Journal of Convex Analysis
Volume 19 (2012), No. 1, 281–294

A General Lower Bound for the Relaxation of an Optimal Design Problem with a General Quadratic Cost Functional, and a General Linear State Equation*

U. Fidalgo Prieto†
Departamento de Matemática,
Universidade de Aveiro, Aveiro, Portugal
ulisesfidalgoprieto@yahoo.es

Pablo Pedregal
Departamento de Matemáticas,
U. Castilla La Mancha, 13071 Ciudad Real, Spain
pablo.pedregal@uclm.es

Received: November 16, 2010

Recently, several particular problems in optimal design have been analyzed by using tools from non-convex, variational problems. As many of those have similarities, but also different features, we pretend to look at a full family of problems that includes most of those particular situations. Specifically, we examine an optimal design problem where anisotropy and/or non-ellipticity is permitted both in the state law, and the cost functional, which is quadratic in the gradient. In this generality, we are able to provide a general lower bound for the relaxed integrand (effective behavior) which is valid in all of these situations. Our philosophy, which has been introduced and implemented in simpler situations, leads to an elementary semi-definite mathematical programming problem for matrices depending on various parameters, that are precisely the variables for the relaxed problem. We also explore when this lower bound may turn out to be exact, and formulate a conjecture for the underlying relaxed problem.

1991 Mathematics Subject Classification: Primary 49J45

1. Introduction

Let us consider a material confined into a bounded domain $\Omega \subset \mathbb{R}^N$, $N > 1$. The medium is obtained by mixing two constituents with different electric permitivity and conductivity. Let Q_0 and Q_1 denote the two $N \times N$ symmetric matrices of electric permitivity corresponding to each phase. For each phase, we also denote by L_j, $j = 0, 1$, the (anisotropic) $N \times N$ symmetric matrix of conductivity. Set $0 \leq t_1 \leq 1$ the proportion of the constituent 1 into the mixture. This material occupies a space in the physical domain Ω which we denote by $E \subset \Omega$. Regarding

*Supported by MTM2010-19739 and MTM2009-12740-C03-01 from MICINN (Spain).
†The work of the first author was also supported by grant SFRH/BPD/62947/2009 from Fundação para a Ciência e a Tecnología, PT2009-0031, Acciones Integradas Portugal, Ministerio de Ciencia y Tecnología.

ISSN 0944-6532 / $ 2.50

the set E as our design variable, we introduce its characteristic function $\mathcal{X}_E$:

$$\mathcal{X}_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{if } x \notin E. \end{cases} \tag{1}$$

Hence the matrix of conductivity corresponding to the material as a whole is $L = \mathcal{X}_E L_1 + (1 - \mathcal{X}_E)L_0$ where

$$\int_E dx = \int_\Omega \mathcal{X}_E(x)dx = t_1 \int_\Omega dx = t_1|\Omega|.$$

We will also put $\mathcal{X}$ with no reference to the set E for simplicity.

Let us suppose that the amount of electric charge that goes into Ω is equal to what comes from it, which means that inside the material there is no source or sink of current. Such requirement determines a stable state, which may be expressed in terms of the electrostatic potential function, $u : \Omega \to \mathbb{R}$, as follows

$$\text{div } [L\nabla u(x)] \equiv \text{ div } \{[\mathcal{X}(x)L_1 + (1 - \mathcal{X}(x))L_0]\nabla u(x)\} = 0 \text{ in } \Omega, \tag{2}$$

where ∇u is the gradient of u. In addition, u is assumed to satisfy appropriate boundary conditions on $\partial\Omega$ in the form $u = u_0$.

Having fixed the electrostatic state through its equation of state (2), our purpose consists in finding a suitable distribution of the constituents into the domain Ω that minimizes the electrostatic energy

$$I(\mathcal{X}) = \int_\Omega \left[\mathcal{X}(\nabla u)^\top Q_1 \nabla u + (1 - \mathcal{X})(\nabla u)^\top Q_0 \nabla u\right](x)dx. \tag{3}$$

Altogether, we seek to solve the following optimal design problem

$$(\mathcal{P}) \begin{cases} \text{Minimize in } \mathcal{X}: & I(\mathcal{X}) = \int_\Omega \left[\mathcal{X}(\nabla u)^\top Q_1 \nabla u + (1 - \mathcal{X})(\nabla u)^\top Q_0 \nabla u\right](x)dx \\ \text{subject to} & \text{div } [\mathcal{X} L_1 \nabla u + (1 - \mathcal{X})L_0 \nabla u](x) = 0 \text{ in } \Omega, \\ & u_0 = u \text{ on } \partial\Omega, \ \int_\Omega \mathcal{X}_E(x)dx = t_1|\Omega|. \end{cases}$$

As it is well-known (see for instance the textbook [1] or [8]), this sort of problems typically lacks optimal solutions which means that the infimum may only be achieved by a sequence of more and more intricate subsets E_j of Ω. Under such circumstances, relaxation should be performed. It consists in looking for another minimization problem $(\mathcal{R})$ for which there does exist an optimal solution, this minimum has the same value as the infimum of $(\mathcal{P})$, and, more importantly, the optimal solution of the relaxed problem encodes the information about (some) minimizing sequences for the original problem. Unfortunately, in many situations of interest, exact descriptions for $(\mathcal{R})$ are not available, and only bounds can be made explicit ([1]). The main result in this note consists in finding a lower bound of the infimum corresponding to $(\mathcal{P})$, or equivalently, the minimum of its relaxed problem $(\mathcal{R})$. See also [6] for a comprehensive analysis of these problems from the point of view of homogenization and mechanics, as well as an updated lists of references and contributions to this field.

