

TOPOLOGICAL OPTIMIZATION VIA COST PENALIZATION

CORNEL MARIUS MUREA — DAN TIBA

Dedicated to the memory of Prof. Dr. Ioan I. Vrabie

ABSTRACT. We consider general shape optimization problems governed by Dirichlet boundary value problems. The proposed approach may be extended to other boundary conditions as well. It is based on a recent representation result for implicitly defined manifolds, due to the authors, and it is formulated as an optimal control problem. The discretized approximating problem is introduced and we give an explicit construction of the associated discrete gradient. Some numerical examples are also indicated.

1. Introduction

Shape optimization is a relatively young branch of mathematics, with important modern applications in engineering and design. Certain optimization problems in mechanics, thickness optimization for plate or rods, geometric optimization of shells, curved rods, drag minimization in fluid mechanics, etc. are some examples. Many appear naturally in the form of control by coefficients problems, due to the formulation of the mechanical models, with the geometric characteristics entering the coefficients of the differential operators. See [15, Chapter 6], where such questions are discussed in details.

It is the aim of this article to develop an optimal control approach, using penalization methods, to general shape optimization problems as investigated in [20], [23], [5], [10], [8], etc. We underline that our methodology allows simultaneous topological and boundary variations.

2010 *Mathematics Subject Classification*. Primary: 49K20, 65K10; Secondary: 49M30.

Key words and phrases. Geometric optimization; optimal design; topological variations; optimal control methods; discrete gradient.

Here, we fix our attention to the Dirichlet boundary conditions and we study the typical problem (denoted by (\mathcal{P})):

$$(1.1) \quad \min_{\Omega} \int_E j(\mathbf{x}, y_{\Omega}(\mathbf{x})) d\mathbf{x},$$

$$(1.2) \quad -\Delta y_{\Omega} = f \quad \text{in } \Omega,$$

$$(1.3) \quad y_{\Omega} = 0 \quad \text{on } \partial\Omega,$$

where $E \subset\subset D \subset \mathbb{R}^2$ are given bounded domains, D is of class $\mathcal{C}^{1,1}$ and the minimization parameter, the unknown domain Ω , satisfies $E \subset \Omega \subset D$ and other possible conditions defining a class of admissible domains. Notice that the case of dimension two is of interest in shape optimization. Moreover, $f \in L^2(D)$, $j : D \times \mathbb{R} \rightarrow \mathbb{R}$ is some Carathéodory mapping. More assumptions or constraints will be imposed later. Other boundary conditions or differential operators may be handled as well via this control approach and we shall examine such questions in a subsequent paper.

For fundamental properties and methods in optimal control theory, we quote [12], [7], [2], [16]. The problem (1.1)–(1.3) and its approximation are strongly non convex and challenging both from the numerical and theoretical points of view. The investigation from this paper continues the one in [26] and is essentially based on the recent implicit parametrization method as developed in [25], [18], [24], that provides an efficient analytic representation of the unknown domains.

The Hamiltonian approach to implicitly defined manifolds will be briefly recalled together with other preliminaries in Section 2. The precise formulation of the problem and its approximation is analyzed in Section 3 together with its differentiability properties. In Section 4, we study the discretized version and find the general form of the discrete gradient. The last section is devoted to some numerical experiments, using this paper approach.

The method studied in this paper has a certain complexity due to the use of Hamiltonian systems and its main advantage is the possibility to extend it to other boundary conditions or boundary observation problems. This will be performed in a subsequent article.

2. Preliminaries

Consider the Hamiltonian system

$$(2.1) \quad x_1'(t) = -\frac{\partial g}{\partial x_2}(x_1(t), x_2(t)), \quad t \in I,$$

$$(2.2) \quad x_2'(t) = \frac{\partial g}{\partial x_1}(x_1(t), x_2(t)), \quad t \in I,$$

$$(2.3) \quad (x_1(0), x_2(0)) = (x_1^0, x_2^0),$$

where $g: D \rightarrow \mathbb{R}$ is in $\mathcal{C}^1(\overline{D})$, $(x_1^0, x_2^0) \in D$ and I is the local existence interval for (2.1)–(2.3), around the origin, obtained via the Peano theorem. The conservation property [24] of the Hamiltonian gives:

PROPOSITION 2.1. *We have*

$$(2.4) \quad g(x_1(t), x_2(t)) = g(x_1^0, x_2^0), \quad t \in I.$$

In the sequel, we assume that

$$(2.5) \quad g(x_1^0, x_2^0) = 0, \quad \nabla g(x_1^0, x_2^0) \neq 0.$$

Under condition (2.5), i.e. in the noncritical case, it is known that the solution of (2.1)–(2.3) is also unique (by applying the implicit functions theorem to (2.4), see [4]).

REMARK 2.2. In higher dimension, iterated Hamiltonian systems were introduced in [25] and uniqueness and regularity properties are proved. Some relevant examples in dimension three are discussed in [18]. In the critical case, generalized solutions can be obtained [24], [25].

We define now the family \mathcal{F} of admissible functions $g \in \mathcal{C}^2(\overline{D})$ that satisfy the conditions:

$$(2.6) \quad g(x_1, x_2) > 0, \quad \text{on } \partial D,$$

$$(2.7) \quad |\nabla g(x_1, x_2)| > 0, \quad \text{on } \mathcal{G} = \{(x_1, x_2) \in D : g(x_1, x_2) = 0\},$$

$$(2.8) \quad g(x_1, x_2) < 0, \quad \text{on } \overline{E}.$$

Condition (2.6) says that $\mathcal{G} \cap \partial D = \emptyset$ and condition (2.7) is an extension of (2.5). In fact, it is related to the hypothesis on the non existence of equilibrium points in the Poincaré–Bendixson theorem, see [11, Chapter 10], and the same is valid for the next proposition. The family \mathcal{F} defined by (2.6)–(2.8) is obviously very rich, but it is not “closed” (we have strict inequalities). Our approach here, gives a descent algorithm for the shape optimization problem (\mathcal{P}) and existence of optimal shapes is not discussed.

Following [26], we have the following two propositions:

PROPOSITION 2.3. *Under hypotheses (2.6), (2.7), \mathcal{G} is a finite union of closed curves of class \mathcal{C}^2 , without self intersections, parametrized by (2.1)–(2.3), when some initial point (x_1^0, x_2^0) is chosen on each component of \mathcal{G} .*

If $r \in \mathcal{F}$ as well, we define the perturbed set

$$(2.9) \quad \mathcal{G}_\lambda = \{(x_1, x_2) \in D : (g + \lambda r)(x_1, x_2) = 0, \lambda \in \mathbb{R}\}.$$

We also introduce the neighbourhood V_ε , $\varepsilon > 0$

$$(2.10) \quad V_\varepsilon = \{(x_1, x_2) \in D : d[(x_1, x_2), \mathcal{G}] < \varepsilon\},$$

where $d[(x_1, x_2), \mathcal{G}]$ is the distance from a point to \mathcal{G} .

PROPOSITION 2.4. *If $\varepsilon > 0$ is small enough, there is $\lambda(\varepsilon) > 0$ such that, for $|\lambda| < \lambda(\varepsilon)$ we have $\mathcal{G}_\lambda \subset V_\varepsilon$ and \mathcal{G}_λ is a finite union of class C^2 closed curves.*

REMARK 2.5. The inclusion $\mathcal{G}_\lambda \subset V_\varepsilon$ shows that $\mathcal{G}_\lambda \rightarrow \mathcal{G}$ for $\lambda \rightarrow 0$, in the Hausdorff–Pompeiu metric [15]. In a “small” neighbourhood of each component of \mathcal{G} there is exactly one component of \mathcal{G}_λ if $|\lambda| < \lambda(\varepsilon)$, due to this convergence property and the implicit functions theorem applied in the initial condition of the perturbed Hamiltonian system derived from (2.1)–(2.3).

PROPOSITION 2.6. *Denote by $T_g, T_{g+\lambda r}$ the periods of the trajectories of (2.1)–(2.3), corresponding to $g, g + \lambda r$, respectively. Then $T_{g+\lambda r} \rightarrow T_g$ as $\lambda \rightarrow 0$.*

PROOF. If (x_1, x_2) , respectively $(x_{1\lambda}, x_{2\lambda})$, are the corresponding trajectories of (2.1)–(2.3), respectively, then they are bounded by Proposition 2.4, if $|\lambda| < \lambda(\varepsilon)$. Consequently, ∇g may be assumed Lipschitzian with constant L_g and we have

$$(2.11) \quad |(x_1, x_2) - (x_{1\lambda}, x_{2\lambda})|(t) \leq \lambda C + L_g \int_0^t |(x_1, x_2) - (x_{1\lambda}, x_{2\lambda})| dt,$$

where we also use that $\nabla r(x_{1\lambda}, x_{2\lambda})$ is bounded since $(x_{1\lambda}, x_{2\lambda})$ is bounded on \mathbb{R} . We infer by (2.11) that

$$(2.12) \quad |(x_1, x_2) - (x_{1\lambda}, x_{2\lambda})|(t) \leq \lambda ct, \quad \text{for } t \in \mathbb{R},$$

$$(2.13) \quad |(x'_1, x'_2) - (x'_{1\lambda}, x'_{2\lambda})|(t) \leq \lambda ct, \quad \text{for } t \in \mathbb{R},$$

for $|\lambda| < \lambda(\varepsilon)$ and with some constant independent of λ , by Gronwall lemma.

Both trajectories start from (x_1^0, x_2^0) , surround \bar{E} , have no self intersections (but $(x_{1\lambda}, x_{2\lambda})$ may intersect (x_1, x_2) even on infinity of times). We study them on $[0, jT_g]$, $j < 2$, for instance.

Assume that $(x_{1\lambda}, x_{2\lambda})$ has the period $T_{g+\lambda r} > jT_g$. Since (x_1, x_2) is periodic with period T_g and relations (2.12)–(2.13) show that $(x_{1\lambda}, x_{2\lambda})$ is very close to (x_1, x_2) in every $t \in [0, jT_g]$ it yields that $(x_{1\lambda}, x_{2\lambda})$ is, as well, surrounding \bar{E} at least once. As it may have no self intersections, it yields that $(x_{1\lambda}, x_{2\lambda})$ is a limit cycle around \bar{E} . Such arguments appear in the proof of the Poincaré–Bendixson theorem, [11, Chapter 10]. That is $(x_{1\lambda}, x_{2\lambda})$ cannot be periodic – and this is a false conclusion due to Proposition 2.4.

Consequently, we get

$$(2.14) \quad T_{g+\lambda r} \leq jT_g, \quad |\lambda| < \lambda(\varepsilon).$$

On a subsequence, by (2.14), we obtain $T_{g+\lambda r} \rightarrow T^* \leq jT_g$. We assume that $T^* \neq T_g$. It is clear that $(x_1(T^*), x_2(T^*)) \neq (x_1(T_g), x_2(T_g))$ by the definition of the period. However, relation (2.12) and the related convergence properties give the opposite conclusion. This contradiction shows that $T^* = T_g$ and the convergence is valid on the whole sequence. \square

REMARK 2.7. Usually, the perturbation of a periodic solution may not be periodic and just asymptotic convergence properties are valid, under certain assumptions, Sideris [22]. A natural question, taking into account (2.12)–(2.13), is whether $|T_{g+\lambda r} - T_g| \leq c|\lambda|$, for $c > 0$ independent of λ , $|\lambda| < \lambda(\varepsilon)$.

