

OPTIMIZATION, RELAXATION AND YOUNG MEASURES

PABLO PEDREGAL

ABSTRACT. We review the use of Young measures in analyzing relaxed and generalized formulations for typical problems of optimization including variational principles, optimal control problems, models in materials science, optimal design problems and nonlocal optimization problems.

1. INTRODUCTION

Many problems in science and technology can be formulated in the language of optimization theory where an optimal solution or the best answer to a particular situation is sought. The nature of different problems and the way to solve and analyze them are oftentimes unrelated to each other despite the fact that we are always looking for some optimal solution according to some criterion that in most cases of interest comes in the form of an integral. It is true that what makes the difference between distinct problems is not the cost functional itself but rather the admissibility restrictions or constraints. We would like to make the effort to place many of these optimization problems within the same framework. There is another feature shared by all of the examples we will discuss in this work that makes this job particularly attractive. Under situations of interest, they all lack classical optimal solutions, or at least, the existence of such solutions is far from being straightforward. Our title refers to these two basic common properties of the examples to be examined. We talk about relaxation when we face an optimization problem that does not admit optimal classical solutions. The essential ingredient responsible for this lack of solutions in most of the cases is the oscillatory behavior of minimizing sequences: the more they oscillate, the less energy is required, to the point that microstructural features arise as optimal solutions are sought. Young measures can be used as a tool to organize our ideas about this oscillatory behavior and to deal in a consistent way with oscillations. The success of Young measures in this field is due to the fact that our cost functionals are integrals. The particular problem and the restrictions placed on admissibility will make the analysis harder in some cases, almost impossible in others and more transparent in some others. Our main goal is to treat them all in the same framework and set up relaxation and generalized optimization problems with the help of Young measures, exploiting their properties with regard to integral cost functionals.

We divide the exposition into several sections. Section 2 is concerned with an abstract approach to optimization and relaxation in order to make explicit the

Received by the editors October 1, 1997, and in revised form October 14, 1998.

1991 *Mathematics Subject Classification*. Primary 49J15, 49J45, 73C50, 73K20, 73V25.

Key words and phrases. Integral functionals, oscillatory behavior, generalized optimization problems, local and nonlocal admissibility constraints.

underlying ideas for the examples in subsequent sections. It also motivates the use of Young measures in describing relaxed formulations when the optimization criterion is an integral; at the same time, we will see that one can use Young measures mainly as a tool to organize and comprehend the behavior of sequences of functions with respect to integral functionals. Our point of view on relaxation is motivated by a consistent interest in understanding the properties that make a sequence minimizing for a particular problem and that are shared by all minimizing sequences for the same problem. In particular we would like to set up new, relaxed optimization problems whose generalized solutions have the following remarkable property: in some cases, any sequence generating, in the appropriate sense, such generalized optimal solution will be minimizing for our initial problem; in some others, there will be a selected class of sequences of functions that generate the generalized minimizer that will be minimizing for our original problem. Our aim is to show how Young measures can be used to describe these relaxed formulations of different types of optimization problems. The analysis in the present work, however, does not go deep into each problem. Some of them have already been the subject of much attention (like variational and optimal control problems) from the point of view of Young measures. Much work is still needed to understand others (like optimal design and shape optimization problems). Many of the latter have been successfully tackled with different methods and techniques. References are given below. Young measures furnish an alternative way to treat and analyze optimization problems in a unifying way.

The rest of the paper comprises a series of selected examples where the abstract ideas of Section 2 can be formulated in a more concrete and specific way. The classification of the examples is more or less arbitrary and motivated by organization. More examples could have been added, but the author has restricted himself to problems he has been involved with in one way or another during the last few years. In some of these examples an alternative and more successful approach in some regards can be found in the literature as indicated below. For some examples we have purposely dealt with the one-dimensional situation where the analysis is much simpler and explicit computations are possible in many instances. We would like to stress the fact that the passage to higher dimensional, vector problems is far from being a direct generalization of the lower dimensional situation. We have tried to emphasize the tremendous difficulties attached to this issue, many of which are not fully understood yet. Another appealing feature of generalized optimization principles is that they usually admit solutions under mild technical assumptions, and this in turn justifies the search for generalized optimality conditions that such generalized optima must verify. Although such analysis has been and is being carried out in some of the problems to be examined later, we will not include here explicit references to such questions.

Our basic existence theorem for Young measures is taken from [4]. Many technical points on Young measures including some of the results used in this work can be found in [53]. There exist other approaches to Young measures other than the pioneering work [69], [70], [71]: [2] and [67] are more abstract and general treatments of Young measures, [23] follows an approach based on slicing measures techniques that may be helpful in some ways, [63] focuses also on applications to partial differential equations, [59] is a new book where optimization problems are also studied in terms of Young measures.

Section 3 is an overview of problems in the Calculus of Variations where non-convexity and nonexistence have played an important role in nonlinear elasticity. [5] and [6] are two important basic references where an energetic approach to phase transitions in crystals is for the first time rigorously established. The point of view based on Young measures generated by gradients was studied in [36], [37], [38] and further pursued and developed in [53]. Because of the too many technicalities involved, we merely give here the main results. See also [20], [21] and [44]. Vector variational problems and all the issues associated to Young measures in this context are not important just because of nonlinear elasticity. They are also relevant in other optimization settings like optimal design problems, as we will see. A rather complete summary on all of the important and difficult issues related to nonconvex, vector variational problems can be found in [43] and the references contained in it.

Optimal control problems governed by ordinary differential equations occupy Section 4. As a matter of fact, Young measures were originally introduced with applications to optimal control problems in mind. The literature on this topic is enormous. There are some interesting textbooks like [7], dealing with the general theory, and [27] and [68], which are closer to our point of view on nonexistence and relaxed formulations. More recently, the study of relaxation in control problems in terms of Young measures has been undertaken in [1], [19], [46]. A different but related point of view about relaxation has been considered in [11]. Γ -convergence techniques have also been used to describe and formulate limit problems of a sequence of optimal control problems. See [12] and [26]. In [41] and [42] an attempt is made to study the relationship between different ways of understanding relaxation for optimal control problems. A complete treatment of these problems from the point of view of relaxation is included in this paper. This choice has been essentially made because technicalities related to constraints can be kept to a minimum in comparison with other optimization problems. They are a good example to emphasize Young measure techniques.

Some models of material behavior are considered in Section 5 in the context of ferromagnetism ([29], [30]) where the energy functional consists in two contributions. One is local; the second one is not in the form of a control problem. Although we present the more complex setting of magnetostriction, technicalities advise us to restrict to the simpler case of micromagnetics where the interaction between magnetic and elastic properties is neglected and the body is assumed rigid. An explicit form of the relaxed energy functional in terms of Young measure-valued magnetizations is provided. The computation makes use of the characterization of gradient Young measures in Section 3: the scalar case for the micromagnetics situation, the vector case for the magnetostriction functional. See [50]. Similar results and conclusions have been obtained in [22] and [66].

Section 6 focuses on optimal design problems. This is also a broad subject which we do not pretend to cover. [39] is a fundamental reference to understand the relationship between optimal design problems and general convex hulls of functions. Homogenization techniques have been crucial in understanding these optimization problems. Some of the basic references are [45], [46], [47], [62], [64]. Here, we consider an optimal design problem for two phase conductors, treating first the one dimensional situation. This particular problem has been examined in [16] and [17]. The relaxed formulation based on Γ -convergence has been described again in [46], [47] and [64]. The huge amount of literature about composites is also relevant to this sort of problem. Our aim here is to emphasize how a relaxed formulation of

the problem based on Young measures can be considered, and the relationship of this formulation with quasiconvexity and vector variational problems.

Another optimal design problem treated here is concerned with plates of variable thickness under a given load. This particular problem has been studied in [10] and [40]. A formulation in terms of Young measures has been analyzed in [8] and [9] where particular forms of the relaxation have been given. Our reference for this problem is [48] where the optimal relaxation with as few design variables as possible is given and proved.

We have also included a short discussion about shape optimization problems in a typical format. The main sources for this problem are [13], [14], [15]. In these works relaxation is described in terms of Γ -convergence techniques, and the generalized set of competing objects together with admissibility conditions and extended functionals is very clearly established. We just make here some general remarks that would require further analysis in order to be useful.

Finally, Section 7 gathers a variational nonlocal problem that is written as a double integral. Again relaxation is described in terms of Young measures, but the restrictions placed on them are harder to clarify (even in dimension one) as the appropriate convexity conditions are mixed up with the nonlocal effect and the possible inhomogeneity of the functional. This analysis was carried out in [54].

New tools to study optimization problems, homogenization and partial differential equations have been recently introduced. H-measures ([28], [65]) provide a new way to formulate and understand some nonlinear phenomena in the presence of differential constraints. We have systematically sought to avoid concentration effects in nonlinear functionals in our hypotheses. This is something that Young measures cannot capture. H-measures and varifolds ([25]) can however handle this phenomenon.

2. AN ABSTRACT FRAMEWORK: YOUNG MEASURES

Let

$$I_0 : \mathcal{U} \rightarrow \mathbf{R}^* = \mathbf{R} \cup \{+\infty\}$$

be a particular functional we are interested in, defined on a set \mathcal{U} which could be a subset of some bigger set. We do not assume any structure on \mathcal{U} whatsoever nor any particular property on I_0 . We are concerned with the general, abstract optimization problem

$$(2.1) \quad \text{Find } u \in \mathcal{U} : \quad I_0(u) = m = \inf_{\mathcal{U}} I_0 \in \mathbf{R}.$$

Historically, one of the main optimization problems, at least from an analytical point of view, is the paradigm of the Calculus of Variations, where \mathcal{U} is a subset of functions belonging to some Sobolev space over a regular subset $\Omega \subset \mathbf{R}^N$ verifying prescribed Dirichlet boundary conditions, and

$$I_0(u) = \int_{\Omega} F_0(x, u(x), \nabla u(x)) dx,$$

for a Carathéodory function

$$F_0 : \Omega \times \mathbf{R}^m \times \mathbf{M}^{m \times N} \rightarrow \mathbf{R}^*.$$

In this case the set of competing objects \mathcal{U} is equipped with such structure that permits us to demonstrate existence of solutions of (2.1) under suitable technical

assumptions. The technique is called the direct method of the Calculus of Variations, and it is especially fruitful when some “convexity” property with respect to the gradient variable is assumed on the integrand F_0 . More recently, many different types of optimization problems have attracted the attention of researchers because of their relevance in modern science and technology. Optimal control problems governed by ordinary differential equations or partial differential equations, optimal design problems, shape optimization problems, nonlocal variational problems, and even more classical local variational principles have been analyzed and examined, both analytically and numerically, in the last thirty or forty years. Many of the most challenging and motivating problems involve lack of existence of the basic optimization problem (2.1), and these difficulties have led to new developments in optimization theory. We would like to concentrate on this situation.

It might be interesting for some of our readers to bring to mind once more the typical Bolza problem so that they can have a feeling of what we mean by lack of existence of classical solutions, and how lack of convexity and oscillatory behavior is usually associated to such problems. The functional is

$$I_0(u) = \int_0^1 \left[(u'(x)^2 - 1)^2 + u(x)^2 \right] dx,$$

where $u : (0, 1) \rightarrow \mathbf{R}$ is supposed to be a function in $W^{1,4}(0, 1)$ satisfying the boundary conditions $u(0) = u(1) = 0$. We claim that there is no admissible v such that $I_0(v) = \inf I(u)$. In order to convince ourselves of this assertion, we claim that $m = \inf I(u)$ vanishes. On the one hand it is clear that $m \geq 0$. But on the other, for the sequence of saw-tooth functions $u_j(x) = u(jx)$ where $u(y) = |y|$ for $y \in [-1/2, 1/2]$ extended by periodicity to all of \mathbf{R} , we have $\lim_{j \rightarrow \infty} I(u_j) = 0$. Notice however that $I(v) = 0$ is impossible for a single function. The oscillatory behavior of the sequence of derivatives $\{u'_j(x)\}$ is what we would like to formalize in a way that may enable us to treat qualitatively and quantitatively more complex situations. Observe the lack of convexity of the integrand $\left[(u'(x)^2 - 1)^2 + u(x)^2 \right]$ with respect to the derivative.

