JOURNAL OF FUNCTIONAL ANALYSIS **46**, 102–118 (1982)

Quasiconvexity and Relaxation of Nonconvex Problems in the Calculus of Variations*

B. DACOROGNA

Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912 and Département de Mathématiques, Ecole Polytechnique Fédérale, av. de Cour 61, 1006 Lausanne, Switzerland

Communicated by the Editors

Received March 31, 1981

In this article we study the existence of solutions for (P) $\inf\{\int_\Omega f(\nabla u(x))\,dx;\ u\in u_0+W_0^{1,p}(\Omega;\mathbb{R}^m)\}$, where f satisfies $a+b|F|^p\leqslant f(F)\leqslant c+d|F|^p$ for some $a,c\in\mathbb{R}$, $d\geqslant b>0$ and $p>1$. Let (QP) $\inf\{\int_\Omega Qf(\nabla u(x))\,dx;\ u\in u_0+W_0^{1,p}(\Omega;\mathbb{R}^m)\}$, where $Qf=\sup\{\Phi\colon \Phi\leqslant f,\ \Phi$ quasiconvex$\}$. We show that

Theorem. inf(P) = inf(QP). *More precisely, for every* $\bar u\in u_0+W_0^{1,p}(\Omega;\mathbb{R}^m)$, *there exists* $\{u^s\}_{s=1}^\infty$ *such that*

$$u^s\in u_0+W_0^{1,p}(\Omega;\mathbb{R}^m),$$
$$u^s\rightharpoonup\bar u \text{ weakly in } W^{1,p}(\Omega;\mathbb{R}^m),$$
$$\int_\Omega f(\nabla u^s(x))\,dx\to\int_\Omega Qf(\nabla\bar u(x))\,dx.$$

INTRODUCTION

In this article we are concerned with existence of solutions for

$$\text{(P)}\qquad \inf\left\{\int_\Omega f(\nabla u(x))\,dx;\ u\in u_0+W_0^{1,p}(\Omega;\mathbb{R}^m)\right\},$$

where

(i) Ω is a bounded open set of $\mathbb{R}^n$; $u\colon\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ and hence $\nabla u\in\mathbb{R}^{nm}$,

(ii) $W^{1,p}(\Omega;\mathbb{R}^m)=\{u=(u_1,\dots,u_m)\colon\ u_j\in L^p(\Omega),\ \mathrm{grad}\,u_j\in L_n^p(\Omega)\ j=$

* This research was supported in part by NSF-CME 80-23824 and ARO-DAAG 29-79-C-0161.

0022-1236/82/040102–17$02.00/0

$1, \ldots, m\}$ and $u_0 \in W^{1,p}(\Omega; \mathbb{R}^m)$ is given. By $u \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^m)$ we just mean that $u \in W^{1,p}(\Omega; \mathbb{R}^m)$ and $u = u_0$ on $\partial\Omega$ in a certain generalized sense (see Adams [1] for details),

(iii) $f\colon \mathbb{R}^{nm} \to \mathbb{R}$ is continuous and satisfies, for every $F \in \mathbb{R}^{nm}$, the following coercivity condition

$$a + b\,|F|^p \leqslant f(F) \leqslant c + d\,|F|^p \tag{$*$}$$

for some $a, c \in \mathbb{R}$, $d \geqslant b > 0$ and $p > 1$.

(We will see later that our result allows much more general growth conditions at infinity than $(*)$; we will merely require that f grows faster than linearly at infinity with respect to, at least, one of the subdeterminants of the matrix $F \in \mathbb{R}^{nm}$.)

Usually to prove the existence of solutions for (P) one tries to show that the functional

$$F(u, \Omega) = \int_\Omega f(\nabla u(x))\, dx \tag{0.1}$$

is lower semicontinuous with respect to some type of convergence (for example, weak convergence in $W^{1,p}$), i.e.,

$$\begin{aligned} &F(u, \Omega) \leqslant \liminf F(u_n, \Omega), \\ &u_n \rightharpoonup u \text{ in } W^{1,p}. \end{aligned} \tag{0.2}$$

(Throughout this article we use $\rightharpoonup$ to denote weak convergence.)

Property (0.2) coupled with the coercivity condition $(*)$ leads to the existence of solutions for (P). This method is called the "direct method" (in the calculus of variations). However, in order to ensure (0.2) one has to impose some kind of convexity condition on f. More precisely Morrey [10, 11] has shown that a necessary and sufficient condition on f for F to be lower semicontinuous with respect to weak convergence in $W^{1,p}$ is that f be quasiconvex, i.e.,

$$\int_D f(F + \nabla\xi(x))\, dx \geqslant \int_D f(F)\, dx \tag{0.3}$$

for every bounded open set $D \subset \mathbb{R}^n$, $F \in \mathbb{R}^{nm}$ and $\xi \in C_0^\infty(D; \mathbb{R}^m)$ (the set of C^∞ functions with compact support in D). Condition (0.3) takes a pointwise form in some particular cases as, for example,

(i) if $n = 1$ or $m = 1$, then convexity and quasiconvexity are equivalent concepts. In general (i.e., $n, m > 1$) one only has that if f is convex then it is quasiconvex;

(ii) another example is: if $n = m$ and

$$f(F) = g(\det F) \tag{0.4}$$

then f is quasiconvex if and only if g is convex.

