We associate to it the following observation operator,

$$\mathbf{z} = \mathbf{C} P(x) = (P(x_1^s), \dots, P(x_{N_s}^s)) \in \mathbb{C}^{N_s}.$$

Evaluating pressure at points  $x_i^s$  makes sense as far as they are different from the locations of actuators. Actually we have the following

**Theorem 3.1.** *There exists a unique solution of the Helmholtz problem (3.58) in  $L^2(\Omega_F)$ . Moreover, for every  $\epsilon > 0$ , we have  $P \in C^0(\Omega_F \setminus \cup_{i=1}^{N_a} B_{(x_i^a, \epsilon)})$  with*

$$\|P\|_{C^0(\Omega_F \setminus \cup_{i=1}^{N_a} B_{(x_i^a, \epsilon)})} \leq C_\epsilon \|\mathbf{u}\|.$$

- The *cost function* to be minimized depends on the observation and maybe on the cost of the control itself:

$$J(\mathbf{u}) = \Phi(P(\mathbf{u}), \mathbf{u}) = \frac{1}{2} \|\mathbf{z}\|^2 + \frac{\nu}{2} \|\mathbf{u}\|^2 = \frac{1}{2} \sum_{i=1}^{N_s} |P(\mathbf{u}; x_i^s)|^2 + \frac{\nu}{2} \sum_{i=1}^{N_a} |u_i|^2,$$

where  $\nu \geq 0$  is a weighting factor.

Then the **optimal control problem** is to find  $\mathbf{u}_{op}$  such that

$$J(\mathbf{u}_{op}) = \inf_{\mathbf{u} \in U_{ad}} J(\mathbf{u}). \quad (3.59)$$

Any solution  $\mathbf{u}_{op}$  of this minimization problem will be called *optimal control*.

### 3.3. Existence and uniqueness of an optimal control

Since the control space is finite-dimensional to prove existence of an optimal control is an easy task. It relies upon the fact that the mapping giving the observation from the control, namely,

$$\begin{aligned} \mathbb{C}^{N_a} &\longrightarrow \mathcal{C}^0(\Omega_F \setminus \{x_1^a, \dots, x_{N_a}^a\}) \longrightarrow \mathbb{C}^{N_s} \\ \mathbf{u} &\longmapsto P(\mathbf{u}) \longmapsto \mathbf{z}(\mathbf{u}) = CP(\mathbf{u}) \end{aligned}$$

is affine (and then continuous) whenever the subsets  $\{x_1^a, \dots, x_{N_a}^a\}$  and  $\{x_1^s, \dots, x_{N_s}^s\}$  are disjoint. Therefore the cost function

$$J(\mathbf{u}) = \frac{1}{2} \sum_{i=1}^{N_s} |P(\mathbf{u}; x_i^s)|^2 + \frac{\nu}{2} \sum_{i=1}^{N_a} |u_i|^2,$$

is quadratic. Moreover it is strictly convex under one of the two following assumptions,

- $\nu > 0$ ;
- $\nu \geq 0$  and  $\mathbf{z}(\mathbf{u})$  is one-to-one;

in which case there exists a unique optimal control.

We notice that  $\mathbf{z}(\mathbf{u})$  is one-to-one if and only if the observations corresponding to each single actuator are linearly independent. Of course this implies that the number of microphones must be less or equal than the number of loud-speakers:  $N_s \geq N_a$ .

### 3.4. Numerical methods

Even if the cost function is defined in a finite-dimensional space, it involves the solution of a partial differential equation which has to be obtained by some discretization process as, for instance, finite element methods. This leads to approximate observations  $\{\mathbf{z}_{1h}, \dots, \mathbf{z}_{N_a h}\} \subset \mathbb{C}^{N_s}$  and thereby to an approximate cost function.

Let us consider a triangular (or tetrahedral) mesh of the domain  $\Omega_F$  with maximum diameter  $h$ . We call  $V_h \subset H^1(\Omega_F)$  the finite element space consisting of continuous piecewise linear functions on this mesh and denote by  $P_h$  the corresponding discrete solution of the Helmholtz problem (3.58). Following an analogous scheme to that by Scott [32] or Casas [10] for elliptic problems and using classical results by Mikhlin and Babuška we can show the following error estimate (see [18] for proof).