Assume either we know (2.1) does not admit a solution in \mathcal{U} , because the optimization criterion may favor oscillatory behavior, or we have well-founded doubts about its existence because no known existence result can be applied, or because numerical experiments or other types of considerations seem to indicate the optimality of oscillations; even if we are certain about existence of classical solutions, it might be interesting to investigate if some optimizing sequences might oscillate. Under all these circumstances, our goal is to comprehend the relevance of oscillatory behavior in situations where we are interested in the functional I_0 . The information we would like to have in such cases is about the precise oscillatory behavior of minimizing sequences: what do sequences $\{u_j\}$ with the property $I_0(u_j) \searrow m$ have in common? In order to answer this fundamental question we would like to describe in general terms what we understand by a relaxed formulation of (2.1). This amounts to enlarging the set of competing objects, $\tilde{\mathcal{U}}$, and to defining a new functional \tilde{I}_0 on $\tilde{\mathcal{U}}$ in such a fashion that there exists $\tilde{u} \in \tilde{\mathcal{U}}$ with

$$\tilde{I}_0(\tilde{u}) = m = \inf_{\mathcal{U}} I_0,$$

and, perhaps more importantly, “every” sequence $\{u_j\} \subset \mathcal{U}$ generating in the appropriate sense \tilde{u} will be minimizing for I_0 . In this sense \tilde{u} represents the common

link to all such minimizing sequences (or many of those). The interesting point we have tried to emphasize in this paragraph is that regardless of whether we have precise information about existence or nonexistence, or if we do not have the slightest idea about it, we lose nothing in considering a generalized formulation but we may gain a lot by doing so.

There is a standard, abstract way of defining this relaxed formulation which is a typical construction in functional analysis. The method is too abstract to be useful in specific situations, but still it is interesting to realize that it can always be done regardless of the complexity of the original optimization problem. Consider a class of functionals $I \in \mathcal{I}$ where each I is defined on \mathcal{U} and takes values on \mathbf{R}^* . We assume $I_0 \in \mathcal{I}$. We would like to develop some sort of theory valid for all functionals in \mathcal{I} and not only for I_0 . It would not make much sense to set up things having in mind only I_0 . For instance, it would be reasonable to have enough functionals in mind so as to separate objects in \mathcal{U} : if $u_1, u_2 \in \mathcal{U}$ and $I(u_1) = I(u_2)$ for all $I \in \mathcal{I}$, then $u_1 = u_2$, but this property is not necessary or assumed in what follows. We define an equivalence relation on the class of sequences of \mathcal{U} by saying

$$\{u_j\} \equiv \{v_j\} \quad \text{if and only if} \quad \lim_{j \rightarrow \infty} I(u_j) = \lim_{j \rightarrow \infty} I(v_j), \text{ for all } I \in \mathcal{I}.$$

Let $\tilde{\mathcal{U}}$ be the set of all equivalence classes \tilde{u} . Notice that $\mathcal{U} \subset \tilde{\mathcal{U}}$. Extend I to $\tilde{\mathcal{U}}$ defined on $\tilde{\mathcal{U}}$ by simply putting

$$\tilde{I}(\tilde{u}) = \lim_{j \rightarrow \infty} I(u_j)$$

if $\{u_j\}$ represents \tilde{u} . Trivially this definition is consistent. Finally define a topological structure on $\tilde{\mathcal{U}}$ by saying

$$\tilde{u}_j \rightarrow \tilde{u} \quad \text{if and only if} \quad \lim_{j \rightarrow \infty} \tilde{I}(\tilde{u}_j) = \tilde{I}(\tilde{u}) \text{ for all } \tilde{I} \in \mathcal{I}.$$

$\tilde{\mathcal{U}}$ is thus endowed with the coarser topology that makes all the (extended) functionals in \mathcal{I} continuous. This is also a standard procedure in topology. Notice that the inclusion $\mathcal{U} \subset \tilde{\mathcal{U}}$ is now dense because given $\tilde{u} \in \tilde{\mathcal{U}}$ any representative $\{u_j\}$ considered as a sequence in $\tilde{\mathcal{U}}$ will converge to \tilde{u} in this topology.

Consider now the relaxed problem

$$(2.2) \quad \text{Find } \tilde{u} \in \tilde{\mathcal{U}} : \quad \tilde{I}_0(\tilde{u}) = \tilde{m} = \inf_{\tilde{u}} \tilde{I}_0 \in \mathbf{R}.$$

We can prove a relaxation theorem.

Proposition 2.1. *Assume that I_0 is such that some minimizing sequence in \mathcal{U} is compact when considered as a sequence in $\tilde{\mathcal{U}}$. Then $m = \tilde{m}$ and (2.2) has at least one solution \tilde{u}*

$$\tilde{I}_0(\tilde{u}) = \tilde{m}.$$

Moreover if $\{u_j\}$ converges to \tilde{u} in $\tilde{\mathcal{U}}$, then it is minimizing for I_0 .

Proof. The proof is almost trivial.

1. $\tilde{m} \leq m$ because $\mathcal{U} \subset \tilde{\mathcal{U}}$ and $\tilde{I}|_{\mathcal{U}} = I$ for all $I \in \mathcal{I}$.
2. Given any $\tilde{u} \in \tilde{\mathcal{U}}$, if $\{u_j\}$ is any representative, then $u_j \rightarrow \tilde{u}$ in $\tilde{\mathcal{U}}$ so that $\tilde{I}_0(\tilde{u}) = \lim_{j \rightarrow \infty} I_0(u_j)$ and hence $m \leq \tilde{m}$.
3. Let $\{u_j\}$ be minimizing and compact. By hypothesis, there exists $\tilde{u} \in \tilde{\mathcal{U}}$ such that $u_j \rightarrow \tilde{u}$. In this case \tilde{u} is trivially a solution for (2.2).

4. If v_j is such that $v_j \rightarrow \tilde{u}$, then by definition

$$m = \tilde{I}_0(\tilde{u}) = \lim_{j \rightarrow \infty} I_0(v_j)$$

and v_j is minimizing for I_0 . □

In each particular problem the issue is to find the more useful description of the relaxation based on the features of the situation. Notice also how the choice of the class \mathcal{I} may influence the description of $\tilde{\mathcal{U}}$ and the analysis of the problem.

In most of the cases of interest, the functionals are represented by integrals where integrands depend on the competing functions. The topology induced by such functionals on sets of competing functions is some sort of weak * topology provided we interpret the action of a functional on a function as linear. This leads us to regard each competing function $u(x), x \in \Omega$, as a family of probability measures $\nu_x = \delta_{u(x)}$ considered as a mapping $\nu : \Omega \rightarrow \mathcal{M}$ where \mathcal{M} represents the set of Radon measures supported in some appropriate euclidean space and δ stands for the Dirac mass. In this case, the set of such mappings taking values in \mathcal{M} (which is huge) furnishes a first model for $\tilde{\mathcal{U}}$ when functionals are given by integrals. Any such mapping $\nu = \{\nu_x\}_{x \in \Omega}$ is called a Young measure or a parametrized measure.

Definition 2.2. A mapping $\nu : \Omega \rightarrow \mathcal{M}$ is called a Young measure if each $\nu(x) \equiv \nu_x$ is a probability measure and the function of x

$$\bar{\psi}(x) = \langle \psi, \nu_x \rangle$$

is measurable for any continuous ψ .

The fact that this set of objects provides a model for $\tilde{\mathcal{U}}$ is a consequence of the existence theorem for Young measures ([4], [53]).

Theorem 2.3. Let $\Omega \subset \mathbf{R}^N$ be a measurable set and let $u_j : \Omega \rightarrow \mathbf{R}^m$ be measurable functions such that

$$\sup_j \int_{\Omega} g(|u_j|) dx < \infty,$$

where $g : [0, \infty) \rightarrow [0, \infty]$ is a continuous, nondecreasing function such that $\lim_{t \rightarrow \infty} g(t) = \infty$. There exists a subsequence, not relabeled, and a family of probability measures, $\nu = \{\nu_x\}_{x \in \Omega}$ (the associated parametrized measure), depending measurably on x , with the property that whenever the sequence $\{\psi(x, u_j(x))\}$ is weakly convergent in $L^1(\Omega)$ for any Carathéodory function $\psi(x, \lambda) : \Omega \times \mathbf{R}^m \rightarrow \mathbf{R}^*$, the weak limit is the (measurable) function

$$\bar{\psi}(x) = \int_{\mathbf{R}^m} \psi(x, \lambda) d\nu_x(\lambda).$$

By a Carathéodory function $\psi(x, \lambda) : \Omega \times \mathbf{R}^m \rightarrow \mathbf{R}^*$ we simply mean a function measurable in x and continuous in λ .

Assume the functionals $I(u)$ are given by

$$I(u) = \int_{\Omega} \psi(x, u(x)) dx$$

for suitable Carathéodory functions ψ . Let I_0 be associated to ψ_0 . Equivalence classes are represented by all those sequences that share the same underlying Young measure. I is trivially extended to Young measures by the natural formula

$$\tilde{I}(\nu) = \int_{\Omega} \langle \psi(x, \cdot), \nu_x \rangle dx.$$

By Proposition 2.1 we have Young measure minimizers which are exactly the ones generated by minimizing sequences for I_0 , provided minimizing sequences for I_0 are eligible to have associated Young measures (the boundedness condition required in the existence theorem) and the weak convergence of $\{\psi_0(x, u_j(x))\}$ is established or assumed.

Three important remarks are necessary. The most delicate point in this abstract construction is to isolate the admissibility conditions on Young measures which must be inherited by the restrictions imposed on competing functions for the original optimization problem. This issue is what makes the analysis of specific problems so different from each other and so hard in some situations. Secondly, it is plausible that there might be some kind of hierarchy on the set of minimizing sequences and that they may be classified according to some relevant criterion to measure complexity or optimality. In this case we are confronted with the task of trying to select in some precise or heuristic sense the best relaxation whose generalized solutions are associated to the minimizing sequences with simpler or best structure. Finally, the formulation of the optimization problem must be local in nature because Young measures cannot capture the limiting behavior of nonlocal quantities. The cost functional must be a local, dependent variable of the optimization independent variables. This means that if the optimization problem is nonlocal, we need to find a local formulation before an analysis in terms of Young measures can be carried out. This process typically involves introducing new “independent variables”. The description of relaxation in terms of the unique, initial, independent variable leads us to consider some sort of projection; in this process, information on all minimizing sequences may be lost, and we may recover only some selected minimizing sequences. The different examples that follow will clarify these remarks.

At this stage, the most fundamental question is: how can one go from a generalized Young measure solution to a minimizing sequence for the initial optimization problem, which is after all what motivated our analysis? Or to put it in another way, what information about our original optimization problem is encoded in the Young measure solution of the generalized or relaxed problem? A Young measure is a family of probability measures. As such, each individual member ν_x yields a family of weights and a family of corresponding states in the range of admissible functions for our optimization problem. Any sequence of functions generating the Young measure will have to oscillate “near” x (unless ν_x is trivial, i.e. a delta) among the different states in the support of ν_x with relative frequencies given by the corresponding weights. This is clearly seen in the fundamental recording property of the Young measure associated to a sequence of functions as expressed in Theorem 2.3. The defining property of a Young measure, what identifies it with respect to one of its generating sequences of functions, is its capability of providing all weak limits of compositions with any continuous function. In some problems, any such generating sequence will be minimizing provided it is admissible. In others, the construction of the optimizing sequence is inherent to the passage to a

generalized formulation. Indeed, the explicit form of a relaxed problem is closely connected to the way in which minimizing sequences can be obtained from Young measure solutions, so that this process is part of the analysis around generalized formulations of optimization problems.

Another underlying, relevant point that makes relaxed formulations interesting is that existence of optimal solutions (whatever the optimization problem may be) is the starting point of looking for optimality conditions that such optimal solutions have to verify. The study of optimality conditions may lead us to find or determine such optimal solutions. This is an important, difficult issue largely unexplored in most of the problems that follow as far as the author knows (see [49], [55]). Finally, computational questions about Young measure solutions of optimization problems have not been addressed in a consistent way ([51], [52] are very modest, initial contributions). It is not clear how to set up the numerical framework so as to be helpful in approximating minimizing sequences for the original optimization problem.

