Journal of Convex Analysis
Volume 33 (2026), No. 3&4, 765–780

Optimal Coefficients for Elliptic PDEs

Giuseppe Buttazzo
Dipartimento di Matematica, Università di Pisa, Italy
giuseppe.buttazzo@unipi.it

Juan Casado-Díaz, Faustino Maestre
Dpto. de Ecuaciones Diferenciales y Análisis Numérico
Facultad de Matemáticas, Sevilla, Spain
jcasadod@us.es, fmaestre@us.es

Dedicated to the memory of Hedy Attouch.

Received: July 15, 2024
Accepted: July 20, 2024

We consider an optimization problem related to elliptic PDEs of the form $-\mathrm{div}(a(x)\nabla u) = f$ with Dirichlet boundary condition on a given domain Ω. The coefficient $a(x)$ has to be determined, in a suitable given class of admissible choices, in order to optimize a given criterion. We first deal with the case when the cost is the so-called elastic compliance, and then we discuss the more general case when the problem is written as an optimal control problem.

Keywords: Shape optimization, optimal coefficients, regularity, optimal control problems.

2020 Mathematics Subject Classification: 49Q10, 49J45, 35B65, 35R05, 49K20.

1. Introduction

Various quantities are involved in the study of elliptic PDEs, which we often refer to as *data*. In particular, in several situations the coefficients of an elliptic operator have to be determined in order to optimize a given cost functional. In all the paper Ω is a given bounded open subset of $\mathbb{R}^d$ and $H_0^1(\Omega)$ is the usual Sobolev space of functions with zero boundary trace.

We give here below a presentation of the optimization problem, and in Section 5 we provide some numerical simulation. Other kinds of optimization problems for elliptic PDEs, namely the optimal choice of lower order terms, together with their regularity, have been considered in [4], [18], [22].

1.1. Position of the problem

We consider the problem of minimizing a cost functional of the form

$$J(u) = \int_\Omega j(x,u)\,dx,$$

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

where $j(x,s)$ is a suitable cost integrand with the appropriate growth conditions, and u is the solution of the elliptic equation

$$\begin{cases} -\operatorname{div}\big(a(x)\nabla u\big) = f & \text{in } \Omega \\ u \in H_0^1(\Omega). \end{cases} \tag{1}$$

Here the right-hand side $f \in L^2(\Omega)$ is prescribed, while the coefficient a has to be chosen in a suitable admissible class $\mathbb{A}$ in order to minimize the functional J above. The problem is then an optimal control problem, where u is the state variable, a the control variable, J the cost functional, and (1) is the state equation. This amounts then to the problem

$$\min\big\{J(u)\ :\ u \text{ solves } (1),\ a \in \mathbb{A}\big\}.$$

The admissible class $\mathbb{A}$ is usually given in the form

$$\mathbb{A} = \Big\{a \geq 0\ :\ \int_\Omega \psi(a)\,dx \leq 1\Big\}.$$

A simpler way to impose the constraint on a is to write the problem in the form

$$\min_{a\geq 0}\ \min_{u\in H_0^1(\Omega)} \Big\{\int_\Omega \big(j(x,u) + \lambda\psi(a)\big)\,dx\ :\ u \text{ solves } (1)\Big\}. \tag{2}$$

where $\lambda > 0$ plays the role of a Lagrange multiplier. Moreover, we will assume ψ convex and non-negative. Replacing ψ by $\lambda\psi$ we can also assume $\lambda = 1$.

2. The compliance and energy problems

A particular case of the problem considered above occurs when the cost J is the so-called *compliance*, that is

$$j(x,u) = f(x)u.$$

In this case an easy integration by parts transforms the problem in the max/min problem

$$\max_{a\geq 0}\ \min_{u\in H_0^1(\Omega)} \int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - f(x)u - \frac{1}{2}\psi(a)\Big)\,dx.$$

We first assume that ψ is superlinear, that is

$$\lim_{s\to+\infty} \frac{\psi(s)}{s} = +\infty, \tag{3}$$

and we set $\quad \mathcal{E}(a) = \inf_{u\in C_0^1(\Omega)} \int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - f(x)u - \frac{1}{2}\psi(a)\Big)\,dx.$

Concerning the right-hand side f, in several problems some concentration phenomena for data occur, so we simply require that the right-hand side f is a signed measure.

Theorem 2.1. *Under assumption* (3)*, the functional $\mathcal{E}(a)$ admits a maximizer $a_{opt} \in L^1(\Omega)$, provided the right-hand side f is such that $\mathcal{E}(a) > -\infty$ for at least a coefficient $a \in L^1(\Omega)$.*

Proof. Since for every $u \in C_0^1(\Omega)$ the map

$$a \mapsto \int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - \frac{1}{2}\psi(a)\Big)\,dx - \int u\,df$$

is weakly $L^1(\Omega)$ upper semicontinuous, the functional $\mathcal{E}(a)$ is weakly $L^1(\Omega)$ upper semicontinuous, being the infimum of a family of upper semicontinuous functions. In addition, testing with $u = 0$, we have

$$\mathcal{E}(a) \le -\frac{1}{2}\int_\Omega \psi(a)\,dx.$$

Hence, by the superlinearity of ψ and by the well-known weak $L^1(\Omega)$ compactness theorem, the existence of an optimal coefficient a_{opt} is easily established by means of the direct methods of the calculus of variations. □

The case when ψ has only a linear growth:

$$\lim_{s\to+\infty}\frac{\psi(s)}{s} = k > 0 \tag{4}$$

is more delicate. In fact, in this case a more careful definition of the integrals

$$\int_\Omega |\nabla u|^2\,da \qquad \text{and} \qquad \int_\Omega \psi(a)$$

is needed. We refer to [5] for more details about this case, which has strong links with the theory of optimal transportation, as first shown in [3] and [2]. However, by an argument similar to the one of Theorem 2.1, an optimal coefficient a_{opt} still exists, but in the larger class $\mathscr{M}(\Omega)$ of nonnegative measures on Ω, as stated below.

Theorem 2.2. *Under assumption* (4)*, the functional* $\mathcal{E}(a)$ *admits a maximizer* a_{opt} *in the class* $\mathscr{M}(\Omega)$*, provided the right-hand side* f *is such that* $\mathcal{E}(a) > -\infty$ *for at least a coefficient* $a \in \mathscr{M}(\Omega)$.

It is interesting to characterize the optimal coefficient a_{opt} in terms of some suitable auxiliary variational problem. Thanks to a well-known result from min/max theory that allows to exchange the order of inf and sup, due to the convexity with respect to the variable u and the concavity with respect to the variable a (see for instance [10] and [14]), the initial problem becomes

$$\inf_{u\in C_0^1(\Omega)}\sup_{a\ge 0}\int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - \frac{1}{2}\psi(a)\Big)\,dx - \int u\,df. \tag{5}$$

The supremum with respect to a can be now easily computed:

$$\sup_{a\ge 0}\int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - \frac{1}{2}\psi(a)\Big)\,dx - \int u\,df = \int_\Omega \frac{1}{2}\psi^*\big(|\nabla u|^2\big)\,dx - \int u\,df,$$

where ψ^* is the Legendre-Fenchel conjugate function of ψ.