**Theorem 3.2.** *If  $P$  is the solution of (3.58) then there exists  $h_0 > 0$  such that, for all  $h < h_0$ , the discretized problem has unique solution and furthermore*

$$\|P - P_h\|_{L^2(\Omega_F)} \leq C(x_1^a, \dots, x_{N_a}^a) |\mathbf{u}| h^{2-n/2},$$

where constant  $C(x_1^a, \dots, x_{N_a}^a)$  goes to  $\infty$  as some  $x_i^a$  approaches the boundary  $\partial\Omega_F$ .

Now we write the control problem in matrix form. Let us introduce the vectors,

$$\begin{aligned} \mathbf{z}_0 &= \mathbf{C} P(\mathbf{0}), \\ \mathbf{z}_i &= \mathbf{C} P(\mathbf{e}_i) - \mathbf{z}_0 \text{ with } (\mathbf{e}_i)_j = \delta_{ij}, i, j = 1, \dots, N_a. \end{aligned}$$

Vector  $\mathbf{z}_0$  is the observation corresponding to the system without any control, *i.e.* the pressure at the sensors arising from the primary source only. Vectors  $\mathbf{z}_i$ ,  $i = 1, \dots, N_a$  are the observations when the system is only excited by the  $i$ -th loud-speaker with unit amplitude and without primary source. Let us define matrix  $\mathbf{Z}$  and vector  $\mathbf{b}_0$  by

$$\begin{aligned} (\mathbf{Z})_{ij} &= (\mathbf{z}_j, \mathbf{z}_i), \quad i, j = 1, \dots, N_a, \\ (\mathbf{b}_0)_i &= (\mathbf{z}_0, \mathbf{z}_i), \quad i = 1, \dots, N_a, \end{aligned}$$

where  $(\cdot, \cdot)$  denotes the scalar product in  $\mathbb{C}^{N_a}$ . Then it is easy to see that the optimal control problem is equivalent to the following quadratic programming problem. Find  $\mathbf{u}_{op} \in U_{ad}$  such that

$$J(\mathbf{u}_{op}) = \inf_{\mathbf{u} \in U_{ad}} \frac{1}{2} \{((\mathbf{Z} + \nu \mathbf{I}) \mathbf{u}, \mathbf{u}) + 2\text{Re}(\mathbf{b}_0, \mathbf{u}) + (z_0, z_0)\}. \quad (3.60)$$

Similar definitions of matrix  $\mathbf{Z}_h$  and vector  $\mathbf{b}_{h0}$  for the finite element approximation leads to the discrete optimal control problem: Find  $\mathbf{u}_{hop} \in U_{ad}$  such that

$$J(\mathbf{u}_{hop}) = \inf_{\mathbf{u}_h \in U_{ad}} \frac{1}{2} \{((\mathbf{Z}_h + \nu \mathbf{I}) \mathbf{u}_h, \mathbf{u}_h) + 2\text{Re}(\mathbf{b}_{h0}, \mathbf{u}_h) + (z_{h0}, z_{h0})\}. \quad (3.61)$$

The following error estimate is obtained in [18] by using pointwise error estimates for the state equation.

**Theorem 3.3.** *There exists  $h_0 > 0$  such that, for all  $h < h_0$ ,*

$$\|\mathbf{u}_{op} - \mathbf{u}_{oph}\|_\infty \leq \frac{1}{\alpha_Z - \|\Delta \mathbf{Z}\|_\infty} (\|\Delta \mathbf{Z}\|_\infty \|\mathbf{u}_{op}\|_\infty + \|\delta \mathbf{b}_0\|_\infty) \leq \frac{C}{\widehat{C}_2^n} \ln \left( \frac{\widehat{C}_3}{h} \right) h^2 \|\mathbf{u}_{op}\|_\infty,$$

where,  $\alpha_Z > 0$  is a positive constant such that

$$((\mathbf{Z} + \nu \mathbf{I}) \mathbf{v}, \mathbf{v}) \geq \alpha_Z \|\mathbf{v}\|^2, \quad \forall \mathbf{v} \in U_{ad}$$

where  $\Delta \mathbf{Z} = \mathbf{Z} - \mathbf{Z}_h$  and  $\delta \mathbf{b}_0 = \mathbf{b}_0 - \mathbf{b}_{h0}$ .