For some particular choices of the matrices Q_j and L_j, this kind of optimal design problems have been recently addressed ([2], [3], [5]) covering several circumstances. In most of these, restrictions are placed on the simplicity, properties, and/or relationship between the matrices in the cost and the matrices in the state law, as well as restrictions on dimension. Sometimes, a cost not depending explicitly on ∇u is only considered. In this contribution, we would like to remove all those constraints, and examine a general problem as $(\mathcal{P})$ under minimal assumptions on the structure of the matrices Q_j and L_j. To counterbalance our generality, we are not able to prove a full explicit relaxation result, but only a partial relaxation statement in the form of a general lower bound (Theorem 1.1 below).

Following the procedure described in [2] and [7], we can obtain a relaxed problem $(\mathcal{R})$, by using Young measures generated by sequences of pairs $\{(v_{1,k}, v_{2,k})\}$ for which we have the additional differential information that the divergence of the first component vanishes, while the second component is a gradient. Such class of Young measures, the so-called div-curl Young measures, has been formally introduced in [7].

As a way to place our goal in this contribution in context, suppose that the cost functional in (3) were linear in ∇u

$$I(\mathcal{X}) = \int_\Omega G(x) \cdot \nabla u(x)\, dx$$

for a certain given field $G \in L^2(\Omega; \mathbf{R}^N)$. In such a situation, being the cost functional weak continuous in $H^1(\Omega)$, a full relaxation of problem $(\mathcal{P})$ above (with such a linear cost functional) would amount to relaxing the differential constraint expressed in the state law. As a consequence of the main results in [2], one has that a full relaxation would be

$$\begin{cases} \text{Minimize in} & (t, \nabla u, V) : \displaystyle\int_\Omega G(x) \cdot \nabla u(x)\, dx \\ \text{subject to} & P_t(\nabla u(x), V(x)) \le 0 \text{ a.e. } x \in \Omega, \\ & \text{div } V = 0 \text{ in } \Omega,\ u_0 = u \text{ on } \partial\Omega,\ \displaystyle\int_\Omega t(x) dx = t_1 |\Omega|, \end{cases}$$

where P_t is a second-degree, homogeneous polynomial of its arguments that is completely explicit (see [2]). When the cost functional is quadratic in ∇u, it is no longer weakly continuous, and relaxation must account for the interaction between such cost functional and the state law. Trying to understand this new scenario from this perspective, and asses to what extent one can succeed in finding a relaxation, is our main objective here.

In our situation for problem $(\mathcal{P})$, we denote by $\nu = \{\nu_x\}_{x\in\Omega}$ the div-curl Young measure associated with the sequence $\{v_{1,k}, v_{2,k}\}$ where

$$v_{1,k}(x) = [\mathcal{X}_k(x) L_1 + (1 - \mathcal{X}_k(x)) L_0]\, \nabla u_k(x) \quad \text{and} \quad v_{2,k}(x) = \nabla u_k(x).$$

Since $v_{1,k}$ comes from the state equation then for each $x \in \Omega$ the support of ν_x satisfies

$$\operatorname{supp}(\nu_x) \subset \Lambda_1 \cup \Lambda_0, \quad \text{where } \Lambda_j = \left\{(\lambda, \rho) \in \mathbb{R}^N \times \mathbb{R}^N : \rho = L_j\lambda\right\}, \quad j = 0, 1.$$

Observe that we can write

$$\nu_x = t(x)\nu_{x,1} + (1 - t(x))\nu_{x,0}, \quad t \in [0, 1], \qquad \operatorname{supp}(\nu_{x,j}) \subset \Lambda_j, \quad j = 0, 1.$$

In order to state our main theorem, let us introduce some additional notation

$$\begin{cases} \lambda(x) = t(x)\lambda_1(x) + (1 - t)\lambda_0(x), \\ \rho(x) = tL_1\lambda_1(x) + (1 - t)L_0\lambda_0(x), \end{cases} \quad \text{where } \lambda_j(x) = \int_{\Lambda_j} y\, d\nu_{x,j}^{(1)}(y), \quad j = 0, 1, \tag{4}$$

with $\nu_x^{(1)}$ being the projection of ν_x onto the first copy of $\mathbb{R}^N$ of the product $\mathbb{R}^N \times \mathbb{R}^N$. We caution the reader not to confuse $\nu_x^{(1)}$ with $\nu_{x,1}$. The set of measures $\{\nu_x^{(1)} = t\nu_{x,1}^{(1)} + (1 - t)\nu_{x,0}^{(1)}\}_{x\in\Omega}$ refers to the first copies of the components of $\{\nu_x\}_{x\in\Omega}$, and it is the Young measure corresponding to the sequence $\{v_{1,k}\}$, while for each $x \in \Omega$, $\nu_{x,1}$ is the contribution in the convex combination $\nu_x = t\nu_{x,1} + (1 - t)\nu_{x,0}$ whose support is contained in the first manifold Λ_1.