The auxiliary variational problem is then

$$\inf_{u\in C^1_0(\Omega)}\int_\Omega \frac{1}{2}\psi^*\big(|\nabla u|^2\big)\,dx-\int u\,df. \tag{6}$$

Since $\psi^*(t)\geq t-\psi(1)$ it is easy to see that, at least when $f\in H^{-1}(\Omega)$, the auxiliary variational problem admits a solution $\bar u\in H^1_0(\Omega)$. Moreover, if ψ is strictly increasing, then the function $s\mapsto\psi^*(s^2)$ is strictly convex and therefore $\bar u$ is unique. The optimal coefficient a_{opt} can now be recovered through the optimality condition

$$a_{opt}|\nabla\bar u|^2=\psi(a_{opt})+\psi^*(|\nabla\bar u|^2). \tag{7}$$

Remark 2.3. In some situations it is important to allow the right-hand side f to be singular, for instance with concentrations on regions of lower dimensions. In general we can assume that $f\in\mathscr{M}$, the class of measures with finite mass; even if for some choice of the coefficient a we may have $\mathcal{E}(a)=-\infty$, the optimal compliance problem is still meaningful, because these “bad” coefficients are ruled out by the optimization criterion, consisting in maximizing $\mathcal{E}(a)$. Moreover, all the arguments also apply to the case when, instead of having a boundary Dirichlet condition, we have the Neumann one, assuming as usual that the right-hand side f has zero average. □

Remark 2.4. When $\psi(s)=\gamma s$, using the equivalence between coefficient optimization and optimal transport problem, pointed out in [2], the following summability properties for the optimal coefficient a_{opt} have been obtained (see [12]):

$$\begin{array}{lcl} f\in\mathscr{M} & \Longrightarrow & \mu_{opt}\in\mathscr{M}\ \text{possibly not unique;}\\ f\in L^1(\Omega) & \Longrightarrow & \mu_{opt}\in L^1(\Omega)\ \text{and is unique;}\\ f\in L^p(\Omega) & \Longrightarrow & \mu_{opt}\in L^p(\Omega)\ \text{for every } p\in[1,+\infty]; \end{array}$$

$$\operatorname{spt}(\mu_{opt})\subset\text{convex envelope of}\begin{cases}\operatorname{spt}(f) & \text{in the Neumann case}\\ \operatorname{spt}(f)\cup\partial\Omega & \text{in the Dirichlet case.}\end{cases}$$

In addition, a mild BV and $W^{1,1}$ regularity for μ_{opt} is available in some cases in dimension two. More precisely, when $d=2$ and under some additional assumptions on the regularity of Ω and on the behavior of the datum f, we have (see [13]):

$$\begin{array}{lcl} f\in BV(\Omega)\cap L^\infty(\Omega) & \Longrightarrow & \mu_{opt}\in BV(\Omega),\\ f\in W^{1,1}(\Omega)\cap L^\infty(\Omega) & \Longrightarrow & \mu_{opt}\in W^{1,1}(\Omega). \end{array}$$

In higher dimension. Furthermore, the correspondence between the optimization and transport problems is unclear when the function ψ is nonlinear. □

Example 2.5. Taking $\psi(s)=s^2/2$ and $f\in W^{-1,4/3}(\Omega)$, we obtain the auxiliary variational problem

$$\min\Big\{\int_\Omega\Big(\frac{1}{4}|\nabla u|^4-f(x)u\Big)\,dx\ :\ u\in W^{1,4}_0(\Omega)\Big\},$$

or equivalently the nonlinear PDE

$$-\Delta_4 u=f,\qquad u\in W^{1,4}_0(\Omega),$$

whose unique solution $\bar{u}$ provides the optimal coefficient $a_{opt}(x) = |\nabla \bar{u}(x)|^2$. For instance, if Ω is the unit ball, and $f = 1$ we obtain

$$\bar{u}(x) = \frac{3}{4d^{1/3}}\big(1 - |x|^{4/3}\big), \qquad a_{opt}(x) = \frac{|x|^{2/3}}{d^{2/3}}.$$

Conversely, taking Ω the unit disc in $\mathbb{R}^2$ and $f = \delta_0$ the unit Dirac mass at the origin, gives

$$\bar{u}(x) = \frac{3}{(16\pi)^{1/3}}(1-|x|^{2/3}), \qquad a_{opt}(x) = (2\pi|x|)^{-2/3}. \qquad \square$$

Example 2.6. Taking

$$\psi(s) = \begin{cases} \gamma s & \text{if } \alpha \le s \le \beta \\ +\infty & \text{otherwise,} \end{cases} \tag{8}$$

with $0 < \alpha < \beta$, $\gamma > 0$, we have the auxiliary variational problem

$$\min \Big\{ \int_\Omega \frac{|\nabla u|^2 - \gamma}{2}\Big(\beta 1_{|\nabla u|^2 \ge \gamma} + \alpha 1_{|\nabla u|^2 \le \gamma}\Big) - f(x)u\,dx \;:\; u \in H^1_0(\Omega)\Big\},$$

whose unique solution $\bar{u}$ provides the optimal coefficient $a_{opt} \in L^\infty(\Omega)$. It has been proved in [8] (see also [9]) that, when Ω is of class $C^{1,1}$ and $f \in L^2(\Omega)$, then $\bar{u}$ is in $H^2(\Omega)$ and $\nabla a_{opt} \cdot \nabla \bar{u}$ belongs to $L^2(\Omega)$. $\square$

Another case where (2) reduces to a variational problem is the minimization of the energy, corresponding to

$$j(x, u) = -f(x)u.$$

Similarly to (5), the problem becomes

$$\min_{a \ge 0} \min_{u \in H^1_0(\Omega)} \int_\Omega \Big(\frac{1}{2}a(x)|\nabla u|^2 - f(x)u + \frac{1}{2}\psi(a)\Big)\,dx. \tag{9}$$

which, computing the minimum in a, can be written as

$$\min_{u \in H^1_0(\Omega)} \int_\Omega \Big(-\frac{1}{2}\psi^*\big(-|\nabla u|^2\big) - fu\Big)\,dx. \tag{10}$$

Imposing that

$$\lim_{s \to 0^+} s\psi(s) = \infty,$$

the functional becomes coercive over $W^{1,1}_0(\Omega)$. Assuming also that ψ is decreasing and that $\psi(1/s)$ is convex, we have that the functional in (10) is convex. Therefore, under these assumptions and taking f smooth enough, problem (10) has a solution in $W^{1,1}_0(\Omega)$ (it is not necessarily in $H^1_0(\Omega)$). However, in other situations this functional is not convex and then (10) may not have a solution. To avoid this difficulty it is necessary to deal with a relaxed formulation consisting in replacing the function $\xi \in \mathbb{R}^d \mapsto -\psi(-|\xi|^2)$ by its convex hull.

