Convexity Conditions in the direct methods in the Calculus of variations. Submitted in fulfillment of the requirements of MAM4001

Daniel Mckenzie
Supervisor: Dr Francois Ebobisse

Department of Maths and Applied Maths, University of Cape Town

(Dated: November 23, 2010)

This paper aims to discuss the question of determining whether solutions to problems of the form (.1) exist

$$\min\{\int_{\Omega} f(x,u(x),\nabla u(x))dx : u \in X \text{ and } u|_{\partial\Omega} = u_0\} \tag{.1}$$

where $\Omega \subset \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^m$. Specifically we shall discuss what are known as the direct methods in the calculus of variations. After outlining the concepts from functional analysis, convex analysis and the theory of Sobolev spaces that we shall require, we consider a generalisation of convexity termed *quasiconvexity* which has been shown by Morrey (Morrey, 1952) amongst others to be a necessary and sufficient condition for the existence of solutions to (.1) within a suitable space X, when supplemented with additional coercivity and growth conditions. The condition of quasiconvexity is however hard to verify, and is unsuited to several important applications of the theory, notably to mathematical elasticity, and as such we consider another generalistion of convexity, *polyconvexity* which was succesfully used by Ball ((Ball, 1977)) and others to prove the existence of several practical problems in the theory of hyperelasticity.

Contents

I. INTRODUCTION

Minimisation problems of the form (.1) arise frequently in various areas of applied mathematics. Most notably from the point of view of the direct methods, the problem of determining how a hyperelastic function will deform subject to given boundary conditions can be formulated as a variational problem of the form (.1) (see (Ciarlet, 1988), (Dacorogna, 1989), (Ball, 1977)). Indeed, it was such problems that motivated J.M. Ball to develop the existence theory based on polyconvex functions that we shall discuss later. A full discussion of mathematical elasticity is beyond the scope of this paper however, and we shall content ourselves with analysing (.1) as a problem in the calculus of variations. We shall follow Dacorogna in calling the case where either $m = 1$ or $n = 1$ the scalar case, while the case where $m > 1$ and $n > 1$ shall be deemed the vectorial case. This distinction is warranted as there is a marked difference in the treatment of the two cases. For brevity we shall frequently use the short-hand $I(u) = \int_\Omega f(x, u(x), \nabla u(x))dx$. Under suitable hypotheses for $I(u)$ (it should be Gâteaux differentiable) and u_0 (it should be twice continuously differentiable, i.e. $u_0 \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$) a necessary condition for a solution to (.1) is the Euler condition:

$$u_0 \text{ is a solution to (.1)} \Rightarrow I^{'}(u_0) = \left(-\sum_{i=1}^{n} \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial \xi_{i\alpha}}(x, u_0(x), \nabla u_0(x))\right) + \frac{\partial f}{\partial u_\alpha}(x, u_0(x), \nabla u_0(x))\right)_{1\le\alpha\le m} = 0 \tag{I.1}$$

Where $f = f(x, u, \xi)$ and $u : \mathbb{R}^n \to \mathbb{R}^m$. If the functional $I(u)$ is in addition *convex* over X the reverse implication above holds and (I.1) may be used to find solutions to (.1). However the abovementioned hypotheses can be unnecessarily restrictive. In many cases of of interest $I(u)$ might not be differentiable, or perhaps we wish to look for solutions to (.1) in a 'larger' space than $\mathcal{C}^2(\Omega, \mathbb{R}^m)$. In this paper we shall consider the more general problem of determining whether or not a solution for (.1) exists within a specified space under weaker hypotheses on $I(u)$. Specifically we shall focus on results by Ball (Ball, 1977), Morrey (**??**) and predecessors pertaining to nonlinear elasticity that show that, under several other milder assumptions on $f(x, u, \nabla u)$, either of two generalisations of convexity - polyconvexity or quasiconvexity - are sufficient to guarantee the existence of solutions to (.1). The intuitive idea behind the direct methods is to find a minimising sequence for

$$\inf\{I(u) : u \in X\} \tag{I.2}$$

which is contained in a compact subset of X. Recall that a minimising sequence is a sequence $\{u_n\}$ such that

$$I(u_n) \to \inf\{I(u) : u \in X\} \text{ as } n \to \infty \tag{I.3}$$

and that if $\{u_n\} \subset K \subset X$ where K is compact, then we can extract a convergent subsequence $\{u_{n_l}\}$, $u_{n_l} \to u_0 \in K$. If $I(u)$ is continuous (in fact, as we shall later show, we require only that $I(u)$ satisfies the weaker condition of *lower semicontinuity*) then $I(u_{n_l}) \to I(u_0)$. Additionally, since $\{I(u_{n_l})\}$ is a subsequence of $\{I(u_n)\}$, $I(u_{n_l}) \to \inf\{I(u) : u \in X\}$. Thus $I(u_0) = \inf\{I(u) : u \in X\}$ and we have found a minimum as required. However, difficulties arise since typically X will be a normed space of functions $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, that is infinite dimensional. Using the norm topology in an infinite dimensional space we no longer have that:

$$K \subset X \text{ closed and bounded} \Rightarrow K \text{ compact} \tag{I.4}$$

thus a bounded sequence need not belong to a compact subset. It is for this reason that we consider a coarser topology on X, the *weak topology*, and it is in proving the lower semicontinuity of I with respect to the weak topology that we shall use the quasi- or poly- convexity of f.
This report is structured as follows. In section 2 we review the necessary concepts from functional and convex analysis required. In section 3, we develop the direct methods, focusing specifically how convexity (and its generalizations) relates to the weak lower semicontinuity of I,while in section 4, we combine the results of section 1 and 2 to obtain the existence theorems of Ball and Morrey.

II. PRELIMINARY IDEAS

A. Banach Spaces

The structure with which we shall be primarily dealing is the *Banach space.*

Definition 1 (Banach space)**.** A Banach space is a vector space X over the field $\mathbb{F}$ (in all cases of interest in this paper, $\mathbb{F} = \mathbb{R}$), equipped with a norm $||\cdot||_X$ such that X is complete with respect to the metric induced by $||\cdot||_X$. All Banach spaces are topological vector spaces, that is, the operations of addition and scalar multiplication are continuous with respect to the metric induced by $||\cdot||_X$.

We define the *dual* of X, X^* as the space of all linear continuous maps $x^* : X \to \mathbb{F}$. X^* is also a vector space over $\mathbb{F}$. The *bidual* of X is the space of all continuous linear maps $x^{**} : X^* \to \mathbb{F}$ is again a vector space, which we shall denote as X^{**}. We shall use the notation $\langle x; x^* \rangle$ to denote $x^*(x)$. Using this notation, we see that all $x \in X$ can equally well be considered as maps $X^* \to \mathbb{F}$, and so we can define a map

$$J : X \to X^{**} \tag{II.1}$$

$$x \mapsto \hat{x} \tag{II.2}$$

$$\forall x^* \in X^* : \ \hat{x}(x^*) = \langle x; x^* \rangle \tag{II.3}$$

$$\tag{II.4}$$

It can easily be seen that J is injective, linear and continuous. If J is in addition surjective then X and X^{**} are isomorphic and X is said to be *reflexive*.
Examples of Banach space which we shall use extensively are the Lebesgue spaces.

Definition 2 (Lebesgue space)**.** Suppose $\Omega \subset \mathbb{R}^n$. For $1 \leq p < \infty$, we define the pth Lebesgue space, $L^p(\Omega, \mathbb{R})$ as the space of all (equivalence classes of) [1] functions $u : \Omega \to \mathbb{R}$ such that :

$$||u||_{L^p} \equiv \left(\int_\Omega |f(x)|^p dx \right)^{1/p} < +\infty \tag{II.5}$$

For brevity we shall denote $L^p(\Omega; \mathbb{R})$ as $L^p(\Omega)$ where no ambiguity can arise. It can easily be checked that $L^p(\Omega)$, equipped with the norm $|| \cdot ||_{L^p}$ is a Banach space over $\mathbb{R}$ which is reflexive for $1 < p < \infty$ L^p. We can similarly define L^∞ as the space of all (equivalence classes of) functions $u : \Omega \to \mathbb{R}$ such that:

$$||u||_{L^\infty} \equiv \sup\{|u(x)| : x \in \Omega\} < \infty \tag{II.6}$$

which is also a vector space over $\mathbb{R}$ and complete with respect to its norm.

The dual of $L^p(\Omega)$ has a particularly nice description:

$$(L^p(\Omega))^{'} \cong L^{p^{'}}(\Omega) \tag{II.7}$$

where the $p^{'}$, the conjugate exponent of p, is defined as:

$$p = \begin{cases} \infty & p = 1 \\ \frac{p}{p-1} & 1 < p < \infty \\ 1 & p = \infty \end{cases} \tag{II.8}$$

The above construction can easily be generalised for functions $u : \Omega \to \mathbb{R}^m$ by taking m products of $L^p(\Omega)$:

$$L^p(\Omega, \mathbb{R}^m) = L^p(\Omega, \mathbb{R})^m \tag{II.9}$$

and equipping the resulting space with an appropriate norm, for example

$$||\mathbf{u}|| = \sum_{i=1}^{m} ||u_i(x)||_{L^p} \tag{II.10}$$

And whenever $L^p(\Omega)$ is reflexive, so is $L^p(\Omega, \mathbb{R}^m)$. Before moving on we make an observation about the relationship between $L^p(\Omega)$ spaces for bounded Ω, which we state as a lemma:

Lemma II.1. Suppose $\Omega \subset \mathbb{R}^n$ is bounded and open. Then

$$L^p(\Omega, \mathbb{R}^m) \subset L^q(\Omega, \mathbb{R}^m) \quad \text{for } 1 \leq q \leq p \text{ if } 1 \leq p \leq \infty \tag{II.11}$$

and furthermore there exists an $\alpha \in \mathbb{R}$ such that

$$||u||_{L^q} \leq \alpha ||u||_{L^p} \quad \forall\ u \in L^p(\Omega, \mathbb{R}^m) \tag{II.12}$$

[1] The equivalence relation here is equality almost everywhere: $f \sim g \Leftrightarrow f(x) = g(x)\ \forall x \in \Omega \setminus E$ where E is a set of measure zero

Proof. We shall present the proof for $L^p(\Omega, \mathbb{R})$. The proof for $L^p(\Omega, \mathbb{R}^m)$ follows easily. Suppose $u \in L^p(\Omega, \mathbb{R})$, then

$$||u||_{L^p} = (\int_\Omega |u|^p dx)^{1/p} < \infty \tag{II.13}$$

so

$$\begin{aligned} ||u||_{L^q} &= (\int_\Omega |u|^q dx)^{1/q} && \text{(II.14)} \\ &= (\int_\Omega |1u^q| dx)^{1/q} && \text{(II.15)} \\ &\leq \left((\int_\Omega |u^q|^{p/q} dx)^{q/p} (\int_\Omega |1|^{p/(p-q)} dx)^{(p-q)/p} \right)^{1/q} && \text{(II.16)} \\ &= ||u||_{L^p} \mu(\Omega)^{(p-q)/pq} && \text{(II.17)} \\ &< \infty && \text{(II.18)} \\ & && \text{(II.19)} \end{aligned}$$

Where the third line above follows from Hölders inequality. □

On a Banach space it is possible to generalize the idea of a derivative from finite dimensional calculus:

Definition 3 (The Gâteaux derivative)**.** Suppose X and Y are Banach spaces, and $f : X \to Y$ The directional derivative of f at $x \in X$ in the direction of $y \in X$ is defined as

$$f_x^{'}(y) = \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x+ty) - f(x)}{t} \tag{II.20}$$

f is said to be Gâteaux differentiable at x if the above limit exists for all $y \in X$ and if the map

$$y \mapsto \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x+ty) - f(x)}{t} \tag{II.21}$$

is linear and continuous. The Gâteaux derivative of f at x is then a linear map from X to Y which we denote as f_x. We say that f is Gâteaux differentiable on X if it is Gâteaux differentiable at all $x \in X$.

In most cases of interest in this paper Y shall be $\mathbb{R}$, in which case $f_x^{'} \in X^*$.

B. The adjugate matrix

Given a 3×3 matrix $\mathbf{A}$ we define its *adjugate* as the matrix of the cofactors, D_{ij} of $\mathbf{A}$. Recall that we define the cofactors of $\mathbf{A}$ as

$$\begin{aligned} D_{ij} &= (-1)^{i+j} \det \mathbf{A}_{ij}^{'} && \text{(II.22)} \\ (\operatorname{adj} \mathbf{A})_{ij} &= D_{ij} && \text{(II.23)} \\ & && \text{(II.24)} \end{aligned}$$

where $\mathbf{A}_{ij}^{'}$ denotes the 2×2 matrix obtained from $\mathbf{A}$ by omitting the i-th row and j-th column. By using the definition of the determinant of a two-by two matrix we may write

$$\operatorname{adj} \mathbf{A}_{ij} = A_{i+1,j+1} A_{i+2,j+2} - A_{i+1,j+2} A_{i+2,j+1} \;\; 1 \leq i,j \leq 3 \tag{II.25}$$

Where the indices are to be understood modulo 3, or more concisely, using the Levi-Civita tensor ϵ:

$$(\operatorname{adj} \mathbf{A})_{i\alpha} = \frac{(-1)^{i+\alpha}}{2} \epsilon_{\alpha\beta\gamma} \epsilon_{ijk} a_{\beta j} a_{\gamma k} \tag{II.26}$$

C. The weak topology

Suppose X is a normed space. The weak topology on X is defined as the weakest topology for which all $x^* \in X^*$ remain continuous. It is the topology generated by the family of semi-norms $||x||_{x^*} = |\langle x; x^*\rangle|,\ x^* \in X^*$. The basic open sets of this topology are of the form:

$$\{x \in X : |\langle x; x^*\rangle| \leq \epsilon\} \tag{II.27}$$

For a fixed $x^* \in X^*$. The weak-$*$ topology is defined in a similar manner, it is defined as the coarsest topology on X^* such that all $x \in X$ (when considered as members of the bidual X^{**}) remain continuous. It is generated by the family of semi-norms $||x^*||_x = |\langle x, x^*\rangle|$, and basic open sets are defined in an identical manner as before. A sequence $\{x_n\} \subset X$ is *weak convergent* if it converges with respect to the weak topology. We shall denote this as $x_n \rightharpoonup x$. From the definition of the basic open sets in (II.27) we see that:

$$x_n \rightharpoonup x \Leftrightarrow \langle x_n; x^*\rangle \to \langle x, x^*\rangle \ \forall x^* \in X^* \tag{II.28}$$

A sequence $\{x_n^*\} \subset X^*$ is *weak-$*$ convergent* if it converges with respect to the weak-$*$ topology. This shall be denoted as $x_n^* \rightharpoonup^* x^*$, and again it is not too difficult to see that:

$$x_n^* \rightharpoonup^* x^* \Leftrightarrow \langle x^*; x\rangle \to \langle x^*; x\rangle \forall x \in X \tag{II.29}$$

A few further results regarding weak and weak-$*$ convergent sequences are expressed in the following theorems, the proof of which can be found in any good book on functional analysis.