Our aim is to study the existence of solutions for (P) precisely when f fails to be quasiconvex; then, in general, (P) will not have any solution in $W^{1,p}$, but we will prove that it has one in a "generalized" sense. More precisely we will show that if

$$Qf = \sup_{\Phi} \{\Phi \leqslant f \text{ and } \Phi \text{ quasiconvex}\} \tag{0.5}$$

and if

(QP) $$\inf \left\{ \int_{\Omega} Qf(\nabla u(x))\, dx \colon u \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^m) \right\}$$

((QP) is then called the "relaxed" problem) then

THEOREM. (i) *Problem* (QP) *has a solution* $\bar{u} \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^m)$

(ii) $$\inf(\mathrm{P}) = \inf(\mathrm{QP}); \tag{0.6}$$

in other words: *for every solution* $\bar{u}$ *of* (QP), *there exists a minimizing sequence* $\{u^s\}_{s=1}^{\infty}$ *of* (P), $u^s \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^m)$ *such that*

$$u^s \rightharpoonup \bar{u} \text{ in } W^{1,p}(\Omega; \mathbb{R}^m), \tag{0.7}$$

$$\int_{\Omega} f(\nabla u^s(x))\, dx \to \int_{\Omega} Qf(\nabla \bar{u}(x))\, dx. \tag{0.8}$$

Remark. If f does not satisfy growth condition $(*)$ but, for example in the case $n = m$, f satisfies

$$a + b\,|\det F|^p \leqslant f(F) \leqslant c + d\,|\det F|^p$$

then the second part (ii) of the above theorem is still valid provided we replace (0.7) by

$$\det \nabla u^s \rightharpoonup \det \nabla \bar{u} \text{ in } L^p(\Omega).$$

In view of the above theorem, we can conclude that (P) and (QP) are "equivalent" with respect to weak convergence and that solutions of (QP) are therefore "generalized" solutions of (P). From the point of view of applications such a result is interesting since in many physical applications, only averages of physical quantities are actually measured (thus the importance of weak convergence which measures some kind of averages). More generally Tartar (see [12] and the references quoted there) has pointed out the importance of weak convergence for modelling the relationship between microscopic and macroscopic quantities. In [4] it was shown how a particular case of the above theorem could be applied to the study of equilibrium of gases. Other applications to elasticity can be found also in [9] (using the result of [8]).

The notions of "generalized solutions" and "relaxation" theorems have been introduced by Young among others (see [13] for references) and have been used and developed in geometric measure theory (the notion of varifolds...) and in optimal control theory. The more specific result obtained in this article generalizes those of Ekeland and Téman [7, 8] and those already obtained in [4].

1. Quasiconvexity and Lower Quasiconvex Envelope

In this section we will first give the main properties of quasiconvexity that we will need later; at the end of this section we will turn our attention to some properties of the lower quasiconvex envelope Qf of a given function f (see (0.5)).

We mentioned in the Introduction that quasiconvexity arises naturally when one searches for necessary and sufficient conditions for lower semicontinuity under some type of convergence. We give here one typical result involving weak convergence (see Morrey [10, 11]).

THEOREM 1. *Let $f\colon \mathbb{R}^{nm} \to \mathbb{R}$ be continuous, let Ω be a bounded domain of $\mathbb{R}^n$, and*

$$F(u, \Omega) = \int_\Omega f(\nabla u(x))\, dx. \tag{1.1}$$

If $F(u, \Omega)$ is lower semicontinuous with respect to weak convergence in $W^{1,p}(\Omega; \mathbb{R}^m)$ $p \geqslant 1$ in any Ω, then f is quasiconvex, i.e.,

$$\int_\Omega f(F + \nabla\xi(x))\, dx \geqslant f(F) \operatorname{meas} \Omega \tag{1.2}$$

for any $F \in \mathbb{R}^{nm}$, any bounded domain $\Omega \subset \mathbb{R}^n$ and any $\xi \in C_0^\infty(\Omega; \mathbb{R}^m)$. Conversely if f is quasiconvex and satisfies for some $\mu \in \mathbb{R}$ and $K > 0$

(i) *$f(F) \geqslant \mu$ for every $F \in \mathbb{R}^{nm}$,*

(ii) *$|f(F) - f(G)| \leqslant K(1 + |F|^{p-1} + |G|^{p-1})\ |F - G|$ for every $F, G \in \mathbb{R}^{nm}$,* *then* $F(u, \Omega)$ *is lower semicontinuous with respect to weak convergence in* $W^{1,p}(\Omega; \mathbb{R}^m)$.

Although as seen in Theorem 1 quasiconvexity is a "natural" condition when one works in higher dimension ($n > 1$, $m > 1$) it is not very easy to handle, since it is not in a pointwise form, however, it implies a well known condition the so-called Legendre–Hadamard (or ellipticity) condition (for a proof of the next Theorem see [10, 11] or [2]).

THEOREM 2. *Every quasiconvex function f is rank one convex, i.e.,*

$$f(\lambda F + (1-\lambda) G) \leqslant \lambda f(F) + (1-\lambda) f(G) \tag{1.3}$$

for every $\lambda \in [0, 1]$, $F, G \in \mathbb{R}^{nm}$ *with* $\operatorname{rank}(F - G) \leqslant 1$. *Furthermore if* $f \in C^2(\mathbb{R}^{nm})$ *then* (1.3) *is equivalent to the Legendre–Hadamard condition*

$$\sum_{i,j,\alpha,\beta} \frac{\partial^2 f(F)}{\partial F_{i\alpha}\, \partial F_{j\beta}} \lambda_i \lambda_j \xi_\alpha \xi_\beta \geqslant 0 \qquad \textit{for all} \quad \lambda \in \mathbb{R}^n, \quad \xi \in \mathbb{R}^m. \tag{1.4}$$

Remark. (i) The terminology used by Morrey for quasiconvexity and rank one convexity is somehow confusing; in [10] rank one convexity is called weak quasiconvexity while in [11] quasiconvexity is called strong quasiconvexity and rank one convexity is defined as quasiconvexity. We will use throughout this article the terminology of Ball [2, 3].