3. The optimization problem and its approximation

Starting from the family \mathcal{F} of admissible functions, we define the family \mathcal{O} of admissible domains in the shape optimization problem (1.1)–(1.3) as the connected component of the open set Ω_g , $g \in \mathcal{F}$,

$$(3.1) \quad \Omega_g = \{(x_1, x_2) \in D : g(x_1, x_2) < 0\}$$

that contains E . Clearly $E \subset \Omega_g$ by (2.8). Notice as well that the domain Ω_g defined by (3.1) (we use this notation for the domain as well) is not simply connected, in general. This is the reason why the approach to (1.1)–(1.3) that we discuss here is related to topological optimization in optimal design problems. But, it also combines topological and boundary variations.

The penalized problem, $\varepsilon > 0$, is given by:

$$(3.2) \quad \min_{g \in \mathcal{F}, u \in L^2(D)} \left\{ \int_E j(\mathbf{x}, y_\varepsilon(\mathbf{x})) \, d\mathbf{x} + \frac{1}{\varepsilon} \int_{I_g} (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_g(t)| \, dt \right\}$$

subject to

$$(3.3) \quad -\Delta y_\varepsilon = f + (g + \varepsilon)_+^2 u \quad \text{in } D,$$

$$(3.4) \quad y_\varepsilon = 0 \quad \text{on } \partial D,$$

where $\mathbf{z}_g = (z_g^1, z_g^2)$ satisfies the Hamiltonian system (2.1)–(2.3) in I_g with some $(x_1^0, x_2^0) \in D \setminus \bar{E}$ such that $g(x_1^0, x_2^0) = 0$

$$(3.5) \quad (z_g^1)'(t) = -\frac{\partial g}{\partial x_2}(\mathbf{z}_g(t)) \quad \text{for } t \in I_g,$$

$$(3.6) \quad (z_g^2)'(t) = \frac{\partial g}{\partial x_1}(\mathbf{z}_g(t)) \quad \text{for } t \in I_g,$$

$$(3.7) \quad \mathbf{z}_g(0) = (x_1^0, x_2^0),$$

and $I_g = [0, T_g]$ is the period interval for (3.5)–(3.7), due to Proposition 2.3.

The problem (3.2)–(3.7) is an optimal control problem with controls $g \in \mathcal{F}$ and $u \in L^2(D)$ distributed in D . The state is given by $[y_\varepsilon, z_g^1, z_g^2] \in H^2(D) \times (C^2(I_g))^2$. We also have $y_\varepsilon \in H_0^1(D)$. In case the corresponding domain Ω_g is not simply connected, in (3.7) one has to choose initial conditions on each component of $\partial\Omega_g$ and the penalization term becomes a finite sum due to Proposition 2.3. The method enters the class of fixed domain methods in shape optimization and can be compared with [13], [14], [17]. It is essentially different from the level set method of Osher and Sethian [19], Allaire [1] or the SIMP approach of Bendsoe and Sigmund [3]. From the computational point of view, it is easy to find initial

condition (3.7) on each component of \mathcal{G} and the corresponding period intervals I_g associated to (3.5)–(3.7). See the last section as well.

We have the following general suboptimality property:

PROPOSITION 3.1. *Let $j(\cdot, \cdot)$ be a Carathéodory function on $D \times \mathbb{R}$, bounded from below by a constant. Denote by $[y_n^\varepsilon, g_n^\varepsilon, u_n^\varepsilon]$ a minimizing sequence in (3.2)–(3.7). Then, on a subsequence $n(m)$ the (not necessarily admissible pairs) $[\Omega_{g_{n(m)}^\varepsilon}, y_{n(m)}^\varepsilon]$ give a minimizing sequence in (1.1), $y_{n(m)}^\varepsilon$ satisfies (1.2) in*

$$\{(x_1, x_2) \in D : g(x_1, x_2) < -\varepsilon\}$$

and (1.3) is fulfilled with a perturbation of order $\varepsilon^{1/2}$.

PROOF. Let $[y_{g_m}, g_m] \in H^2(\Omega_{g_m}) \times \mathcal{F}$ be a minimizing sequence in the problem (1.1)–(1.3), (3.1). By the trace theorem, since $\partial\Omega_{g_m}$ and D are at least $C^{1,1}$ under our assumptions, there is $\tilde{y}_{g_m} \in H^2(D) \cap H_0^1(D)$, not unique, such that $\tilde{y}_{g_m} = y_{g_m}$ in Ω_{g_m} . We define the control $u_{g_m} \in L^2(D)$ as following:

$$\begin{aligned} u_{g_m} &= 0, && \text{in } \Omega_{g_m}, \\ u_{g_m} &= -\frac{\Delta \tilde{y}_{g_m} + f}{(g_m + \varepsilon)_+^2}, && \text{in } \partial D \setminus \Omega_{g_m}, \end{aligned}$$

where Ω_{g_m} is the open set defined in (3.1). Notice that on the second line in the above formula, we have no singularity. It is clear that the triple $[\tilde{y}_{g_m}, g_m, u_{g_m}]$ is admissible for the problem (3.2)–(3.7) with the same cost as in the original problem (1.1)–(1.3) since the penalization term in (3.2) is null due to the boundary condition (1.3) satisfied by \tilde{y}_{g_m} . Consequently, there is $n(m)$ sufficient big, such that

$$\begin{aligned} (3.8) \quad & \int_E j(\mathbf{x}, y_{n(m)}^\varepsilon(\mathbf{x})) d\mathbf{x} + \frac{1}{\varepsilon} \int_{I_{g_{n(m)}^\varepsilon}} (y_{n(m)}^\varepsilon(\mathbf{z}_{g_{n(m)}^\varepsilon}(t)))^2 |\mathbf{z}'_{g_{n(m)}^\varepsilon}(t)| dt \\ & \leq \int_E j(\mathbf{x}, \tilde{y}_{g_m}(\mathbf{x})) d\mathbf{x} = \int_E j(\mathbf{x}, y_{g_m}(\mathbf{x})) d\mathbf{x} \rightarrow \inf(\mathcal{P}). \end{aligned}$$

Since j is bounded from below, we get from (3.8):

$$(3.9) \quad \int_{\partial\Omega_{g_m}} (y_{n(m)}^\varepsilon)^2 d\sigma \leq C\varepsilon$$

with C a constant independent of $\varepsilon > 0$. Then, (3.9) shows that (1.3) is fulfilled with a perturbation of order $\varepsilon^{1/2}$. Moreover, again by (3.8), we see the minimizing property of $\{y_{n(m)}^\varepsilon\}$ in the original problem (\mathcal{P}) .

We notice that in the state equation (3.3), the right-hand side coincides with f in the set $\{(x_1, x_2) \in D : g(x_1, x_2) < -\varepsilon\}$, which is an approximation of $\Omega_{g_{n(m)}^\varepsilon}$. Namely, we notice that for any $g \in \mathcal{F}$, the open sets $\{(x_1, x_2) \in D : g(x_1, x_2) < -\varepsilon\}$ form a nondecreasing sequence contained in $\bar{\Omega}_g$, when $\varepsilon \rightarrow 0$. Take (x_1, x_2) such that $g(x_1, x_2) = 0$ and take some sequence $(x_1^n, x_2^n) \rightarrow (x_1, x_2)$,

$(x_1^n, x_2^n) \in \Omega_g$. We have $g(x_1^n, x_2^n) < 0$ by (3.1) and $g(x_1^n, x_2^n) \rightarrow 0$. Moreover, $(x_1^n, x_2^n) \in \Omega_{\varepsilon_n} = \Omega_{g+\varepsilon_n}$, for $\varepsilon_n > 0$ sufficiently small. Consequently, we have the desired convergence property by [15, p. 461]. \square

REMARK 3.2. A detailed study of the approximation properties in the penalized problem is performed in [26], in a slightly different case.

We consider now variations $u + \lambda v$, $g + \lambda r$, where $\lambda \in \mathbb{R}$, $u, v \in L^2(D)$, $g, r \in \mathcal{F}$, $g(x_1^0, x_2^0) = r(x_1^0, x_2^0) = 0$. Notice that $u + \lambda v \in L^2(D)$ and $g + \lambda r \in \mathcal{F}$ for $|\lambda|$ sufficiently small. The conditions (2.6)–(2.8) from the definition of \mathcal{F} are satisfied for $|\lambda|$ sufficiently small (depending on g) due to the Weierstrass theorem and the fact that \bar{E} , ∂D and \mathcal{G} are compacts. Here, we also use Proposition 2.4. Consequently, we assume $|\lambda|$ “small”. We study first the differentiability properties of the state system (3.3)–(3.7):

PROPOSITION 3.3. *The system of variations corresponding to (3.3)–(3.7) is*

$$(3.10) \quad -\Delta q_\varepsilon = (g + \varepsilon)_+^2 v + 2(g + \varepsilon)_+ u r \quad \text{in } D,$$

$$(3.11) \quad q_\varepsilon = 0, \quad \text{on } \partial D,$$

$$(3.12) \quad w'_1 = -\nabla \partial_2 g(\mathbf{z}_g) \cdot \mathbf{w} - \partial_2 r(\mathbf{z}_g) \quad \text{in } I_g,$$

$$(3.13) \quad w'_2 = \nabla \partial_1 g(\mathbf{z}_g) \cdot \mathbf{w} + \partial_1 r(\mathbf{z}_g) \quad \text{in } I_g,$$

$$(3.14) \quad w_1(0) = 0, \quad w_2(0) = 0,$$

where $q_\varepsilon = \lim_{\lambda \rightarrow 0} (y_\varepsilon^\lambda - y_\varepsilon)/\lambda$, $\mathbf{w} = [w_1, w_2] = \lim_{\lambda \rightarrow 0} (\mathbf{z}_{g+\lambda r} - \mathbf{z}_g)/\lambda$ with $y_\varepsilon^\lambda \in H^2(D) \cap H_0^1(D)$ being the solution of (3.3)–(3.4) corresponding to $g + \lambda r$, $u + \lambda v$, and $\mathbf{z}_{g+\lambda r} \in \mathcal{C}^1(I_g)^2$ is the solution of (3.5)–(3.7) corresponding to $g + \lambda r$. The limits exist in the above spaces. We denote by “ \cdot ” the scalar product on \mathbb{R}^2 .

PROOF. We subtract the equations corresponding to y_ε^λ and y_ε and divide by $\lambda \neq 0$, small:

$$(3.15) \quad -\Delta \frac{y_\varepsilon^\lambda - y_\varepsilon}{\lambda} = \frac{1}{\lambda} [(g + \lambda r + \varepsilon)_+^2 (u + \lambda v) - (g + \varepsilon)_+^2 u], \quad \text{in } D,$$

with 0 boundary conditions on ∂D . The regularity conditions on \mathcal{F} and $u, v \in L^2(D)$ give the convergence of the right-hand side in (3.15) to the right-hand side in (3.10) (strongly in $L^2(D)$) via some calculations. Then, by elliptic regularity, we have $(y_\varepsilon^\lambda - y_\varepsilon)/\lambda \rightarrow q_\varepsilon$ strongly in $H^2(D) \cap H_0^1(D)$ and (3.10), (3.11) follows.

For (3.12)–(3.14), the argument is the same as in [24, Proposition 6]. The convergence of the ratio $(\mathbf{z}_{g+\lambda r} - \mathbf{z}_g)/\lambda$ is in $\mathcal{C}^1(I_g)^2$ on the whole sequence $\lambda \rightarrow 0$, due to the uniqueness property of the linear system (3.12)–(3.14). Here, we also use Remark 2.5, on the convergence $\mathcal{G}_\lambda \rightarrow \mathcal{G}$ and the continuity with respect to the perturbations of g in the Hamiltonian system (2.1)–(2.3), according to [24]. \square

REMARK 3.4. We have as well imposed the condition

$$(3.16) \quad g(x_1^0, x_2^0) = 0, \quad \text{for all } g \in \mathcal{F},$$

where $(x_1^0, x_2^0) \in D \setminus E$ is some given point. Similarly, constraints like (3.16) may be imposed on a finite number of points or on some curves in $D \setminus E$ and their geometric meaning is that the boundary $\partial\Omega_g$ of the admissible unknown domains should contain these points, curves, etc.

