The following lecture notes have been written for a European Summer School on Optimal Shape Design, which was held in Tróia, Portugal, from June 1 to June 6, 1998. It was organized by Centro Internacional de Matemática (C.I.M.), Coimbra, Portugal, and Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, Italy. It was financially supported by European Union through the T.M.R. Programme and partially funded by C.I.M. and C.I.M.E..

The organizing committee consisted of Arrigo CELLINA (C.I.M.E and Universitá di Milano, Italy), and António ORNELAS (C.I.M. and Universidade de Évora, Portugal). They had set the objectives of the course as "To show to young mathematicians how to deal with a class of problems coming from engineering, and how to solve them using theoretical and numerical algorithms. To present to young engineers the best theoretical methods that mathematicians have recently developped, and to show how in fact these methods can be used to solve difficult engineering problems."

The other lecturers were
Bernd KAWOHL, Köln, Germany: Some Nonconvex Shape Optimization Problems
Olivier PIRONNEAU, Paris VI, France: Mesh Adaptation for Optimal Shape Design
Piero VILLAGIO, Pisa, Italy: Explicit Solution in Elastic Optimization
Jean Paul ZOLESIO, INRIA, France: Optimal Shape Design: Theory, Models, Numerical Algorithms

The proceedings of the summer school will be published by Springer.

An Introduction to the Homogenization Method in Optimal Design

Luc TARTAR

Department of Mathematical Sciences
CARNEGIE-MELLON University
Pittsburgh, PA 15213-3890, U.S.A.

Table of Contents

0. Introduction

This course describes Optimization/Optimal Design problems which lead to Homogenization questions, together with the method to treat them that François MURAT and I had developed, mostly in the 70s. I adopt a chronological point of view, which I find best for describing how some new techniques (which are often misattributed nowadays) were introduced for overcoming some difficulties that we had encountered, or for generalizing a particular result that we had found useful in our search.

There are many ways to practice research in Mathematics, and I hope that the description of the way we thought in front of new questions could help some to experience for themselves the extraordinary feeling of discovery; it must be said, however, that one rarely finds oneself in exactly the same situations that others

had experienced before. What one finds may also have been known to others, a fact that Ennio DE GIORGI had resumed in "Chi cerca trova, chi ricerca ritrova"[1].

In a few occasions it had probably been useful that I did not know the approach that others had followed before, as the one that I created appeared to be different and to offer new possibilities: the knowledge of an old method often makes it difficult to invent a new one, and this is an important obstacle in research.

Although the difference between exploration and exploitation is quite obvious for any oil engineer, many researchers in Mathematics spend their life exploiting methods invented by others, mimicking what they have read or heard. There is nothing really wrong about that behaviour for those who understand their role in the scientific community as that of soldiers of an army working for the benefit of the whole community. However, it is a potentially destructive behaviour that some pretend to have invented the ideas that they have read or that they have been taught directly by more creative people, even if they simply misattribute them to some of their friends instead, because too often they have not completely understood the whole potential of the ideas that they use, and they may transmit twisted informations with the main effect of misleading part of the army, and playing therefore the same role as traitors.

Having been raised in a religious background, I do not attribute my mathematical ability to my own efforts and therefore I consider it a duty to put the gift that I was given to the service of others. I have noticed that religious teachers from the past seem to have liked using parables, stories so simple that their students asked about their meaning, and oral tradition then transmitted us the initial story, the questions of the students, and one "example" given by the master; because it was so easy to remember, the teaching had therefore been transmitted intact by people who did not even understand it, unaware of the innumerable applications that the little story contained. Some mathematicians behave in this way too, writing general theorems but giving only one or two examples, or they may simply teach a general theory by explaining one example in detail, thinking that every trained mathematician will automatically see what the general idea is. I have done that "mistake" often, and after inventing a method applicable to all variational problems, I had been quite upset when I had found written that I had only solved the case of a diffusion equation, wondering how someone writing such a stupid statement could consider himself a mathematician, and why his coauthors had not jumped out of their seat at such an idiotic remark that could well be attributed to them too. Jealousy, the observation that mathematical ability is too sparsely distributed, or the saying "au pays des aveugles les borgnes sont rois",[2] come to my mind for explaining the behaviour of those who try to mislead students out of the right path, but it may well be the result of their complete lack of moral education.

For some mysterious reason, François MURAT and I were the first who put together the various pieces of the theory that I will describe, and I will explain what other people had done according to my information. I had taught similar lectures in 1983, before some people embarked in a systematic campaign of misattribution of ideas and results: everyone with a minimal mathematical ability may quickly identify the names of those who have indulged in intentional misleading, but this does not mean that they have not themselves had some genuine idea, in which case I will mention their name for that.

1. A counter-example of François MURAT

In 1970, François MURAT worked on an academic problem of optimization that had been proposed by Jacques-Louis LIONS [Li1], and he found that it had no solution [Mu1]. His result was unexpected,[3] and as we were sharing an office in Jussieu, we started a long and fruitful collaboration, first discussing some

---

[1] From the sentence "Ask and it will be given to you; seek and you will find; knock and the door will be opened to you." in the gospels (Matthew 7:7, Luke 11:9), the middle part has given rise to the French saying "Qui cherche trouve", equivalent to the Italian "Chi cerca trova". The play on the prefix, "re" in French, "ri" in Italian, does not work as well in English (one could replace seek and find by search and discover in order to use research and rediscover).

[2] In the land of the blind, a one-eyed man is king.

[3] We had not heard about the work that Laurence C. YOUNG had done in the 40s [Yo]. I had first met him in 1971 in Madison, but I only learned many years after that he was the inventor of the objects which I was still using under the name of parametrized measures in my HERIOT-WATT lectures in 1978 [Ta8]; it was Ronald DIPERNA who then insisted that they should be called YOUNG measures.

generalizations of his first idea [Mu2], and then embarking on the exploration that led us to (re)discover[4] the general theory of "Homogenization"[5] [Ta1], [Ta2], [Ta3], [Mu3]. The initial problem that François MURAT considered was to minimize the cost function

$$J(a) = \int_0^1 |y(x) - 1 - x^2|^2 \, dx, \tag{1.1}$$

when the state y solves the equation of state

$$-\frac{d}{dx}\left(a\frac{dy}{dx}\right) + a\, y = 0 \text{ in } (0,1), \ y(0) = 1, \ y(1) = 2; \ y \in H^1(0,1), \tag{1.2}$$

and the control a lies in the following admissible control set

$$\mathcal{A} = \{a | a \in L^\infty(0,1), \alpha \le a \le \beta \text{ a.e. in } (0,1)\}, \tag{1.3}$$

and as he tried to apply the direct method of the Calculus of Variations, he noticed that for a sequence $a^n \in \mathcal{A}$ such that

$$a^n \rightharpoonup a_+ \text{ and } \frac{1}{a^n} \rightharpoonup \frac{1}{a_-} \text{ in } L^\infty(0,1) \text{ weak } \star, \tag{1.4}$$

then the corresponding sequence of solutions y_n of (1.2) converges in $H^1(0,1)$ weak to the solution y_∞ of

$$-\frac{d}{dx}\left(a_-\frac{dy_\infty}{dx}\right) + a_+\, y_\infty = 0 \text{ in } (0,1), \ y_\infty(0) = 1, \ y_\infty(1) = 2; \ y_\infty \in H^1(0,1), \tag{1.5}$$

with

$$a^n \frac{dy_n}{dx} \to a_- \frac{dy_\infty}{dx} \text{ in } L^2(0,1) \text{ strong}, \tag{1.6}$$

and

$$J(a^n) \to \widetilde{J}(a_-, a_+) = \int_0^1 |y_\infty(x) - 1 - x^2|^2 \, dx. \tag{1.7}$$

Indeed, y_n is bounded in $H^1(0,1)$ and $v_n = a^n \frac{dy_n}{dx}$ is bounded in $L^2(0,1)$, but as its derivative is $a^n\, y_n$ which is bounded in $L^2(0,1)$, v_n is actually bounded in $H^1(0,1)$. Therefore one can extract a subsequence such that y_m converges in $H^1(0,1)$ weak and in $L^2(0,1)$ strong to y_∞, and v_m converges in $L^2(0,1)$ strong to v_∞. Then $a^m\, y_m$ and $\frac{dy_m}{dx} = \frac{1}{a^m}v_m$ converge in $L^2(0,1)$ weak, respectively to $a_+ y_\infty$ and to $\frac{dy_\infty}{dx} = \frac{1}{a_-}v_\infty$, showing (1.6), and therefore (1.5); the fact that y_∞ is uniquely determined by (1.5) shows that the extraction of a subsequence is not necessary.

For $\alpha = \frac{\sqrt{2}-1}{\sqrt{2}}, \beta = \frac{\sqrt{2}+1}{\sqrt{2}}$, François MURAT used the particular sequence

$$a^n(x) = \begin{cases} 1 - \sqrt{\frac{1}{2} - \frac{x^2}{6}} & \text{when } x \in \left(\frac{2k}{2n}, \frac{2k+1}{2n}\right), k = 0, \ldots, n-1 \\ 1 + \sqrt{\frac{1}{2} - \frac{x^2}{6}} & \text{when } x \in \left(\frac{2k+1}{2n}, \frac{2k+2}{2n}\right), k = 0, \ldots, n-1, \end{cases} \tag{1.8}$$

---

[4] We were not aware of the earlier work on G-convergence of Sergio SPAGNOLO [Sp1], [Sp2], and his work with Antonio MARINO [Ma&Sp], or the work of Tullio ZOLEZZI [Zo]. It was only after having developed our own approach, which François MURAT named H-convergence in [Mu3], that we became aware of these works and that of Ennio DE GIORGI and Sergio SPAGNOLO [DG&Sp].

[5] As in the title of my PECCOT lectures in 1977, I have adopted the term Homogenization, first introduced by Ivo BABUŠKA [Ba], for describing our general approach of H-convergence; of course, no constraints of periodicity are considered. Some authors automatically associate the term with periodic structures, probably because they have not understood the general framework introduced by Sergio SPAGNOLO or by François MURAT and me!

corresponding to

$$a_- = \frac{1}{2} + \frac{x^2}{6}; \ a_+ = 1; \ y_\infty = 1 + x^2 \text{ in } (0,1), \tag{1.9}$$

showing that $\inf_{a \in \mathcal{A}} J(a) = 0$. He checked then that it was not possible to have $y = 1 + x^2$ in (1.2) for some $a \in \mathcal{A}$, by considering (1.2) as a differential equation for a, and noticing that all nonzero solutions are unbounded, as they are $a = \frac{C}{\sqrt{x}} exp\left(\frac{x^2}{4}\right)$.

We were naturally led to characterize all the possible pairs (a_-, a_+) which could appear in (1.4) and we found

$$\alpha \leq a_-(x) \leq a_+(x) \leq \frac{a_-(x)(\alpha + \beta) - \alpha \beta}{a_-(x)} \leq \beta \text{ a.e. } x \in (0,1), \tag{1.10}$$

or equivalently

$$\frac{1}{a_+(x)} \leq \frac{1}{a_-(x)} \leq \frac{\alpha + \beta - a_+(x)}{\alpha \beta} \text{ a.e. } x \in (0,1). \tag{1.11}$$

Our proof of (1.11) easily extended to the following more general situation,[6] and the characterization (1.11) corresponds to using the following Lemma, with a and $\frac{1}{a}$ being the components of U, K being the piece of hyperbola $U_1 U_2 = 1$ with $\alpha \leq U_1 \leq \beta$.

Lemma 1: Let $U^{(n)}$ be a sequence of measurable functions from an open set $\Omega \subset R^N$ into R^p satisfying $U^{(n)} \rightharpoonup U^{(\infty)}$ in $L^\infty(\Omega; R^p)$ weak $\star$ and $U^{(n)}(x) \in K$ a.e. $x \in \Omega$. For a bounded set K,[7] the characterization of all the possible limits $U^{(\infty)}$ is $U^{(\infty)}(x) \in \overline{conv(K)}$, the closed convex hull of K, a.e. $x \in \Omega$.
Proof: The closed convex hull of K is the intersection of all the closed half spaces which contain K, and a closed half space H_+ has an equation $\{\lambda | \lambda \in R^p, L(\lambda) \geq 0\}$ for some nonconstant affine function L, and if H_+ contains K one has $L(U^n) \geq 0$ a.e. $x \in \Omega$ and therefore $L(U^\infty) \geq 0$ a.e. $x \in \Omega$, i.e. $U^\infty(x) \in H_+$ a.e. $x \in \Omega$; the conclusion follows, if one is careful to write the closed convex hull as a countable intersection of closed half spaces containing K.
Let $V \in L^\infty(\Omega; R^p)$ be such that $V(x) \in clconv(K)$ a.e. $x \in \Omega$. For each m, one can cut R^p into small cubes of size $\frac{1}{m}$ and choose a point of $conv(K)$, the convex hull of K, in each cube intersecting $clconv(K)$ and that helps creating a function $W^{(m)} \in L^\infty(\Omega; R^p)$ such that $|V - W^{(m)}| \leq \frac{1}{m}$ a.e. $x \in \Omega$ and $W^{(m)}$ takes only a finite number of values in $conv(K)$. On a measurable subset ω of Ω where $W^{(m)}$ is constant, we want to construct a sequence of functions converging in $L^\infty(\omega; R^p)$ weak $\star$ to $W^{(m)}$ and taking their values in K, and putting these functions together will create a sequence converging in $L^\infty(\Omega; R^p)$ weak $\star$ to $W^{(m)}$, and then, as the weak $\star$ topology of $L^\infty(\Omega; R^p)$ is metrizable on bounded sets, this will ensure that one can approach V in that topology.
Let $W^{(m)} = \lambda \in conv(K)$ on ω, so that $\lambda = \sum_i \theta^i k^i$, with $k^i \in K$ and the sum is finite (with all $\theta^i \geq 0$, $\sum_i \theta^i = 1$). We cut now ω into measurable pieces of diameter at most $\frac{1}{n}$, then partition each of such pieces E into measurable subsets E^i with $meas(E^i) = \theta^i meas(E)$ and define the function Z^n to be equal to k^i on each such E^i. The claim is then that as n tends to ∞, the sequence Z^n converges in $L^\infty(\omega; R^p)$ weak $\star$ to λ; as Z^n only takes a finite number of values in K it is bounded, and it is enough to check that for every continuous function φ with compact support $\int_\omega \varphi Z^n dx \to \int_\omega \varphi \lambda dx$. As φ is uniformly continuous $|\varphi(x) - \varphi(y)| \leq \varepsilon$ when $|x - y| \leq \frac{1}{n}$, so if $e \in E$ one has $|\varphi(e) \int_E Z^n dx - \int_E \varphi Z^n dx| \leq \varepsilon M meas(E)$ and $|\varphi(e) \int_E \lambda dx -$

---

[6] As we had learned that weak convergence is not adapted to nonlinear problems, we were surprised to have found such a simple characterization, and I called our common thesis advisor, Jacques-Louis LIONS, to ask him if this was not known already. He suggested that I ask Ivar EKELAND, who told me that it had been implicitly used in some work of CASTAING and was related to a classical result of LYAPUNOV, valid for a set endowed with a nonnegative measure without atoms; indeed our proof extended easily to such a general case. I only learned in 1975 from Zvi ARTSTEIN about his simple proof of LYAPUNOV's result [Ar].

[7] If K is unbounded, one denotes K_M the elements of K of norm $\leq M$, and the characterization is that there exists M with $U^\infty(x) \in \overline{conv(K_M)}$ a.e. $x \in \Omega$. The functions $U^{(\infty)}$ such that $U^{(\infty)}(x) \in \overline{conv(K)}$ correspond to the weak $\star$ closure in $L^\infty(\Omega; R^p)$ of the functions taking (a.e.) their values in K, and one can easily find cases where these two sets of functions are different, showing that the weak $\star$ topology is not metrizable on these unbounded sets.

$\int_E \varphi \, \lambda \, dx| \leq \varepsilon \, M \, meas(E)$ where M is a bound of the norm of R^p on K, but as $\int_E Z^n \, dx = \sum_i \int_{E_i} k^i \, dx = \sum_i \theta^i \, meas(E)k^i = meas(E)\lambda = \int_E \lambda \, dx$, one deduces $|\int_\omega \varphi \, Z^n \, dx - \int_\omega \varphi \, \lambda \, dx| \leq 2\varepsilon \, M \, meas(\omega)$.

We left aside the construction of the E^i from E, which is easy for sets in R^N. $\widetilde{L}$ being a nonzero linear function on R^N, the measure of $E \bigcap \{x|x \in R^N, \widetilde{L}(x) \geq t\}$ is a continuous function of t which grows from 0 to $meas(E)$ and one obtains the desired partition of E by cutting E by suitable hyperplanes $\widetilde{L}^{-1}(t_i)$. The construction can be generalized when Ω is any set equipped with a measure without atoms, as stated by a classical result of LYAPUNOV (in 1975, I learned from Zvi ARTSTEIN a method of his giving a simple proof of that result as well as bang-bang results in control theory [Ar]).∎

2. The independent discoveries of others

Jean-Louis ARMAND had studied at Ecole Polytechnique in Paris a year ahead of me, but we only met almost fifteen years after graduating, not so much because we were then part time lecturers at Ecole Polytechnique in Palaiseau (he in Mechanics, I in Mathematics), but because he had learned about my work by going to visit Konstantin LUR'IE, in Leningrad. Jean-Louis ARMAND had been computing some Optimal Design problems and he had been puzzled by the fact that in the meetings that he attended, various engineers were showing results which were quite different, although they were supposed to solve the same problem; however, he seemed to have been alone in thinking that this was the sign of a serious theoretical difficulty. He had tried to find explanations in the litterature, and he had discovered an article by K. LUR'IE which seemed relevant; he had then traveled to visit him in Leningrad and, having been told about my work there, he had contacted me after his return to Paris and he had mentioned to me what K. LUR'IE had done and what he had told him.

K. LUR'IE had extended some ideas of PONTRYAGUIN to partial differential equations, and he had been able to obtain better necessary conditions of optimality than those given by a classical method.[8] However, he had been quite puzzled when he had discovered a situation with no function satisfying his necessary conditions [Lu]. I do not know if K. LUR'IE had already obtained the right intuition about what was going on before finding my article [Ta2], but he had mentioned to Jean-Louis ARMAND that it was in [Ta2] that he had found the missing ideas that he needed. I had not given the detail of my work with François MURAT in [Ta2], but I had mentioned the work on G-convergence of Sergio SPAGNOLO [Sp1], [Sp2], as well as the work of Henri SANCHEZ-PALENCIA [S-P1], [S-P2] , which had helped us understand that our work had something to do with the question of effective properties of mixtures; K. LUR'IE had then coined the term G-closure for describing the set of all possible effective tensors of admissible mixtures. I will describe later the intuitive ideas behind the necessary conditions of optimality obtained by K. LUR'IE, but it is important to understand that the reason why we were able to go further was that we were rediscovering and extending the ideas of Laurence C. YOUNG, while K. LUR'IE was following the ideas of PONTRYAGUIN and he could hardly have realized that he had taken the wrong track. As pointed out by Laurence C. YOUNG [Yo], if one has obtained some necessary conditions of optimality for an optimization problem, and one finds that only one function satisfies them, one cannot even assert that it is the solution of the problem, unless of course one has already proved that there exists at least one solution of the problem. PERRON's paradox seems too naive an example,[9] but the point of view of PONTRYAGUIN becomes indeed useless if the problem at hand has no solution. On the contrary, the point of view of Laurence C. YOUNG is adapted to that kind of situation, and it creates a relaxed problem which answers two questions: first it explains how the minimizing sequences may have their limit outside the initial space in the case of nonexistence of solutions, then it does give the necessary conditions obtained by following the point of view of PONTRYAGUIN, as they are just part of the necessary conditions of optimality for the relaxed problem. I will describe these questions on an elementary

---

[8] It is an idea going back to HADAMARD to push an interface along its normal for computing derivatives of functionals, in order to obtain necessary conditions of optimality for example. Although many applications of that idea may be formal, François MURAT and Jacques SIMON have spent some time giving the method a rigourous framework [Mu&Si]. However, the classical approach only gives conditions that must be satisfied along the interface, while K. LUR'IE's approach as well as ours give conditions which must be satisfied everywhere.

[9] The paradox quoted in [Yo] is as follows: let n be the largest integer; a necessary condition of optimality is $n \geq n^2$, which leaves only two candidates, 0 or 1; therefore the largest integer must be 1!

model.

Once we had become aware of the work of Sergio SPAGNOLO, we did find there some ideas which we had not thought about, and we checked that we could handle them with our methods, but for the question of Optimal Design which was our initial motivation we needed more precise results. For example, Antonio MARINO and Sergio SPAGNOLO had shown that in order to obtain all the materials with an effective tensor equal to a general symmetric tensor having its eigenvalues between α and β, it was sufficient to mix isotropic materials with tensor γI with $\gamma \in [\alpha', \beta']$ for some $\alpha' > 0$ and $\beta' < \infty$ [Ma&Sp], but in order to compute necessary conditions of optimality for questions of Optimal Design we needed a more precise characterization, and we were able to obtain one in dimension 2. As we will see the question of optimal bounds has two sides, one where one must prove inequalities that must be satisfied by every mixture using given proportions, and we had found a method for doing that, and one where one must build particular mixtures and compute their effective tensors, and we used repeated layering for that, as Antonio MARINO and Sergio SPAGNOLO had done, but they had not addressed the first question.

I am not sure if we had found reference to the work of Sergio SPAGNOLO before reading some work of Tullio ZOLEZZI, and one of his articles had puzzled us for a while, as we thought that one of his theorems contradicted some of ours [Zo]. The puzzling theorem stated that if a sequence a^n converges weakly to a_+ in $L^\infty(\Omega)$ then the corresponding sequence of solutions u_n converges weakly to the solution associated to a_+. François MURAT thought that some nuance in Italian might have tricked us in mistranslating what was meant, but as we were pondering if "debolmente" could mean anything else than weakly, it suddenly appeared that our mistake had been to read correctly weakly and to interpret it incorrectly as weakly $\star$, as indeed it was the first time that we had seen any use of the weak topology of $L^\infty(\Omega)$ in a concrete situation; we understood then why there was a reference to an article of Alexandre GROTHENDIECK, who had shown that convergence in $L^\infty(\Omega)$ weak implies strong convergence in $L^p_{loc}(\Omega)$ for every finite p.

In the early 70s, Jean CÉA and his team in Nice had been performing some numerical computations for similar problems of Optimal Design than the ones which I had been studying with François MURAT. We had been aware of some work by Denise CHENAIS [Ch], which means that if one imposes some kind of regularity condition on an interface between two materials then the set of corresponding characteristic functions belongs to a compact subset of $L^p(\Omega)$ for $p < \infty$, and therefore a classical optimal solution exists; our work had suggested that if one does not impose such a condition there may not exist any classical solution, in which case one has to use the generalized solutions that we had been studying, corresponding to mixtures.

In 1974, after my talk containing the necessary conditions of optimality described in [Ta2], I had not been able to convince Jean CÉA that the apparition of mixtures was a real possibility that one had to consider, and he might have been mistaken because of a result which he had obtained a few years before with K. MALANOWSKI [Cé&Ma], corresponding to (4.1)/(4.3) with $g(x, u, a) = f(x)u$, for which a special simplification occurs and a classical solution exists. For a given triangulation, one of the numerical methods that Jean CÉA had tested consists in choosing each triangle to be entirely made of only one material, and I could not convince him that if one refines triangulations enough one may start to see oscillations and that our analysis is important for that reason. Had the computers been more powerful in those days, he might indeed have discovered numerical oscillations in refining his triangulations, but the cost would have been prohibitive at the time and only coarse triangulations were used. The classical way for obtaining necessary conditions of optimality by pushing the interface in the direction of its normal, an idea of J. HADAMARD that is often only used at a formal level but which François MURAT and Jacques SIMON have put into a rigourous framework [Mu&Si], gives conditions that must be satisfied on the interface, while I was obtaining necessary conditions which are valid everywhere; in his method J. CÉA could switch any triangle from one material to another, and therefore there was no real interface in his approach and this led him to some kind of discretized necessary condition valid everywhere, and he might have been mistaken by what he may have seen as a similarity between our results.

After some discussions with Guy CHAVENT, who was studying the related problem of identifying the local permeability of an oil field from measurements at various points, I had proposed a numerical approach for solving numerically the type of optimization problem that we had been studying, but the numerical method that I had proposed, and that one of his students had implemented, did not work well at all. As we knew that the optimal mixture that we were looking for could be obtained locally as a layered medium, I had chosen to parametrize the various possible mixtures with a proportion $\theta \in [0, 1]$ and with an angle describing

the orientation of layers (as I was considering a 2-dimensional problem), but that method appeared to be quite unstable because when θ is 0 or 1, i.e. the material is isotropic, the orientation of the layers is not well defined. I did not try another numerical method, but I had learned that even when the solution of an optimization problem is on the boundary of a set, it might not be a good idea to move only along the boundary of this set in order to find the solution, and a better approach could be to cut through the set in order to arrive more quickly at the interesting points on the boundary.

In June 1980 in New York, Robert V. KOHN had told me that he had been to a meeting organized by Jean CÉA and E.J. HAUG, and as I knew that the approach of my work with François MURAT had probably not been mentioned at this meeting, I taught him our ideas about how Homogenization problems appear in Optimal Design problems.

3. An elementary model problem

This model problem was mentioned to me by Ivar EKELAND, and I believe that his answer involved huge abstract compactified spaces, while my solution is entirely based on our Lemma 1 (which corresponds to a small compactification).

We want to minimize the cost function

$$J(u) = \int_0^T \left(|y|^2 - |u|^2 \right) dt, \tag{3.1}$$

where the control u belongs to

$$U_{ad} = \{u \,|\, u \in L^\infty(0, T), \ -1 \le u(t) \le 1, \ \text{a.e. } t \in (0, T)\}, \tag{3.2}$$

and the state y is defined by the equation of state

$$\frac{dy}{dt} = u \ \text{a.e. on } (0, T); \ y(0) = 0. \tag{3.3}$$

We can consider here classical necessary conditions of optimality because U_{ad} is convex; the map $u \mapsto y$ is affine continuous and therefore the map $u \mapsto J$ is quadratic continuous, and therefore FRÉCHET differentiable, but the following arguments are valid in cases where only GÂTEAUX differentiability holds. Let $u_* \in U_{ad}$, corresponding to a state y_*, and let $\delta u \in L^\infty(0, T)$ be an admissible direction at u, i.e. $u = u_* + \varepsilon \, \delta u \in U_{ad}$ for $\varepsilon > 0$ small. Then $y = y_* + \varepsilon \, \delta y$ and $J(u) = J(u*) + \varepsilon \, \delta J + o(\varepsilon)$, where

$$\frac{d(\delta y)}{dt} = \delta u; \ \delta y(0) = 0, \tag{3.4}$$

and

$$\delta J = 2 \int_0^T \left(y_* \, \delta y - u_* \, \delta u \right) dt \tag{3.5}$$

and the classical necessary conditions of optimality consist in writing that $\delta J \ge 0$ for all admissible δu. In order to eliminate δy so that δJ is expressed only in terms of δu, one introduces the adjoint state[10] p_* by

$$-\frac{dp_*}{dt} = y_*; \ p_*(T) = 0, \tag{3.6}$$

and a simple integration by parts gives

$$\int_0^T y_* \, \delta y \, dt = -\int_0^T \frac{dp_*}{dt} \, \delta y \, dt = \int_0^T p_* \, \frac{d\delta y}{dt} \, dt = \int_0^T p_* \, \delta u \, dt, \tag{3.7}$$

---

[10] I do not know who introduced that notion. It does play an important role in PONTRYAGUIN's approach, and he may have introduced it.

and therefore

$$\delta J = 2 \int_0^T (p_* - u_*)\delta u \, dt. \tag{3.8}$$

The admissibility of δu means that $\delta u \geq 0$ where $u_* = -1$, $\delta u \leq 0$ where $u_* = +1$, and δu arbitrary where $-1 < u_* < 1$ (one first works on the set of points where $-1 + \eta \leq u \leq 1 - \eta$ for $\eta > 0$ and then one takes the union of these sets for all $\eta > 0$), and therefore one immediately deduces the classical necessary conditions of optimality

$$\begin{cases} u_* = -1 \text{ implies } p_* - u_* \geq 0, \text{ i.e. } p_* \geq -1 \\ -1 < u_* < +1 \text{ implies } p_* - u_* = 0, \text{ i.e. } p_* = u_* \\ u_* = +1 \text{ implies } p_* - u_* \leq 0, \text{ i.e. } p_* \leq +1, \end{cases} \tag{3.9}$$

which can be read as giving u_* as the following multivalued function of p_*

$$\begin{cases} p_* < -1 \text{ implies } u_* = +1 \\ -1 \leq p_* \leq +1 \text{ implies } u_* \in \{-1, p_*, +1\} \\ p_* > +1 \text{ implies } u_* = -1. \end{cases} \tag{3.10}$$

One can notice that the system of these classical necessary conditions of optimality, i.e. (3.3), (3.6) and (3.9)/(3.10), has at least the solution $u_* = 0$ on $(0, T)$, corresponding to $y_* = p_* = 0$ on $(0, T)$.

The point of view of PONTRYAGUIN for obtaining necessary conditions of optimality consists in comparing u_* to another control $w \in U_{ad}$ by noticing that any control which jumps from u_* to w is admissible. In the language of Functional Analysis it means

$$u = (1 - \chi)u_* + \chi w \in U_{ad} \text{ for every characteristic function } \chi \text{ of a measurable subset of } (0, T), \tag{3.11}$$

and it is then natural to consider a sequence χ_n of characteristic functions such that

$$\chi_n \rightharpoonup \theta \text{ in } L^\infty(0, T) \text{ weak } \star, \tag{3.12}$$

with $0 \leq \theta \leq 1$ a.e. in $(0, T)$; Lemma 1 also tells us that any such θ can be obtained in this way, as can be checked easily directly. One notices then that the corresponding functions y_n, which satisfy a uniform LIPSCHITZ condition, converge uniformly to y_∞ solution of

$$\frac{dy_\infty}{dt} = (1 - \theta)u_* + \theta w \text{ in } (0, T); \ y_\infty(0) = 0, \tag{3.13}$$

and, using the fact that $F\big((1 - \chi)u_* + \chi w\big) = (1 - \chi)F(u_*) + \chi F(w)$ for every function F and every characteristic function χ, one deduces that $J(u_n)$ converges to $\widetilde{J}(\theta)$ given by

$$\widetilde{J}(\theta) = \int_0^T \left(|y_\infty|^2 - (1 - \theta)|u_*|^2 - \theta|w|^2\right) dt. \tag{3.14}$$

If J attains its minimum on U_{ad} at u_*, one deduces that $\widetilde{J}(0) = J(u_*) \leq \lim_n J(u_n) = \widetilde{J}(\theta)$, and therefore $\widetilde{J}$ attains its minimum at 0. One writes then the classical necessary conditions of optimality for $\widetilde{J}$, noticing that admissibility for $\delta\theta$ means $\delta\theta \geq 0$, that δy solves

$$\frac{d\delta y}{dt} = (w - u_*)\delta\theta; \ \delta y(0) = 0, \tag{3.15}$$

and, as $\theta = 0$ corresponds to $y_\infty = y_*$, that

$$\delta\widetilde{J} = \int_0^T \left(2y_* \, \delta y + (|u_*|^2 - |w|^2)\delta\theta\right) dt. \tag{3.16}$$

Using the same p_* as defined in (3.6), the integration by parts gives a different result because the equation for δy is different

$$\int_0^T 2y_* \, \delta y \, dt = -2 \int_0^T \frac{dp_*}{dt} \, \delta y \, dt = \int_0^T 2p_* \frac{d\delta y}{dt} \, dt = \int_0^T 2p_*(w - u_*)\delta\theta \, dt, \tag{3.17}$$

and therefore

$$\delta \widetilde{J} = \int_0^T \Big(2p_*(w - u_*) + (|u_*|^2 - |w|^2) \Big) \delta\theta \, dt, \tag{3.18}$$

and the necessary conditions of optimality for $\widetilde{J}$ become

$$2p_*(w - u_*) + (|u_*|^2 - |w|^2) \geq 0 \text{ a.e. on } (0, T). \tag{3.19}$$

It is only now that one lets w vary in U_{ad} and, taking advantage of the fact that p_* is independent of the choice of w, (3.19) means

$$2p_* \, u_* - |u_*|^2 = \inf_{-1 \leq w \leq 1} \big(2p_* \, w - |w|^2 \big) \text{ a.e. on } (0, T)$$
$$= -2|p_*| - 1 \text{ a.e. on } (0, T), \tag{3.20}$$

and therefore

$$\begin{cases} u_* = -1 \text{ implies } p_* \geq 0 \\ -1 < u_* < +1 \text{ does not occur} \\ u_* = +1 \text{ implies } p_* \leq 0, \end{cases} \tag{3.21}$$

or

$$\begin{cases} p_* < 0 \text{ implies } u_* = +1 \\ p_* = 0 \text{ implies } u_* = \pm 1 \\ p_* > 0 \text{ implies } u_* = -1, \end{cases} \tag{3.22}$$

which are obviously more restrictive than (3.9)/(3.10).

The choice of the model problem comes from the simple direct observation that it has no solution, as I will show later. This fact by itself does not tell much about the existence of solutions for the system of necessary conditions of optimality, but the analysis of the relaxed problem that I will also introduce later will have as a consequence that no such solution exists.

However, one can see by a direct computation that there is no function $u_* \in U_{ad}$ for which the necessary conditions of optimality (3.21)/(3.22) hold, with y_* defined by (3.3) and p_* defined by (3.6). Indeed

$$0 = \int_0^T \frac{d(y_* \, p_*)}{dt} \, dt = \int_0^T \big(u_* \, p_* - |y_*|^2 \big) \, dt = -\int_0^T \big(|p_*| + |y_*|^2 \big) \, dt, \tag{3.23}$$

shows that one must have $y_* = p_* = 0$ a.e. in $(0, T)$, but $y_* = 0$ is incompatible with the condition $u_* = \pm 1$ a.e. in $(0, T)$.

The point of view of PONTRYAGUIN usually gives stronger necessary conditions of optimality than the classical ones.[11] If the necessary conditions of optimality (either the classical ones or those of PONTRYAGUIN) have no solution, then the minimization problem cannot have any solution, but there is then no obvious explanation of what minimizing sequences are doing, for example. Of course, the proof of the necessary

---

[11] If the problem is convex, it gives the same conditions as the classical ones. If the controls are imposed to take values in a discrete set, there is no natural differentiable path from one control to another and therefore one cannot obtain any classical necessary conditions of optimality. However, even for a convex set of admissible functions, if the equation of state has the form $y' = A(y, u)$ and the cost function has the form $J(u) = \int_0^T B(y, u) \, dt$, one requires differentiability of A and B in both y and u in order to obtain the classical necessary conditions of optimality, while one only requires differentiability of A and B in y for obtaining the necessary conditions of optimality of PONTRYAGUIN.

conditions contains a hint about oscillating sequences,[12] and it is Laurence C. YOUNG's point of view to study directly such oscillating sequences, in order to create a relaxed problem.[13]

In order to show directly that our minimization problem has no solution, one first notices that

$$J(u) > -T \text{ for all } u \in U_{ad}, \tag{3.24}$$

because $|y|^2 - |u|^2 \geq -1$ a.e. on $(0, T)$ for every $u \in U_{ad}$ implies $J(u) \geq -T$, and because one cannot have $J(u) = -T$, which would require both $y = 0$ and $|u| = 1$ a.e. on $(0, T)$, in contradiction with the fact that $y = 0$ a.e. on $(0, T)$ implies $u = 0$ a.e. on $(0, T)$. Then one notices that

$$u_n \rightharpoonup 0 \text{ in } L^\infty(0, T) \text{ weak } \star \text{ and } |u_n| = 1 \text{ a.e. in } (0, T) \text{ imply } J(u_n) \to -T, \tag{3.25}$$

as $u_n \rightharpoonup 0$ in $L^\infty(0, T)$ weak $\star$ implies that y_n converges uniformly to 0; an example of such a sequence u_n belonging to U_{ad} is defined by $u_n(t) = sign(\cos n\, t)$ on $(0, T)$.

One sees also that any minimizing sequence, i.e. any sequence $u_n \in U_{ad}$ such that $J(u_n) \to -T = \inf_{u \in U_{ad}} J(u)$, must be such that $y_n \to 0$ in $L^2(0, T)$ strong and $u_n^2 \to 1$ in $L^1(0, T)$ strong . Because U_{ad} is bounded in $L^\infty(0, T)$, $y_n \to 0$ in $L^2(0, T)$ strong is equivalent to $u_n \rightharpoonup 0$ in $L^\infty(0, T)$ weak $\star$, and because $|u_n| \leq 1$ a.e. in $(0, T)$, $u_n^2 \to 1$ in $L^1(0, T)$ strong is equivalent to $u_n^2 \rightharpoonup 1$ in $L^\infty(0, T)$ weak $\star$.

The same analysis shows that if a sequence $u_n \in U_{ad}$ converges in $L^\infty(0, T)$ weak $\star$ to u, one can deduce that y_n converges uniformly to y given by (3.3), but one cannot deduce what the limit of $J(u_n)$ is. However, if one knows that

$$\begin{aligned} u_n \rightharpoonup u \text{ in } L^\infty(0, T) \text{ weak } \star \\ u_n^2 \rightharpoonup v \text{ in } L^\infty(0, T) \text{ weak } \star, \end{aligned} \tag{3.26}$$

then,

$$J(u_n) \to \widetilde{\mathcal{J}}(u, v) = \int_0^T \left(|y|^2 - v \right) dt. \tag{3.27}$$

Lemma 1 characterizes the pairs (u, v) that one can obtain by (3.26) for a sequence $u_n \in U_{ad}$, choosing for K the piece of parabola $U_2 = U_1^2$ with $-1 \leq U_1 \leq 1$, and one can therefore introduce a relaxed problem defined on

$$\widetilde{U}_{ad} = \{ (u, v) | -1 \leq u \leq 1; \ u^2 \leq v \leq 1 \text{ a.e. in } (0, T) \}, \tag{3.28}$$

the state y still being given by (3.3), and the cost function $\widetilde{\mathcal{J}}$ being given by the formula in (3.27). The original problem is a subset of the new one as

$$\begin{aligned} u \in U_{ad} \text{ if and only if } (u, u^2) \in \widetilde{U}_{ad} \\ J(u) = \widetilde{\mathcal{J}}(u, u^2) \text{ for all } u \in U_{ad}. \end{aligned} \tag{3.29}$$

By Lemma 1, for every $(u, v) \in \widetilde{U}_{ad}$ there exists a sequence u_n with $u_n \rightharpoonup u$ and $u_n^2 \rightharpoonup v$ in $L^\infty(0, T)$ weak $\star$, which imply $J(u_n) \to \widetilde{\mathcal{J}}(u, v)$, and using (3.29) one deduces that

$$\inf_{u \in U_{ad}} J(u) = \inf_{(u, v) \in \widetilde{U}_{ad}} \widetilde{\mathcal{J}}(u, v) \tag{3.30}$$

---

[12] The original proof of PONTRYAGUIN's principle, which Jacques-Louis LIONS had asked me to read in the late 60s, contains no Functional Analysis at all, but the idea of switching quickly from u_* to w is explicit there. I found the proof shown above much later, probably in the late 70s or early 80s, and it might have been used in this way before.

[13] Again, I am not really sure about who introduced the term relaxation, but I have probably heard it first in the seminar PALLU DE LA BARRIÈRE at IRIA in the late 60s, when I also heard about parametrized measures, now named YOUNG measures.

and
$$u_* \text{ minimizes } J \text{ on } U_{ad} \text{ if and only if } (u_*, u_*^2) \text{ minimizes } \widetilde{\mathcal{J}} \text{ on } \widetilde{U}_{ad}. \tag{3.31}$$

Let us look for the classical necessary conditions of optimality for $(u_*, v_*) \in \widetilde{U}_{ad}$. Actually these conditions will be sufficient conditions of optimality as $\widetilde{U}_{ad}$ is convex and $\widetilde{\mathcal{J}}$ is a convex function; as $\widetilde{\mathcal{J}}$ is strictly convex in u, u_* is known in advance to be unique, but although $\widetilde{\mathcal{J}}$ is affine in v, one deduces that $v_* = 1$ a.e. on $(0, T)$ from the observation that

$$\widetilde{\mathcal{J}}(u, v) \leq \widetilde{\mathcal{J}}(u, 1) = K(u) - T \text{ with } K(u) = \int_0^T |y|^2 \, dt, \tag{3.32}$$

with equality if and only if $v = 1$ a.e. on $(0, T)$. Obviously K attains its minimum at $u_* = 0$, but let us forget for a moment that $\widetilde{\mathcal{J}}$ attains its minimum only at $(0, 1)$, and let us check what the necessary conditions of optimality for $\widetilde{\mathcal{J}}$ are at an arbitrary point $(u_*, v_*) \in \widetilde{U}_{ad}$. For an admissible direction $(\delta u, \delta v)$, δy is still given by (3.4), but

$$\delta \widetilde{\mathcal{J}} = \int_0^T (2y_* \, \delta y - \delta v) \, dt, \tag{3.33}$$

which, using p_* defined by (3.6) and the integration by parts (3.7) gives

$$\delta \widetilde{\mathcal{J}} = \int_0^T (2p_* \, \delta u - \delta v) \, dt. \tag{3.34}$$

As $\widetilde{U}_{ad}$ is convex, it is equivalent to restrict attention to the admissible directions of the form $(\delta u, \delta v) = (w - u_*, w^2 - v_*)\delta \eta$ with $w \in U_{ad}$ and $\delta \eta \in L^\infty(0, T)$ with $\delta \eta \geq 0$ a.e. in $(0, T)$, and therefore the necessary conditions of optimality can be read as

$$2p_*(w - u_*) - (w^2 - v_*) \geq 0 \text{ a.e. on } (0, T), \tag{3.35}$$

which in the case $v_* = u*^2$ coincide with the PONTRYAGUIN necessary conditions of optimality (3.19). Instead of (3.20), (3.35) implies

$$2p_* \, u_* - v_* = -2|p_*| - 1 \text{ a.e. on } (0, T), \tag{3.36}$$

and therefore
$$u_* \in -sign(p_*); \; v_* = 1 \text{ a.e. on } (0, T), \tag{3.37}$$

and the system of necessary conditions (3.3), (3.6), (3.37) gives then $u_* = 0$.

Instead of the above relaxed problem, I could have used a set much bigger than $\widetilde{U}_{ad}$ by introducing the set of YOUNG measures, which describe the possible weak $\star$ limits of sequences $F(u_n)$ for all continuous functions F. It would have appeared then that only the limits in $L^\infty(0, T)$ weak $\star$ of u_n and u_n^2 were important, i.e. the only useful functions F are the identity id, together with id^2. Therefore starting with a relaxed problem which is too big for the problem at hand, one has not lost information but one carries some unnecessary information; one can reduce the size of the relaxed problem by getting rid of a part of that unnecessary information.

The preceding analysis has identified a topology which is adapted to the initial problem, namely that defined by (3.26). As the weak $\star$ topology of $L^\infty(0, T)$ is metrizable on bounded sets, one can define it using a distance d_0 for the set U_{ad}, and (3.26) corresponds to using the new topology associated to the distance d_1 defined by $d_1(f, g) = d_0(f, g) + d_0(f^2, g^2)$. The set U_{ad} is not complete for the metric d_1, but Lemma 1 describes the completion of U_{ad} which is $\widetilde{U}_{ad}$, and it also shows that $\widetilde{U}_{ad}$ is compact for the metric d_1. The function J is uniformly continuous for d_1, and therefore it extends in a unique way to the completion, and this extension is the function $\widetilde{\mathcal{J}}$ defined in (3.27).

The preceding construction also fulfills the following requirements for a relaxed problem.