Theorem II.2. *Let X be a Banach space.*

1. *Suppose $x_n \rightharpoonup x$, then there exists $K > 0$ such that $||x_n||_X \leq K\ \forall n \in \mathbb{N}$ and furthermore $||x||_X \leq \liminf_{n\to\infty} ||x_n||_X$*

2. *Suppose $x_n^* \rightharpoonup^* x^*$, then there exists $K > 0$ such that $||x_n^*||_{X^*} \leq K\ \forall n \in \mathbb{N}$ and furthermore $||x^*|| \leq \liminf_{n\to\infty} ||x_n^*||_{X^*}$*

3. *If $x_n \to x$ then $x_n \rightharpoonup x$*

4. *If $x_n^* \to x^*$ then $x_n^* \rightharpoonup^* x^*$*

5. *If $F : X \to \mathbb{R}$ is weak-continuous, then F is norm-continuous*

6. *if $F : X^* \to \mathbb{R}$ is weak-$*$-continuous, then F is norm-continuous*

Theorem II.3 (The Banach-Alaoglu theorem)**.** *The closed unit ball of X^* (which we shall henceforth denote as B_{X^*}) is weak-$*$ compact*

Theorem II.4. *If X is a reflexive normed space then the weak topology on X coincides with the weak-$*$ topology on X^{**}.*

We can now justify why the weak topology is so important for the direct methods. By combining theorems (II.3) and (II.4) we see that if X is a reflexive Banach space, then every bounded sequence $\{x_n\}$ is contained within a copy of the unit ball KB_X which is a weak-compact set, thus $\{x_n\}$ has a weakly-convergent subsequence [2] Before moving on, we note one more result about weak convergent sequences, which links weak convergence with the notion of convexity.

---

[2] A subtlety here is that for a topology τ the notions of compactness and sequential compactness need not coincide. However we are saved by the Eberlein-Šmulian theorem, which states that if X is a Banach space then $A \subset X$ is compact if and only if A is sequentially compact (Whitley, 1967).

Theorem II.5 (Mazur's Lemma). *Suppose that X is a Banach space and $x_\nu \rightharpoonup x$ in X. Then for all $\epsilon \geq 0$ there exists an $n = n(\epsilon)$ (i.e. n depends on ϵ) and scalars $\{\alpha_i(\epsilon) : i = 1 \ldots, n\}$ (which will also depend on ϵ) with $\sum_{i=1}^{n} \alpha_i = 1$ such that*

$$||x - \sum_{i=1}^{n} \alpha_i x_i|| \leq \epsilon \tag{II.30}$$

Proof. A proof can be found in (Ekeland and Temam, 1972) □

Let us now consider the continuity of $I(u)$.

Definition 4. 1. Let V be a topological space. A function $J : V \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous if its epigraph:

$$\text{epi}\, J = \{(v, a) \in V \times \mathbb{R} : j(v) \leq a\} \tag{II.31}$$

is closed in the space $V \times \mathbb{R}$

2. J as defined above is *sequentially* lower semicontinuous if:

$$u_k \to u_0 \in V \Rightarrow J(u) \leq \lim \inf_{k\to\infty} J(u_k) \tag{II.32}$$

Note that if J is lower semicontinuous then J is also sequentially lower semicontinuous, but the converse is true only if V is metrisable. The weaker condition of sequential weak lower semicontinuity will suffice for our purposes.

D. Sobolev Spaces

As mentioned, we wish to relax the restrictions on the space of functions in which we are looking for a solution to (.1). The natural spaces to consider are the *Sobolev spaces*. We shall present here only the results necessary for developing the direct methods, and in general these will be presented without proof . For a more thorough introduction the reader is referred to (Adams, 1975) or (Kuttler, 2006).
Suppose that $\Omega \subset \mathbb{R}^n$ is open. We denote by $\mathcal{D}(\Omega)$ the set of all real-valued, infinitely differentiable functions defined on Ω with compact support contained in Ω. With addition and scalar multiplication defined in a natural manner:

$$(f + g)(x) = f(x) + g(x) \tag{II.33}$$
$$(\lambda f)(x) = \lambda f(x) \tag{II.34}$$
$$\tag{II.35}$$

$\mathcal{D}(\Omega)$ becomes a vector space, and by equipping it with the inductive limit topology it is in fact a topological vector space. Thus we may define its dual, $\mathcal{D}^{'}(\Omega)$. Elements of $\mathcal{D}^{'}(\Omega)$ are known as *distributions*.
$\mathcal{D}^{'}(\Omega)$ is in fact a very large space. We can get an idea of this by noting that if $f : \Omega \to \mathbb{R}$ is locally integrable (that is, for all $K \subset \Omega$, K compact, $f \in L^1(K)$ then

$$T_f : \mathcal{D}(\Omega) \to \mathbb{R} \tag{II.36}$$
$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)dx \tag{II.37}$$
$$\tag{II.38}$$

is well defined (since the support of φ is some compact set K) and $T_f \in \mathcal{D}^{'}(\Omega)$. Since all $f \in L^p(\Omega)$ $1 \leq p \leq \infty$ are locally integrable, in some sense $\mathcal{D}^{'}(\Omega)$ contains all L^p spaces.We now define a derivative of sorts for distributions, often referred to as the generalised, or weak derivative.

Definition 5 (The weak derivative). Suppose $\Omega \subset \mathbb{R}^n$ is open and $T \in \mathcal{D}^{'}(\Omega)$ then $\partial_i T \in \mathcal{D}^{'}(\Omega)$ is defined as follows:

$$\partial_i T(\varphi) = -T(\partial_i \varphi) \quad \text{where } \partial_i = \frac{\partial}{\partial x_i} \tag{II.39}$$

Remark II.6. 1. If f is a locally integrable function that is in addition differentiable then

$$\partial_i T_f(\varphi) = -\int_\Omega f(x)\partial_i\varphi(x)dx = \int_\Omega \partial_i f(x)\varphi(x)dx = T_{\partial_i f}(\varphi) \tag{II.40}$$

Where the second equality follows from an application of integration by parts and the fact that φ has compact support contained in Ω. Hence the weak derivative coincides with the regular derivative where both are defined.

2. The weak derivative greatly expands the set of things which are differentiable. For example we can now calculate the derivative of pathological functions such as the Heaviside step function:

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{II.41}$$

$$\partial\theta(\varphi) = -\int_{-\infty}^{\infty} \theta\partial_x\varphi(x)dx = -\int_0^{\infty} \partial_x\varphi dx = [\varphi(x)]_0^{\infty} = \varphi(0) = \int_{-\infty}^{\infty} \delta(x)\varphi(x)dx = T_\delta(\varphi) \tag{II.42}$$

i.e. the weak derivative of the Heaviside step function is the Dirac delta function.

We define higher order derivatives in a natural manner:

$$\partial_i\partial_j T(\varphi) = -(\partial_i T(\partial_j\varphi)) = (-1)^2 T(\partial_i\partial_j\varphi) \tag{II.43}$$

and more generally if $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is a multi index:

$$\partial_\alpha T(\varphi) = (-1)^{|\alpha|}T(\partial_\alpha\varphi) = T(\partial_1^{\alpha_1}\partial_2^{\alpha_2}\ldots\partial_k^{\alpha_k}\varphi) \tag{II.44}$$

where

$$|\alpha| = \sum_{i=1}^{k}\alpha_i \tag{II.45}$$

Since all φ under consideration are infinitely differentiable it is easy to see that weak partial derivatives *always* commute.
Now suppose $u \in L^p(\Omega)$. For an arbitrary multi index α, consider $\partial_\alpha T_u(\varphi)$ as defined above and suppose there exists a $g \in L^p(\Omega)$ such that

$$\partial_\alpha T_u(\varphi) = (-1)^{|\alpha|}\int u(x)\partial_\alpha\varphi dx = \int g\varphi dx = T_g(\varphi)\ \forall\varphi \in \mathcal{D}(\Omega) \tag{II.46}$$

By a lemma in (Kuttler, 2006) if such a g exists it is necessarily unique, thus we write $g = \partial_\alpha u \in L^p(\Omega)$. This leads us to the definition of *Sobolev spaces*

Definition 6 (Sobolev Spaces)**.** Suppose $\Omega \subset \mathbb{R}^n$ is open. Then the Sobolev space $W^{m,p}(\Omega)$ is defined as the set of all functions $u \in L^p(\Omega)$ such that for all multi indices α, $|\alpha| \leq m$, $\partial_\alpha \in L^p$. We equip $W^{m,p}(\Omega)$ with the norm

$$||u||_{m,p} \equiv \left(\int_\Omega \sum_{|\alpha|\leq m} |\partial_\alpha u|^p\right)^{1/p} \tag{II.47}$$

although since we are effectively placing a norm on a subspace of k products of the Banach space $L^p(\Omega)$ where k is the number of multi indices smaller than or equal to m any other 'product norm' such as :

$$||u||_{W^{m,p}} \equiv \sum_{|\alpha|\leq m} ||\partial_\alpha||_{L^p} \tag{II.48}$$

will induce the same topology. We shall denote as $W_0^{m,p}(\Omega)$ the closure of $\mathcal{D}(\Omega)$ in $W^{m,p}(\Omega)$ with respect to the norm $||\cdot||_{W^{m,p}}$

If $u : \mathbb{R}^m \to \mathbb{R}^n$, we shall frequently use the notation

$$\nabla u = \begin{pmatrix} \partial_1 u^1 & \dots & \partial_m u^1 \\ \vdots & \ddots & \vdots \\ \partial_1 u^n & \dots & \partial_m u^n \end{pmatrix} \tag{II.49}$$

and if $u \in W^{1,p}(\Omega)$ then the partial derivatives $\partial_i u^j$ are to be understood in a weak sense.

Remark II.7. 1. $W^{m,p}(\Omega)$ is a Banach space for all $m \in \mathbb{N}$ $m \geq 0$ $1 \leq p \leq \infty$, is separable for $1 \leq p < \infty$ and reflexive if $1 < p < \infty$.

2. By replacing $L^p(\Omega)$ with $L^p(\Omega, \mathbb{R}^k)$ wherever necessary we may define $W^{m,p}(\Omega, \mathbb{R}^k)$ and equip it with the norm

$$||\mathbf{u}||_{m,p} \equiv \left(\int_\Omega \sum_{|\alpha| \leq m} \sum_{l=1}^{k} |\partial_\alpha u_l|^p dx \right)^{1/p} \tag{II.50}$$

and as in the case of L^p spaces whenever $W^{m,p}(\Omega, \mathbb{R})$ is reflexive, so is $W^{m,p}(\Omega, \mathbb{R}^k)$.

3. We shall not attempt to describe $(W^{m,p}(\Omega))^{'}$ here (for this the reader is referred to (Adams, 1975)). However we shall note that for every $L \in (W^{m,p}(\Omega))^{'}$ we can find $\{v_\alpha\}_{0 \leq \alpha \leq k} \subset L^{p^{'}}$ where k is the total number of multi indices smaller than m such that

$$L(u) = \sum_{0 \leq |\alpha| \leq m} \int_\Omega \partial_\alpha u v_\alpha dx \tag{II.51}$$

and we may define weak convergence of sequences in Sobolev space as follows:

$$u_\nu \rightharpoonup u \text{ in } W^{m,p} \Leftrightarrow \partial_\alpha u_\nu \rightharpoonup \partial_\alpha u \in L^p(\Omega) \ 0 \leq |\alpha| \leq m \tag{II.52}$$

4. Note that if $||u||_{m,p} < \infty$ then it is obvious that $||u||_p < \infty$. Hence if $u \in W^{m,p}(\Omega)$ it is also in $L^p(\Omega)$. Conversely, it is easy to construct functions $u \in L^p(\Omega)$ which are not in $W^{m,p}(\Omega)$. For example take $\Omega = (0, 1)$ and $u = 1/\sqrt{x}$. then

$$\int_0^1 u dx = 2\sqrt{x}|_0^1 = 2 \tag{II.53}$$

but

$$\int_0^1 \partial_x u dx = 1/\sqrt{x}|_0^1 = \infty \tag{II.54}$$

Another useful feature of Sobolev spaces is their relation to the space of affine functions and the space of smooth functions, as described in the next two theorems. Recall that we define the space of affine functions as follows:

Definition 7 (Affine functions)**.** Suppose $\Omega \subset \mathbb{R}^n$ is bounded. $u : \Omega \to \mathbb{R}$ is affine if there exists a finite collection of disjoint open sets $\{\Omega_1, \dots, \Omega_n\}$ such that:

1. $\bar{\Omega} = \cup_{i=1}^n \bar{\Omega}_i$

2. ∇u is constant on each Ω_i

The set of all affine functions on Ω shall be denoted as $Aff(\Omega)$, and the set of all $u \in Aff(\Omega)$ such that $u|_{\partial\Omega} = 0$ shall be denoted as $Aff_0(\Omega)$.

Theorem II.8. *Suppose $\Omega \subset \mathbb{R}^n$ is bounded with Lipschitz boundary (that is, $\partial\Omega$ is the image of a Lipschitz continuous function).*

1. *Let $u \in W_0^{1,\alpha}(\Omega)$ with $1 \leq \alpha < \infty$, then there exists a sequence $\{u_\nu\} \subset Aff_0(\Omega)$ such that*

$$u_\nu \to u \text{ in } W^{1,\alpha}(\Omega) \tag{II.55}$$

2. *Let $u \in W^{1,\infty}(\Omega)$, then there exists a sequence $\{u_\nu\} \subset Aff_0(\Omega)$ such that*

$$u_\nu \to u \text{ in } W^{1,\alpha}(\Omega) \text{ for every } 1 \leq \alpha < \infty \tag{II.56}$$

$$\text{and } ||\nabla u_\nu||_{L^\infty} \leq K||\nabla u||_{L^\infty} \tag{II.57}$$

$$\tag{II.58}$$

where $K \in \mathbb{R}$ is a constant.

