Grégoire ALLAIRE

Commissariat à l'Energie Atomique
DRN/DMT/SERMA
C.E.N. Saclay
F-91191 Gif Sur Yvette (FRANCE)

An Introduction to Homogenization and Bounds on Effective Properties Applied to Optimal Shape Design

School on Homogenization
ICTP, Trieste, September 6–17, 1993

CONTENTS

Introduction.

These lecture notes are a brief introduction to homogenization methods, and especially the theory of optimal bounds on effective properties of composite materials, in the context of optimal shape design. Shape optimization is already a huge field by itself, so these notes can merely give a flavor of that extremely wide and active area (let us mention a few reference books : [6], [21], [22], [23]). Of course, this is reflected in the many different methods which have been devised for analysing and solving problems in shape optimization. Since this course takes place in a "School on Homogenization", we focus on only one of these, the so-called homogenization method. Even so, it would take a whole book to cover that topic. A fortiori, these notes, corresponding to a four-hour course intended for graduate students, cannot possibly be more than a short initiation into optimal shape design using homogenization. Students (or readers) are assumed to already have a basic knowledge in homogenization (for example, the compactness theorem in G-convergence) and in the mathematical theory of composite materials (cf. the courses of A. Defranceschi [12] and L. Gibiansky [15]).

For many newcomers in that field, the association of homogenization and shape optimization seems a little weird or unnatural. However, for some of the "founding fathers" of homogenization, shape optimization was indeed their main motivation. (Remark that in the mid-seventies the expression "shape optimization" was sometimes replaced by "optimal control", the control being a domain.) The interrelation between homogenization, shape optimization, and relaxation in the calculus of variations has been recognized quite early by many authors (see [18], [19], [20]). At first sight, the use of homogenization in shape optimization seems to be just a trick to prove existence of optimal designs. But its importance goes far beyond this purely mathematical aspect, and recent contributions have demonstrated its usefulness for numerical computations (see [3], [7], and the proceedings [8]). Although these notes will not cover any numerical aspects of this problem, we just want to emphasize here the interest of that method from the point of view of industrial applications.

Section 1 briefly introduces some notions of shape optimization, and analyses a scalar model problem concerning heat conduction. Section 2 is devoted to the application of G-convergence to the relaxation of the model problem. In section 3, its relaxed formulation is computed with the help of the Hashin-Shtrikman variational principle. Finally, section 4 presents a more involved problem of shape optimization in elasticity.

1) Optimal shape design : a model problem.

Shape optimization is a branch of the calculus of variations, where the class of admissible solutions is a set of characteristic functions of domains (i.e. shapes). In this context, admissible solutions are also called admissible designs. A typical problem of shape optimization is to minimize a functional (i.e. a function $E(\chi)$, sometimes called a cost function or an energy function) over this set of admissible designs (i.e. characteristic functions χ which take the value 1 in their corresponding domain and 0 elsewhere).

There are two well-known difficulties associated with this type of problem when one tries to apply the so-called standard method of the calculus of variations. This method works in two steps : first, existence of a solution is proved by considering a minimizing sequence and applying some kind of lower semi-continuity result ; second, optimality conditions (also known as the Euler equation) are sought, which give a useful characterization of solutions. The first difficulty in shape optimization is that no solution may exist in the postulated class of admissible designs. More precisely, a minimizing sequence of characteristic functions usually does not converge to a characteristic function, but rather to a density function (taking all possible values between 0 and 1). In other words, a sequence of nearly optimal shapes can escape from the class of admissible designs, and converge in a larger class (including, for example, composite materials). The second difficulty is linked with the optimality conditions obtained by deriving the cost function. Unfortunately, in the general case of shape optimization, one cannot do variations of the functional $E(\chi)$ since the sum of two characteristic functions is usually not a characteristic function itself (with the notable, but limited, exception of variations along the boundary normal, see e.g. [23]). Thus, optimality conditions cannot be obtained if the original space of admissible designs is not enlarged.

To remove these two obstacles, the so-called relaxation procedure can be used : firstly, generalized solutions are defined, and secondly, the cost function is extended to this new class of admissible generalized solutions. This extension is precisely constructed in such a way that it is lower semi-continuous, thus implying the existence of an optimal generalized solution. The extended cost function is called the relaxed cost function. Of course, the space of generalized solutions must be larger than the space of classical solutions, but not too large to retain some knowledge of the minimizing sequences of the original problem. In other words, a relaxed formulation must satisfy the following conditions :

(1) generalized solutions include classical solutions, and the relaxed cost function is equal to the original one on classical solutions,

(2) the minimum values of the original and relaxed minimization problems are equal,

(3) there exists a solution of the relaxed problem, and any such solution is the limit of a minimizing sequence of the original one.

This idea of relaxation goes back to the work of L.C. Young [25], and has been much developed since (see e.g. [1], [10], [11], [13]). However, we do not intend to follow directly this general pattern here, but rather proceed in a more constructive, and physically sound, way by using homogenization.

Roughly speaking, homogenization is a natural and systematic method for computing a relaxed formulation for problems where the cost function $E(\chi)$ is defined through a state equation, which is a linear partial differential equation posed in the domain corresponding to the characteristic function χ. Furthermore, homogenization gives a physical meaning to the relaxation process by associating to minimizing sequences and generalized designs the concrete notions of infinitely fine mixtures and composite materials. This situation occurs for a large class of problems, but to keep things in a reasonably small format, we focus on a model problem, defined below. The connections between homogenization and relaxation for shape optimization have been explored notably in [18], [19], and [20].

We now introduce a simple model problem in the context of heat conduction. We have two materials at our disposal, with conductivities a_1 and a_2, the second one being a better conductor than the first one, i.e. $0 < a_1 < a_2$. They fill a given bounded domain Ω which is submitted to a fixed heat flux $\vec{J}_0$ on its boundary $\partial\Omega$ satisfying the thermal equilibrium condition $\int_{\partial\Omega} \vec{J}_0.\vec{n} = 0$. We assume that a_2, being a better conductor than a_1, is also more expensive, so that there is a constraint on the total amount of a_2, i.e. a_2 can occupy only a proportion α of Ω (with $0 \leq \alpha \leq 1$). The problem is now to find the best arrangement of a_1 and a_2 in Ω, which minimizes the heat energy stored in Ω (this is taken as a global measure of its conductivity).

