ESAIM: Control, Optimisation and Calculus of Variations August 2003, Vol. 9, 437–448
DOI: 10.1051/cocv:2003021

SOME REMARKS ON EXISTENCE RESULTS FOR OPTIMAL BOUNDARY CONTROL PROBLEMS

Pablo Pedregal[1]

Abstract. An optimal control problem when controls act on the boundary can also be understood as a variational principle under differential constraints and no restrictions on boundary and/or initial values. From this perspective, some existence theorems can be proved when cost functionals depend on the gradient of the state. We treat the case of elliptic and non-elliptic second order state laws only in the two-dimensional situation. Our results are based on deep facts about gradient Young measures.

Mathematics Subject Classification. 49J20, 49J45.

Received June 12, 2002. Revised October 17, 2002.

1. Introduction

Some papers have recently examined optimal design problems from a variational perspective [9,10] leading to some specific issues and problems in the Calculus of Variations. We would like to explore here the possibilities of this approach for optimal control problems, and in particular, for optimal boundary control problems where controls act on (a part of) the boundary of the domain. Specifically we would like to

$$\text{Minimize} \quad I(y) = \int_\Omega W(x, u(x), \nabla u(x))\, \mathrm{d}x$$

where the state u is determined from the control y through the differential law

$$\operatorname{div}\,(H(x, u(x), \nabla u(x))) = 0 \quad \text{in} \quad \Omega, \tag{1.1}$$

and appropriate boundary conditions (BC) on $\partial\Omega$

$$\mathrm{BC}[y] \text{ for } u \quad \text{on} \quad \partial\Omega, \tag{1.2}$$

where the control y is defined on (a part of) $\partial\Omega$. We have written BC[y] to stress that controls y enter into the boundary conditions, in an unspecified way, to determine u. For future reference, we will identify this problem as (P). The class of admissible controls consists of all y's that can be obtained, by restriction on $\partial\Omega$, from all (weak) solutions of (1.1) in some appropriate Sobolev space $W^{1,p}(\Omega)$. Our approach does not require to be more specific about the class of competing controls. In fact we will say that a pair (y, u) is admissible for

Keywords and phrases. Boundary controls, vector variational problems, gradient Young measures.

[1] ETSI Industriales, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain;
e-mail: `Pablo.Pedregal@uclm.es`

© EDP Sciences, SMAI 2003

(P) if $u \in W^{1,p}(\Omega)$ is the solution of (1.1) and (1.2). Notice that certain types of parabolic and hyperbolic equations are also included in the form of our problem, so that controls y can also incorporate initial conditions. Our analysis here is restricted to the two-dimensional situation where $\Omega \subset \mathbf{R}^2$ is a simply-connected domain, possibly unbounded. Our main results refer to some situations, under various sets of structural hypotheses, where existence of optimal solutions for (P) can be shown to exist.

It is interesting to realize that it is not inexcusable to fully specify how controls act on boundary and/or initial conditions. In fact, our perspective focuses on the problem

$$\text{Minimize} \quad J(u) = \int_\Omega W(x, u(x), \nabla u(x))\, \mathrm{d}x$$

subject to

$$\text{div}\,(H(x, u(x), \nabla u(x))) = 0 \quad \text{in} \quad \Omega.$$

Optimal solutions of this variational problem will determine, by restriction, optimal boundary controls regardless of the particular form in which controls act. From this point of view, we will analyze these variational problems under differential restrictions but no boundary condition is enforced. We will identify this variational problem as $(\tilde{P})$ and would take for granted all necessary ingredients on (P) so that both (P) and $(\tilde{P})$ are equivalent. Thus optimal solutions for (P) will be sought by looking for optimal solutions for $(\tilde{P})$. We will keep however the formulation of results in the form of optimal boundary control problems as in our model problem (P).

As pointed out above, Ω is assumed to be a regular, simply-connected, possibly unbounded domain in $\mathbf{R}^2$, and

$$H : \Omega \times \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2, \quad W : \Omega \times \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}$$

are Carathéodory maps for which some solutions for (1.1) exist.

Optimal control problems governed by partial differential equations is a fundamental field in applied mathematics with an astonishing number of applications in science and engineering. During the last decades, this kind of problems have received a lot of attention, mainly emphasizing the importance of optimality conditions in situations where existence of optimal solutions was not really an issue. For many of these examples, cost functionals would not depend explicitly on derivatives of states. The literature on optimal control of distributed parameter systems is overwhelming. We simply cite here a recent collection of papers where one can find many issues and directions of current research in this area [2].