### 3.5. Numerical results

In this paragraph we present some numerical results for a 3D test. In order to assess the effect of noise control we use the following measure of attenuation,

$$\text{Attenuation (dB)} = 10 \log \left( \frac{\sum_{i=1}^{N_s} |P(\mathbf{u}; x_i^s)|^2}{\sum_{i=1}^{N_s} |P(\mathbf{0}; x_i^s)|^2} \right).$$

The data of the problem are

- $\Omega_F = [0, 1] \times [0, 1] \times [0, 1]$ ;
- $\omega = 2$ ,  $c = 1$ ,  $\rho = 1$   $\nu = 0$ ;
- primary source of noise:  $G(x, y, 0) = e^{i 2\pi y}$ , on  $z = 0$ ;
- one loud-speaker located at  $x_1^a = (4/6, 4/6, 4/6)$ ;
- two microphones located at  $x_1^s = (1/6, 2/6, 1/6)$  and  $x_2^s = (5/6, 1/6, 1/2)$ ;
- the wall  $z = 1$  has an acoustic impedance for  $\omega = 2$  given by  $Z(2) = 300 + 10^3 i$ . The rest of the walls are rigid.

In Figure 7 the coarsest mesh is shown. The optimal control has been computed for several mesh refinements and then an extrapolated optimal control has been determined. This is used to compute the relative errors of both its real and imaginary parts. They are shown in Figure 8 where an order two of convergence is observed, as predicted by theoretical results.

For  $h = 1/24$ , the computed attenuation is  $-0.75$  dB. Figures 9 and 10 show the modulus of the complex pressure field without control and with control, respectively, on the plane containing the actuator and the two sensors. Figure 11 represents the attenuation field on the same plane. In this case we can observe that there exists zones where noise is reinforced, that is, where primary and secondary sources interfere in a constructive way.

### 3.6. Optimal location of actuators

In the previous paragraph the position of both sensors (microphones) and actuators (loud-speakers) were given. Then we used the complex amplitudes of actuators as a control variable and determine their optimal values with the objective of minimizing the pressure level at points where sensors were located. Now we assume that the positions of actuators can be chosen in certain subsets of the domain and determine those that minimize the same objective function as above when the complex amplitudes with respect to these positions are optimal. This is an important problem when a system of active control of sound has to be implemented to reduce noise in an enclosure. It can also be formulated as an optimal control problem. The difficulty now is that dependency of the state with respect to the additional control variables (the positions of actuators) is no longer affine. Thus, the cost function may have many local minima and therefore gradient-like methods are not suitable to solve the problem. Moreover, in practical situations, the number of feasible locations is finite and hence the optimization problem is an integer programming problem. In these cases, one can use, for instance, genetic or simulated annealing algorithms. Once the optimal locations have been determined it will be possible to improve them in some given neighborhoods by performing a local minimization using the gradient of the cost function with respect to the location of the actuators. The reader is referred to [18] for a full analysis of this problem;

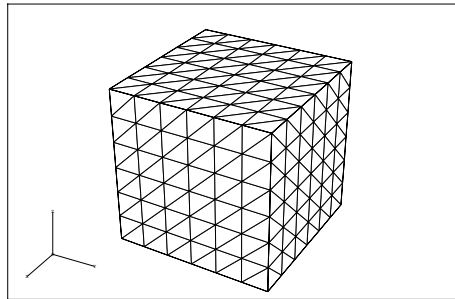


FIGURE 7. Coarsest mesh ( $h = 1/6$ ).