To give a precise mathematical definition of this problem, we first give the form of the so-called state equation. Denoting by T the temperature, and by $\vec{J}$ the heat flux, it reads

$$\begin{cases} \vec{J}(x) = a(x)\,\nabla T(x) \quad with \;\; a(x) = [1-\chi(x)]a_1 + \chi(x)a_2 \\ div\;\vec{J} = 0 \quad in\;\Omega \\ \vec{J}.\vec{n} = \vec{J}_0.\vec{n} \quad on\;\partial\Omega \end{cases} \tag{1.1}$$

where $\chi(x)$ is the characteristic function of the subdomain occupied by a_2 (a measurable function on Ω satisfying $\chi(x) = 0$, or 1 a.e.). Then, the energy (or cost) function is

$$\int_\Omega a(x)\nabla T(x).\nabla T(x)\,dx = \int_\Omega a(x)^{-1}\vec{J}(x).\vec{J}(x)\,dx \tag{1.2}$$

which is to be minimized under the constraint

$$\int_\Omega \chi(x)\,dx \leq \alpha\,|\Omega|. \tag{1.3}$$

To further simplify the presentation, we introduce a positive Lagrange parameter λ in order to add the constraint (1.3) to the energy function (1.2). This gives the new cost function to be minimized without any constraint

$$E(\chi) = \int_\Omega a(x)^{-1}\vec{J}(x).\vec{J}(x)\,dx + \lambda\int_\Omega \chi(x)\,dx. \tag{1.4}$$

The minimization of (1.4) is our model problem of shape optimization. (In the degenerate case $a_1 = 0$, we really optimize the shape of the domain containing a_2 with zero-flux boundary condition.) Remark that the compatibility condition $\int_{\partial\Omega} \vec{J}_0.\vec{n} = 0$ is always satisfied for a divergence-free function $\vec{J}_0 \in L^2(\Omega)^N$, and that equation (1.1) has a unique solution $T \in H^1(\Omega)/\mathbb{R}$ for any measurable characteristic function χ. This implies that $E(\chi)$ is well defined for any such χ. Nevertheless, we claim that the minimization of (1.4) is ill-posed, i.e. that $E(\chi)$ has usually no minimizer among characteristic functions. To support that claim, we give an equivalent form of (1.4) which turns out not to be lower semicontinuous.

Proposition 1.1.

The minimization of (1.4) is equivalent to the minimization of

$$\int_\Omega Min\left[a_1^{-1}\vec{j}.\vec{j}\ ,\ a_2^{-1}\vec{j}.\vec{j} + \lambda\right]\ dx \tag{1.5}$$

among admissible currents $\vec{j}$, i.e. satisfying

$$\begin{cases} div\ \vec{j} = 0 & in\ \Omega \\ \vec{j}.\vec{n} = \vec{J}_0.\vec{n} & on\ \partial\Omega. \end{cases} \tag{1.6}$$

Namely, any minimizer $\vec{j}^*$ of (1.5) (if any) yields a minimizer χ^* of (1.4), related by

$$\chi^*(x) = 1 \quad \text{if } a_2^{-1}\vec{j}^*.\vec{j}^* + \lambda \le a_1^{-1}\vec{j}^*.\vec{j}^*\ ,\ \chi^*(x) = 0 \quad \text{if } not. \tag{1.7}$$

Reciprocally, the solution $\vec{J}^*$ of equation (1.1), associated to a minimizer χ^* of (1.4) (if any), is a minimizer of (1.5).

Proof.

By the dual variational principle, the solution $\vec{J}$ of equation (1.1), for a fixed χ, is the minimizer of

$$\int_\Omega a(x)^{-1}\vec{J}(x).\vec{J}(x)\ dx = \mathrm{Min}_{\begin{cases} div\ \vec{j} = 0 \ \ in\ \Omega \\ \vec{j}.\vec{n} = \vec{J}_0.\vec{n} \ \ on\ \partial\Omega \end{cases}} \int_\Omega a(x)^{-1}\vec{j}(x).\vec{j}(x)\ dx. \tag{1.8}$$

Inserting (1.8) in (1.4), it is perfectly legitimate to interchange the order of minimizations in χ and $\vec{j}$. For fixed $\vec{j}$, the minimization in χ is pointwise and gives the integrand in (1.5). In view of (1.8), the interrelation between the minimizers χ^* and $\vec{j}^*$ is now obvious. This completes the proof of Proposition 1.1.

The integrand in (1.5) is clearly not convex if $\lambda > 0$. Consequently, (1.5) is not lower semi-continuous by a classical result (at least in 2-D, see e.g. Chapter X in [13]), that won't be detailed here. The upshot is this : the model problem (1.4) may have no solution ; thus, we need to introduce its relaxation, but we don't want to compute it with the general arguments of the relaxation theory, rather we use the homogenization theory (see section 2). The reason for not using the general theory is that it works only for scalar state equation, as considered in our model problem. In particular, it does not apply to the other model problem introduced in section 4, where the state equation is actually a system of equations (elasticity). On the contrary, the homogenization method goes through exactly in the same way for that problem (but the computations are a little more tricky).

Before introducing homogenization in the next section, we give the final result for the formulation (1.5) of our model problem.

Theorem 1.2.

The relaxed formulation of (1.5) is the minimization, under the same constraint (1.6), of the relaxed energy

$$\int_\Omega D(\vec{j}(x))\, dx \tag{1.9}$$

with

$$D(\vec{j}) = \begin{cases} a_2^{-1}\vec{j}.\vec{j} + \lambda & \text{if } |\vec{j}| \geq a_2\left(\dfrac{\lambda}{a_2-a_1}\right)^{1/2} \\ 2\left(\dfrac{\lambda}{a_2-a_1}\right)^{1/2}|\vec{j}| - \dfrac{\lambda a_1}{a_2-a_1} & \text{if } a_2\left(\dfrac{\lambda}{a_2-a_1}\right)^{1/2} > |\vec{j}| > a_1\left(\dfrac{\lambda}{a_2-a_1}\right)^{1/2} \\ a_1^{-1}\vec{j}.\vec{j} & \text{if } a_1\left(\dfrac{\lambda}{a_2-a_1}\right)^{1/2} \geq |\vec{j}| \end{cases}$$

Of course, this relaxed formulation satisfies the condition (1)-(3) defined above. Remark that the integrand of (1.9) is just the convexification of the integrand of (1.5) as can be expected (see e.g. Chapter X in [13]). However, relaxation is not always synonymous of convexification : in the vector-valued case the relaxed integrand is the quasi-convexification (see section 4 for references).