Our main point in this note is to stress that in order to achieve the existence of optimal solutions, there is an important and interesting interplay between the structure of the underlying state differential law and the convexity properties of the integrand with respect to the derivative variable. The necessity of this interaction has been known, even in much more general circumstances, for many years, and can be roughly stated by saying that W should be convex along the characteristic cone defined by the state equation. In particular, the techniques of compensated compactness [4–6, 14, 15] focus on deriving profound weak continuity results under differential constraints (see also [7]). Here we would like to show the sufficiency of this "characteristic convexity" in some specific situations, that in turn translate into existence theorems for optimal boundary control problems.

Our two main theorems are the following. The first one deals with elliptic or monotone operators. The second one treats the case of linear, non-monotone operators. We also provide some insight on why the linearity must be, in a sense, an essential part of the second theorem (Sect. 4).

Theorem 1.1. *Let $\Omega \subset \mathbf{R}^2$ be a regular, bounded, simply-connected domain and*

$$W : \Omega \times \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R},$$
$$H : \Omega \times \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2,$$

be Carathéodory mappings such that

$$c\left(|A|^p + |u|^p - 1\right) \le W(x,u,A),$$
$$|H(x,u,A)| \le c\left(|A|^q + 1\right),$$

for every pair (x,u)*,* $p > q \ge 1$ *and* $c > 0$*. Suppose further that*

$$(A - B)\cdot(H(x,u,A) - H(x,u,B)) = 0 \quad \textit{implies} \quad A = B,$$

for every pair (x,u)*. Then problem* (P) *admits optimal solutions, i.e. there are optimal pairs* (y,u) *where* $u \in W^{1,p}(\Omega)$*.*

Notice that there is no convexity whatsoever assumed on W.

Suppose now that

$$H(x,u,A) = f(x,u) + G(x,u)A$$

and

$$f : \Omega \times \mathbf{R} \to \mathbf{R}^2, \quad G : \Omega \times \mathbf{R} \to \mathbf{M}^{2\times 2},$$

are Carathéodory mappings. Suppose that

$$\|G(x,u)\| \le M < +\infty,$$
$$|f(x,u)| \le c\left(|u|^p + f_0(x)\right),$$

for positive constants M and c, and $f_0 \in L^1(\Omega)$. Moreover

$$W : \Omega \times \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}$$

is also a Carathéodory function for which the lower bound

$$c\left(|A|^p + |u|^p - 1\right) \le W(x,u,A),$$

for all pairs (x,u) is valid where $p > 1$, $c > 0$.

Theorem 1.2. *Under the previous hypotheses on* H *and* W*, if for every fixed* (x,u)*,* $W(x,u,\cdot)$ *is convex along the directions given by the vectors* n *such that*

$$n \cdot G(x,u)n = 0,$$

then the associated problem (P) *admits optimal pairs* (y,u) *where* $u \in W^{1,p}(\Omega)$*.*

Clearly, the interesting situations occur when such set of vectors n is not negligible. Notice that the situation where the state differential law changes type is included.

Section 2 is dedicated to several preliminary facts about gradient Young measures [8] as well as to how one can understand the original optimal control problem as a vector variational situation. Sections 3 and 4 treat the monotone and the non-monotone cases, respectively. The fundamental results that will enable us to prove the theorems above are contained in [3] and [11].

2. Some preliminaries

The dependence of W and H on the variables (x,u) is somehow irrelevant if they both depend explicitly on the gradient variable. This remark is clear when examining variational principles by means of gradient Young measures (see [8]). For this reason, and for the sake of simplicity, we will drop the dependence of H and W on the variables (x, u) in our initial discussion. Thus we take

$$H : \mathbf{R}^2 \to \mathbf{R}^2, \quad W : \mathbf{R}^2 \to \mathbf{R}$$

to be continuous maps.

By introducing a new independent field v, we can transform our problem into a vector variational situation as follows. Write

$$H(\nabla u(x)) + T\nabla v(x) = 0 \quad \text{in} \quad \Omega \tag{2.1}$$

where T is the $\pi/2$, counterclockwise rotation in the plane, and collect u and v into a single vector field $U = (u, v)$. This U will be the variable of our new, equivalent variational problem. The equality

$$H(\nabla U^{(1)}(x)) + T\nabla U^{(2)}(x) = 0$$

must be enforced always in Ω. Define a new density

$$\varphi : \mathbf{M}^{2\times 2} \to \mathbf{R}^* = \mathbf{R} \cup \{+\infty\}$$

by setting

$$\varphi(A) = \begin{cases} W(A^{(1)}), & A \in \Gamma, \\ +\infty, & \text{else}, \end{cases} \tag{2.2}$$

where Γ is the manifold associated with equation (2.1)

$$\Gamma = \left\{ A \in \mathbf{M}^{2\times 2} : A^{(2)} = TH\left(A^{(1)}\right)\right\}\cdot$$

$A^{(i)}$ is the i-th row of A. It is elementary to check that the initial problem $(\tilde{P})$ is equivalent to the vector variational problem

$$\text{Minimize} \quad J(U) = \int_\Omega \varphi(\nabla U(x))\,\mathrm{d}x \tag{2.3}$$

subject to no further restrictions.