An initial problem is to minimize a function F on a set X, [14] a relaxed problem of it is a topological space $\widetilde{X}$, a mapping j from X into $\widetilde{X}$ and a lower semicontinuous function $\widetilde{F}$ defined on $\widetilde{X}$ satisfying the following properties

i) $\widetilde{F}\big(j(x)\big) = F(x)$ for all x in X,

ii) $j(X)$ is dense in $\widetilde{X}$,

iii) For every $y \in \widetilde{X}$ there exists a sequence x_n in X such that $j(x_n)$ converges in $\widetilde{X}$ to y, and $\widetilde{F}(y) = \lim_n F(x_n)$ in the case where the topology of $\widetilde{X}$ is metrizable. If the topology of $\widetilde{X}$ is not metrizable, for every $\varepsilon > 0$ and every neighbourhood V of y in $\widetilde{X}$ there exists $x \in X$ with $j(x) \in V$ and $|\widetilde{F}(y) - F(x)| \leq \varepsilon$.

By i) the relaxed problem contains a copy of the initial one; if ii) was not satisfied one would restrict oneself to the closure of $j(X)$; if iii) was not satisfied, one would replace $\widetilde{F}$ by a larger function by taking the lower semi-continuous envelope of the function equal to $F \circ j^{-1}$ on $j(X)$ and $+\infty$ on $\widetilde{X} \setminus j(X)$, eventually removing the point where this envelope is $+\infty$ if one wants to work with functions which are finite everywhere.

I could have used a smaller set than $\widetilde{U}_{ad}$ by working on U_{ad} but minimizing the function $K(u) - T$, and this would be consistent with the observation (3.32), and although one can still compute $\inf_{u \in U_{ad}} J(u)$ from this new problem, it cannot help discover what is the adequate topology for the initial problem, as defined by (3.26). This new problem is not a relaxed problem of the initial one, as i) is not satisfied, but the function $u \mapsto K(u) - T$ is the Γ-limit of J, i.e. the lower semi-continuous envelope of J, for a particular topology, the $L^\infty(0, T)$ weak $\star$ topology. Indeed, $u_n \rightharpoonup u$ in $L^\infty(0, T)$ weak $\star$ implies $\liminf_n J(u_n) \geq K(u) - T$ and for every $u \in U_{ad}$ there exists a sequence $u_n \in U_{ad}$ with $u_n \rightharpoonup u$ and $u_n^2 \rightharpoonup 1$ in $L^\infty(0, T)$ weak $\star$, and therefore $J(u_n) \to \widetilde{\mathcal{J}}(u, 1) = K(u) - T$.

This last point suggests to be careful with the use of Γ-limits, which are not always the right objects to characterize if one has not choosen the right topology, which might not belong to the list of usual topologies that one is accustomed to use. Another reason to be careful lies in the difference between G-convergence and H-convergence,[15] the latter being more general and adapted to most of the situations occuring in Continuum Mechanics/Physics. As many problems of Optimal Design come from engineering, and often involve Elasticity, it is worth mentioning not only the inadequacy of linearized Elasticity, but the inadequacy of the Γ-convergence approach, which is not the same as Homogenization, to questions of Elasticity. Although an intensive propaganda has made many mathematicians believe that Nature minimizes Energy, it is obviously not so, and one must remember that "conservation of Energy" is the First Principle of Thermodynamics, which no one doubts (of course, one has to include all forms of Energy, including Heat, but in cases where some energy can be stored and released later, one might have to be careful in writing the balance of Energy). Unfortunately Thermodynamics should be called Thermostatics as it only deals with questions at equilibrium, and its Second Principle does not explain what are the possible exchanges of Energy under its different forms and only postulates the result of these exchanges, but it would be quite naive to believe that in an elastic material equilibrium is obtained instantaneously.[16]

---

[14] If X has already a topology, one may forget about it, as j is usually not continuous from X into $\widetilde{X}$.

[15] When I first heard Ennio DE GIORGI talk about Γ-convergence at the seminar of Jacques-Louis LIONS at Collège de France around 1977, I understood it as a natural generalization of his earlier work on G-convergence with Sergio SPAGNOLO [DG&Sp], but I had been impressed by the application that he had mentioned that energy localized on a surface could appear as the Γ-limit of a three-dimensional problem. The natural association which immediately came to my mind was that surface tension in liquids should be determined from three-dimensional laws, and that one should extend the idea of Γ-convergence to evolution problems in order to study that question.

[16] I heard a talk of Joseph KELLER at a meeting of the Institute for Mathematics and its Applications in Minneapolis in 1985, in which he explained damping in real elastic materials by the presence of inhomogeneities together with the effect of geometric nonlinearity. Elastic waves are scattered by the various inclusions in the material, or by the grain boundaries in the case of a polycrystal, but without the nonlinearity of geometric origin in the strain-stress relation there would be no coupling between different modes, and no possible explanation about why Energy gets trapped in higher and higher frequencies, which is the reason why one thinks that one has attained a macroscopic equilibrium.

4. H-convergence

After his one-dimensional counter-example, François MURAT had looked at the more general situation where u is the solution of

$$\mathcal{A}u = -div\Big(a\, grad(u)\Big) = f \text{ in } \Omega,\ u \in H_0^1(\Omega), \tag{4.1}$$

Ω being a bounded open set of R^N, with $f \in L^2(\Omega),$[17] and

$$a \in A_{ad} = \{a | a \in L^\infty(\Omega), \alpha \le a \le \beta \text{ a.e. in } (\Omega)\}, \tag{4.2}$$

with $\alpha > 0$ (so that by LAX-MILGRAM lemma, the operator $\mathcal{A}$ is an isomorphism from $H_0^1(\Omega)$ onto its dual $H^{-1}(\Omega)$), with the intention of minimizing a functional J of the form

$$J(a) = \int_\Omega g\Big(x, u(x), a(x)\Big)\, dx, \tag{4.3}$$

and it is important in the sequel that $grad(u)$ does not enter explicitly in the functional, although some special dependence in $grad(u)$ can be allowed, like $g(u)grad(u)$, $g(u)a\, grad(u)$ or $g(u)\big(a\, grad(u).grad(u)\big)$, with some smoothness and growth properties imposed on g. He had noticed that he could solve explicitly the special case of "layered" media, involving sequences $a^n \in A_{ad}$ depending on x_1 alone for example, and as for (1.4) he assumed that for an interval I such that $\Omega \subset I \times R^{N-1}$

$$a^n \rightharpoonup a_+;\ \frac{1}{a^n} \rightharpoonup \frac{1}{a_-} \text{ in } L^\infty(I) \text{ weak } \star, \tag{4.4}$$

and he had deduced that

$$u_n \rightharpoonup u_\infty \text{ in } H_0^1(\Omega) \text{ weak and } L^2(\Omega) \text{ strong}$$
$$a^n \frac{\partial u_n}{\partial x_1} \rightharpoonup a_- \frac{\partial u_\infty}{\partial x_1} \text{ in } L^2(\Omega) \text{ weak} \tag{4.5}$$
$$a^n \frac{\partial u_n}{\partial x_j} \rightharpoonup a_+ \frac{\partial u_\infty}{\partial x_j} \text{ in } L^2(\Omega) \text{ weak for } j = 2, \ldots, N,$$

and therefore that u_∞ is the solution of

$$\mathcal{A}^{eff} u_\infty = -\sum_{i,j=1,\ldots,N} \frac{\partial}{\partial x_i}\left(A_{ij}^{eff} \frac{\partial u_\infty}{\partial x_j}\right) = f \text{ in } \Omega;\ u_\infty \in H_0^1(\Omega), \tag{4.6}$$

with

$$A_{11}^{eff} = a_-$$
$$A_{jj}^{eff} = a_+, j = 2, \ldots, N \tag{4.7}$$
$$A_{ij}^{eff} = 0, i \ne j.$$

Of course, one first extracts a subsequence for which $u_n \rightharpoonup u_\infty$ in $H_0^1(\Omega)$ weak and $L^2(\Omega)$ strong, and as (4.5) has a unique solution, all the sequence converges. For $j \ne 1$ one has $a^n \frac{\partial u_n}{\partial x_j} = \frac{\partial(a^n u_n)}{\partial x_j}$ and $a^n u_n \rightharpoonup a_+ u_\infty$

---

[17] More generally $f \in H^{-1}(\Omega)$, the dual of $H_0^1(\Omega)$. $H_0^1(\Omega)$ is the closure of smooth functions with compact support in Ω in the SOBOLEV space $H^1(\Omega)$, consisting of functions in $L^2(\Omega)$ having each of their partial derivatives in $L^2(\Omega)$. As Ω is bounded, POINCARÉ inequality holds on $H_0^1(\Omega)$, i.e. $\int_\Omega |u|^2\, dx \le C \int_\Omega |grad(u)|^2|\, dx$. $H_0^1(\Omega)$ is compactly imbedded in $L^2(\Omega)$ and $L^2(\Omega)$ is compactly imbedded in $H^{-1}(\Omega)$. $H^{-1}(\Omega)$ consists of distributions in Ω which are sums of derivatives of functions in $L^2(\Omega)$. If the boundary $\partial\Omega$ of Ω is smooth, there is a notion of trace on the boundary for functions in $H^1(\Omega)$, and $H_0^1(\Omega)$ consists then of those functions in $H^1(\Omega)$ having trace 0 on the boundary.

in $L^2(\Omega)$ weak because $u_n \to u_\infty$ in $L^2(\Omega)$ strong. For computing the limit of $a^n \frac{\partial u_n}{\partial x_1}$,[18] François MURAT used a compactness argument that we had learned from Jacques-Louis LIONS [Li2]: writing $D_i^n = a^n \frac{\partial u_n}{\partial x_i}$, he assumed that $f \in L^2(\Omega)$, and for an interval I in x_1 and a cube ω in $(x_2, \ldots, x_N)$ such that $I \times \omega \subset \Omega$ he observed that $div\, D^n = f$ implies that $\frac{\partial D_1^n}{\partial x_1}$ is bounded in $L^2(I; H^{-1}(\omega))$, and as D_1^n is bounded in $L^2(I; L^2(\omega))$ and the injection of $L^2(\omega)$ into $H^{-1}(\omega)$ is compact, the compactness argument implies that D_1^n stays in a compact set of $L^2(I; H^{-1}(\omega))$. A subsequence D_1^m converges in $L^2(I; H^{-1}(\omega))$ strong to D_1^∞ and therefore for $\varphi \in H_0^1(\omega)$ the function v^m defined in I by $v^m(x_1) = \int_\omega D_1^m(x_1, x_2, \ldots, x_N) \varphi(x_2, \ldots, x_N)\, dx_2 \ldots dx_N$ converges in $L^2(I)$ strong to v^∞ defined in a similar way from D_1^∞, and therefore $\frac{1}{a^m} v^m$ converges in $L^2(\omega)$ weak to $\frac{1}{a_-} v^\infty$, and this implies that $D_1^\infty = a_- \frac{\partial u_\infty}{\partial x_1}$ in $I \times \omega$; varying the cube $I \times \omega$ in Ω gives the same result in Ω.

I noticed later that the analysis of effective properties of layered media can be greatly simplified by using the Div-Curl lemma,[19] which we only proved in 1974 after having completed our basic approach described here, in an attempt to unify the cases for which we were able to compute explicitly the effective coefficient A^{eff}.[20]

Lemma 2: (Div-Curl lemma) Let Ω be an open set of R^N. Let

$$E^n \rightharpoonup E^\infty \text{ in } L^2_{loc}(\Omega; R^N) \text{ weak}$$
$$D^n \rightharpoonup D^\infty \text{ in } L^2_{loc}(\Omega; R^N) \text{ weak}$$
$$div\, D^n \text{ stays in a compact set of } H^{-1}_{loc}(\Omega) \text{ strong}$$
$$curl\, E^n \text{ stays in a compact set of } H^{-1}_{loc}(\Omega; R^{N(N-1)/2}) \text{ strong,} \tag{4.8}$$

then

$$\int_\Omega \Big(\sum_{i=1}^N E_i^n D_i^n \Big) \varphi\, dx \to \int_\Omega \Big(\sum_{i=1}^N E_i^\infty D_i^\infty \Big) \varphi\, dx \text{ for every } \varphi \in C_c(\Omega), \tag{4.9}$$

where $curl\, E$ denotes the lists of all $\frac{\partial E_i}{\partial x_j} - \frac{\partial E_i}{\partial x_j}$ and $C_c(\Omega)$ is the space of continuous functions with compact support in Ω.[21]∎

In 1974,[22] our first proof involved localization in x, FOURIER transform, LAGRANGE formula and PLANCHEREL formula, but in the case where $E^n = grad(u_n)$ with u_n converging weakly to u_∞ in $H^1_{loc}(\Omega)$,

---

[18] He had considered the more general case where $a^n(x) = \prod_i f_n^i(x_i)$ with $0 < \alpha_i \leq f_n^i \leq \beta_i$, with $f_n^i \rightharpoonup f_+^i$ and $\frac{1}{f_n^i} \rightharpoonup \frac{1}{f_-^i}$ in L^∞ weak $\star$, and he had found that $a^n \frac{\partial u_n}{\partial x_j} \rightharpoonup A_{jj}^{eff} \frac{\partial u_\infty}{\partial x_j}$ in $L^2(\Omega)$ weak, with $A_{ii}^{eff} = f_-^i \prod_{j \neq i} f_+^j$ for $i = 1, \ldots, N$, and $A_{ij}^{eff} = 0$ for $i \neq j$.

[19] Some, who want to avoid mentioning the Div-Curl lemma or the more general theory of Compensated Compactness which I also developed with François MURAT in 1976, do not hesitate to lengthen their proofs in order to use only older methods. The correct behaviour in Mathematics is to mention the shortest proof even if one does not follow it, usually because the writer finds it too difficult for himself/herself, and assumes that it would be the same for the reader. Failing to mention such generalizations is a good way to slow down the progress of Science. In this course, the general Compensated Compactness theory (and the theory of H-measures which I developed in the late 80s) will only be used in describing methods for obtaining bounds on effective coefficients.

[20] After learning the term Homogenization, introduced by Ivo BABUŠKA, we called these limiting coefficients "homogenized" coefficients, but after learning the term effective, from George PAPANICOLAOU, I decided to adopt it; it is often used by physicists, even if they have almost never defined it correctly.

[21] One cannot use for φ the characteristic function of a smooth set, for example, but I have noticed that one can develop the theory of H-measures with test functions in $L^\infty \bigcap VMO$, by using a commutation lemma of COIFMAN, ROCHBERG and WEISS, and therefore one can use $\varphi \in L^\infty \bigcap VMO$ in the Div-Curl lemma.

[22] Joel ROBBIN taught me afterwards how to interpret the Div-Curl lemma in terms of differential forms, and he showed me another proof, based on HODGE decomposition. In 1976, François MURAT and I developped the Compensated Compactness theory, following our original proof using FOURIER transform, and PLANCHEREL formula (Proposition 30, Theorem 31).

there is an easier proof by integration by parts, which we had already used: for $\varphi \in C_c^1(\Omega)$ one has $\int_\Omega (\sum_i E_i^n D_i^n) \varphi \, dx = -\langle div(D^n), u_n \varphi \rangle - \int_\Omega u_n (\sum_i D_i^n \partial_i \varphi) \, dx$ and one passes easily to the limit as $u_n \varphi$ converges in $H_0^1(\Omega)$ weak to $u_\infty \varphi$ and u_n converges in $L_{loc}^2(\Omega)$ strong to u_∞. Most applications of the Div-Curl lemma correspond to E^n being a gradient, but I will describe later the Compensated Compactness theorem (Theorem 31), which I also used for questions of bounds for effective coefficients.

As a corollary of the Div-Curl lemma, if a sequence of functions ψ_n which only depends upon x_1 converges in $L_{loc}^2(R)$ weak to ψ_∞, then $\psi_n D_1^n$ converges in $L_{loc}^2(\Omega)$ weak to $\psi_\infty D_1^\infty$ if ψ_n is bounded in L^∞; as a simplification, I say that if D^n satisfies the hypothesis in Lemma 2, then D_1^n does not oscillate in x_1.[23] Before we had solved the general problem, extending the notion of G-convergence introduced by Sergio SPAGNOLO but unaware of his work yet, François MURAT had obtained an explicit formula for the effective coefficients of a layered media in the general anisotropic case; later, in the Spring 1975, Louis NIRENBERG had shown me a preprint of W. MCCONNELL (maybe a preliminary version of [MC]), who had derived the general formula for layered media in linearized Elasticity, and it was the analogue of what François MURAT had done, but with more technical computations of Linear Algebra; a few years after I noticed a general approach for obtaining the effective behaviour of layered media,[24] based on the preceding corollary. The result of François MURAT can be stated as

$$E^n \rightharpoonup E^\infty \text{ in } L_{loc}^2(\Omega; R^N) \text{ weak}; \ D^n = A^n(x_1)E^n \rightharpoonup D^\infty \text{ in } L_{loc}^2(\Omega; R^N) \text{ weak}$$
$$div \, D^n \text{ and the components of } curl \, E^n \text{ stay in a compact set of } H_{loc}^{-1}(\Omega) \text{ strong} \qquad (4.10)$$
$$\text{imply } D^\infty = A^{eff}(x_1)E^\infty \text{ a.e. in } \Omega,$$

where

$$\frac{1}{A_{11}^n} \rightharpoonup \frac{1}{A_{11}^{eff}} \text{ in } L^\infty(\Omega) \text{ weak } \star$$

$$\frac{A_{1i}^n}{A_{11}^n} \rightharpoonup \frac{A_{1i}^{eff}}{A_{11}^{eff}} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for } i = 2, \ldots, N$$

$$\frac{A_{i1}^n}{A_{11}^n} \rightharpoonup \frac{A_{i1}^{eff}}{A_{11}^{eff}} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for } i = 2, \ldots, N \qquad (4.11)$$

$$A_{ij}^n - \frac{A_{i1}^n A_{1j}^n}{A_{11}^n} \rightharpoonup A_{ij}^{eff} - \frac{A_{i1}^{eff} A_{1j}^{eff}}{A_{11}^{eff}} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for } i,j = 2, \ldots, N.$$

The A^n are not assumed to be symmetric, and the proof actually shows that uniform ellipticity of the A^n is not necessary: in the case of layered media in x_1, the result holds if there exists $\alpha > 0$ such that $A_{11}^n \geq \alpha$ a.e. in Ω for all n. The basic idea is that D_1^n does not oscillate in x_1, but also $E_2^n, \ldots, E_N^n$, because of the information on $\partial_1 E_j^n - \partial_j E_1^n$ for $j \geq 2$; one forms the vector G^n with the "good" components $D_1^n, E_2^n, \ldots, E_N^n$, and one expresses the vector O^n of the "oscillating" components $E_1^n, D_2^n, \ldots, D_N^n$, and one has $O^n = B^n(x_1)G^n$ with $B^n = \Phi(A^n)$, and Φ is a well defined (involutive) nonlinear transformation; the corollary of the Div-Curl lemma shows that the weak limit of $B^n G^n$ is $B^\infty G^\infty$, and therefore the explanation

---

[23] More precisely, a sequence v_n converging weakly in $L_{loc}^2(\Omega)$ does not oscillate in a direction ξ_0 if the H-measures associated with subsequences do not charge the direction ξ_0; a consequence is that $v_n f_n((\xi_0 \cdot))$ converges weakly to $v_\infty f_\infty((\xi_0 \cdot))$ if f_n converges weakly to f_∞ in $L_{loc}^2(R)$.

[24] In 1979, working with Georges DUVAUT as consultants for INRIA, we had been asked about an industrial application using layers of steel and rubber. I already knew the method shown here in the linear case, and I explained how to use it for nonlinear Elasticity, but I pointed out that there was no general theory of Homogenization for nonlinear Elasticity (this is still true, as the results based on Γ-convergence do not answer the right questions).

of (4.11) is that whenever A^n only depends upon x_1 and $A_{11}^n \geq \alpha > 0$,

$$D^n = A^n E^n \text{ is the same as } \begin{pmatrix} E_1^n \\ D_2^n \\ \vdots \\ D_N^n \end{pmatrix} = \Phi(A^n) \begin{pmatrix} D_1^n \\ E_2^n \\ \vdots \\ E_N^n \end{pmatrix} \tag{4.12}$$

$$\Phi(A^n) \rightharpoonup \Phi(A^{eff}) \text{ in } L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right) \text{ weak } \star .$$

In the case of linearized Elasticity, the relation is $\sigma_{ij}^n = \sum_{kl} C_{ijkl}^n(x_1)\varepsilon_{kl}^n$, with $\varepsilon_{kl}^n = (\partial_k u_l^n + \partial_l u_k^n)/2$, σ^n is the symmetric CAUCHY stress tensor, and one may always assume that $C_{ijkl}^n = C_{ijlk}^n = C_{jikl}^n$ for all i, j, k, l, but one does not need to use $C_{ijkl}^n = C_{klij}^n$ for all i, j, k, l (which correspond to hyperelastic materials); the equilibrium equations are $\sum_j \partial_j \sigma_{ij}^n = f_i$ for all i; for a direction ξ one defines the acoustic tensor $A^n(\xi)$ by $A_{ik}^n(\xi) = \sum_{jl} C_{ijkl}^n \xi_j \xi_l$ for all i, k, and in the case of layered media in x_1, the formula for layers holds if there exists $\alpha > 0$ such that $(A^n(e^1)\lambda.\lambda) \geq \alpha|\lambda|^2$ a.e. in Ω for all $\lambda \in R^N$ and all n; in that case the components of the good vector G^n are the σ_{i1}^n and σ_{1i}^n for all i and the ε_{kl}^n for $k, l \geq 2$, and the components of the oscillating vector O^n are the ε_{i1}^n and ε_{1i}^n for all i and the σ_{kl}^n for $k, l \geq 2$; one rewrites the constitutive relation $O_{ij}^n = \sum_{kl} \Gamma_{ijkl}^n G_{kl}^n$ and the formulas for layered material in linearized Elasticity that MCCONNELL had derived consist in writing $\Gamma^n \rightharpoonup \Gamma^{eff}$.[25]

Sergio SPAGNOLO had introduced the notion of G-convergence in the late 60s, and unaware of it François MURAT and I had introduced a slightly different concept in the early 70s,[26] for which the name H-convergence was coined much later.[27] We were interested in sequences satisfying

$$u_n \rightharpoonup u_\infty \text{ in } H_{loc}^1(\Omega) \text{ weak}$$
$$- div\left(A^n \, grad(u_n)\right) = f_n \to f \text{ in } H_{loc}^{-1} \text{ strong,} \tag{4.13}$$

where A^n satisfies

$$A^n \text{ bounded in } L^\infty\left(\Omega; \, \mathcal{L}(R^N, R^N)\right); \; (A^n \lambda.\lambda) \geq \alpha|\lambda|^2 \text{ a.e. in } \Omega, \text{ for all } \lambda \in R^N \text{ and all } n, \tag{4.14}$$

with $\alpha > 0$; the bounds on u_n were deduced from an application of LAX-MILGRAM lemma, using specific boundary conditions (we had started with DIRICHLET conditions, but after having read that Sergio SPAGNOLO had noticed that A^{eff} is the same for different boundary conditions, we checked that this

---

[25] In the case of nonlinear Elasticity, the stress tensor used is the PIOLA-KIRCHHOFF stress tensor, which is not symmetric, and the strain F is defined by $F_{ij} = \delta_{ij} + \partial_j u_i$ where $u(x)$ is the displacement from the initial position x of a material point to its final position $x + u(x)$; for a material like steel which breaks if one stretches it more than 10%, F lies near the set of rotations $SO(3)$, but in linearized Elasticity one postulates (often wrongly) that it lies near I; in nonlinear Elasticity, the components of the good vector G^n are the σ_{i1}^n for all i and the F_{ij}^n for all i but only for $j \geq 2$, and the components of the oscillating vector O^n are the F_{i1}^n for all i and the σ_{ij}^n for all i but only for $j \geq 2$; in order to be able to compute the constitutive relation as $O^n = \Psi^n(G^n)$ in the case of hyperelastic materials, a natural condition to impose for the stored energy is the uniform convexity in all directions of the form $a \otimes e^1$.

[26] I had met Sergio SPAGNOLO at a CIME course in Varenna in 1970, and he had asked me if my interpolation results had something to do with his own results, but as soon as he had mentioned that he did not assume any regularity for the coefficients in his work I could tell him that what I had done could not help; however, I did not get a clear idea of what his results were.

[27] The name was choosen by François MURAT in the lectures that he gave in Algiers [Mu3], shortly after I had taught my PECCOT lectures in the Spring 1977, where under the title "Homogénéisation dans les équations aux dérivées partielles" I had described my method of oscillating test functions in Homogenization and the Compensated Compactness theorem, but the notion of H-convergence was indeed clear from our early work.

was also clear in our framework). I used notations from Electrostatics, denoting $E^n = grad(u_n)$ and $D^n = A^n \, grad(u_n)$, and after having extracted a subsequence so that $D^n \rightharpoonup D^\infty$ in $L^2_{loc}(\Omega; R^N)$ weak, the question was to identify what D^∞ was. If one showed that there exists A^{eff} such that $D^\infty = A^{eff} \, E^\infty$, then with usual boundary conditions u_∞ would be the solution of a similar boundary value problem, and this was what Sergio SPAGNOLO had done in the symmetric case; in the nonsymmetric case the knowledge of the inverse mapping $f \mapsto u_\infty$ does not characterize what A^{eff} is, but as nonsymmetric problems do not occur so often in applications, the main advantage of our approach is that after I had introduced my method of oscillating test functions,[28] it generalizes easily to all sort of linear partial differential equations or systems.[29]

In the early 70s, we had started by an abstract elliptic framework, where V is a real separable HILBERT space (corresponding to $H^1_0(\Omega)$ in the concrete example that we had in mind), using $|| \cdot ||$ for the norm on V, $|| \cdot ||_*$ for the norm on V', and $\langle \cdot, \cdot \rangle$ for the duality product between V' and V, and we had considered a bounded sequence $\mathcal{A}_n \in \mathcal{L}(V, V')$ (corresponding to $\mathcal{A}_n u = -div\big(A^n \, grad(u)\big)$ in our example), satisfying a uniform V-ellipticity condition (corresponding to $(A^n(x)\xi.\xi) \geq \alpha|\xi|^2$, $|A^n(x)\xi| \leq M|\xi|$ for all $\xi \in R^N$, a.e. $x \in \Omega$ in our example), i.e. one assumes that there exist $0 < \alpha \leq M < \infty$ such that

$$\langle \mathcal{A}_n u, u \rangle \geq \alpha||u||^2 \text{ and } ||\mathcal{A}_n u||_* \leq M||u|| \text{ for all } u \in V. \tag{4.15}$$

By LAX-MILGRAM lemma, each $\mathcal{A}_n$ is an isomorphism from V onto V', and the first basic result in this abstract framework is the following lemma.

Lemma 3: There exists a subsequence $\mathcal{A}_m$ and a linear continuous operator $\mathcal{A}_{eff}$ from V into V' such that for every $f \in V'$, the sequence of solutions u_m of $\mathcal{A}_m u_m = f$ converges in V weak to the solution u_∞ of $\mathcal{A}_{eff} u_\infty = f$, and $\mathcal{A}_{eff}$ satisfies

$$\langle \mathcal{A}_{eff} u, u \rangle \geq \alpha||u||^2 \text{ and } ||\mathcal{A}_{eff} u||_* \leq \frac{M^2}{\alpha}||u|| \text{ for all } u \in V, \tag{4.16}$$

but $\frac{M^2}{\alpha}$ can be replaced by M if all the operators $\mathcal{A}_n$ are symmetric.
Proof: One has $||(\mathcal{A}_n)^{-1}||_{\mathcal{L}(V',V)} \leq \frac{1}{\alpha}$ as $\mathcal{A}_n \, u_n = f$ implies $\alpha||u_n||^2 \leq \langle \mathcal{A}_n u_n, u_n \rangle = \langle f, u_n \rangle \leq ||f||_* ||u_n||$, so that $||u_n|| \leq \frac{1}{\alpha}||f||_*$. One can extract a subsequence u_m converging in V weak to u_∞, and repeating this extraction for f belonging to a countable dense family F of V' and using a diagonal subsequence, one can extract a subsequence $\mathcal{A}_m$ such that for every $f \in F$, the sequence $u_m = (\mathcal{A}_m)^{-1} f$ converges in V weak to a limit $S(f)$; the sequence $(\mathcal{A}_m)^{-1}$ being uniformly bounded and F being dense, $(\mathcal{A}_m)^{-1} f$ converges in V weak to a limit $S(f)$ for every $f \in V'$. S is a linear continuous operator from V' into V, with $||S f|| \leq \frac{1}{\alpha}||f||_*$ for all $f \in V'$, and in order to show that S is invertible the mere fact that the operators $(\mathcal{A}_n)^{-1}$ are uniformly bounded is not sufficient, because in any infinite dimensional HILBERT space, one can construct a sequence of symmetric surjective isometries converging weakly to 0 (in $L^2(0,1)$ for example, one can take the multiplication by $sign\big(\sin(n \cdot)\big)$). However, the ellipticity condition prevents this difficulty. One notices that $\langle S f, f \rangle = \lim_m \langle u_m, f \rangle = \lim_m \langle \mathcal{A}_m \, u_m, u_m \rangle \geq \alpha \liminf_m ||u_m||^2$ (and $\leq \frac{1}{M}||f||^2_*$ in the symmetric case),

---

[28] The method was discovered independently by Leon SIMON [Si]; his student McCONNELL had only done the layered case in linearized Elasticity, and he had looked himself at the problem in a general way; it was the referee, probably from the Italian school, who had mentioned to him my work. I first wrote about my method in [Ta3], but I had explained it to Jacques-Louis LIONS in the Fall 1975 in Marseille [Li3]; he had been convinced by Ivo BABUŠKA of the importance in engineering of problems with a periodic structure, but he had not thought of asking about the proof of our results which I had mentioned in a meeting that he had organized in June 1974 [Ta2], either to François MURAT who had stayed in Paris, or to me who had spent the year in Madison, where he actually came in the Spring.
[29] As I taught it in my PECCOT lectures, my method extends easily to some monotone situations, but as I mentioned at a meeting in Rio de Janeiro a few months after [Ta7], I could not find a good setting for Homogenization of nonlinear Elasticity. This is still the case, and the spreading error of mistaking Γ-convergence for Homogenization seems to come from the fact that those who commit it had not paid much attention to the difference between G-convergence and H-convergence.

and as $M||u_m|| \geq ||\mathcal{A}_m u_m||_* = ||f||_*$, one finds $\langle S\,f, f\rangle \geq \frac{\alpha}{M^2}||f||_*^2$, and therefore S is invertible by LAX-MILGRAM lemma and its inverse $\mathcal{A}_{eff}$ has a norm bounded by $\frac{M^2}{\alpha}$. As u_m converges in V weak to $S\,f$, one has $\liminf_m ||u_m||^2 \geq ||S\,f||^2$, so that $\langle S\,f, f\rangle \geq \alpha||S\,f||^2$ for every $f \in V'$, or equivalently, as we know now that S is invertible, $\langle \mathcal{A}_{eff}\,v, v\rangle \geq \alpha||v||^2$ for every $v \in V$.■

Of course, as most results in Functional Analysis, this lemma only gives a general framework and does not help much for identifying $\mathcal{A}_{eff}$ in concrete cases,[30] but one often uses the information that $\mathcal{A}_{eff}$ is invertible so that one can choose $f \in V'$ for which the weak limit u_∞ is any element of V prescribed in advance.

In our concrete example, the problem of G-convergence consists in showing that u_∞ solves an equation $-div\big(A^{eff}\,grad(u_\infty)\big) = f$, while the problem of H-convergence consists in showing that $A^m\,grad(u_m)$ converges in $L^2(\Omega; R^N)$ weak to $A^{eff}\,grad(u_\infty)$, and this is a different question in nonsymmetric situations, because if one adds to A^n a constant antisymmetric matrix B (small enough in norm in order to retain the ellipticity condition), one will not change the operator $\mathcal{A}_n$ and therefore the preceding abstract result cannot help identify the precise matrix A^{eff}, as it only uses the operators $\mathcal{A}_n$.

Lemma 3 points to a technical difficulty in the nonsymmetric case, because we started with a bound M for $\mathcal{A}_n$ and ended with the greater bound $\frac{M^2}{\alpha}$ for $\mathcal{A}_{eff}$, and this is solved by defining differently the bounds on the coefficients A^n.

Definition 4: For $0 < \alpha \leq \beta < \infty$, $M(\alpha, \beta; \Omega)$ will denote the set of $A \in L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ satisfying $(A(x)\xi.\xi) \geq \alpha|\xi|^2$ and $(A(x)\xi.\xi) \geq \frac{1}{\beta}|A(x)\xi|^2$ for all $\xi \in R^N$ (or equivalently $(A^{-1}(x)\xi.\xi) \geq \frac{1}{\beta}|\xi|^2$ for all $\xi \in R^N$), a.e. $x \in \Omega$.[31] If A is independent of $x \in \Omega$, one writes $A \in M(\alpha, \beta)$.■

The reason for using the sets $M(\alpha, \beta; \Omega)$ is that they are compact for the topology of H-convergence which I define now.

Definition 5: We will say that a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converges to $A^{eff} \in M(\alpha', \beta'; \Omega)$ for some $0 < \alpha' \leq \beta' < \infty$,[32] if for every $f \in H^{-1}(\Omega)$, the sequence of solutions $u_n \in H_0^1(\Omega)$ of $-div\big(A^n\,grad(u_n)\big) = f$ converges in $H_0^1(\Omega)$ weak to u_∞, and the corresponding sequence $A^n\,grad(u_n)$ converges in $L^2(\Omega; R^N)$ weak to $A^{eff}\,grad(u_\infty)$; u_∞ is therefore the solution of $-div\big(A^{eff}\,grad(u_\infty)\big) = f$ in Ω.■

H-convergence comes from a topology on $X = \bigcup_{n \geq 1} M\big(\frac{1}{n}, n; \Omega\big)$, union of all $M(\alpha, \beta; \Omega)$, which is the coarsest topology that makes a list of maps continuous. For $f \in H^{-1}(\Omega)$ one such map is $A \mapsto u$ from X into $H_0^1(\Omega)$ weak and another one is $A \mapsto A\,grad(u)$ from X into $L^2(\Omega; R^N)$ weak, where u is the solution of $-div\big(A\,grad(u)\big) = f$. When one restricts that topology to $M(\alpha, \beta; \Omega)$, it is equivalent to consider only f belonging to a countable bounded set whose combinations are dense in $H^{-1}(\Omega)$; then u and $A\,grad(u)$ belong to bounded sets respectively of $H_0^1(\Omega)$ and $L^2(\Omega; R^N)$ which are metrizable for their respectuve weak topology, and therefore the restriction of that topology to $M(\alpha, \beta; \Omega)$ is defined by a countable number of semi-distances and is therefore defined by a semi-distance. That it is actually a distance can be seen by showing uniqueness of the limit: if a sequence A^n H-converges to both A^{eff} and to B^{eff}, then one deduces that $A^{eff}\,grad(u_\infty) = B^{eff}\,grad(u_\infty)$ a.e. $x \in \Omega$ for every $f \in H^{-1}(\Omega)$, and therefore for every $u_\infty \in H_0^1(\Omega)$; choosing then u_∞ to coincide successively with x_j, $j = 1, \ldots, N$, on an open subset ω with compact closure in Ω, one must have $A^{eff} = B^{eff}$ a.e. $x \in \omega$. Of course, one never needs much from this topology, but some arguments do make use of the fact that $M(\alpha, \beta; \Omega)$ is metrizable.

However, if one wants to let α tend to 0, like some people do for domains with holes, one must be very careful because one cannot use arguments based on metrizability in that case.

Theorem 6: For any sequence $A^n \in M(\alpha, \beta; \Omega)$ there exists a subsequence A^m and an element $A^{eff} \in M(\alpha, \beta; \Omega)$ such that A^m H-converges to A^{eff}.

---

[30] Even when all the operators $\mathcal{A}_n$ are differential operators, it may happen that $\mathcal{A}_{eff}$ is not a differential operator and in some cases nonlocal integral corrections must be taken into account.

[31] If $A \in \mathcal{L}(R^N, R^N)$ satisfies $(A\,\xi.\xi) \geq \frac{1}{\beta}|A\,\xi|^2$ for all $\xi \in R^N$, then $|A\,\xi| \leq \beta|\xi|$ for all $\xi \in R^N$, but if $A \in \mathcal{L}(R^N, R^N)$ satisfies $(A\,\xi.\xi) \geq \alpha|\xi|^2$ and $|A\,\xi| \leq M|\xi|$ for all $\xi \in R^N$, then one can only deduce $(A\,\xi.\xi) \geq \frac{\alpha}{M^2}|A\,\xi|^2$, if A is not symmetric, while of course one has $(A\,\xi.\xi) \geq \frac{1}{M}|A\,\xi|^2$ if A is symmetric.

[32] Theorem 6 shows that one can take $\alpha' = \alpha$ and $\beta' = \beta$.

Proof: Using the same argument than in Lemma 3, F being a countable dense set of $H^{-1}(\Omega)$, we can extract a subsequence A^m such that for every $f \in F$ the sequence $u_m \in H_0^1(\Omega)$ of solutions of $-div(A^m\, grad(u_m)) = f$ converges in $H_0^1(\Omega)$ weak to $u_\infty = S(f)$ and $A^m\, grad(u_m)$ converges in $L^2(\Omega; R^N)$ weak to $R(f)$; the same is true then for all $f \in H^{-1}(\Omega)$, the operator S is invertible, and $R(f) = C\, u_\infty$ where C is a linear continuous operator from $H_0^1(\Omega)$ into $L^2(\Omega; R^N)$. It remains to show that C is local, of the form $C\, v = A^{eff}\, grad(v)$ for all $v \in H_0^1(\Omega)$, and that $A^{eff} \in M(\alpha, \beta; \Omega)$. We first show that for all $v \in H_0^1(\Omega)$ one has $(C\, v.grad(v)) \geq \alpha|grad(v)|^2$ and $(C\, v.grad(v)) \geq \frac{1}{\beta}|C\, v|^2$ a.e. $x \in \Omega$.

For $v \in H_0^1(\Omega)$, let $f = -div(C\, v)$, so that $u_\infty = v$, and let φ be a smooth function so that we may use $\varphi\, u_m$ and $\varphi\, v$ as test functions. One gets $\langle f, \varphi\, u_m \rangle = \int_\Omega (A^m\, grad(u_m).\varphi\, grad(u_m) + u_m\, grad(\varphi))\, dx$, and as u_m converges strongly to v in $L^2(\Omega)$ because $H_0^1(\Omega)$ is compactly imbedded into $L^2(\Omega)$, one deduces that $\langle f, \varphi\, v \rangle = \lim_m \int_\Omega \varphi(A^m\, grad(u_m).grad(u_m))\, dx + \int_\Omega (C\, v.v\, grad(\varphi))\, dx$, but $\langle f, \varphi\, v \rangle = \int_\Omega (C\, v.\varphi\, grad(v) + v\, grad(\varphi))\, dx$, and therefore one deduces that for every smooth function φ one has[33]

$$\int_\Omega \varphi\Big(A^m\, grad(u_m).grad(u_m)\Big)\, dx \to \int_\Omega \varphi\Big(C\, v.grad(v)\Big)\, dx. \tag{4.17}$$

Choosing now φ to be nonnegative, and using the first part of the definition of $M(\alpha, \beta; \Omega)$, we deduce that

$$\int_\Omega \varphi\Big(C\, v.grad(v)\Big)\, dx \geq \alpha \liminf_m \int_\Omega \varphi|grad(u_m)|^2\, dx \geq \alpha \int_\Omega \varphi|grad(v)|^2\, dx, \tag{4.18}$$

where the second inequality follows from the fact that $grad(u_m)$ converges in $L^2(\Omega; R^N)$ weak to $grad(v)$, and as this inequality holds for all smooth nonnegative functions φ, one obtains

$$\Big(C\, v.grad(v)\Big) \geq \alpha|grad(v)|^2 \text{ a.e. } x \in \Omega, \text{ for every } v \in H_0^1(\Omega). \tag{4.19}$$

Using the second part of the definition of $M(\alpha, \beta; \Omega)$, we deduce that

$$\int_\Omega \varphi\Big(C\, v.grad(v)\Big)\, dx \geq \frac{1}{\beta} \liminf_m \int_\Omega \varphi|A^m\, grad(u_m)|^2\, dx \geq \alpha \int_\Omega \varphi|C\, v|^2\, dx, \tag{4.20}$$

as $A^m\, grad(u_m)$ converges in $L^2(\Omega; R^N)$ weak to $C\, v$, so that

$$\Big(C\, v.grad(v)\Big) \geq \frac{1}{\beta}|C\, v|^2 \text{ a.e. } x \in \Omega, \text{ for every } v \in H_0^1(\Omega). \tag{4.21}$$

From (4.21) one deduces

$$|C\, v| \leq \beta|grad(v)| \text{ a.e. } x \in \Omega, \text{ for every } v \in H_0^1(\Omega), \tag{4.22}$$

and as C is linear, (4.22) implies that

$$\text{if } grad(v) = grad(w) \text{ a.e. in an open subset } \omega, \text{ then } C\, v = C\, w \text{ a.e. in } \omega. \tag{4.23}$$

Writing Ω as the union of an increasing sequence ω_k of open subsets with compact closure in Ω, we define A^{eff} in the following way: for $\xi \in R^N$, we choose $v_k \in H_0^1(\Omega)$ such that $grad(v_k) = \xi$ on ω_k, and we define $A^{eff}\, \xi$ on ω_k to be the restriction of $C(v_k)$ to ω_k; this defines $A^{eff}\, \xi$ as a measurable function in Ω because $C\, v_k$ and $C\, v_l$ coincide on $\omega_k \cap \omega_l$ by (4.23); (4.23) also implies that A^{eff} is linear in ξ. If $w \in H_0^1(\Omega)$ is piecewise affine so that $grad(w)$ is piecewise constant, then (4.23) implies that $C\, w = A^{eff}\, grad(w)$ a.e. $x \in \Omega$. As piecewise affine functions are dense in $H_0^1(\Omega)$, for each $v \in H_0^1(\Omega)$ there is a sequence w_j of piecewise affine functions such that $grad(w_j)$ converges strongly in $L^2(\Omega; R^N)$ to $grad(v)$, and as $|C\, v - A^{eff}\, grad w_k| = |C\, v - C\, w_j| \leq \beta|grad(v - w_j)|$ a.e. $x \in \Omega$, one deduces $C\, v = A^{eff}\, grad(v)$

---

[33] One could deduce (4.17) directly from the Div-Curl lemma, but I am showing how we first argued, and we had proved this result before discovering the Div-Curl lemma.

a.e. $x \in \Omega$. Having shown that $C\,v = A^{eff}\,grad(v)$ for a measurable A^{eff}, (4.19) and (4.21) imply that $A^{eff} \in M(\alpha, \beta; \Omega)$ as one can take v to be any affine function in an open subset ω with compact closure in Ω.$\blacksquare$

At a meeting in Roma in the Spring of 1974,[34] I had apparently upset Ennio DE GIORGI by my claim that my method (which was actually the joint work with François MURAT) was more general than the method developed by the Italian school. I thought that MEYERS's regularity theorem which Sergio SPAGNOLO was using in his proof was based on the maximum principle [Me], but Ennio DE GIORGI had told me that it was not restricted to second order equations. Actually, it would have been difficult for me at that time to explain how to perform all the computations for higher order nonnecessarily symmetric equations or to systems like linearized Elasticity, and even more difficult to explain what to do for nonlinear elliptic equations, but less than a year after I had noticed that most of the important known properties of Homogenization of second order variational elliptic equations in divergence form could be obtained through repeated applications of the Div-Curl lemma, and the extension to linear elliptic systems in a variational framework (and some simple nonlinear systems of monotone type) really became straightforward. As my method, which has been wrongly called the "energy method" and which I prefer to call the "method of oscillating test functions", uses only a variational structure, it can be extended with minor changes[35] to most of the linear partial differential equations of Continuum Mechanics (not much is understood for nonlinear equations).