2) G-convergence and relaxation.

The first part of this section recalls fundamental notions of G-convergence in the particular context of the model problem of section 1. (For proofs and generalizations, we refer to the course of A. Defranceschi [12] and the references therein.) In a second part, G-convergence is applied to the relaxation of the model problem. We keep the notations of section 1, i.e. Ω is a bounded domain of $\mathbb{R}^N$, a_1 and a_2 are two conductivities satisfying $0 < a_1 < a_2$, and χ_ε is a sequence of measurable characteristic functions indexed by a positive real ε. For this sequence χ_ε, we define an associated sequence of conductivities

$$a_\varepsilon(x) = [1-\chi_\varepsilon(x)]a_1 + \chi_\varepsilon(x)a_2. \tag{2.1}$$

We denote by $L^2(\Omega,div)$ the space of divergence-free functions in $L^2(\Omega)^N$. For a given $\vec{J}_0$ in $L^2(\Omega,div)$, let T_ε be the unique solution in $H^1(\Omega)/\mathbb{R}$ of the model problem

$$\begin{cases} \vec{J}_\varepsilon(x) = a_\varepsilon(x)\,\nabla T_\varepsilon(x) \\ div\ \vec{J}_\varepsilon = 0 \quad in\ \Omega \\ \vec{J}_\varepsilon.\vec{n} = \vec{J}_0.\vec{n} \quad on\ \partial\Omega. \end{cases} \tag{2.2}$$

Definition 2.1.

The sequence of conductivities a_ε is said to G-converge to an homogenized or effective conductivity tensor A^* if, for any $\vec{J}_0$ in $L^2(\Omega,div)$, the associated sequence of solutions T_ε converges weakly in $H^1(\Omega)/\mathbb{R}$ to the solution T of the homogenized problem

$$\begin{cases} \vec{J}(x) = A^*(x)\,\nabla T(x) \\ div\ \vec{J} = 0 \quad in\ \Omega \\ \vec{J}.\vec{n} = \vec{J}_0.\vec{n} \quad on\ \partial\Omega. \end{cases} \tag{2.3}$$

This definition makes sense because of the following compactness theorem.

Theorem 2.2.

From any sequence of conductivities a_ε, defined by (2.1), we can extract a subsequence, and there exists an effective tensor A^*, such that this subsequence G-converges to A^*.

Of course, G-convergence can be defined for more general sequences a_ε, and more general elliptic problems than (2.2). There are a number of properties of G-convergence that we recall in the next proposition.

Proposition 2.3.

Any G-limit A^* of a sequence a_ε is a symmetric, positive definite, bounded, and measurable matrix on Ω. Furthermore, it does not depend on Ω, and on the type of boundary condition in (2.2). Finally, the sequence of energies associated to equation (2.2) converge to the homogenized energy of (2.3), i.e.

$$\int_\Omega a_\varepsilon(x)^{-1}\vec{J}_\varepsilon(x).\vec{J}_\varepsilon(x)\,dx \;\to\; \int_\Omega A^*(x)^{-1}\vec{J}(x).\vec{J}(x)\,dx. \tag{2.4}$$

By remarking that the set of measurable characteristic functions on Ω is weakly compact "star" in the set of measurable densities on Ω (i.e. functions which take their values in the whole interval [0;1]), we easily obtain the following corollary of Theorem 2.2.

Corollary 2.4.

From any sequence of characteristic functions χ_ε and conductivities a_ε, we can extract a subsequence (still denoted by ε), and there exist a density θ and an effective tensor A^*, such that, for this subsequence, one has

$$\chi_\varepsilon(x) \rightharpoonup \theta(x) \text{ in } L^\infty(\Omega;[0;1]) \text{ weak } *, \text{ and } a_\varepsilon \text{ G-converges to } A^*. \tag{2.5}$$

For any density $\theta(x) \in L^\infty(\Omega;[0;1])$, the precise set of associated G-limits $A^*(x)$ is denoted by $G_{\theta(x)}$ and is called the G-closure of a_1 and a_2 with density $\theta(x)$. The following theorem of G. Dal Maso and R. Kohn indicates that there is a pointwise (or local) definition of the G-closure and that it is enough to consider effective tensors that arise by periodic homogenization of a_1 and a_2.

Theorem 2.5.

Denote by P_θ the set of all effective tensors obtained by periodic homogenization of a_1 and a_2 in proportions $(1-\theta)$ and θ. Then, any possible effective tensor $A^*(x) \in G_{\theta(x)}$ belongs to the closure of P_θ where $\theta(x) = \theta$ (almost everywhere in Ω). Furthermore, any tensor $A(x)$, such that $A(x) \in \bar{P}_{\theta(x)}$, is a G-limit, i.e. belongs to $G_{\theta(x)}$.

We have now finished with results on G-convergence. Let us comment a little on the so-called G-closure problem, i.e. finding the set of all possible effective conductivity tensors obtained by mixing a_1 and a_2. Since the work of F. Murat and L. Tartar [20], [24], and K. Lurie and A. Cherkaev [19], a complete answer for our model problem is available, i.e. an algebraic closed form of the G-closure is known (see the course of L. Gibiansky [15] for details). We won't use it here. Rather, we will use a partial knowledge of the G-closure furnished by so-called optimal bounds on effective properties (see section 3). The justification of our method is that we ultimately want to generalize our result to the model problem of section 4 in elasticity. Since the G-closure is unfortunately still unknown for elasticity, our method is the only available route to relaxation in this case.

We turn to the relaxation of the shape optimization model problem (1.4).

Theorem 2.6.