Notice that the quantities λ_0 and λ_1 depend on the variables λ and ρ through the solution of the linear system in (4) as follows

$$\lambda_1(x) = \frac{1}{t} (L_1 - L_0)^{-1} (\rho(x) - L_0\lambda(x)) \quad \text{and}$$
$$\lambda_0(x) = \frac{1}{1 - t} (L_0 - L_1)^{-1} (\rho(x) - L_1\lambda(x)).$$

For each real σ, we define the set

$$\mathcal{A}_\sigma = \{\gamma \in \mathbb{R} : \text{ there is } j \in \{0, 1\} \text{ with } L_j \text{ having an eigenvalue of the same sign as } \sigma \text{ and } \det(Q_j - \gamma L_j) = 0\}.$$

For $\sigma = \lambda^\top\rho - t\lambda_1^\top L_1\lambda_1 - (1 - t)\lambda_0^\top L_0\lambda_0$, put $\overline{\gamma} = \overline{\gamma}(t, \rho, \lambda)$ so that

$$\overline{\gamma}\sigma = \min_\gamma \{\gamma\sigma : \gamma \in \mathcal{A}_\sigma\}.$$

Consider the problem

$$(\mathcal{S}) \quad \begin{cases} \text{Minimize in } (t, \rho = V, \lambda = \nabla u) : & \displaystyle\int_\Omega \left(\overline{\gamma}\sigma + t\lambda_1^\top Q_1\lambda_1 + (1 - t)\lambda_0^\top Q_0\lambda_0\right) dx, \\ \text{subject to:} & 0 \le t(x) \le 1, \ \displaystyle\int_\Omega t(x)dx = t_1|\Omega|, \\ & \operatorname{div} V(x) = 0 \text{ in } \Omega, \ u = u_0 \text{ on } \partial\Omega. \end{cases}$$

Note that $\overline{\gamma}$, σ, λ_1, and λ_0 depend upon $t(x)$, $V(x)$, and $\nabla u(x)$.

Theorem 1.1. *Suppose that the four matrices L_j, Q_j are symmetric, and $L_1 - L_0$ is a non-singular matrix. Then the variational problem $(\mathcal{S})$ is a sub-relaxation of $(\mathcal{P})$ in the sense that*

- *it admits optimal solutions;*
- *the minimum is a lower bound for the infimum of* $(\mathcal{P})$.

The proof of this result is based on a certain explicit semi-definite programming problem for matrices ([9]), after a suitable reformulation of the problem. The insistence on making the relaxed formulation as explicit as possible, and not only described in abstract terms, leads us to enlarge the class of competing measures in a more manageable way, at the expense of being contented with a partial relaxation, which still, we believe, keeps the clue to optimal solutions of the problem (see the final section). As a matter of fact, in some particular cases one can show that this lower bound is exact. In such situations, numerical approximations can some times be implemented based on this analysis. For example, this is so in the elliptic case when $L_j = \alpha_j$ identity, $\alpha_j > 0$ ([3]). Also in a hyperbolic setting where the matrices L_j have eigenvalues of different signs ([5]).

The paper is organized as follows. In Section 2, we recall the relevant facts about div-curl Young measures. These are mainly taken from [7]. Section 3 pushes the computation of the relaxation until it leads to a suitable semi-definite programming problem in the appropriate variables. The analysis of the constraint set for such mathematical programming problems is the content of Section 4. The main part of the paper is Section 5, where we prove our lower bound (Theorem 1.1). In Section 6, we include some remarks on when this lower bound may become exact, and some final remarks are written in Section 7.

2. Reformulation in terms of Young Measures

In this section, we adopt the notation of [7]. For a domain Ω in $\mathbb{R}^N$, we take a sequence of functions $\{v_k\} \subset L^\infty(\Omega)$, $v_k : \Omega \subset \mathbb{R}^N \to \mathbb{R}^m$, which converges weakly $*$ to a function $v \in L^\infty(\Omega)$ ($v_k \overset{*}{\rightharpoonup} v$). The fundamental theorem for Young measures ensures that there exists a family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ associated with the sequence of functions $\{v_k\}$ where $\operatorname{supp}(\nu_x) \subset \mathbb{R}^m$, such that for any continuous function $\phi : \mathbb{R}^m \to \mathbb{R}$, the mesurable function

$$\bar{\phi}(x) = \int_{\mathbb{R}^m} \phi(\lambda) d\nu_x(\lambda) = \langle \phi, \nu_x \rangle,$$

is the weak $*$ limit in $L^\infty(\Omega)$ of the sequence $\{\phi(v_k)\}$. This means that for all $g \in L^1(\Omega)$, we have that

$$\lim_{k\to\infty} \int_\Omega \phi(u_k(x))g(x)dx = \int_\Omega \bar{\phi}(v(x))g(x)dx = \int_\Omega g(x) \int_{\mathbb{R}^m} \phi(\lambda)d\nu_x(\lambda).$$

The family of measures $\nu = \{\nu_x\}_{x\in\Omega}$ is called the parametrized measure or Young measure associated with the sequence $\{v_k\}$.

In the present paper, we care about a class of parametrized measures whose elements are known as div-curl Young measures. In order to justify the main property of this class of measures, we need to recall the classical div-curl lemma ([1]).

Lemma 2.1 (Div-curl Lemma). *Let* $(\{v_{1,k}\},\{v_{2,k}\})$ *be a pair of sequences of functions in* $L^2(\Omega)$. *Each one of the two sequences converges weakly in* $L^2(\Omega)$ *to a function* v_j *and* $\{\operatorname{div} v_{1,k}\}$, $\{\nabla v_{2,k}\}$ *converge in* $H^{-1}(\Omega)$ *to* $\operatorname{div} v_1$ *and* ∇v_2, *respectively. Then the product* $v_{1,j}v_{2,j} \rightharpoonup v_1 v_2$ *in the sense of distributions.*

We define div-curl Young measures as follows.

Definition 2.2. A parametrized measure $\nu = \{\nu_x\}_{x\in\Omega}$, supported in $\mathbb{R}^{m\times N} \times \mathbb{R}^{m\times N}$, is called a ($L^2$-) div-curl Young measure, if it can be generated by a sequence

$$\{(v_{1,j}, \nabla v_{2,j})\}, \quad v_{1,j} : \Omega \to \mathbb{R}^{m\times N}, \ v_{2,j} : \Omega \to \mathbb{R}^m$$

with $\operatorname{div} v_{1,j} = 0$ for all j.