One starts with the same abstract analysis, Lemma 3 and the beginning of the proof of Theorem 6; one extracts a subsequence A^m for which there is a linear continuous operator C from $H_0^1(\Omega)$ into $L^2(\Omega; R^N)$ such that for every $f \in H^{-1}(\Omega)$ the sequence of solutions $u_m \in H_0^1(\Omega)$ of $-div\big(A^m\,grad(u_m)\big) = f$ converges in $H_0^1(\Omega)$ weak to u_∞ and $A^m\,grad(u_m)$ converges in $L^2(\Omega; R^N)$ weak to $R(f) = C(u_\infty)$. One uses then the Div-Curl lemma to give a new proof that C is a local operator of the form $C(v) = A^{eff}\,grad(v)$ with $A^{eff} \in L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$.

One constructs a sequence of oscillating test functions v_m satisfying

$$-div\Big((A^m)^T grad(v_m)\Big) \text{ converges in } H_{loc}^{-1}(\Omega) \text{ strong,} \tag{4.24}$$

where $(A^m)^T$ is the transposed operator of A^m and

$$v_m \rightharpoonup v_\infty \text{ in } H^1(\Omega) \text{ weak; } (A^m)^T grad(v_m) \rightharpoonup w_\infty \text{ in } L^2(\Omega; R^N) \text{ weak,} \tag{4.25}$$

and one passes to the limit in the identity

$$\Big(A^m\,grad(u_m).grad(v_m)\Big) = \Big(grad(u_m).(A^m)^T grad(v_m)\Big). \tag{4.26}$$

The Div-Curl lemma applies to the left side of (4.26) which converges in the sense of measures[36] to $\big(C(u_\infty).grad(v_\infty)\big)$, because $div\big(A^m\,grad(u_m)\big)$ is a fixed element of $H^{-1}(\Omega)$ and $grad(v_m)$ converges in $L^2(\Omega; R^N)$ weak to $grad(v_\infty)$; the Div-Curl lemma also applies to the right side of (4.26) which converges in the sense of measures to $(grad(u_\infty).w_\infty)$, because of (4.24) and (4.25), and this shows

$$\big(C(u_\infty).grad(v_\infty)\big) = (grad(u_\infty).w_\infty) \text{ a.e. in } \Omega. \tag{4.27}$$

---

[34] Umberto MOSCO had insisted that I should write something before I left, and instead of visiting Roma, I stayed in the hotel, so nicely located above Piazza di Spagna, writing [Ta1] and another short description concerning quasi-variational inequalities.

[35] Becoming a mathematician requires some ability with abstract concepts, and it is part of the training to check that one can apply a general method to particular examples, and it depends upon one's taste and one's own scientific stature to decide if it was an exercise or something worth publishing.

[36] A sequence s_n converges to s_∞ in the sense of distributions in Ω if for every $\varphi \in C_c^\infty(\Omega)$, the space of indefinitely differentiable functions with compact support in Ω, one has $\langle s_n, \varphi \rangle \to \langle s_\infty, \varphi \rangle$; the s_n must be distributions and according to the theory of distributions of Laurent SCHWARTZ the limit s_∞ is automatically a distribution. The sequence s_n converges to s_∞ in the sense of measures if the preceding convergence holds for every $\varphi \in C_c(\Omega)$, the space of continuous functions with compact support in Ω; the s_n must be RADON measures and s_∞ is automatically a RADON measure. If $s_n \in L_{loc}^1(\Omega)$, then $\langle s_n, \varphi \rangle$ means $\int_\Omega s_n\,\varphi\,dx$.

One constructs a sequence v_m satisfying (4.24) and (4.25) by first choosing an open set Ω' containing the closure of Ω, then extending A^m in $\Omega' \setminus \Omega$ for example by $A^m(x) = \alpha I$ for $x \in \Omega' \setminus \Omega$, and then choosing $v_m \in H_0^1(\Omega')$ solution of

$$-div\Big((A^m)^T \, grad(v_m)\Big) = g \text{ in } \Omega', \tag{4.28}$$

for some $g \in H^{-1}(\Omega')$. One obtains a sequence v_m bounded in $H_0^1(\Omega')$ and therefore its restriction to Ω is bounded in $H^1(\Omega)$, and a subsequence satisfies (4.24) and (4.25). By Lemma 3 one can choose $g \in H^{-1}(\Omega')$ such that v_∞ is any arbitrary element of $H_0^1(\Omega')$, and in particular for each $j = 1, \ldots, N$, there exists $g_j \in H^{-1}(\Omega')$ such that $v_\infty = x_j$ a.e. in Ω, and using these N choices of g_j, (4.27) means $C \, u_\infty = A^{eff} \, grad(u_\infty)$ for some $A^{eff} \in L^2\big(\Omega; \mathcal{L}(R^N, R^N)\big)$.

This method quickly gives only an intermediate result and is not very good for questions of bounds, and it must emphasized that questions of bounds are not yet very well understood for general equations or systems. As pointed out by François MURAT, one can easily show that $A^{eff} \in L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ by using the following lemma based on the continuity of C, but the information $A^{eff} \in M(\alpha, \beta; \Omega)$ is not so natural in this approach.

Lemma 7: If $M \in L^2\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ and the operator C defined by $C(v) = M \, grad(v)$ for all $v \in H_0^1(\Omega)$ is a linear continuous operator from $H_0^1(\Omega)$ into $L^2(\Omega; R^N)$ of norm $\leq \gamma$, then one has $M \in L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ and $||M(x)||_{\mathcal{L}(R^N, R^N)} \leq \gamma$ a.e. in Ω.
Proof: Let $\xi \in R^N \setminus 0$ and $\varphi \in C_c^1(\Omega)$, the space of functions of class C^1 with compact support in Ω. Define φ_n by $\varphi_n(x) = \varphi(x)\frac{\sin n(\xi.x)}{n}$ for $x \in \Omega$, which gives a bounded sequence in $H_0^1(\Omega)$ with $\lim_n ||grad(\varphi_n)||_{L^2(\Omega; R^N)} = \frac{1}{\sqrt{2}}||\xi \, \varphi||_{L^2(\Omega; R^N)} = \frac{|\xi|}{\sqrt{2}}||\varphi||_{L^2(\Omega)}$. $C(\varphi_n) = M\big(\frac{\sin n(\xi.\cdot)}{n} grad(\varphi) + \varphi \cos n(\xi.\cdot)\xi\big)$, and therefore $\lim_n ||C(\varphi_n)||_{L^2(\Omega; R^N)} = \frac{1}{\sqrt{2}}||\varphi \, M \, \xi||_{L^2(\Omega)}$, because φ and $grad(\varphi)$ are bounded, so that $\varphi \, M \, \xi$ and $M \, grad(\varphi)$ belongs to $L^2(\Omega; R^N)$. Therefore one deduces that $||\varphi \, M \, \xi||_{L^2(\Omega)} \leq \gamma|\xi| \, ||\varphi||_{L^2(\Omega)}$ for every $\varphi \in C_c^1(\Omega)$, and therefore for every $\varphi \in L^2(\Omega)$ by density, and this means $||M \, \xi||_{L^\infty(\Omega; R^N)} \leq \gamma|\xi|$, and as this is valid for all ξ, the lemma is proved.∎

Using the same approach,[37] one can derive a few useful properties of H-convergence, the main tool remaining the Div-Curl lemma.

Proposition 8: If a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converges to A^{eff}, then the transposed sequence $(A^n)^T$ H-converges to $(A^{eff})^T$. In particular if a sequence A^n H-converges to A^{eff} and if $A^n(x)$ is symmetric a.e. $x \in \Omega$ for all n, then $A^{eff}(x)$ is symmetric a.e. $x \in \Omega$.
Proof: $A \in M(\alpha, \beta; \Omega)$ implies (and therefore is equivalent to) $A^T \in M(\alpha, \beta; \Omega)$ as $A \in M(\alpha, \beta; \Omega)$ means $(A(x)\xi.\xi) \geq \alpha|\xi|^2$ and $(A^{-1}(x)\xi.\xi) \geq \frac{1}{\beta}|\xi|^2$ for all $\xi \in R^N$, a.e. $x \in \Omega$, and as $(A^{-1})^T = (A^T)^{-1}$, this is the same as $(A^T(x)\xi.\xi) \geq \alpha|\xi|^2$ and $\big((A^T)^{-1}(x)\xi.\xi\big) \geq \frac{1}{\beta}|\xi|^2$ for all $\xi \in R^N$, a.e. $x \in \Omega$. By Theorem 6 a subsequence $(A^m)^T$ H-converges to B^{eff}. For $f, g \in H^{-1}(\Omega)$, let us define the sequences $u_m, v_m \in H_0^1(\Omega)$ by

$$-div\Big(A^m \, grad(u_m)\Big) = f, \quad -div\Big((A^m)^T \, grad(v_m)\Big) = g \text{ in } \Omega, \tag{4.29}$$

so that u_m and v_m converge in $H_0^1(\Omega)$ weak respectively to u_∞ and v_∞, $A^m \, grad(u_m)$ and $(A^m)^T grad(v_m)$ converge in $L^2(\Omega; R^N)$ weak respectively to $A^{eff} \, grad(u_\infty)$ and $B^{eff} grad(v_\infty)$, and one can use the Div-Curl lemma to take the limit of the identity

$$\Big(A^m \, grad(u_m).grad(v_m)\Big) = \Big(grad(u_m).(A^m)^T grad(v_m)\Big), \tag{4.30}$$

and obtain

$$\Big(A^{eff} \, grad(u_\infty).grad(v_\infty)\Big) = \Big(grad(u_\infty).B^{eff} grad(v_\infty)\Big), \tag{4.31}$$

---

[37] The preceding method is not adapted to nonlinear problems, but I developed a variant where the oscillating test functions satisfy the initial equation instead of the transposed equation, and it extends to the case of monotone operators.

and as u_∞ and v_∞ can be arbitrary elements of $H_0^1(\Omega)$ by Lemma 3, this implies $(A^{eff})^T = B^{eff}$ a.e. in Ω. The second part of the Proposition results from uniqueness of H-limits.∎

The next result shows that H-convergence inside Ω is not related to any particular boundary condition imposed on $\partial\Omega$.

Proposition 9: If a sequence $A^n \in M(\alpha,\beta;\Omega)$ H-converges to A^{eff} and a sequence u_n converges in $H_{loc}^1(\Omega)$ weak to u_∞, and $div\big(A^n\,grad(u_n)\big)$ belongs to a compact set of $H_{loc}^{-1}(\Omega)$ strong, then the sequence $A^n\,grad(u_n)$ converges to $A^{eff}\,grad(u_\infty)$ in $L_{loc}^2(\Omega;R^N)$ weak.

Proof: Let $\varphi \in C_c^1(\Omega)$ so that $\varphi\,u_n$ converges in $H_0^1(\Omega)$ weak to $\varphi\,u_\infty$, $\varphi\,grad(u_n)$ converges in $L^2(\Omega;R^N)$ weak to $\varphi\,grad(u_\infty)$; $curl\big(\varphi\,grad(u_n)\big)$ has its components bounded in $L^2(\Omega)$, as they are of the form $\frac{\partial\varphi}{\partial x_j}\frac{\partial u_n}{\partial x_k} - \frac{\partial\varphi}{\partial x_k}\frac{\partial u_n}{\partial x_j}$. As $div\big(\varphi\,A^n\,grad(u_n)\big) = \varphi\,div\big(A^n\,grad(u_n)\big) + (A^n\,grad(u_n).grad(\varphi))$, it belongs to a compact set of $H^{-1}(\Omega)$ strong, as multiplication by φ maps $H_{loc}^{-1}(\Omega)$ into $H^{-1}(\Omega)$ and $(A^n\,grad(u_n).grad(\varphi))$ is bounded in $L^2(\Omega)$. One extracts a subsequence such that $\varphi\,A^m\,grad(u_m)$ converges in $L^2(\Omega;R^N)$ weak to w_∞. For $f \in H^{-1}(\Omega)$, one defines $v_n \in H_0^1(\Omega)$ by $-div\big((A^n)^T\,grad(v_n)\big) = f$, so that v_n converges in $H_0^1(\Omega)$ weak to v_∞ and $(A^n)^T\,grad(v_n)$ converges in $L^2(\Omega;R^N)$ weak to $(A^{eff})^T\,grad(v_\infty)$ by Proposition 8. One then passes to the limit in both sides of $\big(\varphi\,A^m\,grad(u_m).grad(v_m)\big) = \big(\varphi\,grad(u_m).(A^m)^T\,grad(v_m)\big)$ by using the Div-Curl lemma, and one obtains the relation $(w_\infty.grad(v_\infty)) = \big(\varphi\,grad(u_\infty).(A^{eff})^T\,grad(v_\infty)\big)$ a.e. in Ω. As v_∞ is an arbitrary element of $H_0^1(\Omega)$ by Lemma 3, $w_\infty = \varphi\,A^{eff}\,grad(u_\infty)$ a.e. in Ω, and as φ is arbitrary in $C_c^1(\Omega)$ and the limit does not depend upon which subsequence has been chosen, one deduces that all the sequence $A^n\,grad(u_n)$ converges in $L_{loc}^2(\Omega;R^N)$ weak to $A^{eff}\,grad(u_\infty)$.∎

In the preceding proof, $div\big(A^n\,grad(\varphi\,u_n)\big)$ may not belong to a compact set of $H^{-1}(\Omega)$ strong as it is $div\big(\varphi\,A^n\,grad(u_n)\big) + div\big(u_n\,A^n\,grad(\varphi)\big)$ and $div\big(\varphi\,A^n\,grad(u_n)\big)$ does indeed belong to a compact set of $H^{-1}(\Omega)$ strong as was already used, but it is not clear if $div\big(u_n\,A^n\,grad(\varphi)\big)$ does, because $u_n\,A^n\,grad(\varphi)$ may only converge in $L^2(\Omega;R^N)$ weak. We have used then the complete form of the Div-Curl lemma and not only the special case where one only considers gradients.

Proposition 9 expresses that the boundary conditions used for u_n are not so important, as long as the solutions stay bounded, as had been noticed by Sergio SPAGNOLO in the case of G-convergence. We did define H-convergence by using DIRICHLET conditions, but the result inside Ω would be the same for other boundary conditions, if one can apply LAX-MILGRAM lemma for existence as we need to start by using Lemma 3. Using DIRICHLET conditions has the advantage that no smoothness assumption is necessary for the boundary of Ω. What happens on the boundary $\partial\Omega$ may depend upon the particular boundary condition used; the particular cases of nonhomogeneous DIRICHLET conditions, NEUMANN conditions, and other variational conditions can all be considered at once in the framework of variational inequalities, allowing actually some nonlinearity in the boundary conditions (the nonlinearity inside Ω is a different matter).

The next result states that H-convergence has a local character, extending the corresponding result of Sergio SPAGNOLO for G-convergence.

Proposition 10: If a sequence $A^n \in M(\alpha,\beta;\Omega)$ H-converges to A^{eff}, and ω is an open subset of Ω, then the sequence $M^n = A^n\big|_\omega$ of the restrictions of A^n to ω H-converges to $M^{eff} = A^{eff}\big|_\omega$. Therefore if a sequence $B^n \in M(\alpha,\beta;\Omega)$ H-converges to B^{eff} and $A^n = B^n$ for all n, a.e. $x \in \omega$, then $A^{eff} = B^{eff}$ a.e. $x \in \omega$.

Proof: If all A^n belong to $M(\alpha,\beta;\Omega)$, then all M^n belong to $M(\alpha,\beta;\omega)$ and by Theorem 6 a subsequence M^m H-converges to some $M^{eff} \in M(\alpha,\beta;\omega)$. For $f \in H^{-1}(\omega)$ and $g \in H^{-1}(\Omega)$, let us solve $-div\big(M^m\,grad(u_m)\big) = f$ in ω and $-div\big((A^m)^T\,grad(v_m)\big) = g$ in Ω so that u_m converges in $H_0^1(\omega)$ weak to u_∞, $M^m\,grad(u_m)$ converges in $L^2(\omega;R^N)$ weak to $M^{eff}\,grad(u_\infty)$, v_m converges in $H_0^1(\Omega)$ weak to v_∞, and $(A^m)^T\,grad(v_m)$ converges in $L^2(\Omega;R^N)$ weak to $(A^{eff})^T\,grad(v_\infty)$. Extending u_n and u_∞ by 0 in $\Omega \setminus \omega$, one can apply the Div-Curl lemma in ω for the left side and in Ω for the right side of the equality $\big(M^m\,grad(u_m).grad(v_m)\big) = \big(grad(u_m).(A^m)^T\,grad(v_m)\big)$, and one obtains $\big(M^{eff}\,grad(u_\infty).grad(v_\infty)\big) = \big(grad(u_\infty).(A^{eff})^T\,grad(v_\infty)\big)$ a.e. in ω. As by Lemma 3 u_∞ can be arbitrary in $H_0^1(\omega)$ and v_∞ arbitrary in $H_0^1(\Omega)$, one deduces that $M^{eff} = A^{eff}$ a.e. in ω; as the H-limit is independent of the subsequence used, all the sequence M^n H-converges to $A^{eff}\big|_\omega$.∎

Actually if for a measurable subset ω of Ω, one has $A^n = B^n$ for all n, a.e. $x \in \omega$, and the sequences $A^n, B^n \in M(\alpha,\beta;\Omega)$ H-converge respectively in Ω to A^{eff}, B^{eff}, then one has $A^{eff} = B^{eff}$ a.e. $x \in \omega$.

This can be proved by applying the same regularity theorem of MEYERS [Me] that Sergio SPAGNOLO used in the symmetric case. It is equivalent to prove that $A^{eff} = B^{eff}$ a.e. in $\omega(\varepsilon)$ for each $\varepsilon > 0$, where $\omega(\varepsilon)$ is the set of points of ω at a distance at least ε from $\partial\Omega$. Defining v_n as above but choosing $f \in H^{-1}(\Omega)$ and $u_n \in H_0^1(\Omega)$ instead, the problem is to use $\chi_{\omega(\varepsilon)}$, the characteristic function of $\omega(\varepsilon)$, as a test function in the Div-Curl lemma in Ω. For obtaining that result one first takes $f, g \in W^{-1,p}(\Omega)$ with $p > 2$, and MEYERS's regularity theorem tells that $grad(u_n)$ and $grad(v_n)$ stay bounded in $L^{q(\varepsilon)}\big(\omega(\varepsilon)\big)$ for some $q(\varepsilon) \in (2, p]$, and therefore $\big(A^n\, grad(u_n).grad(v_n)\big)$ and $\big(grad(u_n).(B^n)^T\, grad(v_n)\big)$ (which are equal on ω) are bounded in $L^{q(\varepsilon)/2}\big(\omega(\varepsilon)\big)$, and converge in $L^{q(\varepsilon)/2}\big(\omega(\varepsilon)\big)$ weak to $\big(A^{eff}\, grad(u_\infty).grad(v_\infty)\big)$ and $\big(grad(u_\infty).(B^{eff})^T\, grad(v_\infty)\big)$ which are then equal a.e. in $\omega(\varepsilon)$. As $W^{-1,p}(\Omega)$ is dense in $H^{-1}(\Omega)$, one can pass to the limit in this equality in $\omega(\varepsilon)$ and obtain it for arbitrary $f, g \in H^{-1}(\Omega)$, i.e. for arbitrary $u_\infty, v_\infty \in H_0^1(\Omega)$, and that gives $A^{eff} = B^{eff}$ a.e. in $\omega(\varepsilon)$, and therefore a.e. in ω.

The argument of Proposition 10 is variational and extends therefore to all variational situations, while in order to extend the preceding argument to a general situation one would have to prove a regularity theorem like MEYERS's one, and I do not know if this has been done; it has been checked by Jacques-Louis LIONS that the analogous statement is valid for some linearized Elasticity systems.

5. Bounds on effective coefficients: first method

In the case $A^n = a^n I$, which we had investigated first, we knew that the $L^\infty(\Omega)$ weak $\star$ limits of a^n and $\frac{1}{a^n}$, denoted respectively by a_+ and $\frac{1}{a_-}$, are needed for expressing A^{eff} in the case where a^n only depends upon one variable. We had obtained sequences $E^n = grad(u_n)$ and $D^n = a^n\, E^n = a^n\, grad(u_n)$ converging in $L^2(\Omega; R^N)$ weak, respectively to E^∞ and D^∞, and the analogue of (4.17) had told us that $a^n |E^n|^2$ was converging in the sense of measures to $(D^\infty.E^\infty)$. We had then decided to look at the convex hull in R^{2N+3} of the set

$$K = \left\{ \left(E, a\, E, a|E|^2, a, \frac{1}{a} \right) \big| E \in R^N, a \in [\alpha, \beta] \right\}, \tag{5.1}$$

with the goal of investigating what could be deduced if D^∞ and E^∞ satisfy the property

$$\left(E^\infty, D^\infty, (D^\infty.E^\infty), a_+, \frac{1}{a_-} \right) \in clconv(K). \tag{5.2}$$

I will show the necessary computations, which we did not carry out exactly as (5.34)/(5.40) in the early 70s, as we had noticed that a simple argument of convexity showed that

$$a_-I \leq A^{eff} \leq a_+I. \tag{5.3}$$

Lemma 11: If a sequence v_n converges in $L^2(\Omega; R^N)$ weak to v_∞, if $M^n \in M(\alpha, \beta; \Omega)$ is symmetric a.e. $x \in \Omega$ and $(M^n)^{-1}$ converges in $L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ weak $\star$ to $(M_-)^{-1}$, then for every nonnegative continuous function φ with compact support in Ω, one has

$$\liminf_n \int_\Omega \varphi(M^n\, v_n.v_n)\, dx \geq \int_\Omega \varphi(M_-\, v_\infty.v_\infty)\, dx, \tag{5.4}$$

i.e. if $(M^n\, v_n.v_n)$ converges to a RADON measure ν in the sense of measures (i.e. weakly $\star$), then $\nu \geq (M_-\, v_\infty.v_\infty)$ in the sense of measures in Ω.
Proof: If $\mathcal{L}_{s+}(R^N, R^N)$ denotes the convex cone of symmetric positive operators from R^N into itself, Lemma 11 follows from the fact that $(P, v) \mapsto (P^{-1}\, v.v)$ is convex on $\mathcal{L}_{s+}(R^N, R^N) \times R^N$. Indeed, one has

$$(P^{-1}\, v.v) = (P_0^{-1}\, v_0.v_0) + 2(P_0^{-1}\, v_0.v - v_0) - (P_0^{-1}(P - P_0)P_0^{-1}\, v_0.v_0) + R \tag{5.5}$$

and the remainder R is nonnegative for every $P_0 \in \mathcal{L}_{s+}(R^N, R^N)$ and $v_0 \in R^N$, as an explicit computation shows that

$$R = \Big(P(P^{-1}\, v - P_0^{-1}\, v_0).(P^{-1}\, v - P_0^{-1}\, v_0) \Big). \tag{5.6}$$

∎

As an application of Lemma 11, we can deduce upper bounds as well as lower bounds for A^{eff}, improving then the result of Theorem 6; the bounds are expressed in terms of the weak $\star$ limit of A^n and the weak $\star$ limit of $(A^n)^{-1}$.

Proposition 12: Assume that a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converges to A^{eff}, if $(A^n)^T(x) = A^n(x)$ a.e. $x \in \Omega$ for all n, and satisfies

$$A^n \rightharpoonup A_+;\ (A^n)^{-1} \rightharpoonup (A_-)^{-1} \text{ in } L^\infty \left(\Omega; \mathcal{L}(R^N, R^N) \right) \text{ weak } \star. \tag{5.7}$$

Then one has

$$A_- \leq A^{eff} \leq A_+ \text{ a.e. } x \in \Omega. \tag{5.8}$$

Proof: In the proof of Theorem 6 we have constructed a sequence $grad(u_n)$ converging in $L^2(\Omega; R^N)$ weak to $grad(u_\infty)$, and such that $A^n\, grad(u_n)$ converges in $L^2(\Omega; R^N)$ weak to $A^{eff}\, grad(u_\infty)$; moreover we had shown that $\left(A^n\, grad(u_n).grad(u_n) \right)$ converges in the sense of measures to $\left(A^{eff}\, grad(u_\infty).grad(u_\infty) \right)$ (which can be deduced from the Div-Curl lemma).

By using Lemma 11 with $M^n = A^n$ and $v_n = grad(u_n)$ one obtains $\left(A^{eff}\, grad(u_\infty).grad(u_\infty) \right) \geq \left(A_-\, grad(u_\infty).grad(u_\infty) \right)$ in the sense of measures. As both sides of the inequality belong to $L^1(\Omega)$ the inequality is valid a.e. $x \in \Omega$. From the fact that u_∞ can be any element of $H_0^1(\Omega)$, one can choose $grad(u_\infty)$ to be any constant vector λ on an open subset ω with compact closure in Ω, so that one has proved that $(A^{eff}\lambda.\lambda) \geq (A_-\lambda.\lambda)$ for every $\lambda \in R^N$, and therefore $A^{eff} \geq A_-$ a.e. in Ω.

Similarly, $\left(A^{eff}\, grad(u_\infty).grad(u_\infty) \right) \geq \left((A_+)^{-1}A^{eff}\, grad(u_\infty).A^{eff}\, grad(u_\infty) \right)$ in the sense of measures, by applying Lemma 11 with $M^n = (A^n)^{-1}$ and $v_n = A^n\, grad(u_n)$, and as both sides of the inequality belong to $L^1(\Omega)$ the inequality is valid a.e. $x \in \Omega$, and choosing $grad(u_\infty) = \lambda$ on an open subset ω with compact closure in Ω, one obtains $(A^{eff}\lambda.\lambda) \geq (A_+)^{-1}A^{eff}\lambda.A^{eff}\lambda)$ for every $\lambda \in R^N$, and therefore $A^{eff} \geq A^{eff}(A_+)^{-1}A^{eff}$, or equivalently $(A^{eff})^{-1} \geq (A_+)^{-1}$ or $A^{eff} \leq A_+$ a.e. in Ω.∎

There is an important logical point to be emphasized here, as this kind of result may easily be attributed to a few different persons. It would be interesting to check if those who either claim to have proved it before François MURAT and I had proved it in the early 70s, or claim that it had been proved a long time ago by such or such a pioneer in Continuum Mechanics or Physics, could show that there was a clear definition of what one was looking for in any of these "proofs". When one says that something is well known, it only means that it is well known by those who know it well, and in Roma in 1974, I think that Ennio DE GIORGI did not know about such an inequality; I believe that he would have quickly found a proof if I had asked him the question instead of saying that I had already proved the result. It is difficult to explain how anyone could have proved the result before there was a definition of what effective coefficients were, i.e. before the work of Sergio SPAGNOLO in the late 60s or the work of François MURAT and me in the early 70s.

Many would probably argue that they knew about effective coefficients much before there was a definition, and indeed some had a good intuition about that question, but many just had a fuzzy idea of what it was about, and I could observe that at a meeting at the Institute of Mathematics and its Applications in Minneapolis in the Fall of 1995 when one of the speakers challenged the mathematicians by saying that he had proved a result that mathematicians had not proved. Of course that could well happen and I have always been ready to learn from engineers about anything that they may know on interesting scientific matters, but it did not start well because the speaker was working with linearized Elasticity, and if there are still a few mathematicians who do not know about the defects of linearized Elasticity it is not really my fault because I have been a strong advocate of mentioning the known defects of models, and those of linearized Elasticity are well known by now, and I expected a better understanding about this kind of questions from an engineer anyway. The speaker pretended then to have proved bounds for (linear) effective elastic coefficients using inclusions of (linear) elastic materials that were not elliptic, and as he was claiming that his bounds only involved proportions I had mentioned to my neighbour and then loud enough to be heard by all that it was false;[38] indeed my comment was heard by many but after the talk there was only one person in the audience interested in clarifying the question, as John WILLIS came to tell me that the speaker had not really meant to

---

[38] For a diffusion equation, the question is very similar to using a function of one complex variable in the style of the work of David BERGMAN [Be], and extending it for small negative real values, and one cannot

say nonelliptic, because the materials that he wanted to use were actually strongly elliptic (by opposition to very strongly elliptic), and certainly an engineer who does not know the definition of ellipticity should avoid challenging mathematicians in public,[39] but that was not the only problem in the statement. The speaker had only obtained results for DIRICHLET conditions and he would have been in great trouble for showing that his results were local and could be obtained for all kinds of variational boundary conditions. Some people might argue that boundary conditions are not of such importance in Elasticity, as they may have in other questions,[40] but the problem is that the materials used in the mathematical approach do not exist in the real world, and if engineers misuse their knowledge and intuition about the real world by considering unrealistic situations and pretending that they know the mathematical answer to some questions, they may just be wrong: I do not see much reason why engineers should have a better intuition than mathematicians about problems which are completely unrealistic.[41]

According to the work of Sergio SPAGNOLO in the lates 60s or the work of François MURAT and me in the early 70s, homogenized/effective properties are "local" properties of a mixture of materials, and in this course about Optimal Design it is extremely important to use local properties and to avoid any restriction like periodic situations for example, because we are looking for the best design and we should not postulate what we would like it to be. It is a different question to consider "global" properties, like how much energy is located in a container, and some pioneers might have understood effective coefficients only in this restricted way. People who are interested in the question of how much energy a given domain contains without being interested about where this energy is located precisely and how this energy moves around,[42] often drift to quite unrealistic questions, and some still use names like Elasticity for these unrealistic questions, luring a

---

expect to use materials which are not elliptic without imposing something on the interface, as can be checked for the checkerboard pattern, according to a formula of Joseph KELLER [Ke], which George PAPANICOLAOU pointed out to me in 1980 after I had proved the same result. In the early 80s, Stefano MORTOLA and Sergio STEFFÉ had shown that using regularity of interfaces one can indeed use some "materials" with nonelliptic coefficients, but there is a relation between a measure of the regularity of the interface and the amount of nonellipticity allowed [Mo&St1].

[39] He may also have mistaken as mathematicians some people who do not hesitate to publish without correct attribution some results that they have learned from others, without realizing that it may soon become apparent that they do not understand very well the subject that they are talking about; the lack of reaction of the audience might have been a sign that many did not care much about publishing wrong results.

[40] I remember a Chemistry teacher showing us a piece of phosphorus in a container full of oil, and he explained why it was kept in oil by taking a small piece out, and it quickly burst into flames; the boundary conditions on a piece of phosphorus are important because of chemical reactions taking place precisely at the boundary.

[41] After this incident in Minneapolis, I asked my student Sergio GUTIÉRREZ to look if one could extend the theory of Homogenization in linearized Elasticity to some materials which are strongly elliptic but not very strongly elliptic. As he showed as part of his PhD thesis [Gu], a local theory englobing the very strongly elliptic materials and for which the formula of layers is valid cannot englobe any (isotropic) material which is strongly elliptic but not very strongly elliptic, except perhaps for the limiting cases. I considered that result as a fact that it is unlikely that such materials may exist, and I conjectured that the (linearized) evolution equation with a single interface with one of these materials could be ill posed. Mort GURTIN has pointed out that one can obtain some of these materials by linearization around an unstable equilibrium in (nonlinear) Elasticity, and it suggests then that it is unlikely that one could avoid DIRICHLET conditions if one uses some of these materials, except perhaps by putting enormous forces at the boundary in order to avoid these materials to become unstable. It is interesting to notice that the publication of Sergio GUTIÉRREZ's result has created a strange reaction from a referee, who thought that it was contradicting a result on Γ-convergence; I have not been able to convince the editor that this was irrelevant and proved that the referee did not understand what Homogenization is about if he/she thinks that Γ-convergence is Homogenization, that the correctness of the computations of Sergio GUTIÉRREZ is quite easy to check and that there is no reason to impose on him the burden of explaining the errors that others may have committed elsewhere in unrelated subjects.

[42] Many mathematicians still seem to believe in a world described by stationary equations, where energy

few naive mathematicians out of the scientific path.

Let us go back to deriving bounds on effective coefficients, and look at the compatibility of H-convergence with the usual preorder relation on $\mathcal{L}(R^N, R^N)$; let us recall that for $A, B \in \mathcal{L}(R^N, R^N)$, $A \leq B$ means that for every $\xi \in R^N$ one has $(A\,\xi.\xi) \leq (B\xi.\xi)$; this preorder is not an order on $\mathcal{L}(R^N, R^N)$, but it is a partial order if one restricts attention to symmetric operators.

Proposition 13: [DG&Sp] If a sequence $A^n \in M(\alpha, \beta; \Omega)$ satisfies $(A^n)^T = A^n$ for all n, a.e. $x \in \Omega$ and H-converges to A^{eff}, if a sequence u_n converges to u_∞ in $H_0^1(\Omega)$ weak and if $\varphi \geq 0$ in Ω with $\varphi \in C_c(\Omega)$, then one has

$$\liminf_n \int_\Omega \varphi \Big(A^n \, grad(u_n).grad(u_n) \Big) dx \geq \int_\Omega \varphi \Big(A^{eff} \, grad(u_\infty).grad(u_\infty) \Big) dx. \tag{5.9}$$

For every $u_\infty \in H_0^1(\Omega)$ there exists a sequence v_n converging to u_∞ in $H_0^1(\Omega)$ weak and such that for every $\varphi \in C_c(\Omega)$ one has

$$\lim_n \int_\Omega \varphi \Big(A^n \, grad(v_n).grad(v_n) \Big) dx = \int_\Omega \varphi \Big(A^{eff} \, grad(u_\infty).grad(u_\infty) \Big) dx. \tag{5.10}$$

Proof: Let $f = -div\Big(A^{eff} \, grad(u_\infty) \Big)$ and let $v_n \in H_0^1(\Omega)$ be the solution of $-div\Big(A^n \, grad(v_n) \Big) = f$ in Ω, then v_n converges to some v_∞ in $H_0^1(\Omega)$ weak and as v_∞ is solution of $-div\Big(A^{eff} \, grad(v_\infty) \Big) = f$ in Ω, one must have $v_\infty = u_\infty$ a.e. in Ω; therefore v_n converges to u_∞ in $H_0^1(\Omega)$ weak and $A^n \, grad(v_n)$ converges to $A^{eff} \, grad(u_\infty)$ in $L^2(\Omega; R^N)$ weak. Then one computes

$$\liminf_n \int_\Omega \varphi \Big(A^n \Big(grad(u_n) - grad(v_n) \Big).grad(u_n) - grad(v_n) \Big) dx, \tag{5.11}$$

which is a nonnegative number. One term is $\int_\Omega \varphi \big(A^n \, grad(u_n).grad(u_n) \big) dx$ whose $\liminf_n$ is what we are interested in; then, because of the symmetry of A^n, the other terms are $\int_\Omega \varphi \big(A^n \, grad(v_n).grad(-2u_n + v_n) \big) dx$, and the Div-Curl lemma applies so the limit is $-\int_\Omega \varphi \big(A^{eff} \, grad(u_\infty).grad(u_\infty) \big) dx$, and this gives (5.10). Of course, (5.11) is obtained by using the sequence v_n just constructed and applying the Div-Curl lemma.∎

The preceding result is due to Ennio DE GIORGI and Sergio SPAGNOLO [DG&Sp], who were using characteristic functions of measurable sets for φ, because the use of MEYERS's regularity result permits to prove the preceding result for $\varphi \geq 0$, $\varphi \in L^\infty(\Omega)$. Notice that the convexity argument of Lemma 11 gives A_- on the right side of (5.11) instead of A^{eff}, so Proposition 13 is more precise then that Lemma 11 as $A_- \leq A^{eff}$ by Proposition 12.

Proposition 14: If a sequence $A^n \in M(\alpha, \beta; \Omega)$ satisfies $(A^n)^T = A^n$ for all n, a.e. $x \in \Omega$ and H-converges to A^{eff}, if a sequence $B^n \in M(\alpha, \beta; \Omega)$ H-converges to B^{eff} and if $A^n \leq B^n$ a.e. in Ω for all n, then one has $A^{eff} \leq B^{eff}$ a.e. $x \in \Omega$.
Proof: Let $g \in H^{-1}(\Omega)$ and let $v_n \in H_0^1(\Omega)$ be the sequence of solutions of $-div\big(B^n \, grad(v_n) \big) = g$ in Ω, which converges in $H_0^1(\Omega)$ weak to v_∞, and $B^n \, grad(v_n)$ converges in $L^2(\Omega; R^N)$ weak to $B^{eff} \, grad(v_\infty)$. Then for $\varphi \geq 0$, $\varphi \in C_c(\Omega)$, one passes to the limit in the inequality

$$\int_\Omega \varphi \Big(B^n \, grad(v_n).grad(v_n) \Big) dx \geq \int_\Omega \varphi \Big(A^n \, grad(v_n).grad(v_n) \Big) dx. \tag{5.12}$$

The left side converges to $\int_\Omega \varphi \big(B^{eff} \, grad(v_\infty).grad(v_\infty) \big) dx$ by the Div-Curl lemma; the $\liminf_n$ of the right side is $\geq \int_\Omega \varphi \big(A^{eff} \, grad(v_\infty).grad(v_\infty) \big) dx$ by Proposition 13, and one obtains

$$\int_\Omega \varphi \Big(B^{eff} \, grad(v_\infty).grad(v_\infty) \Big) dx \geq \int_\Omega \varphi \Big(A^{eff} \, grad(v_\infty).grad(v_\infty) \Big) dx. \tag{5.13}$$

---

is minimized, as if they had never heard about the First Principle of "conservation of energy", and although it is a difficult question to explain how some energy is transformed into heat and how good the Second Principle is, it certainly seems utterly unrealistic to believe that after a short time every system finds its point of minimum potential energy.

As v_∞ is arbitrary in $H_0^1(\Omega)$, one obtains $\varphi\, B^{eff} \geq \varphi\, A^{eff}$ a.e. in Ω, and varying φ gives the result.∎

The second part of Proposition 13 is valid without any symmetry requirement, but the first part is not always true without symmetry. Similarly, Proposition 14 is not true for a general sequence A^n, even if all the B^n are symmetric instead. Furthermore, in Proposition 12 we compared A^{eff} with A_+, and the symmetry hypothesis on A^n is also important there. Let us construct a counter-example for $N \geq 2$ (as every operator is symmetric if $N = 1$); define A^n by

$$A^n = I + \psi_n(x_1)(e_1 \otimes e_2 - e_2 \otimes e_1), \tag{5.14}$$

where

$$\psi_n \rightharpoonup \Psi_1; \ (\psi_n)^2 \rightharpoonup \Psi_2 \text{ in } L^\infty(R) \text{ weak } \star, \tag{5.15}$$

and choose the sequence ψ_n so that $\Psi_2 > (\Psi_1)^2$. The formula for layers (4.11) shows that

$$A^n \text{ H-converges to } A^{eff} = I + \Psi_1(e_1 \otimes e_2 - e_2 \otimes e_1) + \Big(\Psi_2 - (\Psi_1)^2 \Big) e_2 \otimes e_2. \tag{5.16}$$

As $A_+ = I + \Psi_1(e_1 \otimes e_2 - e_2 \otimes e_1)$, this gives an example where $A^{eff} \leq A_+$ is false. One has $A^n \leq I$ for all n, while one does not have $A^{eff} \leq I$, and one has instead $A^{eff} \geq I$, as it must be from Proposition 14 because one has $A^n \geq I$ for all n. Finally, taking $u_n = u_\infty$ for all n, one has $\int_\Omega \varphi\big(A^n\, grad(u_n).grad(u_n)\big)\, dx = \int_\Omega \varphi| \, grad(u_\infty)|^2\, dx$ but

$$\int_\Omega \Big(A^{eff}\, grad(u_\infty).grad(u_\infty) \Big)\, dx = \int_\Omega \varphi \Big(|grad(u_\infty)|^2 + (\Psi_2 - \Psi_1^2) \Big| \frac{\partial u_\infty}{\partial x_2} \Big|^2 \Big)\, dx \tag{5.17}$$

showing that in this case (5.9) is not true.

Although in this course on Optimal Design all the problems considered are symmetric, I find useful to describe general results valid without symmetry assumptions: it is a good training in Mathematics to learn how to deal with general problems first so that one can easily deduce what to do on simpler problems, and it is usually difficult for those who have been only trained on simple special cases to understand what to do when they encounter a new situation. For that reason, I describe estimates which are useful for questions of perturbations and continuous dependence of H-limits with respect to parameters.