Example 2.7. Related to Example 2.5, we take $\psi(s) = 1/(2s^2)$. Then problem (10) becomes

$$\min_{u \in W^{1,\frac{4}{3}}_0(\Omega)} \int_\Omega \Big(\frac{3}{4}|\nabla u|^{\frac{4}{3}} - fu\Big)\,dx,$$

which has a unique solution $\overline{u}$ if f is in $W^{-1,4}(\Omega)$.

The optimal control is given by $a_{opt} = |\nabla\overline{u}|^{-2/3}$. In this way, if Ω is the unit ball in $\mathbb{R}^d$ and $f = 1$, we get

$$\overline{u}(x) = \frac{1-|x|^4}{4d^3}, \qquad a_{opt}(x) = \frac{d^2}{|x|^2}. \qquad \square$$

Example 2.8. Taking

$$\psi(s) = \begin{cases} \gamma(\beta - s) & \text{if } s \in [\alpha, \beta] \\ +\infty & \text{otherwise,} \end{cases} \tag{11}$$

with $0 < \alpha < \beta$, $\gamma > 0$, we have

$$-\psi^*(-|\xi|^2) = \begin{cases} \beta|\xi|^2 & \text{if } |\xi|^2 \leq \gamma \\ \alpha|\xi|^2 + \gamma(\beta-\alpha) & \text{if } |\xi|^2 > \gamma, \end{cases}$$

which is not convex. Computing its convex hull (see e.g. [15]) we get the relaxed formulation

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \Big(\frac{1}{2}\phi\big(|\nabla u|\big) - fu\Big)\, dx, \tag{12}$$

with
$$\phi(s) = \begin{cases} \beta s^2 & \text{if } s^2 \leq \frac{\alpha}{\beta}\gamma \\ 2\sqrt{\alpha\beta\gamma}\, s - \alpha\gamma & \text{if } \frac{\alpha}{\beta}\gamma \leq s^2 \leq \frac{\beta}{\alpha}\gamma \\ \alpha s^2 + \gamma(\beta-\alpha) & \text{if } \frac{\beta}{\alpha}\gamma \leq s^2. \end{cases}$$

The Euler-Lagrange equation for (12) proves that for a given solution $\overline{u}$, the associated optimal "relaxed control" is given by

$$a_{opt} = \frac{\phi'\big(|\nabla\overline{u}|\big)}{2|\nabla\overline{u}|}.$$

It can be proved that this optimal relaxed control coincides with the optimal relaxed control defined in the following section. Moreover, although a_{opt} and $\nabla\overline{u}$ may not be unique, the function $\overline{\sigma} = a_{opt}\nabla\overline{u}$ is unique.

Taking Ω the unit ball in $\mathbb{R}^d$, $f = 1$, $\gamma < 1/(d^2\alpha\beta)$, and denoting $\tau = d\sqrt{\alpha\beta\gamma}$, we have

$$\overline{u}(x) = \begin{cases} \dfrac{(1-\tau^2)(\beta-\alpha)}{2d\alpha\beta} + \dfrac{1-|x|^2}{2d\beta} & \text{if } |x| < \tau \\ \dfrac{1-|x|^2}{2d\alpha} & \text{if } |x| > \tau, \end{cases} \qquad a_{opt}(x) = \begin{cases} \beta & \text{if } |x| < \tau \\ \alpha & \text{if } |x| > \tau. \end{cases}$$

Remark 2.9. For ψ given by (11), it has been proved in [8] that $\Omega \in C^{1,1}$ and $f \in L^2(\Omega)$ imply that $\overline{\sigma}$ is in $H^1(\Omega)^d$ and that the derivatives of a_{opt} in the orthogonal directions to σ are in $L^2(\Omega)$. $\square$

3. The general problem

In the case of a general optimal control problem of the form

$$\min_{a\geq 0}\min_{u\in H_0^1(\Omega)}\left\{\int_\Omega \big(j(x,u)+\psi(a)\big)\,dx\ :\ u\text{ solves (1)}\right\},\tag{13}$$

the existence of an optimal coefficient a_{opt} may fail, and a solution exists only in a *relaxed* sense. For $0<\alpha<\beta$, a counterexample to the existence of a solution a_{opt} can be found in [20], where

$$\psi(s)=\begin{cases}0 & \text{if } \alpha\leq s\leq\beta\\ +\infty & \text{otherwise,}\end{cases}$$

and
$$j(x,s)=|s-u_0(x)|^2$$

for a suitable function u_0. A different counterexample is illustrated in Section 4.

In order to understand what relaxed solutions are, we have to recall the notion of G-convergence, introduced by De Giorgi and Spagnolo in [11]: a sequence $a_n(x)$ of functions between α and β is said to G-converge to a symmetric $d\times d$ matrix $A(x)$ if for every $f\in L^2(\Omega)$ the solutions u_n of the PDEs

$$-\operatorname{div}\big(a_n\nabla u_n\big)=f,\qquad u_n\in H_0^1(\Omega)$$

converge in $L^2(\Omega)$ to the solution u of the PDE

$$-\operatorname{div}\big(A\nabla u\big)=f,\qquad u\in H_0^1(\Omega).\tag{14}$$

The question becomes now to characterize the G-closure $\overline{\mathbb{A}}$ of the set of coefficients a_n. A complete answer has been given by Murat and Tartar in [21], [23] (see also [17] for the two-dimensional case). They proved that the G-closure $\overline{\mathbb{A}}$ above consists of all symmetric $d\times d$ matrices $A(x)$ whose eigenvalues $\lambda_1(x)\leq\lambda_2(x)\leq\cdots\leq\lambda_d(x)$ are between α and β and satisfy for a suitable $t\in[0,1]$ (depending on x) the following $d+2$ inequalities:

$$\begin{cases}\displaystyle\sum_{1\leq i\leq d}\frac{1}{\lambda_i-\alpha}\leq\frac{1}{\nu_t-\alpha}+\frac{d-1}{\mu_t-\alpha}\\ \displaystyle\sum_{1\leq i\leq d}\frac{1}{\beta-\lambda_i}\leq\frac{1}{\beta-\nu_t}+\frac{d-1}{\beta-\mu_t}\\ \nu_t\leq\lambda_i\leq\mu_t\qquad i=1,\dots,d,\end{cases}$$

being μ_t and ν_t respectively the arithmetic and the harmonic mean of α and β, namely

$$\mu_t=t\alpha+(1-t)\beta,\qquad \nu_t=\Big(\frac{t}{\alpha}+\frac{1-t}{\beta}\Big)^{-1}.$$