PROPOSITION 3.5. Assume that $f \in L^p(D)$, $j(\mathbf{x}, \cdot)$ is of class $C^1(\mathbb{R})$ and bounded, $g \in \mathcal{F}$, $u \in L^p(D)$, $p > 2$ and $y_\varepsilon(z_g(t)) = 0$ in $[0, T_g]$. Then, for any direction $[r, v] \in \mathcal{F} \times L^p(D)$, the derivative of the penalized cost (3.2) is given by:

$$(3.17) \quad \int_E \partial_2 j(\mathbf{x}, y_\varepsilon(\mathbf{x})) q_\varepsilon(\mathbf{x}) \, d\mathbf{x} + \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) q_\varepsilon(\mathbf{z}_g(t)) |\mathbf{z}'_g(t)| \, dt \\ + \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) \nabla y_\varepsilon(\mathbf{z}_g(t)) \cdot \mathbf{w}(t) |\mathbf{z}'_g(t)| \, dt + \frac{1}{\varepsilon} \int_{I_g} (y_\varepsilon(\mathbf{z}_g(t)))^2 \frac{\mathbf{z}'_g(t) \cdot \mathbf{w}'(t)}{|\mathbf{z}'_g(t)|} \, dt$$

where $q_\varepsilon \in W^{2,p}(D) \cap W_0^{1,p}(D)$, $\mathbf{w} \in C^1(I_g)^2$, $\mathbf{z}_g \in C^1(I_g)^2$ satisfy (3.10)–(3.14) and (2.1)–(2.3) respectively, and $I_g = [0, T_g]$ is the period interval for $\mathbf{z}_g(\cdot)$.

PROOF. In the notations of Proposition 3.3, we compute

$$(3.18) \quad \lim_{\lambda \rightarrow 0} \left\{ \frac{1}{\lambda} \int_E [j(\mathbf{x}, y_\varepsilon^\lambda(\mathbf{x})) - j(\mathbf{x}, y_\varepsilon(\mathbf{x}))] \, d\mathbf{x} \right. \\ \left. + \frac{1}{\varepsilon \lambda} \int_{I_g} [(y_\varepsilon^\lambda(\mathbf{z}_{g+\lambda r}(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)| - (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_g(t)|] \, dt \right\}.$$

In (3.18), $\lambda > 0$ is “small” and Proposition 2.4 ensures that $g + \lambda r \in \mathcal{F}$ (see [25] as well). By Proposition 2.3 we know that the trajectories associated to $g + \lambda h$ are periodic, that is the functions in the second integrals are defined on I_g . Moreover, since $f, u \in L^p(D)$, then $y_\varepsilon^\lambda, y_\varepsilon$ defined as in (3.3), (3.4) are in $W^{2,p}(D) \subset C^1(\overline{D})$, by the Sobolev theorem and elliptic regularity. Consequently, all the integrals appearing in (3.17), (3.18) make sense.

Moreover, in (3.18), we have neglected the term

$$L = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda \varepsilon} \int_{T_g}^{T_{g+\lambda r}} y_\varepsilon^\lambda(\mathbf{z}_{g+\lambda r}(t))^2 |\mathbf{z}'_{g+\lambda r}(t)| \, dt \\ = \lim_{\lambda \rightarrow 0} \frac{1}{\lambda \varepsilon} \int_{T_g}^{T_{g+\lambda r}} [y_\varepsilon^\lambda(\mathbf{z}_{g+\lambda r}(t))^2 |\mathbf{z}'_{g+\lambda r}(t)| - y_\varepsilon(\mathbf{z}_g(t))^2 |\mathbf{z}'_g(t)|] \, dt$$

due to the hypothesis on $y_\varepsilon(\mathbf{z}_g(t))$. We can study term by term:

$$\int_{T_g}^{T_{g+\lambda r}} \left[\frac{y_\varepsilon^\lambda(\mathbf{z}_{g+\lambda r}(t))^2 - y_\varepsilon(\mathbf{z}_{g+\lambda r}(t))^2}{\lambda} |\mathbf{z}'_{g+\lambda r}(t)| \right] \, dt, \\ \int_{T_g}^{T_{g+\lambda r}} \left[\frac{y_\varepsilon(\mathbf{z}_{g+\lambda r}(t))^2 - y_\varepsilon(\mathbf{z}_g(t))^2}{\lambda} |\mathbf{z}'_{g+\lambda r}(t)| \right] \, dt,$$

$$\int_{T_g}^{T_{g+\lambda r}} \left[y_\varepsilon(\mathbf{z}_g(t))^2 \frac{|\mathbf{z}'_{g+\lambda r}(t)| - |\mathbf{z}'_g(t)|}{\lambda} \right] dt.$$

By Proposition 3.3, each of the above three integrands are uniformly bounded and their limits can be easily computed, for instance on $[0, 2T_g]$ due to Proposition 2.6. Notice, in the last term, that $|\mathbf{z}'_g(t)| = |\nabla g(\mathbf{z}_g(t))| \neq 0$ due to (3.5)–(3.7), that is we can differentiate here as well.

Again, by Proposition 2.6 and the above uniform boundedness, we infer that each of the above three terms has null limit as $\lambda \rightarrow 0$, i.e. $L = 0$. Consequently, it is enough to study the limit (3.18).

We also have $y_\varepsilon^\lambda \rightarrow y_\varepsilon$ in $C^1(\overline{D})$, for $\lambda \rightarrow 0$, by elliptic regularity. Then, under the assumptions on $j(\cdot, \cdot)$, we get

$$(3.19) \quad \frac{1}{\lambda} \int_E [j(\mathbf{x}, y_\varepsilon^\lambda(\mathbf{x})) - j(\mathbf{x}, y_\varepsilon(\mathbf{x}))] d\mathbf{x} \rightarrow \int_E \partial_2 j(\mathbf{x}, y_\varepsilon(\mathbf{x})) q_\varepsilon(\mathbf{x}) d\mathbf{x}.$$

For the second integral in (3.18), we intercalate certain intermediary terms and we compute their limits for $\lambda \rightarrow 0$:

$$(3.20) \quad \lim_{\lambda \rightarrow 0} \frac{1}{\varepsilon \lambda} \int_{I_g} [(y_\varepsilon^\lambda(\mathbf{z}_{g+\lambda r}(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)| - (y_\varepsilon(\mathbf{z}_{g+\lambda r}(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)|] dt \\ = \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) q_\varepsilon(\mathbf{z}_g(t)) |\mathbf{z}'_g(t)| dt$$

due to the convergence $\mathbf{z}_{g+\lambda r} \rightarrow \mathbf{z}_g$ in $C^1(I_g)^2$ by $g, r \in C^2(\overline{D})$ and the continuity properties in (2.1)–(2.3);

$$(3.21) \quad \lim_{\lambda \rightarrow 0} \frac{1}{\varepsilon \lambda} \int_{I_g} [(y_\varepsilon(\mathbf{z}_{g+\lambda r}(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)| - (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)|] dt \\ = \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) \nabla y_\varepsilon(\mathbf{z}_g(t)) \cdot \mathbf{w}(t) |\mathbf{z}'_g(t)| dt,$$

where $\mathbf{w} = (w_1, w_2)$ satisfies (3.12)–(3.14) and again we use the regularity and the convergence properties in $C^1(D)$, respectively $C^1(I_g)^2$.

$$(3.22) \quad \lim_{\lambda \rightarrow 0} \frac{1}{\varepsilon \lambda} \int_{I_g} [(y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_{g+\lambda r}(t)| - (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_g(t)|] dt \\ = \frac{1}{\varepsilon} \int_{I_g} (y_\varepsilon(\mathbf{z}_g(t)))^2 \frac{\mathbf{z}'_g(t) \cdot \mathbf{w}'(t)}{|\mathbf{z}'_g(t)|} dt,$$

where we recall that $|\mathbf{z}'_g(t)| = \sqrt{(z_g^1)'(t)^2 + (z_g^2)'(t)^2}$ is non zero by (2.7) and the Hamiltonian system, and standard derivation rules may be applied under our regularity conditions. By summing up (3.19)–(3.22), we end the proof of (3.17). \square

REMARK 3.6. In the case that Ω_g is not simply connected, the penalization integral in (3.2) is in fact a finite sum and each of these terms can be handled separately, in the same way as above, due to Proposition 2.4 and Remark 2.5.

One significance of the hypothesis $y_\varepsilon(\mathbf{z}_g(t)) = 0$, is that we may first minimize the penalization term with respect to the control u (this is possible due to the arguments in the proof of Proposition 3.1). Then, the obtained control may be fixed and the minimization with respect to $g \in \mathcal{F}$ is to be performed. In case $T_{g+\lambda r}$ can be evaluated as in Remark 2.7, then the hypothesis can be relaxed to $y_\varepsilon(x_1^0, x_2^0) = 0$ via a variant of the above arguments.

Now, we denote by $A: \mathcal{C}^2(\bar{D}) \times L^p(D) \rightarrow W^{2,p}(D) \cap W_0^{1,p}(D)$ the linear continuous operator given by $r, v \rightarrow q_\varepsilon$, defined in (3.10), (3.11). We also denote by $B: \mathcal{C}^2(\bar{D}) \rightarrow \mathcal{C}^1(I_g)^2$ the linear continuous operator given by (3.12)–(3.14), $Br = [w_1, w_2]$. In these definitions, $g \in \mathcal{C}^2(\bar{D})$ and $u \in L^p(D)$ are fixed. We have:

COROLLARY 3.7. *The relation (3.17) can be rewritten as:*

$$\begin{aligned}
 (3.23) \quad & \int_E \partial_2 j(\mathbf{x}, y_\varepsilon(\mathbf{x})) A(r, v)(\mathbf{x}) \, d\mathbf{x} + \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) A(r, v)(\mathbf{z}_g(t)) |\mathbf{z}'_g(t)| \, dt \\
 & + \frac{2}{\varepsilon} \int_{I_g} y_\varepsilon(\mathbf{z}_g(t)) \nabla y_\varepsilon(\mathbf{z}_g(t)) \cdot Br(t) |\mathbf{z}'_g(t)| \, dt \\
 & + \frac{1}{\varepsilon} \int_{I_g} \frac{(y_\varepsilon(\mathbf{z}_g(t)))^2}{|\mathbf{z}'_g(t)|} \mathbf{z}'_g(t) \cdot [-\partial_2 r, \partial_1 r](\mathbf{z}_g(t)) \, dt \\
 & + \frac{1}{\varepsilon} \int_{I_g} \frac{(y_\varepsilon(\mathbf{z}_g(t)))^2}{|\mathbf{z}'_g(t)|} C(t) \cdot \mathbf{w}(t) \, dt,
 \end{aligned}$$

where the vector $C(t)$ is explained below.