(ii) Summarizing our notions we have the following diagram

$$f \text{ convex} \Rightarrow f \text{quasiconvex} \Rightarrow f \text{ rank one convex;}$$

the converse of these implications being: an open problem for the second one and false for the first one as the example $f(F) = \det F$ shows.

There are some particular cases where quasiconvexity and rank one convexity are known to be equivalent; they are listed below (for a proof see [10, 11] or [6] Theorem 5.6 and 5.7).

THEOREM 3. (i) *If* $n = 1$ *or* $m = 1$, *then convexity, quasiconvexity and rank one convexity are equivalent.*

(ii) *If* $n = m$ *and there exists* $g\colon \mathbb{R} \to \mathbb{R}$ *continuous such that*

$$f(F) = g(\det F) \tag{1.5}$$

then the convexity of g, the quasiconvexity and rank one convexity of f are equivalent.

(iii) *If* $m = n + 1$, $D\colon \mathbb{R}^{n(n+1)} \to \mathbb{R}^{n+1}$ *is such that*

$$D(F) = (D_1(F), \ldots, D_{n+1}(F))$$

with

$$D_k(F) = (-1)^{k+1} \det(\hat{F}_k), \tag{1.6}$$

where $\hat{F}_k$ is the $n \times n$ matrix obtained by suppressing the k-th line in the $n(n+1)$ matrix F, and if $g: \mathbb{R}^{n+1} \to \mathbb{R}$ is continuous and such that

$$f(F) = g(D(F))$$

then the convexity of g, the quasiconvexity and rank one convexity of f are equivalent.

Remark. Part (ii) of the Theorem is just a particular case of a result we will mention later.

Closely related to the notion of quasiconvexity is the notion of null Lagrangians which by definition are functions Φ such that Φ and $-\Phi$ are quasiconvex, they have the following properties (see [2, 3]).

PROPOSITION 4. *The following properties are equivalent*

(i) *Φ is a null Lagrangian.*

(ii) *For every bounded domain $\Omega \subset \mathbb{R}^n$, for every $F \in \mathbb{R}^{nm}$ and for every $\xi \in C_0^\infty(\Omega; \mathbb{R}^m)$*

$$\int_\Omega \Phi(F + \nabla\xi(x))\, dx = \int_\Omega \Phi(F)\, dx. \tag{1.7}$$

(iii) *Φ is rank one affine, i.e.,*

$$\Phi(\lambda F + (1-\lambda)\, G) = \lambda\Phi(F) + (1-\lambda)\, \Phi(G) \tag{1.8}$$

for every $\lambda \in [0, 1]$, $F, G \in \mathbb{R}^{nm}$ with rank $(F - G) \leqslant 1$.

(iv) *Let $\operatorname{adj}_s F$ denotes the matrix of all $s \times s$ $(1 \leqslant s \leqslant \inf\{n, m\})$ subdeterminants of the matrix $F \in \mathbb{R}^{nm}$, then*

$$\Phi(F) = A + \sum_{s=1}^{\inf\{n,m\}} \langle B_s; \operatorname{adj}_s F\rangle_{\sigma(s)}, \tag{1.9}$$

where

$$\sigma(s) = \binom{m}{s}\binom{n}{s} = \frac{m!\, n!}{s!\,(m-s)!\, s!\,(n-s)!},$$

$\langle \cdot, \cdot \rangle_{\sigma(s)}$ denotes the sclar product in $\mathbb{R}^{\sigma(s)}$, $A \in \mathbb{R}$ and $B_s \in \mathbb{R}^{\sigma(s)}$ are constants.

Remark. (i) If $m=n=2$, then (1.9) can be rewritten as

$$\Phi(F)=A+\sum_{i,j=1}^{2} B_1^{ij}F_{ij}+B_2\det F.$$

(ii) The Euler Lagrange equations for the functional $\int_\Omega \Phi(\nabla u(x))\,dx$ reduce to an identity if Φ is a null Lagrangian.

We now turn our attention to lower quasiconvex envelopes; recall that if $f\colon \mathbb{R}^{nm}\to\mathbb{R}$ is continuous, we have defined Qf by

$$Qf=\sup_{\Phi}\{\Phi\leqslant f,\ \Phi \text{ quasiconvex}\}, \tag{1.10}$$

we then have the following theorem

THEOREM 5. *Let $D\subset\mathbb{R}^n$ be a hypercube and for every $F\in\mathbb{R}^{nm}$ let*

$$Q'f(F)=\inf_{\xi}\left\{\frac{1}{\operatorname{meas} D}\int_D f(F+\nabla\xi(x))\,dx\colon \xi\in C_0^\infty(D;\mathbb{R}^m)\right\}. \tag{1.11}$$

Suppose furthermore that f satisfies the following condition

$$a\leqslant f(F)\leqslant b+c\,|F|^p$$

for some $a,b\in\mathbb{R}$, $c\geqslant 0$, $p\geqslant 1$ and for all $F\in\mathbb{R}^{nm}$; then $Q'f$ is continuous and $Q'f=Qf$.