Notice that the div-curl lemma (Lemma 2.1) applies to sequences generating div-curl Young measures.

The following results appeared in [7]. Lemma 2.3 is an immediate consequence of Lemma 2.1, which is the fundamental commutation property. Lemma 2.4 says that for a.e. $a \in \Omega$, each measure ν_a, an individual member of a div-curl measure $\nu = \{\nu_x\}_{x\in\Omega}$, is, in its own right, a homogeneous (not dependent on $x \in \Omega$) div-curl Young measure, and Lemma 2.5 brings a way to recognize a certain class of div-curl Young measures. Finally, Proposition 2.6 is the main result which may serve to reformulate the relaxation of the original problem $(\mathcal{P})$ in terms of div-curl Young measures.

Lemma 2.3. *Let* $\nu = \{\nu_x\}_{x\in\Omega}$ *be a div-curl Young measure, then for a.e.* $x \in \Omega$

$$\int_{\mathbb{R}^{m\times N}\times\mathbb{R}^{m\times N}} \tau^\top \xi \, d\nu_x(\tau,\xi) = \left(\int_{\mathbb{R}^{m\times N}} \tau \, d\nu_x^{(1)}(\tau)\right)^\top \int_{\mathbb{R}^{m\times N}} \xi \, d\nu_x^{(2)}(\xi), \qquad (5)$$

where $\nu_x^{(i)}$, $i = 1, 2$, *are the marginals on the two components, respectively.*

Lemma 2.4. *A parameterized measure* $\nu = \{\nu_x\}_{x\in\Omega}$ *is a div-curl Young measure if and only if:*

- *For a.e.* $x \in \Omega$, *each individual* ν_x *is a homogeneous, div-curl Young measure itself.*
- *There exists a divergence-free vector field* v_1 *in* $L^2(\Omega;\mathbb{R}^{m\times N})$ *and* $v_2 \in H^1(\Omega;\mathbb{R}^m)$, *such that*

$$\operatorname{div}\left(\int_{\mathbb{R}^{m\times N}} \tau \, d\nu_x^{(1)}(\tau)\right) = 0, \qquad \nabla v_2(x) = \int_{\mathbb{R}^{m\times N}} \xi d\, \nu_x^{(2)}(\xi).$$

Lemma 2.5. *Suppose that* ρ_i, λ_i, $i = 1, 0$, *are four* N*-vectors such that*

$$(\rho_0^\top - \rho_1^\top)(\lambda_0 - \lambda_1) = 0. \qquad (6)$$

Then the probability measure

$$\mu = t\delta_{(\rho_1,\lambda_1)} + (1-t)\delta_{(\rho_0,\lambda_0)}$$

is a div-curl Young measure for all $t \in [0,1]$. If ν_1 and ν_0 are two div-curl Young measures with barycenters (ρ_1, λ_1) and (ρ_0, λ_0), respectively, such that (6) *holds, then*

$$\mu = t\nu_1 + (1-t)\nu_0$$

is a div-curl Young measure too for each $t \in [0,1]$.

Proposition 2.6. *A family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ can be generated by a sequence of pairs*

$$\{([\mathcal{X}_k(x)Q_1 + (1-\mathcal{X}_k(x))Q_0]\nabla u_k(x), \nabla u_k(x))\}, \tag{7}$$

with

$$\operatorname{div}(([\mathcal{X}_k(x)Q_1 + (1-\mathcal{X}_k(x))Q_0]\nabla u_k(x)) = 0 \quad \text{in } \Omega,$$

if and only if

- *Each measure ν_x is a div-curl Young measure, which is homogeneous and supported on the set*

$$\Lambda = \Lambda_1 \cup \Lambda_0, \qquad \nu_x = t(x)\nu_{x,1} + (1-t(x))\nu_{x,0},$$

  where each $\nu_{x,j}$, $j = 1, 0$, is supported on Λ_j.
- *There exists a divergence-free vector field v_1 in $L^2(\Omega; \mathbb{R}^{m\times N})$, and a field $v_2 \in H^1(\Omega; \mathbb{R}^m)$ such that*

$$v_1(x) = \int_{\mathbb{R}^{m\times N}} \rho d\nu_x^{(1)}(\rho), \qquad \nabla v_2(x) = \int_{\mathbb{R}^{m\times N}} \lambda d\nu_x^{(2)}(\lambda).$$

3. Reformulation in Terms of Div-Curl Young Measures. A semidefinite programming problem.

In order to reformulate the relaxation of the problem $(\mathcal{P})$ in terms of Young measures, we proceed as in [7]. The method is justified by Proposition 2.6. The relaxation of the original optimal design problem $(\mathcal{P})$ in terms of the family of div-curl Young measures may be written

$$\text{(R)} \quad \text{Minimize in } \nu: \quad \int_\Omega \left[tQ_1 : \int_{\mathbb{R}^N} \xi\xi^\top d\nu_{x,1}^{(1)}(\xi) + (1-t)Q_0 : \int_{\mathbb{R}^N} \xi\xi^\top d\nu_{x,0}^{(1)}(\xi) \right] dx$$

subject to

$$\nu = \{\nu_x\}_{x\in\Omega}, \; \nu_x = t(x)\nu_{x,1} + (1-t(x))\nu_{x,0} \; \text{ is a div-curl Young measure.}$$

$$\operatorname{supp} \nu_{x,j} \subset \Lambda_j, \quad j = 1, 0, \qquad \int_\Omega t(x)dx = t_1|\Omega|.$$

$$\operatorname{div} \int_{\mathbb{R}^{N\times N}} \tau d\nu_x(\xi, \tau) = 0, \quad \text{weakly in } \Omega,$$

$$\nabla u(x) = \int_{\mathbb{R}^{N\times N}} \xi d\nu_x(\xi, \tau), \quad u = u_0 \;\text{ on } \partial\Omega.$$