PROOF. In the last integral in (3.17), we replace $\mathbf{w}'(t)$ by the right-hand side in (3.12), (3.13). We compute:

$$\begin{aligned}
 (3.24) \quad & \mathbf{z}'_g(t) \cdot \mathbf{w}'(t) = \mathbf{z}'_g(t) \cdot [-\nabla \partial_2 g(\mathbf{z}_g(t)) \cdot \mathbf{w}(t) - \partial_2 r(\mathbf{z}_g(t)), \\
 & \quad \quad \quad \nabla \partial_1 g(\mathbf{z}_g(t)) \cdot \mathbf{w}(t) + \partial_1 r(\mathbf{z}_g(t))] \\
 & = \mathbf{z}'_g(t) \cdot [-\partial_2 r(\mathbf{z}_g(t)), \partial_1 r(\mathbf{z}_g(t))] \\
 & \quad + \mathbf{z}'_g(t) \cdot [-\partial_{1,2}^2 g(\mathbf{z}_g(t)) w_1(t), \partial_{1,1}^2 g(\mathbf{z}_g(t)) w_1(t)] \\
 & \quad + \mathbf{z}'_g(t) \cdot [-\partial_{2,2}^2 g(\mathbf{z}_g(t)) w_2(t), \partial_{2,1}^2 g(\mathbf{z}_g(t)) w_2(t)].
 \end{aligned}$$

We denote by $C(t)$ the (known) vector

$$\begin{aligned}
 C(t) = & [- (z_g^1)'(t) \partial_{1,2}^2 g(\mathbf{z}_g(t)) + (z_g^2)'(t) \partial_{1,1}^2 g(\mathbf{z}_g(t)), \\
 & - (z_g^1)'(t) \partial_{2,2}^2 g(\mathbf{z}_g(t)) + (z_g^2)'(t) \partial_{2,1}^2 g(\mathbf{z}_g(t))]
 \end{aligned}$$

and together with (3.24), we get (3.23). □

4. Finite element discretization

We assume that D and E are polygonal. Let \mathcal{T}_h be a triangulation of D with vertices $A_i, i \in I = \{1, \dots, n\}$.

We consider that \mathcal{T}_h is compatible with E , i.e.

$$T \subset \bar{E} \quad \text{or} \quad T \subset \overline{D \setminus E} \quad \text{for all } T \in \mathcal{T}_h,$$

where T designs a triangle of \mathcal{T}_h and h is the size of \mathcal{T}_h . We consider a triangle as a closed set. For simplicity, we employ piecewise linear finite element and we denote

$$\mathbb{W}_h = \{ \varphi_h \in \mathcal{C}(\bar{D}) : \varphi_h|_T \in \mathbb{P}_1(T), \text{ for all } T \in \mathcal{T}_h \}.$$

We use a standard basis of \mathbb{W}_h , $\{ \phi_i \}_{i \in I}$, where ϕ_i is the hat function associated to the vertex A_i , see for example [6], [21]. The finite element approximations of g and u are

$$g_h(\mathbf{x}) = \sum_{i \in I} G_i \phi_i(\mathbf{x}), \quad u_h(\mathbf{x}) = \sum_{i \in I} U_i \phi_i(\mathbf{x}), \quad \text{for all } \mathbf{x} \in \bar{D}.$$

We set the vectors $G = (G_i)_{i \in I}^T \in \mathbb{R}^n$, $U = (U_i)_{i \in I}^T \in \mathbb{R}^n$ and g_h can be identified by G , etc. The function u is in $L^p(D)$, as in Proposition 3.5. Alternatively, for u_h , we can use discontinuous piecewise constant finite element \mathbb{P}_0 . In order to approach $g \in \mathcal{C}^2(\bar{D})$, we can use high order finite elements.

4.1. Discretization of the optimization problem. We introduce

$$\mathbb{V}_h = \{ \varphi_h \in \mathbb{W}_h : \varphi_h = 0 \text{ on } \partial D \},$$

$I_0 = \{ i \in I : A_i \notin \partial D \}$ and $n_0 = \text{card}(I_0)$. The finite element weak formulation of (3.3)–(3.4) is: find $y_h \in \mathbb{V}_h$ such that

$$(4.1) \quad \int_D \nabla y_h \cdot \nabla \varphi_h \, d\mathbf{x} = \int_D (f + (g_h + \varepsilon)_+^2 u_h) \varphi_h \, d\mathbf{x}, \quad \text{for all } \varphi_h \in \mathbb{V}_h.$$

As before, for $y_h(\mathbf{x}) = \sum_{j \in I_0} Y_j \phi_j(\mathbf{x})$, we set $Y = (Y_j)_{j \in I_0}^T \in \mathbb{R}^{n_0}$. In order to obtain the linear system, we take the basis functions $\varphi_h = \phi_i$ in (4.1) for $i \in I_0$. Let us consider the vector

$$F = \left(\int_D f \phi_i \, d\mathbf{x} \right)_{i \in I_0}^T \in \mathbb{R}^{n_0},$$

the $n_0 \times n_0$ matrix K defined by

$$K = (K_{ij})_{i \in I_0, j \in I_0}, \quad K_{ij} = \int_D \nabla \phi_j \cdot \nabla \phi_i \, d\mathbf{x}$$

and the $n_0 \times n$ matrix $B^1(G, \varepsilon)$ defined by

$$B^1(G, \varepsilon) = (B_{ij}^1)_{i \in I_0, j \in I}, \quad B_{ij}^1 = \int_D (g_h + \varepsilon)_+^2 \phi_j \phi_i \, d\mathbf{x}.$$

The matrix K is symmetric, positive definite. The finite element approximation of the state system (3.3)–(3.4) is the linear system:

$$(4.2) \quad KY = F + B^1(G, \varepsilon)U.$$

Now, we shall discretize the objective function (3.2). We denote $I_E = \{i \in I : A_i \in \overline{E}\}$ and $n_E = \text{card}(I_E)$. For the first term of (3.2), we introduce

$$J_1(Y) = \int_E j(\mathbf{x}, y_h(\mathbf{x})) \, d\mathbf{x}.$$

We shall study the second term of (3.2). In order to solve numerically the ODE system (3.5)–(3.7), we use a partition $[t_0, \dots, t_k, \dots, t_m]$ of $[0, T_g]$, with $t_0 = 0$ and $t_m = T_g$. We can use the forward Euler scheme:

$$(4.3) \quad Z_{k+1}^1 = Z_k^1 - (t_{k+1} - t_k) \frac{\partial g_h}{\partial x_2}(Z_k^1, Z_k^2),$$

$$(4.4) \quad Z_{k+1}^2 = Z_k^2 + (t_{k+1} - t_k) \frac{\partial g_h}{\partial x_1}(Z_k^1, Z_k^2),$$

$$(4.5) \quad (Z_0^1, Z_0^2) = (x_1^0, x_2^0),$$

for $k = 0, \dots, m - 2$. We set $Z_k = (Z_k^1, Z_k^2)$ and we impose $Z_m = Z_0$. In fact, Z_k is an approximation of $\mathbf{z}_g(t_k)$. We do not need to stock Z_0 and we set $Z = (Z^1, Z^2) \in \mathbb{R}^m \times \mathbb{R}^m$, with $Z^1 = (Z_k^1)_{1 \leq k \leq m}^T$ and $Z^2 = (Z_k^2)_{1 \leq k \leq m}^T$. In the applications, one can use more performant numerical methods for the ODE's, like explicit Runge–Kutta or backward Euler, but here we want to avoid a too tedious exposition.

Without risk of confusion, we introduce the function $Z: [0, T_g] \rightarrow \mathbb{R}^2$ defined by

$$Z(t) = \frac{t_{k+1} - t}{(t_{k+1} - t_k)} Z_k + \frac{t - t_k}{(t_{k+1} - t_k)} Z_{k+1}, \quad \text{for } t_k \leq t < t_{k+1}$$

for $k = 0, 1, \dots, m - 1$. We have $Z(t_k) = Z_k$ and we can identify the function $Z(\cdot)$ by the vector $Z \in \mathbb{R}^m \times \mathbb{R}^m$. We remark that $Z(\cdot)$ is derivable on each interval (t_k, t_{k+1}) and

$$Z'(t) = \frac{1}{(t_{k+1} - t_k)} (Z_{k+1}^1 - Z_k^1, Z_{k+1}^2 - Z_k^2) \quad \text{for } t_k \leq t < t_{k+1}.$$

We introduce the $n_0 \times n_0$ matrix $N(Z)$ defined by

$$N(Z) = \left(\int_0^{T_g} \phi_j(Z(t)) \phi_i(Z(t)) |Z'(t)| \, dt \right)_{i \in I_0, j \in I_0}$$

and the second term of (3.2) is approached by $Y^T N(Z) Y / \varepsilon$, then the discrete form of the optimization problem (3.2)–(3.7) is

$$(4.6) \quad \min_{G, U \in \mathbb{R}^n} J(G, U) = J_1(Y) + \frac{1}{\varepsilon} Y^T N(Z) Y$$

subject to (4.2). We point out that Y depends on G and U by (4.2) and Z depends on G by (4.3)–(4.5).

4.2. Discretization of the derivative of the objective function. Let r_h, v_h be in \mathbb{W}_h and $R, V \in \mathbb{R}^n$ be the associated vectors. The finite element weak formulation of (3.10)–(3.11) is: find $q_h \in \mathbb{V}_h$ such that

$$(4.7) \quad \int_D \nabla q_h \cdot \nabla \varphi_h \, d\mathbf{x} = \int_D ((g_h + \varepsilon)_+^2 v_h + 2(g_h + \varepsilon)_+ u_h r_h) \varphi_h \, d\mathbf{x},$$

for all $\varphi_h \in \mathbb{V}_h$. Let $Q \in \mathbb{R}^{n_0}$ be the associated vector to q_h and we construct the $n_0 \times n$ matrix $C^1(G, \varepsilon, U)$ defined by

$$C^1(G, \varepsilon, U) = \left(\int_D 2(g_h + \varepsilon)_+ u_h \phi_j \phi_i \, d\mathbf{x} \right)_{i \in I_0, j \in I}.$$

The linear system of (4.7) is

$$(4.8) \quad KQ = B^1(G, \varepsilon)V + C^1(G, \varepsilon, U)R.$$

In order to approximate $\partial_2 j(\mathbf{x}, y_\varepsilon(\mathbf{x}))$, y_ε given by (3.3)–(3.4), we consider the nonlinear application $Y \in \mathbb{R}^{n_0} \rightarrow L(Y) \in \mathbb{R}^{n_E}$ such that

$$\partial_2 j(\mathbf{x}, y_h(\mathbf{x})) = \sum_{i \in I_E} (L(Y))_i \phi_{i|E}(\mathbf{x})$$

where $\phi_{i|E}$ is the restriction of ϕ_i to E . We define the $n_E \times n_0$ matrix M_{ED} defined by

$$M_{ED} = \left(\int_D \phi_i \phi_j \, d\mathbf{x} \right)_{i \in I_E, j \in I_0}.$$

The first term of (3.17) is approached by

$$(4.9) \quad (L(Y))^T M_{ED} Q$$

and the second term of (3.17) is approached by

$$(4.10) \quad \frac{2}{\varepsilon} Y^T N(Z) Q$$

where the matrix $N(Z)$ was introduced in the previous subsection.

Next, we introduce the partial derivative for a piecewise linear function. Let $g_h \in \mathbb{W}_h$ and $G \in \mathbb{R}^n$ its associated vector, i.e. $g_h(\mathbf{x}) = \sum_{i \in I} G_i \phi_i(\mathbf{x})$. Let

$\Pi_h^1 G \in \mathbb{R}^n$ defined by

$$(\Pi_h^1 G)_i = \frac{1}{\sum_{j \in J_i} \text{area}(T_j)} \sum_{j \in J_i} \text{area}(T_j) \partial_1 g_h|_{T_j}$$

where J_i is the set of index j such that the triangle T_j has the vertex A_i . Since g_h is a linear function in each triangle T_j , then $\partial_1 g_h|_{T_j}$ is constant. Similarly, we construct $\Pi_h^2 G \in \mathbb{R}^n$ for ∂_2 . In fact, Π_h^1 and Π_h^2 are two $n \times n$ matrices depending on \mathcal{T}_h . Then, we set

$$\partial_1^h g_h(\mathbf{x}) = \sum_{i \in I} (\Pi_h^1 G)_i \phi_i(\mathbf{x})$$

and similarly for $\partial_2^h g_h$. Finally, we put $\nabla_h g_h = (\partial_1^h g_h, \partial_2^h g_h)$. Since $y_h \in \mathbb{V}_h \subset \mathbb{W}_h$, we can define $\partial_1^h y_h$ and $\partial_2^h y_h$.