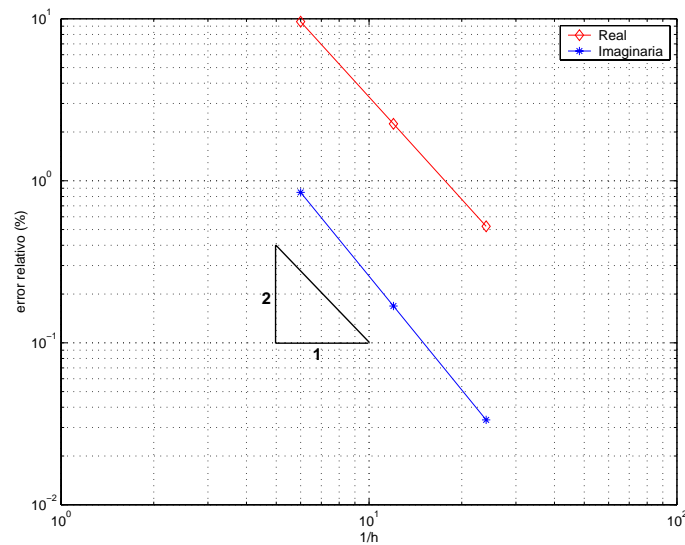


FIGURE 8. Relative error (%) as a function of  $1/h$  in log-log scale.

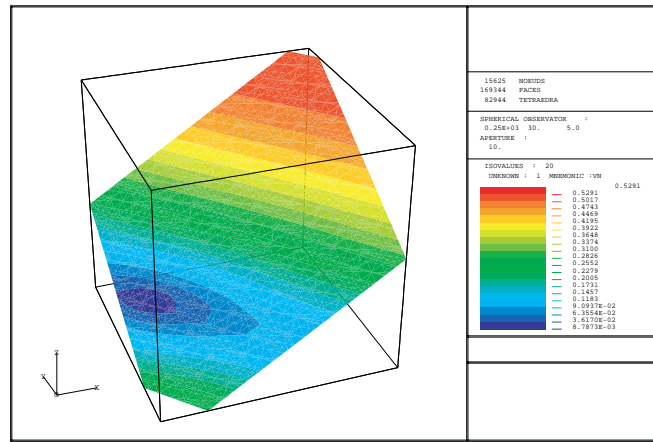


FIGURE 9. Pressure modulus without control for  $h = 1/24$ .

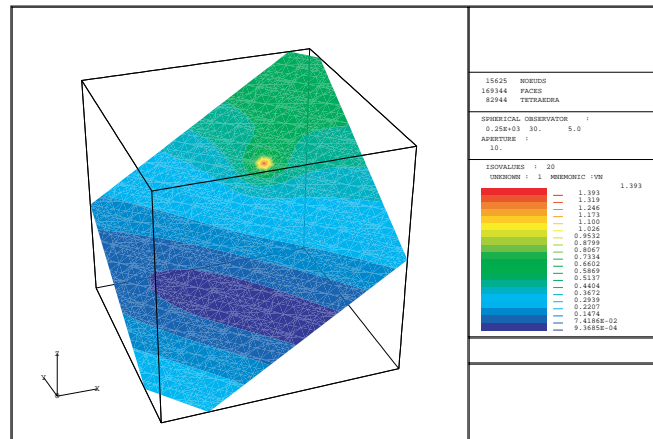


FIGURE 10. Pressure modulus with control for  $h = 1/24$ .

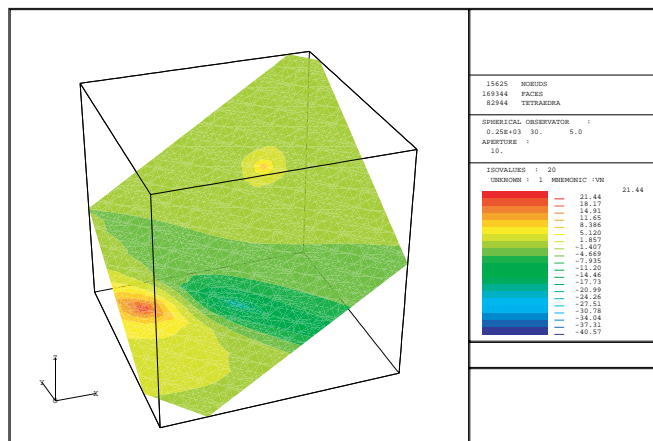


FIGURE 11. Attenuation field for  $h = 1/24$ .

in particular, the first order optimality conditions given below are proved there. They allow computing the gradient of the cost function through an adjoint state.