As we know, and overlooking for the moment the fact that φ is not a Carathéodory density, the main issue concerning existence of optimal solutions for vector variational problems is the quasiconvexity of the integrand φ [1]; or rather, the weak lower semicontinuity of the functional as we are lacking the Carathéodory requirement as well as the typical upper bound on φ. This weak lower semicontinuity depends on the properties of gradient Young measures [8] whose support is contained in Γ. The important point here is the structure of Γ itself which in turn depends on H, as Γ is essentially the "graph" of H, and how this structure translates into properties of gradient Young measures supported on it. This is a profound issue whose significance has been stressed in several occasions [12, 13].

We would like to recall here several basic facts about gradient Young measures which are important when one wants to understand variational principles from this perspective (see [8]). The first one refers to the existence issue: for every bounded sequence in $L^p(\Omega)$, $p > 0$, $\{w_j\}$, there exists a family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ such that for every Carathéodory function

$$\psi : \Omega \times \mathbf{R}^d \to \mathbf{R}$$

for which $\{\psi(x, w_j(x))\}$ is weakly convergent in $L^1(\Omega)$, we have

$$\psi(x, w_j(x)) \rightharpoonup \overline{\psi}(x) = \int_{\mathbf{R}^d} \psi(x, \lambda)\,\mathrm{d}\nu_x(\lambda).$$

In particular, we can apply this representation of weak limits when $\{w_j\}$ is a bounded sequence of gradients. In this case we refer to ν as a gradient Young measure.

The second fact is related to the most basic, non-trivial example of a gradient Young measure supported in matrices. These are called first-order laminates and correspond to the probability measure

$$\nu_x = \nu_0 = t\delta_A + (1-t)\delta_B, \quad t \in (0,1), \text{ a.e. } x \in \Omega,$$

where A and B are matrices such that the difference $A - B$ must be a rank-one matrix. As a matter of fact, there is a whole class of gradient Young measures, identified in general as laminates, which are recursively built by "composing" first-order laminates. This class of gradient Young measures are characterized by Jensen's inequality for all rank-one convex functions.

The third one relates to the behavior of Young measure with respect to concentration effects and can be stated as follows. If $\nu = \{\nu_x\}_{x\in\Omega}$ is the Young measure associated with $\{w_j\}$ then

$$\liminf_{j\to\infty} \int_\Omega \psi(x, w_j(x))\,\mathrm{d}x \geq \int_\Omega \int_{\mathbf{R}^d} \psi(x,\lambda)\,\mathrm{d}\nu_x(\lambda)\,\mathrm{d}x,$$

for every Carathéodory function ψ bounded from below. If $\{\psi(x, w_j(x))\}$ converges weakly in $L^1(\Omega)$ then we actually have equality. This is indeed the defining property of the Young measure. But if the sequence $\{\psi(x, w_j(x))\}$ develops concentrations then we still have the above inequality which is the right one for weak lower semicontinuity.

The following is a localization result stating that a gradient Young measure $\nu = \{\nu_x\}_{x\in\Omega}$ is made up of homogeneous gradient Young measures at a.e. point $x \in \Omega$. If $\nu = \{\nu_x\}_{x\in\Omega}$ is a Young measure generated by a bounded sequence of gradients in $W^{1,p}(\Omega)$ then for a.e. $a \in \Omega$ and for any domain Q, there exists a bounded sequence in $W^{1,p}(Q)$, $\{v_{a,j}\}$, such that the Young measure associated to $\{\nabla v_{a,j}\}$ is ν_a, homogeneous. Moreover each function $v_{a,j}$ can be chosen in such a way that $v_{a,j} - u_{F(a)} \in W_0^{1,p}(Q)$ where $u_F(x)$ is the linear function Fx, and

$$F(a) = \int_{\mathbf{M}} A\,\mathrm{d}\nu_a(A).$$

Finally, we would like to indicate that strong convergence in Sobolev spaces is equivalent to triviality (a delta measure on the gradient of the limit) of the underlying gradient Young measure. This triviality is shown in the components where we have strong convergence. In particular, this is an interesting remark when dealing with sequences of functions of Sobolev spaces jointly with their gradients.

3. Monotone operators

The proofs of our main results in the Introduction rely on appropriate weak lower semicontinuity results. As indicated in the previous section, we treat first the case where W and H do not depend explicitly on (x, u) and

$$c\left(|A|^p - 1\right) \leq W(A),$$
$$|H(A)| \leq c\left(|A|^q + 1\right),$$

for $c > 0$ and $p > q \geq 1$. Take $1 < r < p/q$.