Proof. The proof is decomposed in four steps.

Step 1. We show first that $Q'f$ is continuous. Let $H\in\mathbb{R}^{nm}$ and $\varepsilon>0$ be arbitrary, we then have by definition that there exist $\xi,\varphi\in C_0^\infty(D;\mathbb{R}^m)$ so that (we take D to be the unit hypercube)

$$\left|Q'f(F)-\int_D f(F+\nabla\xi(x))\,dx\right|\leqslant\varepsilon/2, \tag{1.12}$$

$$\left|Q'f(F+H)-\int_D f(F+H+\nabla\varphi(x))\,dx\right|\leqslant\varepsilon/2. \tag{1.13}$$

Since f is continuous, then by choosing $|H|$ small enough we have

$$\left|\int_D f(F+H+\nabla\xi(x))\,dx-\int_D f(F+\nabla\xi(x))\,dx\right|\leqslant\frac{\varepsilon}{2}, \tag{1.14}$$

$$\left|\int_D f(F+H+\nabla\varphi(x))\,dx-\int_D f(F+\nabla\varphi(x))\,dx\right|\leqslant\frac{\varepsilon}{2}. \tag{1.15}$$

Using the definition of $Q'f$ we get

$$Q'f(F) \leqslant \int_D f(F + \nabla\varphi(x))\,dx \tag{1.16}$$

$$Q'f(F+H) \leqslant \int_D f(F + H + \nabla\xi(x))\,dx. \tag{1.17}$$

Using (1.17) and (1.12) we get

$$\begin{aligned} &Q'f(F+H) - Q'f(F) \\ &\quad\leqslant \int_D f(F+H+\nabla\xi(x))\,dx - \int_D f(F+\nabla\xi(x))\,dx + \frac{\varepsilon}{2}; \end{aligned} \tag{1.18}$$

applying (1.14) to (1.18) we obtain

$$Q'f(F+H) - Q'f(F) \leqslant \varepsilon. \tag{1.19}$$

Similarly using (1.16), (1.13) and (1.15) we obtain

$$\begin{aligned} &Q'f(F) - Q'f(F+H) \\ &\quad\leqslant \int_D f(F+\nabla\varphi(x))\,dx - \int_D f(F+H+\nabla\varphi(x))\,dx + \frac{\varepsilon}{2} \leqslant \varepsilon. \end{aligned} \tag{1.20}$$

Thus $Q'f$ is continuous.

Step 2. We next want to show that the definition of $Q'f$ is independent of D. Let D_1 and D_2 be two hypercubes of $\mathbb{R}^n$, then there exist $\lambda > 0$ and $x_0 \in \mathbb{R}^n$ so that $D_2 = x_0 + \lambda D_1$ (up to a rotation). Let for $F \in \mathbb{R}^{nm}$

$$Q_1'f(F) = \inf_{\xi} \left\{ \frac{1}{\operatorname{meas} D_1} \int_{D_1} f(F + \nabla\xi(x))\,dx : \xi \in C_0^\infty(D_1; \mathbb{R}^m) \right\}, \tag{1.21}$$

$$Q_2'f(F) = \inf_{\xi} \left\{ \frac{1}{\operatorname{meas} D_2} \int_{D_2} f(F + \nabla\xi(x))\,dx : \xi \in C_0^\infty(D_2; \mathbb{R}^m) \right\}. \tag{1.22}$$

Then for every $\xi \in C_0^\infty(D_2; \mathbb{R}^m)$ we have

$$Q_2'f(F) \leqslant \frac{1}{\operatorname{meas} D_2} \int_{D_2} f(F + \nabla\xi(x))\,dx \tag{1.23}$$

therefore

$$Q'_2 f(F) \leqslant \frac{1}{\lambda^n \operatorname{meas} D_1} \int_{x_0+\lambda D_1} f(F + \nabla\xi(x))\, dx$$

$$= \frac{1}{\operatorname{meas} D_1} \int_{D_1} f(F + \nabla\xi(x_0 + \lambda y))\, dy. \tag{1.24}$$

Observe that for every $\varphi \in C_0^\infty(D_1; \mathbb{R}^m)$ we have that

$$\xi(y) = \lambda\varphi\left(\frac{y - x_0}{\lambda}\right) \in C_0^\infty(D_2; \mathbb{R}^m). \tag{1.25}$$

Therefore apply (1.24) to the function ξ defined in (1.25) to get

$$Q'_2 f(F) \leqslant \frac{1}{\operatorname{meas} D_1} \int_{D_1} f(F + \nabla\varphi(y))\, dy; \tag{1.26}$$

taking the infinimum in (1.26) over every $\varphi \in C_0^\infty(D_1; \mathbb{R}^m)$ we get

$$Q'_2 f(F) \leqslant Q'_1 f(F).$$

Similarly one gets that $Q'_1 f \leqslant Q'_2 f$.