3. THE CALCULUS OF VARIATIONS

The standard problem of the Calculus of Variations is to find minimizers of the functional

$$I(u) = \int_{\Omega} F(x, u(x), \nabla u(x)) dx$$

where Ω is some domain in \mathbf{R}^N , the functions $u : \Omega \rightarrow \mathbf{R}^m$ belong to $W^{1,p}(\Omega)$, $p > 1$, and they all satisfy Dirichlet boundary conditions: $u - u_0 \in W_0^{1,p}(\Omega)$. The integrand F is assumed to be measurable on x and continuous in both u and ∇u . The property on F that ensures the lower semicontinuity of I with respect to weak convergence in $W^{1,p}(\Omega)$ is the quasiconvexity condition (in the vector case $N, m > 1$), introduced by Morrey in [44]. A function $\varphi : \mathbf{M}^{m \times N} \rightarrow \mathbf{R}$ is said to be quasiconvex if

$$\varphi(A) \leq \frac{1}{|\Omega|} \int_{\Omega} \varphi(A + \nabla w(x)) dx$$

for all matrices $A \in \mathbf{M}^{m \times N}$ and all test functions w . An equivalent, more manageable formulation of the quasiconvexity condition involves the requirement

$$(3.1) \quad \varphi(A) \leq \int_Q \varphi(A + \nabla w(x)) dx$$

where Q is any unit cube in \mathbf{R}^N and w is a Q -periodic vector field ([60]). A function F depending also on (x, u) is quasiconvex if $F(x, u, \cdot)$ is quasiconvex for all pairs (x, u) . Under the typical coerciveness assumption and upper bound

$$(3.2) \quad c(|A|^p - 1) \leq F(x, u, A) \leq C(1 + |u|^p + |A|^p),$$

the initial problem admits a minimizer provided the integrand F enjoys the quasiconvexity condition. See [21], [53].

Energy functionals of this type that do not enjoy the quasiconvexity property have been the object of much attention in recent years mainly motivated by nonlinear elasticity models of crystalline materials. The hypotheses on the integrand (the free energy of the crystal) that are natural from the physical point of view (frame indifference) and symmetry considerations lead to the realization that such energy densities cannot be quasiconvex. This lack of quasiconvexity is in turn responsible

for the oscillatory behavior of minimizing sequences. There are no minimizers in some circumstances, and a very fine alternating microstructure is the result of the minimization process with nonquasiconvex integrands. The variational approach provides in this way an explanation for those fine, observed microstructures ([5], [6]). In our framework for relaxation, we need to focus on Young measures associated to sequences of gradients bounded in $W^{1,p}(\Omega)$ with prescribed boundary data.

Set

$$\mathcal{U} = \left\{ \nabla u : u \in W^{1,p}(\Omega), u - u_0 \in W_0^{1,p}(\Omega) \right\},$$

$$\tilde{\mathcal{U}} = \left\{ \nu = \{\nu_x\}_{x \in \Omega} : \nu \text{ is associated to a sequence in } \mathcal{U} \right\}.$$

The generalized functional will be

$$\tilde{I}(\nu) = \int_{\Omega} \int_{\mathbf{M}^{m \times N}} F(x, u(x), A) d\nu_x(A) dx$$

where

$$\nabla u(x) = \int_{\mathbf{M}^{m \times N}} A d\nu_x(A), \quad u \in \mathcal{U},$$

yields the relationship between u and ν . Under (3.2), Proposition 2.1 gives the existence of generalized minimizers provided that $\{F(x, u_j(x), \nabla u_j(x))\}$ is weakly convergent in $L^1(\Omega)$ for some minimizing sequence. This weak convergence is a quite remarkable fact (see [25], [53]). What is crucial is to understand the nature of such generalized minimizers. This amounts to characterizing Young measures coming from bounded sequences of gradients in $W^{1,p}(\Omega)$. This characterization comes in the form of a duality between quasiconvex functions and gradient Young measures.

Theorem 3.1. $\nu = \{\nu_x\}_{x \in \Omega}$ belongs to $\tilde{\mathcal{U}}$ if and only if

1. $\nabla u(x) = \int_{\mathbf{M}^{m \times N}} A d\nu_x(A)$ for some $u \in \mathcal{U}$;
- 2.

$$\varphi(\nabla u(x)) \leq \int_{\mathbf{M}^{m \times N}} \varphi(A) d\nu_x(A)$$

for a.e. $x \in \Omega$ and every quasiconvex function φ with growth of order p th at infinity;

- 3.

$$\int_{\Omega} \int_{\mathbf{M}^{m \times N}} |A|^p d\nu_x(A) dx < \infty.$$

Moreover, every such ν can be generated by a sequence $\{\nabla u_j\}$ such that $\{|\nabla u_j|^p\}$ is equiintegrable.

The analysis of all these questions related to gradient Young measures is the central focus of [53]. Notice that we are not discussing relaxation in terms of the quasiconvexification of integrands ([20]).

In order to help our readers appreciate the difficulties attached to the verification of the conditions in the previous theorem, let us turn our attention to the scalar case where either m or N is one. In such cases, the quasiconvexity condition reduces

to the usual notion of convexity. The easiest way of convincing ourselves of this fact (if we take $m = 1$) is by considering the vector field

$$w(x) = t \int_0^{x \cdot n} (2\chi(s) - 1) ds$$

where n is any unit vector in \mathbf{R}^N , t is any real and χ is the characteristic function of $(0, 1/2)$ in $(0, 1)$ extended by periodicity to all of \mathbf{R} . If Q is any unit cube in \mathbf{R}^N with one axis along the direction of n , this function w is admissible in (3.1), and its gradient can be easily computed

$$\nabla w(x) = t (2\chi(x \cdot n) - 1) n.$$

After calculating the resulting integral we arrive at

$$\varphi(A) \leq \frac{1}{2}\varphi(A + tn) + \frac{1}{2}\varphi(A - tn).$$

The arbitrariness of t and n imply that φ must be convex. But then requirement 2 in Theorem 3.1 is empty because the classical Jensen's inequality always holds for convex functions and probability measures. This condition is however crucial in the fully vector case when both dimensions m and N are at least 2. In order to convey the difficulties attached to the understanding of this inequality for the vector case, let us try to reproduce the same argument we have used for the scalar case. The test field w as defined above must now incorporate components. The most direct generalization would be to take

$$w(x) = \int_0^{x \cdot n} (2\chi(s) - 1) ds a,$$

where now $a \in \mathbf{R}^m$ is a constant vector. In this case

$$\nabla w(x) = (2\chi(x \cdot n) - 1) a \otimes n,$$

where $a \otimes n$ stands for the rank-one matrix $(a \otimes n)_{ij} = a_i n_j$. The inequality we get is

$$(3.3) \quad \varphi(A) \leq \frac{1}{2}\varphi(A + a \otimes n) + \frac{1}{2}\varphi(A - a \otimes n).$$

Hence we only get convexity along matrices of rank one, and not along any matrix. This is a good indicator that understanding the quasiconvexity property is a very complex issue because it is a genuinely vector concept. A function φ defined on matrices enjoying (3.3) is called rank-one convex to emphasize the fact that such a function is convex only when the property is tested along rank-one matrices. A quasiconvex function must be rank-one convex. Although the converse has been a long-time standing conjecture, it has recently been disproved ([61]) thus providing well-founded evidence of the difficulty associated with the quasiconvexity property and with Jensen's inequality with respect to quasiconvex functions. There is however a large class of nonconvex, quasiconvex functions that are very important in nonlinear elasticity. These are called polyconvex and provide the main source of explicit quasiconvex functions ([3]).

4. OPTIMAL CONTROL PROBLEMS

The problem to be analyzed in this section is to minimize the cost functional

$$(4.1) \quad I(u) = \int_0^T F(t, x(t), u(t)) dt + G(x(T))$$

where

1. $u : (0, T) \rightarrow K \subset \mathbf{R}^m$ is the control that is assumed measurable and takes values on a given closed set K (not assumed bounded);
2. $x : (0, T) \rightarrow \mathbf{R}^d$ is the state of the system governed by the equation of state

$$(4.2) \quad x'(t) = f(t, x(t), u(t)), \quad x(0) = x_0 \in \mathbf{R}^d;$$

the function $f : (0, T) \times \mathbf{R}^d \times \mathbf{R}^m \rightarrow \mathbf{R}^d$ is assumed to be measurable on the variable t , continuous in (x, u) and satisfies a uniform Lipschitz condition with respect to x

$$(4.3) \quad |f(t, x_1, u) - f(t, x_2, u)| \leq \gamma |x_1 - x_2|, \quad \gamma > 0,$$

so that the problem in (4.2) always has a unique solution in $W^{1,1}(0, T)$;

3. the integrand $F : (0, T) \times \mathbf{R}^d \times \mathbf{R}^m \rightarrow \mathbf{R} \cup \{+\infty\}$ is assumed to be continuous on the pair (x, u) and measurable on t ;
4. the function $G : \mathbf{R}^d \rightarrow \mathbf{R}$ is assumed to be continuous.

Let us further set

$$g(u) = \sup \{ |f(t, x, u)| : 0 < t < T, x \in \mathbf{R}^d \}$$

and

$$F_K(t, x, u) = \begin{cases} F(t, x, u), & u \in K, \\ +\infty, & \text{else.} \end{cases}$$

Notice that the function g is continuous with respect to x , even Lipschitz continuous.

It is well-known that this type of optimal control problem, under further suitable technical hypotheses, always has solutions provided that the equation of state is linear on the control variable, the constraint set $K \subset \mathbf{R}^m$ is convex and the integrand F is also convex with respect to the control variable. Indeed the following theorem holds.

Theorem 4.1. *Assume, in addition to the previous hypotheses, the following*

1. *coercivity*

$$(4.4) \quad c(|u|^p - 1) \leq F(t, x, u), \quad p > 1, c > 0;$$

2. F_K is a convex function of u ;
3. we have

$$(4.5) \quad \lim_{|u| \rightarrow \infty} \frac{g(u)}{|u|^p} = 0;$$

4. for $f_1 : (0, T) \times \mathbf{R}^d \rightarrow \mathbf{M}^{m \times d}$ and $f_2 : (0, T) \times \mathbf{R}^d \rightarrow \mathbf{R}^d$,

$$f(t, x, u) = f_1(t, x)u + f_2(t, x).$$

If I is not identically $+\infty$, the optimal control problem has at least one solution.

Under different sets of technical assumptions this result is classical. We would like to describe relaxation of optimal control problems when the existence hypotheses do not hold and nonexistence is expected. As in problems in the calculus of variations, nonexistence (under coercivity assumptions) is always related to oscillatory behavior which in turn is induced by a lack of convexity. In the more general context of optimal control problems, nonlinear state equations by themselves (assuming convexity of K and F) could induce oscillations and nonexistence, although as we shall see some “lack of convexity” must be involved in order for classical solutions to fail to exist.

Notice that according to Theorem 2.3 there is always a Young measure associated to a bounded sequence $\{u_j\}$ in $L^p(\Omega)$ for $p > 1$. Some other basic facts we need include a result giving sufficient conditions in order to have the equiintegrability necessary to be able to represent weak limits in terms of integrals against the Young measure, another result dealing with the situation where we have strong convergence and how this fact gets reflected on the Young measure, and an inequality related to concentration effects ([53]).

Proposition 4.2. *Let $\psi : \Omega \times \mathbf{R}^m \rightarrow \mathbf{R}$ be a Carathéodory function such that*

$$|\psi(x, \lambda)| \leq \tilde{\psi}(|\lambda|), \text{ for a.e. } x \in \Omega,$$

where $\tilde{\psi} \in L^\infty_{loc}(\mathbf{R})$. Let $\{z_j\}$ be a sequence with

$$\sup_j \int_\Omega g(|z_j|) dx = C < \infty,$$

where g is a continuous, non-decreasing function and $\lim_{t \rightarrow \infty} g(t) = \infty$. If

$$\lim_{t \rightarrow \infty} \frac{\tilde{\psi}(t)}{g(t)} = 0,$$

then

$$\psi(x, z_j(x)) \rightharpoonup \bar{\psi}(x) = \int_{\mathbf{R}^m} \psi(x, \lambda) d\nu_x(\lambda) \text{ in } L^1(\Omega),$$

where $\nu = \{\nu_x\}_{x \in \Omega}$ is the parametrized measure associated to the sequence $\{z_j\}$ (or possibly to a subsequence).