EXAMPLE 4.1. We shall give a simple example to understand the discrete derivative of \mathbb{W}_h functions. We consider the square $[A_1 A_2 A_4 A_3]$ of vertices $A_1 = (0, 0)$, $A_2 = (1, 0)$, $A_4 = (1, 1)$, $A_3 = (0, 1)$ and the triangulation of two triangles $T_1 = [A_1 A_2 A_4]$ and $T_2 = [A_1 A_4 A_3]$. We shall present the discrete derivative of the hat function

$$\phi_4(x_1, x_2) = \begin{cases} x_2 & \text{in } T_1, \\ x_1 & \text{in } T_2. \end{cases}$$

We have $J_1 = \{1, 2\}$ and

$$\begin{aligned} (\Pi_h^1 \phi_4)_1 &= \frac{1}{\text{area}(T_1) + \text{area}(T_2)} (\text{area}(T_1) \partial_1 \phi_4|_{T_1} + \text{area}(T_2) \partial_1 \phi_4|_{T_2}) \\ &= \frac{1}{1/2 + 1/2} (1/2 \times 0 + 1/2 \times 1) = 1/2. \end{aligned}$$

Similarly, $J_2 = \{1\}$, $J_3 = \{2\}$, $J_4 = \{1, 2\}$,

$$(\Pi_h^1 \phi_4)_2 = \frac{1}{1/2} (1/2 \times 0) = 0,$$

$$(\Pi_h^1 \phi_4)_3 = \frac{1}{1/2} (1/2 \times 1) = 1/2,$$

$$(\Pi_h^1 \phi_4)_4 = \frac{1}{1/2 + 1/2} (1/2 \times 0 + 1/2 \times 1) = 1/2$$

then

$$\partial_1^h \phi_4(x_1, x_2) = 1/2 \times \phi_1(x_1, x_2) + 0 \times \phi_2(x_1, x_2) + 1 \times \phi_3(x_1, x_2) + 1/2 \times \phi_4(x_1, x_2).$$

In order to solve the ODE system (3.12)–(3.14), we use the forward Euler scheme on the same partition as for (4.3)–(4.5):

$$(4.11) \quad W_{k+1}^1 = W_k^1 - (t_{k+1} - t_k) \nabla_h \partial_2^h g_h(Z_k) \cdot (W_k^1, W_k^2) - (t_{k+1} - t_k) \partial_2^h r_h(Z_k),$$

$$(4.12) \quad W_{k+1}^2 = W_k^2 + (t_{k+1} - t_k) \nabla_h \partial_1^h g_h(Z_k) \cdot (W_k^1, W_k^2) + (t_{k+1} - t_k) \partial_1^h r_h(Z_k),$$

$$(4.13) \quad W_0^1 = 0, \quad W_0^2 = 0,$$

for $k = 0, \dots, m - 1$. We set $W_k = (W_k^1, W_k^2)$ and now we have $W_m \neq W_0$ generally. In fact, W_k is an approximation of $\mathbf{w}(t_k)$. We do not need to stock W_0 and we set $W = (W^1, W^2) \in \mathbb{R}^m \times \mathbb{R}^m$, with $W^1 = (W_k^1)_{1 \leq k \leq m}^T$ and $W^2 = (W_k^2)_{1 \leq k \leq m}^T$. As mentioned before, we can use more performant numerical methods for the ODE, like the explicit Runge–Kutta or the backward Euler procedures.

We construct $W : [0, T_g] \rightarrow \mathbb{R}^2$ in the same way as for $Z(t)$

$$W(t) = \frac{t_{k+1} - t}{(t_{k+1} - t_k)} W_k + \frac{t - t_k}{(t_{k+1} - t_k)} W_{k+1}, \quad t_k \leq t < t_{k+1}$$

for $k = 0, 1, \dots, m - 1$. We have

$$W(t_k) = W_k \quad \text{and} \quad W'(t) = \frac{1}{(t_{k+1} - t_k)} (W_{k+1}^1 - W_k^1, W_{k+1}^2 - W_k^2)$$

for $t_k \leq t < t_{k+1}$. If ψ_k is the one-dimensional piecewise linear hat function associated to the point t_k of the partition $[t_0, \dots, t_k, \dots, t_m]$, we can write equivalently

$$W(t) = \sum_{k=0}^m W_k \psi_k(t) \quad \text{for } t \in [0, T_g].$$

The third term of (3.17) is approached by

$$\begin{aligned} (4.14) \quad & \frac{2}{\varepsilon} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} y_h(Z(t)) \nabla_h y_h(Z(t)) \cdot W(t) |Z'(t)| dt \\ &= \frac{2}{\varepsilon} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} y_h(Z(t)) \nabla_h y_h(Z(t)) \\ & \quad \cdot (W_k \psi_k(t) + W_{k+1} \psi_{k+1}(t)) |Z'(t)| dt \\ &= \frac{2}{\varepsilon} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} y_h(Z(t)) \partial_1^h y_h(Z(t)) \\ & \quad \cdot (W_k^1 \psi_k(t) + W_{k+1}^1 \psi_{k+1}(t)) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \\ & \quad + \frac{2}{\varepsilon} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} y_h(Z(t)) \partial_2^h y_h(Z(t)) \\ & \quad \cdot (W_k^2 \psi_k(t) + W_{k+1}^2 \psi_{k+1}(t)) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \end{aligned}$$

where $|Z_k Z_{k+1}|$ is the length of the segment in \mathbb{R}^2 with ends Z_k and Z_{k+1} . We have

$$\begin{aligned} (4.15) \quad & \int_{t_k}^{t_{k+1}} y_h(Z(t)) \partial_1^h y_h(Z(t)) (W_k^1 \psi_k(t) + W_{k+1}^1 \psi_{k+1}(t)) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \\ &= W_k^1 \int_{t_k}^{t_{k+1}} \left(\sum_{i \in I_0} Y_i \phi_i(Z(t)) \right) \\ & \quad \cdot \left(\sum_{j \in I} (\Pi_h^1 Y)_j \phi_j(Z(t)) \right) \psi_k(t) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \end{aligned}$$

$$\begin{aligned}
 &+ W_{k+1}^1 \int_{t_k}^{t_{k+1}} \left(\sum_{i \in I_0} Y_i \phi_i(Z(t)) \right) \\
 &\quad \cdot \left(\sum_{j \in I} (\Pi_h^1 Y)_j \phi_j(Z(t)) \right) \psi_{k+1}(t) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt.
 \end{aligned}$$

We introduce the $n_0 \times n$ matrices $N_k^{[k,k+1]}(Z)$ and $N_{k+1}^{[k,k+1]}(Z)$ defined by

$$\begin{aligned}
 N_k^{[k,k+1]}(Z) &= \left(\int_{t_k}^{t_{k+1}} \phi_i(Z(t)) \phi_j(Z(t)) \psi_k(t) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \right)_{i \in I_0, j \in I}, \\
 N_{k+1}^{[k,k+1]}(Z) &= \left(\int_{t_k}^{t_{k+1}} \phi_i(Z(t)) \phi_j(Z(t)) \psi_{k+1}(t) \frac{|Z_k Z_{k+1}|}{(t_{k+1} - t_k)} dt \right)_{i \in I_0, j \in I},
 \end{aligned}$$

then (4.15) can be rewritten as

$$Y^T (W_k^1 N_k^{[k,k+1]}(Z) + W_{k+1}^1 N_{k+1}^{[k,k+1]}(Z)) (\Pi_h^1 Y)$$

and finally, the third term of (3.17) is approached by

$$\begin{aligned}
 (4.16) \quad &\frac{2}{\varepsilon} Y^T \sum_{k=0}^{m-1} (W_k^1 N_k^{[k,k+1]}(Z) + W_{k+1}^1 N_{k+1}^{[k,k+1]}(Z)) (\Pi_h^1 Y) \\
 &+ \frac{2}{\varepsilon} Y^T \sum_{k=0}^{m-1} (W_k^2 N_k^{[k,k+1]}(Z) + W_{k+1}^2 N_{k+1}^{[k,k+1]}(Z)) (\Pi_h^2 Y).
 \end{aligned}$$

We can introduce the linear operators $T^1(Z)$ and $T^2(Z)$ by

$$\begin{aligned}
 (4.17) \quad &W^1 \in \mathbb{R}^m \rightarrow T^1(Z)W^1 = \sum_{k=0}^{m-1} (W_k^1 N_k^{[k,k+1]}(Z) + W_{k+1}^1 N_{k+1}^{[k,k+1]}(Z)), \\
 &W^2 \in \mathbb{R}^m \rightarrow T^2(Z)W^2 = \sum_{k=0}^{m-1} (W_k^2 N_k^{[k,k+1]}(Z) + W_{k+1}^2 N_{k+1}^{[k,k+1]}(Z)),
 \end{aligned}$$

then (4.16) can be rewritten as

$$(4.18) \quad \frac{2}{\varepsilon} Y^T (T^1(Z)W^1) (\Pi_h^1 Y) + \frac{2}{\varepsilon} Y^T (T^2(Z)W^2) (\Pi_h^2 Y).$$

The fourth term of term of (3.17) is approached by

$$(4.19) \quad \frac{1}{\varepsilon} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left(\sum_{i \in I_0} Y_i \phi_i(Z(t)) \right) \left(\sum_{j \in I_0} Y_j \phi_j(Z(t)) \right) \frac{Z'(t) \cdot W'(t)}{|Z'(t)|} dt.$$

But $Z'(t)$ and $W'(t)$ are constants for $t_k \leq t < t_{k+1}$, then

$$\frac{Z'(t) \cdot W'(t)}{|Z'(t)|} = \frac{(Z_{k+1}^1 - Z_k^1, Z_{k+1}^2 - Z_k^2) \cdot (W_{k+1}^1 - W_k^1, W_{k+1}^2 - W_k^2)}{(t_{k+1} - t_k) |Z_k Z_{k+1}|}$$

where $|Z_k Z_{k+1}|$ is the length of the segment in \mathbb{R}^2 with ends Z_k and Z_{k+1} . We introduce the $n_0 \times n_0$ matrix $R_k(Z)$ defined by

$$R_k(Z) = \left(\int_{t_k}^{t_{k+1}} \phi_i(Z(t))\phi_j(Z(t))|Z'(t)| dt \right)_{i \in I_0, j \in I_0}$$

and the linear operators $T^3(Z)$

$$(4.20) \quad W \in \mathbb{R}^m \times \mathbb{R}^m \rightarrow T^3(Z)W,$$

$$T^3(Z)W = \sum_{k=0}^{m-1} \frac{(Z_{k+1}^1 - Z_k^1, Z_{k+1}^2 - Z_k^2) \cdot (W_{k+1}^1 - W_k^1, W_{k+1}^2 - W_k^2)}{|Z_k Z_{k+1}|^2} R_k(Z).$$

The (4.19) can be rewritten as

$$(4.21) \quad \frac{1}{\varepsilon} Y^T (T^3(Z)W)Y.$$

The study of this subsection can be resumed as follows:

PROPOSITION 4.2. *The discret version of (3.17) is*

$$(4.22) \quad dJ_{(G,U)}(R, V) = (L(Y))^T M_{ED}Q + \frac{2}{\varepsilon} Y^T N(Z)Q$$

$$+ \frac{2}{\varepsilon} Y^T (T^1(Z)W^1) (\Pi_h^1 Y) + \frac{2}{\varepsilon} Y^T (T^2(Z)W^2) (\Pi_h^2 Y) + \frac{1}{\varepsilon} Y^T (T^3(Z)W)Y,$$

which represents the derivative of J at (G, U) in the direction (R, V) .