The relaxed formulation of (1.4) is the minimization of the relaxed energy

$$\tilde{E}(\theta,A^*) = \int_\Omega A^*(x)^{-1}\vec{J}(x).\vec{J}(x)\,dx + \lambda \int_\Omega \theta(x)\,dx, \tag{2.6}$$

where $\vec{J}(x)$ is the solution of the homogenized problem (2.3). The minimization of (2,5) takes place over all densities $\theta(x) \in L^\infty(\Omega;[0;1])$ and all effective tensors $A^*(x) \in G_{\theta(x)}$.

Proof.

Let χ_ε be a minimizing sequence of (1.4). By applying Corollary 2.4 to the corresponding sequence of conductivities a_ε, up to a subsequence one has

$$\chi_\varepsilon(x) \rightharpoonup \tilde{\theta}(x) \text{ in } L^\infty(\Omega;[0;1]) \text{ weak } *, \text{ and } a_\varepsilon \text{ G-converges to } \tilde{A}^*.$$

Passing to the limit in the state equation (2.2) yields the relaxed state equation (2.3). Furthermore, Proposition 2.3 implies that

$$E(\chi_\varepsilon) \to \tilde{E}(\tilde{\theta},\tilde{A}^*).$$

Since, by definition, any couple (θ, A^*) in $L^\infty(\Omega;[0;1])\times G_\theta$ is attained as a limit of some sequence of characteristic functions, the above limit $(\tilde{\theta}, \tilde{A}^*)$ is a minimizer of (2.6). Thus, (2.6) satisfies all the required properties of a relaxed problem for (1.4) : it has a solution, its minimum value is also that of (1.4), and its minimizers are attained by minimizing sequences of (1.4).

Theorem 2.6 gives the relaxation of our model problem in shape optimization, in a form which is not suitable for our purpose. Indeed for numerical computations, it requires the knowledge of the G-closure to minimize (2.6) on that precise set. Thus, in the same spirit of Proposition 1.1, we give an equivalent formulation of (2.6).

Proposition 2.7.

The minimization of (2.6) is equivalent to the minimization of

$$\int_\Omega Min_{\begin{cases} 0 \leq \theta \leq 1 \\ A^* \in G_\theta \end{cases}} \left[A^{*-1} \vec{j}(x).\vec{j}(x) + \lambda\theta(x) \right] dx \tag{2.7}$$

among admissible currents $\vec{j}$, i.e. satisfying

$$\begin{cases} div\ \vec{j} = 0 & in\ \Omega \\ \vec{j}.\vec{n} = \vec{J}_0.\vec{n} & on\ \partial\Omega. \end{cases} \tag{2.8}$$

Proof.

By using the dual variational principle as in Proposition 1.1, we can interchange the order of minimization in $\vec{j}$ and in (θ, A^*). By virtue of Theorem 2.5 the definition of G_θ is pointwise (as is the constraint $0 \leq \theta \leq 1$). Thus, the minimization in (θ, A^*) can be interchanged with the integration on Ω, yielding (2.7).

The integrand in the new formulation (2.7) is not very explicit, and still involves the G-closure G_θ. However, the main interest of (2.7) is that, even if we don't know G_θ, we can explicitly compute the minimum in A^* of the dual energy $<A^{*-1}\vec{j}.\vec{j}>$ for a fixed field $\vec{j}$. The result is especially simple.

Proposition 2.8.

When homogenizing materials a_1 and a_2 in proportions $(1-\theta)$ *and* θ *respectively, the minimum value of the effective dual energy is*

$$\text{Min}_{A^* \in G_\theta} < A^{*-1}\vec{j}.\vec{j} > \; = \; < \left((1-\theta)a_1 + \theta a_2 \right)^{-1} \vec{j}.\vec{j} >. \tag{2.9}$$

The minimum value in (2.9) is a so-called optimal lower bound on the dual energy. Proposition 2.8 will be proved in the next section with the help of the Hashin-Shtrikman variational principle, which is a very general and powerful tool for computing such bounds. After this minimization in A^*, it remains an obvious 1-D minimization in θ to compute an explicit form of the integrand in (2.7). This is done in the next Proposition, which implies Theorem 1.2.

Proposition 2.9.

Let m be a constant defined by $m = \left(\dfrac{\lambda}{a_2 - a_1}\right)^{1/2}$. *Denoting by $\tilde{\theta}$ the optimal value of the density, the integrand in (2.7) is*

$$< a_2^{-1}\vec{j}.\vec{j} > + \lambda \qquad \textit{if } |\vec{j}| \geq m a_2 \text{ and } \tilde{\theta} = 1 \tag{2.10}$$

$$2m\,|\vec{j}| - m^2 a_1 \qquad \text{if } m a_2 > |\vec{j}| > m a_1 \text{ and } \tilde{\theta} = \frac{m^{-1}|\vec{j}| - a_1}{a_2 - a_1} \tag{2.11}$$

$$< a_1^{-1}\vec{j}.\vec{j} > \qquad \text{if } m a_1 \geq |\vec{j}| \text{ and } \tilde{\theta} = 0. \tag{2.12}$$

Proof.

We have to minimize in θ the function

$$\left(a_1 + \theta(a_2 - a_1)\right)^{-1} |\vec{j}|^2 + \lambda\theta.$$

The optimal value of θ is easily seen to be

$$\tilde{\theta} = \frac{m^{-1}|\vec{j}| - a_1}{a_2 - a_1}.$$

The different regimes (2.10)-(2.12) arise because of the constraint $0 \leq \tilde{\theta} \leq 1$. This easy calculation is left to the reader.

Remark that the original problem (1.4) of shape optimization has been relaxed in a minimization of a non-linear dual energy defined by (2.10)-(2.12). However, this relaxed formulation is still a problem of shape optimization, thanks to the optimality condition for the density in Proposition 2.9. Numerically, one process by computing the minimizer $\vec{J}$ of (2.7), then recovering an optimal shape, defined by its density $\tilde{\theta}$. Remark also that where the heat flux $\vec{j}$ is small, the bad conductor a_1 is chosen ($\tilde{\theta} = 0$), where it is large, the good one a_2 is preferred ($\tilde{\theta} = 1$), and for intermediate values, composites arise.

3) The Hashin-Shtrikman variational principle.

This section is devoted to the proof of Proposition 2.8, i.e. to the computation of an optimal lower bound on the effective dual energy $< A^{*-1}\vec{j},\vec{j} >$. To this end, we introduce the Hashin-Shtrikman variational principle (see their original paper [17] or the course of L. Gibiansky [15] ; here, we follow the lines of [4]). We begin with the definition of such an optimal bound.