Step 3. Let D be the unit hypercube; we now want to show that for every $F \in \mathbb{R}^{nm}$ and for every $\varphi \in C_0^\infty(D; \mathbb{R}^m)$ we have

$$\int_D Q'f(F + \nabla\varphi(y))\, dy \geqslant \int_D Q'f(F)\, dy = Q'f(F). \tag{1.27}$$

For every $\varepsilon > 0$ we can approximate $\varphi \in C_0^\infty(D; \mathbb{R}^m)$ by a piecewise affine function $\psi \in W_0^{1,\infty}(D; \mathbb{R}^m)$ (see Proposition 2.1 p. 286 in [8]) such that

$$\left| \int_D Q'f(F + \nabla\varphi(x))\, dx - \int_D Q'f(F + \nabla\psi(x))\, dx \right| \leqslant \frac{\varepsilon}{3}. \tag{1.28}$$

Since ψ is piecewise affine we may decompose D into open subsets D_i, $1 \leqslant i \leqslant N$, and find $A_i \in \mathbb{R}^{nm}$ so that

$$\nabla\psi = A_i \qquad \text{in } D_i$$

and

$$\int_D Q'f(F + \nabla\psi(y))\, dy = \sum_{i=1}^{N} \operatorname{meas} D_i Q'f(F + A_i). \tag{1.29}$$

We now use the definition of $Q'f$ to get $\xi_i \in C_0^\infty(D_i; \mathbb{R}^m)$ $1 \leqslant i \leqslant N$ such that

$$Q'f(F + A_i) + \frac{\varepsilon}{3} \geqslant \frac{1}{\operatorname{meas} D_i} \int_{D_i} f(F + A_i + \nabla \xi_i(x))\, dx, \quad 1 \leqslant i \leqslant N. \tag{1.30}$$

We then define $\chi: D \to \mathbb{R}^m$ by

$$\chi(x) = \psi(x) + \xi_i(x) \qquad \text{for} \quad x \in D_i. \tag{1.31}$$

Observe that $\chi \in W_0^{1,\infty}(D; \mathbb{R}^m)$; therefore there exists $Z \in C_0^\infty(D; \mathbb{R}^m)$ so that

$$\left| \int_D f(F + \nabla \chi(x))\, dx - \int_D f(F + \nabla Z(x))\, dx \right| \leqslant \frac{\varepsilon}{3}. \tag{1.32}$$

Combining (1.29) and (1.30) we get

$$\sum_{i=1}^N \int_{D_i} f(F + \nabla \chi(x))\, dx \leqslant \int_D Q'f(F + \nabla \psi(x))\, dx + \frac{\varepsilon}{3}, \tag{1.33}$$

i.e.,

$$\int_D f(F + \nabla \chi(x))\, dx \leqslant \int_D Q'f(F + \nabla \psi(x))\, dx + \frac{\varepsilon}{3}.$$

Now use (1.32) and (1.28) in (1.33) to obtain

$$\int_D f(F + \nabla Z(x))\, dx \leqslant \int_D Q'f(F + \nabla \varphi(x))\, dx + \varepsilon; \tag{1.34}$$

since $\varepsilon > 0$ is arbitrary and $Z \in C_0^\infty(D; \mathbb{R}^m)$ we deduce (1.27).

Step 4. We are now able to complete the proof of the theorem.

(i) We want first to show that $Q'f$ is quasiconvex. Let Ω be any bounded open set of $\mathbb{R}^n$, $\varphi \in C_0^\infty(\Omega; \mathbb{R}^m)$ and $F \in \mathbb{R}^{nm}$. We want to prove that

$$\int_\Omega Q'f(F + \nabla \varphi(x))\, dx \geqslant \int_\Omega Q'f(F)\, dx. \tag{1.35}$$

Let D be a hypercube containing Ω and extend φ to be identically 0 in $D - \Omega$ then apply Step 3 (and Step 2 which shows that the definition of $Q'f$ is independent of D) to get

$$\int_{\Omega} Q'f(F+\nabla\varphi(x))\,dx = \int_{D} Q'f(F+\nabla\varphi(x))\,dx - \int_{D-\Omega} Q'f(F)\,dx$$
$$\geqslant \int_{D} Q'f(F)\,dx - \int_{D-\Omega} Q'f(F)\,dx. \tag{1.36}$$

(ii) It finally remains to prove that $Q'f = Qf$. Let $h \leqslant f$ be quasiconvex we then deduce from the definition of the quasiconvexity of h that

$$Q'h = h. \tag{1.37}$$

Since $h \leqslant f$ we also have

$$h = Q'h \leqslant Q'f \leqslant f \tag{1.38}$$

for every quasiconvex function $h \leqslant f$, hence $Qf \leqslant Q'f$. Since $Q'f$ is also quasiconvex we deduce the result. ∎

2. Main Result

We now set the hypotheses and prove the main theorem. The basic assumption we will have to make is that f is continuous and grows faster at infinity than linearly with respect to at least one of the subdeterminants of the matrix $\nabla u \in \mathbb{R}^{nm}$. More precisely we will assume that

(H1) $\Omega \subset \mathbb{R}^n$ is a bounded open set with Lipschitz boundary,

(H2) $f\colon \mathbb{R}^{nm} \to \mathbb{R}$ is continuous and satisfies the following coercivity condition

$$a + \sum_{\nu=1}^{N} b_\nu \, |\Phi_\nu(F)|^{\beta_\nu} \leqslant f(F) \leqslant c + \sum_{\nu=1}^{N} d_\nu \, |\Phi_\nu(F)|^{\beta_\nu} \tag{$*$}$$

for every $F \in \mathbb{R}^{nm}$, for some $a, c \in \mathbb{R}$, $N \geqslant 1$ (an integer), $\beta_\nu > 1$, $d_\nu \geqslant b_\nu > 0$ and where $\Phi_\nu\colon \mathbb{R}^{nm} \to \mathbb{R}$, $\nu = 1,\ldots, N$, are null Lagrangians.