For instance, when $d=2$, the set above is given by the symmetric 2×2 matrices $A(x)$ whose eigenvalues $\lambda_1(x)$ and $\lambda_2(x)$ are between α and β and

$$\frac{\alpha\beta}{\alpha+\beta-\lambda_1(x)}\leq\lambda_2(x)\leq\alpha+\beta-\frac{\alpha\beta}{\lambda_1(x)}\ .$$

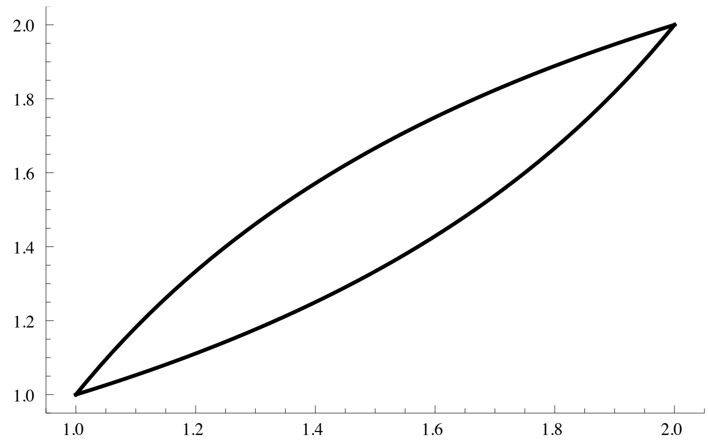

Figure 3.1: Attainable matrices, in the plane (λ_1, λ_2), for $d = 2$, $\alpha = 1$, $\beta = 2$.

As far as we know, an explicit form of the relaxation

$$\Psi(A) = \inf_{a_n \to_G A} \liminf_n \int_\Omega \psi(x, a_n)\, dx;$$

with a general function $\psi(x, a)$ is not known. The case $\psi(x, a) = g(x)a$ has been considered in [6] and [7]. For instance, if ψ is given by (8), denoting by $\overline{\mathbb{A}}$ the G-closure described above, the relaxation $\Psi(A)$ is given by

$$\Psi(A) = \begin{cases} \displaystyle\int_\Omega \gamma \lambda_{max}\big(A(x)\big)\, dx & \text{if } A \in \overline{\mathbb{A}} \\ +\infty & \text{otherwise,} \end{cases}$$

being $\lambda_{max}(A)$ the largest eigenvalue of the $d \times d$ symmetric matrix A. Namely, taking into account that the solution of the state equation

$$-\operatorname{div}\big(A(x)\nabla u\big) = f, \qquad u \in H^1_0(\Omega),$$

does not vary if we replace $A(x)$ by another matrix function $B(x)$ such that

$$A(x)\nabla u = B(x)\nabla u \qquad \text{a.e. in } \Omega,$$

and that for every $\xi \in \mathbb{R}^d$ one has

$$\big\{A\xi : \ A \in \overline{\mathbb{A}}\big\} = \big\{\eta \in \mathbb{R}^d : \ (\eta - \nu_t \xi) \cdot (\eta - \mu_t \xi) \le 0\big\},$$

we have that a relaxation of (13) with $\psi(x, a) = g(x)a$ is given by

$$\min_{\substack{\nu_t I \le A \le \mu_t I \\ 0 \le t \le 1}} \ \min_{u \in H^1_0(\Omega)} \Big\{ \int_\Omega \big(j(x, u) + g(x)\mu_t\big)\, dx \ : \ u \text{ solves } (14)\Big\}, \tag{15}$$

where I denotes the identity matrix.

As an example, we can consider the energy problem $j(x, u) = -f(x)u$ with ψ given by (11). Then, taking into account (9) and (15) the relaxed problem can be written as

$$\min_{0 \le t \le 1} \ \min_{u \in H^1_0(\Omega)} \Big\{ \int_\Omega \big(\tfrac{\nu_t}{2}|\nabla u|^2 - f(x)u - \gamma\mu_t\big)\, dx\Big\}. \tag{16}$$

Using the minimum in t this proves again that u solves (12).

Thanks to (15) we obtain a system of optimality conditions. For this purpose, we assume the function $j(x,s)$ derivable with respect to s with appropriate growth conditions.

Take (t_{opt}, A_{opt}) an optimal solution of the relaxed problem, then for any admissible control (t,A) and $0 \le \varepsilon \le 1$ the control

$$\big(t_{opt} + \varepsilon(t - t_{opt}), A_{opt} + \varepsilon(A - A_{opt})\big)$$

is also admissible. Using it and deriving with respect to ε we conclude that

$$A_{opt}\nabla\bar{u}\cdot\nabla\bar{p} + g(\beta-\alpha)t_{opt} = \max_{\substack{\nu_t I \le A \le \mu_t I \\ 0\le t\le 1}} \Big\{A_{opt}\nabla\bar{u}\cdot\nabla\bar{p} - g(\beta-\alpha)t_{opt}\Big\}, \tag{17}$$

with $\bar{u}, \bar{p}$ the state and adjoint state functions, solutions of

$$-\operatorname{div}\big(A_{opt}\nabla\bar{u}\big) = f, \quad -\operatorname{div}\big(A_{opt}\nabla\bar{p}\big) = \partial_s j(x,\bar{u}), \qquad \bar{u},\bar{p} \in H_0^1(\Omega). \tag{18}$$

Computing the maximum in (17), we obtain the optimality conditions (see for instance [1], [21])

$$\begin{cases} A_{opt}\nabla\bar{u} = \dfrac{\mu_{t_{opt}} + \nu_{t_{opt}}}{2}\nabla\bar{u} + \dfrac{\mu_{t_{opt}} - \nu_{t_{opt}}}{2}\dfrac{|\nabla\bar{u}|}{\nabla\bar{p}|}\nabla\bar{p} & \text{a.e. in } \{\nabla\bar{p} \neq 0\} \\ A_{opt}\nabla\bar{p} = \dfrac{\mu_{t_{opt}} + \nu_{t_{opt}}}{2}\nabla\bar{p} + \dfrac{\mu_{t_{opt}} - \nu_{t_{opt}}}{2}\dfrac{|\nabla\bar{p}|}{|\nabla\bar{u}|}\nabla\bar{u} & \text{a.e. in } \{\nabla\bar{u} \neq 0\}, \end{cases} \tag{19}$$

$$t_{opt} = \begin{cases} 0 & \text{if } g < N^+ - \dfrac{\beta}{\alpha}N^- \\ \dfrac{1}{\beta-\alpha}\Big(\sqrt{\dfrac{\alpha\beta N^-}{N^+ - g}} - \alpha\Big) & \text{if } N^+ - \dfrac{\beta}{\alpha}N^- \le g \le N^+ - \dfrac{\alpha}{\beta}N^- \\ 1 & \text{if } N^+ - \dfrac{\alpha}{\beta}N^- < g, \end{cases} \tag{20}$$

with
$$N^+ = \frac{|\nabla u||\nabla p| + \nabla u\cdot\nabla p}{2}, \qquad N^- = \frac{|\nabla u||\nabla p| - \nabla u\cdot\nabla p}{2}. \tag{21}$$