(Dacorogna, 1989) pg27.

The space of all smooth functions is also dense in $W^{m,p}(\Omega)$:

Theorem II.9 (Meyer-Serrin)**.** *Suppose $\Omega \subset \mathbb{R}^n$ is an open connected set . If $m \in \mathbb{Z}$ and $1 \leq p < \infty$ then:*

$$\overline{\mathcal{C}^\infty(\Omega) \cap W^{m,p}(\Omega)} = W^{m,p}(\Omega) \tag{II.59}$$

with the closure being taken with respect to the Sobolev norm $|| \cdot ||_{m,p}$. (Ciarlet, 1988) pg 279

Proof. A proof can be found in (Kuttler, 2006) □

We remarked above that $W^{m,p}(\Omega) \subset L^q(\Omega)$, where $p = q$. This observation is made precise as well as generalised to certain cases where $p \neq q$ by the following embedding theorem, due in part to Sobolev, Rellich and Koňdrasov. We shall not present this theorem in its full generality, but rather consider only special cases that will be relevant later. First, however, recall that a normed space X is said to be *embedded* in another normed space Y if there exists an isomorphism:

$$j : X \to j(X) \subset Y \tag{II.60}$$

and we write

$$X \hookrightarrow Y \tag{II.61}$$

The embedding is said to be *compact* if j is a compact linear map.

Theorem II.10 (Sobolev, Rellich and Koňdrasov embedding theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open and connected set and let $1 \leq p \leq \infty$*

- *If $1 \leq p < n$ then*

$$W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \text{ for every } 1 \leq q \leq \frac{np}{n-p} \tag{II.62}$$

  and the embedding is compact for every $1 \leq q < \frac{np}{n-p}$

- *If $p = n$ then*

$$W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \text{ for every } 1 \leq q < \infty \tag{II.63}$$

  and the embedding is compact.

- *If $p > n$, then*

$$W^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega}) \tag{II.64}$$

and the embedding is compact.

Remark II.11. 1. The above results can be stated in full generality for arbitrary $W^{m,p}(\Omega)$ (see for example (Ciarlet, 1988)).

2. The fact that the embeddings are compact is particularly useful. To see this, recall that a compact operator $\mathcal{K} : X \to Y$ maps weakly convergent sequences into convergent sequences. So, considering the second case in the theorem as an example, we have that

$$u_\nu \rightharpoonup u \text{ in } W^{1,p}(\Omega) \Rightarrow u_\nu \to u \text{ in } L^q(\Omega) \ 1 \leq q < \infty \tag{II.65}$$

We shall use this result repeatedly later.

A final result regarding Sobolev spaces allows us to bound the L^p-norm of an arbitrary function in $W^{1,p}(\Omega)$ by a (fixed) multiple of the L^p norm of its derivative

Theorem II.12 (Poincaré inequality)**.** *Suppose that $\Omega \subset \mathbb{R}^n$ is open and bounded, and that $1 \leq p < \infty$, then:*

$$||u||_{L^p} \leq K||\nabla u||_{L^p} \tag{II.66}$$

for every $u \in W^{1,p}(\Omega)$ and for some $K > 0$, constant.

Remark II.13. A more general statement of the above theorem is possible, requiring only that Ω is open and of *finite width* (that is, it lies between two parallel hyperplanes) (See (Ciarlet, 1988)).

E. Convexity

Recall that a function $f : X \to \mathbb{R}$ is said to be convex if:

$$f(\lambda x + (1-\lambda)y \leq \lambda f(x) + (1-\lambda)f(y) \ \lambda \in [0,1], \ \forall x, y \in X \tag{II.67}$$

Note that this can easily be extended to:

$$f(\sum_{i=1}^{N} \lambda_i x_i) \leq \sum_{i=1}^{N} \lambda_i f(x_i) \tag{II.68}$$

where $\sum_{i=1}^{N} \lambda_i = 1$ and $\lambda_i \geq 0 \ \forall i$. The theory of convex functions is particularly useful in proving the weak lower semicontinuity of our functional $I(u)$. In the vectorial case convexity is an unnecessarily strong condition, and we work instead with the related notions of poly- and quasi- convexity.

Definition 8. 1. Let $f : \mathbb{R}^{nm} \to \mathbb{R}$ be a Borel measurable and locally integrable function. f is said to be quasiconvex if:

$$f(A) \leq \frac{1}{\mu(D)} \int_D f(A + \nabla\varphi(x))dx \tag{II.69}$$

for every bounded domain $D \subset \mathbb{R}^n$ every $A \in \mathbb{R}^{nm}$ and every $\varphi \in W_0^{1,\infty}(D; \mathbb{R}^m)$. $\mu(D)$ is the measure of D:

$$\mu(D) = \int_D dx \tag{II.70}$$

2. Let $f : \mathbb{R}^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$. f is said to be *polyconvex* if there exists a $g : \mathbb{R}^{3\times 3+3\times 3+1} \to \bar{\mathbb{R}}$ which is convex such that:

$$f(A) = g(T(A)) \quad \forall A \in \mathbb{R}^{3\times 3} \tag{II.71}$$

where T is defined as follows:

$$T : \mathbb{R}^{3\times 3} \to \mathbb{R}^{3\times 3+3\times 3+1} \tag{II.72}$$
$$T : A \mapsto (A, \operatorname{adj} A, \det A) \tag{II.73}$$
$$\tag{II.74}$$

Remark II.14. 1. Polyconvexity can be defined for functions $f : \mathbb{R}^{nm} \to \bar{\mathbb{R}}$ for arbitrary m,n (see for example (Dacorogna, 1989)) although the definition is somewhat long-winded and obscures the motivation for introducing the concept. Since we are developing the direct methods for application to three-dimensional elasticity, the definition given above shall suffice.

2. Note that while polyconvexity can be defined for functions taking values in $\bar{\mathbb{R}}$, the notion of quasiconvexity does not generalise to this case. (See (Dacorogna, 1989), pg. 100 for discussion).

3. If (II.69) holds for one domain D, then it in fact holds for all domains. This can easily be shown:

Proof. Suppose that (II.69) holds for a domain D. We wish to show that for arbitrary $E \subset \mathbb{R}^n$:

$$\int_E f(A + \nabla\varphi(x))dx \geq f(A)\mu(A) \quad \forall A \in \mathbb{R}^{nm},\ \varphi \in W_0^{1,\infty}(E, \mathbb{R}^m) \tag{II.75}$$

We can always find $x_0 \in \mathbb{R}^n$ and $\epsilon > 0$ such that $x_0 + \epsilon E \subset D$. Define $\psi(x)$ as follows:

$$\psi(x) = \begin{cases} \epsilon\varphi(\frac{x-x_0}{\epsilon} & \text{if } x \in x_0 + \epsilon E \\ 0 & \text{otherwise} \end{cases} \tag{II.76}$$

Then $\psi \in W^{1,\infty}(D, \mathbb{R}^m)$ and

$$\int_E f(A + \nabla\phi(y))dy = \epsilon^{-n} \int_{x_0+\epsilon E} f(A + \nabla\phi(\frac{x - x_0}{\epsilon}))dx \tag{II.77}$$
$$= \epsilon^{-n} \left(\int_D f(A + \nabla\psi(x))dx - f(A)\mu(D - (x_0 + \epsilon E)) \right) \tag{II.78}$$
$$\tag{II.79}$$

Where we have made the substitution:

$$y = \frac{x - x_0}{\epsilon} \Rightarrow dy = dy_1 \dots dy_n = (\frac{dx_1}{\epsilon}) \dots (\frac{dx^n}{\epsilon} = \epsilon^{-n}dx$$

Now by assumption $\int_D f(A + \nabla\psi) \geq \mu(D)f(A)$. Also:

$$\mu(D - (x_0 + \epsilon E)) = \int_{D-(x_0+\epsilon E)} dx = \mu(D) - \epsilon^{-n}\mu(E)$$

Hence

$$\int_E f(A + \nabla\varphi(y))dy \geq \epsilon^{-n} \left(\mu(D)f(A) - f(A)(\mu(D) - \epsilon^n\mu(E))\right) \tag{II.80}$$
$$= f(A)\mu(E) \tag{II.81}$$

as required. ∎

The various notions of convexity are related as follows:

Theorem II.15. *1. Suppose $f : \mathbb{R}^{nm} \to \mathbb{R}$, then:*

$$f \text{ convex} \Rightarrow f \text{ polyconvex} \Rightarrow f \text{quasiconvex} \tag{II.82}$$

2. Suppose $f : \mathbb{R}^{nm} \to \bar{\mathbb{R}}$, then:

$$f \text{ convex} \Rightarrow f \text{ polyconvex} \tag{II.83}$$

3. If $m = 1$ or $n = 1$ then all these notions are equivalent.

Proof. A proof can be found in (Dacorogna, 1989). □

Note that the third point in the above theorem hints at why the scalar case needs to be treated separately from the vectorial case. Before we study the concepts of poly- and quasiconvexity further let us recall a few concepts from convex analysis, all of which can be found in (Dacorogna, 1989).

Theorem II.16. *The following are equivalent:*

1. *f is convex*

2. epi f *is convex*

Proof. a proof can be found in (Ekeland and Temam, 1972) □

Definition 9. 1. A *hyperplane* is a set of the form

$$H = \{x \in X : f(x) = \alpha\} \tag{II.84}$$

where f is a linear functional $f : X \to \mathbb{R}$ and $\alpha \in \mathbb{R}$.

2. A hyperplane H as defined above *separates* (respectively *separates strictly*) two sets $A, B \subset X$ if $f(x) \leq \alpha \; \forall x \in A$ and $f(x) \geq \alpha \; \forall x \in B$ (respectively if $f(x) < \alpha \; \forall x \in A$ and $f(x) > \alpha \; \forall x \in B$)

Proposition 1. A hyperplane $H = \{x \in X : f(x) = \alpha\}$ is closed if and only if f is continuous. That is, if $f \in X^*$.

We now state without proof an important theorem linking convex sets and hyperplanes that will be useful later (its proofs can be found in any standard text on functional analysis)

Theorem II.17 (Hahn-Banach (Geometric version))**.** *Let X be a Banach space.*

1. *Let $A, B \subset X$ be non-empty, disjoint and convex. Let A be open, then there exists a closed hyperplane which separates A and B.*

2. *Let $A, B \subset X$ be non-empty, disjoint and convex. Let A be closed and B compact, then there exists a closed hyperplane which strictly separates A and B.*

3. *Every closed convex set is the intersection of the closed half spaces which contain it.*

If f is convex and in addition Gâteaux differentiable then the convexity condition has a nice interpretation in terms of the Gâteaux derivative, which is presented in (II.18).

Theorem II.18. *Suppose X is a Banach space and $f : X \to \mathbb{R}$ is Gâteaux differentiable. Then the following are equivalent.*

1. *f is convex.*

2. *for every $x, y \in X$*

$$f(y) \geq f(x) + \langle y - x; f^{'}(x)\rangle \tag{II.85}$$

Proof. A proof can be found in Dacorogna (Dacorogna, 1989) ☐

We can extend this result to convex functions which are merely continuous, not necessarily Gâteaux differentiable

Proposition 2. Suppose $f : X \to \mathbb{R}$ continuous and convex. For every $x \in X$ there exists an $x^* \in X^*$ (referred to as the subgradient of f at x) such that

$$f(y) \geq f(x) + \langle y - x; x^*\rangle \quad \forall y \in X \tag{II.86}$$

In the case where X is a Banach space and $f(x)$ is Gâteaux differentiable, $f^{'}(x) = x^*$.

Proof. A proof can be found in (Dacorogna, 1989) ☐

III. THE DIRECT METHODS

Let us formalise the approach outlined in the introduction.

Theorem III.1. *Suppose X is a reflexive Banach space and $I : X \to \bar{\mathbb{R}}$ is sequentially weakly lower semicontinuous (s.w.l.s.c, and this abbreviation shall be used consistently from here on) and coercive over X. That is:*

$$I(u) \geq \alpha||u|| + \beta \tag{III.1}$$

for some $\alpha > 0$, $\beta \in \mathbb{R}$. Suppose also that there exists $\tilde{u} \in X$ with $I(\tilde{u}) < \infty$. Then

$$\inf\{I(u) : u \in X\} \tag{III.2}$$

has at least one solution.

Remark III.2.
1. The coercivity condition ensures that $I(u) \to \infty$ as $||u|| \to \infty$, which means that the infimum in (III.2) is finite. This is generally quite a mild restriction in that many functionals of interest are coercive and this property is easy to check.

2. The assumption of the existence of such a $\tilde{u}$ is merely to rule out the degenerate case where $I(u) \equiv +\infty \; \forall u \in X$

Proof. By the coercivity assumption we have that:

$$\begin{aligned} \beta &\leq I(u) - \alpha||u|| && \text{(III.3)}\\ &\leq I(u) \; \forall u \in X && \text{(III.4)}\\ &\leq \inf\{I(u) : u \in X\} && \text{(III.5)}\\ & && \text{(III.6)} \end{aligned}$$

Now let $\{u_k\}$ be a minimizing sequence for (III.2). Again from the coercivity we have that:

$$||u_k|| \leq \frac{1}{\alpha}(I(u_k) - \beta) \quad \forall k \tag{III.7}$$

Where by (III.6) the right hand side of (III.7) is non-negative for all k. Now observe that since $\{I(u_k)\}$ is a convergent sequence of real numbers it is bounded, and hence the sequence $\{||u_k||\}$ is bounded. So by the discussion in section 2C we may extract a weakly convergent subsequence $\{u_{k_l}\}$ such that $u_{k_l} \rightharpoonup \bar{u}$ and $I(u_{k_l}) \to \inf\{I(u) : u \in X\}$. Now by the s.w.l.s.c of I and the definition of an infimum:

$$\inf\{I(u) : u \in X\} \leq I(\bar{u}) \leq \liminf_{u_{k_l} \rightharpoonup \bar{u}} I(u_{k_l}) = \inf\{I(u) : u \in X\} \tag{III.8}$$

Hence $\bar{u}$ is a solution to (III.2) ☐

Let us return to the original problem:

$$\min\{\int_\Omega f(x,u(x),\nabla u(x))dx : u \in X \text{ and } u|_{\partial\Omega} = u_0\} \tag{III.9}$$

$$u:\Omega\to\mathbb{R}^m \quad \Omega\subset\mathbb{R}^n \tag{III.10}$$

$$\tag{III.11}$$

By choosing X appropriately (for example $X = W^{1,p}(\Omega;\mathbb{R}^m)$) we can always ensure that it is a reflexive Banach space. Furthermore, in most cases of interest in nonlinear elasticity, where f represents an appropriate stored energy function for our elastic material, I will be coercive. The hard part then is determining whether or not I is weakly lower semicontinuous. We shall now examine three distinct cases where the convexity, polyconvexity, or quasiconvexity of f will be shown to be both necessary and sufficient for the s.w.l.s.c. of I, and our approach shall be as follows.