Definition 3.1.

A lower bound on the effective dual energy $< A^{*-1}\vec{j},\vec{j} >$ is a function f depending only on θ, a_1, a_2, and $\vec{j}$ such that, for any effective tensor $A^* \in G_\theta$

$$< A^{*-1}\vec{j},\vec{j} > \;\geq\; f(\theta, a_1, a_2, \vec{j}). \tag{3.1}$$

This bound is called optimal if, for any value of θ, a_1, a_2, *and* $\vec{j}$, *one can find a particular effective tensor* A^* *for which there is equality in (3.1). The corresponding microstructure (i.e. the arrangement of* a_1 *and* a_2 *in this composite) is also called optimal.*

By virtue of Theorem 2.5, it is enough to establish the bound (3.1) for effective tensors obtained by periodic homogenization. This allows us to use the convenient following formula.

Proposition 3.2.

Let A^* *be the homogenized tensor obtained by homogenization of* a_1 *and* a_2, *distributed in the periodic cell* $Y = [0;1]^N$ *with characteristic functions* $[1-\chi(y)]$ *and* $\chi(y)$ *respectively. Then, for any constant vector* $\vec{j}$, A^* *is characterized by*

$$< A^{*-1}\vec{j},\vec{j} > \;=\; \mathrm{Min}_{\vec{\phi}} \int_Y < \left((1-\chi(y))a_1^{-1} + \chi(y)a_2^{-1}\right)(\vec{j}+\vec{\phi}(y)),(\vec{j}+\vec{\phi}(y)) > dy \tag{3.2}$$

where the minimization is subject to the constraints

$$\vec{\phi} \in H^1_\#(Y)^N \;,\; \int_Y \vec{\phi}(y)\, dy \;=\; 0 \;,\; div\, \vec{\phi} \;=\; 0 \quad in\ Y.$$

(For a proof of that well-known result, see e.g. [9].) Since formula (3.2) involves periodic functions on the unit cube, it is tempting to use Fourier analysis to evaluate it. This is indeed the main idea behind the Hashin-Shtrikman variational principle that we can now state.

Theorem 3.3

Let A^* *be an effective tensor obtained by homogenization of* a_1 *and* a_2 *(with* $0 < a_1 < a_2$*), in proportions* $(1-\theta)$ *and* θ *respectively. The Hashin-Shtrikman variational principle is the following lower bound*

$$< A^{*-1}\vec{j},\vec{j} > \;\geq\; < a_2^{-1}\vec{j},\vec{j} > + (1-\theta)\ \mathrm{Max}_{\vec{\eta}} \left[2<\vec{j},\vec{\eta}> - <(a_1^{-1}-a_2^{-1})^{-1}\vec{\eta},\vec{\eta}> - \theta g(\vec{\eta})\right] \tag{3.3}$$

where the maximum is taken over all constant vectors $\vec{\eta}$, *and* $g(\vec{\eta})$ *is the so-called non-local term defined by*

$$g(\vec{\eta}) \;=\; a_2\ \mathrm{Sup}_{\vec{k}} \left[|\vec{\eta}|^2 - \frac{(\vec{\eta}.\vec{k})^2}{|\vec{k}|^2} \right] \tag{3.4}$$

where the supremum is taken over all vectors $\vec{k}$ *with integer components (*$\vec{k}$ *is actually the Fourier variable corresponding to* y*).*

Proof.

We start from the definition of A^* (3.2). Adding and subtracting the reference energy $< a_2^{-1}(\vec{j}+\vec{\phi}(y)),(\vec{j}+\vec{\phi}(y)) >$ gives

$$< A^{*-1}\vec{j},\vec{j} > \;=\; \mathrm{Min}_{\vec{\phi}} \left[\int_Y (1-\chi(y))< [a_1^{-1}-a_2^{-1}](\vec{j}+\vec{\phi}),(\vec{j}+\vec{\phi}) > dy \;+\; \int_Y < a_2^{-1}(\vec{j}+\vec{\phi}),(\vec{j}+\vec{\phi}) > dy \right]. \tag{3.5}$$

Let us rewrite the first term in the right hand side of (3.5). Since $a_1 < a_2$, by convex duality we have

$$\int_Y < (1-\chi(y))[a_1^{-1}-a_2^{-1}](\vec{j}+\vec{\phi}),(\vec{j}+\vec{\phi}) > dy \;=\; \operatorname{Sup}_{\vec{\eta}(y)} \int_Y (1-\chi(y))\Big[2< \vec{\eta},(\vec{j}+\vec{\phi}) > - < [a_1^{-1}-a_2^{-1}]^{-1}\vec{\eta},\vec{\eta} >\Big]\, dy. \tag{3.6}$$

Here, $\vec{\eta}(y)$ ranges over periodic vector fields. One can get an inequality by making a special choice of $\vec{\eta}(y)$ in (3.6). We take $\vec{\eta}(y)$ being constant in the domain of integration in (3.6). By integration, this implies

$$\int_Y < (1-\chi(y))[a_1^{-1}-a_2^{-1}](\vec{j}+\vec{\phi}),(\vec{j}+\vec{\phi}) > dy \;\geq\; 2(1-\theta)< \vec{\eta},\vec{j} >$$
$$- (1-\theta)< [a_1^{-1}-a_2^{-1}]^{-1}\vec{\eta},\vec{\eta} > + 2\int_Y (1-\chi(y))< \vec{\eta},\vec{\phi}(y) > dy$$

for any constant vector $\vec{\eta}$. Substitution into (3.5) yields, after a bit of simplification

$$< A^{*-1}\vec{j},\vec{j} > \;\geq\; < a_2^{-1}\vec{j},\vec{j} > + 2(1-\theta)< \vec{\eta},\vec{j} > - (1-\theta)< [a_1^{-1}-a_2^{-1}]^{-1}\vec{\eta},\vec{\eta} > \tag{3.7}$$
$$+ \operatorname{Min}_{\vec{\phi}} \int_Y \Big[< a_2^{-1}\vec{\phi}(y),\vec{\phi}(y) > - 2\chi(y)< \vec{\eta},\vec{\phi}(y) >\Big]\, dy.$$