4. Nonexistence of an optimal coefficient

In this section we provide a counterexample to the existence of an optimal coefficient a_{opt} for (2). We take $\Omega = B(0,1)$ the unit ball in $\mathbb{R}^d$, and

$$\begin{cases} f = 1 \text{ the right-hand side,} \\ j(x_1,...,x_d,s) = (1+\varepsilon x_1)s \text{ with } \varepsilon > 0, \\ \psi(s) = \begin{cases} \tau^2 s & \text{if } s \in [1,2] \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } 0 < \tau < \dfrac{1}{d}. \end{cases} \tag{22}$$

Let us prove the existence of $\varepsilon_0 > 0$ such that (2) has no solution for $0 < \varepsilon < \varepsilon_0$.

First we observe that for $\varepsilon = 0$ problem (2) is a particular case of the compliance problem considered in Section 2. By (6) we get that the state function u_0 associated to an optimal control a_0 is the unique solution of (6) with

$$\psi^*(s) = \begin{cases} s - \tau^2 & \text{if } s < \tau^2 \\ 2(s-\tau^2) & \text{if } s > \tau^2. \end{cases}$$

By uniqueness u_0 is invariant by rotations and then is a radial function $u_0(r)$. Combined with (7), this implies

$$u_0'(r) = \begin{cases} -\frac{r}{d} & \text{if } r < d\tau \\ -\tau & \text{if } d\tau \le r \le 2d\tau \\ -\frac{r}{2d} & \text{if } 2\tau d < r, \end{cases} \qquad a_0(r) = \begin{cases} 1 & \text{if } r < d\tau \\ \frac{r}{d\tau} & \text{if } d\tau < r < 2d\tau \\ 2 & \text{if } r > 2d\tau. \end{cases} \tag{23}$$

On the other hand, recalling that the solution u of the state equation in (1) satisfies $a\nabla u = \sigma$, with σ the solution of

$$\min\Big\{\int_\Omega \frac{|\zeta|^2}{a}\,dx \ :\ -\operatorname{div}\zeta = 1\Big\},$$

and that

$$\int_\Omega u\,dx = \int_\Omega \frac{|\sigma|^2}{a}\,dx,$$

we get that (2) with $\varepsilon = 0$ is also equivalent to

$$\min_{1\le a\le 2}\ \min_{-\operatorname{div}\sigma=1}\int_\Omega\Big(\frac{|\sigma|^2}{a} + \tau^2 a\Big)\,dx.$$

Taking the minimum with respect to a, this provides

$$a(x) = \begin{cases} 1 & \text{if } |\sigma(x)| < \tau \\ \frac{|\sigma(x)|}{\tau} & \text{if } \tau \le |\sigma(x)| \le 2\tau \\ 2 & \text{if } |\sigma(x)| > 2\tau, \end{cases}$$

with σ a solution of

$$\min_{-\operatorname{div}\sigma=1}\int_\Omega \Upsilon(|\sigma|)\,dx, \qquad \Upsilon(s) = \begin{cases} s^2+\tau^2 & \text{if } 0 \le s < \tau \\ 2\tau s & \text{if } \tau \le s \le 2\tau \\ \frac{s^2}{2} + 2\tau^2 & \text{if } s > 2\tau. \end{cases} \tag{24}$$

By what proved above, this problem has a unique solution $\sigma_0(x) = -x/d$.

Let us now prove the non-existence result. Arguing by contradiction, we assume there exist $\varepsilon_k > 0$ tending to zero and a_{ε_k} solution of (2). We denote by u_{ε_k} the corresponding state function. Since the state equation does not depend on ε, we have

$$\int_\Omega \big((1+\varepsilon_k x_1)u_{\varepsilon_k} + \tau^2 a_{\varepsilon_k}\big)\,dx \le \int_\Omega \big((1+\varepsilon_k x_1)u_0 + \tau^2 a_0\big)\,dx. \tag{25}$$

Using also
$$\int_\Omega \big(u_0+\tau^2 a_0\big)\,dx=\int_\Omega \Upsilon(|\sigma_0|)\,dx,$$
and that σ_0 solves (24), we deduce from (25)
$$0\le \int_\Omega \big(\Upsilon(|\sigma_\varepsilon|)-\Upsilon(|\sigma_0|)\big)dx\le \varepsilon\int_\Omega x_1(u_0-u_\varepsilon)\,dx. \tag{26}$$
Taking into account that u_{ε_k} solves (1) with $a=a_{\varepsilon_k}$, we know ([11]) that there exist a subsequence, still denoted by ε_k, and $A\in L^\infty(\Omega)^{d\times d}$ symmetric such that
$$I\le A\le 2I \text{ a.e. in } \Omega,\qquad a_{\varepsilon_k}I\overset{G}{\rightharpoonup} A,$$
$$u_{\varepsilon_k}\rightharpoonup u \text{ in } H^1_0(\Omega),\qquad \sigma_{\varepsilon_k}\rightharpoonup A\nabla u \text{ in } L^2(\Omega)^d,\qquad -\operatorname{div}(A\nabla u)=1 \text{ in } \Omega.$$
Coming back to (26) and taking into account the convexity of Υ, we have that
$$\int_\Omega \Upsilon(|A\nabla u|)\,dx=\min_{-\operatorname{div}\zeta=1}\int_\Omega \Upsilon(|\zeta|)\,dx.$$
Thus, $A\nabla u$ is a solution of (24) and then u is a solution of (2). By uniqueness this proves that
$$u=u_0,\qquad A=a_0I,\qquad \sigma=\sigma_0.$$
Now, we consider p_ε the adjoint state defined as the solution of
$$\begin{cases}-\operatorname{div}(a_{\varepsilon_k}\nabla p_{\varepsilon_k})=1+\varepsilon_k x_1 & \text{in } \Omega\\ p_{\varepsilon_k}=0 & \text{on } \partial\Omega.\end{cases}$$
Then $z_{\varepsilon_k}=(p_{\varepsilon_k}-u_{\varepsilon_k})/\varepsilon_k$ solves
$$\begin{cases}-\operatorname{div}(a_{\varepsilon_k}\nabla z_{\varepsilon_k})=x_1 & \text{in } \Omega\\ z_{\varepsilon_k}=0 & \text{on } \partial\Omega.\end{cases}$$
By the G-convergence of $a_{\varepsilon_k}I$, we have
$$z_{\varepsilon_k}\rightharpoonup z \text{ in } H^1_0(\Omega),\qquad a_{\varepsilon_k}\nabla z_{\varepsilon_k}\rightharpoonup A\nabla z \text{ in } L^2(\Omega)^d,$$
with z the solution of
$$\begin{cases}-\operatorname{div}(A\nabla z)=x_1 & \text{in } \Omega\\ z=0 & \text{on } \partial\Omega.\end{cases}$$
Since we are assuming a_{ε_k} a solution of (2) and then of the relaxed problem (15), we deduce from (19) that the matrix with columns $\nabla u_{\varepsilon_k},\nabla p_{\varepsilon_k}$ has rank one. By Morrey's theorem relative to the weak converges of the Jacobian ([19]), which in our case reduces to
$$\begin{aligned}&\partial_i u_{\varepsilon_k}\partial_j z_{\varepsilon_k}-\partial_i z_{\varepsilon_k}\partial_j u_{\varepsilon_k}=\partial_i(u_{\varepsilon_k}\partial_j z_{\varepsilon_k})-\partial_j(u_{\varepsilon_k}\partial_i z_{\varepsilon_k})\\ &\rightharpoonup \partial_i(u\partial_j z)-\partial_j(u_0\partial_i z)=\partial_i u\partial_j z-\partial_i z\partial_j u_0\quad \text{in } W^{-1,1}(\Omega),\end{aligned}$$