Notice that

$$Q_i : \xi\xi^\top = \xi^\top Q_i \xi \in \mathbb{R},$$

and that we have a unique component $m = 1$. To simplify the notation, we will drop the dependence on $x \in \Omega$. For example, we are going to write ν, ρ, λ, instead of

$$\nu_x, \quad \int_{\mathbb{R}^N} \tau d\nu_x(\xi, \tau), \quad \nabla u(x),$$

respectively.

Let us take ν, a div-curl Young measure, supported in the set

$$\Lambda = \Lambda_1 \cup \Lambda_0,$$

where

$$\Lambda_j = \{(\xi, \tau) \in \mathbb{R}^N \times \mathbb{R}^N : \tau = L_j \xi\}, \quad j = 1, 0.$$

Hence we may decompose $\nu = t\nu_1 + (1-t)\nu_0$, where ν_j is a probability measure (most likely not a div-curl Young measure itself) supported in Λ_j.

If we put

$$\lambda_j = \int_{\mathbb{R}^N} \xi d\nu_j^{(1)}(\xi), \quad j = 1, 0,$$

then

$$\lambda = t\lambda_1 + (1-t)\lambda_0, \qquad \rho = tL_1\lambda_1 + (1-t)L_0\lambda_0.$$

From these two identities, we can express λ_j in terms of λ and ρ. Namely,

$$\lambda_1 = \frac{1}{t}(L_0 - L_1)^{-1}(\rho - L_1\lambda), \qquad \lambda_0 = \frac{1}{1-t}(L_0 - L_1)^{-1}(L_0\lambda - \rho). \tag{8}$$

We have assumed above that $L_0 - L_1$ is not a singular matrix. The commutation property with the inner product on div-curl Young measures (5) yields

$$\int_\Lambda \xi^\top \tau d\nu(\xi, \tau) = \lambda^\top \rho.$$

But the integral on the left-hand side can be recast in the form

$$\begin{aligned} &t \int_{\mathbb{R}^N} \xi^\top L_1 \xi d\nu_1^{(1)}(\xi) + (1-t) \int_{\mathbb{R}^N} \xi^\top L_0 \xi d\nu_0^{(1)}(\xi) \\ &= tL_1 : \int_{\mathbb{R}^N} \xi\xi^\top d\nu_1^{(1)}(\xi) + (1-t)L_0 : \int_{\mathbb{R}^N} \xi\xi^\top d\nu_0^{(1)}(\xi). \end{aligned}$$

To find a lower bound we are going to retain just the relevant property expressed in the commutation just indicated, so that we regard feasible measures ν as Young measures which satisfy this commutation property, but are not necessarily a div-curl Young measure.

Let us introduce at this stage some further notation. Let us denote

$$X_j = \int_{\mathbb{R}^N} \xi\xi^\top d\nu_j^{(1)}(\xi), \quad j = 1, 0.$$

a convex combination of symmetric rank-one matrices. It is well known that

$$X_j \geq \lambda_j \lambda_j^\top, \quad j = 1, 0$$

in the usual sense of symmetric matrices.

On the other hand, the cost can clearly be written as

$$t \operatorname{tr}(Q_1 X_1) + (1 - t) \operatorname{tr}(Q_0 X_0),$$

where tr is the usual trace operator for matrices.

Therefore, in seeking a lower bound, we are led to consider the mathematical programming problem, which we are going to designate by $(\widetilde{R})$,

$$\begin{array}{ll} \text{Minimize in } (X_1, X_0): & t \operatorname{tr}(Q_1 X_1) + (1 - t) \operatorname{tr}(Q_0 X_0) \\ \text{subject to} & \lambda^\top \rho = t \operatorname{tr}(L_1 X_1) + (1 - t) \operatorname{tr}(L_0 X_0), \\ & \qquad X_j - \lambda_j \lambda_j^\top \geq 0, \ j = 1, 0. \end{array}$$

After the changes of variables $X_j = X_j - \lambda_j \lambda_j^\top$, $j = 1, 0$, we become interested in exploring

$$(\widetilde{R}) \quad \begin{array}{ll} \text{Minimize in } (X_1, X_0): & t \operatorname{tr}(Q_1 X_1) + (1 - t) \operatorname{tr}(Q_0 X_0) + t \lambda_1^\top Q_1 \lambda_1 \\ & \qquad + (1 - t) \lambda_0^\top Q_0 \lambda_0 \\ \text{subject to} & \sigma = t \operatorname{tr}(L_1 X_1) + (1 - t) \operatorname{tr}(L_0 X_0), \ X_j \geq 0, \ j = 0, 1, \end{array}$$

with $\sigma = \lambda^\top \rho - t \lambda_1^\top L_1 \lambda_1 - (1 - t) \lambda_0^\top L_0 \lambda_0$, independent of (X_0, X_1).