Theorem 3.1. *Assume H is such that*

$$(A - B)\cdot(H(A) - H(B)) = 0 \quad \textit{implies} \quad A = B, \tag{3.1}$$

then regardless of the convexity properties of W, J in (2.3) is weak lower semicontinuous in $W^{1,r}(\Omega)$.

Notice that hypothesis (3.1) amounts to the monotonicity of the corresponding differential operator.

The proof of this result makes use of the following fundamental fact [11].

Theorem 3.2. *Let*

$$\nabla U_j : \Omega \to \mathbf{M}^{2\times 2}$$

be a uniformly bounded sequence of gradients in $W^{1,p}(\Omega)$*,* $p > 1$*. Let* K *be a closed, connected subset of* $\mathbf{M}^{2\times 2}$ *such that* $\det(X - Y) \neq 0$ *for any two distinct* $X, Y \in K$ *and suppose*

$$\operatorname{dist}(\nabla U_j(x), K) \to 0, \quad j \to \infty,$$

for a.e. $x \in \Omega$*. Then the sequence* $\{\nabla U_j\}$ *is compact in* $L^q(\Omega)$ *for every* $1 \leq q < \infty$*.*

For the proof of Theorem 3.1, take $K = \Gamma$ and consider

$$\nabla U_j = \begin{pmatrix} \nabla u_j \\ \nabla v_j \end{pmatrix},$$
$$\nabla v_j(x) = TH(\nabla u_j(x)), \quad \text{a.e. } x \in \Omega.$$

The fact that Ω is bounded, $r < p/q \leq p$, the bounds assumed on W and H, and this last equation, enable us to ensure that $\{U_j\}$ is uniformly bounded in $W^{1,r}(\Omega)$ (in fact $\{\nabla U_j\}$ is equiintegrable in $L^r(\Omega)$), and hence a (sub)sequence converges weakly in this Sobolev space to some U. Obviously, if ∇U_j does not satisfy the equation (or for a full subsequence of it) then the weak lower semicontinuity is trivial since $J(U_j) = +\infty$. We would like to show that, regardless of the convexity of W, we have

$$J(U) \leq \liminf_{j\to\infty} J(U_j),$$

the weak lower semicontinuity property for J. Because of the form of ∇U_j, we have

$$\operatorname{dist}(\nabla U_j(x), \Gamma) = 0,$$

for all j and a.e. $x \in \Omega$. Furthermore, it is elementary to check that the difference of two distinct matrices in Γ can never be a rank-one matrix since

$$\det\left(\begin{pmatrix} A^{(1)} \\ TH\left(A^{(1)}\right) \end{pmatrix} - \begin{pmatrix} B^{(1)} \\ TH\left(B^{(1)}\right) \end{pmatrix}\right) = \left(A^{(1)} - B^{(1)}\right) \cdot \left(H\left(A^{(1)}\right) - H\left(B^{(1)}\right)\right),$$

and by our hypothesis this can never be zero if $A^{(1)}$ is different from $B^{(1)}$. From Theorem 3.2 we conclude that $\{\nabla U_j\}$ converges strongly in any Lebesgue space and this, in turn, implies, by Fatou's lemma, that

$$J(U) \leq \lim_{j\to\infty} J(U_j)$$

possibly for a suitable subsequence. Notice that this compactness property amounts to saying that the underlying gradient Young measure is trivial.

The proof of Theorem 1.1 is now standard, and follows along the lines of the direct method of the Calculus of Variations [1,8]. We will show that the variational problem $(\tilde{P})$ admits optimal solutions.

Proof of Theorem 1.1. We again use the notation

$$J(U) = \int_\Omega \varphi(x, U(x), \nabla U(x))\,\mathrm{d}x$$

for

$$\varphi(x, U, A) = \begin{cases} W(x, U^{(1)}, A^{(1)}), & \text{if } A \in \Gamma_{x,U^{(1)}}, \\ +\infty, & \text{else,} \end{cases}$$

where this time

$$\Gamma_{x,u} = \left\{ A \in \mathbf{M}^{2\times 2} : A^{(2)} = TH\left(x, u, A^{(1)}\right) \right\}.$$