**Theorem 3.4.** *Let  $B_{ad} \subset \mathbb{R}^N$  the set of admissible locations for actuators. If  $(\mathbf{u}_{op}, \mathbf{b}_{op}) \in U_{ad} \times B_{ad}$  is an optimal control and  $P$  the corresponding state, then there exists an adjoint state  $R$  such that the following equations hold,*

**State equation:**

$$\left\{ \begin{array}{ll} -\Delta P - \left(\frac{\omega}{c}\right)^2 P = \sum_{i=1}^{N_a} u_{op_i} \delta_{b_{op_i}} & \text{in } \Omega_F, \\ \frac{\partial P}{\partial \mathbf{n}} = \frac{i\omega\rho}{Z(\omega)} P & \text{on } \Gamma_Z \subset \partial\Omega_F, \\ \frac{\partial P}{\partial \mathbf{n}} = G & \text{on } \Gamma_N \subset \partial\Omega_F. \end{array} \right. \quad (3.62)$$

**Adjoint state equation:**

$$\left\{ \begin{array}{ll} -\Delta R - \left(\frac{\omega}{c}\right)^2 R = \sum_{i=1}^{N_s} P(x_i^s) \delta_{x_i^s} & \text{in } \Omega_F, \\ \frac{\partial R}{\partial \mathbf{n}} = \frac{-i\omega\rho}{Z(\omega)} R & \text{on } \Gamma_Z \subset \partial\Omega_F, \\ \frac{\partial R}{\partial \mathbf{n}} = 0 & \text{on } \Gamma_N \subset \partial\Omega_F. \end{array} \right. \quad (3.63)$$

**Euler's inequality**

$$\operatorname{Re} \left\{ \sum_{i=1}^{N_a} (R(b_{op_i}) + \nu u_{op_i}, u_i - u_{op_i}) \right\} + \sum_{i=1}^{N_a} \operatorname{Re} \{ \bar{u}_{op_i} \nabla R(b_{op_i}) \} \cdot (b_i - b_{op_i}) \geq 0, \quad \forall (\mathbf{u}, \mathbf{b}) \in U_{ad} \times B_{ad}. \quad (3.64)$$

The author is indebted to P. Gamallo, A. Martínez, C. Rodríguez and M.E. Vázquez-Méndez for their help during the preparation of this article.

## REFERENCES

- [1] L. Álvarez-Vázquez and A. Martínez, Modelling and control of natural convection in canned foods. *IMA J. Appl. Math.* **63** (1999) 246-265.
- [2] K.H. Baek and S.J. Elliot, Natural algorithms for choosing source locations in active control systems. *J. Sound Vibr.* **186** (1995) 245-267.
- [3] Beranek and Ver, *Noise and vibration control engineering. Principles and applications*. John Wiley and Sons, New York (1992).
- [4] A. Bermúdez, Mathematical techniques for some environmental problems related to water pollution control, in *Mathematics, Climate and Environment*, edited by J.I. Díaz, J.-L. Lions. Masson, Paris (1993).
- [5] A. Bermúdez and A. Martínez, A state constrained optimal control problem related to the sterilization of canned foods. *Automatica. The IFAC Journal* **30** (1994) 319-329.
- [6] A. Bermúdez, A. Martínez and C. Rodríguez, Un problème de contrôle ponctuel lié à l'emplacement optimal d'émissaires d'évacuation sous-marine. *C. R. Acad. Sci. Paris Sér. I Math.* **313** (1991) 515-518.
- [7] A. Bermúdez, C. Rodríguez and M.A. Vilar, Solving shallow water equations by a mixed implicit finite element method. *IMA J. Num. Anal.* **11** (1991) 79-97.
- [8] A. Bermúdez and C. Saguez, Optimal control of a Signorini problem. *SIAM J. Control Optim.* **25** (1987) 576-582.
- [9] J.F. Bonnans and E. Casas, Contrôle de systèmes elliptiques semilinéaires comportant des contraintes distribuées sur l'état, in *Nonlinear partial differential equations and their applications*, edited by H. Brezis and J.-L. Lions. Pitman (1988).
- [10] E. Casas,  $L^2$  estimates for the finite element method for the Dirichlet problem with singular data. *Numer. Math.* **47** (1985) 627-632.
- [11] E. Casas, Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.* **24** (1986) 1309-1318.
- [12] E. Casas, Pontryagin's principle for state constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.* **35** (1997) 1297-1327.