4. Analysis of the constraint set for $(\mathcal{R})$

A first important issue is to check the non-emptiness of the constraint set for our mathematical programming problem

$$\sigma = t \operatorname{tr}(L_1 X_1) + (1 - t) \operatorname{tr}(L_0 X_0), \quad X_j \geq 0, \ j = 0, 1. \tag{9}$$

What are the constraints that the triplet (t, λ, ρ) should verify so that a pair of matrices X_1 and X_0 can be found so that (9) is verified with

$$\sigma = \lambda^\top \rho - t \lambda_1^\top L_1 \lambda_1 - (1 - t) \lambda_0^\top L_0 \lambda_0 \tag{10}$$

and λ_j given in (8)?

Since for each $j = 0, 1$, L_j is a symmetric matrix, we can always find a diagonal matrix D_j such that $L_j = P_j^\top D_j P_j$, where P_j is an orthogonal matrix. Substituting this into the above identity (9), we obtain that

$$\sigma = t \operatorname{tr}(D_1 P_1 X_1 P_1^\top) + (1 - t) \operatorname{tr}(D_0 P_0 X_0 P_0^\top). \tag{11}$$

Let us analyze three different cases:

1. $L_0, L_1 \geq 0$;

2. $L_0, L_1 \leq 0$;
3. one of the two matrices has positive and negative eigenvalues, or the sign of the eigenvalues for the two of them is different.

In the first one, notice that saying $L_0, L_1 \geq 0$ is equivalent to ensuring that $D_0, D_1 \geq 0$. Hence if we take into account that $X_j \geq 0$, $j = 0, 1$, the diagonal components of each $P_j X_j P_j^\top$, $j = 0, 1$, are non-negative. That implies that the right-hand side on the equality must be non-negative, hence a necessary condition to find matrices X_j, $j = 0, 1$ which satisfy (11) is that

$$\sigma = \lambda^\top \rho - t\lambda_1^\top L_1 \lambda_1 - (1-t)\lambda_0^\top L_0 \lambda_0 \geq 0.$$

Furthermore this is a sufficient condition. Since $L_1 - L_0$ is not a singular matrix, one of the two matrices L_1 and L_0 is not the zero matrix. Without loss of generality, we can assume that $L_1 \neq 0$. Let us choose a positive element $d_s > 0$, of the diagonal matrix D_1, whose diagonal positition is $1 \leq s \leq N$. Fixing $\sigma \geq 0$, the relation (11) always holds for the matrices

$$X_0 = 0 \quad \text{and} \quad X_1 = \frac{\sigma}{t d_s} e_s e_s^\top, \tag{12}$$

where e_s is the $s-$th column of the identity matrix of size N.

In a similar way, case (2) can be treated so that (9) represents a non-empty set of matrices if and only if (t, λ, ρ) are such that $\sigma \leq 0$.

In the third case, we can always find matrices X_j, $j = 1, 0$, for which the constraints (9) and (11) are satisfied without any condition on σ. Suppose that $\sigma > 0$. We do not lose generality if we assume, as above, that $d_s > 0$. Hence, it is easy to see that the matrices X_1 and X_0 as in (12) satisfy (11). When $\sigma < 0$, we proceed in a similar way, but choosing $d_s < 0$.

5. Lower Bound for Solutions of $(\mathcal{R})$.

Let us start by obtaining a lower bound through the solution of $(\widetilde{R})$. We introduce some new notation. For X_j, $j = 1, 0$, we put

$$X_j = \sum_{\ell=1}^{N} v_{j,\ell} v_{j,\ell}^\top, \quad j = 0, 1, \tag{13}$$

where for each $j = 0, 1$, and $\ell = 1, \ldots, N$ $v_{j,\ell} \in \mathbb{R}^N$. Hence, problem $(\widetilde{R})$ can be reformulated in the following terms

Minimize in $(v_{1,1}, \ldots, v_{1,N}, v_{0,1}, \ldots, v_{0,N})$:

$$t \sum_{\ell=1}^{N} v_{1,\ell}^\top Q_1 v_{1,\ell} + (1-t) \sum_{\ell=1}^{N} v_{0,\ell}^\top Q_0 v_{0,\ell} + t\lambda_1^\top Q_1 \lambda_1 + (1-t)\lambda_0^\top Q_0 \lambda_0$$

subject to

$$\sigma = t \sum_{\ell=1}^{N} v_{1,\ell}^\top L_1 v_{1,\ell} + (1-t) \sum_{\ell=1}^{N} v_{0,\ell}^\top L_0 v_{0,\ell}.$$

According to the Lagrange's multiplier method, we define the augmented function by setting

$$\begin{aligned}&\Psi(v_{1,1},\dots,v_{1,N},v_{0,1},\dots,v_{0,N},\gamma)\\&=t\sum_{\ell=1}^{N}v_{1,\ell}^{\top}Q_1v_{1,\ell}+(1-t)\sum_{\ell=1}^{N}v_{0,\ell}^{\top}Q_0v_{0,\ell}+t\lambda_1^{\top}Q_1\lambda_1+(1-t)\lambda_0^{\top}Q_0\lambda_0\\&\qquad-\gamma\left(t\sum_{\ell=1}^{N}v_{1,\ell}^{\top}L_1v_{1,\ell}+(1-t)\sum_{\ell=1}^{N}v_{0,\ell}^{\top}L_0v_{0,\ell}-\sigma\right).\end{aligned}$$

We therefore arrive at the following requirements

$$\frac{\partial\Psi}{\partial v_{i,\ell}}=2h_j(Q_j-\gamma L_j)v_{j,\ell}=0,\quad j=0,1,\ \ell=1,\dots,N. \tag{14}$$

where $h_1=t$ and $h_0=1-t$. We also need that

$$\frac{\partial\Psi}{\partial\gamma}=t\sum_{\ell=1}^{N}v_{1,\ell}^{\top}L_1v_{1,\ell}+(1-t)\sum_{\ell=1}^{N}v_{0,\ell}^{\top}L_0v_{0,\ell}-\sigma=0. \tag{15}$$