Due to the equivalence of $(\tilde{P})$ with this variational reformulation, we will show that the functional J admits (global) minimizers by examining coercivity and weak lower semicontinuity. Let $\{U_j\}$ be a minimizing sequence for J. By the coercivity of W, this sequence is bounded in $W^{1,p}(\Omega)$. As before, the equation

$$H(x, U_j^{(1)}(x), \nabla U_j^{(1)}(x)) + T\nabla U_j^{(2)}(x) = 0,$$

together with the boundedness of Ω and the bounds assumed on H, let us conclude that the full vector gradients $\{\nabla U_j\}$ is equiintegrable in $L^r(\Omega)$. Let $\nu = \{\nu_x\}_{x\in\Omega}$ be the associated gradient Young measure. It is clear that for a.e. $x \in \Omega$ the support of ν_x is contained in the manifold $\Gamma_{x,u(x)}$ where u is the weak limit of u_j (weak in $W^{1,p}(\Omega)$ and strong in $L^p(\Omega)$). For this, it suffices to consider the Young measure, $\mu = \{\mu_x\}_{x\in\Omega}$, associated with $\left\{(U_j^{(1)}, \nabla U_j)\right\}$. By the strong convergence of $U_j^{(1)}$ (recall the comments at the end of Sect. 2), we can write

$$\mu_x = \delta_{U^{(1)}(x)} \otimes \nu_x,$$

and by using the sequence of functions

$$\left| H(x, U_j^{(1)}(x), \nabla U_j^{(1)}(x)) + T\nabla U_j^{(2)}(x) \right|^2$$

which vanishes identically, it is elementary to conclude the property on the support of ν_x.

On the other hand, each individual member ν_x for fixed $x \in \Omega$ is a (homogeneous) gradient Young measure generated by a bounded sequence in $W^{1,r}(\Omega)$ (see again the comments in Sect. 2). By our monotonicity hypothesis, and applying Theorem 3.2 to this sequence of gradients for each fixed $x \in \Omega$, we conclude that ν_x is trivial

$$\nu_x = \delta_{\nabla U(x)} \quad \text{a.e. } x \in \Omega,$$

where U is the weak limit in $W^{1,r}(\Omega)$ of the pairs of gradients. This implies that the convergence to U is in fact strong in $W^{1,r}(\Omega)$ (recall that $\{\nabla U_j\}$ is equiintegrable in $L^r(\Omega)$), and in particular it is strong in $W^{1,p}(\Omega)$ for the first component $\left\{\nabla U_j^{(1)}\right\}$. This leads us to conclude that, again as in Theorem 3.1,

$$J(U) \leq \lim_{j\to\infty} J(U_j)$$

and this proves the theorem since

$$\nabla U(x) \in \Gamma_{x,U^{(1)}(x)}$$

for a.e. $x \in \Omega$ and so $U^{(1)}$ is a minimizer for $(\tilde{P})$ in $W^{1,p}(\Omega)$. □

Examples where Theorem 1.1 can be applied include all typical linear and non-linear elliptic operators.

4. The non-monotone case

When the monotonicity hypothesis does not hold, then weak lower semicontinuity of J amounts to "linearity" (or flatness) of H and "convexity" of W in the following sense. We drop the dependence on (x, u).

Proposition 4.1. *Assume that the functional J in (2.3) is weak lower semicontinuous in $W^{1,\infty}(\Omega)$. If A, B are different vectors in $\mathbf{R}^2$ such that*

$$(A - B) \cdot (H(A) - H(B)) = 0$$

then

$$H(tA + (1-t)B) = tH(A) + (1-t)H(B),$$
$$W(tA + (1-t)B) \leq tW(A) + (1-t)W(B),$$

for all $t \in [0,1]$.

Proof. By the same computations made earlier, the probability measure

$$\mu = t\delta_{\overline{A}} + (1-t)\delta_{\overline{B}},$$

where

$$\overline{A} = \begin{pmatrix} A \\ TH(A), \end{pmatrix} \quad \overline{B} = \begin{pmatrix} B \\ TH(B) \end{pmatrix}$$

belong to Γ, is a laminate supported in Γ according to our discussion in Section 2. Therefore we must have, if the functional J in (2.3) is weak lower semicontinuous, that

$$\varphi\left(\langle \mathrm{id}, \mu\rangle\right) \leq \langle \varphi, \mu\rangle,$$

where φ is the integrand for J defined in (2.2). This is nothing but Jensen's inequality for the homogeneous gradient Young measure μ. Hence we obtain

$$\varphi\begin{pmatrix} tA + (1-t)B \\ T\left(tH(A) + (1-t)H(B)\right) \end{pmatrix} \leq tW(A) + (1-t)W(B).$$