we have that ∇z is parallel a.e. to ∇u_0, and that z solves

$$\begin{cases} -\operatorname{div}\big(a_0\nabla z\big) = x_1 & \text{in } \Omega \\ z = 0 & \text{on } \partial\Omega. \end{cases}$$

However, the solution of this problems is given by $z(x) = h(|x|)x_1$ with h the unique solution of

$$\begin{cases} -\big(a_0 h'\big)' - \dfrac{d+1}{r} a_0 h' - \dfrac{a_0'}{r} h = 1 & \text{in } (0,1) \\ h(1) = 0, \qquad \displaystyle\int_0^1 r^{d+1}|h'|^2 dr < \infty. \end{cases}$$

Thus z is not a radial function and ∇z is not parallel to ∇u_0. $\square$

5. Numerical simulations

In this section we present some numerical experiments for the resolution of problems of the kind of (2) in the 2-d case. We solve numerically the problems showed in Examples 2.5 and 2.6 for compliance optimization and the example given by problem (9) with ψ defined in (11) for energy optimization.

In the case of the compliance problem we put

$$J(a) = \int_\Omega \big(f(x)u + \psi(a)\big)dx,$$

then, having in mind that for u solution of the state equation (1), it is easy to compute that

$$\frac{dJ(a)}{da} \cdot \tilde{a} = \int_\Omega \tilde{a}\left(\psi'(a) - |\nabla u|^2\right) dx.$$

We will apply a gradient descent method with projection into the appropriate subspace functions a such that $\psi(a) < \infty$. For instance in Example 2.6 ψ is finite where $a \in [\alpha, \beta]$. The algorithm is the following:

- Initialization: choose an admisible function a_0.
- for $k \geq 0$, iterate until convergence as follow:
  - compute u_k solution of (1) for $a = a_k$.
  - compute $\bar{a}_k = -(\psi'(a) - |\nabla u|^2)$ descent direction associated to u_k.
  - update the function a_k: $a_{k+1} = P_\psi(a_k + \epsilon_k \bar{a}_k)$, where P_ψ is a projection operator associated to the set $\{a : \psi(a) < \infty\}$, and where ϵ_k is small enough to ensure the decrease of the cost function.
- Stop if convergence: $\dfrac{|J(a_k) - J(a_{k-1})|}{|J(a_0)|} < tol$, for $tol > 0$ small.

In the case of the energy problem, we use a similar algorithm based on formulation (16).

On the other hand, we are interested also in showing the numerical evidence of the non-existence of an optimal coefficient for the general case. We propose to solve the relaxed formulation of (2), given by (15), in $\Omega = B(0,1)$ unit disc of $\mathbb{R}^2$, and with the data given in (22). By convex minimization and following the system of optimality, we propose the following algorithm to compute (t_{opt}, A_{opt}).

- Initialization: choose an admissible (t_0, A_0) such that we have $0 \leq t_0 \leq 1$ and $\nu_t I \leq A_0 \leq \mu_t I$.
- For $k \geq 0$, iterate until convergence as follows:
  – compute the solutions u_k and p_k of (18) for $A_{opt} = A_k$. Then, we define N^+, N^- by (21);
  – compute $\hat{t}$ given by (20);
  – compute $\hat{A}$ defined by (19), considering $\hat{u} = u_k$, $\hat{p} = p_k$, $t_{opt} = \hat{t}$ and such that the spectrum of $\hat{A}$ belongs to $[\nu_{t_k}, \mu_{t_k}]$;
  – for $\varepsilon_k \in (0, 1]$, update the function (t_k, A_k) as:

$$t_{k+1} = t_k + \varepsilon_k(\hat{t} - t_k), \qquad A_{k+1} = A_k + \varepsilon_k(\hat{A} - A_k).$$

- Stop if convergence: $\dfrac{|J(t_{k+1}, A_{k+1}) - J(t_k, A_k)|}{|J(t_0, A_0)|} < tol$, for $tol > 0$ small and J corresponding with the cost function in (15) .

We now show some numerical experiments based on the algorithms described above. The computation has been carried out using the free software FreeFem++ v4.5 ([16], available in http://www.freefem.org). The picture of figures are made in Paraview 5.10.1 (available at https://www.kitware.com/open-source/# paraview), which is free too. We use P_1-Lagrange finite element approximations for u and p, solutions of the state and adjoint state equations respectively, and P_0-Lagrange finite element approximations for control variables, a for scalar problems or (t, A) for the matrix problems. For all simulations where the parameters α and β appear, we consider a normalized value $\alpha = 1$, and $\beta = 2$.