1. We first show that if

$$f:\Omega\times\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R} \tag{III.12}$$

and I is s.w.l.s.c. then f is *quasiconvex*. This gives us the necessity of quasiconvexity in the vectorial case.

2. We specialise the above result to the case where either m or n equals 1 (what we have previously called the scalar case) and note that in this case (see (II.15))

$$f \text{ convex} \Leftrightarrow f \text{ quasiconvex} \tag{III.13}$$

This gives the necessity of convexity in the scalar case.

3. We then show that if

$$f:\Omega\times\mathbb{R}^m\times\mathbb{R}^N\to\bar{\mathbb{R}} \tag{III.14}$$

$$J(u,\xi) = \int_\Omega f(x,u,\xi)dx \tag{III.15}$$

and $f(x,u,\cdot)$ is convex (as well as satisfying some other milder conditions that we shall specify later) and

$$u_\nu \to \bar{u} \quad \text{in } L^p(\Omega;\mathbb{R}^m)\ p\geq 1 \tag{III.16}$$

$$\xi_\nu \rightharpoonup \bar{\xi} \text{ in } L^q(\Omega;\mathbb{R}^N)\ q\geq 1 \tag{III.17}$$

$$\tag{III.18}$$

Then

$$\liminf_{\nu\to\infty} J(u_\nu,\xi_\nu) \geq J(\bar{u},\bar{\xi}) \tag{III.19}$$

4. Specialising this result to the case where $\xi_\nu = \nabla u$ gives us the sufficiency of convexity (in both scalar and vectorial cases)

5. Following this we shall show that in the vectorial case if f is defined as in (III.12) then the *quasiconvexity* of f is a sufficient condition for I to be s.w.l.s.c. Note that although this is a more general result than the above in the sense that quasiconvexity is a weaker notion than convexity, we have restricted ourselves to functions f which can only take finite values.

6. Finally, since in nonlinear elasticity it is frequently useful to consider functions which are allowed to take the value ∞, we use result 3 to show that the *polyconvexity* is a sufficient condition for the s.w.l.s.c of I when

$$f:\Omega\times\mathbb{R}^m\times\mathbb{R}^n\to\bar{\mathbb{R}} \tag{III.20}$$

First, however we state an elementary observation about Sobolev spaces that shall be useful later.

Lemma III.3. Suppose that $\Omega \subset \mathbb{R}^n$ is bounded and $u_n \rightharpoonup^* u$ in $W^{1,\infty}(\Omega, \mathbb{R}^m)$. Then $u_n \rightharpoonup u$ in $W^{1,p}(\Omega, \mathbb{R}^m)$ for all $1 \leq p < \infty$.

Proof. We shall prove the lemma for $W^{1,\infty}(\Omega, \mathbb{R})$. The proof for the general case is similar. Suppose $u_n \rightharpoonup^* u$ in $W^{1,\infty}(\Omega, \mathbb{R})$

$$\int_\Omega \partial_\alpha u_n \varphi dx \to \int_\Omega \partial_\alpha u \varphi dx \quad \forall \varphi \in L^1(\Omega) \quad \alpha = 0, \dots, n \tag{III.21}$$

In particular, since $L^{p'}(\Omega) \subset L^1(\Omega)$ by lemma II.1 we have that

$$\int_\Omega \partial_\alpha u_n \varphi dx \to \int_\Omega \partial_\alpha u \varphi dx \quad \forall \varphi \in L^p(\Omega) \quad \alpha = 0, \dots, n \tag{III.22}$$

so $u_n \rightharpoonup u$ in $W^{1,p}(\Omega, \mathbb{R})$. □

An immediate consequence of the above lemma is that if $I : W^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R}$ is s.w.l.s.c. then $I : W^{1,\infty}(\Omega; \mathbb{R}^m) \to \mathbb{R}$ is s.w.l.s.c. since for an arbitrary weakly convergent sequence in $W^{1,\infty}$ that weak-star converges to $u \in W^{1,\infty}$, $u_n \rightharpoonup u$ in $W^{1,p}$ and by the s.w.l.s.c of I

$$\liminf_{n \to \infty} I(u_n) \geq I(u) \tag{III.23}$$

as required.

A. The necessity of quasiconvexity

Theorem III.4. *Suppose $\Omega \subset \mathbb{R}^n$ is open and bounded, and that $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \to \mathbb{R}$ is continuous and satisfies*

$$|f(x, u, \xi)| \leq a(x, |u|, |\xi|) \quad \forall \ (x, u, \xi) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \tag{III.24}$$

where a is increasing with respect to $|u|$ and $|\xi|$ and is locally integrable with respect to x. If I is weak $*$ lower semicontinuous in $W^{1,\infty}(\Omega, \mathbb{R}^m)$ then $f(x, u, \cdot)$ is quasiconvex.

Remark III.5. The idea behind this theorem is originally due to Tonelli (Tonelli, 1921) and has since been improved upon by numerous authors. See (Dacorogna, 1989) for a complete list.

Before we prove the theorem, we need a lemma:

Lemma III.6. Let $D = \prod_{i=1}^n (a_i, b_i)$ (i.e. an open n-dimensional cuboid) and $f \in L_p(\Omega)$ $1 \leq p \leq \infty$. Extend f by periodicity from Ω to $\mathbb{R}^n$ and let

$$f_\nu(x) = f(\nu x) \tag{III.25}$$

then

$$f_\nu \rightharpoonup \bar{f} = \frac{1}{\mu(\Omega)} \int_\Omega f(x) dx \quad \text{in } L_p(\Omega) \text{ if } 1 \leq p < \infty \tag{III.26}$$

$$f_\nu \rightharpoonup^* \bar{f} \quad \text{if } p = \infty \tag{III.27}$$

$$\tag{III.28}$$

Proof. A proof can be found in (Dacorogna, 1989) □

Proof. We shall follow the proof contained in (Dacorogna, 1989), although this is based on earlier work by Meyers (Meyers, 1965) and Morrey (Morrey, 1952). We simplify the proof slightly by proving it for the case where $0 \in \Omega$. The case for a general region Ω is then easily obtained by a translation.
We wish to show that for every cube $D \subset \Omega$, every $(x_0, u_0, \xi_0) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm}$ and for every $\varphi \in W_0^{1,\infty}(D; \mathbb{R}^m)$

$$\int_D f(x_0, u_0, \xi_0 + \nabla \varphi(y)) dy \geq \mu(D) f(x_0, u_0, \xi_0) \tag{III.29}$$

So consider such a cube $D = \{\mathbf{x} = (x^1, x^2, \ldots, x^n) \in \mathbf{R}^n : 0 < x^i < \alpha,\ i = 1, 2, \ldots, n\}$ and an arbitrary $(x_0, u_0, \xi_0) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm}$. We then translate and shrink D such that one corner lies at x_0 and the cube is still contained in Ω:

$$Q_h = x_0 + \frac{1}{h} D = \{\mathbf{x} \in \mathbb{R}^n : x_0^i < x^i < x_0^i + \frac{\alpha}{h},\ i = 1, \ldots, n\} \tag{III.30}$$

where h is a sufficiently large integer. Now suppose $\varphi \in W_0^{1,\infty}(D; \mathbb{R}^m)$. Extend it by periodicity from D to $\mathbf{R}^n$ and define:

$$\varphi_{\nu h}(x) = \begin{cases} \frac{1}{\nu h}\varphi(\nu h(x - x_0)) & \text{if } x \in Q_h \\ 0 & \text{if } x \notin Q_h \end{cases} \tag{III.31}$$

For fixed h we have, by lemma (III.6), that $\varphi_{\nu h} \rightharpoonup^* 0$ in $W^{1,\infty}(\Omega, \mathbb{R}^m)$ and that $\varphi_{\nu h} \rightharpoonup^* 0$ in $W^{1,\infty}(Q_h, \mathbb{R}^m)$. Now define $\bar{u} = u_0 + \xi_0(x - x_0)$ and the sequence $\{u_\nu\}$ by:

$$u_\nu(x) = \bar{u}(x) + \varphi_{\nu h}(x) \tag{III.32}$$

We now split Q_h into ν^n cubes, $Q^\nu_{h,j}$, of side length $\alpha/\nu h$ and denote by x_j the corner of $Q^\nu_{h,j}$ closest to x_0. So

$$Q_h = \bigcup_{j=0}^{\nu^n - 1} Q^\nu_{h,j} = \bigcup_{j=0}^{\nu^n - 1} (x_j + \frac{1}{\nu h} D) \tag{III.33}$$

Now

$$I(u_\nu) = \int_\Omega f(x, u_\nu(x), \nabla u_\nu(x)) dx \tag{III.34}$$

$$= \int_{\Omega - Q_h} f(x, \bar{u}(x), \nabla \bar{u}(x)) dx + \int_{Q_h} f(x, u_\nu(x), \nabla u_\nu(x)) dx \tag{III.35}$$

$$= \int_{\Omega - Q_h} f(x, \bar{u}(x), \nabla \bar{u}(x)) dx + \sum_{j=0}^{v^n - 1} \int_{Q^\nu_{h,j}} f(x, u_\nu(x), \nabla u_\nu(x)) dx \tag{III.36}$$

$$= \int_{\Omega - Q_h} f(x, \bar{u}(x), \nabla \bar{u}(x)) dx + \sum_{j=0}^{v^n - 1} \int_{Q^\nu_{h,j}} f(x_j, \bar{u}(x_j), \nabla u_\nu(x)) dx \tag{III.37}$$

$$+ \sum_{j=0}^{\nu^{n-1}} \int_{Q^\nu_{h,j}} \left[f(x, u_\nu(x), \nabla u_\nu(x)) - f(x_j, \bar{u}(x_j), \nabla u_\nu(x)) \right] dx \tag{III.38}$$

$$\tag{III.39}$$

Let us examine the third term, which we may rewrite as

$$\int_{Q_h} F_\nu(x) dx \tag{III.40}$$

$$F_\nu(x) = \sum_{j=0}^{\nu^{n-1}} \left[f(x, u_\nu(x), \nabla u_\nu(x)) - f(x_j, \bar{u}(x_j), \nabla u_\nu(x)) \right] \chi_{Q^\nu_{h,j}}(x) \tag{III.41}$$

$$\tag{III.42}$$

Where $\chi_{Q^\nu_{h,j}}$ is the indicator function, defined in the usual manner:

$$\chi_{Q^\nu_{h,j}}(x) = \begin{cases} 1 & x \text{ in } Q^\nu_{h,j} \\ 0 & \text{otherwise} \end{cases} \tag{III.43}$$

We shall first show that for all $x \in \Omega,\ F(x) \to 0$. Do to this we shall make use of the following facts:

$$u_\nu \rightharpoonup^* \bar{u} \text{ in } W^{1,\infty} \Rightarrow u_\nu \to \bar{u} \text{ in } L^\infty \tag{III.44}$$

$$u_\nu \in C(\bar{\Omega})\ \forall \nu \tag{III.45}$$

Both of which can be deduced from the Sobolev, Rellich and Koňdrasov embedding theorem. By (III.44) we know there exists an $M \in \mathbb{R}$ such that

$$||u_\nu||_{L^\infty} \leq M \;\forall\; \nu \tag{III.46}$$

and since $\nabla u_\nu \rightharpoonup^* \nabla u$ in $L^\infty(\Omega;\mathbb{R}^n)$ there exists a $K \in \mathbb{R}$ such that

$$||\nabla u_\nu||_{L^\infty} \leq K \;\forall\; \nu \tag{III.47}$$

Now consider the set

$$A = \Omega \times \{v \in \mathbb{R}^n \;:\; ||v||_{\mathbb{R}^n} \leq M\} \times \{\xi \in \mathbb{R}^{nm} \;:\; ||\xi||_{\mathbb{R}^{nm}} \leq K\} \subset \Omega \times \mathbb{R}^n \times \mathbb{R}^{nm} \tag{III.48}$$

By the Heine-Borel theorem and the Tychonoff compactness theorem we deduce that A is relatively compact. Since $f(\cdot,\cdot,\cdot)$ is continuous it follows that f is uniformly continuous on A. That is, for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $(x,u,\xi),\; (y,v,\eta) \;\in A$

$$||(x,u,\xi) - (y,v,\eta)|| \leq \delta \Rightarrow |f(x,u,\xi) - f(y,v,\eta)| \leq \epsilon \tag{III.49}$$

Now consider $(x, u_\nu(x), \nabla u_\nu(x)) \in A$ and $(x_j, \bar{u}(x_j), \nabla u_\nu(x)) \in A$. Then:

$$\begin{aligned} ||(x,u_\nu(x),\nabla u_\nu(x)) - (x_j,\bar{u}(x_j),\nabla u_\nu(x))|| &= ||x - x_j||_{\mathbb{R}^m} + ||u_\nu(x) - \bar{u}(x_j)||_{\mathbb{R}^n} && \text{(III.50)} \\ &\leq ||x - x_j|| + ||u_\nu(x) - u_\nu(x_j)|| + ||u_\nu(x_j) - \bar{u}(x_j)|| && \text{(III.51)} \\ & && \text{(III.52)} \end{aligned}$$