- [13] J.F. Bonnans, *An introduction to Newton type algorithms for nonlinearly constrained optimization problems*. Birkhauser-Verlag, Basel, *Internat. Ser. Numer. Math.* **87** (1989) 1-17.
- [14] E. Casas and C. Pola, *PLCBAS User's Guide VERSION 1.1*. Computación 1. Universidad de Cantabria, Santander, Spain (1989).
- [15] P.G. Ciarlet, Basic error estimates for elliptic problems, in *Handbook of Numerical Analysis*, Vol. II, edited by P.G. Ciarlet and J.-L. Lions. North-Holland (1991).
- [16] E. Di Benedetto, On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **13** (1986) 487-535.
- [17] I. Ekeland and R. Temam, *Convex analysis and variational problems*. North-Holland, Amsterdam (1976).
- [18] P. Gamallo, *Contribución al estudio matemático de problemas de simulación y control activo del ruido*, Ph. Thesis. Universidade de Santiago de Compostela, Spain (2002).
- [19] J. Herskovits, A two stage feasible directions algorithm for nonlinear constrained optimization. *Math. Programming* **36** (1986) 19-38.
- [20] J. Herskovits, A feasible directions interior point technique for nonlinear optimization. *J. Optim. Theory Appl.* **99** (1998) 121-146.
- [21] J.B. Hiriart-Urruty and C. Lemarechal, *Convex analysis and Minimization Algorithms*. Springer-Verlag, Berlin, Heidelberg (1993).
- [22] B. Hu and J. Yong, Pontriagin maximum principle for semilinear and quasilinear parabolic equations with pointwise state constraints. *SIAM J. Control Optim.* **33** (1995) 1857-1880.
- [23] O.A. Ladyzhenskaya, V.A. Solonnikov and N.N. Uraltseva, *Linear and quasilinear equations of parabolic type*. Amer. Math. Soc., Providence, *Transl. Math. Monogr.* **23** (1968).
- [24] J.-L. Lions, *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*. Dunod, Paris (1968).
- [25] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, Paris (1969).
- [26] P.A. Nelson and S.J. Elliot, *Active Control of Sound*. Academic Press, London (1999).
- [27] G.I. Marchuk, *Mathematical models in environmental problems*. North Holland, Amsterdam (1986).
- [28] A. Martínez, C. Rodríguez and M.E. Vázquez-Méndez, Theoretical and numerical analysis of an optimal control problem related to waste-water treatment. *SIAM J. Control Optim.* **38** (2000) 1534-1553.
- [29] C. Olin Ball and F.C.W. Olson, *Sterilization in food technology*. Mc Graw Hill, New York (1957).
- [30] R.I. Pérez Martín, J.R. Banga and J.M. Gallardo, *Simulation of thermal processes in tuna can manufacture*. Instituto de Investigaciones Marinas (C.S.I.C.), Vigo, Spain (1989).
- [31] E.R. Panier, A.L. Tits and J. Herskovits, A QP-Free, Globally Convergent, Locally Superlinearly Convergent Algorithm for Inequality Constrained Optimization. *SIAM J. Control Optim.* **26** (1988) 788-810.
- [32] R. Scott, Finite element convergence for singular data. *Numer. Math.* **21** (1973) 317-327.
- [33] M.E. Vázquez-Méndez, *Contribución a la resolución numérica de modelos para el estudio de la contaminación de aguas*. Master thesis. Dept. Matemática Aplicada. Univ. Santiago de Compostela, Spain (1992).