Proposition 4.3. *Let $z_j = (x_j, u_j) : \Omega \rightarrow \mathbf{R}^d \times \mathbf{R}^m$ be a bounded sequence in $L^p(\Omega)$ such that $\{x_j\}$ converges strongly to x in $L^p(\Omega)$. If $\mu = \{\mu_t\}_{t \in \Omega}$ is the parametrized measure associated to $\{z_j\}$, then $\mu_t = \delta_{x(t)} \otimes \nu_t$ a.e. $t \in \Omega$, where $\{\nu_t\}_{t \in \Omega}$ is the parametrized measure corresponding to $\{u_j\}$.*

Proposition 4.4. *If $\{z_j\}$ is a sequence of measurable functions with associated parametrized measure $\nu = \{\nu_x\}_{x \in \Omega}$, then*

$$\liminf_{j \rightarrow \infty} \int_E \psi(x, z_j(x)) dx \geq \int_E \int_{\mathbf{R}^m} \psi(x, \lambda) d\nu_x(\lambda) dx,$$

for every Carathéodory function ψ , bounded from below, and every measurable subset $E \subset \Omega$.

We now undertake the proof of Theorem 4.1 by Young measure techniques because the main ingredients of a relaxed formulation are present in it. Let \mathcal{U} be the set of admissible controls

$$\mathcal{U} = \{u : u \in L^p(0, T), u(t) \in K\},$$

and let

$$m = \inf \{I(u) : u \in \mathcal{U}\} \in \mathbf{R}.$$

Let $\{u_j\}$ be any minimizing sequence for the optimal control problem so that $I(u_j) \searrow m$, and let x_j be the associated states

$$x_j'(t) = f(t, x_j(t), u_j(t)), \quad x_j(0) = x_0.$$

By the coercivity assumed on F the sequence $\{u_j\}$ is bounded in $L^p(0, T)$. By Theorem 2.3, let $\nu = \{\nu_t\}_{t \in (0, T)}$ be the associated (possibly to some subsequence) Young measure. On the other hand

$$\begin{aligned} |x_j(t) - x_0| &= \left| \int_0^t f(s, x_j(s), u_j(s)) ds \right| \\ &\leq \int_0^t g(u_j(s)) ds \\ &\leq C \|u_j\|_{L^p(0, T)}^p, \end{aligned}$$

because of the behavior of g at infinity. This estimate implies that x_j is bounded in $L^\infty(0, T)$. Let, according to Theorem 2.3, $\mu = \{\mu_t\}_{t \in (0, T)}$ be the joint Young measure associated to the sequence of pairs $\{(x_j, u_j)\}$. By Proposition 4.2 we can conclude that the sequence $\{f(t, x_j(t), u_j(t))\}$ converges weakly in $L^1(0, T)$ and thus its weak limit is

$$\bar{f}(t) = \int_{\mathbf{R}^d \times \mathbf{R}^m} f(t, \lambda_1, \lambda_2) d\mu_t(\lambda_1, \lambda_2).$$

Define $x(t)$ by

$$x(t) = x_0 + \int_0^t \bar{f}(s) ds.$$

Then

$$x_j(t) - x(t) = \int_0^t [f(s, x_j(s), u_j(s)) - \bar{f}(s)] ds$$

tends to 0 for a.e. $t \in (0, T)$. This pointwise convergence together with the uniform bounds implies in particular that $x_j \rightarrow x$ in $L^p(0, T)$. Consequently, by Proposition 4.3 we obtain

$$\mu_t = \delta_{x(t)} \otimes \nu_t$$

for a.e. $t \in (0, T)$. In this case

$$\begin{aligned} \bar{f}(t) &= \int_{\mathbf{R}^m} f(t, x(t), \lambda) d\nu_t(\lambda), \\ x'(t) &= \int_{\mathbf{R}^m} f(t, x(t), \lambda) d\nu_t(\lambda), \quad x(0) = x_0. \end{aligned}$$

Finally by Proposition 4.4 and Jensen's inequality (F_K is convex)

$$\begin{aligned} m &= \lim_{j \rightarrow \infty} \int_0^T F(t, x_j(t), u_j(t)) dt \\ &\geq \int_0^T \int_K F(t, x(t), \lambda) d\nu_t(\lambda) dt \\ &\geq \int_0^T F(t, x(t), u(t)) dt, \end{aligned}$$

where $u(t)$, the first moment of ν_t , is the weak limit of u_j . Because the convexity of F_K entails the convexity of K , u is an admissible control. In order to conclude the proof of the theorem we need to show that x and u are related through the equation of state. This is the case if and only if

$$f(t, x(t), u(t)) = \int_K f(t, x(t), \lambda) d\nu_t(\lambda),$$

and this can only happen (for every possible family ν_t) if f is linear on the control variable u . \square

After the proof of the linear case, it seems natural that a generalized Young measure solution, ν , of the optimal control problem should give rise to the state x satisfying the differential law

$$(4.6) \quad x'(t) = \int_K f(t, x(t), \lambda) d\nu_t(\lambda), \quad x(0) = x_0.$$

This initial value problem is well-posed because the function

$$(t, x) \mapsto \int_K f(t, x(t), \lambda) d\nu_t(\lambda)$$

verifies the Lipschitz condition in x necessary for finding a solution.

On the other hand we define the cost \tilde{I} of such family of probability measures by

$$\tilde{I}(\nu) = \int_0^T \int_K F(t, x(t), \lambda) d\nu_t(\lambda) dt,$$

where as usual $x(t)$ is the solution of (4.6). The admissibility conditions on the competing Young measures must include some reference to the set K . Indeed it is an easy exercise to prove the following.

Proposition 4.5. *A Young measure $\nu = \{\nu_x\}_{x \in \Omega}$ has its support contained in a set K if and only if it can be generated by a sequence $\{u_j\}$ with*

$$|\{u_j \notin K\}| \rightarrow 0, \quad \text{as } j \rightarrow \infty.$$

On the other hand, Theorem 3.1 for the nongradient, scalar case yields the only restriction on the integrability of the integral against the p th power. Hence

$$\tilde{\mathcal{U}} = \left\{ \nu = \{\nu_t\}_{t \in (0, T)} : \text{supp}(\nu_t) \subset K, \text{ a.e. } t \in (0, T), \int_0^T \int_K |\lambda|^p d\nu_t(\lambda) dt < \infty \right\}.$$

Theorem 4.6. *If in addition to the coercivity condition (4.4) we have the upper bound*

$$F(t, x, u) \leq h(x) (1 + |u|^p)$$

where $h \in L_{loc}^\infty(\mathbf{R}^d)$, then

$$\inf \{I(u) : u \in \mathcal{U}\} = \min \{\tilde{I}(\nu) : \nu \in \tilde{\mathcal{U}}\}.$$

Notice the use of min in the generalized problem to emphasize the existence of a solution.

Proof. The proof is essentially contained in that of Proposition 2.1. It only requires us to take care of some technicalities. If for $u \in \mathcal{U}$ we take $\nu_t = \delta_{u(t)}$, then trivially $\nu \in \tilde{\mathcal{U}}$ and $I(u) = \tilde{I}(\nu)$. This implies that $\tilde{m} \leq m$, if m and \tilde{m} stand for the two infima in the statement of the theorem, respectively.

Assume that $\nu \in \tilde{\mathcal{U}}$ is given. By Theorem 3.1 for the nongradient case there exists a sequence of admissible controls $\{u_j\}$ such that $\{|u_j|^p\}$ is equiintegrable. Let x_j be the solution of the problem

$$x_j'(t) = f(t, x_j(t), u_j(t)), \quad x_j(0) = x_0.$$

As in the proof of the linear case, Theorem 4.1, we conclude that x_j converges strongly in $L^p(0, T)$ to x , which is a solution of (4.6) for ν the Young measure associated to $\{u_j\}$. Moreover the joint Young measure associated to the sequence of pairs $\{(x_j, u_j)\}$ is $\delta_{x(t)} \otimes \nu_t$. Because of the upper bound on F and the equiintegrability of $\{|u_j|^p\}$ we have that $\{F(t, x_j(t), u_j(t))\}$ is also equiintegrable and hence

$$\lim_{j \rightarrow \infty} \int_0^T F(t, x_j(t), u_j(t)) dt = \int_0^T \int_K F(t, x(t), \lambda) d\nu_t(\lambda) dt = \tilde{I}(\nu).$$

Because of the arbitrariness of $\nu \in \tilde{\mathcal{U}}$ we obtain $m = \tilde{m}$.

In order to show that there is a minimizer in $\tilde{\mathcal{U}}$, let $\{u_j\}$ be minimizing for I and ν the associated Young measure. As before, and due to Proposition 4.4,

$$m = \lim_{j \rightarrow \infty} \int_0^T F(t, x_j(t), u_j(t)) dt \geq \int_0^T \int_K F(t, x(t), \lambda) d\nu_t(\lambda) dt = \tilde{I}(\nu) \geq m.$$

Therefore ν is a minimizer for \tilde{I} . \square

Based on this generalized formulation, one can try to recover existence of classical solutions a posteriori. One important such existence result (typically called Filippov's theorem) assumes, roughly speaking, the convexity of the image of the set of controls under the right-hand side of the equation of state. More general statements involve the convexity of the sets

$$(4.7) \quad \mathcal{A}_{(t,x)} = \{(\alpha, f(t, x, u)) : \alpha \geq F(t, x, u), u \in K\}$$

for each pair (t, x) ([18]). One can prove this existence result by examining generalized optimality conditions that such Young measure solutions must verify. Even in the absence of such convexity conditions, existence of classical solutions can be achieved by showing that optimal paths do not intersect the "nonexistence" set for the problem. See [49]. Optimal control problems governed by partial differential equations are much more complex. The next two sections are good examples of this.

5. SOME MODELS IN MATERIALS SCIENCE

Our example in this section refers to an energy model for magnetostriction where interactions of elastic and magnetic properties are the goal of the analysis. The energy functional is

$$I(m, y) = \int_{\Omega} \varphi_0(m(y(x)), \nabla y(x)) dx + \frac{1}{2} \int_{\mathbf{R}^2} |\nabla u(z)|^2 dz,$$

where

$$\begin{aligned} y &= y_0 \text{ on } \partial\Omega, \\ |\det(\nabla y(x)) m(y(x))| &= 1 \text{ a.e. } x \in \Omega, \end{aligned}$$

and

$$\operatorname{div}(-\nabla u + m\chi_{y(\Omega)}) = 0, \quad \text{in } H^{-1}(\mathbf{R}^2).$$

y_0 is a smooth, one-to-one deformation onto $y_0(\Omega)$. The vector field m is the magnetization over the deformed configuration $y_0(\Omega)$, and the constraint involving the determinant of the deformation gradient reflects a typical saturation hypothesis. u is the associated potential. See [30], [31], [32], [33], [34], [35].

After a change of variables and of notation, the functional can be rewritten as

$$I(m, y) = \int_{\mathbf{R}^2} \varphi(x, m(x), \nabla y(x)) dx + \frac{1}{2} \int_{\mathbf{R}^2} |\nabla u(x)|^2 dx,$$

where

$$\varphi(x, \lambda, A) = \begin{cases} 0, & x \in \mathbf{R}^2 \setminus \tilde{\Omega}, \lambda = 0, A = \nabla(Ty_0)^{-1}(x), \\ |\lambda| \varphi_0(\lambda, A^{-1}), & x \in \tilde{\Omega}, |\lambda| = \det A > 0, \\ +\infty, & \text{else.} \end{cases}$$

In considering Young measures associated to sequences $\{(m_j, \nabla y_j)\}$ we immediately encounter the difficulty attached to the fact that $\{\nabla u_j\}$ are determined in a nonlocal fashion through the differential constraint. What this means is that the second contribution to the energy cannot be expressed in terms of the Young measure corresponding to a minimizing sequence of pairs $\{(m_j, \nabla y_j)\}$.