Since the right-hand side is finite, we conclude that the matrix in the left-hand side ought to belong to Γ because this is the set where φ is finite, and thus the linearity claimed on H in the statement of the theorem holds. In addition

$$\varphi\begin{pmatrix} tA + (1-t)B \\ TH\left(tA + (1-t)B\right) \end{pmatrix} = W(tA + (1-t)B),$$

and this yields the convexity of W. □

Because of the previous fact, we will restrict attention henceforth to the linear, non-elliptic case, so that H will be identified with a non-positive definite 2×2-matrix. We do not mean however to cover, without loss of generality, all cases. The previous proposition is simply a justification on why we restrict attention to this simplified case. The differential restriction reads

$$\mathrm{div}\,(H\nabla u) = 0 \quad \text{in} \quad \Omega,$$

or equivalently

$$H\nabla u + T\nabla v = 0.$$

It is elementary to check that rank-one directions contained in the manifold Γ are associated with vectors n such that

$$n \cdot Hn = 0.$$

Moreover, the number of directions n such that $n \cdot Hn = 0$ is at most two and equal to the rank of H (under the further hypotheses assumed on H). It is indeed well-known that if there are three vectors with this property, then H is a multiple of a rotation of angle $\pi/2$ and in this case the differential law is identically satisfied for any function u.

Theorem 4.2. *Let H be a matrix as described above. The functional J in (2.3) is weak lower semicontinuous in $W^{1,p}(\Omega)$, $p > 1$, if and only if the density W is (separately) convex along the directions given by the vectors n such that $n \cdot Hn = 0$.*

Proof. The necessity of the convexity on W is very well-known and in fact has already been indicated in Proposition 4.1. Let us examine the sufficiency of this convexity.

Under our linearity assumption on H, we know that Γ is a two-dimensional, linear subspace containing at most two independent rank-one directions. Let these matrices be $a_i \otimes n_i$, $i = 1, 2$, where we can take $|a_i| = |n_i| = 1$. Three situations may occur:

1. the two sets of vectors $\{a_1, a_2\}$ and $\{n_1, n_2\}$ are linearly independent;
2. the vectors a_i are dependent;
3. the vectors n_i are dependent.

The last situation is easy to deal with since in this case all matrices in Γ are of the form $a \otimes n$ for $a \in \mathbf{R}^2$. Any probability measure supported in Γ will be a laminate and hence a gradient Young measure. The integrand defined in (2.2) will in fact be convex provided W is convex along the direction given by n.

The second case cannot happen since Γ would be the set of all vectors $a \otimes n$ for fixed a and $n \in \mathbf{R}^2$. This would imply that $n \cdot Hn = 0$ for all n and this would take us back to H being a multiple of T.

For the case where the two vectors n_1 and n_2 are independent, as well as $\{a_1, a_2\}$, we can define a non-singular, linear transformation $\tilde{H}$ by putting $\tilde{H}(a_i) = n_i$, $i = 1, 2$. In this way $\tilde{H}(\Gamma)$ is the subspace of diagonal matrices with respect to the basis $n_i \otimes n_j$, $i, j = 1, 2$. A suitable change of independent variables would take us to the standard subspace of diagonal matrices. Let $\{U_j\}$ be a bounded sequence in $W^{1,p}(\Omega)$ converging weakly to U. If

$$\liminf_{j\to\infty} J(U_j) = +\infty$$

there is nothing to prove. Let us assume that the previous liminf is finite. Because of the definition of the density φ for J, we must have that $\nabla U_j(x)$ belongs to Γ a.e. $x \in \Omega$. Consider further the sequence of functions

$$\tilde{U}_j = \tilde{H} U_j.$$

This is also a bounded sequence in $W^{1,p}(\Omega)$. Let $\nu = \{\nu_x\}_{x\in\Omega}$ be the gradient Young measure associated with a suitable subsequence which we do not bother to relabel. Because of our choice of $\tilde{H}$, the support of ν_x is contained in the subspace of diagonal matrices for a.e. $x \in \Omega$. But we know that all such gradient Young measures are laminates [3], so that, since $\tilde{W}(A) = W(\tilde{H}^{-1}A)$ is separately convex,

$$\tilde{\varphi}(A) = \begin{cases} \tilde{W}(A^{(1)}), & A \in \Gamma, \\ +\infty, & \text{else}, \end{cases}$$

is rank-one convex. By our remarks in Section 2 about Jensen's inequality for laminates and rank-one convex functions, we have

$$\tilde{\varphi}\left(\langle \mathrm{id}, \nu_x \rangle\right) \le \langle \tilde{\varphi}, \nu_x \rangle,$$

for a.e. $x \in \Omega$. On the other hand $\langle \mathrm{id}, \nu_x \rangle$ is the weak limit of $\left\{\tilde{U}_j\right\}$ which is $\tilde{H}U$, and integrating the above inequality over Ω, we obtain

$$\int_\Omega \varphi(\nabla U(x))\,\mathrm{d}x \le \int_\Omega \int_{\mathbf{M}^{2\times 2}} \varphi(\tilde{H}^{-1}A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x.$$