Since $||x - x_j|| \leq \frac{1}{\nu h}$ by choosing ν large enough we may make

$$||x - x_j|| \leq \frac{\delta}{3} \tag{III.53}$$

Also, by (III.45), choosing ν larger if necessary we may make

$$||u_\nu(x) - u_\nu(x_j)|| \leq \frac{\delta}{3} \tag{III.54}$$

Finally, by (III.44), by choosing ν larger still if necessary we have that

$$||u_\nu(x_j) - \bar{u}(x_j)|| \leq ||u_\nu - \bar{u}||_{L^\infty} \leq \frac{\delta}{3} \tag{III.55}$$

So we have that for every $x \in \Omega$, for ν sufficiently large

$$||(x,u_\nu(x),\nabla u_\nu(x) - (x_j,\bar{u}(x_j),\nabla u_\nu(x_j))|| \leq \delta \tag{III.56}$$

and so by (III.49)

$$|f(x,u_\nu(x),\nabla u_\nu(x) - f(x_j,\bar{u}(x_j),\nabla u_\nu(x_j))| \leq \epsilon \tag{III.57}$$

Now since for a fixed x all $\chi_{Q^\nu_{h,j}}$ in the sum in (III.41) bar one will equal zero:

$$F_\nu(x) \leq \epsilon \tag{III.58}$$

thus $F_\nu \to 0$ pointwise as required. Finally, since for all $x \in Q_h$

$$|F_\nu(x)| = |\sum_{j=0}^{\nu^{n-1}} [f(x,u_\nu(x),\nabla u_\nu(x)) - f(x_j,\bar{u}(x_j),\nabla u_\nu(x))]\, \chi_{Q^\nu_{h,j}}(x)| \leq 2a(x,M,K) \tag{III.59}$$

(where we have used the fact that a is increasing) and a is locally integrable with respect to x, we may apply the Lebesgue dominated convergence theorem to obtain:

$$\lim_{\nu\to\infty} \int_{Q_h} F_\nu dx = \int_{Q_h} \lim_{\nu\to\infty} F_\nu dx = 0 \tag{III.60}$$

Let us consider the second term in (III.38).

$$\sum_{j=0}^{v^n-1}\int_{Q^\nu_{h,j}} f(x_j,\bar{u}(x_j),\nabla\bar{u}(x))dx \tag{III.61}$$
$$=\sum_{j=0}^{v^n-1}\int_{x_j+\frac{1}{\nu h}D} f(x_j,u_0+\xi_0(x_j-x_0),\xi_0+\nabla\varphi(\nu h(x-x_0)))dx \tag{III.62}$$
$$=\sum_{j=0}^{v^n-1}\frac{1}{(\nu h)^n}\int_D f(x_j,u_0(x_j-x_0),\xi_0+\nabla\varphi(y+\nu h(x_j-x_0))) \tag{III.63}$$
$$=\sum_{j=0}^{v^n-1}\frac{1}{(\nu h)^n}\int_D f(x_j,u_0(x_j-x_0),\xi_0+\nabla\varphi(y))dy \tag{III.64}$$
$$=\frac{1}{\mu(D)}\sum_{j=0}^{v^n-1}\int_D f(x_j,u_0(x_j-x_0),\xi_0+\nabla\varphi(y))dy\frac{\mu(D)}{(\nu h)^n} \tag{III.65}$$
$$\tag{III.66}$$

where in (III.63) we have made the substitution $y=\nu h(x-x_j)$ and recalled that $dx=dx^1dx^2\dots dx^n$ so

$$dy=\frac{1}{(\nu h)^n}dx \tag{III.67}$$

and in (III.64) we have used the periodicity of $\varphi(x)$ to deduce that

$$\nabla\varphi(y+\nu h(x_j-x_0))=\nabla\varphi(y) \tag{III.68}$$

Finally, let us take the limit $\nu\to\infty$ in (III.65):

$$\lim_{\nu\to\infty}\frac{1}{\mu(D)}\sum_{j=0}^{v^n-1}\int_D f(x_j,u_0(x_j-x_0),\xi_0+\nabla\varphi(y))dy\frac{\mu(D)}{(\nu h)^n} \tag{III.69}$$
$$=\lim_{\nu\to\infty}\frac{1}{\mu(D)}\sum_{j=0}^{v^n-1}\int_D f(x_j,u_0(x_j-x_0),\xi_0+\nabla\varphi(y))dy\mu(Q^\nu_{h,j}) \tag{III.70}$$
$$=\frac{1}{\mu(D)}\int_{Q_h}\int_D f(x_j,u_0(x-x_0),\xi_0+\nabla\varphi(y))dydx \tag{III.71}$$
$$\tag{III.72}$$

Putting all the above together and using the weak $*$ lower semicontinuity of I we have

$$\liminf_{\nu\to\infty} I(u_\nu)=\int_{\Omega-Q_h} f(x,\bar{u}(x),\nabla\bar{u}(x))dx+\frac{1}{\mu(D)}\int_{Q_h}\int_D f(x_j,u_0(x-x_0),\xi_0+\nabla\phi(y))dydx \tag{III.73}$$
$$\geq I(\bar{u})=\int_\Omega f(x,\bar{u}(x),\nabla\bar{u}(x))dx \tag{III.74}$$
$$\tag{III.75}$$

And so:

$$\frac{1}{\mu(Q_h)}\int_{Q_h}\int_D f(x,u_0+\xi_0(x-x_0),\xi_0+\nabla\phi(y))dydx\geq\frac{\mu(D)}{\mu(Q_h)}\int_{Q_h} f(x,u_0+\xi_0(x-x_0),\xi_0)dx \tag{III.76}$$

We recognise the term $\frac{1}{\mu(Q_h)}\int_{Q_h}\int_D f(x_0,u_0+\xi_0(x-x_0),\xi_0+\nabla\phi(y))dydx$ as the average value of the function $F(x)=\int_D f(x_0,u_0+\xi_0(x-x_0),\xi_0+\nabla\phi(y))dy$ on Q_h. Since f is continuous, F will be continuous and a moments reflection reveals that as $h\to\infty$ (and therefore $\mu(Q_h)\to 0$) the average value of $F(x)$ approaches $F(x_0)$, since by construction x_0 is contained in Q_h $\forall h$, thus

$$\lim_{h\to\infty}\frac{1}{\mu(Q_h)}\int_{Q_h}\int_D f(x,u_0+\xi_0(x-x_0),\xi_0+\nabla\phi(y))dydx=\int_D f(x_0,u_0,\xi_0+\nabla\phi(y))dy \tag{III.77}$$

A similar argument yields that

$$\lim_{h\to\infty} \frac{\mu(D)}{\mu(Q_h)} \int_{Q_h} f(x, u_0 + \xi_0(x - x_0), \xi_0)dx = f(x_0, u_0, \xi_0) \tag{III.78}$$

So taking the limit as $h \to \infty$ on both sides of (III.76) we obtain:

$$\int_D f(x_0, u_0, \xi_0 + \nabla\phi(y))dy \geq f(x_0, u_0, \xi_0) \tag{III.79}$$

as required. □

B. The sufficiency of convexity

Before we state the promised theorem, we first define a class of functions, the *Carathéodory* functions, which allows us to state the result slightly more generally.

Definition 10. Suppose $\Omega \subset \mathbb{R}^n$ is open and

$$f : \Omega \times \mathbb{R}^m \times \mathbb{R}^N \to \bar{\mathbb{R}} \tag{III.80}$$

f is a Carathéodory function if

1. $f(x, \cdot, \cdot)$ is continuous for almost every $x \in \Omega$

2. $f(\cdot, u, \xi)$ is measurable in x for every $(u, \xi) \in \mathbb{R}^m \times \mathbb{R}^N$

A theorem on Carathéodory functions which we shall use, but not prove is the *Scorza-Dragoni* theorem.

Theorem III.7 (The Scorza-Dragoni theorem). *Suppose*

$$f : \Omega \times \mathbb{R}^m \times \mathbb{R}^N \to \bar{\mathbb{R}} \tag{III.81}$$

Then the following two conditions are equivalent:

1. *f is a Carathéodory function*

2. *for every compact set $K \subset \Omega$ and every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset K$ such that $\mu(K \setminus K_\epsilon) \leq \epsilon$ and f restricted to $K_\epsilon \times \mathbb{R}^m \times \mathbb{R}^N \to \bar{\mathbb{R}}$ is continuous.*

Proof. A proof can be found in (Ekeland and Temam, 1972) □

We shall also require the following lemma

Lemma III.8. Suppose that X is a Banach space and $I : X \to \mathbb{R}$ is convex and lower semicontinuous. Then I is weak lower semicontinuous.

Proof. Suppose $u_\nu \rightharpoonup u$ in X. We aim to show that

$$L = \liminf_{\nu\to\infty} I(u_\nu) \geq I(u) \tag{III.82}$$

Passing to a subsequence if necessary we may assume that

$$L = \lim_{\nu\to\infty} I(u_\nu) \tag{III.83}$$

we can disregard the degenerate cases $L = +\infty$ (as then the result would be trivially true) and $L = -\infty$ because I being convex and lower semicontinuous for arbitary $z_0 \in X$ there exists a $z_0^* \in X^*$ such that

$$I(u_\nu) \geq \langle u_\nu - z_0; z_0^* \rangle \tag{III.84}$$

$$\Rightarrow I(u_\nu) \geq \langle u_\nu; z_0^* \rangle + \alpha = ||z_0^*||\langle u_\nu; \frac{z_0^*}{||z_0^*||}\rangle + \alpha \tag{III.85}$$

$$\tag{III.86}$$

where $\alpha \in \mathbb{R}$ is independent of ν. Now since $|\langle u_\nu; \frac{z_0^*}{||z_0^*||}\rangle| \leq ||u_\nu||$ we have

$$\langle u_\nu; \frac{z_0^*}{||z_0^*||}\rangle \geq -||u_\nu|| \geq -K \tag{III.87}$$

where K is independent of ν (c.f. theorem (II.2)). Hence $I(u_\nu)$ is bounded from below and so $L > -\infty$. Now for all $\epsilon > 0$, choose $N = N(\epsilon)$ sufficiently large such that

$$I(u_\nu) \leq L + \epsilon \quad \forall \nu \geq N \tag{III.88}$$

We now consider the subsequence $\{u_\nu\}_{\nu=N}^{\infty}$ and for all $k \in \mathbb{N}$ we apply Mazur's lemma to obtain an n_k, α_i^k with $N \leq i \leq n_k$ such that

$$\sum_{i=1}^{n_k} \alpha_{i=1}^k = 1 \quad \text{and} \tag{III.89}$$

$$||u - \sum_{i=N}^{n_k} \alpha_i^k u_i|| \leq \epsilon/k \tag{III.90}$$

(III.91)

Denoting $v_k = \sum_{i=N}^{n_k} \alpha_i^k u_i$ we have a sequence of convex combinations of u_ν that *strong* converge to u. We may now use the lower semicontinuity of I to obtain

$$I(u) \leq \liminf_{k\to\infty} I(v_k) \tag{III.92}$$

and we may certainly find an k^* such that $\liminf_{k\to\infty} I(v_k) \leq I(v_k^*)$ thus:

$$\begin{aligned} I(u) &\leq I(v_{k^*}) && \text{(III.93)}\\ &= I(\sum_{i=N}^{k^*} \alpha_i^{k^*}) && \text{(III.94)}\\ &\leq \sum_{i=N}^{k^*} \alpha_I^{k^*} I(u_i) \quad \text{(by the convexity of } I) && \text{(III.95)}\\ &\leq \max\{I(u_i) : N \leq i \leq k^*\} && \text{(III.96)}\\ &\leq L + \epsilon && \text{(III.97)}\\ & && \text{(III.98)} \end{aligned}$$

Since ϵ was arbitrary, we conclude that $I(u) \leq L$, as required. □

Theorem III.9. *Suppose $\Omega \subset \mathbb{R}^n$ is bounded and open, and that $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^N \to \bar{\mathbb{R}}$ is Carathéodory and for almost every $x \in \Omega$ there exists $a(x) \in L^{q'}(\Omega;\mathbb{R}^N)$ (where q' is the conjugate index of q, to be defined below), and $b(x) \in L^1(\Omega;\mathbb{R})$ such that:*

$$f(x, u, \xi) \geq \langle a(x); \xi\rangle + b(x) \quad \forall (u, \xi) \in \mathbb{R}^m \times \mathbb{R}^N \tag{III.99}$$

Define J as

$$J(u, \xi) = \int_\Omega f(x, u(x), \xi(x))dx \tag{III.100}$$

If f is in addition convex, then if

$$\begin{cases} u_\nu \to \bar{u} & \text{in } L^p(\Omega;\mathbb{R}^m) \;\; p \geq 1 \\ \xi_\nu \rightharpoonup \bar{\xi} & \text{in } L^q(\Omega;\mathbb{R}^N) \;\; q \geq 1 \end{cases} \tag{III.101}$$

then $\liminf_{\nu\to\infty} J(u_\nu, \xi_\nu) \geq J(\bar{u}, \bar{\xi})$.

Remark III.10. 1. Let us set $\xi = \nabla u$ and $N = nm$ and consider a sequence $u_\nu \rightharpoonup u$ in $W^{1,r}(\Omega;\mathbb{R}^m)$, then as a consequence of the Rellich-Kondrašov embedding theorem we have

$$\begin{cases} u_\nu \to u \text{ in } L^p(\Omega;\mathbb{R}^m) & \text{for every } 1 \le p < \frac{nr}{n-r} \text{ if } \ 1 \le r < n \\ u_\nu \to u \text{ in } L^p(\Omega;\mathbb{R}^m) & \text{for every } 1 \le p \le \infty \text{ if } r = n \\ u_\nu \to u & \text{in } C(\bar{\Omega}) \text{ if } r > n \end{cases} \tag{III.102}$$

so by choosing p appropriately, and setting $q = r$ we have that

$$\liminf_{\nu\to\infty} I(u_\nu) = \liminf_{\nu\to\infty} J(u_\nu, \nabla u_\nu) \ge J(u, \nabla u) = I(u) \tag{III.103}$$

Hence the convexity of f is sufficient for the weak lower semicontinuity of I.

2. Stating the above theorem for functions which are Carathéodory as opposed to continuous is not an unnecessary generalization. Indeed there are many cases in elasticity where one considers a body made up of various smaller bodies glued together. In this case the stored energy function f need not be continuous at the boundaries of the smaller boundaries. Hence f will be Carathéodory, but not continuous.