Example. (i) The case where f satisfies a condition

$$a + b\,|F|^\beta \leqslant f(F) \leqslant c + d\,|F|^\beta$$

is a particular case of (H2), it suffices to choose $N = nm$, $\beta_\nu = \beta > 1$, $d_\nu = d \geqslant b_\nu = b > 0$ for $\nu = 1,\ldots, N$ and for $F = (F_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m}$

$$\begin{cases} \Phi_1(F) = F_{11}, ..., \Phi_n(F) = F_{n1} \\ \Phi_{n+1}(F) = F_{12}, ..., \Phi_{2n}(F) = F_{n2} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad ..., \Phi_{nm}(F) = F_{nm}. \end{cases}$$

(ii) Another case contained in (H2) is the case where $n = m$ and f satisfies

$$a + b\,|\det F|^\beta \leqslant f(F) \leqslant c + d\,|\det F|^\beta,$$

then choose $N = 1$, $\Phi_\nu(F) = \det F$.

Remarks. (i) Hypotheses (H1) and (H2) are equivalent to those considered by Ekeland and Témam [8, Chap. X], where functionals of the form $\int_\Omega f(\text{gradu}(x))\,dx$, $u: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ (i.e., $m = 1$) are studied, and to those considered in [4], where the functionals have the form $\int_\Omega f(\Phi(\nabla u(x)))\,dx$, $u: \Omega \subset \mathbb{R}^3 \to \mathbb{R}^3$ and Φ is a null Lagrangian. However, the coercivity condition $(*)$ is too strong to include the case of parametric problems, in particular that considered in [5], where the functional has the form $\int_\Omega f(D(\nabla u(x)))\,dx$, $u: \Omega \subset \mathbb{R}^n \to \mathbb{R}^{n+1}$ and D is as in Theorem 3; in these problems $(*)$ is satisfied only for $\beta_\nu = 1$, $\nu = 1,..., N$ (while in (H2), $\beta_\nu > 1$), since f is positively homogeneous of degree 1.

(ii) Observe that

$$h(F) = a + \sum_{\nu=1}^{N} b_\nu\,|\Phi_\nu(F)|^{\beta_\nu}$$

is quasiconvex and hence

$$h \leqslant Qf \leqslant f.$$

RELAXATION THEOREM. *Under the above hypotheses we have that, for every $u \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ with $u = u_0$ on $\partial\Omega(u_0 \in W^{1,\infty}(\Omega; \mathbb{R}^m))$ there exists $\{u^s\}_{s=1}^\infty$; $u^s \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ such that*

(i) *$u^s = u_0$ on $\partial\Omega$ for every $s \geqslant 1$,*

(ii) *$\Phi_\nu(\nabla u^s) \rightharpoonup \Phi_\nu(\nabla u)$ in $L^{\beta_\nu}(\Omega)$, $\nu = 1,..., N$, as $s \to \infty$,*

(iii) *$\int_\Omega f(\nabla u^s(x))\,dx \to \int_\Omega Qf(\nabla u(x))\,dx$ as $s \to \infty$.*

Remark. Observe that if f satisfies the coercivity condition of Example 1 above, i.e.,

$$a + b\,|F|^\beta \leqslant f(F) \leqslant c + d\,|F|^\beta,$$

then the conclusion (ii) of the above theorem means that

$$u^s \rightharpoonup u \qquad \text{in } W^{1,\beta}(\Omega; \mathbb{R}^m).$$

Proof of Theorem. Fix $\eta > 0$ and first observe that there is no loss of generality if we suppose u piecewise affine in Ω. Otherwise we may find $O \subset \Omega$, an open set, and $w \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ such that (see Prop. 2.9, Chap. X in [8])

$$\text{meas}(\Omega - O) \leqslant \eta, \tag{2.1}$$

$$w \text{ is piecewise affine in } O, \tag{2.2}$$

$$w = u_0 \qquad \text{on } \partial\Omega, \tag{2.3}$$

$$|w(x) - u(x)| \leqslant \eta \qquad \text{for all} \quad x \in \Omega, \tag{2.4}$$

$$\|\nabla w - \nabla u\|_{L^p} \leqslant \eta \qquad \text{for all} \quad p \geqslant 1. \tag{2.5}$$

So if we can prove the theorem for w and O, by extending u^s outside O by $u^s = u$ we will have proved the theorem for every $u \in W^{1,\infty}(\Omega; \mathbb{R}^m)$.