By the third fact about Young measures mentioned at the end of Section 2, and bearing in mind the definition of $\tilde{W}$,

$$\int_\Omega \int_{\mathbf{M}^{2\times 2}} \varphi(\tilde{H}^{-1}A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x \le \lim_{j\to\infty} \int_\Omega \varphi(\nabla U_j(x))\,\mathrm{d}x.$$

Putting together these two inequalities we arrive at

$$\int_\Omega \varphi(\nabla U(x))\,\mathrm{d}x \le \lim_{j\to\infty}\int_\Omega \varphi(\nabla U_j(x))\,\mathrm{d}x,$$

the weak lower semicontinuity result claimed in the statement. □

Based on this weak lower semicontinuity result, the proof of Theorem 1.2 is straightforward along the lines of the direct method of the Calculus of Variations. As in the proof of Theorem 1.1, what matters is the dependence on $(\nabla u, \nabla v)$. In general, the differential constraint reads

$$\operatorname{div}\,(f(x,u(x)) + G(x,u(x))\nabla u) = 0 \quad \text{in} \quad \Omega,$$

or equivalently

$$f(x,u(x)) + G(x,u(x))\nabla u + T\nabla v = 0.$$

Proof of Theorem 1.2. As in the case of Theorem 1.1, we let

$$J(U) = \int_\Omega \varphi(x, U(x), \nabla U(x))\,\mathrm{d}x$$

where

$$\varphi(x,U,A) = \begin{cases} W(x,U^{(1)},A^{(1)}), & \text{if } A\in\Gamma_{x,U^{(1)}},\\ +\infty, & \text{else,}\end{cases}$$

and

$$\Gamma_{x,u} = \left\{A\in\mathbf{M}^{2\times 2} : f(x,u) + G(x,u)A^{(1)} + TA^{(2)} = 0\right\}\cdot$$

Due to the equivalence of the initial problem with this variational reformulation, we will show that the functional J admits (global) minimizers by examining coercivity and weak lower semicontinuity. Let $\{u_j\}$ be minimizing for J. By the coercivity assumed on W, this sequence is uniformly bounded in $W^{1,p}(\Omega)$. The pointwise constraint

$$f(x,u_j(x)) + G(x,u_j(x))\nabla u_j(x) + T\nabla v_j(x) = 0$$

together with the bounds we have on f and G and the uniform boundedness of ∇u_j in $L^p(\Omega)$, allow us to conclude that (after substracting appropriate constants if necessary) $\{v_j\}$ is also uniformly bounded in $W^{1,p}(\Omega)$. Let $\nu = \{\nu_x\}_{x\in\Omega}$ be the gradient Young measure corresponding to the sequence of pairs of gradients $\{(\nabla u_j, \nabla v_j)\}$. As before, it is clear that the support of ν_x is contained in $\Gamma_{x,u(x)}$ where u is the weak limit of u_j (strong in $L^p(\Omega)$). By our remarks in the proof of Theorem 4.2, we have that each ν_x is a laminate and $\varphi(x,U,A)$ is rank-one convex with respect to A (because $W(x,u,A)$ is convex along rank-one directions contained in $\Gamma_{x,u}$). We finish the proof as in Theorem 4.2. If we put $U_j = (u_j, v_j)$, $U = (u,v)$ its weak limit in $W^{1,p}(\Omega)$, then

$$\lim_{j\to\infty}\int_\Omega \varphi(x,U_j(x),\nabla U_j(x))\,\mathrm{d}x \ge \int_\Omega\int_{\mathbf{M}^{2\times 2}} \varphi(x,U(x),A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x \ge \int_\Omega \varphi(x,U(x),\nabla U(x))\,\mathrm{d}x.$$

The last inequality is Jensen's inequality for laminates and rank-one convex functions. We see that $u = U^{(1)}$ is an optimal solution for our initial problem. □

Typical corollaries of this result for the linear heat and wave equations in dimension one are included here as an illustration without entering into technicalities about the appropriate functional spaces for these evolution equations.

Consider the one-dimensional, linear heat equation

$$u_t(t,x) - u_{xx}(t,x) = v(t,x), \quad 0<x<L, t>0.$$

We would like to

$$\text{Minimize} \quad J(u) = \int_0^\infty \int_0^L W(t,x,u(t,x),u_t(t,x),u_x(t,x))\,\mathrm{d}x\,\mathrm{d}t$$

among all solutions of the preceding heat equation. In this situation $\Omega = (0,\infty) \times (0,L)$.