Lemma 15: Let $A \in M(\alpha, \beta; \Omega)$ and $D \in L^\infty\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$ with

$$||D||_{L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)} \leq \delta < \alpha, \tag{5.18}$$

then

$$A + D \in M\Big(\alpha - \delta, \frac{\alpha\, \beta - \delta^2}{\alpha - \delta}; \Omega \Big). \tag{5.19}$$

Proof: Of course $\big((A+D)\xi.\xi\big) \geq \alpha|\xi|^2 - |D\,\xi|.|\xi| \geq (\alpha - \delta)|\xi|^2$. If A and D are symmetric, one has immediately $A + D \in M(\alpha - \delta, \beta + \delta; \Omega)$, but in the general case the replacement for β requires more technical computations. One first notices that $(A\,\xi.\xi) \geq \frac{1}{\beta}|A\,\xi|^2$ means $\Big| A\,\xi - \frac{\beta}{2}\xi \Big| \leq \frac{\beta}{2}|\xi|$, and drawing a picture in a Euclidean plane containing ξ and $A\,\xi$ helps understand how to obtain the above bound and also see why it is optimal. Analytically, defining L by $2L = \frac{\alpha\,\beta - \delta^2}{\alpha - \delta}$ one wants to show that for all ξ one has $|(A + D)\xi - L\,\xi| \leq L|\xi|$, and this will be a consequence of $|A\,\xi - L\,\xi| \leq (L - \delta)|\xi|$, i.e. of $|A\,\xi|^2 - 2L(A\,\xi.\xi) \leq (-2\delta\, L + \delta^2)|\xi|^2$, and by the definition of L one has $-2\delta\, L + \delta^2 = (\beta - 2L)\alpha$; then one notices that $|A\,\xi|^2 - 2L(A\,\xi.\xi) \leq (\beta - 2L)(A\,\xi.\xi)$ which is $\leq (\beta - 2L)\alpha|\xi|^2$ as $\beta - 2L \leq 0$.∎

Proposition 16: If sequences $A^n \in M(\alpha, \beta; \Omega)$ and $B^n \in M(\alpha', \beta'; \Omega)$ H-converge respectively to A^{eff} and B^{eff}, and $|B^n - A^n|_{\mathcal{L}(R^N, R^N)} \leq \varepsilon$ a.e. $x \in \Omega$ for all n, then

$$||B^{eff} - A^{eff}||_{L^\infty(\Omega; \mathcal{L}(R^N, R^N))} \leq \varepsilon \sqrt{\frac{\beta\, \beta'}{\alpha\, \alpha'}}. \tag{5.20}$$

Proof: For $f, g \in H^{-1}(\Omega)$, one solves

$$-div\left(A^n \, grad(u_n)\right) = f \text{ in } \Omega; \ -div\left((B^n)^T \, grad(v_n)\right) = g \text{ in } \Omega, \tag{5.21}$$

so u_n, v_n converge in $H_0^1(\Omega)$ weak respectively to u_∞, v_∞, and $A^n \, grad(u_n)$ and $(B^n)^T \, grad(v_n)$ converge in $L^2(\Omega; R^N)$ weak respectively to $A^{eff} \, grad(u_\infty)$ and $(B^{eff})^T \, grad(v_\infty)$. By the Div-Curl lemma one knows that $\left(A^n \, grad(u_n).grad(v_n)\right)$ and $\left(grad(u_n).(B^n)^T \, grad(v_n)\right)$ converge in the sense of measures respectively to $\left(A^{eff} \, grad(u_\infty), grad(v_\infty)\right)$ and $\left(grad(u_\infty).(B^{eff})^T \, grad(v_\infty)\right)$, and so for every $\varphi \in C_c(\Omega)$ one has

$$\lim_n \int_\Omega \varphi\left((B^n - A^n) \, grad(u_n).grad(v_n)\right) dx = \int_\Omega \varphi\left((B^{eff} - A^{eff}) \, grad(u_\infty).grad(v_\infty)\right) dx. \tag{5.22}$$

Choosing moreover $\varphi \geq 0$, and defining

$$X = \left|\int_\Omega \varphi\left((B^{eff} - A^{eff}) \, grad(u_\infty).grad(v_\infty)\right) dx\right| \tag{5.23}$$

one deduces

$$X \leq \varepsilon \limsup_n \int_\Omega \varphi|grad(u_n)| \, |grad(v_n)| \, dx. \tag{5.24}$$

Then $|grad(u_n)| \, |grad(v_n)| \leq a \, \alpha|grad(u_n)|^2 + b \, \alpha'|grad(v_n)|^2$ when $4a \, b \, \alpha \, \alpha' \geq 1$, and as $A^n \in M(\alpha, \beta; \Omega)$ and $B^n \in M(\alpha', \beta'; \Omega)$ one has

$$X \leq \varepsilon \limsup_n \int_\Omega \varphi\left[a\left(A^n \, grad(u_n).grad(u_n)\right) + b\left(B^n \, grad(v_n).grad(v_n)\right)\right] dx, \tag{5.25}$$

which gives

$$X \leq \varepsilon \int_\Omega \varphi\left[a\left(A^{eff} \, grad(u_\infty).grad(u_\infty)\right) + b\left(B^{eff} \, grad(v_\infty).grad(v_\infty)\right)\right] dx, \tag{5.26}$$

and therefore

$$X \leq \varepsilon \, a \, \beta \int_\Omega \varphi|grad(u_\infty)|^2 \, dx + \varepsilon \, b \, \beta' \int_\Omega \varphi|grad(v_\infty)|^2 \, dx. \tag{5.27}$$

As this inequality is true for every $\varphi \geq 0$, $\varphi \in C_c(\Omega)$, one deduces

$$\left|\left((B^{eff} - A^{eff}) \, grad(u_\infty).grad(v_\infty)\right)\right| \leq \varepsilon\left(a \, \beta|grad(u_\infty)|^2 + b \, \beta'|grad(v_\infty)|^2\right) \tag{5.28}$$

a.e. in Ω, and optimizing on rationals a and b satisfying $4a \, b \, \alpha \, \alpha' \geq 1$, one obtains

$$\left|\left((B^{eff} - A^{eff}) \, grad(u_\infty).grad(v_\infty)\right)\right| \leq \varepsilon\sqrt{\frac{\beta \, \beta'}{\alpha \, \alpha'}}|grad(u_\infty)| \, |grad(v_\infty)|, \text{ in } \Omega, \tag{5.29}$$

and as u_∞ and v_∞ are arbitrary, one obtains (5.20).∎

Proposition 17: Let P be an open set of R^p. Let A^n be a sequence defined on $\Omega \times P$, such that $A^n(\cdot, p) \in M(\alpha, \beta; \Omega)$ for each $p \in P$, and such that the mappings $p \to A^n(\cdot, p)$ are of class C^k (or real analytic) from P into $L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right)$, with bounds of derivatives up to order k independent of n. Then there exists a subsequence A^m such that for every $p \in P$ the sequence $A^m(\cdot, p)$ H-converges to $A^{eff}(\cdot, p)$ and $p \to A^{eff}(\cdot, p)$ is of class C^k (or real analytic) from P into $L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right)$.

Proof: One considers a countable dense set Π of P and, using a diagonal subsequence, one extracts a subsequence A^m such that for every $p \in \Pi$ the sequence $A^m(\cdot, p)$ H-converges to a limit $A^{eff}(\cdot, p)$. Using the fact that A is uniformly continuous on compact subsets of P and Proposition 16, one deduces then that $p \to A^{eff}(\cdot, p)$ is continuous from P into $L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right)$ and that for every $p \in P$ the sequence $A^m(\cdot, p)$ H-converges to $A^{eff}(\cdot, p)$.

Defining the operators $\mathcal{A}_m(p)$ from $V = H_0^1(\Omega)$ into $V' = H^{-1}(\Omega)$ by $\mathcal{A}_m(p)v = -div\big(A^m(\cdot,p)\,grad(v)\big)$, one finds that the mappings $p \to \mathcal{A}_m(p)$ are of class C^k (or real analytic) from P to $\mathcal{L}(V,V')$ and similarly $p \to \big(\mathcal{A}_m(p)\big)^{-1}$ are of class C^k (or real analytic) from P to $\mathcal{L}(V',V)$, and finally the operators R_m defined by $R_m v = A^m(\cdot,p)\,grad(v_m)$ with v_m defined by $\mathcal{A}_m(v_m) = \mathcal{A}_{eff}(v)$ are of class C^k (or real analytic) from P into $\mathcal{L}\big(V; L^2(\Omega; R^N)\big)$; all the bounds of derivatives up to order k being independent of m so that the limit inherits of the same bounds and as $R^{eff}v = A^{eff}(\cdot,p)\,grad(v)$, one deduces therefore that $p \to A^{eff}(\cdot,p)$ is of class C^k (or real analytic) from P into $L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$.∎

As I mentioned, some of the preceding results will not be used in this course but they serve as a more general training on questions of Homogenization, but let us now come back to the method that François MURAT and I were following in the early 70s in order to obtain some information on effective coefficients: we had computed the convex hull of K defined in (5.1), we had used the relation (5.2) (where a first version of the Div-Curl lemma had been used), and we concluded that A^{eff} must satisfy (5.3); then we had noticed a quicker way to prove the result by the convexity argument of Lemma 11, but I will describe in (5.34)/(5.40) how to carry out the computations of the convex hull in a way that will be useful for further generalizations. We were considering the special case where $a^n = \chi_n \alpha + (1 - \chi_n)\beta$, with $0 < \alpha < \beta < \infty$, and χ_n being a sequence of characteristic functions converging in $L^\infty(\Omega)$ weak $\star$ to θ, so that $\theta(x)$ represent the proportion of the material of conductivity α near the point x; in that case we denote

$$\lambda_+(\theta) = \theta\,\alpha + (1-\theta)\beta; \quad \frac{1}{\lambda_-(\theta)} = \frac{\theta}{\alpha} + \frac{1-\theta}{\beta}, \tag{5.30}$$

and the analogue of Proposition 12 told us that the eigenvalues of A^{eff} were in the interval $[\lambda_-(\theta), \lambda_+(\theta)]$. In dimension $N = 2$, the eigenvalues λ_1, λ_2 of A^{eff} must satisfy

$$\lambda_-(\theta) = \frac{\alpha\,\beta}{\theta\,\beta + (1-\theta)\alpha} \le \lambda_1, \lambda_2 \le \theta\,\alpha + (1-\theta)\beta = \lambda_+(\theta), \tag{5.31}$$

defining a square in the (λ_1, λ_2) plane, and varying θ between 0 and 1, the union of these squares gave us the constraint that any mixture of two isotropic materials with conductivity α and β must satisfy the inequalities

$$\alpha \le \frac{\alpha\,\beta}{\alpha + \beta - \max\{\lambda_1, \lambda_2\}} \le \min\{\lambda_1, \lambda_2\} \le \max\{\lambda_1, \lambda_2\} \le \alpha + \beta - \frac{\alpha\,\beta}{\min\{\lambda_1, \lambda_2\}} \le \beta, \tag{5.32}$$

and we showed that the characterization (5.32) is optimal (while the characterization (5.31) is not optimal). For this we used the formula for layers, which in dimension N creates a tensor A^{eff} with one eigenvalue equal to $\lambda_-(\theta)$ with the eigenvector orthogonal to the layers, and $(N-1)$ eigenvalues equal to $\lambda_+(\theta)$ with the eigenvectors parallel to the layers, so that if $N = 2$ all the points on the boundary of the set (5.32) are obtained by layered materials. Then we took one anisotropic material with eigenvalues (γ, δ) and changing the orientation of the material gave two diagonal tensors $A_1 = (\gamma, \delta)$ and $A_2 = (\delta, \gamma)$, and using layers orthogonal to the x_1 axis, using proportion η of material A_1 and proportion $(1 - \eta)$ of material A_2 gave materials with diagonal tensors $\big(\frac{\gamma\,\delta}{\eta\,\delta + (1-\eta)\gamma}, \eta\,\delta + (1 - \eta)\gamma\big)$, having therefore determinant equal to $\gamma\,\delta$, so that by taking $\gamma = \lambda_-(\theta)$ and $\delta = \lambda_+(\theta)$ showed that the piece of equilateral hyperbola joining the two (non isotropic) diagonal corners of the square defined by (5.31) were attainable by mixtures and the union of these pieces of hyperbolas covered the set (5.32) (independently, Alain BAMBERGER had noticed that in dimension $N = 2$ if one mixes materials with $det(A) = c$ then one has $det(A^{eff}) = c$.

In doing so, we had followed the intuition that if one mixes a few materials which themselves have been obtained as mixtures of some initial materials, then the result can be obtained by mixing directly the initial materials in an adapted way. Mathematically, it is here that the metrizability property mentioned before is important: one tries to define the closure for a metrizable topology of a set containing the tensors of the form $(\chi\,\alpha + (1-\chi)\beta)I$ with χ being the characteristic function of an arbitrary measurable set (or of a smoother set like an open set, for example), and one identifies then some first generation sets contained in the sequential closure of the initial set, then one identifies some second generation sets contained in the sequential closure of some first generation sets, and one repeats the process finitely many times, and because the topology is

metrizable every set constructed is included in the sequential closure of the initial set. However, the local character of H-convergence proved in Proposition 10 has also been used: if ω_j, $j \in J$, is a countable collection of disjoint open[43] sets such that Ω is the union of all the ω_j plus a set of measure 0, and if for each j one has a sequence $A_{jn} \in M(\alpha, \beta; \omega_j)$ which H-converges to A_j^{eff}, then one can glue these pieces together, defining $A_n = \sum_j \chi_{\omega_j} A_{jn}$, which belongs to $M(\alpha, \beta; \Omega)$ and by Theorem 6 a subsequence H-converges to some A^{eff}, and Proposition 10 asserts that the restriction of A^{eff} to ω_j is A_j^{eff}, and therefore all the sequence H-converges to $A^{eff} = \sum_j \chi_{\omega_j} A_j^{eff}$. The preceding arguments enabled us to create in the desired H-closure (or G-closure, as we were working with symmetric tensors) any measurable tensor taking a constant value belonging to the set (5.32) for each of the disjoint open sets ω_j, and the conclusion came from the remark that if a sequence $A^n \in M(\alpha, \beta; \Omega)$ converges almost everywhere to A^∞ then it H-converges to A^∞.

This is the result which I used in 1974 in order to compute necessary conditions of optimality for classical solutions [Ta2]. Although François MURAT and I had followed the same type of construction that Antonio MARINO and Sergio SPAGNOLO had used in [Ma&Sp], the reason why we had been able to go further was that we had obtained the necessary condition of Proposition 12, which had given us in dimension $N = 2$ the conditions (5.31) and (5.32). We were not able at the time to obtain the characterization in dimension $N \geq 3$, or even in the case $N = 2$ we could not find the optimal characterization improving (5.31), when one imposes to use given proportions. The first step towards the solution of these more general questions was my introduction at the end of 1977 of a new method for obtaining bounds for effective coefficients [Ta6]; this method makes use of the notion of correctors in Homogenization and it requires the choice of adapted functionals for which one checks the hypotheses by applying the Compensated Compactness theory. Before describing these new ingredients, I want to show what a more precise analysis of (5.2) with the definition (5.1) gives.[44]

In the case $A^n = a^n I$ which we had investigated first, we knew that the $L^\infty(\Omega)$ weak $\star$ limits of a^n and $\frac{1}{a^n}$, denoted respectively by a_+ and $\frac{1}{a_-}$, were needed in the case where a^n only depends upon one linear combination of coordinates. We had constructed sequences $E^n = grad(u_n)$ and $D^n = a^n E^n = a^n grad(u_n)$ converging in $L^2(\Omega; R^N)$ weak, respectively to E^∞ and D^∞, and we had shown by an integration by parts (instead of the Div-Curl lemma which we discovered later) that $a^n |E^n|^2$ converges in the sense of measures to $(D^\infty . E^\infty)$. Therefore we decided to look at the convex hull in R^{2N+3} of the set K defined by (5.1), and one may wonder if it changes much to add other functions of a^n to the list and look at the convex hull of the subset $\widetilde{K}$ of R^{2N+k+1} defined by

$$\widetilde{K} = \left\{ \left(E, a\, E, a|E|^2, f_1(a), \ldots, f_k(a) \right) | E \in R^N, a \in [\alpha, \beta] \right\}, \tag{5.33}$$

and $f_1, \ldots, f_k$ are k given continuous functions on $[\alpha, \beta]$. The computations will show that the $L^\infty(\Omega)$ weak $\star$ limits of a^n and $\frac{1}{a^n}$ appear naturally. One characterization of the closed convex hull of $\widetilde{K}$ requires considering quantities of the form $(E.u_1) + (a\, E.u_2) + C_0\, a\, |E|^2 + \sum_{i=1}^k C_i f_i(a)$, where $u_1, u_2 \in R^N$ and $C_0, \ldots, C_k \in R$, in order to compute their infimum for $E \in R^N$ and $a \in [\alpha, \beta]$. One has then to consider the infimum of $(E.u_1) + (a\, E.u_2) + C_0\, a\, |E|^2$ for $E \in R^N$, and this infimum is $-\infty$ if $C_0 < 0$, or if $C_0 = 0$ and either u_1 or u_2 is not 0; therefore the typical formula is

$$\min_{E \in R^N} \left(a|E|^2 - 2(E.v) - 2(a\, E.w) \right) = -\frac{|v + a\, w|^2}{a} = -\frac{|v|^2}{a} - 2(v.w) - a\, |w|^2 \text{ for } v, w \in R^N. \tag{5.34}$$

Taking the limit of

$$a^n |E^n|^2 - 2(E^n.v) - 2(a^n\, E^n.w) \geq -\frac{|v|^2}{a^n} - 2(v.w) - a^n\, |w|^2, \tag{5.35}$$

---

[43] In his work on G-convergence, Sergio SPAGNOLO uses MEYERS's regularity theorem [Me], and he can use disjoint measurable sets.

[44] Although we could have done the following computations in the early 70s, I only noticed Lemma 18 and Lemma 19 while I was working on a set of lecture notes for my CBMS-NSF course in Santa Cruz in the Summer 1993 (I have abandoned this project since), and for preparing my lecture for a meeting in Nice in 1995, where I applied our method for the case of mixing arbitrary anisotropic materials [Ta14]; I will describe later this extension (Lemma 42), which is based on Lemma 18.

one obtains

$$(D^\infty.E^\infty) - 2(E^\infty.v) - 2(D^\infty.w) \geq -\frac{|v|^2}{a_-} - 2(v.w) - a_+|w|^2, \tag{5.36}$$

and that inequality is true for every $v, w \in R^N$. The best choice for v and w in (5.36) is obtained by solving the system

$$\frac{v}{a_-} + w = E^\infty$$
$$v + a_+ w = D^\infty, \tag{5.37}$$

and there is a difficulty if $a_- = a_+$, as one needs to have $D^\infty = a_+ E^\infty$; this is not surprising as one always has $a_- \leq a_+$, with equality if and only if a^n converges in $L^1_{loc}(\Omega)$ strong to a_+ (i.e. in $L^p_{loc}(\Omega)$ strong for every $p < \infty$, because a^n is bounded in $L^\infty(\Omega)$). If $a_- < a_+$, the solution of (5.37) is given by

$$v = \frac{a_-(a_+ E^\infty - D^\infty)}{a_+ - a_-}$$
$$w = \frac{D^\infty - a_- E^\infty}{a_+ - a_-}, \tag{5.38}$$

and (5.34) leads to

$$(a_+ - a_-)(D^\infty.E^\infty) - a_-\Big(E^\infty.(a_+ E^\infty - D^\infty)\Big) - \Big(D^\infty.(D^\infty - a_- E^\infty)\Big) \geq 0, \tag{5.39}$$

i.e.

$$(D^\infty - a_- E^\infty.D^\infty - a_+ E^\infty) \leq 0, \tag{5.40}$$

which means that D^∞ belongs to the closed ball with diameter $[a_- E^\infty, a_+ E_\infty]$, and the formula is still valid if $a_- = a_+$. Having shown already that $D^\infty = A^{eff} E^\infty$ for a symmetric matrix A^{eff}, the validity of (5.40) for every $E^\infty \in R^N$ is equivalent to A^{eff} having all its eigenvalues between a_- and a_+ (defining $M = A^{eff} - \frac{\alpha+\beta}{2} I$, the condition becomes $(M z.z) \leq \frac{\beta-\alpha}{2}|z|^2$ for all $z \in R^N$, equivalent to M having all its eigenvalues with modulus $\leq \frac{\beta-\alpha}{2}$).

A slightly different point of view for dealing with (5.40) is that if $0 < b \leq a < \infty$ and two vectors $E, D \in R^N$ satisfy $(D - a E.D - b E) \leq 0$, then there exists a symmetric B having its eigenvalues between b and a such that $D = B E$. Indeed, assuming $b < a$, let $\gamma = \frac{a+b}{2}, \delta = \frac{a-b}{2}$ and $F = \frac{1}{\delta}(D - \gamma E)$, one has $|F| \leq |E|$ and one wants to find a symmetric C of norm ≤ 1, equal to $\frac{1}{\delta}(B - \gamma I)$, such that $C E = F$, and there is actually such a C satisfying $||C|| \leq \frac{|F|}{|E|}$ if $E \neq 0$; in the case of two unit vectors e_1, e_2 which are not parallel, the basic construction for finding a symmetric contraction mapping e_1 onto e_2 is to consider the symmetry which has $e_1 \pm e_2$ as eigenvectors with eigenvalues ± 1 and if $N > 2$ the subspace orthogonal to e_1 and e_2 as eigenspace with an eigenvalue between -1 and $+1$. If $N \geq 2$, the preceding construction shows that if $E \neq 0$ and $(D - a E.D - b E) = 0$ then $D = B E$ for a symmetric B having one eigenvalue b and the $N - 1$ other eigenvalues equal to a, and such a case appears when one uses layerings.

This helps proving the following result.

Lemma 18: If $0 < \alpha' \leq b^n \leq a^n \leq \beta' < \infty$ a.e. in Ω, $E^n, D^n \in L^2(\Omega; R^N)$ for all n, and

$$(D^n - a^n E^n.D^n - b^n E^n) \leq 0 \text{ a.e. in } \Omega, \tag{5.41}$$

with

$$a^n \rightharpoonup a^\infty \text{ in } L^\infty(\Omega) \text{ weak } \star,$$
$$\frac{1}{b^n} \rightharpoonup \frac{1}{b^\infty} \text{ in } L^\infty(\Omega) \text{ weak } \star,$$
$$E^n \rightharpoonup E^\infty \text{ in } L^2(\Omega; R^N) \text{ weak}, \tag{5.42}$$
$$D^n \rightharpoonup D^\infty \text{ in } L^2(\Omega; R^N) \text{ weak},$$
$$(E^n.D^n) \rightharpoonup (E^\infty.D^\infty) \text{ in the sense of measures},$$

31

then one has

$$(D^\infty - a^\infty\, E^\infty.D^\infty - b^\infty\, E^\infty) \le 0 \text{ a.e. in } \Omega. \tag{5.43}$$

Proof: By the preceding analysis, one has $D^n = B^n\, E^n$ a.e. in Ω, with B^n symmetric and having its eigenvalues in $[b^n, a^n]$. Therefore for $v, w \in R^N$ one has

$$(D^n.E^n) - 2(E^n.v) - 2(D^n.w) = (B^n\, E^n.E^n) - 2(E^n.v + B^n\, w) \ge -\Big((B^n)^{-1}(v + B^n\, w).(v + B^n\, w)\Big)$$

$$= -\Big((B^n)^{-1}v.v\Big) - 2(v.w) - (B^n\, w.w) \ge -\frac{1}{b^n}|v|^2 - 2(v.w) - a^n|w|^2, \tag{5.44}$$

a.e. in Ω; after using test functions $\varphi \in C_c(\Omega)$ with $\varphi \ge 0$ in Ω, one obtains

$$(D^\infty.E^\infty) - 2(E^\infty.v) - 2(D^\infty.w) \ge -\frac{1}{b^\infty}|v|^2 - 2(v.w) - a^\infty|w|^2 \text{ a.e. in } \Omega, \tag{5.45}$$

and as this is the same inequality than (5.36) with a_- replaced by b^∞ and a_+ replaced by a^∞, one deduces the analogue of (5.40), which is (5.43).■

Lemma 18 can also be derived as a consequence of the following result.

Lemma 19: Define the real function Φ on $R^N \times R^N \times (0, \infty) \times (0, \infty)$ by

$$\Phi(E, D, a, b) = \begin{cases} \frac{1}{a-b}(D - a\, E.D - b\, E) \text{ if } b < a, \\ 0 \text{ if } b = a \text{ and } D = a\, E, \\ +\infty \text{ otherwise,} \end{cases} \tag{5.46}$$

then Φ is a convex function in $\big(E, D, (E.D), a, \frac{1}{b}\big)$, and more precisely

$$\Phi(E, D, a, b) = \sup_{v,w \in R^N} \Big(-(D.E) + 2(E.v) + 2(D.w) - \frac{1}{b}|v|^2 - 2(v.w) - a|w|^2\Big). \tag{5.47}$$

Proof: Indeed in the case $0 < b < a$ the quadratic form $-\frac{1}{b}|v|^2 - 2(v.w) - a|w|^2$ is negative definite, and the supremum is attained when (v, w) solves the analogue of (5.37), $\frac{v}{b} + w = E$ and $v + a\, w = D$, which gives the analogue of (5.38), $v = \frac{b(a\, E - D)}{a-b}$ and $w = \frac{D - b\, E}{a-b}$, and the value of the supremum is the quantity defined in (5.46). If $b > a > 0$ the quadratic form is not definite and the supremum is $+\infty$. If $b = a > 0$, the quantity to maximize is $-(D.E) + 2(D - a\, E.w) + 2(E.v + a\, w) - \frac{1}{a}|v + a\, w|^2$, and the supremum is $+\infty$ if $D - a\, E \ne 0$, and if $D = a\, E$ it is $-(D.E) + a|E|^2$, which is 0 in that case.■

6. Correctors in Homogenization

In 1975, I heard Ivo BABUŠKA mention the importance of amplification factors: in real elastic materials there is a threshold above which nonelastic effects (plastic deformation or fracture) usually appear. In a mixture it is not the average stress which is relevant but the local stress, and therefore one must know an amplification factor for computing the local stresses from the average stress; I did keep this comment in mind when I defined my correctors.[45] I first consider the case of layered media, for which one can prove a stronger result than for the general case of H-convergence.[46]

---

[45] I had then heard Jacques-Louis LIONS describe his computations for the periodically modulated case, which he had studied with Alain BENSOUSSAN and George PAPANICOLAOU [Be&Li&Pa], and I found that his notation with χ_{ij} created an unnecessary chaos with indices, which I decided to avoid.

[46] I did the general framework in 1975 or 1976, but I only noticed later the stronger result for the case of layered media, because of a lecture at a meeting in Luminy in the Summer 1993 [Ta13], where I considered functionals depending upon the gradient, and for the reasons mentioned in footnote 44.

Proposition 20: Let a sequence $A^n \in M(\alpha, \beta; \Omega)$ be such that it only depends upon x_1 and that it H-converges to A^{eff}. If a sequence u_n converges in $H^1_{loc}(\Omega)$ weak to u_∞ and $div\big(A^n\,grad(u_n)\big)$ stays in a compact of $H^{-1}_{loc}(\Omega)$ strong, then

$$\Big(A^n\,grad(u_n)\Big)_1 \to \Big(A^{eff}\,grad(u_\infty)\Big)_1 \text{ in } L^2_{loc}(\Omega) \text{ strong,}$$
$$\frac{\partial u_n}{\partial x_i} \to \frac{\partial u_\infty}{\partial x_i} \text{ in } L^2_{loc}(\Omega) \text{ strong, for } i = 2, \ldots, N. \tag{6.1}$$

If one defines the sequence $P_n \in L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ by

$$(P_n)_{11} = \frac{(A^{eff})_{11}}{(A^n)_{11}},$$
$$(P_n)_{1j} = \frac{(A^{eff})_{1j}}{(A^{eff})_{11}} - \frac{(A^n)_{1j}}{(A^n)_{11}} \text{ for } j = 2, \ldots, N, \tag{6.2}$$
$$(P_n)_{ij} = \delta_{ij} \text{ for } i = 2, \ldots, N, \text{ and } j = 1, \ldots, N,$$

then one has

$$grad(u_n) - P_n\,grad(u_\infty) \to 0 \text{ in } L^2_{loc}(\Omega; R^N) \text{ strong.} \tag{6.3}$$

Proof: If one denotes $E^n = grad(u_n)$ and $D^n = A^n\,grad(u_n)$ and if one uses the vector G^n and the tensor $B^n = \Phi(A^n)$ introduced after (4.11), the statement (6.1) means that G^n converges in $L^2_{loc}(\Omega; R^N)$ strong to G^∞. In order to prove this statement, one first notices that

$$\Big(B^n(G^n - G^\infty).G^n - G^\infty\Big) \text{ converges to 0 in the sense of measures.} \tag{6.4}$$

Indeed, $(B^n\,G^n.G^n) = (D^n.E^n)$ which converges in the sense of measures to $(D^\infty.E^\infty) = (B^\infty\,G^\infty.G^\infty)$ by the Div-Curl lemma, and as G^n does not oscillate in x_1 and B^n only depends upon x_1 and converges in $L^\infty\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ weak $\star$ to $B^\infty = \Phi(A^{eff})$, both $(B^n\,G^n.G^\infty)$ and $(B^n\,G^\infty.G^n)$ converge in $L^1_{loc}(\Omega)$ weak to $(B^\infty\,G^\infty.G^\infty)$. Then one notices that there exists $\gamma > 0$ such that

$$(B^n\,\lambda.\lambda) \geq \gamma|\lambda|^2 \text{ for all } \lambda \in R^N, \tag{6.5}$$

as one may take $\gamma = \frac{\alpha}{\beta^2+1}$ for example, as $(B^n\,G^n.G^n) = (A^n\,E^n.E^n) \geq \alpha|E^n|^2$ and $|G^n|^2 \leq |D^n|^2 + |E^n|^2 \leq (\beta^2+1)|E^n|^2$. Then (6.1) follows from writing

$$\frac{\partial u_n}{\partial x_1} = \frac{1}{(A^n)_{11}}\Big[\Big(A^n\,grad(u_n)\Big)_1 - \sum_{j=2}^N (A^n)_{1j}\frac{\partial u_n}{\partial x_j}\Big], \tag{6.6}$$

and using the fact that $\frac{1}{(A^n)_{11}}$ is uniformly bounded by $\frac{1}{\alpha}$.∎

I have mentioned after (4.11) that the formula for computing the effective properties of a layered material is valid under a much weaker hypothesis than the one used for the general theory of H-convergence, namely A^n bounded and $A^n_{11} \geq \alpha > 0$ a.e. in Ω for layers in the direction x_1.[47] The main reason for using LAX-MILGRAM lemma in the general framework is that it enables to construct sequences $E^n = grad(u_n)$ converging in $L^2(\omega, R^N)$ weak to a constant vector with $D^n = (A^n)^T E^n$ such that $div(D^n)$ stays in a compact of $H^{-1}_{loc}(\Omega)$; such an abstract construction is not needed in the case of layers because one can immediately write down explicitly a similar sequence; indeed taking G to be a constant vector and defining $O^n = B^n(x_1)G$, gives a vector E^n which is a gradient as $(E^n)_1$ only depends upon x_1 and E_i is constant for

---

[47] It applies to the hyperbolic equation $\frac{\partial}{\partial t}\Big(\rho_n(x)\frac{\partial u_n}{\partial t}\Big) - \frac{\partial}{\partial x}\Big(a^n(x)\frac{\partial u_n}{\partial x}\Big) = f$ if $a^n \geq \alpha$ and $\rho_n \geq \alpha$ a.e. for all n: if ρ_n converges to ρ_∞ and $\frac{1}{a^n}$ converges in $L^\infty(\Omega)$ weak $\star$ to $\frac{1}{a^{eff}}$, then the effective equation has the same form, with ρ_n and a^n replaced by ρ_∞ and a^{eff}.

$i = 2, \ldots, N$, and the vector $D^n = A^n(x_1)E^n$ has divergence 0 as $(D^n)_1$ is constant and $(D^n)_i$ only depends upon x_1 for $i = 2, \ldots, N$; the sequence E^n indeed converges in $L^2(\Omega; R^N)$ weak to a limit E^∞, and E^∞ is not constant as $(E^\infty)_1$ may indeed be a nonconstant function of x_1, but it is not so important that the limit be constant as what is necessary for the argument to work is that one can construct N such sequences with limits which are linearly independent, and this is true here.

However, the strong convergence result of Proposition 20 cannot be true without assuming an ellipticity condition: let A be a constant tensor with $A_{11} > 0$ and $(A\,\xi.\xi) = 0$ for some $\xi \neq 0$ (and $\xi \neq e_1$, as $A_{11} > 0$), then the sequence u_n defined by $u_n(x) = \frac{1}{n}\sin\big(n(\xi.x)\big)$ satisfies $div\big(A\,grad(u_n)\big) = 0$ and G^n converges in $L^2(\Omega; R^N)$ weak to 0 but it does not converge in $L^2_{loc}(\Omega; R^N)$ strong. It is therefore natural to assume that $(A\,\xi.\xi)$ does not vanish and the hypothesis $A^n \in M(\alpha, \beta; \Omega)$ appears then as a natural restriction when one wants the argument to apply for every direction of layers.

In the general framework of H-convergence, one cannot prove a result as strong as Proposition 20, and the basic result is the following.

Theorem 21: Let a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converge to A^{eff}. Then there is a subsequence A^m and an associated sequence P^m of correctors such that

$$
\begin{aligned}
&P^m \rightharpoonup I \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,} \\
&A^m\, P^m \rightharpoonup A^{eff} \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,} \\
&\mathrm{curl}(P^m\,\lambda) = 0 \text{ in } \Omega \text{ for all } \lambda \in R^N, \\
&div(A^m\, P^m\,\lambda) \text{ stays in a compact of } H^{-1}_{loc}(\Omega) \text{ strong, for all } \lambda \in R^N.
\end{aligned}
\tag{6.7}
$$

For any sequence u_m converging in $H^1_{loc}(\Omega)$ weak to u_∞ with $div\big(A^m\, grad(u_m)\big)$ staying in a compact of $H^{-1}_{loc}(\Omega)$ strong, one has

$$
grad(u_m) - P^m\, grad(u_\infty) \to 0 \text{ in } L^1_{loc}(\Omega; R^N) \text{ strong.}
\tag{6.8}
$$

Proof: For an open set Ω' of R^N containing $\overline{\Omega}$, one extends A^n by $\alpha\,I$ in $\Omega' \setminus \overline{\Omega}$ and one extracts a subsequence A^m which H-converges to a limit on Ω'; one also denotes this limit A^{eff}, and by Proposition 10 it must be an extension to Ω' of the H-limit already defined on Ω. Then for $i = 1, \ldots, N$, one chooses a function $\varphi_i \in H^1_0(\Omega')$ such that $grad(\varphi_i) = e_i$ on Ω, and one defines $P^m\, e_i = grad(v_m)$ in Ω, where $v_m \in H^1_0(\Omega')$ is the solution of $div\big(A^m\, grad(v_m) - A^{eff}\, grad(\varphi_i)\big) = 0$ in Ω'. By this construction v_m converges in $H^1_0(\Omega')$ weak to v_∞, solution of $div\big(A^{eff}\, grad(v_\infty) - A^{eff}\, grad(\varphi_i)\big) = 0$ in Ω', i.e. $v_\infty = \varphi_i$, and therefore $grad(v_m)$ and $A^m\, grad(v_m)$ converge in $L^2(\Omega'; R^N)$ weak, respectively to $grad(\varphi_i)$ and $A^{eff}\, grad(\varphi_i)$, i.e. $P^m\, e_i$ and $A^m\, P^m\, e_i$ converge in $L^2(\Omega; R^N)$ weak, respectively to e_i and $A^{eff}\, e_i$. By repeating this construction for $i = 1, \ldots, N$, one obtains a sequence P^m satisfying (6.7). Actually the construction gives P^m satisfying a more precise condition than (6.7), as one has $div(A^m\, P^m\, \lambda - A^{eff}\,\lambda) = 0$ in Ω for all $\lambda \in R^N$, but it is useful to impose only (6.7) as there may be slightly different definitions for P^n that may not satisfy this supplementary requirement.[48] The sequence $P^m\, grad(u_\infty)$ is bounded in $L^1(\Omega; R^N)$ because the sequence of correctors P^m is bounded in $L^2\big(\Omega; \mathcal{L}(R^N, R^N)\big)$. In order to prove (6.8), one chooses $g \in C(\Omega; R^N)$, $\varphi \in C_c(\Omega)$, and one computes the limit of

$$
X_m = \int_\Omega \varphi\Big(A^m(grad(u_m) - P^m\, g).grad(u_m) - P^m\, g\Big)\, dx.
\tag{6.9}
$$

Writing $g = \sum_k g_k e_k$, one expands the integrand in X_m and by (6.7) the Div-Curl lemma applies to each term: $\big(A^m\, grad(u_m).grad(u_m)\big)$ converges in the sense of measures to $\big(A^{eff}\, grad(u_\infty).grad(u_\infty)\big)$, and similarly $\big(A^m\, grad(u_m).P^m\, e_l\big)$ converges to $\big(A^{eff}\, grad(u_\infty).e_l\big)$, $\big(A^m\, P^m\, e_k.grad(u_m)\big)$ converges to

---

[48] In the case of layers for example, it is more natural to look for P^m depending only upon x_1, and the correctors defined in (6.2) satisfy $div(A^m\, P^m\,\lambda) = 0$ in Ω, even if $div(A^{eff}\,\lambda)$ does depend upon x_1. In the periodic case, it is more natural to ask for P^m to be periodic.

$\left(A^{eff} e_k.grad(u_\infty)\right)$ and $(A^m P^m e_k.P^m e_l)$ converges to $(A^{eff} e_k.e_l)$; as each g_k is continuous and φ has compact support, one deduces that

$$X_m \to X_\infty = \int_\Omega \varphi\left(A^{eff}(grad(u_\infty) - g).grad(u_\infty) - g\right) dx. \tag{6.10}$$

If $u_\infty \in C^1(\Omega)$ one can take $g = grad(u_\infty)$, and one deduces that $X_m \to 0$; by taking $0 \le \varphi \le 1$ and $\varphi = 1$ on a compact K of Ω, one deduces that $grad(u_m) - P^m grad(u_\infty) \to 0$ in $L^2(K; R^N)$ strong for every compact of Ω. If $u_\infty \in H^1(\Omega)$, one cannot use $g = grad(u_\infty)$ in general, and therefore one approaches $grad(u_\infty)$ by $g \in C(\Omega; R^N)$ in order to have

$$\|grad(u_\infty) - g\|_{L^2(\Omega; R^N)} \le \varepsilon, \tag{6.11}$$

implying

$$X_\infty \le \beta \int_\Omega |grad(u_\infty) - g|^2 dx \le \beta\,\varepsilon^2. \tag{6.12}$$

By (6.10) and (6.12) one has

$$\limsup_m \int_K \alpha|grad(u_m) - P^m g|^2 dx \le \beta\,\varepsilon^2, \tag{6.13}$$

from which one deduces

$$\limsup_m \int_K |grad(u_m) - P^m g|\, dx \le \varepsilon\sqrt{\frac{\beta\,meas(K)}{\alpha}}. \tag{6.14}$$

Using (6.11) one deduces that

$$\limsup_m \int_K |grad(u_m) - P^m grad(u_\infty)|\, dx \le \varepsilon\sqrt{\frac{\beta\,meas(K)}{\alpha}} + C\,\varepsilon, \tag{6.15}$$

where C is an upper bound for the norm of P^m in $L^2\left(\Omega; \mathcal{L}(R^N, R^N)\right)$; therefore $grad(u_m) - P^m grad(u_\infty) \to 0$ in $L^1(K; R^N)$ strong and as K is an arbitrary compact of Ω, one obtains the desired result (6.8).∎

Using better integrability property for $grad(u_\infty)$, and MEYERS's regularity theorem [Me], one can prove that $grad(u_m) - P^m grad(u_\infty)$ converges in $L^p_{loc}(\Omega; R^N)$ strong to 0 for some $p > 1$. If P^m is bounded in $L^q_{loc}\left(\Omega; \mathcal{L}(R^N, R^N)\right)$ for some $q > 2$, as can be shown using MEYERS's regularity theorem [Me], or directly as in the case of layers where one can take $q = \infty$, and $grad(u_\infty) \in L^r(\Omega; R^N)$ for some $r \ge 2$, then one can take $g \in L^s(\Omega; R^N)$ in (6.9) and (6.10) with $s = \frac{2q}{q-2}$, and if g is near $grad(u_\infty)$ in $L^r(\Omega; R^N)$ or equal to $grad(u_\infty)$ if $r \ge s$, then $P^m(grad(u_\infty) - g)$ is small in $L^t_{loc}(\Omega; R^N)$ with $t = \frac{qr}{q+r}$; one can then choose $p = \min\{s, t\}$.

Although the following results will not be used in this course, it is important to realize that correctors are important even for finding the effective equation for similar equations obtained by adding lower order terms; for simplicity, I will ignore the advantage of using MEYERS's regularity theorem [Me].

Proposition 22: Let a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converge to A^{eff}, let c^n be a sequence bounded in $L^p(\Omega; R^N)$ with $p > N$ if $N \ge 2$, $p = 2$ if $N = 1$, and assume that a sequence u_n converges in $H^1_{loc}(\Omega)$ weak to u_∞ and satisfies

$$-div\left(A^n grad(u_n)\right) + \left(c^n.grad(u_n)\right) \to f \text{ in } H^{-1}_{loc}(\Omega) \text{ strong.} \tag{6.16}$$

Then u_∞ satisfies the equation

$$-div\left(A^{eff} grad(u_\infty)\right) + \left(c^{eff}.grad(u_\infty)\right) = f \text{ in } \Omega, \tag{6.17}$$

with an effective coefficient c^{eff} such that for a subsequence

$$(P^m)^T c^m \rightharpoonup c^{eff} \text{ in } L^{2p/(p+2)} \text{ weak if } N \geq 2, \text{ in the sense of measures if } N = 1, \qquad (6.18)$$

for a sequence of correctors P^m.

Proof: One extracts a subsequence such that $(P^m)^T c^m$ converges in $L^{2p/(p+2)}(\Omega)$ weak to c^{eff} if $N \geq 2$ and in the sense of measures if N = 1.[49] Using SOBOLEV's imbedding theorem, $\left(c^m.grad(u_m)\right)$ stays in a compact of $H_{loc}^{-1}(\Omega)$ and therefore Theorem 21 implies that $grad(u_m) - P^m \, grad(u_\infty)$ converges in $L_{loc}^1(\Omega; R^N)$ strong to 0, and one wants to prove that

$$\left(c^m.grad(u_m)\right) \rightharpoonup \left(c^{eff}.grad(u_\infty)\right) \text{ in } L^{2p/(p+2)} \text{ weak if } N \geq 2, \text{ in the sense of measures if } N = 1.$$
$$(6.19)$$

Indeed $(c^m.P^m \, g)$ converges to $(c^{eff}.g)$ if $g \in C(\Omega; R^N)$, and if g satisfies (6.11), then (6.13) implies that both $(c^m.grad(u_m) - P^m \, g)$ and $(c^{eff}.grad(u_\infty) - g)$ have a small norm in L^1, so that one deduces (6.19).∎

One could add a term $d^n u_n$ in the equation, with d^n being a bounded sequence in $L^q(\Omega)$ with $q > \frac{N}{2}$ for $N \geq 2$, $q = 1$ for $N = 1$. The term $d^n u_n$ stays in a compact of $H_{loc}^{-1}(\Omega)$ strong and a subsequence converges to $d^\infty u_\infty$ if for that subsequence d^m converges in $L^q(\Omega)$ weak to d^∞ for $N \geq 2$ or in the sense of measures if $N = 1$. Without loss of generality, this term can be put into the right hand side converging in $H_{loc}^{-1}(\Omega)$ strong to a known limit.

I conclude by some computations of François MURAT giving other properties of the correctors and showing how to treat cases with terms converging only in $H_{loc}^{-1}(\Omega)$ weak.

Proposition 23: Let a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converge to A^{eff}, let b^n be a sequence bounded in $L^2(\Omega; R^N)$ and let u_n be a sequence converging in $H_{loc}^1(\Omega)$ weak to u_∞ and satisfying

$$-div\left(A^n \, grad(u_n) + b^n\right) \to f \text{ in } H_{loc}^{-1}(\Omega) \text{ strong.} \qquad (6.20)$$

Then u_∞ satisfies

$$-div\left(A^{eff} \, grad(u_\infty) + b^{eff}\right) = f \text{ in } \Omega, \qquad (6.21)$$

with an effective term $b^{eff} \in L^2(\Omega; R^N)$ such that

$$(\Pi^m)^T b^m \rightharpoonup b^{eff} \text{ in the sense of measures,} \qquad (6.22)$$

for a subsequence for which Π^m denotes the corresponding correctors associated to $(A^m)^T$.[50]

Proof: One extracts a subsequence such that (6.22) holds and

$$A^m \, grad(u_m) + b^m \text{ converges in } L^2(\Omega; R^N) \text{ weak to } \xi, \qquad (6.23)$$

and one shows that

$$\xi = A^{eff} \, grad(u_\infty) + b^{eff} \text{ in } \Omega. \qquad (6.24)$$

For $v_\infty \in C_c^1(\Omega)$, let $v_n \in H_0^1(\Omega)$ be the sequence of solutions of

$$div\left((A^m)^T \, grad(v_m) - (A^{eff})^T \, grad(v_\infty)\right) = 0 \text{ in } \Omega, \qquad (6.25)$$

so that

$$v_m \rightharpoonup v_\infty \text{ in } H_0^1(\Omega) \text{ weak,}$$
$$(A^m)^T \, grad(v_m) \rightharpoonup (A^{eff})^T \, grad(v_\infty) \text{ in } L^2(\Omega; R^N) \text{ weak,} \qquad (6.26)$$
$$grad(v_m) - \Pi^m \, grad(v_\infty) \to 0 \text{ in } L_{loc}^2(\Omega; R^N) \text{ strong,}$$

---

[49] There could be different subsequences of $(P^m)^T c^m$ converging to different limits, but Proposition 22 shows that all these limits give the same value for $\left(c^{eff}.grad(u_\infty)\right)$.