Proof. We may assume, without loss of generality, that $f \ge 0$, since if this is not the case we may replace f with $\hat{f}$:

$$\hat{f}(x,u,\xi) = f(x,u,\xi) - \langle a(x);\xi\rangle - b(x) \ge 0 \tag{III.104}$$

and J with $\hat{J}$:

$$\hat{J}(u,\xi) = \int_\Omega \hat{f}(x,u,\xi)dx \tag{III.105}$$

$$= \int_\Omega \left(f(x,u,\xi) - \langle a(x);\xi\rangle - b(x)\right)dx \tag{III.106}$$

and since $\xi_\nu \rightharpoonup \xi \Rightarrow \langle a(x),\xi_\nu\rangle \to \langle a(x),\xi\rangle$ if

$$\liminf_{\nu\to\infty} \hat{J}(u_\nu,\xi_\nu) \ge \hat{J}(u,\xi) \tag{III.107}$$

it follows that

$$\liminf_{\nu\to\infty} J(u_\nu,\xi_\nu) \ge J(u,\xi) \tag{III.108}$$

Now denoting $L = \liminf_{\nu\to\infty} J(u_\nu,\xi_\nu)$ we note that $L \ge 0$ since $f \ge 0$, and we may safely assume $L < \infty$ (or else the theorem follows trivially). Passing to a subsequence if necessary, we may assume that $L = \lim_{\nu\to\infty} J(u_\nu,\xi_\nu)$.
We shall show later that for any fixed $\epsilon > 0$, there exists a measurable set $\Omega_\epsilon \subset \Omega$, a $\nu_\epsilon \in \mathbb{N}$ and subsequences $\{u_{\nu_j}\}$ and $\{\xi_{\nu_j}\}$ such that for every $\nu_j \ge \nu_\epsilon$

$$\mu(\Omega - \Omega_\epsilon) \le \epsilon \tag{III.109}$$

$$\int_{\Omega_\epsilon} \left|f(x,u_{\nu_j}(x),\xi_{\nu_j}(x)) - f(x,u(x),\xi_{\nu_j}(x))\right| dx < \epsilon\mu(\Omega) \tag{III.110}$$

but for now we shall assume the existence of such Ω_ϵ and ν_ϵ.
Define $\chi_\epsilon(x)$ as follows:

$$\chi_\epsilon(x) = \begin{cases} 1 & x \in \Omega_\epsilon \\ 0 & x \in \Omega - \Omega_\epsilon \end{cases} \tag{III.111}$$

and define g and G as

$$g(x,\eta) = \chi_\epsilon(x) f(x,u(x),\eta(x)) \tag{III.112}$$

$$G(\eta) = \int_\Omega g(x,\eta(x))dx \tag{III.113}$$

then g is a Carathéodory function (since f is a Carathéodory function and $\chi_\epsilon(x)$ is measurable and continuous for almost all $x \in \Omega$) and g is convex for almost all $x \in \Omega$. It follows that G is convex and lower semicontinuous over $L^q(\Omega;\mathbb{R}^N)$ and so by lemma (III.B) it is weak lower semicontinuous over $L^q(\Omega;\mathbb{R}^N)$, i.e.

$$\liminf_{\nu_j\to\infty} G(\xi_{\nu_j}) = \int_\Omega \chi_\epsilon(x) f(x,u(x),\xi_{\nu_j}(x)dx \tag{III.114}$$
$$\geq G(\xi) \tag{III.115}$$
$$= \int_\Omega \chi_\epsilon(x) f(x,u(x),\xi(x))dx \tag{III.116}$$

Now

$$\int_\Omega \chi_\epsilon f(x,u_{\nu_j}(x),\xi_{\nu_j})dx = \int_\Omega \chi_\epsilon f(x,u(x),\xi_{\nu_j})dx - \int_\Omega \chi_\epsilon \left(f(x,u(x),\xi_{\nu_j}) - f(x,u_{\nu_j}(x),\xi_{\nu_j})\right) \tag{III.117}$$
$$\geq \int_\Omega \chi_\epsilon f(x,u(x),\xi_{\nu_j})dx - \int_\Omega \chi_\epsilon \left|f(x,u(x),\xi_{\nu_j}) - f(x,u_{\nu_j}(x),\xi_{\nu_j})\right| \tag{III.118}$$
$$\geq \int_\Omega \chi_\epsilon f(x,u(x),\xi_{\nu_j})dx - \epsilon\mu(\Omega) \text{ for } \nu_j \geq \nu_\epsilon \tag{III.119}$$

and since $f \geq 0$ we have that

$$\int_\Omega \chi_\epsilon f dx = \int_{\Omega_\epsilon} f dx \leq \int_\Omega f dx \tag{III.120}$$

So using the above and (III.119) we obtain

$$\int_\Omega f(x,u_{\nu_j}(x),\xi_{\nu_j}(x)dx \geq \int_\Omega \chi_\epsilon(x) f(x,u(x),\xi_{\nu_j}(x)dx - \epsilon\mu(\Omega) \tag{III.121}$$

and taking the limit inferior on both sides and applying (III.116):

$$L = \liminf_{\nu\to\infty} \int_\Omega f(x,u_{\nu_j}(x),\xi_{\nu_j}(x)dx \geq \int_\Omega \chi_\epsilon(x) f(x,u(x),\xi(x))dx - \epsilon\mu(\Omega) \tag{III.122}$$

where $\epsilon > 0$ is arbitrary but fixed. Now if we set $\epsilon = 1/n$ and define

$$h_n(x) = \chi_{1/n}(x) f(x,u(x),\xi(x)) \tag{III.123}$$

then $\{h_n(x) : n \in \mathbb{N}\}$ is a sequence of measurable functions dominated by a measurable function f. Hence in sending $n \to \infty$, we may apply Lebesgue's dominated convergence theorem:

$$L \geq \lim_{n\to\infty} \left(\int_\Omega h_n(x)dx - \frac{1}{n}\mu(\Omega) \right) \tag{III.124}$$
$$= \int_\Omega \lim_{n\to\infty} h_n(x)dx \tag{III.125}$$
$$= \int_\Omega f(x,u(x),\xi(x))dx \tag{III.126}$$

It remains to show that we can find a set Ω_ϵ and an integer ν_ϵ as alluded to in (III.109) and (III.110). So, fix $\epsilon > 0$. Since

$$u_\nu \to u \text{ in } L^p(\Omega) \Rightarrow ||u_\nu||_{L^q} \leq K \text{ for some } K \text{ independent of } \nu \tag{III.127}$$
$$\xi_\nu \rightharpoonup \xi \text{ in } L^q(\Omega) \Rightarrow ||\xi_\nu||_{L^p} \leq K^{'} \text{ for some } K^{'} \text{ independent of } \nu \tag{III.128}$$

there exists an $M_\epsilon > 0$ such that if

$$K^1_{\epsilon,\nu} = \{x \in \Omega : |u(x)|, |u_\nu(x)| > M_\epsilon\} \tag{III.129}$$
$$K^2_{\epsilon,\nu} = \{x \in \Omega : |\xi_\nu(x)| > M_\epsilon\} \tag{III.130}$$

then $\mu(K^1_{\epsilon,\nu}), \mu(K^2_{\epsilon,\nu}) < \epsilon/6\ \forall \nu$ and hence if

$$\Omega^1_{\epsilon,\nu} = \Omega - (K^1_{\epsilon,\nu} \cup K^2_{\epsilon,\nu}) \tag{III.131}$$

then

$$\mu(\Omega - \Omega^1_{\epsilon,\nu}) < \epsilon/3 \tag{III.132}$$

Now, by assumption f is Carathéodory. So, by the Scorza-Dragoni theorem (cf. (III.B)) we obtain a compact set $\Omega^2_{\epsilon,\nu} \subset \Omega^1_{\epsilon,\nu}$ such that

$$\mu(\Omega^1_{\epsilon,\nu} - \Omega^2_{\epsilon,\nu}) < \epsilon/3 \tag{III.133}$$

and f restricted to $\Omega^2_{\epsilon,\nu} \times \mathbb{R}^m \times \mathbb{R}^N$ is continuous. If we restrict f further to the compact set

$$A = \Omega^2_{\epsilon,\nu} \times \{u \in \mathbb{R}^m : |u| \leq M_\epsilon\} \times \{\xi \in \mathbb{R}^N : |\xi| \leq M_\epsilon\} \tag{III.134}$$

then f is uniformly continuous on A thus we may find a $\delta(\epsilon) > 0$ such that

$$\begin{gathered} |u - v| \leq \delta(\epsilon) \Rightarrow |f(x, u, \xi) - f(x, v, \xi)| < \epsilon \\ \forall\ x \in \Omega^2_{\epsilon,\nu},\ \ |v|, |u|, |\xi| \leq M_\epsilon \end{gathered} \tag{III.135}$$

Using the fact that $u_\nu \to u$ in $L^p(\Omega)$ we may find $\nu_\epsilon = \nu_{\delta(\epsilon)}$ such that if $\Omega^3_{\epsilon,\nu} = \{x \in \Omega : |u_\nu(x) - u(x)| < \delta(\epsilon)\}$ then

$$\mu(\Omega - \Omega^3_{\epsilon,\nu}) < \epsilon/3 \quad \forall \nu \geq \nu_\epsilon \tag{III.136}$$

Let $\Omega_{\epsilon,\nu} = \Omega^2_{\epsilon,\nu} \cap \Omega^3_{\epsilon,\nu}$. Then by (III.132), (III.133) and (III.136):

$$\mu(\Omega - \Omega_{\epsilon,\nu}) < \epsilon \quad \forall \nu \geq \nu_\epsilon \tag{III.137}$$

and by definition, for all $x \in \Omega_{\epsilon,\nu}$ $|u_\nu(x)|, |u(x)|, |\xi_\nu(x)| \leq M_\epsilon$ and $|u_\nu(x) - u(x)| \leq \delta(\epsilon)$. So, by (III.135)

$$\int_{\Omega_{\epsilon,\nu}} |f(x, u_\nu(x), \xi_\nu(x)) - f(x, u(x), \xi_\nu(x))| dx \leq \int_{\Omega_{\epsilon,\nu}} \epsilon dx \leq \epsilon\mu(\Omega)\ \ \forall\ \nu \geq \nu_\epsilon \tag{III.138}$$

Since (III.137) and (III.138) hold for arbitrary ϵ, they hold for $\epsilon_j = \epsilon/2^j$. That is, for all j there exists a set Ω_j and an integer ν_{ϵ_j} such that

$$\mu(\Omega - \Omega_j) < \epsilon/2^j \tag{III.139}$$

$$\int_{\Omega_j|} f(x, u_\nu(x), \xi_\nu(x) - f(x, u(x), \xi(x))| dx \leq \frac{\epsilon}{2^j}\mu(\Omega) \quad \forall\ \nu \geq \nu_{\epsilon_j} \tag{III.140}$$

So taking any subsequence $\{\nu_j\}$ which is increasing and such that $\nu_j \geq \nu_{\epsilon_j}$ for all j and setting

$$\Omega_\epsilon = \bigcap_{j=1}^{\infty} \Omega_j \tag{III.141}$$

we have that

$$\mu(\Omega - \Omega_\epsilon) \leq \sum_{j=1}^{\infty} \mu(\Omega - \Omega_j) \tag{III.142}$$

$$\leq \sum_{j=1}^{\infty} \epsilon/2^j = \epsilon \tag{III.143}$$

and for all $\nu_j \geq \nu_\epsilon$

$$\int_\Omega |f(x, u_{\nu_j}(x), \xi_{\nu_j}(x)) - f(x, u(x), \xi_{\nu_j}(x))| dx \leq \epsilon \tag{III.144}$$

as required.

□

C. The sufficiency of quasiconvexity

Let us now consider how we might obtain the weak lower semicontinuity of I using the assumption that f is quasiconvex, which in the vectorial case is strictly weaker than f convex. For clarity we consider the case where $f = f(\nabla u)$, the case for more general f being an extension of the proof for the special case (See (Dacorogna, 1989)). Before we do this however we need to define some growth conditions to place on f, as well as prove a lemma which we shall need in the proof.

Definition 11. Let $f : \mathbb{R}^{nm} \to \mathbb{R}$. f is said to satisfy growth condition

1. f satisfies growth condition (C_∞) if there exists $\eta : \mathbb{R} \to \mathbb{R}$ continuous and increasing, such that:

$$|f(A)| \leq \eta(|A|)\ \forall A \in \mathbb{R}^{nm} \tag{III.145}$$

2. f satisfies growth conditon C_p for $1 < p < \infty$ if

$$-\alpha(1 + |A|^q) \leq f(A) \leq \alpha(1 + |A|^p) \tag{III.146}$$

   where $\alpha \geq 0$ and $1 \leq q < p$

3. f satisfies growth condition C_1 if

$$|f(A)| \leq \alpha(1 + |A|)\ \forall A \in \mathbb{R}^{nm} \tag{III.147}$$

   where $\alpha \geq 0$.

Lemma III.11. Let $f : \mathbb{R}^N \to \mathbb{R}$ be convex in each variable and let:

$$|f(x)| \leq \alpha(1 + |x|^p)\ \forall x \in \mathbb{R}^N \tag{III.148}$$

where $\alpha \geq 0$ and $p \geq 1$. Then there exists $\beta \geq 0$ such that:

$$|f(x) - f(y)| \leq \beta(1 + |x|^{p-1} + |y|^{p-1})|x - y|\ \forall x, y \in \mathbb{R}^N \tag{III.149}$$

Proof. A proof may be found in (Dacorogna, 1989) □

We may now state the theorem.

Theorem III.12. *Suppose that $f : \mathbb{R}^{nm} \to \mathbb{R}$ is quasiconvex and satisfies growth condition (C_p). Let $\Omega \subset \mathbb{R}^n$ be open and bounded. If $I(u)$ is defined as*

$$I(u, \Omega) = \int_\Omega f(\nabla u(x))dx \tag{III.150}$$

then $I(u, \Omega)$ is s.w.l.s.c. over $W^{1,p}(\Omega, \mathbb{R}^m)$ (s.w.s.l.s.c. in the case where $p = \infty$)

Remark III.13. Note that there is an asymmetry in the definition of growth condition C_p for $1 < p < \infty$. We cannot in general allow $p = q$ and expect the above theorem to still hold. Indeed, in (Dacorogna, 1989), Dacorogna produces a counterexample (originally due to Ball and Murat) showing that even in the simplest possible case of $m = n = 2$ it is possible to find a bounded open set Ω and a quasiconvex function $f(\nabla u)$ satisfying a growth condition of the form:

$$-\alpha(1 + |A|^2) \leq f(A) \leq \alpha(1 + |A|^2) \tag{III.151}$$

such that $I(u) = \int_\Omega f(\nabla u)dx$ is not weakly lower semicontinuous.