Corollary 4.3. *If $W(t,x,u,u_t,u_x)$ is a Carathéodory function, convex with respect to u_t when all other variables are fixed, and such that*

$$c\,(|(u_t,u_x)|^p - 1) \le W(t,x,u,u_t,u_x)$$

for some $p > 1$ and $c > 0$, then there are optimal solutions (in $W^{1,p}(\Omega)$) for the corresponding optimal boundary control problem.

Observe that this situation fits in the framework of Theorem 1.2 for the choice

$$f(t,x,u) = (u, -V(t,x)), \quad G(t,x,u) = \begin{pmatrix} 0 & 0 \\ 0 & -1, \end{pmatrix}$$

where $v = \dfrac{\partial V}{\partial x}$ and $V \in W^{1,1}(\Omega)$.

For the wave equation

$$u_{tt}(t,x) - u_{xx}(t,x) = v(t,x), \quad 0 < x < L, 0 < t,$$

if we want to

$$\text{Minimize} \quad J(u) = \int_0^\infty \int_0^L W(t,x,u(t,x),u_t(t,x),u_x(t,x))\,\mathrm{d}x\,\mathrm{d}t,$$

among all solutions of the wave equation, we can rely on the following result.

Corollary 4.4. *If $W(t,x,u,u_t,u_x)$ is a Carathéodory function, separately convex with respect to u_t and u_x when all other variables are fixed, and such that*

$$c\,(|(u_t,u_x)|^p - 1) \le W(t,x,u,u_t,u_x)$$

for some $p > 1$ and $c > 0$, then there are optimal solutions (in $W^{1,p}(\Omega)$) for the corresponding optimal boundary control problem.

This time we choose

$$f(t,x,u) = (0, -V(t,x)), \quad G(t,x,u) = \begin{pmatrix} 1 & 0 \\ 0 & -1, \end{pmatrix},$$

where $v = \dfrac{\partial V}{\partial x}$ and $V \in W^{1,1}(\Omega)$.

It is important to remark that this separate convexity is much weaker than the joint convexity with respect to both variables.

A finer analysis would be required to clarify existence of optimal solutions on typical spaces of solutions for the heat and wave equations as indicated before.

This work is supported by research projects BFM2001-0738 of the MCyT and GC-02-001 of Castilla-La Mancha (Spain). The criticism of two anonymous referees is greatly appreciated.

References

[1] B. Dacorogna, *Direct methods in the Calculus of Variations*. Springer (1989).
[2] K.H. Hoffmann, G. Leugerin and F. Tröltzsch, *Optimal Control of Partial Differential Equations*. Birkhäuser, Basel, *ISNM* **133** (2000).
[3] S. Müller, Rank-one convexity implies quasiconvexity on diagonal matrices. *Intern. Math. Res. Notices* **20** (1999) 1087-1095.
[4] F. Murat, Compacité par compensation. *Ann. Scuola Norm. Sup. Pisa Sci. Fis. Mat. (IV)* **5** (1978) 489-507.

[5] F. Murat, Compacité par compensation II, in *Recent Methods in Nonlinear Analysis Proceedings*, edited by E. De Giorgi, E. Magenes and U. Mosco. Pitagora, Bologna (1979) 245-256.

[6] F. Murat, A survey on compensated compactness, in *Contributions to the modern calculus of variations*, edited by L. Cesari. Pitman (1987) 145-183.

[7] P. Pedregal, Weak continuity and weak lower semicontinuity for some compensation operators. *Proc. Roy. Soc. Edinburgh Sect. A* **113** (1989) 267-279.

[8] P. Pedregal, *Parametrized Measures and Variational Principles*. Birkhäuser, Basel (1997).

[9] P. Pedregal, Optimal design and constrained quasiconvexity. *SIAM J. Math. Anal.* **32** (2000) 854-869.

[10] P. Pedregal, Fully explicit quasiconvexification of the mean-square deviation of the gradient of the state in optimal design. *ERA-AMS* **7** (2001) 72-78.

[11] V. Sverak, On Tartar's conjecture. *Inst. H. Poincaré Anal. Non Linéaire* **10** (1993) 405-412.

[12] V. Sverak, *On regularity for the Monge-Ampère equation*. Preprint (1993).

[13] V. Sverak, *Lower semicontinuity of variational integrals and compensated compactness*, edited by S.D. Chatterji. Birkhäuser, *Proc. ICM* **2** (1994) 1153-1158.

[14] L. Tartar, Compensated compactness and applications to partal differential equations, in *Nonlinear analysis and mechanics: Heriot–Watt Symposium*, Vol. IV, edited by R. Knops. *Pitman Res. Notes Math.* **39** (1979) 136-212.

[15] L. Tartar, The compensated compactness method applied to systems of conservation laws, in *Systems of Nonlinear Partial Differential Eq.*, edited by J.M. Ball. Riedel (1983).