In this full generality, it is almost impossible to examine relaxation. Something can be said under important simplifying assumptions which on the other hand are close to familiar hypotheses in the linear theories of magnetostriction. Such conditions refer to the structure of the energy density φ . Since the discussion would take a technical route, we refer the reader to [56] and concentrate now for the sake of illustration on the simpler case of micromagnetics where we neglect elastic features. We assume our body is rigid, and then the energy functional is given by

$$I(u) = \int_{\Omega} \varphi_0(u(x)) dx + \frac{1}{2} \int_{\mathbf{R}^2} |\nabla y(x)|^2 dx,$$

where

$$(5.1) \quad |u| = 1, \quad \operatorname{div}(-\nabla y + u\chi_{\Omega}) = 0 \text{ in } H^{-1}(\mathbf{R}^2).$$

In this case the nonconvex constraint $|u| = 1$ forces us to look for a relaxed formulation despite the fact that φ_0 might be convex. Since the second contribution to the energy is not a local quantity of u , before proceeding with an analysis based on Young measures, we are in need of a local formulation of the problem. This can be

done in our case by taking advantage of the structure of the differential law (5.1). In dimension $N = 2$, we have the equivalence

$$\operatorname{div}(-\nabla u + m) = 0 \quad \text{if and only if} \quad m(x) = \nabla u(x) + T\nabla v(x),$$

where T is a $\pi/2$ rotation in the plane. The nonlocality expressed by the differential equation can thus be handled as a local condition with the pairs $\{(\nabla u(x), \nabla v(x))\}$, a vector gradient. In this way, the energy functional can be written and analyzed as a local functional on $\{(\nabla u(x), \nabla v(x))\}$. This analysis would yield information on all minimizing sequences of pairs $\{(\nabla u_j(x), \nabla v_j(x))\}$. However, since the relevant vector field is the magnetization m , it is interesting to express our conclusions in terms of Young measures associated to sequences of magnetization. In this sense, we would like to explicitly compute the generalized functional

$$I(\nu) = \inf \left\{ \liminf_{j \rightarrow \infty} I(m_j) : \{m_j\} \text{ generates } \nu \right\},$$

where ν is the Young measure associated to sequences $\{m_j\}$. Let us emphasize once again that for local problems the explicit form of $I(\nu)$ in terms of ν is straightforward. For nonlocal problems, the computation of $I(\nu)$ is not so. Indeed, if we take two different sequences $\{m_j\}$ and $\{m'_j\}$ giving rise to the same Young measure $\nu = \{\nu_x\}_{x \in \Omega}$, since the functional also depends upon y_j and y'_j through the differential constraint (5.1), there is no guarantee (and indeed it is not true in most of the cases) that

$$\lim_{j \rightarrow \infty} I(m_j) = \lim_{j \rightarrow \infty} I(m'_j),$$

despite the fact that both generate the same ν . Since the particular form of the energy functional does not play a role in our analysis, we will assume

$$I(m) = \int_{\Omega} \varphi(x, u(x), m(x)) dx + \int_{\Omega} \psi(x, u(x), \nabla u(x)) dx,$$

where $m \in L^2(\Omega)$ and $u \in H_0^1(\Omega)$ are coupled by the nonlocal, differential constraint

$$\operatorname{div}(-\nabla u + m) = 0, \quad \text{in } H^{-1}(\Omega).$$

In covering the second step of the computation of $I(\nu)$, we will find a way of generating sequences $\{m_j\}$ such that $\lim_{j \rightarrow \infty} I(m_j) = I(\nu)$. When ν is a minimizer, such sequence will be minimizing. Another issue of interest is to detect all minimizing sequences for I . If the hypotheses on the density ψ are strengthened to incorporate strict convexity with respect to the gradient variable, then all minimizing sequences are essentially the ones obtained through the next theorem ([56]).

Theorem 5.1. *Under the notation and assumptions introduced so far, if ψ is convex with respect to the gradient variable and*

$$c(|m|^2 - 1) \leq \varphi(x, u, m) \leq C(|m|^2 + 1), \quad C > c > 0,$$

then

$$(5.2) \quad I(\nu) = \int_{\Omega} \int_{\mathbf{R}^2} \varphi(x, u(x), \lambda) d\nu_x(\lambda) dx + \int_{\Omega} \psi(x, u(x), \nabla u(x)) dx,$$

where ν and u are coupled again by the differential equation

$$\begin{aligned} \operatorname{div}(-\nabla u + m) &= 0, \quad \text{in } H_0^{-1}(\Omega), \\ m(x) &= \int_{\mathbf{R}^2} \lambda \, d\nu_x(\lambda). \end{aligned}$$

Proof. Let $\tilde{I}(\nu)$ be the right-hand side in (5.2), for the time being. Due to the weak continuity of the operator

$$m \mapsto u \quad \text{where} \quad \operatorname{div}(-\nabla u + m) = 0,$$

it is very easy to see because of the convexity of ψ and Proposition 4.4 that

$$I(\nu) \geq \tilde{I}(\nu).$$

Notice that we have to use Proposition 4.3 in order to deal with the variable u in both terms of I .

Let ν be given, and let $m(x) \in L^2(\Omega)$ be its first moment

$$m(x) = \int_{\mathbf{R}^2} \lambda \, d\nu_x(\lambda).$$

Write, solving the differential constraint,

$$m(x) = \nabla u(x) + T\nabla v(x).$$

For each $x \in \Omega$ define the probability measure μ_x through the identity

$$\langle \chi, \mu_x \rangle = \int_{\mathbf{R}^2} \chi(T^{-1}\lambda - T^{-1}\nabla u(x)) \, d\nu_x(\lambda).$$

We claim that the family $\mu = \{\mu_x\}_{x \in \Omega}$ is a gradient Young measure. According to Theorem 3.1 (scalar case), all we have to check is that the first moment is the gradient of an $H^1(\Omega)$ -function and that the integral of the second moment is finite. In fact, it is elementary to compute

$$\langle \lambda, \mu_x \rangle = \nabla v(x),$$

and

$$\int_{\Omega} \int_{\mathbf{R}^2} |\lambda|^2 \, d\mu_x(\lambda) \, dx \leq \int_{\Omega} \left(\int_{\mathbf{R}^2} |\lambda|^2 \, d\nu_x(\lambda) + |\nabla u(x)|^2 \right) \, dx < \infty.$$

Therefore there exists a bounded sequence in $H^1(\Omega)$, $\{v_j\}$, such that $\{\nabla v_j(x)\}$ generates μ and $\{|\nabla v_j|^2\}$ is equiintegrable. Put

$$m_j(x) = \nabla u(x) + T\nabla v_j(x), \quad x \in \Omega.$$

We claim that m_j generates ν : if $\chi(x, \lambda)$ is any smooth function

$$\begin{aligned} \lim_{j \rightarrow \infty} \int_{\Omega} \chi(x, m_j(x)) \, dx &= \lim_{j \rightarrow \infty} \int_{\Omega} \chi(x, \nabla u(x) + T\nabla v_j(x)) \, dx \\ &= \int_{\Omega} \int_{\mathbf{R}^2} \chi(x, \nabla u(x) + T\lambda) \, d\mu_x(\lambda) \, dx \\ &= \int_{\Omega} \int_{\mathbf{R}^2} \chi(x, \lambda) \, d\nu_x(\lambda) \, dx. \end{aligned}$$

Finally, because of the choice of m_j we have

$$\operatorname{div}(-\nabla u + m_j) = 0, \quad \text{for all } j,$$

so that it is easy to have $\lim_{j \rightarrow \infty} I(m_j) = \tilde{I}(\nu)$. It is important to bear in mind the upper bound on φ and the equiintegrability of $\{|m_j|^2\}$ in order to pass to the limit. \square

This proof essentially consists in specifying how to go from a given Young measure ν to a generating sequence yielding the value $I(\nu)$ in the limit :

1. compute the first moment

$$m(x) = \int_{\mathbf{R}^2} \lambda \, d\nu_x(\lambda);$$

2. solve the differential constraint

$$\operatorname{div}(-\nabla u + m) = 0;$$

3. look at the family of probability measures

$$\langle \chi, \mu_x \rangle = \int_{\mathbf{R}^2} \chi(T^{-1}\lambda - T^{-1}\nabla u(x)) \, d\nu_x(\lambda),$$

and find a generating sequence of gradients $\{\nabla v_j\}$;

4. the sequence

$$m_j(x) = \nabla u(x) + T\nabla v_j(x),$$

will provide the value of $I(\nu)$: $I(\nu) = \lim_{j \rightarrow \infty} I(m_j)$.

It is interesting to notice that for the magnetostriction case the underlying ideas are essentially the same. However, the proof for the analogous result relies on the characterization of gradient Young measures for the vector case.

6. OPTIMAL DESIGN PROBLEMS

Our first optimal design example is also an optimal control problem governed by a second order ordinary differential equation completed by boundary values. For this reason it has the flavor of an optimal design problem. It is a one-dimensional simplified version of an important optimal design problem concerned with two phase conductors that we will briefly discuss afterwards.

Suppose $\alpha > 0$ and $f \in L^2(0, 1)$ are given. Let $F(x, u, y)$ be continuous in (u, y) and measurable in x . Let

$$\mathcal{U} = \{u \in L^1(0, 1) : u \geq \alpha\}$$

be the set of competing functions for our optimization problem, and let our functional be defined by

$$I(u) = \int_0^1 F(x, u(x), y(x)) \, dx$$

where $y \in H_0^1(0, 1)$ is the solution of $-(uy')' = f$. The solution of this boundary problem can be given in closed form in terms of u and f by

(6.1)

$$y(x) = - \int_0^x \left(\frac{1}{u(s)} \int_0^s f(\tau) \, d\tau \right) ds + \frac{\int_0^x \frac{1}{u(s)} \, ds}{\int_0^1 \frac{1}{u(s)} \, ds} \int_0^1 \left(\frac{1}{u(s)} \int_0^s f(\tau) \, d\tau \right) ds.$$

If the chosen, second order equation would have been $-(uy')' + uy = f$, the solution could not be written as explicitly as in (6.1), but the same ideas that follow hold almost without change. Assume that $\{u_j\}$ is a sequence in \mathcal{U} . The sequence of inverses $\{1/u_j\}$ is bounded in $L^\infty(0, 1)$ (by $1/\alpha$), and it generates a Young measure $\nu = \{\nu_x\}_{x \in (0,1)}$ supported in the interval $[0, 1/\alpha]$. If we let

$$\frac{1}{u(x)} = \int_0^{1/\alpha} \lambda d\nu_x(\lambda)$$

so that $1/u_j \xrightarrow{*} 1/u$ in $L^\infty(\mathbf{R})$, then by (6.1) $y_j \rightarrow y$ strongly where y is the corresponding solution of the boundary problem associated to u . By Proposition 4.3 we have

$$\lim_{j \rightarrow \infty} I(u_j) = \lim_{j \rightarrow \infty} \int_0^1 F(x, u_j(x), y_j(x)) dx = \int_0^1 \int_0^{1/\alpha} F(x, 1/\lambda, y(x)) d\nu_x(\lambda) dx.$$

If the function

$$\tilde{F}(x, u, y) = F(x, 1/u, y)$$

is convex in u , then by Jensen's inequality we obtain

$$\int_0^1 F(x, u(x), y(x)) dx \leq \int_0^1 \int_0^{1/\alpha} F(x, 1/\lambda, y(x)) d\nu_x(\lambda) dx = \lim_{j \rightarrow \infty} I(u_j).$$

If $\{u_j\}$ is minimizing, we conclude that u is a minimizer of the problem. However if the convexity of \tilde{F} fails to be true, a relaxation should be considered. Let

$$\tilde{\mathcal{U}} = \left\{ \nu = \{\nu_x\}_{x \in (0,1)} : \text{supp}(\nu_x) \subset [0, 1/\alpha], \text{ a.e. } x \in (0, 1) \right\}$$

and define

$$\tilde{I}(\nu) = \int_0^1 \int_0^{1/\alpha} F(x, 1/\lambda, y(x)) d\nu_x(\lambda) dx$$

where y is the solution of the boundary value problem

$$-(uy')' = f, \quad y \in H_0^1(0, 1), \quad u(x) = \left(\int_0^{1/\alpha} \lambda d\nu_x(\lambda) \right)^{-1}.$$

Because in this situation we are working with L^∞ -bounds and the weak convergence of sequences is not an issue, Proposition 2.1 can be applied and generating sequences of the generalized minimizer will be minimizing for the original optimization problem.