Since u is piecewise affine in Ω we may decompose Ω into open sets Δ_i $1 \leqslant i \leqslant I$ so that ∇u is constant in Δ_i. Then decompose Δ_i into small hypercubes R_i^p $1 \leqslant p \leqslant P_i$ so that

$$\left| \int_{\Delta_i - \bigcup_{p=1}^{P_i} R_i^p} Qf(\nabla u(x))\, dx \right| \leqslant \frac{\eta}{3I}, \tag{2.6}$$

$$\left| \int_{\Delta_i - \bigcup_{p=1}^{P_i} R_i^p} f(\nabla u(x))\, dx \right| \leqslant \frac{\eta}{3I}. \tag{2.7}$$

But on each R_i^p we have

$$Qf(\nabla u) = \inf_\xi \left\{ \frac{1}{\text{meas } R_i^p} \int_{R_i^p} f(\nabla u + \nabla \xi(x))\, dx \colon \xi \in C_0^\infty(R_i^p; \mathbb{R}^m) \right\}.$$

Therefore there exists $\{\xi_q\}_{q=1}^\infty$, $\xi_q \in C_0^\infty(R_i^p; \mathbb{R}^m)$ so that for q large enough

$$\left| \int_{R_i^p} f(\nabla u + \nabla \xi_q(x))\, dx - \int_{R_i^p} Qf(\nabla u)\, dx \right| \leqslant \frac{\eta}{3IP_i}. \tag{2.8}$$

Now extend ξ_q by periodicity (in each variable) from R_i^p to the whole of $\mathbb{R}^n$ and let for s an integer

$$\Psi_{q,s}(x) = \frac{1}{s}\, \xi_q(sx). \tag{2.9}$$

We obtain from (2.9) that

$$\Psi_{q,s}(x)=0 \quad \text{if } x\in \partial R_i^p \text{ (since } \xi_q \text{ is periodic and } \xi_q\in C_0^\infty(R_i^p;\mathbb{R}^m)) \tag{2.10}$$

$$\nabla \Psi_{q,s}(x)=\nabla\xi_q(sx). \tag{2.11}$$

We now use the coercivity condition (H2) on f (and Qf); from (*) we deduce that $\{\Phi_\nu(\nabla u+\nabla\xi_q)\}_{q=1}^\infty$ is bounded in $L^{\beta_\nu}(R_i^p)$ (and $\beta_\nu>1$) for every $\nu=1,\dots,N$. Therefore there exist $\delta_\nu\in L^{\beta_\nu}(R_i^p)$ so that (up to a subsequence)

$$\Phi_\nu(\nabla u+\nabla\xi_q(x)) \rightharpoonup \delta_\nu(x) \quad \text{in} \quad L^{\beta_\nu}(R_i^p), \quad \nu=1,\dots,N \text{ as } q\to\infty. \tag{2.12}$$

Observe also that since Φ_ν is a null Lagrangian and $\xi_q\in C_0^\infty(R_i^p;\mathbb{R}^m)$ we have

$$\int_{R_i^p}\delta_\nu(x)\,dx=\lim_{q\to\infty}\int_{R_i^p}\Phi_\nu(\nabla u+\nabla\xi_q(x))\,dx=\int_{R_i^p}\Phi_\nu(\nabla u), \quad \nu=1,\dots,N. \tag{2.13}$$

As we extended ξ_q by periodicity (in (2.9)), we do it for δ_ν (in each variable) from R_i^p to the whole of $\mathbb{R}^n$; we therefore deduce that

$$\delta_\nu(sx) \rightharpoonup \frac{1}{\text{meas } R_i^p}\int_{R_i^p}\delta_\nu(x)\,dx=\Phi_\nu(\nabla u) \quad \text{in } L^{\beta_\nu}(R_i^p),\ \nu=1,\dots,N. \tag{2.14}$$

We then take the diagonal sequence of (2.12) and (2.14), to get

$$\Phi_\nu(\nabla u+\nabla\xi_s(sx)) \rightharpoonup \Phi_\nu(\nabla u) \quad \text{in} \quad L^{\beta_\nu}(R_i^p),\ \nu=1,\dots,N. \tag{2.15}$$

Summarizing (2.9), (2.10) and (2.15) and defining for $x\in R_i^p$

$$u^s(x)=u(x)+\Psi_{s,s}(x)=u(x)+\frac{1}{s}\,\xi_s(sx), \tag{2.16}$$

we have obtained

$$u^s(x)=u(x) \quad \text{for} \quad x\in\partial R_i^p, \tag{2.17}$$

$$\Phi_\nu(\nabla u^s) \rightharpoonup \Phi_\nu(\nabla u) \quad \text{in} \quad L^{\beta_\nu}(R_i^p),\ \nu=1,\dots,N. \tag{2.18}$$

We therefore have defined u^s on R_i^p, and hence on $\bigcup_{p=1}^{P_i} R_i^p$ (using (2.17)). By letting $u^s=u$ on $\Delta_i-\bigcup_{p=1}^{P_i} R_i^p$, we have then constructed u^s on Δ_i $1\leqslant i\leqslant I$, and thus on $\Omega=\bigcup_{i=1}^I \Delta_i$. Obviously from the construction of u^s and from (2.18), u^s has the required properties (i) and (ii); so it only remains to prove (iii)

$$\left|\int_\Omega Qf(\nabla u(x))\,dx - \int_\Omega f(\nabla u^s(x))\,dx\right|$$
$$\leqslant \sum_{i=1}^{I} \left|\int_{\Delta_i} Qf(\nabla u)\,dx - \int_{\Delta_i} f(\nabla u^s(x))\,dx\right|$$
$$\leqslant \sum_{i=1}^{I} \left\{ \left|\int_{\Delta_i - \bigcup_{p=1}^{P_i} R_i^p} Qf(\nabla u)\,dx\right| + \left|\int_{\Delta_i - \bigcup_{p=1}^{P_i} R_i^p} f(\nabla u^s(x))\,dx\right| \right\}$$
$$+ \sum_{i=1}^{I} \sum_{p=1}^{P_i} \left|\int_{R_i^p} Qf(\nabla u)\,dx - \int_{R_i^p} f(\nabla u^s(x))\,dx\right|. \tag{2.19}$$