[50] From the explicit computation (6.2) of P^n in the case of layers, one sees that Π^m is not in general equal to $(P^m)^T$; of course, if $(A^n)^T = A^n$ for all n one can choose $\Pi^m = P^m$.

and one computes the limit of $\big(A^m\,grad(u_m) + b^m.grad(v_m)\big)$, which is $\big(\xi.grad(v_\infty)\big)$, by using the Div-Curl lemma. As $\big(grad(u_m).(A^m)^T\,grad(v_m)\big)$ has limit $\big(grad(u_\infty).(A^{eff})^T\,grad(v_\infty)\big)$ and as $\big(b^m.grad(v_m)\big)$ has the same limit as $\big(b^m.\Pi^m\,grad(v_\infty)\big)$, i.e. $\big(b^{eff}.grad(v_\infty)\big)$, one has proved that

$$\Big(\xi - A^{eff}\,grad(u_\infty) - b^{eff}.grad(v_\infty)\Big) = 0 \text{ in } \Omega, \tag{6.27}$$

and by density of $C_c^1(\Omega)$ in $H_0^1(\Omega)$, one obtains (6.24).∎

François MURAT also noticed that although P^m may only be bounded in $L^p\big(\Omega; \mathcal{L}(R^N, R^N)\big)$ for some $p < \infty$ (with $p \in (2, \infty)$ if one uses MEYERS's regularity theorem [Me], or $p = 2$ if one does not use it), it is nevertheless true that for $q \in [2, \infty]$ and any sequence β_n bounded in $L^q(\Omega)$ and any $i, j = 1, \ldots, N$, all the limits of subsequences of $(P^m)_{ij}\,\beta_m$ in the sense of measures actually belong to $L^q(\Omega)$. For $q = 2$ this is what (6.22) asserts for Π^m instead of P^m by taking $b^m = \beta_m e_j$. For $q > 2$, assume that $(P^m)_{ij}\,\beta_m$ converges in $L^{2q/(q+2)}(\Omega)$ weak to ξ and β_m^2 converges in $L^{q/2}(\Omega)$ weak to β_∞^2 with $\beta_\infty \in L^q(\Omega)$; then one uses the fact that $(A^m\,P^m\,\lambda.P^m\,\lambda)$ converges in the sense of measures to $(A^{eff}\,\lambda.\lambda)$ by the Div-Curl lemma, and therefore the limit of any $(P^m)_{ij}^2$ in the sense of measures belongs to $L^\infty(\Omega)$. For $\varphi \in C_c(\Omega)$ with $-1 \le \varphi \le 1$ in Ω, one has

$$\int_\Omega \varphi\,\xi\,dx = \lim_m \int_\Omega \varphi(P^m)_{ij}\,\beta_m\,dx \le \lim_m \int_\Omega |\varphi|\Big(\frac{\varepsilon(x)}{2}(P^m)_{ij}^2 + \frac{1}{2\varepsilon(x)}\beta_m^2\Big)\,dx \tag{6.28}$$

for every function $\varepsilon > 0$; therefore $\int_\Omega \varphi\,\xi\,dx \le \int_\Omega |\varphi|\Big(\frac{\varepsilon}{2}C^2 + \frac{1}{2\varepsilon}\beta_\infty^2\Big)\,dx$ and then taking the infimum in ε gives $\int_\Omega \varphi\,\xi\,dx \le \int_\Omega |\varphi|C\beta_\infty\,dx$, i.e. $|\xi| \le C\beta_\infty$ a.e. $\in \Omega$ by varying φ.

Proposition 24: Under the hypotheses of Proposition 23, one has

$$grad(u_m) - P^m\,grad(u_\infty) - r_m \to 0 \text{ in } L_{loc}^1(\Omega) \text{ strong}, \tag{6.29}$$

for some r_m (constructed explicitly) which satisfies

$$\begin{aligned} &r_m \rightharpoonup 0 \text{ in } L^2(\Omega; R^N) \text{ weak},\\ &A^m\,r_m + b^m \rightharpoonup b^{eff} \text{ in } L^2(\Omega; R^N) \text{ weak}. \end{aligned} \tag{6.30}$$

Proof: Let $\rho_n \in H_0^1(\Omega)$ be the solution of

$$div\Big(A^n\,grad(\rho_n) + b^n\Big) = 0 \text{ in } \Omega, \tag{6.31}$$

so that ρ_n is bounded in $H_0^1(\Omega)$ and by Proposition 23 a subsequence ρ_m converges in $H_0^1(\Omega)$ weak to ρ_∞, solution of

$$div\Big(A^{eff}\,grad(\rho_\infty) + b^{eff}\Big) = 0 \text{ in } \Omega. \tag{6.32}$$

Then one notices that $div\big(A^m\,grad(u_m) - A^m\,grad(\rho_m)\big) \to f$ in $H_{loc}^{-1}(\Omega)$ strong so that $grad(u_m) - grad(\rho_m) - P^m\big(grad(u_\infty) - grad(\rho_\infty)\big) \to 0$ in $L_{loc}^1(\Omega; R^N)$ strong, and therefore one has (6.29) and (6.30) by taking

$$r_m = grad(\rho_m) - P^m\,grad(\rho_\infty). \tag{6.33}$$

∎

As a corollary, if a sequence $A^n \in M(\alpha, \beta; \Omega)$ H-converges to A^{eff}, if b^n is bounded in $L^2(\Omega; R^N)$, if c^n is bounded in $L^p(\Omega; R^N)$ with $p > N$ if $N \ge 2$, $p = 2$ if $N = 1$, and if u_n is a sequence converging in $H_{loc}^1(\Omega)$ weak to u_∞ and satisfies

$$-div\Big(A^n\,grad(u_n) + b^n\Big) + \Big(c^n.grad(u_n)\Big) \to f \text{ in } H_{loc}^{-1}(\Omega) \text{ strong}, \tag{6.34}$$

then, using the definitions of c^{eff} and b^{eff} given by (6.18) and (6.22), u_∞ satisfies

$$-div\Big(A^{eff}\,grad(u_\infty) + b^{eff}\Big) + \Big(c^{eff}.grad(u_\infty)\Big) + e_\infty = f \text{ in } \Omega, \tag{6.35}$$

37

where

$$(c^n . r_n) \rightharpoonup e_\infty \text{ in } L^{2p/(p+2)}(\Omega) \text{ weak.} \tag{6.36}$$

As many seem to believe that Homogenization means periodicity, it is important to notice that in the framework of G-convergence that Sergio SPAGNOLO had developped in the late 60s or in the framework of H-convergence that François MURAT and I had developed in the early 70s, there were no conditions of periodicity. As François MURAT and I were looking at questions of Optimal Design, there was no reason for thinking that periodicity had anything to do with our problem, and when we discovered that Henri SANCHEZ-PALENCIA had been working on asymptotic methods for periodic structures [S-P1], [S-P2], it helped us understand that what we had been doing was related to effective properties of mixtures, but it was not more useful for our purpose. In the Fall of 1974, after I had described my work in Madison, Carl DE BOOR had mentioned some work by Ivo BABUŠKA; this work was restricted to some engineering applications where periodicity is natural, and when I first met Ivo BABUŠKA in the Spring 1975 [Ba], I did learn from him about some practical questions, quite unrelated to those that we were interested in our work.[51] In the Fall of 1975, at a IUTAM meeting in Marseille, I learned that Jacques-Louis LIONS had been convinced by Ivo BABUŠKA of the importance of Homogenization for periodic structures and had worked with Alain BENSOUSSAN and George PAPANICOLAOU, and I showed him my method of oscillating test functions associated with the Div-Curl lemma, and the first mention of it appears then in the article which he wrote for the proceedings [Li3]. It was only on the occasion of my lectures on our method at Bréau-sans-Nappe in the Summer 1983 [Mu&Ta1] that George PAPANICOLAOU told me that he finally understood why I had insisted so much about working without periodicity assumptions. Although I taught about general questions of Homogenization in my PECCOT lectures in the Spring 1977, many who attended these lectures but specialized in questions with periodic structures seem to have forgotten to either quote that they were using my method or that my method was not restricted to periodic situations: it might be for that reason that Olga OLEINIK rediscovered my method by considering first quasi-periodic situations and then general situations.

Some people, who seem to try to avoid mentioning either the name of Sergio SPAGNOLO for the introduction of G-convergence in the late 60s or the names of François MURAT and me for the introduction of H-convergence in the early 70s, often state that it is enough to consider periodic media; they may be unaware that such a statement is perfectly meaningless for someone who does not know that there exists a general theory; they may not realize either that for those who know about the general theory it clearly shows that they have been unable to understand the general framework. It seems that many who started by studying the special case of periodic structures have had some trouble learning about the general framework, while for all those who have started by learning the general framework, the case of periodic structures appears as the following simple exercise.

In the periodic setting, one starts with a period cell Y, generated by N linearly independent vectors $y_1, \ldots, y_N$, of R^N i.e.

$$Y = \left\{ y \mid y \in R^N, y = \sum_{i=1}^N \xi_i \, y_i, 0 \le \xi_i \le 1 \text{ for } i = 1, \ldots, N \right\}, \tag{6.37}$$

and one says that a function g defined on R^N is Y-periodic if

$$g(y + y_i) = g(y) \text{ a.e. } y \in R^N, \text{ for } i = 1, \ldots, N. \tag{6.38}$$

For $A \in M(\alpha, \beta; R^N)$ and Y-periodic, one defines A^n by

$$A^n(x) = A\left(\frac{x}{\varepsilon_n} \right) \text{ a.e. } x \in \Omega, \tag{6.39}$$

---

[51] I could imagine some real situations where our work could be useful, at least after we would have made some progress on the question of characterization of effective coefficients. I think that it was on this occasion that I learned from Ivo BABUŠKA about the importance of amplification factors for stress, but I do not recall ever hearing him mention that the defects of linearized Elasticity were quite worse for mixtures than for homogeneous materials, and I only realized that many years after.

where ε_n tends to 0.

Proposition 25: The whole sequence A^n defined by (6.39) H-converges to a constant A^{eff}, independent of the particular sequence ε_n used, and A^{eff} can be computed in the following way. For $\lambda \in R^N$, let $w_\lambda \in H^1_{loc}(R^N)$ be the Y-periodic solution (defined up to addition of a constant) of

$$div\Big(A(grad(w_\lambda) + \lambda)\Big) = 0 \text{ in } R^N, \qquad (6.40)$$

and let $P \in H^1_{loc}\big(R^N; \mathcal{L}(R^N, R^N)\big)$ be the Y-periodic function defined by

$$P\lambda = grad(w_\lambda) + \lambda \text{ a.e. in } R^N. \qquad (6.41)$$

Then

$$A^{eff}\,\lambda = \frac{1}{meas(Y)} \int_Y A(grad(w_\lambda) + \lambda)\,dy \text{ for every } \lambda \in R^N, \qquad (6.42)$$

and a sequence of corrector is defined by

$$P^n(x) = P\Big(\frac{x}{\varepsilon_n}\Big) \text{ a.e. } x \in R^N. \qquad (6.43)$$

Proof: The sequence u_n defined by

$$u_n(x) = (\lambda.x) + \varepsilon_n w_\lambda\Big(\frac{x}{\varepsilon_n}\Big), \text{ a.e. } x \in R^N, \qquad (6.44)$$

satisfies $u_n \in H^1_{loc}(R^N)$ and $div\big(A^n\,grad(u_n)\big) = 0$ in R^N. The sequence u_n converges in $H^1_{loc}(R^N)$ weak to u_∞, defined by $u_\infty(x) = (\lambda.x)$, $grad(u_n)$ is the rescaled version of $\lambda + grad(w_\lambda)$ which is Y-periodic and therefore it converges in $L^2_{loc}(R^N; R^N)$ weak to its average on the period cell Y, i.e. to $grad(u_\infty) = \lambda$, and $A^n\,grad(u_n) \in L^2_{loc}(R^N; R^N)$ is the rescaled version of $A\big(\lambda + grad(w_\lambda)\big)$ which is Y-periodic and therefore it converges in $L^2_{loc}(R^N; R^N)$ weak to its average on Y, i.e. to the value $A^{eff}\,\lambda$ as defined by (6.42); using N linearly independent $\lambda \in R^N$ characterizes the H-limit of A^n as A^{eff}.

As $grad(u_n) = P^n\,grad(u_\infty)$ a.e., and P_n satisfies the conditions (6.7), the sequence P_n gives acceptable correctors.∎

Once correctors had become natural objects for studying Homogenization, it was very natural to use them for obtaining bounds on effective coefficients.

7. Bounds on effective coefficients: second method

A first difference between this new method that I introduced at the end of 1977 in [Ta7] and the preceding one that I had used with François MURAT in the early 70s based on (5.1) and (5.2) (after having used an earlier version of the Div-Curl lemma), is that instead of considering one sequence of solutions one considers N linearly independent sequences of solutions which are the columns of the sequence of correctors P^n. A second difference is that the Div-Curl lemma had to be replaced by the more general theory of Compensated Compactness that I had just developed with François MURAT in the meantime [Mu4], [Ta4], [Ta5], [Ta6],

[Ta8].[52] If $A^n \in M(\alpha, \beta; \Omega)$ H-converges to A^{eff}, then any sequence of correctors P^m has the property that

$$P^m \rightharpoonup P^\infty = I \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,}$$
$$curl(P^m \lambda) \text{ stays in a compact of } H_{loc}^{-1}\Big(\Omega; \mathcal{L}_a(R^N, R^N)\Big) \text{ strong for all } \lambda \in R^N, \tag{7.1}$$

where $\mathcal{L}_a(R^N, R^N)$ is the space of antisymmetric matrices, and if one defines the sequence Q^m by

$$Q^m = A^m P^m, \tag{7.2}$$

then Q^m has the property that

$$Q^m \rightharpoonup Q^\infty = A^{eff} \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,}$$
$$div(Q^m \lambda) \text{ stays in a compact of } H_{loc}^{-1}(\Omega) \text{ strong for all } \lambda \in R^N. \tag{7.3}$$

Of course each column of P^m plays the role of a vector E^m and each column of Q^m plays the role of a vector D^m for which the Div-Curl lemma applies, and this means that $(Q^m)^T P^m$ converges in the sense of measures to $(Q^\infty)^T P^\infty = (A^{eff})^T$, but the Compensated Compactness theorem creates a few other interesting inequalities. While I was visiting the Mathematics Research Center in Madison in the Fall 1977, I had found a crucial additive to the Compensated Compactness theorem, as I had discovered a way to use a formal computation based on "entropies" for passing to the limit in general systems.[53] As I was wondering what I should talk about at a meeting in Versailles in December 1977, and I had thought of improving bounds on effective coefficients, it was then natural that I tried to use more general functionals, not necessarily quadratic.

Theorem 26: Assume that F is a continuous function on $\mathcal{L}(R^N, R^N) \times \mathcal{L}(R^N, R^N)$ which has the property

---

[52] Jacques-Louis LIONS had asked François MURAT to generalize our Div-Curl lemma, and he had given him an article by SCHULENBERGER and WILCOX which he thought related. François MURAT first proved a bilinear theorem: a sequence U^n converged weakly to U^∞ and satisfied a list of differential constraints, another sequence V^n converged weakly to V^∞ and satisfied another list of differential constraints, and he characterized which bilinear forms B had the property that $B(U^n, V^n)$ automatically converged in the sense of measures to $B(U^\infty, V^\infty)$. I told him that the bilinear setting looked artificial and that a quadratic setting was more natural: for a sequence U^n converging weakly to U^∞ and satisfying a list of differential constraints, he then characterized which quadratic forms Q are such that $Q(U^n)$ automatically converges in the sense of measures to $Q(U^\infty)$. While he was giving a talk about his results at the seminar that Jacques-Louis LIONS was organizing at Institut Henri POINCARÉ, it suddenly occurred to me that the right question was to look at quadratic forms Q such that if $Q(U^n)$ converges in the sense of measures to ν then one automatically has $\nu \geq Q(U^\infty)$ and before the end of the talk I had checked that the same method that we had used for the Div-Curl lemma gave me the right characterization, and I did not even need the hypothesis of constant rank that François MURAT had to impose, because of a slightly different method of proof.

[53] This is the improvement which I call the Compensated Compactness Method, on which I based my lectures at HERIOT-WATT University in the Summer 1978 [Ta8]. Of course, my framework was never restricted to hyperbolic systems, and I had already explained in [Ta5] how to use it for minimization problems, and I had described again the same example in [Ta8] in order to show that my approach based on characterizing YOUNG measures associated to a given list of differential constraints was better than the programme that others preferred of looking only at sequentially weakly lower semi-continuous functionals. Of course, "entropies" were never specific to hyperbolic situations, and before discussing the case of hyperbolic systems, I had explained how "entropies" explain the sequential weak continuity of Jacobian determinants of size larger than 2 as examples of the Compensated Compactness theorem. In [Ta6] I had advocated a different fact, that "entropy conditions" were also necessary for stationary solutions of Elasticity.

that

$$\widetilde{P}^m \rightharpoonup \widetilde{P}^\infty \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,}$$

$$\widetilde{Q}^m \rightharpoonup \widetilde{Q}^\infty \text{ in } L^2\Big(\Omega; \mathcal{L}(R^N, R^N)\Big) \text{ weak,}$$

$$curl(\widetilde{P}^m \lambda) \text{ stays in a compact of } H_{loc}^{-1}\Big(\Omega; \mathcal{L}_a(R^N, R^N)\Big) \text{ strong for all } \lambda \in R^N, \quad (7.4)$$

$$div(\widetilde{Q}^m \lambda) \text{ stays in a compact of } H_{loc}^{-1}(\Omega) \text{ strong for all } \lambda \in R^N,$$

imply

$$\liminf_{m \to \infty} \int_\Omega F(\widetilde{P}^m, \widetilde{Q}^m)\varphi \, dx \geq \int_\Omega F(\widetilde{P}^\infty, \widetilde{Q}^\infty)\varphi \, dx \text{ for all } \varphi \in C_c(\Omega), \varphi \geq 0. \quad (7.5)$$

One defines the function g on $\mathcal{L}(R^N, R^N)$, possibly taking the value $+\infty$, by

$$g(A) = \sup_{P \in \mathcal{L}(R^N, R^N)} F(P, A\,P). \quad (7.6)$$

Then if $A^n \in M(\alpha, \beta; \Omega)$ H-converges to A^{eff}, then one has

$$\liminf_{n \to \infty} \int_\Omega g(A^n)\varphi \, dx \geq \int_\Omega g(A^{eff})\varphi \, dx \text{ for all } \varphi \in C_c(\Omega), \varphi \geq 0. \quad (7.7)$$

Proof: Of course, one assumes that the left side of (7.7) is $< +\infty$, one extracts a subsequence A^m for which $\liminf_m$ is a limit and a sequence of correctors P^m exists. For $X \in C^1\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$, the sequences $\widetilde{P}^m = P^m X$ and $\widetilde{Q}^m = Q^m X$ satisfy (7.4) with $\widetilde{P}^\infty = X$ and $\widetilde{Q}^\infty = A^{eff} X$, and therefore by (7.5) one has

$$\liminf_{m \to \infty} \int_\Omega F(P^m X, A^m P^m X)\varphi \, dx \geq \int_\Omega F(X, A^{eff} X)\varphi \, dx \text{ for all } \varphi \in C_c(\Omega), \varphi \geq 0, \quad (7.8)$$

and as $F(P^m X, A^m P^m X) \leq g(A^m)$ by (7.6), one deduces that

$$\liminf_{m \to \infty} \int_\Omega g(A^m)\varphi \, dx \geq \int_\Omega F(X, A^{eff} X)\varphi \, dx \text{ for all } X \in C^1\Big(\Omega; \mathcal{L}(R^N, R^N)\Big), \quad (7.9)$$

and for all $\varphi \in C_c(\Omega)$, $\varphi \geq 0$. For $X \in L^\infty\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$, there exists a sequence $X_n \in C^1\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$ such that X_n stays bounded in $L^\infty\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$ and converges a.e. to X, and by LEBESGUE dominated convergence theorem $F(X_n, A^{eff} X_n)$ converges in $L^1(\Omega)$ strong to $F(X, A^{eff} X)$ and therefore (7.9) is true for all $X \in L^\infty\Big(\Omega; \mathcal{L}(R^N, R^N)\Big)$.

For $r < \infty$ let

$$g_r(A) = \sup_{||P|| \leq r} F(P, A\,P), \quad (7.10)$$

which is continuous, as F is uniformly continuous on bounded sets. For $\varepsilon > 0$ let M_ε be a measurable function taking only a finite number of distinct values in $\mathcal{L}(R^N, R^N)$ and such that $||M_\varepsilon - A^{eff}|| \leq \varepsilon$ a.e. in Ω. Then one can choose a measurable X_ε taking only a finite number of distinct values in $\mathcal{L}(R^N, R^N)$, such that $||X_\varepsilon|| \leq r$ and $F(X_\varepsilon, M_\varepsilon X_\varepsilon) = g_r(M_\varepsilon)$ a.e. in Ω, and as $g_r(M_\varepsilon)$ converges uniformly to $g_r(A^{eff})$ as ε tends to 0, one deduces from the inequality (7.9) for X_ε that one has

$$\liminf_{m \to \infty} \int_\Omega g(A^m)\varphi \, dx \geq \int_\Omega g_r(A^{eff})\varphi \, dx \text{ for all } r < +\infty \text{ and all } \varphi \in C_c(\Omega), \varphi \geq 0.. \quad (7.11)$$

Then $g_r(A^{eff})$ increases and converges to $g(A^{eff})$ as r increases to $+\infty$, and one deduces (7.7) by BEPPO-LEVI's theorem.∎

Of course, the (quadratic) theorem of Compensated Compactness, which I will state and prove a little later, provides an analytic characterization of all the homogeneous quadratic functions F which are such that (7.4) implies (7.5), namely it is true if and only if

$$F(\eta \otimes \xi, Q_\xi) \geq 0 \text{ for all } \eta, \xi \in R^N \text{ and all } Q_\xi \in \mathcal{L}(R^N, R^N) \text{ satisfying } Q_\xi \xi = 0. \quad (7.12)$$

41

It was only in June 1980, while I was visiting the COURANT Institute at New York University, that I tried to find which F would be suitable for the case of mixing isotropic materials, restricting myself to the case where A^{eff} would also be isotropic, i.e. equal to $a^{eff} I$, and I decided then to look for functions F satisfying (7.12) which would also be invariant under a change of orthonormal basis. Of course as a consequence of the Div-Curl lemma the functions $F_{ij}^{\pm}(P,Q) = \pm(Q P^T)_{ij} = \pm \sum_k Q_{ik} P_{jk}$ do satisfy (7.12), and therefore $F^{\pm}(P,Q) = \pm trace(Q P^T)$ give two such invariant functions F satisfying (7.12). As $trace(P^T P)$, $(trace(P))^2$, $trace(Q^T Q)$ and $(trace(Q))^2$ are invariant under a change of orthonormal basis, I checked which linear combinations of these particular functions would satisfy (7.12). It is obvious that $F_1(P) = trace(P^T P) - (trace(P))^2$ does satisfy (7.12), because if $P = \xi \otimes \eta$ then $trace(P^T P) = |\xi|^2 |\eta|^2$ and $trace(P) = (\xi.\eta)$, and therefore $trace(P^T P) \geq (trace(P))^2$ by CAUCHY-SCHWARZ's inequality. Then I found that $F_2(Q) = (N-1)trace(Q^T Q) - (trace(Q))^2$ also satisfies (7.12), by applying the following lemma to Q_ξ whose rank is at most $N-1$.

Lemma 27: If $M \in \mathcal{L}(R^N, R^N)$ then

$$rank(M) \, trace(M^T M) - (trace(M))^2 \geq 0. \tag{7.13}$$

Proof: If $rank(M) = k$, one chooses an orthogonal basis such that the range of M is spanned by the first k vectors of the basis, and then $trace(M) = \sum_i M_{ii}$ and $trace(M^T M) = \sum_{i,j} M_{ij}^2 \geq \sum_i M_{ii}^2$, which is $\geq \frac{1}{k}(\sum_i M_{ii})^2$ by CAUCHY-SCHWARZ's inequality.∎

Among the combinations of these particular functions, I quickly selected two simple ones, corresponding to the following two lemmas. In June 1980, I only computed $g(A)$ for $A = \lambda I$, but as François MURAT suggested in the Fall that the same functionals would also give an optimal result for anisotropic A, we did together the computations for general symmetric A, and I show this general computation below.

Lemma 28: If

$$F_1(P,Q) = \alpha \Big[trace(P^T P) - (trace(P))^2 \Big] - trace(Q P^T) + 2 trace(P), \tag{7.14}$$

then for $A \in \mathcal{L}(R^N, R^N)$ with $A^T = A$ and $A \geq \alpha I$, and denoting $\lambda_1, \ldots, \lambda_N$, the eigenvalues of A, one has

$$g_1(A) = \frac{\tau}{1 + \alpha\,\tau}, \text{ with } \tau = \sum_{j=1}^{N} \frac{1}{\lambda_j - \alpha}. \tag{7.15}$$

Proof: Of course, if α is an eigenvalue of A then $\tau = \infty$ and $g_1(A) = \frac{1}{\alpha}$. One chooses an orthonormal basis where A is diagonal, and the form of $F_1(P, A\,P)$ is unchanged, and one must compute

$$\sup_P \Big(\alpha \sum_{i,j=1}^N P_{ij}^2 - \alpha(\sum_{i=1}^N P_{ii})^2 - \sum_{i,j=1}^N \lambda_i P_{ij}^2 + 2 \sum_{i=1}^N P_{ii} \Big), \tag{7.16}$$

and for $i \neq j$ a good choice for P_{ij} is 0 (it does not really matter what P_{ij} is if $\lambda_i = \alpha$), and one must then compute

$$\sup_P \Big(\sum_{i=1}^N (\alpha - \lambda_i)P_{ii}^2 - \alpha(\sum_{i=1}^N P_{ii})^2 + 2 \sum_{i=1}^N P_{ii} \Big). \tag{7.17}$$

If $\sum_i P_{ii}$ is a given value t, then in the case where $\lambda_i > \alpha$ for all i, maximizing $\sum_i (\alpha - \lambda_i)P_{ii}^2$ is obtained by taking $P_{ii} = \frac{C}{\lambda_i - \alpha}$ for all i, so that $t = C\tau$, and one finds t by maximizing $-C^2\,\tau - \alpha\,t^2 + 2t$, i.e. by maximizing $-\frac{t^2}{\tau} - \alpha\,t^2 + 2t$, which gives the value of t and the maximum equal to $\frac{\tau}{1+\alpha\,\tau}$. If $\lambda_i = \alpha$ for some i, the best is to take $P_{ii} = t$ and $P_{jj} = 0$ for $j \neq i$, and then the best value of t and the maximum are equal to $\frac{1}{\alpha}$.∎

Lemma 29: If

$$F_2(P,Q) = (N-1)trace(Q^T Q) - \Big(trace(Q)\Big)^2 - \beta(N-1)trace(Q\,P^T) + 2trace(Q), \qquad (7.18)$$

then for $A \in \mathcal{L}(R^N, R^N)$ with $A^T = A$ and $A \leq \beta\,I$, and denoting $\lambda_1, \ldots, \lambda_N$ the eigenvalues of A, one has

$$g_2(A) = \frac{\sigma}{\sigma + N - 1}, \text{ with } \sigma = \sum_{j=1}^{N} \frac{\lambda_j}{\beta - \lambda_j}. \qquad (7.19)$$

Proof: Of course, if β is an eigenvalue of A then $\sigma = \infty$ and $g_2(A) = 1$. One chooses an orthonormal basis where A is diagonal, and the form of $F_2(P, A\,P)$ is unchanged, and one must compute

$$\sup_P \Big((N-1) \sum_{i,j=1}^{N} \lambda_i^2 P_{ij}^2 - (\sum_{i=1}^{N} \lambda_i\,P_{ii})^2 - \beta(N-1) \sum_{i,j=1}^{N} \lambda_i\,P_{ij}^2 + 2 \sum_{i=1}^{N} \lambda_i\,P_{ii} \Big), \qquad (7.20)$$

and for $i \neq j$ a good choice for P_{ij} is 0 (it does not really matter what P_{ij} is if $\lambda_i = \beta$), and one must then compute

$$\sup_P \Big((N-1) \sum_{i=1}^{N} (\lambda_i - \beta)\lambda_i\,P_{ii}^2 - (\sum_{i=1}^{N} \lambda_i\,P_{ii})^2 + 2 \sum_{i=1}^{N} \lambda_i\,P_{ii} \Big). \qquad (7.21)$$

If $\sum_i \lambda_i\,P_{ii}$ is a given value s, then in the case where $\lambda_i < \beta$ for all i, maximizing $\sum_i (\lambda_i - \beta)\lambda_i\,P_{ii}^2$ is obtained by taking $P_{ii} = \frac{C}{\beta - \lambda_i}$ for all i, so that $s = C\,\sigma$, and one finds s by maximizing $-(N-1)C^2\,\sigma - s^2 + 2s$, i.e. by maximizing $-\frac{N-1}{\sigma} s^2 - s^2 + 2s$, which gives the value of s and the maximum equal to $\frac{\sigma}{\sigma + N - 1}$. If $\lambda_i = \beta$ for some i, the best is to take $P_{ii} = \frac{s}{\lambda_i}$ and $P_{jj} = 0$ for $j \neq i$, and then the best value of s and the maximum are equal to 1.∎

    Of course, I had also considered more general combinations like

$$F_3(P,Q) = -trace(Q\,P^T) + a\Big[trace(P^T\,P) - \Big(trace(P)\Big)^2\Big] + b\Big[(N-1)trace(Q^T\,Q) - \Big(trace(Q)\Big)^2\Big]$$
$$+ 2c\,trace(P) + 2d\,trace(Q) \text{ with } a, b \geq 0, \qquad (7.22)$$

for which the computation of $g_3(\gamma\,I)$ requires to compute

$$\sup_P \Big[(-\gamma + a + b(N-1)\gamma^2)trace(P^T\,P) - (a + b\gamma^2)\Big(trace(P)\Big)^2 + 2(c + \gamma\,d)trace(P) \Big]. \qquad (7.23)$$

In order to have $g_3(\gamma\,I) < +\infty$, one needs to have $-\gamma + a + b(N-1)\gamma^2 \leq 0$, and one can then choose all non diagonal coefficients of P equal to 0; for $trace(P)$ given one wants to minimize $trace(P^T\,P)$, and therefore one only considers $P = p\,I$, and one wants then to maximize $\big(-\gamma + a + b(N-1)\gamma^2 - N(a + b\gamma^2)\big)p^2 + 2(c + \gamma\,d)p)$, and one obtains

$$g_3(\gamma\,I) = \frac{(c + \gamma\,d)^2}{(N-1)a + \gamma + b\,\gamma^2} \text{ if } a, b \geq 0 \text{ and } -\gamma + a + b(N-1)\gamma^2 \leq 0. \qquad (7.24)$$

As it was not so easy to handle, I had choosen the simplification of considering either $b = d = 0$, which corresponds to Lemma 28, or $a = c = 0$, which corresponds to Lemma 29. I was interested in characterizing the possible effective tensors A^{eff} of mixtures obtained by using proportion θ of an isotropic material with tensor $\alpha\,I$ and proportion $1 - \theta$ of an isotropic material with tensor $\beta\,I$, i.e. I considered $A^n = (\chi_n\,\alpha + (1 - \chi_n)\beta)I$ with a sequence of characteristic functions χ_n converging in $L^\infty(\Omega)$ weak $\star$ to θ, and A^n H-converging to A^{eff}. I already knew (5.3); in order to show explicitly the dependence in θ, (5.3) means that the eigenvalues $\lambda_1, \ldots, \lambda_N$ of A^{eff} satisfy

$$\lambda_-(\theta) \leq \lambda_j \leq \lambda_+(\theta), j = 1, \ldots, N \text{ a.e. in } \Omega, \qquad (7.25)$$

where, as in (5.3)

$$\lambda_+(\theta) = \theta\,\alpha + (1-\theta)\beta, \quad \frac{1}{\lambda_-(\theta)} = \frac{\theta}{\alpha} + \frac{1-\theta}{\beta}, \tag{7.26}$$

Theorem 26 asserts that

$$g(A^{eff}) \le \theta\,g(\alpha\,I) + (1-\theta)g(\beta\,I) \text{ a.e. in } \Omega, \tag{7.27}$$

whenever g is associated to a function F for which (7.4) implies (7.5). For the particular function g_1 given by Lemma 28, one has $g_1(\alpha\,I) = \frac{1}{\alpha}$ and $g_1(\beta\,I) = \frac{N/(\beta-\alpha)}{1+\alpha\,N/(\beta-\alpha)} = \frac{N}{(N-1)\alpha+\beta}$, and therefore (7.27) means that $\frac{\tau^{eff}}{1+\alpha\,\tau^{eff}} \le \frac{\theta}{\alpha} + \frac{(1-\theta)N}{(N-1)\alpha+\beta} = \frac{(N-\theta)\alpha+\theta\,\beta}{\alpha\big((N-1)\alpha+\beta\big)}$, which gives for τ^{eff} the upper bound

$$\tau^{eff} = \sum_{j=1}^{N} \frac{1}{\lambda_j - \alpha} \le \frac{(N-\theta)\alpha + \beta}{(1-\theta)\alpha(\beta-\alpha)}. \tag{7.28}$$

Equality occurs for the case of layers, which according to (4.11) corresponds to A^{eff} having one eigenvalue equal to $\lambda_-(\theta)$ and the $N-1$ others equal to $\lambda_+(\theta)$, i.e

$$\frac{1}{\lambda_-(\theta) - \alpha} + \frac{N-1}{\lambda_+(\theta) - \alpha} = \frac{(N-\theta)\alpha + \beta}{(1-\theta)\alpha(\beta-\alpha)}. \tag{7.29}$$

For the particular function g_2 given by Lemma 29, one has $g_2(\alpha\,I) = \frac{N\,\alpha/(\beta-\alpha)}{N\,\alpha/(\beta-\alpha)+N-1} = \frac{N\,\alpha}{\alpha+(N-1)\beta}$ and $g_2(\beta\,I) = 1$, and therefore (7.27) means that $\frac{\sigma^{eff}}{\sigma^{eff}+N-1} \le \frac{\theta\,N\,\alpha}{\alpha+(N-1)\beta} + (1-\theta) = \frac{(\theta\,N+1-\theta)\alpha+(1-\theta)(N-1)\beta}{\alpha+(N-1)\beta}$, which gives for σ^{eff} the upper bound

$$\sigma^{eff} = \sum_{j=1}^{N} \frac{\lambda_j}{\beta - \lambda_j} \le \frac{(\theta\,N + 1 - \theta)\alpha + (1-\theta)(N-1)\beta}{\theta(\beta-\alpha)}, \tag{7.30}$$

and equality occurs for the case of layers, i.e

$$\frac{1}{\beta - \lambda_-(\theta)} + \frac{N-1}{\beta - \lambda_+(\theta)} = \frac{(\theta\,N + 1 - \theta)\alpha + (1-\theta)(N-1)\beta}{\theta(\beta-\alpha)}. \tag{7.31}$$

I discuss now the basic result of Compensated Compactness theory, which has been used for Lemma 29 through the condition (7.12); Lemma 28 is more easy, and actually follows from the Div-Curl lemma, stated in (4.8)/(4.9) in Lemma 2, and mostly used in the case of gradients for which a simple proof by integration by parts has been shown, except for an application in Proposition 9. The necessity of a condition like (7.12) is easy and the general result is not even restricted to quadratic functionals [Ta8].

Proposition 30: Assume that Ω is an open subset of R^N, $A_{ijk}, i = 1, \ldots, q, j = 1, \ldots, p, k = 1, \ldots, N$, are real constants and F is a continuous real function on R^p such that, whenever

$$\begin{aligned} & U^n \rightharpoonup U^\infty \text{ in } L^\infty(\Omega; R^p) \text{ weak } \star \\ & F(U^n) \rightharpoonup V^\infty \text{ in } L^\infty(\Omega) \text{ weak } \star \\ & \sum_{j=1}^{p}\sum_{k=1}^{N} A_{ijk}\frac{\partial U_j^n}{\partial x_k} = 0 \text{ for } i = 1, \ldots, q, \end{aligned} \tag{7.32}$$

one can deduce that

$$V^\infty \ge F(U^\infty) \text{ a.e. in } \Omega. \tag{7.33}$$

Then F is Λ-convex, i.e.

$$t \mapsto F(a + t\lambda) \text{ is convex for all } a \in R^p \text{ and all } \lambda \in \Lambda, \tag{7.34}$$

44

where
$$\Lambda = \left\{ \lambda | \lambda \in R^p : \text{ there exists } \xi \in R^N \setminus 0, \sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \lambda_j \xi_k = 0 \text{ for } i = 1, \dots, q \right\}. \tag{7.35}$$

Proof: Let $\lambda \in \Lambda$ and $\xi \in R^N \setminus 0$ satisfy the condition in (7.35), then if one takes
$$U^n(x) = a + \lambda f^n\Big((\xi.x)\Big), \tag{7.36}$$
with f^n smooth, one has
$$\sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \frac{\partial U_j^n}{\partial x_k} = \Big(\sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \lambda_j \xi_k \Big)(f^n)'\Big((\xi.x)\Big) = 0. \tag{7.37}$$

One chooses a sequence χ_n of characteristic functions converging in $L^\infty(R)$ weak $\star$ to θ and a regularization $f^n = \rho_n \star \chi_n$ such that $f^n - \chi_n$ converges almost everywhere to 0, one has
$$\begin{aligned} U^n &\approx \chi_n(a + \lambda) + (1 - \chi_n)a, \ F(U^n) \approx \chi_n F(a + \lambda) + (1 - \chi_n)F(a) \\ U^\infty &= \theta(a + \lambda) + (1 - \theta)a, \ V^\infty = \theta F(a + \lambda) + (1 - \theta)F(a). \end{aligned} \tag{7.38}$$

By hypothesis, one has $V^\infty \geq F(U^\infty)$ and (7.34) follows by varying $\theta \in (0, 1)$, $a \in R^N$, and $\lambda \in \Lambda$.■

The sufficiency of a condition like (7.12) comes from applying a general result valid for quadratic functionals, which I often call the quadratic theorem of Compensated Compactness [Ta8].

Theorem 31: Let Q be a real homogeneous quadratic form on R^p which is Λ-convex, with Λ defined in (7.35), or equivalently
$$Q(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda. \tag{7.39}$$
If
$$\begin{aligned} &U^n \rightharpoonup U^\infty \text{ in } L^2_{loc}(\Omega; R^p) \text{ weak} \\ &Q(U^n) \rightharpoonup \nu \text{ in the sense of measures} \\ &\sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk} \frac{\partial U_j^n}{\partial x_k} \text{ stays in a compact of } H^{-1}_{loc}(\Omega) \text{ strong for } i = 1, \dots, q, \end{aligned} \tag{7.40}$$
then one has
$$\nu \geq Q(U^\infty) \text{ in the sense of measures.} \tag{7.41}$$

Proof: $U^n - U^\infty$ satisfies (7.40) with U^∞ replaced by 0, and ν replaced by $\nu - Q(U^\infty)$, and one may then assume that $U^\infty = 0$ with the goal of proving that $\nu \geq 0$. For $\varphi \in C^1_c(\Omega)$, let $W^n = \varphi U^n$, which is extended by 0 outside Ω and let us prove that
$$\liminf_{n \to \infty} \int_{R^N} Q(W^n) \, dx \geq 0. \tag{7.42}$$

This shows that $\langle \nu, \varphi^2 \rangle \geq 0$ for all $\varphi \in C^1_c(\Omega)$, and by density for all $\varphi \in C_c(\Omega)$, and as every nonnegative function in $C_c(\Omega)$ is a square, one deduces that $\nu \geq 0$ in the sense of measures.

If $Q(U) = \sum_{ij} q_{ij} U_i U_j$ with $q_{ij} = q_{ji}$ for all $i, j = 1, \dots, p$, I still denote Q the Hermitian extension to C^p, i.e. $Q(U) = \sum_{ij} q_{ij} U_i \overline{U_j}$, and by PLANCHEREL formula, (7.42) is equivalent to
$$\liminf_{n \to \infty} \int_{R^N} Q(\mathcal{F}W^n) \, d\xi \geq 0, \tag{7.43}$$

where $\mathcal{F}$ denotes FOURIER transform, for which I use Laurent SCHWARTZ's notations
$$\mathcal{F}W^n(\xi) = \int_{R^N} W^n(x) \, e^{-2i\pi(x.\xi)} \, dx. \tag{7.44}$$

45

One can replace Q by $\Re Q$ in (7.43), because the integral in (7.42) and therefore in (7.43) is real, and one notices that (7.39) is equivalent to

$$\Re Q(\lambda) \geq 0 \text{ for all } \lambda \in \Lambda + i\,\Lambda \subset C^p. \tag{7.45}$$

As W^n converges in $L^2(R^N; R^p)$ weak to 0 and keeps its support in a fixed compact set K of R^N, its FOURIER transform converges pointwise to 0 and is uniformly bounded and therefore by LEBESGUE dominated convergence theorem it converges in $L^2_{loc}(R^N)$ strong to 0, and the problem for proving (7.43) lies in the behaviour of $\mathcal{F}W^n$ at infinity. Information at infinity is given by the partial differential equations satisfied by W^n, and because $\sum_j \sum_k A_{ijk} \frac{\partial W_j^n}{\partial x_k}$ must converge in $H^{-1}(R^N)$ strong to 0 for $i = 1, \ldots, q$, one deduces that

$$\sum_{i=1}^{q} \int_{R^N} \frac{1}{1+|\xi|^2} \Big| \sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk}\, \mathcal{F}W_j^n(\xi)\xi_k \Big|^2 \, d\xi \to 0. \tag{7.46}$$

For $|\xi|$ large $\frac{\xi_k}{\sqrt{1+|\xi|^2}} \approx \frac{\xi_k}{|\xi|}$, and (7.46) tells that near infinity $\mathcal{F}W^n$ is near $\Lambda + i\,\Lambda$, where $\Re Q \geq 0$, and a proof of (7.43) follows from the fact that for every $\varepsilon > 0$ there exists C_ε such that

$$\Re Q(Z) \geq -\varepsilon |Z|^2 - C_\varepsilon \sum_{i=1}^{q} \Big| \sum_{j=1}^{p} \sum_{k=1}^{N} A_{ijk}\, Z_j \frac{\xi_k}{|\xi|} \Big|^2 \text{ for all } Z \in C^p, \xi \in R^N \setminus 0. \tag{7.47}$$

Applying (7.47) to $Z = \mathcal{F}W^n(\xi)$ and integrating in ξ for $|\xi| \geq 1$, gives a lower bound for $\int_{|\xi| \geq 1} \Re Q(\mathcal{F}W^n)\, d\xi$ where the coefficient of $-\varepsilon$ is bounded as W^n is bounded in $L^2(R^N; R^p)$ and the coefficient of C_ε tends to 0 by (7.46), and therefore one deduces that $\liminf_n \int_{R^N} \Re Q(\mathcal{F}W^n)\, d\xi \geq -M\,\varepsilon$, and letting ε tend to 0 proves (7.43). The inequality (7.47) is proved by contradiction: if there exists $\varepsilon_0 > 0$ and a sequence $Z^n \in C^N, \xi^n \in R^N \setminus 0$ such that $\Re Q(Z^n) < -\varepsilon_0 |Z^n|^2 - n \sum_i \Big| \sum_{jk} A_{ijk}\, Z_j^n \frac{\xi_k^n}{|\xi^n|} \Big|^2$, then after normalizing Z^n to $|Z^n| = 1$, and extracting a subsequence such that Z^n converges to Z^∞ and $\eta^n = \frac{\xi^n}{|\xi^n|}$ converges to η_∞, one finds that $\Re Q(Z^\infty) \leq -\varepsilon_0$, which contradicts the fact that (Z^∞, η^∞) satisfies the conditions of the definition of Λ in (7.35), showing that $Z^\infty \in \Lambda + i\,\Lambda$ and therefore implying $\Re Q(Z^\infty) \geq 0$ by (7.45).∎

8. Computation of effective coefficients

I computed the bounds (7.28) and (7.30) in the Fall 1980 with François MURAT, but in June 1980 in New York I had only done the corresponding computations for the case where $A^{eff} = a^{eff} I$, and having shown my new bounds to George PAPANICOLAOU he had suggested that I compare them with the HASHIN-SHTRIKMAN bounds [Ha&Sh], which I was hearing about for the first time. As for the first method for obtaining bounds for effective coefficients that I had used before with François MURAT, that second method which I had developed in [Ta7] was not restricted to symmetric tensors, and I could not understand what would replace in more general problems the particular minimization formulation that Zvi HASHIN and S. SHTRIKMAN had used, but there was an obvious gap in their "proof", and at the time I could not find a mathematical argument which could explain their computation.[54] However, the bounds which I had just found were indeed the same as the formal bounds that they had derived, and I had therefore given the first mathematical proof that the HASHIN-SHTRIKMAN bounds are indeed valid for mixtures of two isotropic materials in the case where the effective tensor is isotropic. I had not yet thought of showing that my bounds were optimal and could be attained (which I would have tried with the method of successive layerings, which was the only simple explicit construction that I knew), and as the construction of coated spheres that Zvi HASHIN and S. SHTRIKMAN had used was clear enough to me, I easily transformed it into a correct

---

[54] In their argument, Zvi HASHIN and S. SHTRIKMAN used something that did not make any (mathematical) sense at the time, and it seems now related to H-measures, which I only introduced in the late 80s for a different purpose [Ta12]; the new method for deriving bounds which I wrote in [Ta12], generalizing my earlier approach of [Ta7], has actually some analogy with the argument of Zvi HASHIN and S. SHTRIKMAN, but it is not restricted to minimization problems.

mathematical argument, but I did not try to compare with what the repeated formula for layerings would have given.