A more realistic situation might incorporate two main features not considered in the previous setting: the control u can only take on two selected, nonnegative values α or β , and the cost functional could also depend upon y' . For the one-dimensional situation the analysis can proceed based on (6.1), making some adjustments on the previous arguments. In order to motivate the two-dimensional situation, we briefly describe a different perspective on the problem. Although it may sound a little bit artificial in the one-dimensional setting, we believe it is fruitful in the higher dimensional case. For the characteristic function χ of a subinterval of $(0, 1)$, put $u = \chi\alpha + (1 - \chi)\beta$ and

$$I(\chi) = \int_0^1 F(x, u(x), y(x), y'(x)) dx,$$

where $-(uy')' = f$, $y \in H_0^1(0, 1)$ and $F : (0, 1) \times \mathbf{R}^3 \rightarrow \mathbf{R}^*$ is a suitable cost density. Because the unique values α and β are allowed for u , we can also write

$$I(\chi) = \int_0^1 [\chi(x)F_\alpha(x, y(x), y'(x)) + (1 - \chi(x))F_\beta(x, y(x), y'(x))] dx.$$

On the other hand if $-U'' = f$, $U \in H_0^1(0, 1)$, then $-(uy' + U')' = 0$ and thus $u(x)y'(x) + U'(x)$ has to be constant throughout $(0, 1)$. If we define a density $\varphi : (0, 1) \times \mathbf{R}^3 \rightarrow \mathbf{R}^*$ by

$$\varphi(x, y, y', k) = \begin{cases} F_\alpha(x, y, y'), & \alpha y' + U'(x) = k, \\ F_\beta(x, y, y'), & \beta y' + U'(x) = k, \\ +\infty, & \text{else,} \end{cases}$$

then our optimal design criterion can be expressed through φ by

$$I(y, k) = \int_0^1 \varphi(x, y(x), y'(x), k) dx,$$

where $y \in H_0^1(0, 1)$ and k must be constant in $(0, 1)$. The main advantage of this perspective is that the pairs (y, k) are independent; there is no differential constraint involved. Or if you like, the differential, nonlocal constraint has been recorded in the fact that k should be constant throughout $(0, 1)$.

Instead of pursuing further the analysis of the one-dimensional case, we would rather move to the two-dimensional situation. Suppose we are given two numbers $\alpha, \beta > 0$ and a function $f \in L^2(\Omega)$ where Ω is some regular domain in \mathbf{R}^2 . Our set of competing objects \mathcal{U} is now the class of all measurable subsets A of Ω . To such a set $A \subset \Omega$ we associate the function $u = \alpha\chi_A + \beta\chi_{\Omega \setminus A}$ and solve the boundary value problem

$$(6.2) \quad -\operatorname{div}(u\nabla y) = f, \quad y \in H_0^1(\Omega).$$

The cost is then computed through the integral

$$I(A) = \int_\Omega F(x, u(x), y(x), \nabla y(x)) dx,$$

where F is some given integrand. Notice that by Poincaré's inequality the L^2 -norm of the gradient of the solution of (6.2) is always bounded by a constant depending on α, β and f and consequently sequences of solutions corresponding to sequences of sets A_j have their gradients always uniformly bounded in $L^2(\Omega)$.

If we assume that

$$\lim_{|A| \rightarrow \infty} \frac{|F(x, u, y, \lambda)|}{|\lambda|^2} = 0$$

uniformly in (x, u, y) , for a given sequence of sets $\{A_j\}$ we can write by Proposition 4.2

$$\lim_{j \rightarrow \infty} I(A_j) = \int_\Omega \int_{\mathbf{R} \times \mathbf{R}^2} F(x, u, y(x), \lambda) d\mu_x(u, \lambda) dx$$

where $\mu = \{\mu_x\}_{x \in \Omega}$ is the joint Young measure associated to $\{(u_j, \nabla y_j)\}$, y_j is the solution of (6.2) corresponding to u_j and

$$\nabla y(x) = \int_{\mathbf{R} \times \mathbf{R}^2} \lambda d\mu_x(u, \lambda)$$

is the weak limit in $H^1(\Omega)$ of $\{\nabla y_j\}$. Hence if we define

$$\tilde{I}(\mu) = \int_{\Omega} \int_{\mathbf{R} \times \mathbf{R}^2} F(x, u, y(x), \lambda) d\mu_x(u, \lambda) dx$$

for such μ 's, we will have generalized minimizers in the form of Young measures associated to sequences $\{(u_j, \nabla y_j)\}$. A different and much harder question is to characterize such family of Young measures.

An alternative to circumvent these difficulties consists in expressing the nonlocal, differential constraint in a local way as suggested in the one-dimensional problem. Indeed if $U \in H_0^1(\Omega)$ is the solution of the problem

$$-\Delta U = f, \quad \text{in } H^{-1}(\Omega),$$

then $\operatorname{div}(-u\nabla y + \nabla U) = 0$. Therefore if T is a $\pi/2$ -rotation, there exists a stream function $v \in H^1(\Omega)$ such that $\nabla U = u\nabla y + T\nabla v$. This identity enables us to think of the pair $w = (u, v) : \Omega \rightarrow \mathbf{R}^2$ as our "free" variable, a vector field in \mathbf{R}^2 . If

$$I(A) = \int_{\Omega} [\chi_A(x)F_{\alpha}(x, y(x), \nabla y(x)) + (1 - \chi_A(x))F_{\beta}(x, y(x), \nabla y(x))] dx,$$

and we define

$$\varphi(x, w, Y) = \begin{cases} F_{\alpha}(x, w^{(1)}, Y^{(1)}), & \alpha Y^{(1)} + TY^{(2)} = \nabla U(x), \\ F_{\beta}(x, w^{(1)}, Y^{(1)}), & \beta Y^{(1)} + TY^{(2)} = \nabla U(x), \\ +\infty, & \text{else,} \end{cases}$$

where $Y^{(i)}$ designates the i -th row of Y and $w^{(i)}$ the i -th component of w , then the cost functional can be rewritten in the following terms:

$$I(w) = \int_{\Omega} \varphi(x, w(x), \nabla w(x)) dx,$$

where $w \in H^1(\Omega)$, but $w^{(1)} \in H_0^1(\Omega)$. The definition of φ is ambiguous for matrices of the form

$$\begin{pmatrix} 0 \\ T^{-1}\nabla U(x) \end{pmatrix},$$

but this can be easily overcome. A more important issue is how to treat a typical volume constraint

$$\frac{|A|}{|\Omega|} = \frac{1}{|\Omega|} \int_{\Omega} \chi_A(x) dx = \lambda.$$

If we let

$$\psi(x, w, Y) = \begin{cases} 1, & \alpha Y^{(1)} + TY^{(2)} = \nabla U(x), \\ +\infty, & \text{else,} \end{cases}$$

we must ask for the condition

$$\frac{1}{|\Omega|} \int_{\Omega} \psi(x, w(x), \nabla w(x)) dx = \lambda.$$

This approach leads us to consider the envelope

$$\Psi(x, w, Y, t) = \frac{1}{|\Omega|} \inf \left\{ \int_{\Omega} \varphi(x, w, Y + \nabla W(y)) dy : W \in W_0^{1,\infty}(\Omega), \int_{\Omega} \psi(x, w, Y + \nabla W(y)) dy = t|\Omega| \right\}.$$

This is some sort of constrained quasiconvexity. It can be considered as a relaxation of the original optimal design problem. The constitutive assumption that permits us to show existence of an optimal solution for a variational principle written in terms of Ψ is the joint convexity with respect to (Y, t) (see [24]). The analysis of this envelope, and in particular its definition in terms of gradient Young measures and its joint convexity, is the goal of [57].

Shape optimization problems are again optimal control or optimal design problems with some specific properties. A typical such problem consists in finding the “best” shape according to some specific criterion.

Assume we are given a set $\Omega \subset \mathbf{R}^N$ and a function $f \in L^2(\Omega)$. Let \mathcal{U} stand for the class of all possible shapes in Ω , that is to say

$$\mathcal{U} = \{A \subset \Omega : A \text{ is regular}\}, \quad |A| \text{ given,}$$

and define the cost of each admissible shape A as

$$I(A) = \int_{\Omega} F(x, y_A(x), \nabla y_A(x)) dx$$

where $y_A \in H_0^1(A)$ is the solution of $-\Delta y_A = f$ in A .

This basic shape optimization problem can be reformulated in a more transparent way in the following terms. If we solve the problem $-\Delta y = f$ in Ω , $y \in H_0^1(\Omega)$ and put $u_A = y_A - y$ for each $A \in \mathcal{U}$, then u_A is harmonic in A and $u_A = y$ on ∂A . If we further set $u_A = y$ outside A , then given $y \in H_0^1(\Omega)$ and $A \in \mathcal{U}$, we find that u_A consists in replacing y in A by the harmonic function that agrees with y in ∂A . Then

$$I(A) = \int_{\Omega} F(x, u_A(x), \nabla u_A(x)) dx,$$

possibly with a different integrand F . Notice that because u_A is harmonic when it does not coincide with y , we always have the estimate

$$\int_{\Omega} |\nabla u_A(x)|^2 dx \leq \int_{\Omega} |\nabla y(x)|^2 dx$$

for any $A \in \mathcal{U}$.

Assume we have a minimizing sequence of optimal shapes $\{A_j\}$, and for simplicity set $u_j \equiv u_{A_j}$. Assume further that

$$\lim_{|\lambda| \rightarrow \infty} \frac{|F(x, u, \lambda)|}{|\lambda|^2} = 0$$

uniformly in (x, u) . If $\nu = \{\nu_x\}_{x \in \Omega}$ is the Young measure associated with $\{\nabla u_j\}$, then because of the above bound and Proposition 4.2 we will have the representation

$$\lim_{j \rightarrow \infty} I(A_j) = \int_{\Omega} \int_{\mathbf{R}^N} F(x, u(x), \lambda) d\nu_x(\lambda) dx,$$

where $\{u_j\}$ converges weakly in $H^1(\Omega)$ and strongly in $L^2(\Omega)$ to u . In this way we establish existence of optimal solutions within the class of Young measures associated to sequences of gradients coming from sequences of subsets of Ω . In order to pursue further the analysis in this framework one would need to characterize all those families of Young measures. This looks like a hard problem even in dimension one.

We finally consider an optimal design problem for plates of variable thickness. The functional to be minimized is the compliance of the plate defined to be the work done by a given load F and is regarded as a function of the half-thickness u ,

$$(6.3) \quad I(u) = \int_{\Omega} Fw \, dx.$$

The deflection or vertical displacement w obeys the fourth order elliptic equation

$$(6.4) \quad \sum_{\alpha, \beta, \gamma, \delta=1}^2 \frac{\partial^2}{\partial x_{\alpha} \partial x_{\beta}} \left(M_{\alpha\beta\gamma\delta} \frac{\partial^2 w}{\partial x_{\gamma} \partial x_{\delta}} \right) = F,$$

where $F \in L^2(\Omega)$ is the vertical load on the plate and

$$M_{\alpha\beta\gamma\delta} = \frac{2}{3} u^3(x_1, x_2) B_{\alpha\beta\gamma\delta}.$$

u is the thickness and $B_{\alpha\beta\gamma\delta}$ is a constant tensor. We restrict ourselves to the case where the thickness u is in fact a function of x_1 alone and x_1 belongs to the interval

$$(a, b) = \{x_1 \in \mathbf{R} : \text{there exists some } x_2 \in \mathbf{R} \text{ with } (x_1, x_2) \in \Omega\}.$$

We further assume that the nonzero components of $B_{\alpha\beta\gamma\delta}$ are

$$\begin{aligned} B_{1111} &= B_{2222} = \frac{E}{1-r^2}, \\ B_{1122} &= B_{2211} = \frac{Er}{1-r^2}, \\ B_{1212} &= B_{1221} = B_{2112} = B_{2121} = \frac{E}{2(1+r)}, \end{aligned}$$

where E and r are constants (the Young's modulus and the Poisson ratio, respectively). Equation (6.4) is completed with the boundary conditions

$$(6.5) \quad w = \frac{\partial w}{\partial n} = 0, \quad \text{on } \partial\Omega.$$

The admissibility conditions on the half-thicknesses u that may compete in (6.3) are the following:

$$\mathcal{U} = \left\{ u \in L^{\infty}(a, b) : u_{\min} \leq u(x) \leq u_{\max}, \int_{\Omega} u(x) \, dx = V_0 \right\},$$

where u_{\min} , u_{\max} and V_0 are prescribed a priori in a consistent way

$$0 < u_{\min} |\Omega| < V_0 < u_{\max} |\Omega|.$$

The basic feature of this optimization problem is the lack of minimizers. This has been sufficiently emphasized in the references cited in the Introduction. Let $\tilde{\mathcal{U}}$ be the set of Young measures associated to sequences $\{u_k\}$ of half-thicknesses. The only restriction we have on such families is the support and the volume integral

$$\tilde{\mathcal{U}} = \left\{ \mu = \{\mu_x\}_{x \in (a, b)} : \text{supp } \mu_x \subset Q = [u_{\min}, u_{\max}] \text{ a.e. } x \in (a, b), \int_a^b \int_Q \lambda \, d\mu_x(\lambda) \, dx = V_0 \right\}.$$