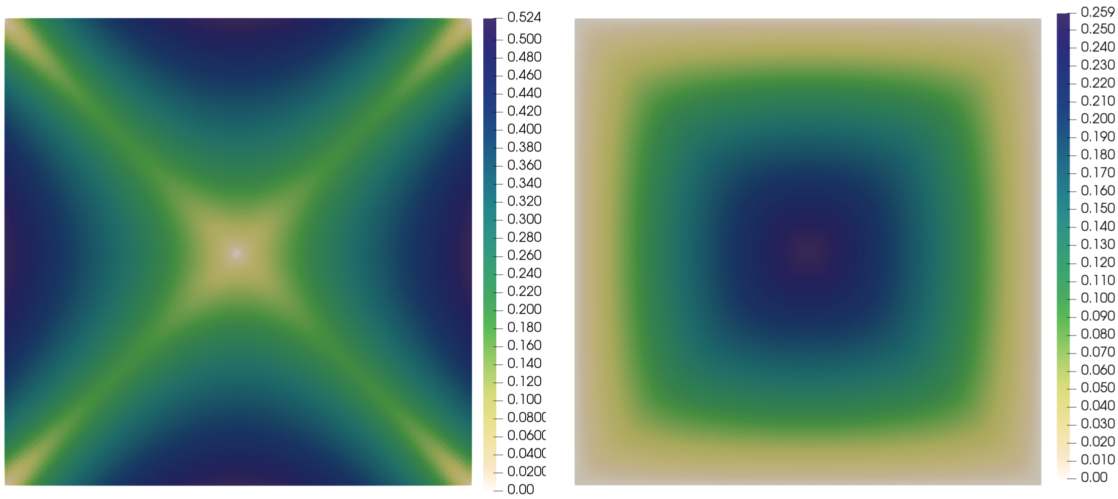

Figure 5.1: Example 1.1 Optimal a (left), and optimal u (right).

Example 5.1. We consider two cases in the framework of compliance optimization, i.e., $j(x, u) = f(x)u$. In Example 2.5 for $\psi(s) = s^2/2$ we provided an explicit solution when Ω is the unit ball and $f \equiv 1$. Here we solve numerically this problem in the non-radial case, considering the square $\Omega = [0, 1]^2$. In Figure 5.1 we show the computed optimal solutions, the optimal density a on the left, and the optimal function u on the right.

The second case corresponds to ψ given by (8). We solve numerically this problem in the cube $\Omega = [0,1]^2$, and considering the Lagrange multiplier $\gamma = 0.01141$ in order to work with a volumen constraint of 50% of each phase α and β. In Figure 5.2 we show the computed optimal solutions, the optimal density a on the left, and optimal function u on the right. $\square$

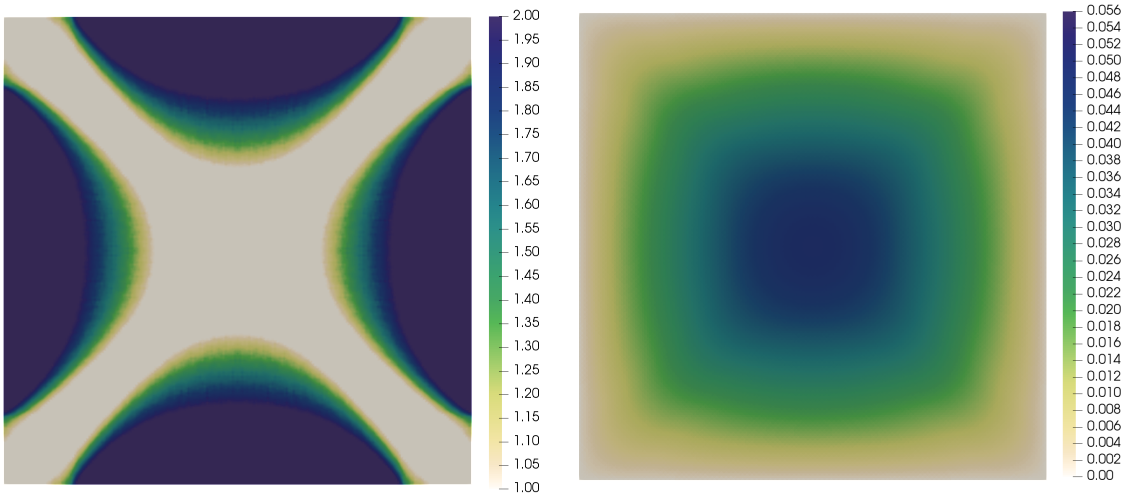

Figure 5.2: Example 1.2 Optimal a (left), and optimal u (right).

Example 5.2. We consider a case in the framework of energy optimization, i.e., $j(x,u) = -f(x)u$. We assume $f \equiv 1$ and ψ defined as (11) with $\gamma = 0.0142$ in order to assure a volume constraint of 50% of each phase α and β and we deal with the relaxed formulation (16). In Figure 5.3 we show the computed optimal solutions, the optimal density a on the left, and optimal function u on the right. $\square$

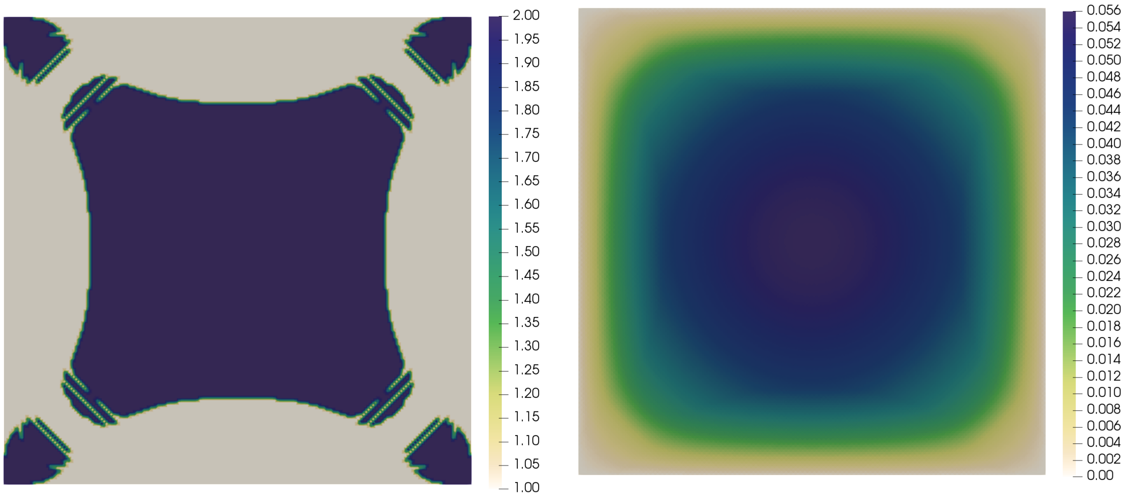

Figure 5.3: Example 2. Optimal a (left), and optimal u (right).

Example 5.3. In the last case we show a numerical evidence of non existence of an optimal density a_{opt} for problem of kind of (2). We follow the counterexample showed in Section 4. We consider $\Omega = B(0,1)$ the unit ball in $\mathbb{R}^2$ and the rest of data as in (22) with $\tau = 0.23539$. We have solved the relaxed formulation of problem (2) searching an optimal density t_{opt} and an optimal matrix A_{opt}. Firstly, we have considered the problem with $\varepsilon = 0$. In this case, the computed optimal matrix is $A_{opt} = \mu_{t_{opt}} I$, a scalar matrix. We show in Figure 5.4 on the left the optimal value of $\mu_{t_{opt}}$ corresponding to $\varepsilon = 0$. As hoped, it agrees with the function a_0 defined by (23). On the other hand, we consider the problem with $\varepsilon = 0.5$. In this case the computed optimal matrix A_{opt} is not scalar. We show in Figure 5.4 on the right the ratio λ_2/λ_1 with λ_i, $i = 1, 2$ the eigenvalues of A_{opt}. Observe that this ratio is not identically one, and therefore A_{opt} is a non-isotropic matrix.