When I went back to Paris after spending the Summer at the Mathematics Research Center in Madison, I showed my computations to François MURAT an he suggested that the same functionals might also give optimal bounds for anisotropic effective tensors, and therefore we computed (7.28) and (7.30) and we tried to show that the bounds were attained by a construction of coated confocal ellipsoids. Edward FRAENKEL was visiting Paris in the Fall of 1980, and as I had mentioned to him our plan, he had given us some advice about the way to compute with ellipsoids, but we could not follow precisely what he had told us. We tried then families of general surfaces, and in order to simplify a very technical computation, we made a simplifying assumption, and that gave us exactly the case of confocal ellipsoids, but using different formulas than the ones that Edward FRAENKEL had advocated. Indeed the set of bounds (7.25), (7.28), (7.30) gave the characterization of the effective tensors of mixtures using exactly proportion θ of an isotropic material with tensor $\alpha\,I$ and proportion $1-\theta$ of an isotropic material with tensor $\beta\,I$. I will not describe the computations for confocal ellipsoids, for which I refer to [Ta9], as I will describe a simpler approach later.

I described our results at a meeting at New York University in June 1981, and they gave the missing link in the method that I had partially described in 1974 [Ta2]. As I suggested that the case of mixing more than two isotropic materials would probably be very similar with a construction like the HASHIN-SHTRIKMAN coated spheres in the isotropic case, with materials of increasing conductivity from inside out or from outside in depending upon which bound was considered, I was surprised to hear a comment by a young participant that even for three materials it was not so, and that hiding the best conductor in the middle was sometimes giving a better effective conductivity than if the best conductor was put outside. The comment was coming from a young Australian physicist, still a graduate student at the time, who has since imposed himself as the best specialist for questions of bounds of effective coefficients, Graeme MILTON.[55]

In the Spring 1982, I gave an introductory course to Homogenization at Ecole Polytechnique, and as I thought that our construction with confocal ellipsoids could not be avoided, I had asked two students, Philippe BRAIDY and Didier POUILLOUX, to make a numerical study comparing the materials that we had constructed by using confocal ellipsoids and those that could be constructed by successive layerings, which was the method that we had used for the results quoted in [Ta2]. Contrary to my mistaken expectations, they reported that the numerical computations showed that the two sets were the same, and a few days after they had a proof of it, using N layerings in orthogonal directions, where in each layering the direction orthogonal to the layers is a common eigenvector for the two materials being mixed [Br&Po]. I immediately checked that the repeated layering construction has the same restricted "generalized BERGMAN function" than our construction with confocal ellipsoids. For example, in my interpretation of the construction with coated spheres of Zvi HASHIN and S. SHTRIKMAN in June 1980, for a parameter $\theta \in [0,1]$ I used a sequence of VITALI coverings with smaller and smaller coated spheres, all showing the same proportion of volume between the inside spheres and the outside spherical coats; the geometry being given I imagined all the interior spheres filled with an isotropic material with tensor $\alpha\,I$ and all the outside spherical coats filled with an isotropic material with tensor $\beta\,I$, and then I showed that the sequence A^n defined in this way H-converges to $\Phi_0(\alpha,\beta)\,I$, where $\Phi_0(\alpha,\beta)$ is one of the two HASHIN-SHTRIKMAN bounds, corresponding to equality in (7.28).[56] This is exactly the type of functions used by David BERGMAN when one mixes two

---

[55] In the early 90s, while I visited Graeme MILTON in New York, he gave me a physical explanation of why it is sometime good to hide the best conductor available; he argued that if one has a spherical core of a very poor conductor, the electric current tries to avoid it and that creates a high concentration of field lines near the surface of the poor conductor and therefore it is there that the best conductor is more useful.

[56] My proof relied on the fact that, despite the huge arbitrariness in the choice of the VITALI coverings, for any $\lambda \in R^N$ I could write explicit solutions of $div\big(A^n\,grad(u_n)\big) = 0$ for which $grad(u_n)$ converges weakly to λ and I could compute the limit of $A^n\,grad(u_n)$. My construction was local, and I followed the computation that I had just read in the article of Zvi HASHIN and S. SHTRIKMAN [Ha&Sh], i.e. the explicit solution of $div\big(A\,grad(v)\big) = 0$ in a coated sphere domain with affine boundary conditions on the outside coat; this computation is a by-product of the formula for the change in electric field created by an isotropic spherical conducting inclusion in an infinite isotropic medium, a classical formula for physicists who associate it with

47

isotropic conductors and one expects the resulting effective material to be isotropic whatever the ratio of the two conductivities are, and therefore one has $\Phi_0(\alpha, \beta) = \alpha\, F\!\left(\frac{\beta}{\alpha}\right)$. David BERGMAN made the important observation that F extends to the complex upper half plane into a holomorphic function satisfying $\Im\big(F(z)\big) \geq 0$ [Be].[57] What I call a "generalized BERGMAN function" is a similar situation where a sequence of geometries is given for mixing r materials with proportions $\theta_1, \ldots, \theta_r$, and if the r materials used have tensors $M_1, \ldots, M_r$, then the resulting effective tensor is $\Phi(M_1, \ldots, M_r)$. One assumes that for $j = 1, \ldots, r$, $M_j \in M(\alpha_j, \beta_j; \Omega)$, and that one has r sequences of characteristic functions of measurable sets from a partition χ_j^n, $j = 1, \ldots, r$, satisfying $\chi_j^n \chi_k^n = 0$ for $j \neq k$, $\sum_j \chi_j^n = 1$ a.e. in Ω, and one assumes that χ_j^n converges in $L^\infty(\Omega)$ weak $\star$ to θ_j for $j = 1, \ldots, r$; then one uses Proposition 17 in order to show that there is a subsequence for which for all such $M_1, \ldots, M_r$, $\sum_j \chi_j^m M_j$ H-converges to an element $\Phi(M_1, \ldots, M_r)$ of $M(\alpha, \beta; \Omega)$ with $0 < \alpha = \min\{\alpha_1, \ldots, \alpha_r\} \leq \beta = \max\{\beta_1, \ldots, \beta_r\} < \infty$. I use the qualificative restricted for expressing the fact that one restricts attention to a special class of $M_1, \ldots, M_r$, for example isotropic tensors $m_1\, I, \ldots, m_r\, I$,[58] and I do not know how to compute the generalized BERGMAN function for a geometry of coated spheres or confocal ellipsoids, and it may be dependent of other properties of the VITALI covering used.[59] In their computation, Philippe BRAIDY and Didier POUILLOUX had used the same method that Antonio MARINO and Sergio SPAGNOLO or François MURAT and I had used in the early 70s, and the reiteration of the layering formula was simple enough because at each step the direction orthogonal to the layers was a common eigenvector of both (symmetric) tensors which were mixed, and actually the two tensors had a common basis of eigenvectors. During the Spring 1983, while I was visiting the Mathematical Sciences Research Institute in Berkeley, I tried to compute the formula for mixing arbitrary materials in arbitrary directions, having in mind to reiterate the procedure. I wanted to rewrite formula (4.11) in a more intrinsic way, and I could easily deduce what the formula (4.11) would become if I used layers orthogonal to a vector e, i.e. A^n depending only upon $(x.e)$, but that did not change much, and it was a different idea that simplified the computation. Using layers orthogonal to e for mixing two materials with tensors A and B, with respective proportions θ and $1 - \theta$, the simplification came by considering θ small, and because the formula appeared to have the form $B + \theta\, F(A, B, e) + o(\theta)$, it suggested to write a differential equation $B' = F(A, B, e)$ and integrate it. In other terms, for e fixed, increasing the proportion of A from 0 to 1 creates a curve going from B to A in the space of matrices, and I first computed that curve by considering it as the trajectory of a differential equation, which was easy to write down. One can first rewrite formula

---

various names, some as famous as MAXWELL, but it must have been known to GAUSS and to DIRICHLET, who seems to be credited for a similar formula for an ellipsoid (and he may therefore have known the formulas that François MURAT and I (re)discovered for our construction with coated ellipsoids).

[57] The idea may have been used before, and I think that I had heard such an idea attributed to PRAGER. In dimension $N = 2$, the function F also satisfies the relation $F(z) F\!\left(\frac{1}{z}\right) = 1$ by an argument of Joseph KELLER [Ke], and in the early 80s Graeme MILTON showed me that all such functions can be obtained.

[58] Ken GOLDEN and George PAPANICOLAOU have studied functions of r complex variables F appearing when one imposes the restriction $\Phi(m_1\, I, \ldots, m_r\, I) = F(m_1, \ldots, m_r) I$ for all $m_1, \ldots, m_r > 0$.

[59] The computations of Zvi HASHIN and S. SHTRIKMAN for coated spheres and diffusion equation consists in looking for solutions of the form $x_j f(r)$ and one finds that f must satisfy a differential equation; they also used the same construction of coated spheres for linearized Elasticity with isotropic materials, and they could compute the effective bulk modulus because it corresponds to applying a uniform pressure and the displacement has the form $x\, g(r)$, and one finds that g must satisfy a differential equation. I do not know how to compute the effective shear modulus for the geometry of coated spheres, and it may depend upon which VITALI covering is used (if I understood correctly what Graeme MILTON told me a few years ago, he knew that it does depend upon the covering). Gilles FRANCFORT and François MURAT have computed in [Fr&Mu] the complete effective elasticity tensors of mixtures, but following the method of multiple layerings, adapting the extension that I had given in [Ta9] of the computation of Philippe BRAIDY and Didier POUILLOUX.

(4.11) for layers orthogonal to e.

$$\frac{1}{(A^n\,e.e)} \rightharpoonup \frac{1}{(A^{eff}\,e.e)} \text{ in } L^\infty(\Omega) \text{ weak } \star$$

$$\frac{(A^n\,f.e)}{(A^n\,e.e)} \rightharpoonup \frac{(A^{eff}\,f.e)}{(A^{eff}\,e.e)} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for every } f \perp e$$

$$\frac{(A^n\,e.g)}{(A^n\,e.e)} \rightharpoonup \frac{(A^{eff}\,e.g)}{(A^{eff}\,e.e)} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for every } g \perp e$$

$$(A^n\,f.g) - \frac{(A^n\,f.e)(A^n\,e.g)}{(A^n\,e.e)} \rightharpoonup (A^{eff}\,f.g) - \frac{(A^{eff}\,f.e)(A^{eff}\,e.g)}{(A^{eff}\,e.e)} \text{ in } L^\infty(\Omega) \text{ weak } \star \text{ for every } f \perp e, g \perp e,$$

$$(8.1)$$

where I have used the Euclidean structure of R^N.[60] Mixing A with a small proportion θ and B with proportion $1 - \theta$ in layers orthogonal to e gives then

$$\frac{1}{(A^{eff}\,e.e)} = \frac{1-\theta}{(B\,e.e)} + \frac{\theta}{(A\,e.e)}$$

$$(A^{eff}\,e.e) = (B\,e.e) + \theta\Big[(B\,e.e) - \frac{(B\,e.e)^2}{(A\,e.e)}\Big] + o(\theta),$$

$$(8.2)$$

and then for f and g orthogonal to e

$$\frac{(A^{eff}\,f.e)}{(A^{eff}\,e.e)} = \frac{(1-\theta)(B\,f.e)}{(B\,e.e)} + \frac{\theta(A\,f.e)}{(A\,e.e)}$$

$$(A^{eff}\,f.e) = (B\,f.e) + \theta\Big[\frac{(A\,f.e)(B\,e.e)}{(A\,e.e)} - \frac{(B\,f.e)(B\,e.e)}{(A\,e.e)}\Big] + o(\theta)$$

$$(8.3)$$

$$\frac{(A^{eff}\,e.g)}{(A^{eff}\,e.e)} = \frac{(1-\theta)(B\,e.g)}{(B\,e.e)} + \frac{\theta(A\,e.g)}{(A\,e.e)}$$

$$(A^{eff}\,e.g) = (B\,e.g) + \theta\Big[\frac{(A\,e.g)(B\,e.e)}{(A\,e.e)} - \frac{(B\,e.g)(B\,e.e)}{(A\,e.e)}\Big] + o(\theta)$$

$$(8.4)$$

$$(A^{eff}\,f.g) - \frac{(A^{eff}\,f.e)(A^{eff}\,e.g)}{(A^{eff}\,e.e)} = (1-\theta)(B\,f.g) - \frac{(1-\theta)(B\,f.e)(B\,e.g)}{(B\,e.e)} + \theta(A\,f.g) - \frac{\theta(A\,f.e)(A\,e.g)}{(A\,e.e)}$$

$$(A^{eff}\,f.g) = (B\,f.g) + \theta\Big[(A\,f.g) - (B\,f.g) - \frac{\big((B\,f.e) - (A\,f.e)\big)\big((B\,e.g) - (A\,e.g)\big)}{(A\,e.e)}\Big] + o(\theta).$$

$$(8.5)$$

The form of (8.5) suggests that one has

$$A^{eff} = B + \theta\Big[A - B - (B - A)\frac{e \otimes e}{(A\,e.e)}(B - A)\Big] + o(\theta),$$

$$(8.6)$$

and indeed this is compatible with (8.2)/(8.4). When e and A are given, formula (8.6) corresponds to a differential equation

$$M' = A - M - (M - A)\frac{e \otimes e}{(A\,e.e)}(M - A).$$

$$(8.7)$$

The integral curve corresponds to the formula for layering with a material with tensor A, with layers orthogonal to e. Formula (8.1), which corresponds to the first lines of (8.2)/(8.5) when one mixes two materials with tensors A and B says that the integral curves become straight lines if one performs the change of variable

---

[60] It can be avoided by denoting $\mathcal{E}$ the ambient vector space, taking e as an element of the dual $\mathcal{E}'$, and considering the tensors A, B, as elements of $\mathcal{L}(\mathcal{E}', \mathcal{E})$, as well as $e \otimes e$ which appears in some formulas.

$A \mapsto \left(\frac{1}{(A\,e.e)}, \frac{(A\,f.e)}{(A\,e.e)}, \frac{(A\,e.g)}{(A\,e.e)}, (A\,f.g) - \frac{(A\,f.e)(A\,e.g)}{(A\,e.e)}\right)$ when f, g span the subspace orthogonal to e. In the early 70s, we knew that if one does not pay attention to the proportions used of various materials, formula (4.11) means that the set of effective tensors has the property that all its images by maps like the one mentioned above are automatically convex. This condition gives a geometric characterization of the sets that one cannot enlarge by layering, at least for the case where one is not allowed to rotate the materials used, in the case where one starts with some anisotropic materials; in realistic problems, one must also allow for rotations of the materials used, i.e. the set must be stable by mappings $A \mapsto P^T A P$ for $P \in SO(N)$, with $N = 2$ or $N = 3$ usually.

Assuming that $M - A$ is invertible, (8.7) can be written as

$$[(M-A)^{-1}]' = -(M-A)^{-1}A'(M-A)^{-1} = (M-A)^{-1} + \frac{e \otimes e}{(A\,e.e)}, \tag{8.8}$$

whch is a linear equation in $(M-A)^{-1}$. Using τ as variable, and assuming that $\tau = 0$ corresponds to B, the solution of (8.8) is

$$(M-A)^{-1} = -\frac{e \otimes e}{(A\,e.e)} + e^{\tau}\left((B-A)^{-1} + \frac{e \otimes e}{(A\,e.e)}\right), \tag{8.9}$$

and if M corresponds to having used proportion $\eta(\tau)$ of A and $1 - \eta(\tau)$ of B, then for θ small $\eta(\tau + \theta) = \theta + (1-\theta)\eta(\tau) + o(\theta)$ gives $\eta' = 1 - \eta$ and therefore $\eta = 1 - e^{-\tau}$ or equivalently $e^{\tau} = \frac{1}{1-\eta}$ for proportion η of A, giving

$$(M-A)^{-1} = \frac{(B-A)^{-1}}{1-\eta} + \frac{\eta}{1-\eta}\frac{e \otimes e}{(A\,e.e)} \text{ for proportion } \eta \text{ of } A. \tag{8.10}$$

If $(B-A)z = 0$ for a nonzero vector z, then (8.7) shows that $(M-A)z = 0$, and in this case one must reinterpret (8.10). Of course, exchanging the role of A and B and changing η into $1 - \eta$, (8.10) is replaced by

$$(M-B)^{-1} = \frac{(A-B)^{-1}}{\eta} + \frac{1-\eta}{\eta}\frac{e \otimes e}{(B\,e.e)} \text{ for proportion } \eta \text{ of } A. \tag{8.11}$$

With formula (8.10) at hand I could easily reiterate the layering process with various directions of layers, with the condition that each layering uses the material with tensor A, and it gave the following generalization of the formula which had been obtained by Philippe BRAIDY and Didier POUILLOUX in the special case where A and B have a common basis of eigenvectors and each e is one of these common eigenvectors.

Proposition 32: For $\eta \in (0,1)$, let $\xi_1, \ldots, \xi_p$ be p positive numbers with $\sum_j \xi_j = 1 - \eta$, let $e_1, \ldots, e_p$ be p nonzero vectors of R^N, then using proportion η of material with tensor A and proportion $1 - \eta$ of material with tensor B, one can construct by multiple layerings the material with tensor M such that

$$(M-B)^{-1} = \frac{(A-B)^{-1}}{\eta} + \frac{1}{\eta}\left(\sum_{j=1}^{p} \xi_j \frac{e_j \otimes e_j}{(B\,e_j.e_j)}\right). \tag{8.12}$$

Proof: Of course, one assumes that $B - A$ is invertible, as the formula must be reinterpreted if $B - A$ is not invertible. One starts from $M_0 = A$ and by induction one constructs M_j by layering M_{j-1} and B in proportions η_j and $1 - \eta_j$, with layers orthogonal to e_j. Formula (8.11) gives

$$(M_j - B)^{-1} = \frac{(M_{j-1} - B)^{-1}}{\eta_j} + \frac{1-\eta_j}{\eta_j}\frac{e \otimes e}{(B\,e.e)} \text{ for } j = 1, \ldots, p, \tag{8.13}$$

which is adapted to reiteration and provides (8.12) with

$$\begin{aligned}
\eta &= \eta_1 \cdots \eta_p \\
\xi_1 &= 1 - \eta_1, \xi_j = \eta_1 \cdots \eta_{j-1}(1 - \eta_j) \text{ for } j = 1, \ldots, p,
\end{aligned} \tag{8.14}$$

which gives $\xi_1 + \ldots + \xi_j = 1 - \eta_1 \cdots \eta_j$ for $j = 1, \ldots, p$, and this defines in a unique way η_j for $j = 1, \ldots, p$.∎

The preceding computations did not require any symmetry assumption for A or B. The characterization of the sum $\sum_j \xi_j \frac{e_j \otimes e_j}{(B e_j . e_j)}$ for all $\xi_j > 0$ with sum $1 - \eta$ and all nonzero vectors e_j depends only on the symmetric part of B (and of η).

Lemma 33: If B is symmetric positive definite then for $\xi_1, \ldots, \xi_p > 0$ and nonzero vectors $e_1, \ldots, e_p$, one has

$$\sum_{j=1}^{p} \xi_j \frac{e_j \otimes e_j}{(B e_j . e_j)} = B^{-1/2} K B^{-1/2}, \text{ with } K \text{ symmetric nonnegative and } trace(K) = \sum_{j=1}^{p} \xi_j, \qquad (8.15)$$

and conversely any such K can be obtained in this way.

Proof: Putting $e_j = B^{-1/2} f_j$ for $j = 1, \ldots, p$, one has $K = \sum_j \xi_j \frac{f_j \otimes f_j}{|f_j|^2}$, and each $\frac{f_j \otimes f_j}{|f_j|^2}$ is a nonnegative symmetric tensor with trace 1, and (8.15) follows. Conversely if K is a symmetric nonnegative tensor with trace equal to S, then there is an orthonormal basis of eigenvectors $f_1, \ldots, f_N$, with $K f_j = \kappa_j f_j$ and $\kappa_j \geq 0$ for $j = 1, \ldots, N$, and $\sum_j \kappa_j = S$, so that $K = \sum_j \kappa_j f_j \otimes f_j$. ∎

Using Proposition 32 and Lemma 33, with $A = \alpha I$ and $B = \beta I$, one can construct materials with a symmetric tensor M with eigenvalues $\lambda_1, \ldots, \lambda_N$, and (8.12) and Lemma 33 mean that

$$\frac{1}{\lambda_j - \beta} \geq \frac{1}{\eta(\alpha - \beta)} \text{ for } j = 1, \ldots, N$$
$$\sum_{j=1}^{N} \frac{1}{\lambda_j - \beta} = \frac{N}{\eta(\alpha - \beta)} + \frac{1 - \eta}{\eta \beta}, \qquad (8.16)$$

i.e. $\lambda_j \leq \lambda_+(\eta)$ for $j = 1, \ldots, N$, and equality in (7.30), which implies $\lambda_j \geq \lambda_-(\eta)$ for $j = 1, \ldots, N$, because of (7.31). Exchanging the roles of A and B one can obtain another part of the boundary of possible effective tensors with equality in (7.28), and filling the interior of the set is then easy.

After I had mentioned these new results to Robert KOHN, who was also visiting MSRI at the time, he wondered if one could find a more direct proof, and I therefore proved again the formulas (8.10)/(8.11) directly.

Lemma 34: Mixing materials with tensor A and B with respective proportions η and $1 - \eta$ in layers orthogonal to e gives an effective tensor A^{eff} given by

$$A^{eff} = \eta A + (1 - \eta)B - \eta(1 - \eta)(B - A)\frac{e \otimes e}{(1 - \eta)(A e . e) + \eta(B e . e)}(B - A). \qquad (8.17)$$

Proof: One considers a sequence of characteristic functions χ_n converging in $L^\infty(R)$ weak $\star$ to η and depending only upon $(x.e)$, and one chooses $A^n = \chi_n A + (1 - \chi_n)B$. For an arbitrary vector $E^\infty \in R^N$, one constructs a sequence $E^n = grad(u_n)$ converging in $L^2_{loc}(R^N; R^N)$ weak to E^∞, depending only upon $(x.e)$ and satisfying $div(A^n grad(u_n)) = 0$, and one computes the limit in $L^2_{loc}(R^N; R^N)$ weak of $D^n = A^n grad(u_n)$, which will be $D^\infty = A^{eff} E^\infty$, with A^{eff} given by (8.17).

One looks for $E_A, E_B \in R^N$ such that one can take

$$E^n = \chi_n E_A + (1 - \chi_n)E_B$$
$$D^n = \chi_n A E_A + (1 - \chi_n)B E_B \qquad (8.18)$$
$$\eta E_A + (1 - \eta)E_B = E^\infty,$$

and the constraints $curl(E^n) = div(D^n) = 0$ become

$$E_B - E_A = c e$$
$$(B E_B - A E_A . e) = 0, \qquad (8.19)$$

51

and then one should have
$$\eta\, A\, E_A + (1-\eta)B\, E_B = A^{eff}\, E^\infty. \tag{8.20}$$

One chooses then
$$E_A = E^\infty + c_A\, e;\ E_B = E^\infty + c_B\, e;\ \eta\, c_A + (1-\eta)c_B = 0, \tag{8.21}$$

and (8.19) requires that
$$\Big((B-A)E^\infty.e\Big) + c_B(B\, e.e) - c_A(A\, e.e) = 0, \tag{8.22}$$

and (8.21)/(8.22) give
$$\begin{aligned}
\Big((1-\eta)(A\, e.e) + \eta(B\, e.e)\Big)c_A &= (1-\eta)\Big((B-A)E^\infty.e\Big)\\
\Big((1-\eta)(A\, e.e) + \eta(B\, e.e)\Big)c_B &= -\eta\Big((B-A)E^\infty.e\Big),
\end{aligned} \tag{8.23}$$

and therefore (8.20) becomes
$$A^{eff}\, E^\infty = (\eta\, A + (1-\eta)B)E^\infty + \frac{\Big((B-A)E^\infty.e\Big)}{(1-\eta)(A\, e.e) + \eta(B\, e.e)}\Big(\eta(1-\eta)A\, e - \eta(1-\eta)B\, e\Big), \tag{8.24}$$

and as (8.24) is true for every $E^\infty \in R^N$, one deduces formula (8.17) for A^{eff}.■

One deduces then (8.10)/(8.11) from (8.17) by applying a result of Linear Algebra.

Lemma 35: If $M \in \mathcal{L}(\mathcal{E}, \mathcal{F})$ is invertible, and if $a \in \mathcal{F}, b \in \mathcal{E}'$, then $M + a \otimes b$ is invertible if $(M^{-1}a.b) \neq -1$ and
$$(M + a \otimes b)^{-1} = M^{-1} - \frac{1}{1 + (M^{-1}a.b)}M^{-1}(a \otimes b)M^{-1}. \tag{8.25}$$

Proof: One wants to solve $(M + a \otimes b)x = y$, i.e. $M\, x + a(b.x) = y$, and therefore $x = M^{-1}\, y - t\, M^{-1}\, a$ with $t = (b.x)$, but one needs then to have $t = (b.M^{-1}\, y) - t(M^{-1}\, a.b)$, which is possible because $(M^{-1}\, a.b) \neq -1$, and gives $x = M^{-1}\, y - M^{-1}\, a\frac{(b.M^{-1}\, y)}{1+(M^{-1}\, a.b)}$, and as y is arbitrary it gives (8.25).■

A few years ago, working with François MURAT on the relation between YOUNG measures and H-measures,[61] we computed the analog of formula (8.17) when one mixes r different materials.

Lemma 36: Mixing r materials with tensors $M_1, \ldots, M_r$, with respective proportions $\eta_1, \ldots, \eta_r$, in layers orthogonal to e, gives an effective tensor M^{eff} given by
$$\begin{aligned}
M^{eff} &= \sum_{i=1}^{r} \eta_i\, M_i - \sum_{1 \leq i < j \leq r} \eta_i\eta_j(M_i - M_j)R_{ij}(M_i - M_j)\\
R_{ij} &= \frac{1}{(M_i\, e.e)}\frac{e \otimes e}{H}\frac{1}{(M_j\, e.e)} \text{ for } i,j = 1, \ldots, r\\
H &= \sum_{k=1}^{r} \frac{\eta_k}{(M_k\, e.e)}.
\end{aligned} \tag{8.26}$$

Proof: As for the proof of Lemma 34, one uses
$$E_i = E^\infty + c_i\, e \text{ in the layers of material } \#i, \text{ for } i = 1, \ldots, r, \text{ and } \sum_{i=1}^{r} \eta_i\, c_i = 0, \tag{8.27}$$

---

[61] I have used in various publications not related to the purpose of this course our construction of admissible pairs of a YOUNG measure and a H-measure associated with a sequence, the first time at a meeting for the 600th anniversary of the University of Ferrara in 1991. Our redaction of these results is still a draft which we have not looked at for years, but I have nevertheless given it to a few persons, and I wonder how many will have claimed our results as theirs.

and one must have $(M_i\,E_i.e) = (M_j\,E_j.e)$ if there is an interface between material $\#i$ and material $\#j$, and therefore there exists a constant C such that

$$(M_i\,E_i.e) = C \text{ for } i = 1, \ldots, r. \tag{8.28}$$

With the definition of $E_i, i = 1, \ldots, r$, (8.28) implies

$$c_i = \frac{C - (M_i\,E^\infty.e)}{(M_i\,e.e)} \text{ for } i = 1, \ldots, r, \tag{8.29}$$

and the condition $\sum_i \eta_i c_i = 0$ gives

$$H\,C = \sum_{i=1}^{r} \eta_i \frac{(M_i\,E^\infty.e)}{(M_i\,e.e)}, \tag{8.30}$$

with H given in (8.26). Using (8.28) one obtains

$$H\,(M_i\,e.e)c_i = \Big(\sum_{j=1}^{r} \eta_j \frac{(M_j\,E^\infty.e)}{(M_j\,e.e)}\Big) - \frac{(M_i\,E^\infty.e)}{(M_i\,e.e)} = \sum_{j=1}^{r} \eta_j \frac{\big((M_j - M_i)\,E^\infty.e\big)}{(M_j\,e.e)} \text{ for } i = 1, \ldots, r. \tag{8.31}$$

This gives

$$
\begin{aligned}
M^{eff}\,E^\infty &= (\sum_{i=1}^{r} \eta_i\,M_i)E^\infty + \frac{1}{H}\sum_{i=1}^{r}\frac{\eta_i}{(M_i\,e.e)}\Big(\sum_{j=1}^{r}\eta_j\frac{\big((M_j - M_i)\,E^\infty.e\big)}{(M_j\,e.e)}\Big)M_i\,e \\
&= (\sum_{i=1}^{r} \eta_i\,M_i)E^\infty - \frac{1}{2H}\sum_{i,j=1}^{r}\eta_i\eta_j\frac{\big((M_i - M_j)E^\infty.e\big)}{(M_i\,e.e)(M_j\,e.e)}(M_i - M_j)e \\
&= (\sum_{i=1}^{r} \eta_i\,M_i)E^\infty - \frac{1}{H}\sum_{i<j}\eta_i\eta_j(M_i - M_j)\frac{e \otimes e}{(M_i\,e.e)(M_j\,e.e)}(M_i - M_j)E^\infty,
\end{aligned}
\tag{8.32}
$$

proving (8.26).∎

I have shown the derivation of the differential equation (8.7) because it has some intrinsic interest. I noticed later that by using relaxation techniques related to Lemma 1, one can replace $\frac{e \otimes e}{(A\,e.e)}$ in (8.7) by any convex combination $\sum_j \theta_j \frac{e_j \otimes e_j}{(A\,e_j.e_j)}$, or more generally $\int_{S^{N-1}} \frac{e \otimes e}{(A\,e.e)}\,d\pi(e)$ for a probability measure π on the sphere S^{N-1}, and this gives a differential analogue of formula (8.12).[62]

As I will explain in the next chapter, I also discovered in the Spring 1983 that the characterization obtained with François MURAT was not absolutely necessary for solving the problems that we had in mind. There was an obvious generalization to mixing an arbitrary number of isotropic materials, but there were some technical details for mixing anisotropic materials, and I only noticed the following results much later (see footnote 44), and Lemma 18 played a crucial role. Although the two methods for obtaining bounds on effective coefficients that I have described in chapters 5 and 7 are valid for nonnecessarily symmetric operators (as is the third one that I have mentioned in footnote 54, based on H-measures), the only applications that I know use symmetric operators,[63] and because one allows for arbitrary rotations the set of effective operators

---

[62] One can also use a convex combination in $(A.e)$, and I had hoped that this trick would give more characterization of effective coefficients. I did talk about this method at a meeting in Minneapolis in 1985, but I only mentioned it in writing for a meeting in Los Alamos in 1987.

[63] Around 1984, I generalized a formula of Joseph KELLER valid for dimension $N = 2$ [Ke], and I later learned from Graeme MILTON that he had also discovered the same result (and he had kindly proposed that I cosign an article where he was using these formulas; I do not know the precise reference of his article). He had been led to these formulas by studying HALL effect, and if my memory is correct it is a nonzero value of a magnetic field which is responsible for the appearance of a nonsymmetric tensor, and this is the only instance of a non symmetric situation which I have heard about in real problems. Later I tried to use these formulas with Michel ARTOLA in order to prove a conjecture of Stefano MORTOLA and Sergio STEFFÉ [Mo&St2] (Sergei KOZLOV told me in 1993 in Trieste that the conjecture is false, but he did not provide enough information, so that I do not know if he had proved it or if he just thought that it was wrong).

corresponding to mixtures using precise proportions of each constituent is a set of matrices defined in terms of their eigenvalues, and that point is worth discussing in detail.

One starts from a finite number of materials with symmetric anisotropic tensors $M_i, i = 1, \ldots, r$, and mixing them with local proportions $\eta_i, i = 1, \ldots, r$, with $\sum_i \eta_i = 1$ a.e. in Ω, means that one considers sequences χ_i^n of characteristic functions of disjoint measurable sets for $i = 1, \ldots, r$, i.e. $\chi_i^n \chi_j^n = 0$ whenever $i \neq j$, such that

$$\chi_i^n \rightharpoonup \eta_i \text{ in } L^\infty(\Omega) \text{ weak } \star$$

$$A^n = \sum_{i=1}^r \chi_i^n (R^n)^T M_i R^n, \text{ with } R^n \in SO(N) \text{ a.e. in } \Omega \tag{8.33}$$

$$A^n \text{ H-converges to } A^{eff},$$

and one writes

$$A^{eff} \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) \text{ a.e. in } \Omega, \tag{8.34}$$

and the claim is that the set $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ of effective materials obtained by mixing the materials $M_1, \ldots, M_r$ with "exact proportions" $\eta_1, \ldots, \eta_r$ only depends upon A^{eff} through its eigenvalues $\lambda_1^{eff}, \ldots, \lambda_N^{eff}$, which satisfy

$$(\lambda_1^{eff}, \ldots, \lambda_N^{eff}) \in \Lambda(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) \text{ a.e. in } \Omega, \tag{8.35}$$

where $\Lambda(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ is a subset of R^N which is invariant by permutation of the coordinates, because one has not imposed any rule for ordering the eigenvalues. Like for the discussion following Proposition 17, the statement above is clear from the intuitive understanding of what mixing is about, but I have not given any precise mathematical definition of what the set $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ is yet. In 1983, Robert KOHN had asked me a question showing that he was concerned about a problem of this kind: in my work with François MURAT, obtained for mixing two isotropic materials, we had found a necessary condition of the type $A^{eff} \in S(\theta)$ for a set of matrices $S(\theta)$ which was our candidate for $\mathcal{K}(\theta, 1 - \theta; \alpha I, \beta I)$, and because this set is convex it is easy to approach a function (θ, A) such that $A(x) \in S(\theta(x))$ a.e. $x \in \Omega$ by a piecewise constant function satisfying the same constraint, as after decomposing Ω into small open cubes (plus a set of measure 0), one can replace θ on any such small cube ω by its average $\overline{\theta}$ on ω and replace A by its projection on $S(\overline{\theta})$, giving a new function $(\widetilde{\theta}, \widetilde{A})$ satisfying the same constraint with $\widetilde{\theta}$ piecewise constant; then on each small cube one approaches $\widetilde{A}$ by a piecewise constant function (on much smaller cubes) taking its values in $S(\overline{\theta})$. One also uses the fact that the HAUSDORFF distance from $S(\theta_1)$ to $S(\theta_2)$ is $O(|\theta_1 - \theta_1|)$. However, the convexity of each $S(\theta)$ is not really important, and one can give a precise meaning of (8.34), for the case (8.33) or for other still more general situations, in the following way.

Definition 37: For nonnegative real numbers $\theta_1, \ldots, \theta_r$, with $\sum_i \theta_i = 1$, and $P \in \mathcal{L}(R^N, R^N)$, one says that P belongs to $\mathcal{K}(\theta_1, \ldots, \theta_r; M_1, \ldots, M_r)$ if and only if there exist sequences χ_i^n of characteristic functions of disjoint measurable sets for $i = 1, \ldots, r$, and a sequence of rotations $R^n \in SO(N)$ such that (8.33) holds, and moreover such that there exists $x_* \in \Omega$, LEBESGUE point of A^{eff} and of $\eta_1, \ldots, \eta_r$, with $\eta_i(x_*) = \theta_i$ for $i = 1, \ldots, r$, and $A^{eff}(x_*) = P$.∎

Of course, because the set of LEBESGUE points of any vector valued function is dense, the fact that (8.34) is valid is now merely the statement of Definition 37, but one must show that this definition is consistent with the intuitive idea of mixing materials. The proof makes use of an obvious fact, that H-convergence commutes with translations and dilations. Commuting with translations means that if A^n H-converges to A^{eff} in Ω and if for $a \in R^N$ one has $B^n(x) = A^n(x+a)$ a.e. $x \in \Omega - a$, then B^n H-converges to B^{eff} in $\Omega - a$ and $B^{eff}(x) = A^{eff}(x+a)$ a.e. in $\Omega - a$; commuting with dilations, or rescaling, means that if for $s \neq 0$ one has $C^n(x) = A^n(s\,x)$ a.e. $x \in s^{-1}\,\Omega$, then C^n H-converges to C^{eff} in $s^{-1}\,\Omega$ and $C^{eff}(x) = A^{eff}(s\,x)$ a.e. in $s^{-1}\,\Omega$. More generally, one has the following result about changing variables in H-convergence, identical to the formula first proved by Sergio SPAGNOLO in the case of G-convergence.

Lemma 38: If φ is a diffeomorphism from Ω onto $\varphi(\Omega)$ and A^n H-converges to A^{eff} in Ω, and B^n is defined on $\varphi(\Omega)$ by

$$B^n\big(\varphi(x)\big) = \frac{1}{det\big(\nabla\varphi(x)\big)} \nabla\varphi(x) A^n(x) \nabla\varphi^T(x) \text{ a.e. } x \in \Omega, \tag{8.36}$$

then B^n H-converges in $\varphi(\Omega)$ to B^{eff}, and

$$B^{eff}\big(\varphi(x)\big) = \frac{1}{det\big(\nabla\varphi(x)\big)}\nabla\varphi(x)A^{eff}(x)\nabla\varphi^T(x) \text{ a.e. } x \in \Omega. \tag{8.37}$$

Proof: If $-div\big(A^n\,grad(u_n)\big) = f$ in Ω, one defines v_n in $\varphi(\Omega)$ by $v_n(y) = u_n\big(\varphi^{-1}(y)\big)$ a.e. $y \in \varphi(\Omega)$, or equivalently $u_n(x) = v_n\big(\varphi(x)\big)$ a.e. $x \in \Omega$, so that $grad(u_n)(x) = \nabla\varphi^T(x)\,grad(v_n)\big(\varphi(x)\big)$ a.e. $x \in \Omega$. Writing then the equation in variational form $\int_\Omega\big(A^n\,grad(u_n).grad(w)\big)\,dx = \int_\Omega f\,w\,dx$ for all $w \in C_c^1(\Omega)$, one finds that $-div\big(B^n\,grad(v_n)\big) = g$ in $\varphi(\Omega)$, with B^n given by (8.36) and $g\big(\varphi(x)\big) = \frac{1}{det\big(\nabla\varphi(x)\big)}f(x)$ a.e. $x \in \Omega$, in the case $f \in L^2(\Omega)$, with straightforward generalization in the case $f \in H^{-1}(\Omega)$. Then one relates the weak limits of $grad(v_n)$ and of $B^n\,grad(v_n)$ in $\varphi(\Omega)$ to the weak limits of $grad(u_n)$ and of $A^n\,grad(v_n)$ in Ω, and one finds that B^{eff} is defined by (8.37).∎

Lemma 39: For nonnegative real numbers $\theta_1,\ldots,\theta_r$, with $\sum_i \theta_i = 1$, and $P \in \mathcal{K}(\theta_1,\ldots,\theta_r; M_1,\ldots,M_r)$, there exist sequences χ_i^n of characteristic functions of disjoint measurable sets for $i = 1,\ldots,r$, and a sequence $R^n \in L^\infty\big(\Omega; SO(N)\big)$ such that (8.33) holds, with $\eta_i = \theta_i$ a.e. in Ω for $i = 1,\ldots,r$, and $A^{eff} = P$ a.e. in Ω.
Proof: One proves the Lemma with Ω replaced by a cube Q centered at 0 and large enough to contain Ω; then one restricts the result to Ω, using Proposition 10. By Definition 37 there exists a point x_* and sequences χ_i^n, $i = 1,\ldots,r$, R^n, which are not yet the ones needed in the Lemma, and one obtains the desired ones by translation and rescaling. For an integer k large enough so that $x_* + \frac{1}{k}Q \subset \Omega$, one defines $\chi_i^{n,k}$, $i = 1,\ldots,r$, $R^{n,k}$, $A^{n,k}$ in Q by $\chi_i^{n,k}(x) = \chi_i^n\big(x_* + \frac{x}{k}\big)$ a.e. $x \in Q$, $i = 1,\ldots,r$, $R^{n,k}(x) = R^n\big(x_* + \frac{x}{k}\big)$, $A^{n,k}(x) = A^n\big(x_* + \frac{x}{k}\big)$ a.e. $x \in Q$. By Lemma 38, for k fixed $A^{n,k}$ H-converges in Q to $A^{eff,k}$, defined by $A^{eff,k}(x) = A^{eff}\big(x_* + \frac{x}{k}\big)$ a.e. $x \in Q$, and because x_* is a LEBESGUE point of A^{eff}, $A^{eff,k}$ converges in $L^\infty(Q)$ weak $\star$ and $L^1(Q)$ strong to the constant tensor $P = A^{eff}(x_*)$, and therefore $A^{eff,k}$ H-converges to P in Q. Similarly, for k fixed $\chi_i^{n,k}$ converges in $L^\infty(Q)$ weak $\star$ to $\chi^{\infty,k}$ defined by $\chi^{\infty,k}(x) = \eta_i\big(x_* + \frac{x}{k}\big)$ a.e. $x \in Q$ for $i = 1,\ldots,r$, and because x_* is a LEBESGUE point of each η_i, $\chi^{\infty,k}$ converges in $L^\infty(Q)$ weak $\star$ and $L^1(Q)$ strong to the constant function $\theta_i = \eta_i(x_*)$, for $i = 1,\ldots,r$. Using the metrizability of H-convergence restricted to $M(\alpha,\beta;Q)$, when $0 < \alpha \le \beta < \infty$ have been choosen so that $M_i \in M(\alpha,\beta)$ for $i = 1,\ldots,r$, and the metrizability of $L^\infty(Q)$ weak $\star$ convergence on bounded sets, there exists a diagonal subsequence indexed by n',k' such that $A^{n',k'}$ H-converges in Q to the constant tensor P, and $\chi_i^{n',k'}$ converges in $L^\infty(Q)$ weak $\star$ to the constant function θ_i, for $i = 1,\ldots,r$.∎

Lemma 40: If for $i = 1,\ldots,r$, $\eta_i \in L^\infty(\Omega)$, $\eta_i \ge 0$ a.e. in Ω, with $\sum_i \eta_i = 1$ a.e. in Ω, and $P \in L^\infty\big(\Omega;\mathcal{L}(R^N,R^N)\big)$ with $P(x) \in \mathcal{K}(\eta_1(x),\ldots,\eta_r(x); M_1,\ldots,M_r)$ a.e. $x \in \Omega$, then there exists sequences χ_i^n of characteristic functions of disjoint measurable sets for $i = 1,\ldots,r$, and a sequence $R^n \in L^\infty\big(\Omega; SO(N)\big)$ such that (8.33) holds and $A^{eff} = P$ in Ω.
Proof: Let $g \in L^1(\Omega; R^p)$ and for $\varepsilon > 0$ let ρ_ε be defined by $\rho_\varepsilon(x) = \frac{1}{\varepsilon^N}\rho_1\big(\frac{x}{\varepsilon}\big)$ with $\rho_1 \in L^1(R^N)$, nonnegative and with compact support; then $\int_{\Omega \times R^N}|g(x) - g(x - y)|\rho_\varepsilon(y)\,dx\,dy$ tends to 0 as ε tends to 0 (as it is $\le 2||g||_{L^1}||\rho_1||_{L^1}$, it is enough to prove the result for a dense subspace, and for $g \in C_c(R^N)$ it is immediate as g is uniformly continuous). For ρ_1 the characteristic function of the cube $(0,1)^N$, and $\delta = \frac{1}{k}$ one can then choose ε small enough to have $\int_{|x-z| \le \varepsilon}|g(x) - g(z)|\,dx\,dz \le \delta\,\varepsilon^N$. Then decomposing R^N into disjoint cubes ω_j of size $\frac{\varepsilon}{\sqrt{N}}$ (plus a set of measure 0), one chooses for each cube ω_j a point $z^j \in \omega_j$ such that $\int_{\omega_j}|g(x) - g(z_j)|\,dx \le \frac{2}{\varepsilon^N}\int_{\omega_j \times \omega_j}|g(x) - g(z)|\,dx\,dz \le \frac{2}{\varepsilon^N}\int_{x \in \omega_j, |x-z| \le \varepsilon}|g(x) - g(z)|\,dx\,dz$, and if g^k denotes the function equal to $g(z_j)$ on ω_j, then g is piecewise constant and one has $||g - g^k||_{L^1(R^N)} \le \frac{1}{k}$.