Combining (14) with (15), we obtain that

$$t\sum_{\ell=1}^{N}v_{1,\ell}^{\top}Q_1v_{1,\ell}+(1-t)\sum_{\ell=1}^{N}v_{0,\ell}^{\top}Q_0v_{0,\ell}=\gamma\sigma. \tag{16}$$

Solving the initial problem is equivalent to obtaining the least value which γ is able to attain. From (14), and taking into account that the vectors $v_{j,\ell}$, $j=0,1$, $\ell=1,\dots,N$, cannot all be zero, we conclude that such value of γ coincides with the least value of $\gamma\sigma$ in which one of the two matrix equalities

$$(Q_j-\gamma L_j)v_{j,\ell}=0,\quad j=0,1, \tag{17}$$

holds. When L_1 and L_0 are non-singular matrices, γ is the eigenvalue of the two matrices $Q_jL_j^{-1}$, $j=0,1$, for which $\gamma\sigma$ is minimum. In this way, we have obtained a lower bound for the solutions of $(\mathcal{R})$, which turns out to be

$$LB=\gamma\sigma+t\lambda_1^{\top}Q_1\lambda_1+(1-t)\lambda_0^{\top}Q_0\lambda_0.$$

The end of the proof of Theorem 1.1 amounts to just putting back the dependence on x, and noting that the commutation property used in finding the lower bound, ensures that the integrand in problem $(\mathcal{S})$ is (div-curl)-quasiconvex (see [4]). The situation is equivalent to finding the polyconvexification as a lower bound for the quasiconvexification in the pure curl case.

6. When the lower bound becomes exact

We do not know if the solutions of $(\mathcal{R})$ and $(\widetilde{R})$ are the same, because the commutation property does not always imply that ν is a div-curl Young measure, or at least we do not know if it does. Only when $\sigma=0$, i.e.

$$\lambda^{\top}\rho-t\lambda_1^{\top}L_1\lambda_1-(1-t)\lambda_0^{\top}L_0\lambda_0=0,$$

can we be sure that the underlying measure is truly a div-curl Young measure, and hence the lower bound just obtained is an exact value. This is a direct consequence of Lemma 2.5.

Lemma 6.1. *Suppose that (t, λ, ρ) is such that $\sigma = 0$. Then the exact value of the relaxation (at this point) is*

$$t\lambda_1^\top Q_1\lambda_1 + (1-t)\lambda_0^\top Q_0\lambda_0,$$

and corresponds to the div-curl Young measure

$$t\delta_{(\lambda_1, L_1\lambda_1)} + (1-t)\delta_{(\lambda_0, L_0\lambda_0)}.$$

The proof of this fact is straightforward. Notice that when $\sigma = 0$, the optimal pair of matrices (X_1, X_0) in our semi-definite programming problem can be taken to vanish. This corresponds to the situation where

$$X_j = \int_{\mathbb{R}^N} \xi\xi^\top d\nu_j^{(1)}(\xi), \quad j = 0, 1,$$

is actually equal, after the change of variables indicated at the end of Section 3, to $\lambda_i\lambda_i^\top$. By strict convexity of the trace operator, this can only happen if

$$\nu_j^{(1)} = \delta_{\lambda_j}.$$

There is a number of situations where first-order laminate are always optimal. See [1] and [6] for further specific bibliography on this from the point of view of homogenization.

It is interesting to examine when this lower bound might become an exact value beyond first-order laminates. According to the computations above, and because σ does not have to vanish necessarily, we should look for a measure ν supported (at least) on three mass-points

$$\nu = (1-t)\delta_{(\lambda_0, L_0\lambda_0)} + t\left(s\delta_{(u, L_1u)} + (1-s)\delta_{(w, L_1w)}\right), \quad su + (1-s)w = \lambda_1, \ u - w = v, \tag{18}$$

where $s \in (0, 1)$ and v is the solution in (17) with the least eigenvalue $\gamma\sigma$ (without loss of generality, we may assume that γ corresponds to the pair (L_1, Q_1)). Notice how the optimal vector v furnishes the direction of lamination rather than the actual mass points. This is due to the way in which we have set up the auxiliary semi-definite programming problem. By solving for u and w in terms of v and λ_1, we can write the three-point probability measure as

$$\nu = (1-t)\delta_{(\lambda_0, L_0\lambda_0)} + t\left(s\delta_{(\lambda_1+(1-s)v, L_1\lambda_1+(1-s)L_1v)} + (1-s)\delta_{(\lambda_1-sv, L_1\lambda_1-sL_1v)}\right). \tag{19}$$

If we would like this measure to be a second-order laminate, when the pairs $(\lambda_0, L_0\lambda_0)$ and $(\lambda_1, L_1\lambda_1)$ are not compatible, then, in agreement with Lemma 2.3, the vector v must be a solution of the following system of equations:

$$\begin{cases} i) & 0 = (Q_1 - \overline{\gamma}L_1)v, \\ ii) & \sigma = tv^\top L_1 v, \end{cases} \tag{20}$$

and, in addition, we should have the compatibility conditions for a second-order laminate with three mass points

$$\begin{cases} iii) & (\lambda_1 - \lambda_0 - sv)^\top (L_1\lambda_1 - L_0\lambda_0 - sL_1 v) = 0, \\ iv) & \begin{array}{l} \left(\frac{1-t}{1-t+t(1-s)}\lambda_0 + \frac{t(1-s)}{1-t+t(1-s)}(\lambda_1 - sv) - \lambda_1 + (1-s)v\right)^\top \\ \cdot\left(\frac{1-t}{1-t+t(1-s)}L_0\lambda_0 + \frac{t(1-s)}{1-t+t(1-s)}(L_1\lambda_1 - sL_1 v) - L_1\lambda_1 + (1-s)L_1 v\right) = 0. \end{array} \end{cases} \tag{21}$$

This is an impressive system of equations for the vector v, and the parameter s. In general, we suspect it is an incompatible system of equations. For example, if the vector space which contains all solutions of the equation $0 = (Q_1 - \gamma L_1)v$ has dimension 1, the second equality $\sigma = tv^\top L_1 v$ determines only two vectors v for which it would be impossible to find s so that the other two equations hold. This implies that in general the measure (19) will not be a div-curl Young measure.