PROOF. We get (4.22) just by assembling (4.9), (4.10), (4.18) and (4.21). \square

4.3. Discretization of the formula (3.23). From (4.8), we get

$$Q = K^{-1}B^1(G, \varepsilon)V + K^{-1}C^1(G, \varepsilon, U)R$$

and the discrete version of the operator A in the Corollary 3.7 is

$$(R, V) \in \mathbb{R}^n \times \mathbb{R}^n \rightarrow A^1(R, V) = K^{-1}B^1(G, \varepsilon)V + K^{-1}C^1(G, \varepsilon, U)R.$$

Replacing Q in the first two terms of (4.22), we get

$$(4.23) \quad \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1}B^1(G, \varepsilon)V$$

$$+ \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1}C^1(G, \varepsilon, U)R.$$

We denote

$$\Lambda_1(t) = y_h(Z(t))\nabla y_h(Z(t))|Z'(t)|,$$

$$\Lambda_2(t) = \frac{(y_h(Z(t)))^2}{|Z'(t)|} Z'(t),$$

$$\Lambda_3(t) = \frac{(y_h(Z(t)))^2}{|Z'(t)|} C(t).$$

The third term of (3.23) is approached by

$$\frac{2}{\varepsilon} \int_0^{T_g} \Lambda_1(t) \cdot W(t) dt$$

and using the trapezoidal quadrature formula on each sub-interval $[t_k, t_{k+1}]$, we get

$$(4.24) \quad \frac{1}{\varepsilon} \sum_{k=0}^{m-1} (t_{k+1} - t_k) [\Lambda_1(t_k) \cdot (W_k^1, W_k^2) + \Lambda_1(t_{k+1}) \cdot (W_{k+1}^1, W_{k+1}^2)].$$

Similarly, for the 4th and 5th terms of (3.23), we get

$$(4.25) \quad \frac{1}{2\varepsilon} \sum_{k=0}^{m-1} (t_{k+1} - t_k) [\Lambda_2(t_k) \cdot [-\partial_2^h r_h, \partial_1^h r_h](Z_k) \\ + \Lambda_2(t_{k+1}) \cdot [-\partial_2^h r_h, \partial_1^h r_h](Z_{k+1})]$$

and

$$(4.26) \quad \frac{1}{2\varepsilon} \sum_{k=0}^{m-1} (t_{k+1} - t_k) [\Lambda_3(t_k) \cdot (W_k^1, W_k^2) + \Lambda_3(t_{k+1}) \cdot (W_{k+1}^1, W_{k+1}^2)].$$

In order to write (4.24)–(4.26) shorter, we introduce the vectors:

- $\tilde{\Lambda}_1^1 \in \mathbb{R}^m$ with first components $(t_{k+1} - t_{k-1})\Lambda_1^1(t_k)$, $1 \leq k \leq m-1$ and the last component $(t_m - t_{m-1})\Lambda_1^1(t_m)$,
- $\tilde{\Lambda}_1^2 \in \mathbb{R}^m$ with first components $(t_{k+1} - t_{k-1})\Lambda_1^2(t_k)$, $1 \leq k \leq m-1$ and the last component $(t_m - t_{m-1})\Lambda_1^2(t_m)$,
- $\tilde{\Lambda}_3^1 \in \mathbb{R}^m$ with first components $(t_{k+1} - t_{k-1})\Lambda_3^1(t_k)/2$, $1 \leq k \leq m-1$ and the last component $(t_m - t_{m-1})\Lambda_3^1(t_m)/2$,
- $\tilde{\Lambda}_3^2 \in \mathbb{R}^m$ with first components $(t_{k+1} - t_{k-1})\Lambda_3^2(t_k)/2$, $1 \leq k \leq m-1$ and the last component $(t_m - t_{m-1})\Lambda_3^2(t_m)/2$. Also, we introduce the vectors in \mathbb{R}^n :

$$\tilde{\Lambda}_2^1 = \frac{1}{2} \sum_{0 \leq k \leq m-1} (t_{k+1} - t_k) (\Lambda_2^1(t_k)\Phi(Z_k) + \Lambda_2^1(t_{k+1})\Phi(Z_{k+1})),$$

$$\tilde{\Lambda}_2^2 = \frac{1}{2} \sum_{0 \leq k \leq m-1} (t_{k+1} - t_k) (\Lambda_2^2(t_k)\Phi(Z_k) + \Lambda_2^2(t_{k+1})\Phi(Z_{k+1})),$$

where $\Phi(Z_k) = (\phi_i(Z_k))_{i \in I}^T \in \mathbb{R}^n$.

PROPOSITION 4.3. *The discrete version of the (3.23) is*

$$(4.27) \quad dJ_{(G,U)}(R, V) = \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} B^1(G, \varepsilon) V \\ + \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} C^1(G, \varepsilon, U) R \\ + \frac{1}{\varepsilon} ((\tilde{\Lambda}_1^1)^T W^1 + (\tilde{\Lambda}_1^2)^T W^2)$$

$$\begin{aligned}
 & + \frac{1}{\varepsilon} \left(-(\tilde{\Lambda}_2^1)^T (\Pi_h^2 R) + (\tilde{\Lambda}_2^2)^T (\Pi_h^1 R) \right) \\
 & + \frac{1}{\varepsilon} \left((\tilde{\Lambda}_3^1)^T W^1 + (\tilde{\Lambda}_3^2)^T W^2 \right).
 \end{aligned}$$

PROOF. We obtain (4.27) by summing (4.23)–(4.26). □

Next, we give more details about the relationship between W and R . Let us introduce the 2×2 matrices

$$\begin{aligned}
 M_2(k) &= \begin{pmatrix} 1 - (t_{k+1} - t_k) \partial_1^h \partial_2^h g_h(Z_k) & -(t_{k+1} - t_k) \partial_2^h \partial_2^h g_h(Z_k) \\ (t_{k+1} - t_k) \partial_1^h \partial_1^h g_h(Z_k) & 1 + (t_{k+1} - t_k) \partial_2^h \partial_1^h g_h(Z_k) \end{pmatrix}, \\
 I_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
 \end{aligned}$$

and the $2 \times n$ matrix

$$N_2(k) = \begin{pmatrix} -(t_{k+1} - t_k) \Phi^T(Z_k) \Pi_h^2 \\ (t_{k+1} - t_k) \Phi^T(Z_k) \Pi_h^1 \end{pmatrix}.$$

We remark that M_2 depends on G and Z and N_2 on Z . The system (4.11)–(4.12) can be written as

$$\begin{pmatrix} W_{k+1}^1 \\ W_{k+1}^2 \end{pmatrix} = M_2(k) \begin{pmatrix} W_k^1 \\ W_k^2 \end{pmatrix} + N_2(k) R.$$

PROPOSITION 4.4. *We have the following equality*

$$(4.28) \quad \begin{pmatrix} W_1^1 \\ W_1^2 \\ \vdots \\ W_m^1 \\ W_m^2 \end{pmatrix} = M_{2m} \times \begin{pmatrix} N_2(0) \\ \vdots \\ N_2(m-1) \end{pmatrix} R$$

where at the right-hand side, M_{2m} is a $2m \times 2m$ matrix defined by

$$\begin{pmatrix} I_2 & 0 & \dots & 0 & 0 \\ M_2(1) & I_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ M_2(m-1) \dots M_2(1) & M_2(m-1) \dots M_2(2) & \dots & M_2(m-1) & I_2 \end{pmatrix}$$

and the size of the second matrix, which contains N_2 , is $2m \times n$.

PROOF. From (4.13) and the recurrent relation, we have

$$\begin{aligned} \begin{pmatrix} W_1^1 \\ W_1^2 \end{pmatrix} &= N_2(0)R, \\ \begin{pmatrix} W_2^1 \\ W_2^2 \end{pmatrix} &= M_2(1) \begin{pmatrix} W_1^1 \\ W_1^2 \end{pmatrix} + N_2(1)R = M_2(1)N_2(0)R + N_2(1)R, \\ &\dots\dots\dots \\ \begin{pmatrix} W_{m-1}^1 \\ W_{m-1}^2 \end{pmatrix} &= M_2(m-2) \dots M_2(1)N_2(0)R + M_2(m-2) \dots M_2(2)N_2(1)R \\ &\quad + \dots + M_2(m-2)N_2(m-3)R + N_2(m-2)R, \\ \begin{pmatrix} W_m^1 \\ W_m^2 \end{pmatrix} &= M_2(m-1) \begin{pmatrix} W_{m-1}^1 \\ W_{m-1}^2 \end{pmatrix} + N_2(m-1)R \\ &= M_2(m-1)M_2(m-2) \dots M_2(1)N_2(0)R \\ &\quad + M_2(m-1)M_2(m-2) \dots M_2(2)N_2(1)R \\ &\quad + \dots + M_2(m-1)M_2(m-2)N_2(m-3)R \\ &\quad + M_2(m-1)N_2(m-2)R + N_2(m-1)R, \end{aligned}$$

which gives (4.28). □

Since W depends on R by (4.28), we can introduce the linear operator approximation of B in the Corollary 3.7:

$$R \in \mathbb{R}^n \rightarrow W = (W^1, W^2) = (B^2(G, Z)R, B^3(G, Z)R) \in \mathbb{R}^m \times \mathbb{R}^m.$$

If we denote by ℓ_i the i -th row of the matrix M_{2m} at the right-hand side of (4.28), for $1 \leq i \leq 2m$, then

$$\begin{aligned} B^2(G, Z) &= \begin{pmatrix} \ell_1 \\ \ell_3 \\ \vdots \\ \ell_{2m-1} \end{pmatrix} \begin{pmatrix} N_2(0) \\ N_2(1) \\ \vdots \\ N_2(m-1) \end{pmatrix}, \\ B^3(G, Z) &= \begin{pmatrix} \ell_2 \\ \ell_4 \\ \vdots \\ \ell_{2m} \end{pmatrix} \begin{pmatrix} N_2(0) \\ N_2(1) \\ \vdots \\ N_2(m-1) \end{pmatrix} \end{aligned}$$

and $B^2(G, Z)$, $B^3(G, Z)$ are $m \times n$ matrices. The size of the matrix containing N_2 is $2m \times n$.

4.4. Gradient type algorithm. We start by presenting the algorithm.

STEP 1. Start with $k = 0$, $\varepsilon > 0$ some given “small” parameter and select some initial (G^k, U^k) .

STEP 2. Compute Y^k the solution of (4.2) and Z^k solution of (4.3)–(4.5).

STEP 3. Find (R^k, V^k) such that $dJ_{(G^k, U^k)}(R^k, V^k) < 0$. We say that (R^k, V^k) is a descent direction.

STEP 4. Define $(G^{k+1}, U^{k+1}) = (G^k, U^k) + \lambda_k(R^k, V^k)$, where $\lambda_k > 0$ is obtained via some line search

$$\lambda_k \in \arg \min_{\lambda > 0} J((G^k, U^k) + \lambda(R^k, V^k)).$$

STEP 5. If $|J(G^{k+1}, U^{k+1}) - J(G^k, U^k)|$ is below some prescribed tolerance parameter, then STOP. If not, update $k := k + 1$ and go to STEP 3.

In the STEP 3, we have to provide a descent direction.