One applies the preceding analysis to $g = (\eta_1,\ldots,\eta_r,P)$, and $g^k = (\eta_1^k,\ldots,\eta_r^k,P^k)$ and one has $P^k \in \mathcal{K}(\eta_1^k,\ldots,\eta_r^k; M_1,\ldots,M_r)$ on each cube ω_j. Using Lemma 39, as well as Proposition 10 in order to glue the pieces together, there exist sequences $\chi_i^{n,k}$ of characteristic functions of disjoint measurable sets for $i = 1,\ldots,r$, and a sequence $R^{n,k} \in L^\infty\big(\Omega; SO(N)\big)$ such that (8.33) holds, with η_i replaced by η_i^k for $i = 1,\ldots,r$, and A^{eff} replaced by P^k. As k tends to ∞, η_i^k converges in $L^\infty(\Omega)$ weak $\star$ and $L^1(\Omega)$ strong to η^k for $i = 1,\ldots,r$, and P^k converges in $L^\infty(\Omega)$ weak $\star$ and $L^1(\Omega)$ strong to P, and therefore P^k H-converges to P in Ω. Using the metrizability of H-convergence restricted to $M(\alpha,\beta;Q)$, and the metrizability of $L^\infty(Q)$ weak $\star$ convergence on bounded sets, there exists a diagonal subsequence indexed by n',k' such that $\chi_i^{n',k'}$ converges in $L^\infty(\Omega)$ weak $\star$ to η_i, for $i = 1,\ldots,r$, and $A^{n',k'}$ H-converges in Ω to P.∎

Because of the local character of H-convergence expressed by Proposition 10, there is no mention of a particular open set in Definition 37, which is actually valid without any hypothesis of symmetry and without the use of rotations R^n (depending measurably in x) in (8.33). However, it is precisely the symmetry of M_i, $i = 1, \ldots, r$, and the use of these rotations which implies that (8.34) has the form (8.35), and this is seen by using Lemma 38 with $\varphi(x) = R\,x$ with $R \in SO(N)$: it shows that if $P \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ then $R\,P\,R^T \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ for any $R \in SO(N)$, and as P is symmetric, all tensors with the same eigenvalues than P belong to the set $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, and one deduces (8.35).[64]

I must say that prior to writing these notes I had not been so interested in writing down the detail of the analysis which I just showed, considering that these are uninteresting details which everyone with a good knowledge of measure theory can check, and that the important questions of Homogenization lie elsewhere. I have heard some mention of a work of Gianni DAL MASO and Robert KOHN which might have addressed this question in general, or it may have been the answer to the question that Robert KOHN was asking me in 1983, which was related to linearized Elasticity, which is not a frame indifferent theory, as he already knew; his concern was to identify which are the invariant quantities replacing the eigenvalues in the case of fourth order tensors C_{ijkl}. I do not see any special difficulty in adapting Definition 37 to this case, but as I consider that questions in linearized Elasticity are too often spoiled by unrealistic effects which deprive the mathematical results of much of their value, I have preferred not to be involved in such questions.

The set $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ is not necessarily convex: in dimension $N = 2$ with only one material M_1 symmetric with distinct eigenvalues $\alpha < \beta$, $\mathcal{K}(1; M_1)$ is the set of symmetric matrices with eigenvalues $\lambda_1, \lambda_2 \in [\alpha, \beta]$ with $\lambda_1 \lambda_2 = \alpha\,\beta$. However, some projections of $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ are convex [Ta14], and this is not dependent on symmetry requirements or use of rotations.

Lemma 41: Consider $\eta_1, \ldots, \eta_r$, nonnegative numbers adding up to 1, and $M_1, \ldots, M_r$, tensors from $M(\alpha, \beta)$; then for every set of $N - 1$ vectors $a^1, \ldots, a^{N-1}$ of R^N,

$$\{(P\,a^1, \ldots, P\,a^{N-1}) | P \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)\} \text{ is convex.} \tag{8.38}$$

Moreover

$$P_1, P_2 \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) \text{ and } rank(P_2 - P_1) \le N - 1$$
$$\text{imply } \theta\,P_1 + (1 - \theta)P_2 \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) \text{ for all } \theta \in (0, 1). \tag{8.39}$$

Proof: For $P_1, P_2 \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ and $\theta \in (0, 1)$ one exhibits $P_3 \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ such that

$$P_3\,a^i = \theta\,P_1\,a^i + (1 - \theta)P_2\,a^i \text{ for } i = 1, \ldots, N - 1, \tag{8.40}$$

and this is done by layering materials with tensors P_1 and P_2, respectively with proportions η and $1 - \eta$, and this obviously creates a tensor belonging to $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, and then one uses formula (8.37) of Lemma 34. The nonzero vector e orthogonal to the layers must satisfy

$$\Big((P_2 - P_1)a^i.e\Big) = 0 \text{ for } i = 1, \ldots, N - 1, \tag{8.41}$$

and this is possible because there are only $N - 1$ vectors a_i, but if $P_2 - P_1$ does not have full rank, one can take e orthogonal to the range of $P_2 - P_1$, and (8.39) follows.∎

It means that if one considers the first $N - 1$ columns of elements from $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ in a basis $a^1, \ldots, a^N$, then one has a convex set. If the basis is orthogonal and all M_i are symmetric, then the choice (8.41) for e in the proof of Lemma 41 means that $(P_2 - P_1)e$ is proportional to a^N and the last term

---

[64] The first method for obtaining bounds on effective coefficients, which I developed with François MURAT in the early 70s, and which I have described in chapter 5, is not restricted to symmetric operators, but I do not know what kind of transformations to impose in the nonsymmetric case. I have not read carefully the article of Graeme MILTON mentioned in footnote 63 as the only instance of a nonsymmetric situation that I have heard of in a realistic situation, but knowing that the HALL effect is concerned with electrical currents in thin ribbons, the macroscopic direction of the current is obviously imposed and therefore the situation is not subject to frame indifference.

in the formula (8.37) is then proportional to $a^N \otimes a^N$ and therefore if one considers the $N^2 - 1$ entries of elements P of $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ obtained by deleting the entry P_{NN}, then one has a convex set. Even though $\mathcal{K}(\eta, 1 - \eta; \alpha\, I, \beta\, I)$ is characterized by explicit inequalities (7,25), (7.28), (7.30), I do not know a simple description of what all these convex sets are in this special case, but if one just wants to identify one column of elements of $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, then one has the following result in the symmetric case using rotations [Ta14].

Lemma 42: Consider $\eta_1, \ldots, \eta_r$, nonnegative numbers adding up to 1, and $M_1, \ldots, M_r$, symmetric tensors from $M(\alpha, \beta)$, with eigenvalues $\lambda_j(M_i)$, $i = 1, \ldots, r$, $j = 1, \ldots, N$ and $N \geq 2$; then for every $E \in R^N$,

$$\{(P\, E | P \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)\} \text{ is the closed ball with diameter } [b\, E, a\, E]$$

$$a = \sum_{i=1}^{r} \eta_i \max_j \{\lambda_j(M_i)\}$$

$$\frac{1}{b} = \sum_{i=1}^{r} \frac{\eta_i}{\min_j\{\lambda_j(M_i)\}}. \tag{8.42}$$

Proof: Let $\lambda_-(M_i) = \min_j\{\lambda_j(M_i)\}$ and $\lambda_+(M_i) = \max_j\{\lambda_j(M_i)\}$. Let $e_1, \ldots, e_N$ be an orthonormal basis, and rotate the materials in such a way that it becomes a basis of eigenvector for each $R_i^T M_i R_i$, with $R_i^T M_i R_i e_1 = \lambda_-(M_i)e_1$ and $R_i^T M_i R_i e_N = \lambda_+(M_i)e_N$; layering in direction orthogonal to e_1 these materials $R_i^T M_i R_i$ with proportion η_i gives by formula (4.11) a tensor $P \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ such that $P\, e_1 = b\, e_1$ and $P\, e_N = a\, e_N$. If E belongs to the plane spanned by e_1 and e_N, then $P\, E = b(E.e_1)e_1 + a(E.e_N)e_N$, so that $P\, E - b\, E = (a - b)(E.e_N)e_N$ and $P\, E - a\, E = (b - a)(E.e_1)e_1$, and therefore $(P\, E - b\, E.P\, E - a\, E) = 0$, showing that $P\, E$ belongs to the sphere with diameter $[b\, E, a\, E]$. By choosing all possible orientations, the set of $P\, E$ contains at least the sphere with diameter $[b\, E, a\, E]$, and as by Lemma 41, the set that we are looking for is convex, it must contain the closed ball with diameter $[b\, E, a\, E]$.

That the desired set is included in the closed ball with diameter $[b\, E, a\, E]$ follows from Proposition 12 and Proposition 14 (or from Lemma 18). Assume that $A^n = \sum_i \chi_i^n (R^n)^T M_i R^n$ H-converges to A^{eff} in Ω, where the χ_i^n are characteristic functions of disjoint measurable sets with χ_i^n converging in $L^\infty(\Omega)$ weak $\star$ to η_i for $i = 1, \ldots, r$. As $(R^n)^T M_i R^n$ has eigenvalues between $\lambda_-(M_i)$ and $\lambda_+(M_i)$ then one has $B^n = \left(\sum_i \chi_i^n \lambda_-(M_i)\right)I \leq A^n \leq \left(\sum_i \chi_i^n \lambda_+(M_i)\right) = C^n$, and if one extracts a subsequence such that B^m H-converges to B^{eff} and C^m H-converges to C^{eff}, then by Proposition 14 one has $B^{eff} \leq A^{eff} \leq C^{eff}$. By Proposition 12, an upper bound for C^{eff} is C_+, the limit in $L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right)$ weak $\star$ of C^n, which by (8.42) is $a\, I$, and a lower bound for B^{eff} is B_-, where $(B_-)^{-1}$ is the limit in $L^\infty\left(\Omega; \mathcal{L}(R^N, R^N)\right)$ weak $\star$ of $(B^n)^{-1}$, which by (8.42) is $\frac{1}{b}\, I$. One has then $b\, I \leq A^{eff} \leq a\, I$, and therefore (as was shown before Lemma 18), $A^{eff}\, E$ belongs to the closed ball with diameter $[b\, E, a\, E]$, a.e. in Ω.∎

9. Necessary conditions of optimality: first step

At a meeting in New York in 1981, I had described the characterization found with François MURAT of all effective materials obtained by mixing two isotropic materials, i.e. (7.25) (Proposition 10, described in [Ta1]) and (7.28), (7.30), obtained by following my second method for obtaining bounds satisfied by effective coefficients (Theorem 26, described in [Ta7], and Lemmas 28, 29). It was clear that our program for questions of Optimal Design initialized in the early 70s was completed, but I did not try immediately to extend the computation of necessary conditions that I had done in [Ta2].

During the Spring 1983, I was invited at a Midwest PDE Seminar in Madison, and I heard Michael RENARDY talk about his work with Daniel JOSEPH on POISEUILLE flows of two immiscible fluids.[65] For a cylinder with arbitrary cross section Ω, there are infinitely many possible POISEUILLE flows, but when Ω is a disc one only observes the flow where the less viscous fluid occupies an annular region near the boundary, and they had imagined that it was related to a maximization of dissipation. This situation is described by (4.1) with $f = 1$ (a is the viscosity of the fluid, independent of the direction along the pipe, and the gradient of the pressure is constant, pointing in the direction of the pipe), and one wants to maximize

---

[65] In applications, it either meant melted polymers or crude oil and water, the water being added in small quantity in a pipeline for lubricating it.

$\int_\Omega a|grad(u)|^2\, dx = \int_\Omega u\, dx$, and therefore it corresponds to $g(x, u, a) = -u$ in (4.3). There is indeed a classical solution in the case of a circular cross section, which is as described, but as they were conjecturing the existence of a classical solution in general, I told Michael RENARDY that on the contrary I expected that no classical solution would exist in general.[66] In order to check about this question, I looked at the necessary conditions of optimality, as I had done in [Ta2], but using now the characterization that François MURAT and I had obtained two years before. Surprisingly, I discovered that our precise characterization could be ignored completely, and that the same necessary conditions could be deduced from the crude bounds that we had obtained almost ten years earlier.[67] Assuming that a classical solution exists, and that the interface between the two materials is smooth enough, the necessary conditions of optimality imply that some DIRICHLET conditions and some NEUMANN conditions must be satisfied on the interface, which is quite unlikely in general. On my way back to France I stopped in New York, and discussing about this matter in the lounge at the COURANT Institute, the precise argument for rejecting that possibility was mentioned to me by Joel SPRUCK, who reminded me of a result of James SERRIN that I had heard at a conference in Jerusalem in 1972 (and he told me about a quicker proof by Hans WEINBERGER); this result assumes the domain to be simply connected, and shows that among simply connected domains a classical solution only exists for circular domains.[68]

In July 1983, I described completely our method for a Summer Course CEA - EDF - INRIA on Homogenization at Bréau sans Nappe, which gave François MURAT and I the occasion to write [Mu&Ta1]. As this first text had been written in French, we wrote a summary in English for a conference at Isola d'Elba in the Fall [Mu&Ta2]. In November, I took the occasion of a meeting in Paris dedicated to Ennio DE GIORGI for writing down the details of the characterization obtained with François MURAT in 1980 [Ta9], characterization which I had first described at a meeting in New York in 1981 but only quoted in [Mu&Ta1], and in [Ta9] I gave our initial construction with coated ellipsoids generalizing the idea of using coated spheres that I had learned in the original work of Zvi HASHIN and S. SHTRIKMAN [Ha&Sh], just after it had been pointed out to me by George PAPANICOLAOU, and for the computations with multiple layerings I gave the generalization that I had obtained in the Spring 1983 in Berkeley of the work initially done the year before by Philippe BRAIDY and Didier POUILLOUX [Br&Po].

In the early 70s, François MURAT and I had developed the Homogenization approach in order to describe a relaxed problem, and this is a way to prove existence of generalized solutions. For the sake of generality, I describe here the more general framework that I developed later in [Ta14], instead of the restricted problem of mixing only two isotropic materials, as we had done in [Mu&Ta1]. For a bounded open set Ω of R^N, one considers a mixture of r symmetric possibly anisotropic materials $M_1, \ldots, M_r \in M(\alpha, \beta)$, in finite quantity $\kappa_1, \ldots, \kappa_r$, assuming that $\sum_i \kappa_i \geq meas(\Omega)$, of course. Then one defines

$$A = \sum_{i=1}^r \chi_i\, R^T\, M_i\, R \text{ in } \Omega, \tag{9.1}$$

where χ_i, $i = 1, \ldots, r$, are characteristic functions of disjoint measurable sets with union equal to Ω (up to a set of measure 0), and R is measurable and takes values in $SO(N)$, a.e. in Ω, and one assumes that

$$\int_\Omega \chi_i\, dx \leq \kappa_i \text{ for } i = 1, \ldots, r. \tag{9.2}$$

---

[66] The idea that turbulent flows are trying to optimize something has been suggested before, as I have read in a book by Daniel JOSEPH [Jo], but it is not clear if fluids are trying to optimize something, and what it could be; however, this idea is consistent with the Homogenization approach that in order to optimize some criteria one might have to create adequate microstructures.

[67] As I learned later, this point had already been noticed in the late 70s by RAITUM [Ra].

[68] Of course, when Ω is not a disc this negative result does not tell what configurations will be chosen by a mixture of fluids in a pipe, as the question of what mixtures of fluids really try to optimize must be studied further. Daniel JOSEPH, Michael RENARDY and Yuriko RENARDY have also studied the stability of the solution for circular domains, and for such questions of stability one must go back to the full NAVIER-STOKES equation, of course.

Then one solves the state equation

$$-div\Big(A\,grad(u)\Big) = f \text{ in } \Omega, \text{ with } u \in H_0^1(\Omega),\qquad(9.3)$$

with $f \in H^{-1}(\Omega)$, and I have choosen homogeneous DIRICHLET conditions for u as an example. Finally one wants to minimize the cost function J defined by

$$J(\chi_1,\ldots,\chi_r,R) = \int_\Omega \sum_{i=1}^r \chi_i\, F_i\Big(x,u(x)\Big)\,dx,\qquad(9.4)$$

so that there is no cost for rotating materials, and one adds regularity and growth hypotheses on $F_1,\ldots,F_r$, in order to have

$$u \mapsto F_i(\cdot,u) \text{ is sequentially continuous from } H_0^1(\Omega) \text{ weak into } L^1(\Omega) \text{ strong, for } i=1,\ldots,r,\qquad(9.5)$$

which is usually obtained by assuming that each F_i is a CARATHÉODORY function, and that $\sum_i |F_i(x,v)| \leq \mu(x) + C|v|^p$ for all $v \in R$ a.e. $x \in \Omega$, with $\mu \in L^1(\Omega)$ and $p < \frac{2N}{N-2}$ for $N \geq 2$ (but for $N = 2$ one can assume instead that $\sum_i |F_i(x,v)| \leq \mu(x) + G(|v|)$ and allow some exponential growth for G), and for $N = 1$ one assumes that $\sum_i \sup_{|v|\leq r} |F_i(x,v)| \leq \mu_r(x)$ with $\mu_r \in L^1(\Omega)$ for all r. In the case where $\sum_i \kappa_i > meas(\Omega)$, one can also add in J a term of the form $G(\int_\Omega \chi_1\,dx,\ldots,\int_\Omega \chi_r\,dx)$.

The set of admissible A is not empty, because $\sum_i \kappa_i \geq meas(\Omega)$, and it is included in some $M(\alpha,\beta;\Omega)$, so that all the u obtained belong to a bounded set of $H_0^1(\Omega)$. For a sequence A^n, for example a minimizing sequence, one can extract a subsequence which H-converges to A^{eff}, and also such that for $i = 1,\ldots,r$, the sequence χ_i^n converges in $L^\infty(\Omega)$ weak $\star$ to a nonnegative function η_i, which satisfies and

$$\int_\Omega \eta_i\,dx \leq \kappa_i \text{ for } i = 1,\ldots,r,\qquad(9.6)$$

with $\sum_i \eta_i = 1$ in Ω, and therefore by Definition 37

$$A^{eff} \in \mathcal{K}(\eta_1,\ldots,\eta_r;M_1,\ldots,M_r) \text{ a.e. in } \Omega,\qquad(9.7)$$

and also such that u_n converges in $H_0^1(\Omega)$ weak to u_∞, and therefore (9.5) implies

$$J(\chi_1^n,\ldots,\chi_r^n,R^n) \to \int_\Omega \Big(\sum_{i=1}^r \eta_i F_i(x,u_\infty)\Big)\,dx.\qquad(9.8)$$

By Lemma 40, if r nonnegative functions η_i, $i = 1,\ldots,r$, satisfy $\sum_i \eta_i = 1$ in Ω and (9.6), and $P \in \mathcal{K}(\eta_1,\ldots,\eta_r;M_1,\ldots,M_r)$ a.e. in Ω, there exist sequences χ_i^n of characteristic functions of disjoint measurable sets converging in $L^\infty(\Omega)$ weak $\star$ to η_i for $i = 1,\ldots,r$, and A^n H-converges to P, but there is a small technical point to resolve, because the sequence created in the proof of Lemma 40 might not satisfy (9.2). In that case one has $\int_\Omega \chi_i^n\,dx \leq \kappa_i + \varepsilon_n$ for $i = 1,\ldots,r$, and ε_n tends to 0 because $\int_\Omega \eta_i\,dx \leq \kappa_i$ for $i = 1,\ldots,r$. For each i one can change χ_i^n on a set of measure at most ε_n in order to create sequences $\widetilde{\chi}_i^n$ of characteristic functions of disjoint measurable sets converging in $L^\infty(\Omega)$ weak $\star$ to η_i for $i = 1,\ldots,r$ and satisfying (9.2), and $\widetilde{A}^n = \sum_i \widetilde{\chi}_i^n (R^n)^T M_i R^n$ H-converges to $\widetilde{P}$, and one must show that $\widetilde{P} = P$ a.e. in Ω. As the perturbations are not small in L^∞ norm, one cannot apply Proposition 16, but one can control the effect of small perturbations of the coefficients in L^s norm with $s < \infty$ (with the coefficients staying in $M(\alpha,\beta;\Omega)$) by using MEYERS's regularity theorem [Me]; however, one can also construct a more direct proof based on Proposition 10 as follows. If for $i = 1,\ldots,r$, one chooses a small cube ω_i such that $\int_{\omega_i} \eta_i\,dx > 0$, and one chooses n large enough so that $\int_{\omega_i} \chi_i^n\,dx \geq \varepsilon_n$ for $i = 1,\ldots,r$, then one can manage to make the changes only inside $\bigcup_i \omega_i$, and as $\widetilde{A}^n = A^n$ in $\Omega \setminus (\bigcup_i \overline{\omega_i})$, Proposition 10 implies that $\widetilde{P} = P$ in $\Omega \setminus (\bigcup_i \overline{\omega_i})$; the ω_i can be taken as small as one wants (around a density point of η_i), and therefore $\widetilde{P} = P$ a.e. in Ω.

One has therefore identified a relaxed problem (as defined in the end of chapter 3): the new set of controls is the set of $(\eta_1, \ldots, \eta_r, A)$ satisfying

$$0 \leq \eta_i \text{ a.e. in } \Omega \text{ for } i = 1, \ldots, r, \ \sum_{i=1}^{r} \eta_i = 1 \text{ a.e. in } \Omega$$

$$\int_{\Omega} \eta_i \, dx \leq \kappa_i \text{ for } i = 1, \ldots, r \qquad (9.9)$$

$$A \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) \text{ a.e. in } \Omega,$$

the state u is still given by (9.3), and the cost function $\widetilde{J}$ to be minimized is given by

$$\widetilde{J}(\eta_1, \ldots, \eta_r, A) = \int_{\Omega} \Big[\sum_{i=1}^{r} \eta_i \, F_i\Big(x, u(x)\Big) \Big] \, dx. \qquad (9.10)$$

The initial problem corresponds to the case where each η_i is a characteristic function.

Lemma 43: The function $\widetilde{J}$ given by (9.10) with u defined by (9.3) attains its minimum on the set described by (9.9).
Proof: The topology on the new control set described by (9.9) is the $L^\infty(\Omega)$ weak $\star$ topology for each η_i and the H-convergence topology for A, and this topology is metrizable. The new control set is the completion of the initial control set, and it is compact for this topology (this uses Theorem 6, Lemma 40, and the simple trick shown above for enforcing (9.2), together with classical results of Functional Analysis for dealing with the $L^\infty(\Omega)$ weak $\star$ topology). Moreover, $\widetilde{J}$ is continuous for that topology. A consequence of all these facts is that the minimum of $\widetilde{J}$ is attained.∎

The questions are rather different for more general functionals depending upon $grad(u)$, as we already knew in the early 70s; I have described some very partial results on this question in [Ta13]. However, the property (9.4)/(9.5) is still true for some functionals depending upon $grad(u)$ in a "fake" way. For example, (9.3) implies that $\int_{\Omega} \big(A \, grad(u).grad(u)\big) \, dx = \int_{\Omega} f \, u \, dx$ if $f \in L^2(\Omega)$, so that one should not really say that the first integral depends upon $grad(u)$! There are other situations of this kind like the following ones, where one cannot make $grad(u)$ disappear completely; for simplicity, I do not try to use MEYERS's regularity theorem [Me]. Assume that for some j, one has $J(\chi_1, \ldots, \chi_r, R) = \int_{\Omega} g(u)\frac{\partial u}{\partial x_j} v \, dx$, with g continuous and having limited growth $|g(u)| \leq C(1 + |u|^p)$ and $p < \frac{N}{N-2}$ if $N \geq 2$, with $v \in L^q(\Omega)$ and $p\big(\frac{1}{2} - \frac{1}{N}\big) + \frac{1}{2} + \frac{1}{q} \leq 1$, so that $q < \infty$ and the integral makes sense by an application of SOBOLEV imbedding theorem. If v is not regular, one cannot make $grad(u)$ disappear entirely by integration by parts, but one can decompose v as $v_1 + v_2$ with v_1 small in $L^q(\Omega)$ and $v_2 \in C_c^1(\Omega)$, and one notices that $\int_{\Omega} g(u)\frac{\partial u}{\partial x_j} v_2 \, dx = -\int_{\Omega} G(u)\frac{\partial v_2}{\partial x_j} \, dx$, where $G(s) = \int_0^s g(\sigma) \, d\sigma$; because the last integral is sequentially continuous on $H_0^1(\Omega)$ weak, one finds that the same is true for J as a uniform limit of sequentially weakly continuous functionals. One can prove the same result differently, and it also applies to the case where one has $J(\chi_1, \ldots, \chi_r, R) = \int_{\Omega} g(u)\big(A \, grad(u)\big)_j v \, dx$, with the same hypotheses on g and v. Because the sequences considered are such that $A^n \, grad(u_n)$ converges in $L^2(\Omega; R^N)$ weak to $A^{eff} \, grad(u_\infty)$, it is enough to show that $g(u_n)v$ converges in $L^2(\Omega)$ strong to $g(u_\infty)v$, and using another decomposition of v as $v_3 + v_4$ with v_3 small in $L^q(\Omega)$ and $v_4 \in L^\infty(\Omega)$, it is enough to show that $g(u_n)$ converges in $L^2(\Omega)$ strong to $g(u_\infty)$, and this follows from SOBOLEV imbedding theorem and the fact that the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact.

The next step is to write necessary conditions of optimality, and therefore one wants to use differentiable paths between the candidate to optimality and any other control. One assumes then that each F_i is continuously differentiable in u, that $\frac{\partial F_i(\cdot, u)}{\partial u}$ is a CARATHÉODORY function, and that some growth condition is satisfied so that

$$u \mapsto \frac{\partial F_i(\cdot, u)}{\partial u} \text{ is continuous from } H_0^1(\Omega) \text{ into } L^s(\Omega) \subset H^{-1}(\Omega), \qquad (9.11)$$

for example $\left|\frac{\partial F_i}{\partial u}\right| \leq g(x) + C|u|^q$ with $g \in L^r(\Omega)$ and $r \geq \frac{2N}{N+2}$ if $N \geq 3$ or $r > 1$ if $N = 2$ (and $r = 1$ for $N = 1$), and $q \leq \frac{N+2}{N-2}$ if $N \geq 3$ or $q < \infty$ if $N \leq 2$. The condition of optimality will be simplified by using an adjoint state, defined by

$$-div\Big(A^T \, grad(p)\Big) = \sum_{i=1}^{r} \eta_i \frac{\partial F_i(\cdot, u)}{\partial u} \text{ in } \Omega, \ p \in H_0^1(\Omega), \tag{9.12}$$

and although I am dealing with symmetric A here, I have used A^T in (9.12) in order to show what is needed in a nonsymmetric situation. In 1983, I was checking the conditions of optimality for the special case of mixing two anisotropic materials, using the characterization obtained with François MURAT, i.e. (7.25), (7.28) and (7.30), and as these conditions define a convex set of matrices, I could use straight segments for mixtures corresponding to the same local proportions. I discovered that the precise form of (7.28) and (7.30) is not really important, and generalizing to mixing more than two isotropic materials became an easy exercise, but I only understood the general case of mixing anisotropic materials in given proportions when I observed the property of Lemma 42. For the generalization, which I show directly without starting by the particular case that I had done first, I will use the notation

$$\mathcal{B}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r) = \{A | b \, I \leq A \leq a \, I\}$$

$$\frac{1}{b} = \sum_{i=1}^{r} \frac{\eta_i}{\min_j\{\lambda_j(M_i)\}} \tag{9.13}$$

$$a = \sum_{i=1}^{r} \max_j\{\lambda_j(M_i)\},$$

which defines a convex set $\mathcal{B}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ containing $\mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, but although they are different sets, Lemma 42 tells that some of their projections are identical.

Lemma 44: For $N \geq 2$, let $\eta_1^*, \ldots, \eta_r^*, A^*$ satisfy (9.9), and

$$\widetilde{J}(\eta_1^*, \ldots, \eta_r^*, A^*) \leq \widetilde{J}(\eta_1^*, \ldots, \eta_r^*, A) \text{ for all } A \in \mathcal{K}(\eta_1^*, \ldots, \eta_r^*, M_1, \ldots, M_r), \tag{9.14}$$

with $\widetilde{J}$ defined by (9.10). Then, if u^* and p^* are the corresponding solutions of respectively (9.3) and (9.12), one has

$$\int_{\Omega}\Big(A^* \, grad(u^*).grad(p^*)\Big)\, dx \geq \int_{\Omega}\Big(B \, grad(u^*).grad(p^*)\Big)\, dx \text{ for all } B \in \mathcal{B}(\eta_1^*, \ldots, \eta_r^*, M_1, \ldots, M_r). \tag{9.15}$$

Proof: For $B \in \mathcal{B}(\eta_1^*, \ldots, \eta_r^*; M_1, \ldots, M_r)$, and $\varepsilon \in (0,1)$, one defines

$$A(\varepsilon) = (1 - \varepsilon)A^* + \varepsilon \, B = A^* + \varepsilon \, \delta A, \tag{9.16}$$

and $A(\varepsilon) \in \mathcal{B}(\eta_1^*, \ldots, \eta_r^*; M_1, \ldots, M_r)$ because it is convex and contains $\mathcal{K}(\eta_1^*, \ldots, \eta_r^*, M_1, \ldots, M_r)$. As $A(\varepsilon)$ is analytic in ε with values in $M(\alpha, \beta; \Omega)$, the operator $\mathcal{A}_\varepsilon = -div\big(A(\varepsilon)grad(\cdot)\big)$ is analytic with values in $\mathcal{L}\big(H_0^1(\Omega), H^{-1}(\Omega)\big)$ and invertible by LAX-MILGRAM lemma, and therefore its inverse $\mathcal{A}_\varepsilon^{-1}$ is analytic with values in $\mathcal{L}\big(H^{-1}(\Omega), H_0^1(\Omega)\big)$, so that the corresponding solution $u(\varepsilon)$ is analytic in ε with values in $H_0^1(\Omega)$, and its derivative δu at $\varepsilon = 0$ satisfies

$$u(\varepsilon) = u^* + \varepsilon \, \delta u + o(\varepsilon) \text{ in } H_0^1(\Omega)$$

$$div\Big(A^* \, grad(\delta u) + \delta A \, grad(u^*)\Big) = 0, \tag{9.17}$$

where $\delta A = B - A^*$. The function $\widetilde{J}\big(\eta_1^*, \ldots, \eta_r^*, A(\varepsilon)\big)$ is therefore differentiable in ε, and its derivative $\delta \widetilde{J}$ at $\varepsilon = 0$ satisfies

$$\widetilde{J}\Big(\eta_1^*, \ldots, \eta_r^*, A(\varepsilon)\Big) = \widetilde{J}(\eta_1^*, \ldots, \eta_r^*, A^*) + \varepsilon \, \delta\widetilde{J} + o(\varepsilon)$$

$$\delta\widetilde{J} = \int_{\Omega}\Big(\sum_{i=1}^{r}\eta_i^* \frac{\partial F_i(x, u^*)}{\partial u}\Big)\delta u \, dx, \tag{9.18}$$

61

and using the definition (9.12) of the adjoint state p^* and (9,17), one finds

$$\delta \widetilde{J} = \int_\Omega \left((A^*)^T grad(p^*).grad(\delta u) \right) dx = \int_\Omega \left(A^* grad(\delta u).grad(p^*) \right) dx$$
$$= -\int_\Omega \left(\delta A \, grad(u^*).grad(p^*) \right) dx = \int_\Omega \left((A^* - B) \, grad(u^*).grad(p^*) \right) dx. \tag{9.19}$$

If one had $A(\varepsilon) \in \mathcal{K}(\eta_1^*, \ldots, \eta_r^*, M_1, \ldots, M_r)$ for all $\varepsilon \in (0,1)$, one would have $\widetilde{J}\left(\eta_1^*, \ldots, \eta_r^*, A(\varepsilon) \right) \geq \widetilde{J}(\eta_1^*, \ldots, \eta_r^*, A^*)$ by (9.14), and therefore $\delta \widetilde{J} \geq 0$ by (9.18), and that would give (9.15) for the particular choice of B. However, almost everywhere in Ω, $A(\varepsilon)grad\left(u(\varepsilon)\right)$ belongs to the closed ball with diameter $[b \, grad\left(u(\varepsilon)\right), a \, grad\left(u(\varepsilon)\right)]$, and therefore by Lemma 42

there exists $M(\varepsilon) \in \mathcal{K}(\eta_1^*, \ldots, \eta_r^*, M_1, \ldots, M_r) : \ A(\varepsilon)grad\left(u(\varepsilon)\right) = M(\varepsilon)grad\left(u(\varepsilon)\right)$ a.e. in Ω. (9.20)

Because $A(\varepsilon)$ and $M(\varepsilon)$ create then the same state $u(\varepsilon)$, one has

$$\widetilde{J}\left(\eta_1^*, \ldots, \eta_r^*, A(\varepsilon) \right) = \widetilde{J}\left(\eta_1^*, \ldots, \eta_r^*, M(\varepsilon) \right) \geq \widetilde{J}(\eta_1^*, \ldots, \eta_r^*, A^*), \tag{9.21}$$

and therefore the conclusion $\delta \widetilde{J} \geq 0$ is valid.∎

In the derivation of (9.20), there is a small technical difficulty, because Lemma 42 was given without any dependence upon $x \in \Omega$, and one must then check that there is a measurable $M(\varepsilon)$ satisfying (9.20) ($\varepsilon > 0$ being fixed). Of course, one can make Lemma 42 more precise by constructing an explicit lifting, which maps $\eta_1, \ldots, \eta_r, B \in \mathcal{B}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r), e \in R^N$ to $A \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ with $Ae = Be$, but a natural construction relies on different cases, e being 0 or not, $B\,e$ being parallel to e or not, $B\,e$ being on the sphere with diameter $[b(\eta)e, a(\eta)e]$ or not, and so on, and in each case one only checks continuity of the lifting, so that the constructed A is measurable. One may also avoid this question of measurability altogether by restricting one's attention to specific mixtures obtained by layering in arbitrary directions, and in the end it gives the same condition (because in Lemma 42 the elements on the boundary of the ball are created by layerings). One chooses an orthonormal basis $e_1, \ldots, e_N$, and one orients material $\#i$ so that e_j is an eigenvector for the eigenvalue $\lambda_j(M_i)$, with eigenvalues increasing with j (so that $\lambda_1(M_j) = \min_j \lambda_j(M_i)$ and $\lambda_N(M_i) = \max_j \lambda_j(M_i)$), and then one uses layers orthogonal to e_1, with proportions $\eta_1^*, \ldots, \eta_r^*$, and one obtains a material with tensor

$$P = \sum_{j=1}^N \pi_j \, e_j \otimes e_j \in \mathcal{K}(\eta_1^*, \ldots, \eta_r^*; M_1, \ldots, M_r)$$
$$b = \pi_1 \leq \ldots \leq \pi_N = a \text{ a.e. in } \Omega. \tag{9.22}$$

For a measurable subset ω of Ω one keeps $A(\varepsilon) = A^*$ in $\Omega \backslash \omega$, and in ω one defines $A(\varepsilon)$ by mixing the material with tensor A^* and the material with tensor P, in layers orthogonal to e and with respective proportions $1 - \varepsilon$ and ε. Of course, one uses Lemma 40 for showing that such an $A(\varepsilon)$ is allowed in (9.14), and formula (8.17) gives

$$A(\varepsilon) = A^* + \varepsilon \, \delta A + o(\varepsilon) \text{ in } \omega \text{ with } \delta A = P - A^* - (P - A^*)\frac{e \otimes e}{(P\,e.e)}(P - A^*), \tag{9.23}$$

and this different definition of δA corresponds to a different δu to use in (9.17) and (9.18), and the last equality in (9.19) changes because of the different definition of δA, and by varying ω one deduces that (9.9) and (9.14) imply that for all $e \neq 0$ one has

$$\left(A^* grad(u^*).grad(p^*) \right) \geq \left(P \, grad(u^*).grad(p^*) \right) - \frac{\left((P - A^*)grad(u^*).e \right)\left((P - A^*)grad(p^*).e \right)}{(P\,e.e)}, \tag{9.24}$$

a.e. in Ω, and by choosing $e \neq 0$ orthogonal to $(P - A^*)grad(u^*)$, one finds

$$\Big(A^*\,grad(u^*).grad(p^*)\Big) \geq \Big(P\,grad(u^*).grad(p^*)\Big) \text{ for all } P \text{ satisfying (9.22), a.e. in } \Omega. \qquad (9.25)$$

From (9.15) or (9.25) one deduces the following necessary condition of optimality.

Proposition 45: For $N \geq 2$, let $\eta_1^*, \ldots, \eta_r^*, A^*$ satisfy (9.9) and (9.14), then at almost every point where $|grad(u^*)|\,|grad(p^*)| \neq 0$ one has

$$\text{if } grad(p^*) = c\,grad(u^*) \text{ with } c > 0, \text{ then } A^*\,grad(u^*) = a\,grad(u^*), A^*\,grad(p^*) = a\,grad(p^*), \quad (9.26)$$

$$\text{if } grad(p^*) = c\,grad(u^*) \text{ with } c < 0, \text{ then } A^*\,grad(u^*) = b\,grad(u^*), A^*\,grad(p^*) = b\,grad(p^*), \quad (9.27)$$

$$\text{if } grad(p^*) \text{ is not parallel to } grad(u^*) \text{ then}$$
$$A^*\,grad(u^*) = \frac{a+b}{2}grad(u^*) + \frac{a-b}{2}\frac{|grad(u^*)|}{|grad(p^*)|}grad(p^*)$$
$$A^*\,grad(p^*) = \frac{a+b}{2}grad(p^*) + \frac{a-b}{2}\frac{|grad(p^*)|}{|grad(u^*)|}grad(u^*). \qquad (9.28)$$

Proof: As $|grad(u^*)| \neq 0$, when one changes the basis $e_1, \ldots, e_N$, the vector $P\,grad(u^*)$ spans the sphere with diameter $[b\,grad(u^*), a\,grad(u^*)]$, and the quantity $\Big(P\,grad(u^*).grad(p^*)\Big)$ attains its maximum at a unique point of the sphere; as $A^*\,grad(u^*)$ belongs to the closed ball with diameter $[b\,grad(u^*), a\,grad(u^*)]$, $A^*\,grad(u^*)$ must be equal to the value of $P\,grad(u^*)$ for the P for which the maximum is attained, and this gives the formulas for $A^*\,grad(u^*)$ in (9.26)/(9.28). The corresponding formulas for $A^*\,grad(p^*)$ follow by a simple remark from Linear Algebra, applied to $A^* - \frac{a+b}{2}I$, namely that if M is symmetric with norm $\leq \gamma$ and if for a vector $E \neq 0$ one has $M\,E = F$ and $|F| = \gamma|E|$, then $M\,F = \gamma^2\,E$ (because $|M\,F - \gamma^2\,E|^2 = |M\,F|^2 - 2\gamma^2(F.M\,E) + \gamma^4|E|^2 \leq \gamma^2(\gamma^2|E|^2 - |F|^2|) = 0$).∎

On the set Ω_+ where neither $grad(u^*)$ nor $grad(p^*)$ vanish, one may therefore replace A^* by P which is obtained by layering the materials with the same proportions $\eta_1^*, \ldots, \eta_r^*$, and one has the same cost because one can keep the same u^*, as $A^*\,grad(u^*) = P\,grad(u^*)$ a.e. in Ω_+. On the set Ω_u where $grad(u^*) = 0$, one can also replace A^* by any material without changing $A^*\,grad(u^*)$ (which remains 0), and in particular one can replace A^* by a material obtained by layering the materials with the same proportions. However, the situation is different for the set Ω_p where $grad(p^*) = 0$ and $grad(u^*) \neq 0$, but one can adapt an argument of RAITUM [Ra] for replacing A^* by materials obtained by layering but with different proportions (if $A^*\,grad(u^*)$ is not on the sphere with diameter $[b\,grad(u^*), a\,grad(u^*)]$, it may only belong to the boundary of another sphere obtained by changing b or a, and therefore one needs to change some of the proportions). Let

$$\Omega_c = \Big\{x|x \in \Omega, \Big(A^*grad(u^*) - b\,grad(u^*).A^*grad(u^*) - a\,grad(u^*)\Big) < 0\Big\}, \qquad (9.29)$$

i.e. the set where $A^*grad(u^*)$ cannot be obtained by a layered material so that $grad(u^*) \neq 0$, and the preceding analysis shows that on Ω_c one must have $grad(p^*) = 0$, i.e. $\Omega_c \subset \Omega_p$. One defines $b(\eta)$ and $a(\eta)$ for

$$\eta = (\eta_1, \ldots, \eta_r) \in L^\infty(\Omega_c; R^r) \text{ such that } \eta_i \geq 0 \text{ a.e. in } \Omega_c \text{ for } i = 1, \ldots, r, \sum_{i=1}^{r}\eta_i = 1 \text{ a.e. in } \Omega_c, \quad (9.30)$$

by the formulas

$$\frac{1}{b(\eta)} = \sum_{i=1}^{r}\frac{\eta_i}{\min_j \lambda_j(M_i)}, \ a(\eta) = \sum_{i=1}^{r}\eta_i \max_j \lambda_j(M_i), \qquad (9.31)$$

and one imposes the two constraints

$$\Big(A^*grad(u^*) - a(\eta)grad(u^*).A^*grad(u^*) - b(\eta)grad(u^*)\Big) \leq 0 \text{ a.e. in } \Omega_c, \qquad (9.32)$$

$$\int_{\Omega_c} \eta_i \, dx = \int_{\Omega_c} \eta_i^* \, dx, i = 1, \ldots, r, \tag{9.33}$$

and one wants to minimize

$$J_c(\eta) = \int_{\Omega_c} \left(\sum_{i=1}^{r} \eta_i \, F_i(x, u^*) \right) dx. \tag{9.34}$$

The set of η satisfying (9.30), (9.32)/(9.33) contains η^* (as (9.32) is true by Lemma 42). The condition (9.30) defines a convex set, which is $L^\infty(\Omega_c; R^r)$ weak $\star$ compact; the constraint (9.32) defines a $L^\infty(\Omega_c; R^2)$ weak $\star$ closed convex set in $\left(a, \frac{1}{b}\right)$ by Lemma 19, and by (9.31) the set of η satisfying (9.32) is therefore convex, and $L^\infty(\Omega_c; R^r)$ weak $\star$ closed, and similarly for the constraint (9.33). As J_c is linear and $L^\infty(\Omega_c; R^r)$ weak $\star$ continuous, it attains its minimum on at least one extreme point of the $L^\infty(\Omega_c; R^r)$ weak $\star$ compact convex set defined by (9.30) and (9.32)/(9.33). One shows then that for such an extreme point one must have equality in (9.32) a.e. in Ω_c, by using an argument of Zvi ARTSTEIN [Ar]. Indeed, assume that $\left(A^* grad(u^*) - a(\eta)grad(u^*).A^* grad(u^*) - b(\eta)grad(u^*)\right) \leq -\varepsilon < 0$ on a subset $K \subset \Omega_c$ with positive measure; one may assume that one can find $r + 1$ disjoint subsets $K_k \subset K$ with positive measure and on each such set K_k one can find two distinct indices $i(k), j(k) \in \{1, \ldots, r\}$ such that $\eta_{i(k)}, \eta_{j(k)} \geq \delta_k > 0$ a.e. in K_k (if it was not true, a.e. in K_k there would only be one η_i different from 0 and thus equal to 1, in which case $a(\eta) = b(\eta)$ and one could not be in Ω_c); for $|c_k| \leq \delta_k$, one can add c_k to $\eta_{i(k)}$ in K_k and subtract c_k to $\eta_{j(k)}$ in K_k and still satisfy (9.30), and by restricting a little more $|c_k|$ one can still satisfy (9.32); then one has $r + 1$ small arbitrary constants that one can play with and only r linear constraints (9.33) that they must satisfy, and it leaves the possibility to move both ways in at least one direction while satisfying all the constraints, contradicting the assumption that one has choosen an extreme point. Therefore one has found an optimal solution of the initial problem which can be realized by layerings inside Ω_c.