The last term in (3.7) is the so-called non-local term, which is easily evaluated by means of Fourier analysis. We denote by $\hat{\phi}(k)$ the Fourier components of $\vec{\phi}(y)$, i.e.

$$\vec{\phi}(y) \;=\; \sum_{k\in\mathbb{Z}^N} \hat{\phi}(k)e^{i(k.y)}$$

where $k.\hat{\phi}(k) = 0$, because of the constraint $div\,\vec{\phi} = 0$, and $\hat{\phi}(0) = 0$, because of the constraint $\int_Y \vec{\phi}\, dy = 0$ (see (3.2)). By application of Plancherel's formula, we obtain

$$\int_Y \Big[a_2^{-1} |\vec{\phi}(y)|^2 - 2\chi(y)< \vec{\eta},\vec{\phi}(y) >\Big]\, dy \;=\; \mathrm{Re} \sum_{k\in\mathbb{Z}^N} \Big[a_2^{-1} |\hat{\phi}(k)|^2 - 2\overline{\hat{\chi}(k)}< \vec{\eta},\hat{\phi}(k) >\Big]. \tag{3.8}$$

Minimizing frequency by frequency is easy. Frequency 0 is special : it contributes nothing to (3.8) since $\int_Y \vec{\phi}\, dy = 0$. For $k \neq 0$, taking into account the divergence-free constraint $k.\hat{\phi}(k) = 0$, each term in the right hand side of (3.8) is minimum for

$$\hat{\phi}(k) \;=\; \hat{\chi}(k)a_2\left(\vec{\eta} - \frac{(\vec{\eta}.k)}{|k|^2}k\right).$$

Thus, the non-local term is

$$- \sum_{k\neq 0} |\hat{\chi}(k)|^2 a_2 \left(|\vec{\eta}|^2 - \frac{(\vec{\eta}.k)^2}{|k|^2}\right).$$

It is bounded below by

$$- \ \mathrm{Sup}_{k} \left[|\vec{\eta}|^2 - \frac{(\vec{\eta}.k)^2}{|k|^2} \right] a_2 \sum_{k \neq 0} |\hat{\chi}(k)|^2. \tag{3.9}$$

One can check that (3.9) is exactly $-g(\vec{\eta})$, as defined in formula (3.4), since, by Plancherel's formula

$$\sum_{k \neq 0} |\hat{\chi}(k)|^2 = \int_Y |\chi(y) - \theta|^2 = \theta(1-\theta).$$

Combining (3.7) and (3.9) gives the Hashin-Shtrikman variational principle.

We favor the Hashin-Shtrikman variational principle, because it gives a systematic procedure for checking that the lower bound (3.3) is optimal (in the sense of Definition 3.1).

Proposition 3.4.

The Hashin-Shtrikman lower bound (3.3) is optimal, namely, for any flux $\vec{j}$, there exists a composite material, obtained by a single lamination of a_1 and a_2 in a direction orthogonal to $\vec{j}$, for which equality is attained in (3.3).

Proof.

In the course of the proof of Theorem 3.3, we have used only two inequalities : the first one is a consequence of forcing $\vec{\eta}(y)$ to be constant, the second one comes from the maximization over all frequencies k. We are going to prove that they are actually equalities for a carefully chosen lamination of a_1 and a_2. Up to a change of variables, we can always assume that the flux $\vec{j}$ is parallel to one of the axis of the unit cell Y. Consider a single lamination of a_1 and a_2 in a direction $\vec{e}$ parallel to the axis and orthogonal to $\vec{j}$ (the periodic cell is cut in two subdomains separated by an interface orthogonal to $\vec{e}$). It is an easy algebra exercise (left to the reader) to check that $\vec{\phi}(y) = 0$ is indeed the solution of the cell problem (3.2) for this special microstructure. This implies that, in the dual transformation (3.6), the optimal $\vec{\eta}(y)$ is a constant vector parallel to $\vec{j}$. Thus, we don't get an inequality but an equality by forcing $\vec{\eta}(y)$ to be constant in (3.6). Furthermore, for this special microstructure, the Fourier components $\hat{\chi}(k)$ of the characteristic function of the a_2-domain are zero except when k is parallel to $\vec{e}$. Thus the non-local term is exactly

$$- a_2 \left(|\vec{\eta}|^2 - \frac{(\vec{\eta}.\vec{e})^2}{|\vec{e}|^2} \right) \sum_{k \neq 0} |\hat{\chi}(k)|^2.$$

Since $\vec{e}$ is orthogonal to $\vec{j}$ which is parallel to $\vec{\eta}$, the non-local term becomes

$$- a_2 |\vec{\eta}|^2 \theta(1-\theta)$$

which is nothing than $g(\vec{\eta})$. Thus equality is achieved for this microstructure in (3.3).

Of course, in the case of our simple model problem, formula (3.4) for the non-local term can be further simplified, and the bound furnished by the Hashin-Shtrikman variational principle can be explicitly computed.

Theorem 3.5.

The optimal lower bound (3.3) on the dual effective energy is nothing else than the arithmetic mean bound

$$< A^{*-1}\vec{j},\vec{j} > \ \geq \ < \left((1-\theta)a_1 + \theta a_2\right)^{-1}\vec{j},\vec{j} >. \tag{3.10}$$

Proof.

The maximum in (3.4) is obviously attained for a vector $\vec{k}$ orthogonal to $\vec{\eta}$

$$g(\vec{\eta}) \ = \ a_2 |\vec{\eta}|^2.$$

Thus, (3.3) becomes

$$< A^{*-1}\vec{j},\vec{j} > \ \geq \ < a_2^{-1}\vec{j},\vec{j} > + (1-\theta) \operatorname{Max}_{\vec{\eta}} \left[2<\vec{j},\vec{\eta}> - <\left((a_1^{-1} - a_2^{-1})^{-1} + \theta a_2\right)\vec{\eta},\vec{\eta}>\right].$$

By convex duality, the maximum in $\vec{\eta}$ reduces to

$$< \left((a_1^{-1} - a_2^{-1})^{-1} + \theta a_2\right)^{-1}\vec{j},\vec{j} >.$$

An easy computation leads to the result (3.10).