We present in the following a partial result. Let us introduce a simplified adjoint system: find $p_h \in \mathbb{V}_h$ such that

$$(4.29) \quad \int_D \nabla \varphi_h \cdot \nabla p_h \, d\mathbf{x} = \int_E \partial_2 j(\mathbf{x}, y_h(\mathbf{x})) \varphi_h \, d\mathbf{x} + \frac{2}{\varepsilon} \int_0^{T_g} y_h(Z(t)) \varphi_h(Z(t)) |Z'(t)| \, dt,$$

for all $\varphi_h \in \mathbb{V}_h$ and with $Z(t)$ given by (4.3)–(4.5). We have

$$p_h(\mathbf{x}) = \sum_{i \in I_0} P_i \phi_i(\mathbf{x}) \quad \text{and} \quad P = (P_i)_{i \in I_0}^T \in \mathbb{R}^{n_0}.$$

The linear system associated to (4.29) is

$$KP = M_{ED}^T L(Y) + \frac{2}{\varepsilon} N(Z)Y.$$

We recall that K and $N(Z)$ are symmetric matrices.

PROPOSITION 4.5. *Given $g_h, u_h \in \mathbb{W}_h$, let $y_h \in \mathbb{V}_h$ the solution of (4.1). If $r_h = -p_h u_h$ and $v_h = -p_h$, where $p_h \in \mathbb{V}_h$ is the solution of (4.29), then*

$$(4.30) \quad \int_E \partial_2 j(\mathbf{x}, y_h(\mathbf{x})) q_h \, d\mathbf{x} + \frac{2}{\varepsilon} \int_0^{T_g} y_h(Z(t)) q_h(Z(t)) |Z'(t)| \, dt \leq 0,$$

where $q_h \in \mathbb{V}_h$ is the solution of (4.7) depending on r_h and v_h .

PROOF. Putting $\varphi_h = p_h$ in (4.7) and $\varphi_h = q_h$ in (4.29), we get

$$(4.31) \quad \int_D ((g_h + \varepsilon)_+^2 v_h + 2(g_h + \varepsilon)_+ u_h r_h) p_h \, d\mathbf{x} = \int_D \nabla q_h \cdot \nabla p_h \, d\mathbf{x} = \int_E \partial_2 j(\mathbf{x}, y_h(\mathbf{x})) q_h \, d\mathbf{x} + \frac{2}{\varepsilon} \int_0^{T_g} y_h(Z(t)) q_h(Z(t)) |Z'(t)| \, dt.$$

For $v_h = -p_h$, we have

$$\int_D (g_h + \varepsilon)_+^2 v_h p_h \, d\mathbf{x} = - \int_D (g_h + \varepsilon)_+^2 p_h^2 \, d\mathbf{x} \leq 0$$

and, for $r_h = -p_h u_h$, we have

$$\int_D 2(g_h + \varepsilon)_+ u_h r_h p_h \, d\mathbf{x} = - \int_D 2(g_h + \varepsilon)_+ (u_h p_h)^2 \, d\mathbf{x} \leq 0$$

since $(g_h + \varepsilon)_+ \geq 0$ in D . \square

REMARK 4.6. The left-hand side of (4.30) represents the first two terms of (4.22). We can obtain a similar result as in Proposition 4.5, without using the adjoint system, by taking

$$\begin{aligned} (V^*)^T &= - \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} B^1(G, \varepsilon), \\ (R^*)^T &= - \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} C^1(G, \varepsilon, U), \end{aligned}$$

in place of V and R in (4.23). In this case, (4.23) becomes

$$-\|V^*\|_{\mathbb{R}^n}^2 - \|R^*\|_{\mathbb{R}^n}^2 \leq 0.$$

We point out that $(V^*)^T = -P^T B^1(G, \varepsilon)$ and $(R^*)^T = -P^T C^1(G, \varepsilon, U)$, so (V^*, R^*) is different from the direction given by Proposition 4.5.

Now, we present a descent direction, obtained from the complete gradient of the discrete cost (4.22).

PROPOSITION 4.7. For $(R^{**}, V^*) \in \mathbb{R}^n \times \mathbb{R}^n$ given by

$$\begin{aligned} (V^*)^T &= - \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} B^1(G, \varepsilon) \\ (R^{**})^T &= - \left((L(Y))^T M_{ED} + \frac{2}{\varepsilon} Y^T N(Z) \right) K^{-1} C^1(G, \varepsilon, U) \\ &\quad - \frac{1}{\varepsilon} \left((\tilde{\Lambda}_1^1)^T B^2(G, Z) + (\tilde{\Lambda}_1^2)^T B^3(G, Z) \right) \\ &\quad - \frac{1}{\varepsilon} \left(-(\tilde{\Lambda}_2^1)^T \Pi_h^2 + (\tilde{\Lambda}_2^2)^T \Pi_h^1 \right) \\ &\quad - \frac{1}{\varepsilon} \left((\tilde{\Lambda}_3^1)^T B^2(G, Z) + (\tilde{\Lambda}_3^2)^T B^3(G, Z) \right), \end{aligned}$$

we obtain a descent direction for J at (G, U) .

PROOF. In (4.27), we replace W^1 by $B^2(G, Z)R$ and W^2 by $B^3(G, Z)R$, we obtain that

$$dJ_{(G,U)}(R, V) = -(V^*)^T V - (R^{**})^T R,$$

then

$$dJ_{(G,U)}(R^{**}, V^*) = -\|V^*\|_{\mathbb{R}^n}^2 - \|R^{**}\|_{\mathbb{R}^n}^2 \leq 0. \quad \square$$

5. Numerical tests

Shape optimization problems and their penalization are strongly nonconvex. The computed optimal domain depends on the starting domain, but also on the penalization ε or other numerical parameters. It may be just a local optimal solution.

Moreover, the final computed value of the penalization integral is small, but not null. This allows differences between the optimal computed domain Ω_g and the zero level curves of the computed optimal state y_ε . Consequently, we compare the obtained optimal cost in the penalized problem with the costs in the original problem (1.1)–(1.3) corresponding to the optimal computed domain Ω_g and the zero level curves of y_ε . This is a standard procedure, to inject the approximating optimal solution in the original problem. Notice that in all the experiments, the cost corresponding to Ω_g is the best one, but the differences with respect to the other computed cost values are small. This shows that the rather complex approximation/penalization that we use is reasonable. Its advantage is that it may be used as well in the case of boundary observation or for Neumann boundary conditions and this will be performed in a subsequent paper.

In the examples, we have employed the software FreeFem++, [9].

EXAMPLE 5.1. The computational domain is $D =]-3, 3[\times]-3, 3[$ and the observation zone E is the disk of center $(0, 0)$ and radius 0.5. The load is $f = 1$, $j(g) = (y_\varepsilon - y_d)^2$, where $y_d(x_1, x_2) = -(x_1 - 0.5)^2 - (x_2 - 0.5)^2 + 1/16$, then the cost function (3.2) becomes

$$(5.1) \quad \min_{g \in \mathcal{F}, u \in L^2(D)} J(g, u) = \left\{ \int_E (y_\varepsilon - y_d)^2 \, d\mathbf{x} + \frac{1}{\varepsilon} \int_{I_g} (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_g(t)| \, dt \right\}.$$

The mesh of D has 73786 triangles and 37254 vertices. The penalization parameter is $\varepsilon = 10^{-3}$ and the tolerance parameter for the stopping test at the STEP 5 of the algorithm is $tol = 10^{-6}$. The initial domain is the disk of center $(0, 0)$ and radius 2.5 with a circular hole of center $(-1, -1)$ and radius 0.5.

At the STEP 3 of the algorithm, we use (R^k, V^k) given by Proposition 4.5. At the STEP 4, in order to have $E \subset \Omega_k$, we use a projection \mathcal{P} at the line search

$$\lambda_k \in \arg \min_{\lambda > 0} J(\mathcal{P}(G^k + \lambda R^k), U^k + \lambda V^k)$$

and $G^{k+1} = \mathcal{P}(G^k + \lambda_k R^k)$. If the value of $g_h^k + \lambda r_h^k$ at a vertex from \bar{E} is positive, then we set this value to -0.1 . We recall that the left-hand side of (4.30) represents only the first two terms of (4.22), not the whole gradient.

If r_h, v_h are given by Proposition 4.5 and $\gamma > 0$ is a scaling parameter, then γr_h and v_h verify (4.30), that is they also give a descent direction. We take the scaling parameter for r_h given by $\gamma = 1/\max(r_h)$, that is a normalization of r_h . In this way we avoid the appearance of very high values of the objective function,

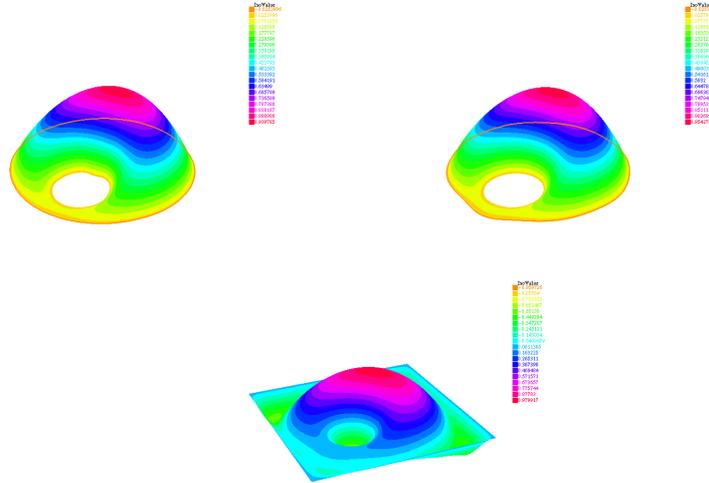


FIGURE 1. The solution of the elliptic problem (1.2)–(1.3) in the domain Ω_g (left), in the domain bounded by the zero level sets of y_ε (right) and the final computed state y_ε in D (bottom).

that may stop the algorithm even in the first iteration. For the line search at the STEP 4, we use $\lambda = \rho^i \lambda_0$, with $\lambda_0 = 1$, $\rho = 0.5$ for $i = 0, 1, \dots, 30$.

The stopping test is obtained for $k = 94$ and some values of the objective function are:

$$J(G^0, U^0) = 33110.5, \quad J(G^{30}, U^{30}) = 54.725, \quad J(G^{94}, U^{94}) = 14.9851.$$

At the final iteration, the first term of the optimal objective function is 1.03796 and $\int_{\partial\Omega_g} y_\varepsilon^2(s) ds = 1.39471 \times 10^{-2}$. We point out that the optimal Ω_g has a hole and the penalization term is a sum of two integrals

$$\int_{\partial\Omega_g} y_\varepsilon^2(s) ds = \sum_{j=1}^2 \int_{I_j} (y_\varepsilon(\mathbf{z}_g(t)))^2 |\mathbf{z}'_g(t)| dt$$

where the integral over I_1 corresponding to the exterior boundary of Ω_g and I_2 to the boundary of the hole. In Figure 1 in the bottom, we can see the computed optimal state y_ε in iteration 94. We also compute the costs $\int_E (y_1 - y_d)^2 d\mathbf{x} = 0.998189$ where y_1 is the solution of the initial elliptic problem (1.2)–(1.3) in the domain Ω_g with g obtained in iteration 94 and $\int_E (y_2 - y_d)^2 d\mathbf{x} = 1.04032$ where y_2 is the solution of the elliptic problem (1.2)–(1.3) in the domain bounded by the zero level sets of y_ε in iteration 94, see Figure 1.