In the particular situation when $Q_1 = \bar{\gamma} L_1$, then the first equality in (20) is satisfied by any vector v, so that it becomes no constraint at all, and the system of equations is not incompatible. This is an indication that in this case, we might be able to show that the lower bound is in fact the exact value. Indeed, matrices Q_j do not play any role in this situation, and conditions so that (19) is a second-order laminate depend only on the matrices L_j. Notice that when one takes $s = 1/(2-t)$, the condition $iv)$ in (21) for the vector v is reduced to $iii)$. Hence, finding v is equivalent to solving the following system of quadratic equations

$$\begin{cases} ii) & \sigma = tv^\top L_1 v, \\ iii) & (\lambda_1 - \lambda_0 - sv)^\top (L_1\lambda_1 - L_0\lambda_0 - sL_1 v) = 0. \end{cases}$$

To treat in general this system of quadratic equations for v and s (even without the first equation in (20)) seems too complicated to perform explicitly. But at least we know that in some more particular situations these second order laminates have been found: when the matrices L_j are positive multiples of the identity ([3]), and in a hyperbolic situation when the matrices L_j reproduce a wave equation for two materials ([5]). In these works, those laminates were found directly by hand.

7. Final remarks

Our analysis in this contribution has focused on the optimal design problem

$$\text{Minimize in } \mathcal{X}: \quad I(\mathcal{X}) = \int_\Omega \left[\mathcal{X}(\nabla u)^\top Q_1 \nabla u + (1-\mathcal{X})(\nabla u)^\top Q_0 \nabla u\right](x)dx$$

subject to

$$\operatorname{div}\{[\mathcal{X}(x)L_1 + (1-\mathcal{X}(x))L_0]\nabla u(x)\} = 0, \quad \text{in } \Omega,$$

$u = u_0$ on $\partial\Omega$, and

$$\int_\Omega \mathcal{X}(x)\,dx = t_1|\Omega|.$$

$\mathcal{X}$ is a characteristic function indicating the location of the L_1-material. By means of an appropriate reformulation, a relaxation of this problem can be recast in the form

$$\text{Minimize in } (t, V, \nabla u): \quad \int_\Omega W(t(x), V(x), \nabla u(x))\, dx$$

subject to

$$u = u_0 \text{ on } \partial\Omega, \qquad \operatorname{div} V = 0 \text{ in } \Omega, \qquad 0 \le t(x) \le 1, \qquad \int_\Omega t(x)\, dx = t_1|\Omega|,$$

where a lower bound for W in the form $LB : \mathbf{R} \times \mathbf{R}^N \times \mathbf{R}^N \to \mathbf{R} \cup \{+\infty\}$ is given explicitly. Under some additional hypotheses on the structure of the different matrices, this lower bound becomes exact.

What is interesting is that the relaxed integrand W and the lower bound LB coincide (regardless of additional assumptions) in a certain subset of $\mathbf{R}\times\mathbf{R}^N\times\mathbf{R}^N$. This is the set of pairs that correspond to first-order laminates as indicated in Section 6. Based on previous experience in more specific situations, we conjecture that in fact the support of optimal solutions for the relaxed problem is actually contained in this special set. If this turns out to be true, then the numerical approximation of microstructures becomes much more manageable. This has already been explored in some particular instances ([3]).

References

[1] G. Allaire: Shape Optimization by the Homogenization Method, Applied Mathematical Sciences 146, Springer, New York (2001).

[2] O. Boussaid, P. Pedregal: Quasiconvexification of sets in optimal design, Calc. Var. Partial Differ. Equ. 34(2) (2009) 139–152.

[3] A. Donoso, P. Pedregal: Optimal design of 2-D conducting graded materials by minimizing quadratic functionals in the field, Struct. Multidisc. Optim. 30 (2005) 360–367.

[4] I. Fonseca, S. Müller: A-quasiconvexity, lower semicontinuity and Young measures, SIAM J. Math. Anal. 30(6) (1999) 1355–1390.

[5] F. Maestre, A. Münch, P. Pedregal: Optimal design under the one-dimensional wave equation, Interfaces Free Bound. 10 (2008) 87–117.

[6] G. W. Milton: The Theory of Composites, Cambridge Monographs on Applied and Computational Mathematics 6, Cambridge University Press, Cambridge (2002).

[7] P. Pedregal: Div-curl Young measures and optimal design in any dimension, Rev. Mat. Complut. 20(1) (2007) 239–255.

[8] L. Tartar: An introduction to the homogenization method in optimal design, in: Optimal Shape Design (Tróia, 1998), A. Cellina et al. (ed.), Springer Lecture Notes in Math. 1740, Springer, Berlin (2000) 47–156.

[9] L. Vandenberghe, S. Boyd: Semidefinite programming, SIAM Rev. 38(1) (1996) 49–95.