Proof of Proposition 2.8.

By combining Proposition 3.4 and Theorem 3.5, the proof is immediate. Since the bound (3.10) is optimal, the minimum of $< A^{*-1}\vec{j},\vec{j} >$ is precisely the right hand side of (3.10).

Remark 3.6.

We have actually proved more than Proposition 2.8, namely we have exhibited a special class of composite materials which are optimal : the laminated composites. This type of composites play an important role in the theory of optimal bounds on effective properties (see [4], [5], [14], [16]). However, they are not the only class of optimal composites (cf. the concentric spheres, or ellipsoids, construction [17], [24]). These optimal composites give also an insight of the geometry of minimizing sequences for the original problem (1.4). In this particular case, a minimizing sequence is obtained by simply considering an optimal lamination of the two materials, but with a finite length scale going to zero.

4) Another model problem in elasticity.

In this section we introduce another model problem, similar to that of section 1, but much more difficult to analyse since the state equation is now a system of equations from elasticity. However, the strategy for solving this problem is completely parallel to that presented in the previous sections. Just some computations are a little more involved... Consequently, we content ourselves in giving the main results without any proof (for details see [3]). Another difficulty comes from the physical motivation of this problem which is to find an optimal shape rather than an optimal arrangement of two materials (as in section 1). This is modeled by considering one of the two materials as being degenerate (i.e. holes, or void). This leads to serious mathematical difficulties that won't be discussed here. Thus, this section must be regarded as an illustration of the homogenization method for shape optimization in a context where applications are numerous (see e.g. [3], [2], [7], [8]).

Let us explain the physical motivation of this problem. The usual goal in structural optimization is to find the "best" structure which is, at the same time, of minimal weight and of maximum strength. Here, we consider a model problem of this type, in the context of linear elasticity with a single loading configuration. For simplicity, we work in two space dimensions, but most part of the analysis can be carried away in three space dimensions. We begin with a plane bounded domain Ω, occupied by a linearly elastic material with isotropic Hooke's law A, and loaded on its boundary by some known force $\vec{f}$. Admissible designs are obtained by removing a subset $H \subset \Omega$, consisting of one or more holes (the new boundaries created this way are traction-free). The holes H are actually the degenerate limit of a second material whose Hooke's law is going to zero. We recall that a Hooke's law is a fourth-order tensor acting on symmetric matrices (it plays the role of conductivity in this problem). An isotropic Hooke's law A is defined by two positive reals κ and μ (the bulk and shear moduli, respectively), and for any symmetric matrix ξ, it satisfies

$$A\xi = 2\mu\xi + (\kappa - \frac{2\mu}{N})(tr\,\xi)I_2, \tag{4.1}$$

where I_2 is the identity matrix. The state equation is the system of elasticity equations

$$\begin{cases} \sigma = Ae(\vec{u}), \quad e(\vec{u}) = \frac{1}{2}(\nabla\vec{u} + {}^t\nabla\vec{u}) \\ div\ \sigma = 0 \quad in\ \Omega\backslash H \\ \sigma.\vec{n} = \vec{f} \quad on\ \partial\Omega, \ \ \sigma.\vec{n} = 0 \quad on\ \partial H, \end{cases} \tag{4.2}$$

where the unknown $\vec{u}$ is the displacement vector, $e(\vec{u})$ and σ are symmetric matrices (the strain, and the stress, respectively). The compliance is defined as the work done by the load, or equivalently as the primal, or dual, energy

$$c(\Omega\backslash H) = \int_{\partial\Omega} f.u = \int_{\Omega\backslash H} \langle Ae(u), e(u)\rangle = \int_{\Omega\backslash H} \langle A^{-1}\sigma, \sigma\rangle . \tag{4.3}$$

Introducing a positive Lagrange multiplier λ, the goal is to minimize, over admissible designs $\Omega\backslash H$, the weighted sum of the compliance and the weight, i.e.

$$Min_H \left[c(\Omega\backslash H) + \lambda |\Omega\backslash H| \right] . \tag{4.4}$$

Problem (4.4) is the equivalent of the conductivity problem (1.4) for elasticity. We now give the equivalent of Proposition 1.1.

Proposition 4.1.

The minimization problem (4.4) is equivalent to the minimization of

$$\int_\Omega G(\tau)\, dx \ , \quad with \ \ G(\tau) = \begin{cases} 0 & \text{if } \tau = 0 \\ <A^{-1}\tau,\tau> + \lambda & \text{if } not \end{cases} \tag{4.5}$$

among admissible stresses τ, *i.e. satisfying*

$$\begin{cases} div \ \tau \ = \ 0 & in \ \Omega \\ \tau.\vec{n} \ = \ \vec{f} & on \ \partial\Omega. \end{cases} \tag{4.6}$$

Namely, to any minimizer of (4.5) corresponds a minimizer of (4.4), and reciprocally.

Following is the equivalent of Theorem 1.2.

Theorem 4.2.

The relaxed formulation of (4.5) is the minimization, under the same constraint (4.6), of the relaxed energy

$$\int_\Omega \tilde{G}(\tau)\, dx \tag{4.7}$$

with

$$\tilde{G}(\tau) \ = \ \begin{cases} < A^{-1}\tau,\tau > \ + \ \lambda & \text{if } \rho \geq 1 \\ < A^{-1}\tau,\tau > \ + \ \lambda\, \rho(2-\rho) & \text{if } \rho < 1 \end{cases} \tag{4.8}$$

and

$$\rho \ = \ \left(\frac{\kappa+\mu}{4\kappa\mu} \right)^{1/2} \lambda^{-1/2} \left(|\tau_1| \ + |\tau_2| \right)$$

where τ_1 *and* τ_2 *are the two eigenvalues of* τ.

The relaxed formulation (4.7) satisfies the condition (1)-(3) defined in section 1. Remark that the integrand (4.8) is not the convexification of the integrand of (4.5), but rather its quasi-convexification (see [3]). As for Theorem 1.2, the proof of Theorem 4.2 is based on G-convergence and on an optimal lower bound on the dual effective energy. Recall that, in the case of elasticity, the G-closure is unknown, which implies that this optimal lower bound is crucial. To enlighten the differences between the conductivity and elasticity cases, we state the equivalent of Theorem 3.3 and 3.5.