EXAMPLE 5.2. The domains D , E and the mesh of D are the same as in Example 5.1. For $f = 4$ and $y_d(x_1, x_2) = -x_1^2 - x_2^2 + 1$, we have the exact

optimal state $y = y_d$ defined in the disk of center $(0, 0)$ and radius 1, that gives an optimal domain of the problem (1.1)–(1.3).

We have used $\varepsilon = 10^{-1}$ and the starting configuration: the disk of center $(0, 0)$ and radius 2.5 with the circular hole of center $(-1, -1)$ and radius 0.5. We use (R^k, V^k) given by Proposition 4.5. The parameters for the line search and γ are the same as in the precedent example.

The stopping test is obtained for $k = 64$. The initial and the final computed values of the objective function are 5368.84 and 11.2311. We obtain a local minimum that is different from the above global solution. The first term of the final computed objective function is 0.472856. The term $\int_{\partial\Omega_g} y_\varepsilon^2(s) ds$ is 1.07583 and it was computed over the exterior boundary as well as over the boundaries of two holes. The length of the total boundary of the optimal domain is 23.9714 and of the initial domain is $2\pi(2.5 + 0.5) = 18.8495$.

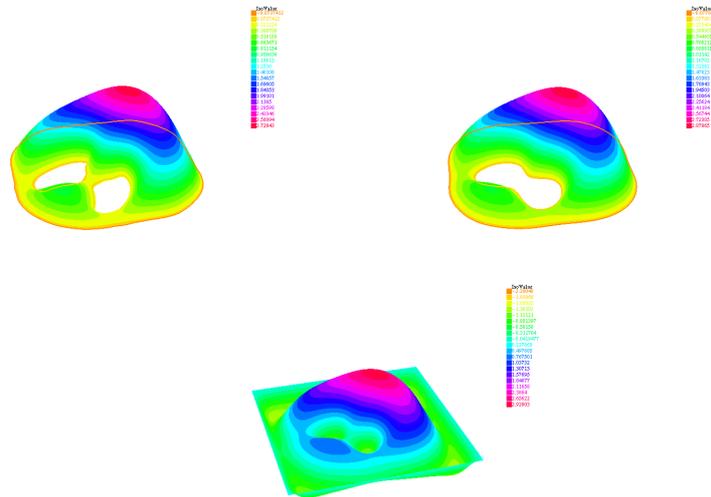


FIGURE 2. The numerical solution of the elliptic problem (1.2)–(1.3) in the optimal domain Ω_g (left), in the domain bounded by the zero level sets of y_ε (right) and the computed optimal state y_ε (bottom).

The domain changes its topology. The computed optimal state y_ε is presented in Figure 2 in the bottom. At the left, we show y_1 the solution of the elliptic problem (1.2)–(1.3) in the domain Ω_g which gives $\int_E (y_1 - y_d)^2 dx = 0.295178$, at the right we show y_2 the solution of the elliptic problem (1.2)–(1.3) in the domain bounded by the zero level sets of y_ε , which gives $\int_E (y_2 - y_d)^2 dx = 0.471788$.

EXAMPLE 5.3. We have also used the descent direction given by Proposition 4.7, for the starting configuration the disk of center $(0, 0)$ and radius 1.5, $\varepsilon = 10^{-1}$, $\gamma = 1/\|r_h\|_\infty$ and a mesh of D of 32446 triangles and 16464 vertices. For solving the ODE systems (4.3)–(4.5) and (4.11)–(4.13) we use $m = 30$.

At the initial iteration, we have $\int_E (y_\varepsilon - y_d)^2 d\mathbf{x} = 72.3767$, $\int_{\partial\Omega_g} y_\varepsilon^2(s) ds = 658.459$ and the value of the objective function is $J_0 = 6656.98$. The algorithm stops after 12 iterations and we have at the final iteration $\int_E (y_\varepsilon - y_d)^2 d\mathbf{x} = 1.22861$, $\int_{\partial\Omega_g} y_\varepsilon^2(s) ds = 0.557556$ and the value of the penalized objective function is $J_{12} = 6.80521$. The final domain is a perturbation of the initial one, the circular non-smooth curve in the top, left image of Figure 3. We have $\int_E (y_1 - y_d)^2 d\mathbf{x} = 1.20398$ for y_1 the solution of the elliptic problem (1.2)–(1.3) in the final domain Ω_g and $\int_E (y_2 - y_d)^2 d\mathbf{x} = 1.21767$ for y_2 the solution of the elliptic problem (1.2)–(1.3) in the domain bounded by the zero level sets of y_ε , Figure 3 at the bottom, right.

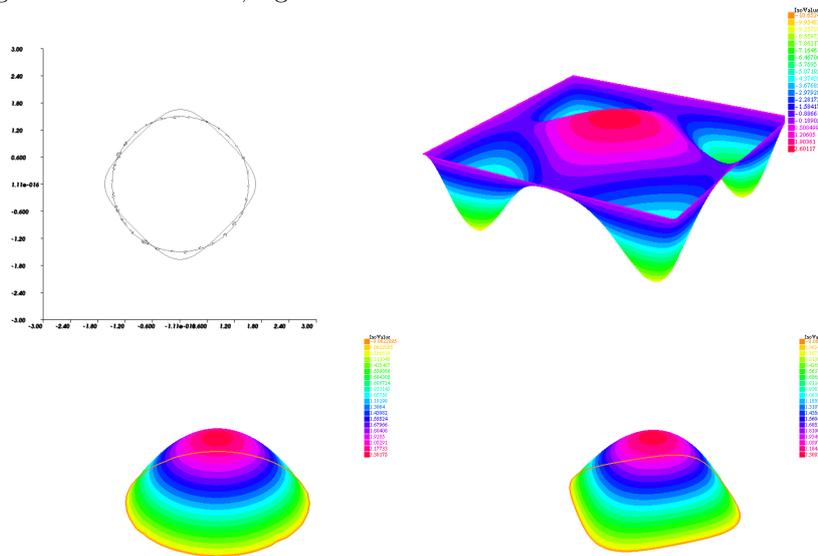


FIGURE 3. The zero level sets of the computed optimal g , y_ε (top, left), the final state y_ε (top, right), the solution of the elliptic problem (1.2)–(1.3) in the domain Ω_g (bottom, left) and in the domain bounded by the zero level sets of y_ε (bottom, right).

Finally, we notice that the hypothesis of Proposition 3.5 is obviously fulfilled by the null level sets of y_ε with a corresponding parametrization. In an approximate sense, it is also fulfilled by the computed optimal domain Ω_g since the penalization integral is “small” in all the examples.

REFERENCES

[1] E.G. ALLAIRE, *Shape optimization by the homogenization method*, Springer, New York, 2002.
 [2] V. BARBU AND TH. PRECUPANU, *Convexity and optimization in Banach spaces*. Mathematics and its Applications, vol. 10, D. Reidel Publishing Co., Dordrecht; Editura Academiei, Bucharest, 1986.

- [3] M. BENDSOE AND O. SIGMUND, *Topology Optimization: Theory, Methods and Application*, 2-nd edition, Engineering Online Library, Springer, Berlin, 2003.
- [4] F. BOUCHUT AND L. DESVILLET, *On two-dimensional Hamiltonian transport equations with continuous coefficients* *Differential Integral Equations* **14** (2001), no. 8, 1015–1024. <https://projecteuclid.org/euclid.die/1356123178>
- [5] D. BUCUR AND G. BUTTAZZO, *Variational methods in shape optimization problems*. *Progress in Nonlinear Differential Equations and their Applications*, vol. 65, Birkhäuser Boston, Inc., Boston, MA, 2005.
- [6] P.G. CIARLET, *The finite element method for elliptic problems*, *Classics in Applied Mathematics*, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.
- [7] F.H. CLARKE, *Optimization and nonsmooth analysis*, *Canadian Mathematical Society Series of Monographs and Advanced Texts*, A Wiley – Interscience Publication, John Wiley & Sons, Inc., New York, 1983.
- [8] M.C. DELFOUR AND J.P. ZOLESIO, *Shapes and Geometries, Analysis, Differential Calculus and Optimization*, SIAM, Philadelphia, 2001.
- [9] F. HECHT, *New development in FreeFem++*, *J. Numer. Math.* **20** (2012), 251–265. <http://www.freefem.org>
- [10] A. HENROT AND M. PIERRE, *Variations et optimisation de formes. Une analyse géométrique*, Springer, 2005.
- [11] M.W. HIRSCH, S. SMALE AND R.L. DEVANEY, *Differential Equations, Dynamical Systems and an Introduction to Chaos*, Elsevier, Academic Press, San Diego, 2004.
- [12] J.-L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Gauthier–Villars, Paris 1969.
- [13] R. MÄKINEN, P. NEITTAANMÄKI AND D. TIBA, *On a fixed domain approach for a shape optimization problem*, *Computational and Applied Mathematics II: Differential Equations*, (W.F. Ames and P.J. van der Houwen, eds) North-Holland, Amsterdam, 1992, pp. 317–326.
- [14] P. NEITTAANMÄKI, A. PENNANEN AND D. TIBA, *Fixed domain approaches in shape optimization problems with Dirichlet boundary conditions*, *Inverse Problems* **25** (2009), 1–18.
- [15] P. NEITTAANMÄKI, J. SPREKELS AND D. TIBA, *Optimization of elliptic systems. Theory and applications*, Springer, New York, 2006.
- [16] P. NEITTAANMÄKI AND D. TIBA, *Optimal control of nonlinear parabolic systems. Theory, algorithms, and applications*, *Monographs and Textbooks in Pure and Applied Mathematics*, vol. 179, Marcel Dekker, Inc., New York, 1994.
- [17] P. NEITTAANMÄKI AND D. TIBA, *Fixed domain approaches in shape optimization problems*, *Inverse Problems*, **28** (2012), 1–35.
- [18] M.R. NICOLAI AND D. TIBA, *Implicit functions and parametrizations in dimension three: generalized solutions*, *Discrete Contin. Dyn. Syst.* **35** (2015), no. 6, 2701–2710.
- [19] S. OSHER AND J. SETHIAN, *Fronts propagating with curvature-dependent speed*, *J. Comput. Phys.* **79** (1988), 12–49.
- [20] O. PIRONNEAU, *Optimal Shape Design for Elliptic Systems*, Springer, Berlin, 1984.
- [21] P.-A. RAVIART AND J.-M. THOMAS, *Introduction à l'Analyse Numérique des Équations aux Dérivées Partielles*, Dunod, 2004.
- [22] T.C. SIDERIS, *Ordinary Differential Equations and Dynamical Systems*, Atlantis Press, Paris, 2013.
- [23] J. SOKOLOWSKI AND J.P. ZOLESIO, *Introduction to Shape Optimization. Shape Sensitivity Analysis*, Springer, Berlin, 1992.

- [24] D. TIBA, *The implicit function theorem and implicit parametrizations*, Ann. Acad. Rom. Sci. Ser. Math. Appl. **5** (2013), no. 1–2, 193–208.
- [25] D. TIBA, *Iterated Hamiltonian type systems and applications* J. Differential Equations **264** (2018), no. 8, 5465–5479.
- [26] D. TIBA, *A penalization approach in shape optimization*, Atti Accad. Peloritana Pericolanti Cl. Sci. Fis. Mat. Natur. **96** (2018), no. 1, A8, DOI: 10.1478/AAPP.961A8.

Manuscript received November 24, 2018

accepted April 30, 2019

CORNEL MARIUS MUREA
Département de Mathématiques
IRIMAS
Université de Haute Alsace, FRANCE
E-mail address: cornel.murea@uha.fr

DAN TIBA
Institute of Mathematics
Romanian Academy
and
Academy of Romanian Scientists
Bucharest, ROMANIA
E-mail address: dan.tiba@imar.ro