We now use (2.6), (2.7) and the fact that $u^s = u$ on $\Delta_i - \bigcup_{p=1}^{P_i} R_i^p$ to get

$$\left|\int_\Omega Qf(\nabla u(x))\,dx - \int_\Omega f(\nabla u^s(x))\,dx\right|$$
$$\leqslant \frac{2\eta}{3} + \sum_{i=1}^{I} \sum_{p=1}^{P_i} \left|\int_{R_i^p} Qf(\nabla u)\,dx - \int_{R_i^p} f(\nabla u^s(x))\,dx\right|. \tag{2.20}$$

Finally observe that

$$\int_{R_i^p} f(\nabla u^s(x))\,dx = \int_{R_i^p} f(\nabla u + \nabla \xi_s(sx))\,dx$$
$$= \frac{1}{s^n} \int_{sR_i^p} f(\nabla u + \nabla \xi_s(y))\,dy$$
$$= \int_{R_i^p} f(\nabla u + \nabla \xi_s(y))\,dy, \tag{2.21}$$

the last equality being a consequence of the periodicity of ξ_s.

Therefore using (2.8) and (2.21) we get

$$\left|\int_{R_i^p} Qf(\nabla u)\,dx - \int_{R_i^p} f(\nabla u^s(x))\,dx\right| \leqslant \frac{\eta}{3IP_i}; \tag{2.22}$$

combining (2.20) and (2.22) we obtain the result. ∎

Conclusions

(i) We deduce from the Theorem that

$$\inf(\mathrm{P}) = \inf(\mathrm{QP})$$

(see the introduction for the definition of (P) and (QP)). However, in general the hypothesis (H2) is not sufficient to guarantee the existence of solutions for (QP); but if f satisfies the following particular form of $(*)$

$$a + b\,|F|^{\beta} \leqslant f(F) \leqslant c + d\,|F|^{\beta} \tag{2.23}$$

with $\beta > 1$, then the direct methods of the calculus of variations imply that (QP) possesses a solution (see [11]). Similarly if $n = m = 3$ and $f(F) = g(\det F)$ then (QP) possesses also a solution (see [4] for details).

(ii) We have seen in the Introduction the interpretation that one can give to the above theorem; some applications of this type of results to elasticity can be found in [4] and in [9].

(iii) The result obtained in the above theorem should be extendable to functionals of the form $\int_{\Omega} f(x, u(x), \nabla u(x))\,dx$ (with appropriate continuity and coercivity conditions on f) since the important dependence is that on ∇u (by Rellich's theorem).

(iv) The proof of the above theorem is shorter and simpler than those in [7, 8] (there the functionals have the form $\int_{\Omega} f(\mathrm{grad} u(x))\,dx$, $u: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ and therefore $Qf = f^{**}$, the lower convex envelope of f) and in [4] (there $f(\nabla u(x)) = g(\Phi(\nabla u(x)))$ with $u: \Omega \subset \mathbb{R}^3 \to \mathbb{R}^3$ and Φ a null Lagrangian, therefore $Qf = g^{**}$, the lower convex envelope of g), although these two results [4, 8] are particular cases of the above theorem. However, they are more constructive since the minimizing sequences $\{u^s\}_{s=1}^{\infty}$ are constructed explicitly.

(v) Finally it is interesting to note that while in [4, 7, 8] it was sufficient to consider ξ_q (see (2.8)) which were finite sums of characteristic functions, it is not sufficient in the general case considered here.

Acknowledgments

I would like to thank W. H. Fleming for useful discussions when writing this paper, especially those related to Theorem 5.

References

1. R. A. Adams, "Sobolev Spaces," Academic Press, New York, 1975.
2. J. M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rational Mech. Anal.* **63** (1977), 337–403.
3. J. M. Ball, Constitutive inequalities and existence theorems in nonlinear elastostatics, *in* "Heriot–Watt Symposium," Vol. 1, Pitman, New York, 1976.
4. B. Dacorogna, A relaxation theorem and its application to the equilibrium of gases, *Arch. Rational Mech. Anal.* (1982).

5. B. DACOROGNA, Minimal hypersurfaces problems in parametric form with nonconvex integrands, *Indiana Math. J.* (1982).
6. B. DACOROGNA, "Weak Continuity and Weak Lower Semicontinuity of Non Linear Functionals," Brown University Reports, L.C.D.S. 81–17, 1981.
7. I. EKELAND, Sur le contrôle optimal de systèmes gouvernés par des équations elliptiques, *J. Funct. Anal.* **3** (1972), 1–62.
8. I. EKELAND AND R. TÉMAM, "Analyse convexe et problèmes variationels," Dunod, Gauthier–Villars, Paris, 1976.
9. M. E. GURTIN AND R. TÉMAM, On the anti-plane shear problem in finite elasticity, *J. Elasticity* **11**, No. 2 (1981).
10. C. B. MORREY, Quasiconvexity and the lower semicontinuity of multiple integrals, *Pacific J. Math.* **2** (1952), 25–53.
11. C. B. MORREY, "Multiple Integrals in the Calculus of Variations," Springer–Verlag, Berlin, 1966.
12. L. TARTAR, Compensated compactness, *in* "Heriot–Watt Symposium," Vol. 4, Pitman, New York, 1978.
13. L. C. YOUNG, "Lectures on the Calculus of Variations and Optimal Control Theory," Saunders, Philadelphia, 1969.