Theorem 4.3.

Let A^* *be an effective Hooke's law obtained by homogenization of the material* A *with holes, in proportions* θ *and* $(1-\theta)$ *respectively. The Hashin-Shtrikman variational principle is the following lower bound*

$$< A^{*-1}\tau,\tau > \ \geq \ < A^{-1}\tau,\tau > + (1-\theta)\ \text{Max}_{\ \varepsilon} \left[2<\tau,\varepsilon> - \theta g(\varepsilon) \right] \tag{4.9}$$

where the maximum is taken over all constant symmetric matrices ε, *and* $g(\varepsilon)$ *is the so-called non-local term defined by*

$$g(\varepsilon) = < A\varepsilon,\varepsilon > - h(A\varepsilon) \tag{4.10}$$

with

$$h(\xi) = \mathrm{Sup}_{|\vec{k}|=1} \left[\frac{1}{\mu}[\,|\xi\vec{k}|^2 - <\xi\vec{k}.\vec{k}>^2] + \frac{1}{2\mu+\kappa-2\mu/N} <\xi\vec{k}.\vec{k}>^2 \right]$$

where the supremum is taken over all unit vectors $\vec{k}$. *The bound (4.9) is optimal and is attained for so-called rank-*N *sequential laminates.*

Furthermore, the right hand side of (4.9) can be computed explicitly in 2-D :

$$< A^{*-1}\tau,\tau > \;\geq\; < A^{-1}\tau,\tau > + \frac{(\kappa+\mu)(1-\theta)}{4\kappa\mu\theta}\left(|\tau_1| + |\tau_2|\right)^2. \tag{4.9}$$

For a proof of Theorem 4.3, we refer to [3].

References :

[1] E. Acerbi, N. Fusco, *Semicontinuity problems in the calculus of variations,* Arch. Rat. Mech. Anal. 86, pp.125-145 (1984).

[2] G. Allaire, *Structural optimization using optimal microstructures,* In "MECAMAT 93 International Seminar on Micromechanics of Materials", Collection de la Direction des Etudes et Recherches d'Electricité de France, Eyrolles, Paris (1993).

[3] G. Allaire, R.V. Kohn, *Optimal Design for Minimum Weight and Compliance in Plane Stress Using Extremal Microstructures,* to appear in Europ. J. Mech. (1993).

[4] G. Allaire, R.V. Kohn, *Optimal Bounds on the Effective Behavior of a Mixture of Two Well-Ordered Elastic Materials,* to appear in Q. Appl. Math..

[5] M. Avellaneda, *Optimal Bounds and Microgeometries for Elastic Two-Phase Composites,* SIAM J. Appl. Math. 47, pp.1216-1228 (1987).

[6] N.V. Banichuk, *Problems and methods of optimal structural design,* Plenum Press, New York (1983).

[7] M. Bendsoe, N. Kikuchi, *Generating Optimal Topologies in Structural Design Using a Homogenization Method,* Comp. Meth. Appl. Mech. Engrg. 71, pp.197-224 (1988).

[8] M.P. Bendsoe and C.A. Soares, *Topology design of structures,* Nato ASI Series E, vol. 227, Kluwer, Dordrecht (1993).

[9] A. Bensoussan, J.L. Lions, G. Papanicolaou, *Asymptotic analysis for periodic structures,* North-Holland (1978).

[10] G. Buttazzo, G. Dal Maso, *Integral representation and relaxation of local functionals,* Nonlinear Anal. Th. Meth. Appl. 9, pp. 515-532 (1985).

[11] B. Dacorogna, *Weak continuity and weak lower semi-continuity of nonlinear functionals,* Lecture Notes in Math. 922, Springer Verlag, New York (1982).

[12] A. Defranceschi, *Lecture notes,* School on Homogenization, ICTP, Trieste, September 6-17, (1993).

[13] I. Ekeland, R. Temam, *Convex analysis and variational problems,* North-Holland, Amsterdam (1976).

[14] G. Francfort, F. Murat, *Homogenization and Optimal Bounds in Linear Elasticity,* Arch. Rat. Mech. Anal. 94, pp.307-334 (1986).

[15] L. Gibiansky, *Lecture notes,* School on Homogenization, ICTP, Trieste, September 6-17, (1993).

[16] L. Gibiansky, A. Cherkaev, *Microstructures of composites of extremal rigidity and exact estimates of provided energy density,* Preprint 1115 (in Russian), Ioffe Physicotechnical Institute, Leningrad (1987). (English translation to appear in "Topics in the mathematical modeling of composite materials", R.V. Kohn ed., series "Progress in Nonlinear Differential Equations and their Applications", Birkhauser, Boston.)

[17] Z. Hashin, S. Shtrikman, *A Variational Approach to the Theory of the Elastic Behavior of Multiphase Materials,* J. Mech. Phys. Solids 11, pp.127-140 (1963).

[18] Kohn R.V. and Strang G., *Optimal Design and Relaxation of Variational Problems I-III,* Comm. Pure Appl. Math. 39, pp.113-137, 139-182, 353-377 (1986).

[19] K. Lurie, A. Cherkaev, A. Fedorov, *Regularization of Optimal Design Problems for Bars and Plates I,II,* J. Optim. Th. Appl. 37, pp.499-521, 523-543 (1982).

[20] F. Murat, L. Tartar, *Calcul des Variations et Homogénéisation,* in Les Méthodes de l'Homogénéisation: Théorie et Applications en Physique, Eyrolles, pp.319-369 (1985).

[21] O. Pironneau, *Optimal shape design for elliptic systems,* Springer Verlag, New York (1984).

[22] G. Rozvany, *Structural design via optimality criteria,* Kluwer, Dordrecht (1989).

[23] J. Sokolowski, J.-P. Zolesio, *Introduction to shape optimization. Shape sensitivity analysis,* Springer Series in Computational Mathematics 16, Springer Verlag, New York (1992).

[24] L. Tartar, *Estimations Fines des Coéfficients Homogénéisés,* Ennio de Giorgi colloquium, P. Krée ed., Pitman Research Notes in Math. 125, pp.168-187 (1985).

[25] L.C. Young, *Lectures on the calculus of variations and optimal control theory,* W.B. Saunders, Philadelphia (1969).