In order to extend the compliance functional in $\tilde{\mathcal{U}}$, the following lemma on H-convergence ([45], [62]) is needed. A fourth order tensor $M(x)$ is said to be orthotropic if the non-vanishing coefficients are M_{1111} , M_{2222} and

$$M_{1122} = M_{2211} = M_{1212} = M_{1221}, \quad M_{2112} = M_{2121}.$$

M is bounded by the constants (d, D) if for every symmetric tensor $t = t_{\alpha\beta}$ we have for every $x \in \Omega$

$$\begin{aligned} d|t|^2 &\leq \sum M_{\alpha\beta\gamma\delta} t_{\alpha\beta} t_{\gamma\delta}, \\ \sum |M_{\alpha\beta\gamma\delta} t_{\alpha\beta}| &\leq D|t| \quad \text{for every } \gamma, \delta. \end{aligned}$$

Lemma 6.1. *Let M^k be a sequence of orthotropic tensors bounded uniformly by (d, D) . Let us assume that*

$$\begin{aligned} (M_{1111}^k)^{-1} &\overset{*}{\rightharpoonup} (M_{1111}^\infty)^{-1}, \\ (M_{1122}^k) (M_{1111}^k)^{-1} &\overset{*}{\rightharpoonup} (M_{1122}^\infty) (M_{1111}^\infty)^{-1}, \\ (M_{2222}^k) - (M_{1122}^k) (M_{1111}^k)^{-1} &\overset{*}{\rightharpoonup} (M_{2222}^\infty) - (M_{1122}^\infty) (M_{1111}^\infty)^{-1}, \\ M_{1212}^k &\overset{*}{\rightharpoonup} M_{1212}^\infty. \end{aligned}$$

If w^k , $1 \leq k \leq \infty$, is the solution of (6.4) and (6.5) corresponding to M^k , then $w^k \rightharpoonup w^\infty$ in $H_0^2(\Omega)$.

$H_0^2(\Omega)$ is the subspace of $H^2(\Omega)$ that satisfies (6.5). The different weak limits we should care about in our case in order to apply the lemma are

$$\begin{aligned} (M_{1111}^k)^{-1} &= \left(\frac{2}{3} u_k^3 \frac{E}{1-r^2} \right)^{-1}, \\ (M_{1122}^k) (M_{1111}^k)^{-1} &= r, \\ (M_{2222}^k) - (M_{1122}^k)^2 (M_{1111}^k)^{-1} &= \frac{2}{3} u_k^3 E, \\ (M_{1212}^k) &= \frac{1}{3} u_k^3 \frac{E}{1+r}. \end{aligned}$$

These weak limits can be represented through the moments of order 3 and -3 of the Young measure, μ , corresponding to the sequence $\{u_k\}$. Hence, if we let

$$(6.6) \quad m(x) = \int_Q \lambda^3 d\mu_x(\lambda), \quad c^{-1}(x) = \int_Q \lambda^{-3} d\mu_x(\lambda)$$

and define

$$\begin{aligned} (M_{1111}^\infty)^{-1} &= \left(\frac{2}{3} c(x) \frac{E}{1-r^2} \right)^{-1}, \\ (M_{1122}^\infty) (M_{1111}^\infty)^{-1} &= r, \\ (6.7) \quad (M_{2222}^\infty) - (M_{1122}^\infty)^2 (M_{1111}^\infty)^{-1} &= \frac{2}{3} m(x) E, \\ (M_{1212}^\infty) &= \frac{1}{3} m(x) \frac{E}{1+r}, \end{aligned}$$

then by Lemma 6.1 (the other hypotheses in the lemma are easily verified in this situation), the displacements w_k corresponding to the tensors M^k in (6.4), (6.5) will

converge weakly to the solution of the same problem corresponding to the tensor M^∞ . We must define the compliance \tilde{I} for elements in $\tilde{\mathcal{U}}$ to be

$$\tilde{I}(\mu) = \int_{\Omega} F w dx,$$

where w is the solution of (6.4), (6.5), with the tensor M depending on μ through (6.6) and (6.7). It is elementary to check that \tilde{I} admits minimizers within the class $\tilde{\mathcal{U}}$.

A crucial observation is that \tilde{I} depends only upon the moments of order 3 and -3 of μ in such a way that if μ^1 and μ^2 have in common these two moments, then $\tilde{I}(\mu^1) = \tilde{I}(\mu^2)$. This fact suggests that indeed the relaxation must be defined on pairs (m, c^{-1}) as in (6.6). Given Q , V_0 , m and c , the problem reduces to seeking a family of probability measures whose support is contained in Q , the volume integral must be V_0 and the moments of order 3 and -3 are m and c^{-1} , respectively. The simplest family satisfying all the restrictions is the following: there exists $\theta(x) \in [0, 1]$, $u(x) \in \mathcal{U}$, $x \in (a, b)$, and $\lambda \in [a, b]$ such that the family of probability measures

$$\nu_x = \begin{cases} \theta(x)\delta_{u_{max}} + (1 - \theta(x))\delta_{u(x)}, & x \in (a, \lambda), \\ \theta(x)\delta_{u_{min}} + (1 - \theta(x))\delta_{u(x)}, & x \in (\lambda, b) \end{cases}$$

represents a minimizer for \tilde{I} in $\tilde{\mathcal{U}}$ ([48]).

7. NONLOCAL PROBLEMS

Although optimal control problems, optimal design problems and shape optimization problems are also nonlocal, we devote this final section to some other problems where nonlocality is expressed in a different or special way. The main example is concerned with the functional

$$I(u) = \int_{\Omega \times \Omega} F(x_1, x_2, u(x_1), u(x_2), \nabla u(x_1), \nabla u(x_2)) dx_1 dx_2,$$

where $u : \Omega \subset \mathbf{R}^N \rightarrow \mathbf{R}^m$ and the density

$$F : \Omega \times \Omega \times \mathbf{R}^m \times \mathbf{R}^m \times \mathbf{M}^{m \times N} \times \mathbf{M}^{m \times N} \rightarrow \mathbf{R}$$

is continuous in the last four variables and measurable in $\Omega \times \Omega$ and satisfies the bounds

$$\begin{aligned} C(|A_1|^p + |A_2|^p + |\lambda_1|^p + |\lambda_2|^p) + c(x_1, x_2) &\leq F(x_1, x_2, \lambda_1, \lambda_2, A_1, A_2) \\ &\leq M(|A_1|^p + |A_2|^p + |\lambda_1|^p + |\lambda_2|^p) + m(x_1, x_2), \end{aligned}$$

$c, m \in L^1(\Omega \times \Omega)$, $0 < C < M$. This type of functional may be, and has been, used to account for long range interactions.

If $u_j \rightharpoonup u$ in $W^{1,p}(\Omega)$ and $\nu = \{\nu_x\}_{x \in \Omega}$ is the corresponding gradient Young measure for a suitable subsequence, the corresponding Young measure, Λ , for the sequence $\{(u_j(x_1), u_j(x_2), \nabla u_j(x_1), \nabla u_j(x_2))\}$ is

$$\Lambda_{(x_1, x_2)} = \delta_{u(x_1)} \otimes \delta_{u(x_2)} \otimes \nu_{x_1} \otimes \nu_{x_2}.$$

The fact that Λ is trivial in the components corresponding to $u_j(x_1)$ and $u_j(x_2)$ reflects the fact that $u_j \rightarrow u$ strongly in $L^p(\Omega)$ (see Proposition 4.3). Young measures associated to bounded sequences of gradients in $W^{1,p}(\Omega)$ have been characterized

in Theorem 3.1 by means of Jensen's inequality for quasiconvex functions. Let $\tilde{\mathcal{U}}$ stand for the set of such gradient Young measures.

We have the following theorem yielding the lower semicontinuous functionals with respect to weak convergence in $W^{1,p}(\Omega)$, and associated to it an existence theorem using the direct method. This last fact is left to the reader.

Theorem 7.1. *The functional I is weak lower semicontinuous in $W^{1,p}(\Omega)$ if and only if*

$$\begin{aligned} \int_{\Omega \times \Omega} \int_{\mathbf{M}^{m \times N} \times \mathbf{M}^{m \times N}} W(x_1, x_2, u(x_1), u(x_2), A_1, A_2) d\nu_{x_1}(A_1) d\nu_{x_2}(A_2) dx_1 dx_2 \\ \geq \int_{\Omega \times \Omega} W(x_1, x_2, u(x_1), u(x_2), \nabla u(x_1), \nabla u(x_2)) dx_1 dx_2, \end{aligned}$$

for all $\nu \in \tilde{\mathcal{U}}$ where

$$\nabla u(x) = \int_{\mathbf{M}^{m \times N}} A d\nu_x(A).$$

It is remarkable how the inequality in this theorem is so complicated to understand. Certainly, a sufficient condition for this weak lower semicontinuity property is the separate quasiconvexity on each of the two gradient variables. However this is far from being necessary. For instance in the scalar, homogeneous, one-dimensional situation where F is just a function of the two derivatives $u'(x_1)$ and $u'(x_2)$, the necessary and sufficient condition can be written in the form

$$\sum_{i,j=1}^{2n} F(\lambda_i, \lambda_j) \geq 4 \sum_{i,j=1}^n F\left(\frac{\lambda_{2i-1} + \lambda_{2i}}{2}, \frac{\lambda_{2j-1} + \lambda_{2j}}{2}\right),$$

for any $n \in \mathbf{N}$ and any choice $\lambda_1, \lambda_2, \dots, \lambda_{2n} \in \mathbf{R}$ ([54]). We have been unable to find a counterexample to the fact that the above inequality implies separate convexity.

Relaxation can now be easily described in terms of Young measures. We extend I on $\tilde{\mathcal{U}}$ by putting

$$\tilde{I}(\nu) = \int_{\Omega \times \Omega} \int_{\mathbf{M}^{m \times N} \times \mathbf{M}^{m \times N}} F(x_1, x_2, u(x_1), u(x_2), A_1, A_2) d\nu_{x_1}(A_1) d\nu_{x_2}(A_2) dx_1 dx_2,$$

where

$$\nabla u(x) = \int_{\mathbf{M}^{m \times N}} A d\nu_x(A).$$

8. EPILOGUE

The message we have tried to convey is that Young measures can be, among others, a convenient tool to treat some aspects of optimization problems. In particular, they may serve as a helpful device to understand the behavior of minimizing sequences from the perspective of optimization. Each particular problem requires a different analysis in which the specific features of the situation must be taken into account if explicit conclusions are sought. Every situation examined in this paper is incomplete in many respects and much remains to be done. As a partial

list of those directions where further improvement is needed, the reader is invited to reflect upon the following issues:

1. A deeper understanding of gradient Young measures. When is a given probability measure supported on matrices gradient? Explicit examples of quasiconvex and rank-one convex functions. How can one compute quasiconvex and rank-one convex hulls of functions, even from a numerical point of view?
2. Variational principles with linear growth, or with growth exponent p under the dimension N . Questions about functions of bounded variation, concentration effects, cavitation and discontinuities in nonlinear elasticity, etc., have not been addressed from the point of view of Young measures.
3. Models of micromagnetics and magnetostriction in dimension three. What is the relevance of the analysis presented here with respect to approximation and numerical issues?
4. Optimal design in dimension three. Construction or approximation of optimizing sequences. What is the relationship of quasiconvexity and constrained quasiconvexity?
5. The analysis of shape optimization problems through Young measure techniques is completely open.
6. Analysis of the optimal design of a genuinely two-dimensional plate under a given load.
7. The convexity for variational problems involving multiple integrals is wide open, even in the simplest case.
8. When a generalized solution is found or shown to exist, the task of computing it or deriving properties that it enjoys through generalized optimality conditions is also an important chapter not explored except in some specific situations. Not much is known in most of the cases discussed here.

ACKNOWLEDGEMENTS

This work has been supported by DGES (Spain) through grants PB96-0534 and HI95-18.