At least one solution of the optimization problem corresponds then to a material obtained by layerings (with varying proportions and directions), and therefore at the end the Homogenization questions seem to disappear from the analysis, and one may wonder if it was really necessary to study Homogenization in the first place.

There is an analogous question about Functional Analysis. One wants to show the existence of a characteristic function $\chi \in L^\infty(\Omega)$ which minimizes $\int_\Omega \chi f \, dx$, where f is given in $L^1(\Omega)$ and where the minimization is only considered for characteristic functions of measurable subsets of Ω satisfying a finite number of constraints $\int_\Omega \chi g_i \, dx = \alpha_i$ for $i = 1, \ldots, m$, where $g_1, \ldots, g_m$ are given in $L^1(\Omega)$; of course, one assumes that at least one such χ exists. The proof of existence that I know was taught to me by Zvi ARTSTEIN in 1975 [Ar],[69] and it consists in minimizing $L(\theta) = \int_\Omega \theta f \, dx$ among functions $\theta \in C = \{\theta | \theta \in L^\infty(\Omega), 0 \leq \theta \leq 1$ a.e. in $\Omega, \int_\Omega \theta g_i \, dx = \alpha_i$ for $i = 1, \ldots, m\}$; as C is a nonempty compact set for the $L^\infty(\Omega)$ weak $\star$ topology (which is metrizable when restricted to C), and L is continuous for this topology, the minimum of L is attained, and because C is convex and L is affine, the minimum is actually attained on a (nonempty) convex compact subset C_0 of C, and C_0 has at least one extreme point θ_0 by KREIN-MILMAN theorem. Finally, Zvi ARTSTEIN argues that θ_0 belongs to a finite dimensional face of the bigger convex set $C_1 = \{\theta | \theta \in L^\infty(\Omega), 0 \leq \theta \leq 1$ a.e. in $\Omega\}$, and using the fact that the LEBESGUE measure has no atoms, he shows that θ_0 must be a characteristic function, which is therefore optimal among characteristic functions from C.

In both situations, one wants to minimize on a set X some function F_i, belonging to a family of such functions indexed by $i \in I$, and one has constructed a compact space $\widehat{X}$ containing X as a dense subspace, and each F_i becomes the restriction to X of a function $\widehat{F}_i$ which is continuous on $\widehat{X}$. Of course, each function $\widehat{F}_i$ attains its minimum on $\widehat{X}$, but one exhibits a subset Y such that $X \subset Y \subset \widehat{X}$ (and equal to X in the second case), which has the property that for each $i \in I$ the set of minima of $\widehat{F}_i$ intersects Y. Of course, as X is dense in $\widehat{X}$, Y is also dense in $\widehat{X}$, and therefore not compact if $Y \neq \widehat{X}$. It is not clear if there exists another topology for which Y is compact with X dense in Y and each F_i is the restriction to X of a function G_i which is lower semi-continuous on Y.

I do not know any obvious abstract framework which explains how to avoid mentioning Functional Analysis or Homogenization in these examples, and although it could be useful to develop special results for

---

[69] The following argument is precisely Zvi ARTSTEIN's proof of LYAPUNOV's theorem.

the growing number of students who hope to use only elementary tools in Mathematics, it seems impossible to maintain such a goal if one is interested in practical problems. It is indeed often emphasized by those who have dealt with practical problems of Optimization, that one is rarely given a very precise function to minimize, and that one must reassert the goal in view of preliminary results. For what concerns Homogenization appearing in the problems that I have been considering, an obvious remark is that in the realization of such an optimal solution including mixtures in some areas, one needs to reconsider the function to be minimized by adding the cost of creating these mixtures.[70]

Actually, as was pointed out later by Robert KOHN, there is another natural class of functionals where the optimal solution seems to require more precise information on optimal bounds, and more general designs than simple layerings. In identification problems, one does a few experiments with the same arrangement of materials, and one tries to estimate some coefficient by using a finite number of measurements. As a simple model of this kind, one considers q equations of the form

$$-div\Big(A\,grad(u_i)\Big) = f_i \text{ in } \Omega, \ u_i \in H_0^1(\Omega), \text{ for } i = 1, \ldots, q, \tag{9.35}$$

and all these equations use the same $A \in M(\alpha, \beta; \Omega)$ but different functions $f_i \in H^{-1}(\Omega), i = 1, \ldots, q$, and q functions $v_i \in H_0^1(\Omega)$ (or simply $v_i \in L^2(\Omega)$) are given for $i = 1, \ldots, q$, corresponding to measurements in a material using an unknown value A_0 that one wants to identify.[71] One idea would be to minimize a cost function J of the form

$$J(A) = \int_\Omega \Big(\sum_{i=1}^q |u_i - v_i|^2\Big)\,dx, \tag{9.36}$$

and if one knows that A is one of the materials $M_1, \ldots, M_r$ rotated in an arbitrary way (a.e. in Ω), one gets a problem where the Homogenization approach will appear, but if one considers functionals of the form

$$J_1(A) = \int_\Omega \Big(\sum_{i=1}^q |grad(u_i - v_i)|^2\Big)\,dx, \tag{9.37}$$

the situation is quite different [Ta13]. In the case of (9.36), minimizing sequences will make some $A^{eff} \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ appear, and there exists then an optimal homogenized A, and one may wonder what its relation with the actual A_0 is, for example if in the case where $M_i = \mu_i\,I$ for $i = 1, \ldots, r$ and therefore A_0 is locally an isotropic material, but one finds that all the optimal A are anisotropic in some regions. Of course, if the measured values were exactly those associated with A_0 when one uses (9.35), then one would have $J(A_0) = 0$, but measurements are not always entirely accurate and this explains why, even when A_0 only takes isotropic values, it may do so on relatively small pieces, and A_0 may well be near an anisotropic A in a distance corresponding to H-convergence (mentioned after Definition 5). If one decides

---

[70] It is unfortunately often the case that some people dealing with questions in Elasticity not only forget to mention the defects of Linearized Elasticity but only consider extremely particular functionals. On the contrary, on the engineering side, as was emphasized by Martin BENDSØE at a meeting in Trieste in 1993, it is important to have an interactive point of view when dealing with Optimal Design problems; one rarely needs to compute with high precision which mixtures will appear in connection with minimizing a particular functional, because one may well abandon this particular functional during the interactive part of the procedure, and one may end up minimizing something different, in principle more adapted to the engineering application.

[71] Practical problems may include elecrical or heat conduction questions or permeability of oil reservoirs, and they do not correspond to homogeneous DIRICHLET data. Instead of using different f_i one may use different nonhomogeneous DIRICHLET data on some parts of the boundary and NEUMANN data on other parts of the boundary, and the different measurements may give the value of u_i or the flux $(A\,grad(u_i).n)$ on various other parts of the boundary. Another important class of problems arising in applications deals with eigenvalues, in which case one may measure a few of the lowest eigenvalues corresponding to some unknown A_0, which one tries to identify from these measurements. It is not difficult however to adapt the methods described in this course (for some academic models) in order to deal with these more realistic variants.

then to minimize (9.36) and one wants to write necessary conditions of optimality, an increase δA creates increases δu_i, $i = 1, \ldots, q$, solutions of

$$-div\Big(A\,grad(\delta u_i) + \delta A\,grad(u_i) \Big) = 0 \text{ in } \Omega, \ \delta u_i \in H_0^1(\Omega), i = 1, \ldots, q, \tag{9.38}$$

and an increase δJ given by

$$\delta J = \int_\Omega 2\Big(\sum_{i=1}^q (u_i - v_i)\delta u_i \Big)\, dx, \tag{9.39}$$

and in order to express δJ in terms of δA, one introduces q adjoint states $p_1, \ldots, p_q$, solutions of

$$-div\Big(A\,grad(p_i) \Big) = u_i - v_i \text{ in } \Omega, \ p_i \in H_0^1(\Omega), i = 1, \ldots, q, \tag{9.40}$$

and the necessary condition of optimality becomes

$$\delta J = -2 \int_\Omega \Big[\sum_{i=1}^q \Big(\delta A\,grad(u_i).grad(p_i) \Big) \Big]\, dx \geq 0 \text{ for admissible } \delta A. \tag{9.41}$$

In the first step of the obtention of necessary conditions that I have described, i.e. keeping the proportions $\eta_1^*, \ldots, \eta_r^*$ fixed, one sees that it would be useful to characterize the set of $\big(A\,grad(u_1), \ldots, A\,grad(u_q) \big)$ for $A \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, but for $q \geq 2$ the analogue of Lemma 42 is not known. Even for the case of mixing two isotropic materials, i.e. $M_1 = \alpha\,I, M_2 = \beta\,I$, for which I had obtained the characterization (7.25), (7.28), (7.30) of $\mathcal{K}(\theta, 1 - \theta; \alpha\,I, \beta\,I)$ with François MURAT in 1980, I do not know a simple characterization of this set, although Lemma 41 asserts that it is convex if $q \leq N - 1$. It was for this case that Robert KOHN had pointed out that the full characterization seems necessary, but I do not know if the necessary condition (9.41) has given rise to a precise analysis of what optimal solutions look like.

10. Necessary conditions of optimality: second step

We can now look at the second step of the derivation of necessary conditions of optimality, which consists in varying $\eta_1, \ldots, \eta_r$, and is therefore more classical. I begin by treating two very special examples, corresponding to $p = u$ or to $p = -u$, where one can avoid Homogenization almost completely. The special examples correspond to minimizing either the cost function J_1 or the cost function J_2 defined by

$$\begin{aligned} J_1(\chi_1, \ldots, \chi_r, R) &= \int_\Omega f\,u\,dx \\ J_2(\chi_1, \ldots, \chi_r, R) &= -\int_\Omega f\,u\,dx, \end{aligned} \tag{10.1}$$

and it is important that in (10.1) one uses the same function f which appears in (9.3). The reason why these two functionals are special is that the definition (9.12) of the adjoint state in the case of minimizing J_1 gives $p = u$, and therefore $A^*\,grad(u^*) = a(\eta^*)grad(u^*)$ by (9.26) for $N \geq 2$, while in the case of minimizing J_2 it gives $p = -u$, and therefore $A^*\,grad(u^*) = b(\eta^*)grad(u^*)$ by (9.27) (which is valid for $N \geq 1$, and also for minimizing J_1 in the case $N = 1$), where the functions a and b are defined by (9.13), repeated as (9.31). Another important fact for the argument is that, because A is symmetric a.e. in Ω, solving (9.3) amounts to a minimization problem, namely

$$Min_{v \in H_0^1(\Omega)} \int_\Omega \Big[\Big(A\,grad(v).grad(v) \Big) - 2f\,v \Big]\, dx \text{ is attained for } v = u \text{ defined by (9.3),}$$

$$\text{and the value of the minimum is } -\int_\Omega f\,u\,dx. \tag{10.2}$$

Lemma 46: The minimization of J_1 is equivalent to the following min-max problem

$$Max_\eta Min_{v \in H_0^1(\Omega)} \int_\Omega \Big(a(\eta)|grad(v)|^2 - 2f\,v \Big)\, dx \text{ if } N \geq 2, \tag{10.3}$$

$$Max_\eta Min_{v \in H_0^1(\Omega)} \int_\Omega \left(b(\eta)|grad(v)|^2 - 2f\,v \right) dx \text{ if } N = 1, \tag{10.4}$$

the set of η on which one maximizes is given by (9.6), repeated in (9.9), and in both cases u^* is defined in a unique way.

Proof: As already mentioned, for any optimal solution u^* one has $p^* = u^*$, and therefore one does not need the argument of RAITUM described after Proposition 45. For $N \geq 2$ one has (9.26) at points where $grad(u^*) \neq 0$, and as it holds automatically at points where $grad(u^*) = 0$, the formula (10.2) gives (10.3). For $N = 1$ one automatically has $A^{eff} = b(\eta)I$ and therefore (9.27) holds, and the formula (10.2) gives (10.4).

That the problems in (10.3) or (10.4) have at least one solution follows from classical min-max theorems, because by using the information $A \in M(\alpha, \beta; \Omega)$ one may restrict the minimization in v to a large enough closed ball of $H_0^1(\Omega)$, which is a compact convex set for the weak topology of $H_0^1(\Omega)$, the set of η is a compact convex set for the $L^\infty(\Omega; R^r)$ weak $\star$ topology, and the functional is convex lower semi-continuous in v and concave upper semi-continuous in η (it is actually linear continuous in the case $N \geq 2$). Moreover, because the functional is strictly convex in v, the solution u^* is unique. ∎

Of course, if without knowing anything about Homogenization one has guessed that the problem (10.3) is a way to solve the initial problem of minimizing J_1, one still has to use Homogenization in order to interpret what a solution η^* means, and that it can be created by layerings. Jean CÉA and K. MALANOWSKI had unknowingly taken advantage of a similar miracle in [Cé&Ma], but they had started from accepting all possible γI with $\alpha \leq \gamma \leq \beta$, and they did not even have to explain by Homogenization the (classical) solution that they had obtained.

Lemma 47: The minimization of J_2 is equivalent to the following minimization problem

$$Min_\eta Min_{v \in H_0^1(\Omega)} \int_\Omega \left(b(\eta)|grad(v)|^2 - 2f\,v \right) dx, \tag{10.5}$$

and the set of η on which one minimizes is given by (9.6), repeated in (9.9), and $b(\eta^*)grad(u^*)$ is defined in a unique way.

Proof: As already mentioned, for any optimal solution u^* one has $p^* = -u^*$, and therefore one does not need the argument of RAITUM described after Proposition 45. For $N \geq 2$ one has (9.27) at points where $grad(u^*) \neq 0$, and as it holds automatically at points where $grad(u^*) = 0$, the formula (10.2) gives (10.5).

That the problem in (10.5) has a solution follows from classical theorems in Optimization, because by using the information $A \in M(\alpha, \beta; \Omega)$ one may restrict the minimization in v to a large enough closed ball of $H_0^1(\Omega)$, which is a compact convex set for the weak topology of $H_0^1(\Omega)$, the set of η is a compact convex set for the $L^\infty(\Omega; R^r)$ weak $\star$ topology, and the functional is convex in (v, η) and lower semi-continuous by Lemma 11 with $M = b(\eta)\,I$; moreover, as $P = M^{-1}$ in formula (5.6), one sees that the functional is strictly convex in $b(\eta)grad(v)$ and therefore $b(\eta^*)grad(u^*)$ is unique. ∎

Of course, if without knowing anything about Homogenization one has guessed that the problem (10.5) is a way to solve the initial problem of minimizing J_2, one still has to use Homogenization in order to interpret what a solution η^* means, and that it can be created by layerings.

These two examples are not always instances of an intermediate problem in the abstract setting mentioned after (9.34), and they do not always fit the framework of relaxed problems described in chapter 3 either, because the set X for which the minimization of J_1 or J_2 is considered (which is the set of characteristic functions $\chi_1, \ldots, \chi_r$ of disjoints sets satisfying (9.2), together with a measurable rotation R so that A is defined by (9.1)) is not always included in the set Y on which (10.3) or (10.4) are considered (which is the set of $\eta_1, \ldots, \eta_r$ satisfying (9.6), and A is $a(\eta)I$ for minimizing J_1 for $N \geq 2$, or $b(\eta)I$ for minimizing J_1 for $N = 1$ or for minimizing J_2). X can only be considered as a subset of Y if one can avoid the use of the rotation R, i.e. in the case where $M_i = \mu_i I$ for $i = 1, \ldots, r$; however, even in this case, the problem is not a completion, except for $N = 1$, where there is a formula for A^{eff}, as $A^{eff} = b(\eta)$.

Coming back to the general case, one wants now to write that the optimal mixture using the proportions $\eta_1^*, \ldots, \eta_r^*$ (with A^* corresponding to a layered material) is better than any other mixture using different proportions $\eta_1, \ldots, \eta_r$, and for this task one follows the arguments used in Lemma 44 (or the variant before

Proposition 45), but one starts now from $A \in \mathcal{K}(\eta_1, \dots, \eta_r; M_1, \dots, M_r)$, with proportions $\eta_1, \dots, \eta_r$ satisfying (9.9). By mixing then A with A^* in an adapted way one will create an admissible differentiable path $\eta(\varepsilon) = \eta^* + \varepsilon \, \delta\eta + o(\varepsilon)$, $A(\varepsilon) = A^* + \varepsilon \, \delta A + o(\varepsilon)$, $u(\varepsilon) = u^* + \varepsilon \, \delta u + o(\varepsilon)$, $\widetilde{J}\big(\eta(\varepsilon), A(\varepsilon)\big) = \widetilde{J}(\eta^*, A^*) + \varepsilon \delta\widetilde{J} + o(\varepsilon)$, with δu and δA still related by (9.17), but with $\delta\widetilde{J}$ given now by

$$\delta\widetilde{J} = \int_\Omega \Big[\sum_{i=1}^r \delta\eta_i \, F_i(x, u^*) + \Big(\sum_{i=1}^r \eta_i^* \frac{\partial F_i(x, u^*)}{\partial u} \Big) \delta u \Big] \, dx, \tag{10.6}$$

and using the definition (9.12) of the adjoint state p^* and (9.17), one finds

$$\delta\widetilde{J} = \int_\Omega \Big[\sum_{i=1}^r \delta\eta_i \, F_i(x, u^*) - \Big(\delta A \, grad(u^*).grad(p^*) \Big) \Big] \, dx. \tag{10.7}$$

Proposition 48: Let $\eta_1^*, \dots, \eta_r^*, A^*$ satisfy (9.9) and

$$\widetilde{J}(\eta_1^*, \dots, \eta_r^*, A^*) \le \widetilde{J}(\eta_1, \dots, \eta_r, A) \text{ for all } \eta_1, \dots, \eta_r, A \text{ satisfying (9.9),} \tag{10.8}$$

with $\widetilde{J}$ defined by (9.10). For $N \ge 2$, (10.8) implies

$$\int_\Omega \Big[\sum_{i=1}^r \eta_i^* \, F_i(x, u^*) - \Big(A^* \, grad(u^*).grad(p^*) \Big) \Big] \, dx \le \int_\Omega \Big[\sum_{i=1}^r \eta_i \, F_i(x, u^*) - \Big(B \, grad(u^*).grad(p^*) \Big) \Big] \, dx$$

for all $\eta_1, \dots, \eta_r$ satisfying (9.9), $B \in \mathcal{B}(\eta_1, \dots, \eta_r; M_1, \dots, M_r)$,

$$\tag{10.9}$$

with $u^*, p^*, \mathcal{B}(\eta_1, \dots, \eta_r; M_1, \dots, M_r)$ defined by (9.3) (using A^*), (9.12) (using A^*), and (9.13). For $N = 1$, one has $a^* = b(\eta^*)$ and (10.8) implies

$$\int_\Omega \Big[\sum_{i=1}^r \Big(F_i(x.u^*) + \frac{(a^*)^2}{\lambda(M_i)} \frac{d\,u^*}{dx} \frac{d\,p^*}{dx} \Big) \eta_i^* \Big] \, dx \le \int_\Omega \Big[\sum_{i=1}^r \Big(F_i(x.u^*) + \frac{(a^*)^2}{\lambda(M_i)} \frac{d\,u^*}{dx} \frac{d\,p^*}{dx} \Big) \eta_i \Big] \, dx \tag{10.10}$$

for all $\eta_1, \dots, \eta_r$ satisfying (9.9).

Proof: Let $\eta_1, \dots, \eta_r$ satisfy (9.9), and $B \in \mathcal{B}(\eta_1, \dots, \eta_r; M_1, \dots, M_r)$. For $\varepsilon \in (0, 1)$, one defines $\eta(\varepsilon)$ by

$$\eta_i(\varepsilon) = (1 - \varepsilon)\eta_i^* + \varepsilon \, \eta_i, \tag{10.11}$$

so that $\eta(\varepsilon)$ satisfies (9.9). Using Lemma 42 (and preceding remarks for constructing measurable liftings), there exists A such that

$$A \, grad(u^*) = B \, grad(u^*), \ A \in \mathcal{K}(\eta_1, \dots, \eta_r; M_1, \dots, M_r), \ \text{a.e.} \ x \in \Omega. \tag{10.12}$$

For a nonvanishing $e \in L^\infty(\Omega; R^N)$ one defines $A(\varepsilon)$ by layering A^* and A with respective proportions $1 - \varepsilon$ and ε, in layers orthogonal to e, and Lemma 34 gives

$$A(\varepsilon) = (1 - \varepsilon)A^* + \varepsilon \, A - \varepsilon(1 - \varepsilon)(A - A^*) \frac{e \otimes e}{\varepsilon (A^* \, e.e) + (1 - \varepsilon)(A \, e.e)} (A - A^*), \tag{10.13}$$

and by construction one has $A(\varepsilon) \in \mathcal{K}(\eta_1(\varepsilon), \dots, \eta_r(\varepsilon); M_1, \dots, M_r)$ a.e. in Ω. Of course, (10.8) implies that $\delta\widetilde{J} \ge 0$, for which one uses (10.7); the value of $\delta\eta_i$ following from (10.11) is $\delta\eta_i = \eta_i - \eta_i^*$, and the value of δA following from (10.13) is $\delta A = A - A^* - (A - A^*) \frac{e \otimes e}{(A\,e.e)} (A - A^*)$, so that $\big(\delta A \, grad(u^*).grad(p^*) \big) = \big((A - A^*)grad(u^*).grad(p^*) \big) - \frac{1}{(A\,e.e)} \big((A - A^*)grad(u^*).e \big) \big((A - A^*)e.grad(p^*) \big)$, and by choosing e orthogonal to $(A - A^*)grad(u^*)$, one has $\big(\delta A \, grad(u^*).grad(p^*) \big) = \big((A - A^*)grad(u^*).grad(p^*) \big)$, which by (10.12) is $\big((B - A^*)grad(u^*).grad(p^*) \big)$, and $\delta\widetilde{J} \ge 0$ is then exactly (10.9).

68

In the case $N = 1$, $A \in \mathcal{K}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ means $A = b(\eta)$, with b defined in (9.13), and therefore $\delta A = \sum_i \frac{\partial b}{\partial \eta_i} \delta \eta_i$; (9.13) implies that $\frac{\partial b}{\partial \eta_i} = -\frac{b^2}{\lambda(M_i)}$ (where $\lambda(M_i)$ is the value of the coefficient for the material M_i), and therefore one has

$$\delta \widetilde{J} = \int_\Omega \Big[\sum_{i=1}^r \Big(F_i(x.u^*) + \frac{(a^*)^2}{\lambda(M_i)} \frac{d\,u^*}{dx} \frac{d\,p^*}{dx} \Big) \delta \eta_i \Big] \, dx, \tag{10.14}$$

from which one immediately deduces (10.10). ∎

Inequality (10.9) consists in minimizing a linear functional on a convex set, and it can be further simplified by noticing that $B \in \mathcal{B}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$ only enters (10.9) through $D = B\,grad(u^*)$, and when B runs through $\mathcal{B}(\eta_1, \ldots, \eta_r; M_1, \ldots, M_r)$, D spans the closed ball of diameter $[b(\eta)\,grad(u^*), a(\eta)\,grad(u^*)]$, and therefore (10.9) is equivalent to

$$\int_\Omega \Big[\sum_{i=1}^r \eta_i^* \, F_i(x, u^*) - \Big(A^* \, grad(u^*).grad(p^*) \Big) \Big] \, dx \le \int_\Omega \Big[\sum_{i=1}^r \eta_i \, F_i(x, u^*) - \Big(D.grad(p^*) \Big) \Big] \, dx \tag{10.15}$$

for all $\eta_1, \ldots, \eta_r$ satisfying (9.9), $\Big(D - b(\eta)\,grad(u^*).D - a(\eta)\,grad(u^*) \Big) \le 0$,

and the convexity of the set of $\eta_1, \ldots, \eta_r, D$ used in (10.15) follows from Lemma 19, and the convexity of the functions a and $\frac{1}{b}$, consequence of their definition in (9.13).

One can go further in the analysis of the necessary conditions obtained, (10.9) or (10.15) for the case $N \ge 2$, or (10.10) for the case $N = 1$, as they consist in minimizing linear functionals on convex sets defined by linear constraints, and LAGRANGE multipliers can then be used for making these conditions more precise, but I will not describe this question here as it is a more classical subject.

11. Conclusion

I have now completed the description of the particular subject of this course, which was to show how questions of Homogenization appear in Optimal Design problems, following the work which I had pioneered in the early 70s with François MURAT.

It is time now that I describe the intuitive ideas behind the necessary conditions of optimality obtained by Konstantin LUR'IE, as they were explained to me in the early 80s by Jean-Louis ARMAND, after he had visited LUR'IE in Leningrad, where he had been told about my work [Ta2]. As I do not read much, I do not know where these ideas have appeared in print,[72] and although the idea is quite natural, it seems hard to transform into sound mathematical estimates.[73] Assume that we consider a problem involving two (isotropic) materials, with imposed global proportions, and that we want to test the optimality of a given classical design, with a smooth interface between the two materials; the classical idea, which goes back to HADAMARD,[74] consists in pushing the interface along its normal of variable amounts (with one constraint related to the global proportion imposed), and computing the change in the cost function leads to a necessary condition of optimality valid along the interface; LUR'IE's first idea was to work away from the interface, taking a small sphere imbedded in one material and an identical one imbedded in the other material and

---

[72] Having learned about some ideas of LUR'IE, I do not need a published reference in order to attribute these ideas to him (and I could not remember of anyone else claiming them as his/hers, although I have not checked the work of Richard DUFFIN, as I mention in footnote 75). I may be alone in thinking that if a new idea is only mentioned orally by someone who does not put it in print immediately it should be attributed to this person, eventually with the names of those who would have found the same idea later but independently, but not with the names of those who had heard about the idea and had put it in print under their name, expecting to acquire fame for an idea which was not theirs.

[73] I wonder then if LUR'IE had been able to carry out these computations in a mathematician's way, or if he had just acquired convincing evidence that some formulas must hold, as nonmathematicians often do.

[74] I have already mentioned the precise analysis along this line of thought, carried out by François MURAT and Jacques SIMON in [Mu&Si].

exchanging their content, and computing the change in the cost function leads to a necessary condition of optimality valid everywhere;[75] LUR'IE's second idea was to consider ellipsoids of the same volume instead of spheres, and his new necessary conditions of optimality were stronger; as he could now play with the orientation of the ellipsoids and the ratio of the axes, he realized that it was better to take them very slender, and in the limit he could understand that layered materials were important for his problem. As I have mentioned earlier, this was a quite good extension of the ideas of PONTRYAGUIN to a setting of partial differential equations, but not quite the right way to discover the analysis which I had performed with François MURAT, which was in some way a good extension of the better ideas of Laurence C. YOUNG.

The description of the method developed in this course, which corresponds in part to results which I had obtained with François MURAT in the 70s, is analogous to what I had already taught in 1983 at the CEA - EDF - INRIA Summer course at Bréau-sans-Nappe, written in [Mu&Ta1], and similar to what I had taught again in 1986 in Durham [Ta11]. I have included here much more of the basic results on Homogenization, which I had only alluded to before, and actually the description of the original method which I had followed with François MURAT in the early 70s had never appeared in print before, the reason being that it had been greatly simplified by my method of oscillating test functions which extends easily to all (linear) variational formulations; because I had decided to use a chronological point of view in this course, in order to show how new ideas had appeared, it was natural that I should describe first our original ideas, even though I had improved them later. I have added a few simplifications, which I had first written in 1995 in [Ta14], and which were therefore not included in my previous courses on the subject.

As I have mentioned earlier, some people have led a campaign of misattribution of my ideas which seems to have intensified around 1983. I may have inadvertently added to the confusion by forgetting to mention the reference [Ta7] of my second method for obtaining bounds on effective coefficients,[76] and I did not think that it had any importance because I could not imagine that there were people ready to steal an idea that they would have heard if they thought that it had not been written yet.[77] Unfortunately, the confusion may also have increased because Graeme MILTON called my method the "translation method",[78] and many have used the possibility of quoting my method by this new name, without attributing it correctly.

I have been told that Alexandre GROTHENDIECK has analyzed in an unpublished book a few ways in which misattribution of ideas is organized,[79] I have not read it and I cannot assert that my observations on this unfortunate aspect of academic behaviour coincide or not with his own. This "book" had first been mentioned to me in 1984 by Jean LERAY, who had pointed out that it was a good sign that my ideas were stolen, and that many who steal others ideas would probably like that some of their ideas be stolen too, as it would prove that they had had some of their own; Jean LERAY also had to face such an adverse behaviour, but if some of his ideas had been "borrowed" by a famous mathematician who had shown enough creativity of his own, for political reasons which were not too dissimilar to those which I had encountered myself more than thirty years later, I have not found myself as fortunate and many who use my ideas without saying it present such a distorted view that one does not have to be a very good student in Mathematics for performing

---

[75] In the late 80s, my late colleague Richard DUFFIN had mentioned to me that he had worked on questions of Optimal Design, and he used to call such necessary conditions a "principle of democracy". Unfortunately, as I knew about such questions, I failed to enquire about his precise results, so that I do not know if his conditions were of the HADAMARD type, or the LUR'IE type, and I do not know when he had first obtained such results.

[76] The reason for not giving the reference of that 1977 conference in Versailles was that the organizers had forgotten to send me a copy of the proceedings, and therefore I did not know the exact reference of my article.

[77] It had not been my intention to hide the existence of a written reference in order to confront later those who were ready to steal my ideas.

[78] I do not find that name so well adapted to what my method is about. On the other hand, when I was a student (in Paris in the late 60s), this precise term was used for describing a method of Louis NIRENBERG for proving regularity of solutions of elliptic equations in smooth domains.

[79] Laurent SCHWARTZ told me recently that the publication of GROTHENDIECK's book, "Récoltes et Semailles", was not possible because of the numerous personal attacks that it contains. It is nevertheless available on the Internet, in Russian translation!

the small detective work of identifying those who have stolen ideas that they do not even understand well enough.

This being said, I must say that I think that the worst sin of a teacher is to induce students in error, and I do consider it actually a minor sin to forget to name the inventor of an idea,[80] but a major sin to give a bad explanation of what an idea is, or to forget mentioning an important idea on a subject. In consequence, if someone would feel such a pressure for avoiding to mention the author of one of the ideas that I have described in this course, it would be better if he/she would start by learning well the content of this course, and then teach an improved version of it.

12. Acknowledgements

13. References

[Ar] ARTSTEIN Z., "Look for the extreme points," SIAM Review **22** (1980), 172–185.

[Ba] BABUŠKA I., "Homogenization and its application. Mathematical and computational problems," *Numerical Solution of Partial Differential Equations-III, (SYNSPADE 1975, College Park MD, May* 1975), *Academic Press, New York*, 1976, 89–116.

[Be&Li&Pa] BENSOUSSAN A. & LIONS J.-L. & PAPANICOLAOU G., *Asymptotic Analysis for Periodic Structures*, Studies in Mathematics and its Applications 5. North-Holland, Amsterdam 1978.

[Be] BERGMAN D., "Bulk physical properties of composite media," *Homogenization methods: theory and applications in physics (Bréau-sans-Nappe, 1983)*, 1–128, *Collect. Dir. Etudes Rech. Elec. France*, 57, *Eyrolles, Paris*, 1985.

[Br&Po] BRAIDY P. & POUILLOUX D., Mémoire d'Option, Ecole Polytechnique, 1982, unpublished.

[Cé&Ha] *Optimization of distributed parameter structures, Vol. I, II, J. CÉA & E. J. HAUG eds.* Proceedings of the NATO Advanced Study Institute on Optimization of Distributed Parameter Structural Systems, Iowa City, May 20-June 4, 1980. NATO advanced Study Institute Series E: Applied Sciences, 49, 50, *Martinus Nijhoff Publishers, The Hague*, 1981.

[Cé&Ma] CÉA J. & MALANOWSKI K., "An example of a max-min problem in partial differential equations," SIAM J. Control **8** 1970, 305–316.

[Ch] CHENAIS D., "Un ensemble de variétés à bord Lipschitziennes dans R^N: compacité, théorème de prolongements dans certains espaces de Sobolev," C. R. Acad. Sci. Paris Sér. A-B **280** (1975), Aiii, A1145–A1147.

[DG&Sp] DE GIORGI E. & SPAGNOLO S., "Sulla convergenza degli integrali dell'energia per operatori ellitici del 2 ordine," Boll. Un. Mat. Ital. (4) **8** (1973), 391–411.

[Fr&Mu] FRANCFORT G. & MURAT F., "Homogenization and Optimal Bounds in Linear Elasticity", Arch. Rational Mech. Anal., **94**, 1986, 307–334.

[Gu] GUTIÉRREZ S., "Laminations in Linearized Elasticity and a Lusin type Theorem for Sobolev Spaces," PhD thesis, CARNEGIE-MELLON University, May 1997.

---

[80] My religious upbringing forbids me to steal, but my memory is not perfect, and I may have forgotten to quote some authors for their ideas. If I realized that I had made such a mistake, either by being told about it or by finding it myself, I would certainly try to give a corrected statement on the next occasion where I would write on the subject (and I hope that a second lapse of memory would not occur at that time).

[Ha&Sh] HASHIN Z. & SHTRIKMAN S., "A variational approach to the theory of effective magnetic permeability of multiphase materials", J. Applied Phys. **33**, (1962) 3125–3131.

[Jo] JOSEPH D., *Stability of fluid motions* I, II, Springer Tracts in Natural Philosophy, Vol. 27, 28, Springer-Verlag, Berlin-New York, 1976.

[Ke] KELLER J., "A theorem on the conductivity of a composite medium," J. Mathematical Phys. **5** (1964) 548–549.

[Li1] LIONS J.-L., *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Dunod - Gauthier-Villars, Paris, 1968.

[Li2] LIONS J.-L. , *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969.

[Li3] LIONS J.-L., "Asymptotic behaviour of solutions of variational inequalities with highly oscillating coefficients," *Applications of methods of functional analysis to problems in mechanics (Joint Sympos., IUTAM/IMU, Marseille, 1975)* , pp. 30–55. Lecture Notes in Math., 503. *Springer, Berlin*, 1976.

[Lu] LUR'IE K. A., "On the optimal distribution of the resistivity tensor of the working substance in a magnetohydrodynamic channel," J. Appl. Mech. (PMM) **34** (1970), 255–274.

[Ma&Sp] MARINO A. & SPAGNOLO S., "Un tipo di approssimazione dell'operator $\sum_{ij} D_i(a_{ij}D_j)$ con operatori $\sum_j D_j(b\,D_j)$," Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **23** (1969), 657–673.

[MC] MCCONNELL W., "On the approximation of elliptic operators with discontinuous coefficients," Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **3** (1976), no. 1, 123–137.

[Me] MEYERS N., "An L^p-estimate for the gradient of solutions of second order elliptic divergence equations," Ann. Scuola Norm. Sup. Pisa (3) **17** (1963), 189–206.

[Mo& St1] MORTOLA S. & STEFFÉ S., Unpublished report, Scuola Normale Superiore, Pisa.

[Mo& St2] MORTOLA S. & STEFFÉ S., "A two-dimensional homogenization problem," Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **78** (1985), no. 3, 77–82.

[Mu1] MURAT F., "Un contre-exemple pour le problème du contrôle dans les coefficients," C. R. Acad. Sci. Paris, Sér. A-B **273** (1971), A708–A711.

[Mu2] MURAT F., "Théorèmes de non-existence pour des problèmes de contrôle dans les coefficients," C. R. Acad. Sci. Paris, Sér. A-B **274** (1972), A395–A398.

[Mu3] MURAT F., "H-convergence," *Séminaire d'analyse fonctionnelle et numérique*, Université d'Alger, 1977-78. Translated into English as MURAT F. & TARTAR L., "H-convergence," *Topics in the mathematical modelling of composite materials*, 21–43, Progr. Nonlinear Differential Equations Appl., 31, *Birkhäuser Boston, Boston, MA*, 1997.

[Mu4] MURAT F., "Compacité par compensation," Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), **5** (1978), 489–507.

[Mu&Si] MURAT F. & SIMON J., "Sur le contrôle par un domaine géométrique," *Publication 76015, Laboratoire d'Analyse Numérique, Université Paris 6*, 1976.

[Mu&Ta1] MURAT F. & TARTAR L., "Calcul des variations et homogénéisation," *Homogenization methods: theory and applications in physics (Bréau-sans-Nappe, 1983)*, 319–369, Collect. Dir. Etudes Rech. Elec. France, 57, *Eyrolles, Paris*, 1985. Translated into English as "Calculus of variations and homogenization," *Topics in the mathematical modelling of composite materials*, 139–173, Prog. Nonlinear Differential Equations Appl., 31, *Birkhäuser Boston, M A*, 1997.

[Mu&Ta2] MURAT F. & TARTAR L., "Optimality conditions and homogenization," *Nonlinear variational problems (Isola d'Elba, 1983)*, 1–8, Res. Notes in Math., 127, *Pitman, Boston, Mass. -London*, 1985.

[Ra] RAITUM U.E., "The extension of extremal problems connected with a linear elliptic equation," Soviet Math. Dokl. **19**, (1978), 1342-1345.

[S-P1] SANCHEZ-PALENCIA E., "Solutions périodiques par rapport aux variables d'espaces et applications," C. R. Acad. Sci. Paris, Sér. A-B **271** (1970), A1129–A1132.

[S-P2] SANCHEZ-PALENCIA E., "Equations aux dérivées partielles dans un type de milieux hétérogènes," C. R. Acad. Sci. Paris, Sér. A-B **272** (1972), A395–A398.

[Si] SIMON L., "On G-convergence of elliptic operators," Indiana Univ. Math. J. **28** (1979), no. 4, 587–594.

[Sp1] SPAGNOLO S., "Sul limite delle soluzioni di problemi di Cauchy relativi all'equazione del calore," Ann. Scuola Norm. Sup. Pisa (3) **21** (1967), 657–699.

[Sp2] SPAGNOLO S., "Sulla convergenza di soluzioni di equazioni paraboliche ed ellitiche," Ann. Scuola Norm. Sup. Pisa (3) **22** (1968), 571–597.

[Ta1] TARTAR L., "Convergence d'opérateurs différentiels," *Atti Giorni Analisi Convessa e Applicazioni* (*Roma*, 1974), 101–104.

[Ta2] TARTAR L., "Problèmes de contrôle des coefficients dans des équations aux dérivées partielles," *Control theory, numerical methods and computer systems modelling (Internat. Sympos., IRIA LABORIA, Rocquencourt, 1974), pp. 420–426. Lecture Notes in Econom. and Math. Systems, Vol.* 107, *Springer, Berlin*, 1975. Translated into English as MURAT F. & TARTAR L., "On the control of coefficients in partial differential equations," *Topics in the mathematical modelling of composite materials*, 1–8, Prog. Nonlinear Differential Equations Appl., 31, *Birkhäuser Boston, MA*, 1997.

[Ta3] TARTAR L., "Quelques remarques sur l'homogénéisation," *Functional analysis and numerical analysis* (*Tokyo and Kyoto*, 1976), 469–481, *Japan Society for the Promotion of Science, Tokyo*, 1978.

[Ta4] TARTAR L., "Weak convergence in nonlinear partial differential equations," *Existence Theory in Nonlinear Elasticity*, 209–218. *The University of Texas at Austin*, 1977.

[Ta5] TARTAR L., "Une nouvelle méthode de résolution d'équations aux dérivées partielles non linéaires," *Journées d'Analyse Non Linéaire (Proc. Conf., Besançon, 1977)*, pp. 228–241, Lecture Notes in Math., 665, *Springer, Berlin*, 1978.

[Ta6] TARTAR L., "Nonlinear constitutive relations and homogenization," *Contemporary developments in continuum mechanics and partial differential equations (Proc. Internat. Sympos., Inst. Mat., Univ. Fed. Rio de Janeiro, Rio de Janeiro)*, pp. 472–484, North-Holland Math. Stud., 30, *North-Holland, Amsterdam-New York*, 1978.

[Ta7] TARTAR L., "Estimations de coefficients homogénéisés," *Computing methods in applied sciences and engineering (Proc. Third Internat. Sympos., Versailles*, 1977), *I*, pp. 364–373, *Lecture Notes in Math.*, 704, *Springer, Berlin*, 1979. Translated into English as "Estimations of homogenized coefficients," *Topics in the mathematical modelling of composite materials*, 9–20, Prog. Nonlinear Differential Equations Appl., 31, *Birkhäuser Boston , MA*, 1997.

[Ta8] TARTAR L., "Compensated compactness and applications to partial differential equations," *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, pp. 136–212. *Res. Notes in Math.*, 39, *Pitman, San Francisco, Calif.*, 1979.

[Ta9] TARTAR L., "Estimations fines des coefficients homogénéisés," *Ennio De Giorgi colloquium (Paris, 1983)*, 168–187, Res. Notes in Math., 125, *Pitman, Boston, MA-London*, 1985.

[Ta10] TARTAR L., "Remarks on homogenization," *Homogenization and effective moduli of materials and media (Minneapolis, Minn., 1984/1985)*, 228–246, IMA Vol. Math. Appl., 1, *Springer, New York-Berlin*, 1986.

[Ta11] TARTAR L., "The appearance of oscillations in optimization problems," *Nonclassical continuum mechanics (Durham, 1986)*, 129–150, London Math. Soc. Lecture Note Ser., 122, *Cambridge Univ. Press, Cambridge-New York*, 1987.

[Ta12] TARTAR L., "H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations," *Proc. Roy. Soc. Edinburgh Sect. A* **115**, (1990), no. 3-4, 193–230.

[Ta13] TARTAR L., "Remarks on Optimal Design Problems," *Calculus of variations, homogenization and continuum mechanics (Marseille, 1993)*, 279–296, Ser. Adv. Math. Appl. Sci., 18, *World Sci. Publishing, River Edge, NJ*, 1994.

[Ta14] TARTAR L., "Remarks on the homogenization method in optimal design problems," *Homogenization and applications to material sciences (Nice, 1995)*, 393–412, GAKUTO Internat. Ser. Math. Sci. Appl., 9, *Gakkokotosho, Tokyo*, 1995 .

[Yo] YOUNG, L. C., *Lectures on the Calculus of Variation and Optimal Control Theory*, (W. B. Saunders, Philadelphia: 1969).

[Zo] ZOLEZZI T., "Teoremi d'esistensa per problemi di controllo ottimo retti da equazioni ellittiche o paraboliche," Rend. Sem. Mat. Padova, **44** (1970), 155–173.