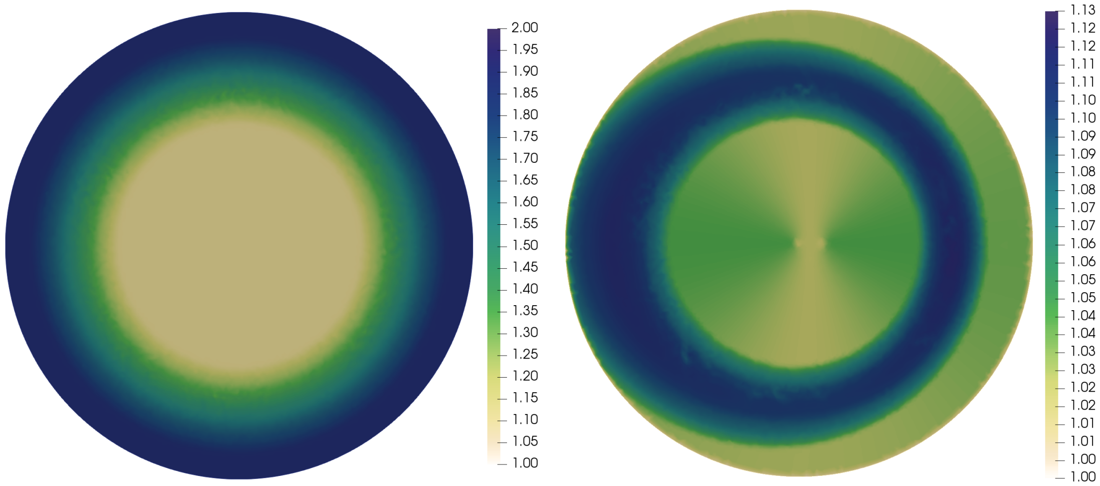

Figure 5.4: Example 3, $\lambda_1 = \lambda_2$ for $\varepsilon = 0$ (left), and λ_2/λ_1 for $\varepsilon = 0.5$.

Acknowledgments. The work of GB is part of the project 2017TEXA3H *"Gradient flows, Optimal Transport and Metric Measure Structures"* funded by the Italian Ministry of Research and University. GB is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

The work of JCD and FM is a part of the FEDER project PID2020-116809GB-I00 of the *Ministerio de Ciencia e Innovación* of the government of Spain.

References

[1] G. Allaire: *Shape Optimization by the Homogenization Method*, Applied Mathematical Sciences Vol. 146, Springer, New York (2002).

[2] G. Bouchitté, G. Buttazzo: *Characterization of optimal shapes and masses through Monge-Kantorovich equation*, J. Eur. Math. Soc. 3 (2001) 139–168.

[3] G. Bouchitté, G. Buttazzo, P. Seppecher: *Shape optimization solutions via Monge-Kantorovich equation*, C. R. Acad. Sci. Paris 324-I (1997) 1185–1191.

[4] G. Buttazzo, J. Casado-Díaz, F. Maestre: *On the regularity of optimal potentials in control problems governed by elliptic equations*, Adv. Calc. Var. 17/4 (2024) 1341–1364.

[5] G. Buttazzo, M.S. Gelli, D. Lučić: *Mass optimization problem with convex cost*, SIAM J. Math. Anal. 55/5 (2023) 5617–5642.

[6] E. Cabib: *A relaxed control problem for two-phase conductors*, Ann. Univ. Ferrara 33 (1987) 207–218.

[7] E. Cabib, G. Dal Maso: *On a class of optimum problems in structural design*, J. Optimization Theory Appl. 56 (1988) 39–65.

[8] J. Casado-Díaz: *Some smoothness results for the optimal design of a two composite material which minimizes the energy*, Calc. Var. Partial Diff. Equations 53 (2015) 649–673.

[9] J. Casado-Díaz: *Smoothness properties for the optimal mixture of two isotropic materials. The compliance and eigenvalue problems*, SIAM J. Cont. Optim. 53 (2015) 2319–2349.

[10] F. H. Clarke: *Multiple integrals of Lipschitz functions in the calculus of variations*, Proc. Amer. Math. Soc. 64 (1977) 260–264.

[11] E. De Giorgi, S. Spagnolo: *Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine*, Boll. Un. Mat. Ital. 8/4 (1973) 391–411.

[12] L. De Pascale, L. C. Evans, A. Pratelli: *Integral estimates for transport densities*, Bull. Lond. Math. Soc. 36/3 (2004) 383–395.

[13] S. Dweik: *BV estimates on the transport density with Dirichlet region on the boundary*, Calc. Var. Partial Diff. Equations 63/7 (2024), art. no. 173, 18 p.

[14] I. Ekeland: *Théorie des Jeux*, Presses Universitaires de France, Paris (1975).

[15] J. Goodman, R. V. Kohn, L. Reyna: *Numerical study of a relaxed variational problem for optimal design*, Comput. Methods Appl. Mech. Engrg. 57 (1986) 107–127.

[16] F. Hecht: *New development in FreeFem++*, J. Numer. Math. 20 (2012) 251–265.

[17] K. A. Lurie, A. V. Cherkaev: *Exact estimates of the conductivity of a binary mixture of isotropic materials*, Proc. Royal Soc. Edinburgh 104 A (1986) 21–38.

[18] R. Monneau, G. S. Weiss: *An unstable elliptic free boundary problem arising in solid combustion*, Duke Math. J. 136/2 (2007) 321–341.

[19] C. B. Morrey: *Multiple Integrals in the Calculus of Variations*, Grundlehren der Mathematischen Wissenschaften 130, Springer, New York (1966).

[20] F. Murat: *Contre-exemples pour divers problèmes ou le contrôle intervient dans les coefficients*, Ann. Mat. Pura Appl. 112 (1977) 49–68.

[21] F. Murat, L. Tartar: *Optimality conditions and homogenization*, in: *Nonlinear Variational Problems*, Int. Workshop Elba/Italy 1983, Research Notes in Mathematics 127, Pitman, London (1985) 1–8.

[22] H. Shahgholian: *The singular set for the composite membrane problem*, Commun. Math. Phys. 271 (2007) 93–101.

[23] L. Tartar: *Estimations fines des coefficients homogénéisés*, in:*Ennio De Giorgi Colloquium*, Paris 1983, Research Notes in Mathematics 125, Pitman, London (1985) 168–187.