REFERENCES

- [1] Balder, E. J. 1984 A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM J. Control and Opt.*, 22, 570-598. MR **85k**:49018
- [2] Balder, E. J. 1995 Lectures on Young Measures, *Cahiers de Ceremade*, 9512.
- [3] Ball, J. M. 1977 Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rat. Mech. Anal.*, 63, 337-403. MR **57**:14788
- [4] Ball, J. M. 1989 A version of the fundamental theorem for Young measures, *PDE's and continuum models of phase transitions*, Lecture Notes in Physics, 344, (Rascle, M., Serre, D., and Slemrod, M., eds.) Springer, 207-215. MR **91c**:49021
- [5] Ball, J. M. and James, R. D. 1987 Fine phase mixtures as minimizers of energy, *Arch. Rat. Mech. Anal.*, 100, 13-52. MR **89c**:80005
- [6] Ball, J. M. and James, R. 1992 Proposed experimental tests of a theory of fine microstructure and the two well problem, *Phil. Trans. R. Soc. London A*, 338, 389-450.
- [7] Barbu, V. 1994 *Mathematical Methods in Optimization of Differential Systems*, Kluwer Academic Publishers. MR **96k**:49002
- [8] Bonnetier, E. and Conca, C. 1993 Relaxation totale d'un problème d'optimisation de plaques, *CRAS Paris*, 317, 931-936. MR **94i**:73068
- [9] Bonnetier, E. and Conca, C. 1994 Approximation of Young measures by functions and application to a problem of optimal design for plates with variable thickness, *Proc. Roy. Soc. Edin., A*, 124, 399-422. MR **95h**:73055

- [10] Bonnetier, E. and Vogelius, M. 1987 Relaxation of a compliance functional for a plate optimization problem, Applications of Multiple Scaling in Mechanics, (P.G. Ciarlet and E. Sanchez-Palencia, eds.), Masson, 31-53. MR **88m**:73030
- [11] Buttazzo, G. 1989 *Semicontinuity, Relaxation and Integral Representations in the Calculus of Variations*, Pitman Res. Notes Math. Ser., 207, Longman, Harlow. MR **91c**:49002
- [12] Buttazzo, G. and Dal Maso, G. 1982 Γ -convergence and optimal control problems, J. Optim. Theor. Appl., 38, 385-407. MR **85b**:49041
- [13] Buttazzo, G. and Dal Maso, G. 1990 Shape optimization for Dirichlet problems: relaxed solutions and optimality conditions, Bull. Amer. Math. Soc., 23, 531-535. MR **91c**:49005
- [14] Buttazzo, G. and Dal Maso, G. 1991 Shape optimization for Dirichlet problems: relaxed formulation and optimality conditions, Appl. Math. Optim., 23, 17-49. MR **92e**:49055
- [15] Buttazzo, G. and Dal Maso, G. 1993 An existence result for a class of shape optimization problems, Arch. Rat. Mech. Anal., 122, 183-195. MR **94i**:49052
- [16] Cabib, E. 1987 A relaxed control problem for two-phase conductors, Ann. Univ. Ferrara Sez. VII, 56, 39-65. MR **89k**:49003
- [17] Cabib, E. and Dal Maso, G. 1988 On a class of optimum problems in structural design, J. Opt. Theory Appl., 56, 39-65. MR **88k**:49007
- [18] Cesari, L. 1983 *Optimization, Theory and Applications*, Springer-Verlag. MR **85c**:49001
- [19] Clarke, F. 1975 Admissible relaxation in variational and control problems, J. Math. Anal. Appl., 51, 557-576. MR **53**:11448
- [20] Dacorogna, B. 1982 Quasiconvexity and relaxation of non convex problems in the calculus of variations, J. Funct. Anal., 46, 102-118. MR **83g**:49025
- [21] Dacorogna, B. 1989 *Direct methods in the Calculus of Variations*, Springer. MR **90e**:49001
- [22] De Simone, A. 1993 Energy minimizers for large ferromagnetic bodies, Arch. Rat. Mech. Anal., 125, 99-143. MR **94j**:82084
- [23] Evans, L. C., 1990 *Weak Convergence Methods for Nonlinear Partial Differential Equations*, CBMS 74, American Mathematical Society. MR **91a**:35009
- [24] Fonseca, I., Kinderlehrer, D., Pedregal, P. 1994 Energy functionals depending on elastic strain and chemical composition, Calc. Var., 2, 283-313. MR **97f**:73011
- [25] Fonseca, I., Müller, S. and Pedregal, P. 1998 Analysis of concentration and oscillation effects generated by gradients I, SIAM J. Math. Anal., 29, n 3, 736-756. CMP 98:11
- [26] Freddi, L. 1994 Rilassamento e convergenza di problemi di controllo ottimo, Preprint del Dipartimento di Matematica dell'Università di Pisa.
- [27] Gamkrelidze, R. V. 1978 *Principles of Optimal Control Theory*, Plenum Press, New York. MR **58**:33350c
- [28] Gerard, P. 1991 Microlocal defect measures, Comm. Part. Diff. Eq., 16 (11), 1761-1794. MR **92k**:35027
- [29] James, R. D. and Kinderlehrer, D. 1990 Frustration in ferromagnetic materials, Cont. Mech. Therm., 2, 215-239. MR **92a**:82132
- [30] James, R. D. and Kinderlehrer, D. 1993 A theory of magnetostriction with application to TbDyFe₂, Phil. Mag., B 68, 237-274.
- [31] James, R. D. and Kinderlehrer, D., Magnetoelastic interactions, *Zeitschrift für Angewandte Mathematik und Mechanik* **76** (Suppl. 2), 1996, pp. 401-404.
- [32] James, R. D., Kinderlehrer, D., and Ma, L., Modeling magnetostrictive microstructure under loading, in Mathematics of Microstructure Evolution (eds. L.-Q Chen, B. Fultz, J. W. Cahn, J. Manning, J. Morral and J. Simmonds), TMS/SIAM, to appear.
- [33] James, R. D. and Müller, S. 1994 Internal variables and fine-scale oscillations in micromagnetics, Cont. Mech. and Therm., 6, 291-336. MR **96a**:82045
- [34] James, R. D. and Wuttig, M. 1996 Magnetostriction of Martensite, Phil. Mag. A, in press.
- [35] James, R. D. and Wuttig, M., 1996 Alternative smart materials, Proceedings SPIE Symposium on "Mathematics and Control in Smart Structures" (ed. V. V. Varadan and J. Chandra), Vol. 2715, 420-426.
- [36] Kinderlehrer, D. and Pedregal, P. 1991 Characterizations of Young measures generated by gradients, Arch. Rat. Mech. Anal., 115, 329-365. MR **92k**:49089
- [37] Kinderlehrer, D. and Pedregal, P. 1992 Weak convergence of integrands and the Young measure representation, SIAM J. Math. Anal., 23, 1-19. MR **92m**:49076
- [38] Kinderlehrer, D. and Pedregal, P. 1994 Gradient Young measures generated by sequences in Sobolev spaces, J. of Geom. Anal., 4, 59-90. MR **95f**:49059

- [39] Kohn, R. V. and Strang, G. 1986 Optimal design and relaxation of variational problems, I, II and III, CPAM, 39, 113-137, 139-182 and 353-377. MR **87d**:49019a; MR **87d**:49019b; MR **87i**:49023
- [40] Kohn, R. V. and Vogelius, M. 1986 Thin plates with varying thickness, and their relation to structural optimization, Homogenization and Effective Moduli of Materials and Media, IMA Volumes 1, (Ericksen, J., Kinderlehrer, D., Kohn, R., Lions, J. L., eds), Springer-Verlag, 126-149. MR **87k**:73044
- [41] Mascolo, E. and Migliaccio, L. 1988 Relaxation in optimal control theory, Ann. Univ. Ferrara Sez. VII, 34, 247-263. MR **90i**:49038
- [42] Mascolo, E. and Migliaccio, L. 1989 Relaxation methods in control theory, App. Math. Opt., 20, 97-103. MR **90c**:49059
- [43] Müller, S. 1998 Variational models for microstructure and phase transitions, Lecture notes no. 2, MPI, Leipzig.
- [44] Morrey, Ch. B., 1952 Quasiconvexity and the lower semicontinuity of multiple integrals, Pacific J. Math., 2, 25-53. MR **14**:992a
- [45] Murat, F. 1977 H-convergence, *Séminaire d'analyse fonctionnelle et numérique de l'Université d'Alger*. CMP 98:07
- [46] Murat, F. and Tartar, L. 1985 Calcul des variations et homogénéisation, in Les méthodes de l'homogénéisation: théorie et applications en physique. Dir. des études et recherches de l'EDF, Eyrolles, Paris, 319-370. MR **87i**:73059
- [47] Murat, F. and Tartar, L. 1985 Optimality conditions and homogenization, Proceedings of Nonlinear Variational Problems, Isola d'Elba 1983, Res. Notes in Math., 127, Pitman, London, 1-8. MR **87b**:49011
- [49] Muñoz, J. and Pedregal, P. 1998 On the relaxation of an optimal design problem for plates, Asympt. Anal., 16, n 2, 125-140. CMP 98:09
- [48] Muñoz, J. and Pedregal, P. 1998 A refinement on existence of solutions to optimal control problems, submitted.
- [50] Pedregal, P. 1994 Relaxation in ferromagnetism: the rigid case, J. of Nonlinear Science, 4, 105-125. MR **95a**:82119
- [51] Pedregal, P. 1995 Numerical approximation of parametrized measures, Numer. Funct. Anal. and Opt., 16, (7 & 8), 1049-1066. MR **96h**:49023
- [52] Pedregal, P. 1996 On the numerical analysis of nonconvex variational problems, Numerische Mathematik, 74, 325-336. MR **97j**:49023
- [53] Pedregal, P. 1997 *Parametrized Measures and Variational Principles*, Birkhäuser, Basel. MR **98e**:49001
- [54] Pedregal, P. 1997 Nonlocal variational principles, Nonlin. Anal., 29, 1379-1392. MR **98m**:49033
- [55] Pedregal, P. 1998 Equilibrium conditions for Young measures, SIAM J. Cont. Opt., 36, n 3, 797-813. CMP 98:11
- [56] Pedregal, P. 1998 Relaxation in magnetostriction, submitted.
- [57] Pedregal, P. 1998 Optimal design and constrained quasiconvexity, preprint.
- [58] Rogers, R. C. 1992 A nonlocal model for the exchange energy in ferromagnetic materials, J. Integral Eq. Appl., 3, 85-127. MR **92i**:73027
- [59] Roubicek, T. 1997 *Relaxation in optimization theory and variational calculus*, W. De Gruyter, Berlin. MR **98e**:49002
- [60] Sverak, V. 1992 New examples of quasiconvex functions, Arch. Rat. Mech. Anal., 119, 293-300. MR **93h**:90072
- [61] Sverak, V. 1992 Rank-one convexity does not imply quasiconvexity, Proc. Roy. Soc. Edinb., 120A, 185-189. MR **93b**:49026
- [62] Tartar, L. 1977 Cours Peccot, Collège de France.
- [63] Tartar, L. 1979 Compensated compactness and applications to partial differential equations, Nonlinear analysis and mechanics: Heriot-Watt Symposium, vol. IV, Knops, R. (ed.), Pitman Res. Notes Math., 39, 136-212. MR **81m**:35014
- [64] Tartar, L. 1985 Estimations fines des coefficients homogénéisés, Ennio De Giorgi Colloquium, Ed. by P. Kree, Res. Notes in Math., 125, Pitman, London, 168-187. MR **89f**:35030
- [65] Tartar, L. 1990 H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations, Proc. Roy. Soc. Edinb., 115A, 193-230. MR **91h**:35042

- [66] Tartar, L. 1992 On mathematical tools for studying partial differential equations of continuum physics: H-measures and Young measures, in *Developments in Partial Differential Equations and Applications to Mathematical Physics*, (eds. Buttazzo, Galdi, Zanghirati), Plenum, New York. MR **94c**:35031
- [67] Valadier M. 1990 Young measures, *Methods of Nonconvex Analysis*, Lect. Notes in Math., 1446, Springer, 152-188. MR **91j**:28006
- [68] Warga, J. 1972 *Optimal Control of Differential and Functional Equations*, Academic Press, New York. MR **51**:8915
- [69] Young, L. C. 1937 Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, classe III*, 30, 212-234.
- [70] Young, L. C. 1942 Generalized surfaces in the calculus of variations, I and II, *Ann. Math.*, 43, 84-103 and 530-544. MR **3**:249a; MR **4**:49d
- [71] Young, L. C. 1969 *Lectures on Calculus of Variations and Optimal Control Theory*, W. B. Saunders. MR **41**:4337

ETSI INDUSTRIALES, UNIVERSIDAD DE CASTILLA-LA MANCHA, 13071 CIUDAD REAL, SPAIN
E-mail address: ppedrega@ind-cr.uclm.es