Proof. We aim to show that

$$\liminf_{u_\nu \rightharpoonup u} I(u_\nu, \Omega) \geq I(u, \Omega) \tag{III.152}$$

for all sequences $\{u_\nu\}$ such that $u_\nu \rightharpoonup u$. Given such a sequence $\{u_\nu\}$ we start by approximating Ω as the union of a finite family of cubes $H_N = \bigcup_s D_s$ such that each cube has side length $1/N$ where N is an integer chosen to be large enough such that

$$\mu(\Omega - H_N) \leq \epsilon \tag{III.153}$$

Let us denote the average value of $\nabla u(x)$ over D_s as A_s:

$$A_s = \frac{1}{\mu(D_s)} \int_{D_s} \nabla u(x) dx \tag{III.154}$$

By choosing N larger if necessary we have that:

$$\sum_s \int_{D_s} |\nabla u(x) - A_s|^p dx < \epsilon \tag{III.155}$$

Let us now consider the difference:

$$I(u_\nu, \Omega) - I(u, \Omega) = \int_\Omega \left(f(\nabla u_\nu(x)) - f(\nabla u(x))\right) dx \tag{III.156}$$

$$= \int_{\Omega - H_N} \left(f(\nabla u_\nu(x)) - f(\nabla u(x))\right) dx + \sum_s \int_{D_s} \left(f(\nabla u + (\nabla u_\nu - \nabla u)) - f(A_s + (\nabla u_\nu - \nabla u)\right) dx \tag{III.157}$$

$$+ \sum_s \int_{D_s} \left(f(A_s + (\nabla u_\nu - \nabla u)) - f(A_s)\right) dx + \sum_s \int_{D_s} \left(f(A_s) - f(\nabla u)\right) dx \tag{III.158}$$

$$= J_1 + J_2 + J_3 + J_4 \tag{III.159}$$

$$\tag{III.160}$$

We need to show that $\liminf_{u_\nu \to u} I(u_\nu, \Omega) - I(u, \Omega) \geq 0$. To do this we shall show that we can make each of the terms J_i as small as we like. More precisely, we shall show that for all $\epsilon > 0$ we may choose N sufficiently large and using the weak convergence of u_ν to u to obtain that $J_i \geq -\epsilon/4$. J_2 shall prove to be the hardest term to estimate, and it is here that we use the quasiconvexity of f.

a. estimating J_1

1. the case where $p = +\infty$

$$J_1 = \int_{\Omega - H_N} \left[f(\nabla u_\nu - f(\nabla u)\right] dx \tag{III.161}$$

$$\geq - \int_{\Omega - H_N} |f(\nabla u_\nu| dx - \int_{\Omega - H_N} |f(\nabla u)| dx \tag{III.162}$$

$$\geq - \int_{\Omega - H_N} \eta(|\nabla u_\nu|) dx - \int_{\Omega - H_N} \eta(|\nabla u|) dx \text{ Since } f \text{ satisfies growth condition } (C_\infty) \tag{III.163}$$

$$\geq -\eta(\sup\{|\nabla u_\nu(x)| : x \in \Omega - H_N\} \mu(\Omega - H_N) - \eta(\sup\{|\nabla u(x)| x \in \Omega - H_N\} \tag{III.164}$$

$$\tag{III.165}$$

Where the final line follows from the fact that η is an increasing function. Finally, using the fact that $u_\nu \rightharpoonup u$ implies that there exists a $K \in \mathbb{R}$ such that

$$\sup\{|\nabla u_\nu(x)| : x \in \Omega - H_N\} \leq \sup\{|\nabla u_\nu(x)| : x \in \Omega\} = ||\nabla u||_{L_\infty} \leq K \tag{III.166}$$

we have that

$$J_1 \geq -\eta(K)\mu(\Omega - H_N) - \eta(||\nabla u||_{L_\infty})\mu(\Omega - H_N) \tag{III.167}$$

and by choosing N larger if necessary we have that $J_1 \geq -\epsilon/3$

2. the case where $1 < p < \infty$

$$J_1 = \int_{\Omega - H_N} [f(\nabla u_\nu) - f(\nabla u)]\, dx \tag{III.168}$$

$$\geq - \int_{\Omega - H_N} [\alpha(1 + |\nabla u_\nu|^q + \alpha(1 + |\nabla u|^p]\, dx \tag{III.169}$$

$$= -\alpha \int_{\Omega - H_N} [2 + |\nabla u|^p]\, dx - \alpha \int_{\Omega - H_N} |\nabla u_\nu|^q dx \tag{III.170}$$

$$\geq -2\alpha\mu(\Omega - H_N) - \int_{\Omega - H_N} |\nabla u|^p dx - \alpha \int_{\Omega - H_N} |\nabla u_\nu|^q dx \tag{III.171}$$

$$\tag{III.172}$$

Since $p < q$ $(\nabla u_\nu)^q \in L^{p/q}(\Omega - H_N)$ and $1 \in L^{p/(p-q)}(\Omega - H_N)$ we may use Hölder's inequality on the second term:

$$\int_{\Omega - H_N} |\nabla u_\nu|^q \leq (\int_{\Omega - H_N} dx)^{(p-q)/p} \left(\int_{\Omega - H_N} |\nabla u_\nu|^p \right)^{q/p} \tag{III.173}$$

$$= (\mu(\Omega - H_N))^{(p-q)/p} \left(\int_{\Omega - H_N} K^p \right)^{q/p} \tag{III.174}$$

$$\leq (\mu(\Omega - H_N))^{(p-q)/p}\, K^q\, (\mu(\Omega - H_N))^{q/p} \tag{III.175}$$

$$= \mu(\Omega - H_N) K^q \tag{III.176}$$

$$\tag{III.177}$$

Where K is the same constant as in (III.166). So for $\mu(\Omega - H_N)$ is sufficiently small we have that

$$J_1 \geq -2\alpha\mu(\Omega - H_N) - \int_{\Omega - H_N} |\nabla u|^p dx - (\mu(\Omega - H_N))\, K^q \tag{III.178}$$

$$\geq -\epsilon/3 \tag{III.179}$$

The case where $p = 1$ may be done in a similar manner.

b. estimating J_2 Let us now place a lower bound on J_2.

$$J_2 = \sum_s \int_{D_s} [f(\nabla u_\nu) - f(A_s + (\nabla u_\nu - \nabla u))]\, dx \tag{III.180}$$

by considering $|J_2|$ and applying lemma (III.11) we obtain:

$$|J_2| \leq \beta \sum_s \int_{D_s} \left(1 + |\nabla u_\nu|^{p-1} + |\nabla u_\nu + A_s - \nabla u|^{p-1}\right) |\nabla u - A_s| dx \tag{III.181}$$

where $\beta \geq 0$ is a fixed constant. Again we use Hölder's inequality repeatedly. We demonstrate the treatment of the second term:

$$\int_{D_s} |\nabla u_\nu|^{p-1} |\nabla u - A_s| dx \leq \left(\int_{D_s} |(\nabla u_\nu)^{p-1}|^{p/(p-1)} \right)^{p-1/p} \left(\int_{D_s} |\nabla u - A_s|^p dx \right)^{1/p} \tag{III.182}$$

So

$$\sum_s \int_{D_s} |\nabla u_\nu|^{p-1} |\nabla u - A_s| dx \leq \sum_s \left(\int_{D_s} |(\nabla u_\nu)^{p-1}|^{p/(p-1)} \right)^{p-1/p} \left(\int_{D_s} |\nabla u - A_s|^p dx \right)^{1/p} \tag{III.183}$$

$$\leq ||\nabla u_\nu||^{p-1}_{L_p(\Omega)} \sum_s \left(\int_{D_s} |\nabla u - A_s|^p \right)^{1/p} \tag{III.184}$$

$$\leq K \sum_s \left(\int_{D_s} |\nabla u - A_s|^p \right)^{1/p} \tag{III.185}$$

$$\tag{III.186}$$

which by the observation in (III.155) we may make arbitrarily small. The other two terms in J_2 my be treated similarly and so we obtain

$$|J_2| \leq \epsilon/3 \tag{III.187}$$

c. estimating J_4

$$J_4 = \sum_s \int_{D_s} [f(A_s) - f(\nabla u)]\, dx \tag{III.188}$$

$$\leq \beta \sum_s \left[\left(\int_{D_s} |A_s|^p \right)^{(p-1)/p} + \left(\int_{D_s} |\nabla u|^p \right)^{(p-1)/p} \right] \left(\int_{D_s} |A_s - \nabla u|^p \right)^{1/p} \tag{III.189}$$

$$\tag{III.190}$$

Using the same combination of lemma (III.11) and Hölder's inequality as in estimating J_2. Again, using (III.155) we may make this arbitrarily small so:

$$|J_4| \leq \epsilon/3 \tag{III.191}$$

To recap, using the fact that $u_\nu \rightharpoonup u$ implies that $||u_\nu|| \leq K$ and by fixing N large enough in order to make

$$\mu(\Omega - H_N) \tag{III.192}$$

$$\text{and } \sum_s \int_{D_s} |\nabla u(x) - A_s|^p dx \tag{III.193}$$

$$\tag{III.194}$$

sufficiently small, we have established a lower bound on $J_1 + J_2 + J_4$ independent of ν:

$$I(u_\nu; \Omega) - I(u; \Omega) \geq -\epsilon + J_3 \tag{III.195}$$

where ϵ is arbitrary. Taking the limit inferior on both sides we have

$$\liminf_{\nu\to\infty} (I(u_\nu;\Omega) - I(u;\Omega)) \geq -\epsilon + \liminf_{\nu\to\infty} \sum_s \int_{D_s} [f(A_s + (\nabla u_\nu - \nabla u)) - f(A_s)] \tag{III.196}$$

$$= -\epsilon + \sum_s \liminf_{\nu\to\infty} \int_{D_s} [f(A_s + (\nabla u_\nu - \nabla u)) - f(A_s)] \tag{III.197}$$

$$\tag{III.198}$$

where we may interchange the order of $\liminf_{\nu\to\infty}$ and $\sum_s$ in (III.196) because the number of terms in the sum depends only on N (which we have fixed) and not on ν.
So, to prove the result, we need to show that

$$\sum_s \liminf_{\nu\to\infty} \int_{D_s} [f(A_s + (\nabla u_\nu - \nabla u)) - f(A_s)] \geq 0 \tag{III.199}$$

Since we are summing over a finite number of terms, it suffices to show that

$$\liminf_{\nu\to\infty} \int_{D_s} f(A_s + (\nabla u_\nu - \nabla u))dx \geq \int_{D_s} f(A_s)dx \tag{III.200}$$

For all s. So, relabeling $v_\nu = u_\nu - u$ and considering a generic cube $D \subset \mathbb{R}^n$ and arbitrary $A \in \mathbb{R}^{nm}$ we shall show that

$$\liminf_{\nu\to\infty} \int_D f(A + \nabla v_\nu)dx \geq \int_D f(A)dx \tag{III.201}$$

where $v_\nu \rightharpoonup 0$ in $W^{1,p}(D;\mathbb{R}^m)$. Note that if it were the case that $v_\nu \rightharpoonup 0$ in $W_0^{1,p}(D;\mathbb{R}^m)$ then (III.201) would follow trivially from the quasiconvexity of f. So, informally speaking, our approach to obtaining (III.201) will be to 'modify'

v_ν to a function v_ν^k that does vanish on the boundary of D, and then show that in the limit $\nu \to \infty$ this modification makes no difference. Let us now show this formally.
Consider a cube $D^0 \subset D$ such that the containment is strict, and let

$$R = \frac{1}{2}\operatorname{dist}(D^0, \delta D) = \inf\{d(x,y) : x \in D^0, y \in \delta D\} \tag{III.202}$$

We of course have $R > 0$ as D^0 is compact and δD is closed. Next we create a finite increasing sequence of cubes $D^1, \ldots, D^K$:

$$D^0 \subset D^1 \subset \ldots \subset D^K \subset D \tag{III.203}$$

such that all of the above inclusions are strict, and

$$\operatorname{dist}(D^0, \delta D^k) = \frac{k}{K}R,\ 1 \le k \le K \tag{III.204}$$

For all $k:\ 1 \le k \le K$ we then choose $\varphi^k \in \mathcal{C}^\infty(D)$ such that

$$0 \le \varphi^k(x) \le 1 \quad \forall x \in D \tag{III.205}$$

$$\varphi^k(x) = \begin{cases} 1 & \text{if } x \in D^{k-1} \\ 0 & \text{if } x \in D - D^k \end{cases} \tag{III.206}$$

$$\tag{III.207}$$

and $|\nabla \varphi^k| \le a\frac{K}{R}$. Such functions can always be found, see for example (Lee, 2003). Denote $v_\nu^k = \varphi^k v_\nu$, then $v_\nu^k \in W_0^{1,p}(D;\mathbb{R}^m), \forall k, \nu$ and so we may use the quasiconvexity of f.

$$\int_D f(A)dx \le \int_D f(A + \nabla v_\nu^k(x))dx \tag{III.208}$$

$$= \int_{D-D^k} f(A)dx + \int_{D^k - D^{k-1}} f(A + \nabla v_\nu^k(x))dx + \int_{D^{k-1}} f(A + \nabla v_\nu(x))dx \tag{III.209}$$

$$\tag{III.210}$$

We rewrite this as

$$\int_{D^k} f(A)dx \le \int_D f(A + \nabla v_\nu(x))dx - \int_{D-D^{k-1}} f(A + \nabla v_\nu(x))dx + \int_{D^k-D^{k-1}} f(A + \nabla v_\nu^k))dx \tag{III.211}$$

$$= \int_D f(A + \nabla v_\nu(x))dx + \alpha_1 + \alpha_2 \tag{III.212}$$

$$\tag{III.213}$$

or

$$\mu(D^k)f(A) \le \int_D f(A + \nabla v_\nu(x))dx + \alpha_1 + \alpha_2 \tag{III.214}$$

we now show that:

1. For all $\epsilon > 0$, choosing R sufficiently large (that is, choosing D^0 sufficiently 'close' to δD) we can ensure that $\alpha_1 \le \epsilon$

2. α_2 is also bounded:

$$\alpha_2 \le \alpha \int_{D^k - D^{k-1}} \left(1 + |A|^p + |\nabla u_\nu|^p + \left(\frac{aK}{R}\right)^p |v_\nu|^p\right) dx \tag{III.215}$$

d. Placing an upper bound on α_1

1. The case where $p = \infty$ If $p = \infty$ we have that $||\nabla u_\nu||_{L^\infty} \leq M$ for all ν. So

$$\alpha_1 \leq \int_{D-D^{k-1}} |f(A) + \nabla u_\nu(x)| dx \tag{III.216}$$
$$\leq \int_{D-D^{k-1}} \eta(|A + \nabla v_\nu|) dx \tag{III.217}$$
$$\leq \int_{D-D^{k-1}} \eta(|A| + |\nabla v_\nu|) dx \tag{III.218}$$
$$\leq \int_{D-D^{k-1}} \eta(|A| + ||\nabla v_\nu||_{L^\infty}) dx \quad \text{since } \eta \text{ is increasing} \tag{III.219}$$
$$\leq \int_{D-D^0} \eta(|A| + M) dx \tag{III.220}$$
$$= \eta(|A| + M)\mu(D - D^0) \tag{III.221}$$
$$\tag{III.222}$$

Now if we choose R sufficiently small (i.e. D^0 sufficiently 'close' to D) we see that we can make $\mu(D - D^0)$ arbitrarily small. Hence

$$\alpha_1 \leq \epsilon \tag{III.223}$$

2. The case where $1 < p < \infty$ Using the growth condition on f we have:

$$\alpha_1 \leq \alpha \int_{D-D^{k-1}} (1 + |A + \nabla v_\nu|^q) \, dx \tag{III.224}$$

Now note that

$$|A + \nabla v_\nu|^q \leq (|A| + |\nabla v_\nu|)^q \tag{III.225}$$
$$\leq (2 \max(|A|, |\nabla v_\nu|))^q \tag{III.226}$$
$$\leq 2^q \max(|A|^q, |\nabla v_\nu|^q) \tag{III.227}$$
$$\leq 2^q (|A|^q + |\nabla v_\nu|^q) \tag{III.228}$$
$$\tag{III.229}$$

So

$$\alpha_1 \leq \alpha^{'} \int_{D-D^{k-1}} (|A|^q + |\nabla v_\nu|^q) \, dx \quad \alpha^{'} = 2^q \alpha \tag{III.230}$$
$$\leq \alpha^{'} \int_{D-D^0} (|A|^q + |\nabla v_\nu|^q) \, dx \tag{III.231}$$
$$\tag{III.232}$$

As in the estimation of J_1, we apply Hölder's inequality and use the fact that $\nabla v_\nu| \leq M$ for some $M \in \mathbb{R}$ to obtain

$$\alpha_1 \leq \alpha^{'}(|A|^q + M^q)\mu(D - D^0) \tag{III.233}$$

where $\alpha^{'}$ is a fixed constant. As in the case where $p = \infty$, we may make $\mu(D - D^0)$ arbitrarily small, and so

$$\alpha_1 \leq \epsilon \tag{III.234}$$

3. The case where $p = 1$ For brevity we shall omit this case, but it can be handled in an analogous manner to the above case

e. Bounding α_2

$$\alpha_2 = \int_{D^k - D^{k-1}} f(A + \nabla \varphi_\nu^k) dx \tag{III.235}$$

for $1 \leq p < \infty$, using the growth condition on f we have:

$$\alpha_2 \leq \alpha \int_{D^k - D^{k-1}} \left(1 + |A + \nabla v_\nu^k(x)|^p\right) dx \tag{III.236}$$

$$\leq \alpha' \int_{D^k - D^{k-1}} \left(1 + |A|^p + |\varphi^k \nabla v_\nu + v_\nu \nabla \varphi^k|^p\right) dx \text{ Using identical reasoning to (III.229) and the definition of } v_\nu^k \tag{III.237}$$

$$\leq \alpha'' \int_{D^k - D^{k-1}} \left(1 + |A|^p + |\nabla v_\nu|^p + |\nabla v_\nu|^p + \left(\frac{aK}{R}\right)^p |v_\nu|^p\right) dx \text{ Again using (III.229) and the definition of } \varphi^k \tag{III.238}$$

$$\tag{III.239}$$

So, returning to (III.212), we have:

$$\int_{D^k} f(A)dx \leq \int_D f(A + \nabla v_\nu)dx + \epsilon + \alpha'' \int_{D^k - D^{k-1}} \left(1 + |A|^p + |\nabla v_\nu|^p + |\nabla v_\nu|^p + \left(\frac{aK}{R}\right)^p |v_\nu|^p\right) dx \tag{III.240}$$

we now take the sum from $k = 1$ tp $k = K$ and using the fact that

$$\int_{D^1 - D^0} + \int_{D^2 - D^1} + \ldots + \int_{D^K - D^{K-1}} = \int_{D^K - D^0} \tag{III.241}$$

we obtain:

$$f(A) \left(\sum_{k=1}^{K} \mu(D^k)\right) \leq K \int_D f(A + \nabla v_\nu)dx + K\epsilon + \alpha'' \int_{D^K - D^0} \left(1 + |A|^p + |\nabla v_\nu|^p + |\nabla v_\nu|^p + \left(\frac{aK}{R}\right)^p |v_\nu|^p\right) dx \tag{III.242}$$

$$\leq K \int_D f(A + \nabla v_\nu)dx + K\epsilon + \gamma + \alpha'' \left(\frac{aK}{R}\right)^p \int_{D^K - D^0} |v_\nu|^p dx \tag{III.243}$$

$$\tag{III.244}$$

where γ is defined as

$$\gamma = \int_D (1 + |A|^p)dx + M \tag{III.245}$$

where M is such that $||\nabla u_\nu||_{L^p}^p \leq M$ for all ν. It can easily be seen that γ is a constant independent of R,K and ν. Now divide (III.243) by K and take the limit inferior as $\nu \to \infty$:

$$f(A)) \left(\frac{1}{K} \sum_{k=1}^{K} \mu(D^k)\right) \leq \epsilon + \gamma/K + \liminf_{\nu \to \infty} \int_D f(A + \nabla v_\nu)dx \tag{III.246}$$

where we have used the fact that if $v_\nu \rightharpoonup 0$ in $W^{1,p}(\Omega; \mathbb{R}^m)$ then as a consequence of the Rellich-Kondrasov embedding theorem $v_\nu \to 0$ in $L^p(\Omega; \mathbb{R}^m)$. Finally sending $K \to \infty$ and $R \to$) (i.e. allowing the boundary of D^0 to approach the boundary of D) we have

$$\frac{1}{K} \sum_{k=1}^{K} \mu(D^k) \to \mu(D) \tag{III.247}$$

and

$$f(A)\mu(D) \leq \epsilon + \liminf_{\nu \to \infty} \int_D f(A + \nabla v_\nu)dx \tag{III.248}$$

But $\epsilon > 0$ is arbitrary, hence we conclude that

$$f(A) \leq \liminf_{\nu \to \infty} \int_D f(A + \nabla v_\nu)dx \tag{III.249}$$

as required. $\square$

D. The sufficiency of Polyconvexity

We now consider the case where

$$f : \Omega \times \mathbb{R}^3 \times \mathbb{R}^9 \to \bar{\mathbb{R}} \tag{III.250}$$

is polyconvex. To simplify our analysis, we shall consider functions f that are independent of u

$$f(x, u, \nabla u) \equiv f(x, \nabla u) \tag{III.251}$$

Recall that if f is polyconvex there exists a convex function

$$F : \Omega \times \mathbb{R}^{19} \to \bar{\mathbb{R}} \tag{III.252}$$

such that

$$f(x, \nabla u) = F(x, \nabla u, \mathrm{adj}(\nabla u), \det(\nabla u)) \quad \forall x \in \Omega \tag{III.253}$$

In order to show that the polyconvexity of f is sufficient to ensure the weak lower semicontinuity of I, it suffices to show that if $u_\nu \rightharpoonup u$ in $W^{1,p}(\Omega; \mathbb{R}^3)$ then

$$(\nabla u_\nu, \mathrm{adj}(\nabla u_\nu), \det(\nabla u_\nu)) \rightharpoonup (\nabla u, \mathrm{adj}(\nabla u), \det(\nabla u)) \quad \text{in } L^q(\Omega, \mathbb{R}^{19}) \tag{III.254}$$

As we may then apply theorem (III.9) to obtain the desired result. Our task then is to determine sufficient conditions on $W^{1,p}(\Omega, \mathbb{R}^3)$ under which the maps

$$u \mapsto \mathrm{adj}(\nabla u) \tag{III.255}$$
$$u \mapsto \det(\nabla u) \tag{III.256}$$
$$\tag{III.257}$$

are (sequentially) weakly continuous. We shall follow Ciarlet's (Ciarlet, 1988) approach in this section, although the results are originally due to Ball (Ball, 1977). First however we need a lemma regarding the continuity of bilinear maps.

Lemma III.14. Suppose that V is a normed space, W is a Banach space and $B : V \times W \to \mathbb{R}$ is a continuous bilinear mapping. Then

$$v_k \to v \text{ and } w_k \rightharpoonup w \Rightarrow B(v_k, w_k) \to B(v, w) \tag{III.258}$$

Proof. A proof may be found in (Ciarlet, 1988) □

Now suppose $u : \mathbb{R}^3 \to \mathbb{R}^3$. Recall that we define the *adjugate* matrix of ∇u as:

$$(\mathrm{adj}(\nabla u))_{i\alpha} = \frac{1}{2}(-1)^{i+\alpha}\epsilon_{\alpha\beta\gamma}\epsilon_{ijk}\partial_j u^\beta \partial_k u^\gamma \tag{III.259}$$

where $\epsilon_{\alpha\beta\gamma}$ is the rank three Levi-Civita tensor. Also, by taking a Laplace expansion along the first row of ∇u, we may write $\det(\nabla u)$ neatly as follows:

$$\det(\nabla u) = \partial_i u^1 (\mathrm{adj}(\nabla u)_{1\alpha} \tag{III.260}$$

where summation over i is implied. For u sufficiently differentiable (i.e. $u \in C^\infty(\Omega)$) we may rewrite $\mathrm{adj}(\nabla u)$ and $\det(\nabla u)$ as follows:

$$(\mathrm{adj}(\nabla u))_{i\alpha} = \frac{1}{2}(-1)^{i+\alpha}\epsilon_{\alpha\beta\gamma}\epsilon_{ijk}\partial_j (u^\beta \partial_k u^\gamma) \tag{III.261}$$
$$\det(\nabla u) = \partial_i (u^1 (\mathrm{adj}(\nabla u)_{1\alpha}) \tag{III.262}$$
$$\tag{III.263}$$

Our first lemma shows that if we interpret the derivatives in the above formulae in a weak sense we may extend them to $u \in W^{1,p}(\Omega; \mathbb{R}^3)$.

Lemma III.15. 1. Suppose Ω is a bounded open set and $u \in W^{1,p}(\Omega;\mathbb{R}^3)$. For each $p \geq 2$ $\operatorname{adj}(\nabla u) \in L^{p/2}(\Omega;\mathbb{R}^9)$, the assignment $u \mapsto \operatorname{adj}(\nabla u)$ is continuous with respect to the norm $||\cdot||_{W^{1,p}(\Omega)}$ and (III.261) holds in $(\mathcal{D}^{'}(\Omega))^9$

2. Again, suppose Ω is a bounded open set and $u \in W^{1,p}(\Omega;\mathbb{R}^3)$. Additionally, suppose that $\operatorname{adj}(\nabla u) \in L^q(\Omega;\mathbb{R}^9)$ where $p^{-1} + q^{-1} \leq 1$ then if

$$s \equiv \frac{pq}{p+q} \tag{III.264}$$

$\det(\nabla u) \in L^s(\Omega)$ and (III.262) holds in $\mathcal{D}^{'}(\Omega)$. Note that if $p \geq 3$ we need place no further restrictions on $\operatorname{adj}(\nabla u)$ since by the first part of the lemma $\operatorname{adj}(\nabla u) \in L^{p/2}(\Omega;\mathbb{R}^9)$ and $p^{-1} + 2p^{-1} = 3/p \leq 1$ as required.

Proof. Suppose $u \in W^{1,p}(\Omega;\mathbb{R}^3)$ then $\partial_i u^\alpha \in L^p(\Omega)$ $\forall 1 \leq i \leq 3,\ 1 \leq \alpha \leq 3$, and by Hölder's inequality:

$$\int_\Omega |\partial_j u^\beta \partial_k u^\gamma|^{p/2} dx = \int_\Omega |(\partial_j u^\beta)^{p/2} (\partial_k u^\gamma)^{p/2}| dx \tag{III.265}$$

$$\leq ||\partial_j u^\beta||^p_{L^p} ||\partial_k u^\gamma||^p_{L^p} \tag{III.266}$$

we have that $\partial_j u^\beta \partial_k u^\gamma \in L^{p/2}(\Omega)$. To show the continuity of the map $u \mapsto \operatorname{adj}(\nabla u)$ it suffices to show that the map $u \mapsto \partial_j u^\beta \partial_k u^\gamma$ is continuous. So suppose that $u_\nu \to u$ in $W^{1,p}(\Omega,\mathbb{R}^3)$

$$\left| \int_\Omega \left(\partial_j u^\beta_\nu \partial_k u^\gamma_\nu - \partial_j u^\beta \partial_k u^\gamma\right) dx \right| \tag{III.267}$$

$$= \left| \int_\Omega \left(\partial_j u^\beta_\nu - \partial_j u^\beta\right) \partial_k u^\gamma_\nu + \partial_j u^\beta \left(\partial_k u^\gamma_\nu - \partial_k u^\gamma\right) s dx \right| \tag{III.268}$$

$$\leq \int_\Omega \left|\left(\partial_j u^\beta_\nu - \partial_j u^\beta\right) \partial_k u^\gamma_\nu\right| dx + \int_\Omega \left|\partial_j u^\beta \left(\partial_k u^\gamma_\nu - \partial_k u^\gamma\right)\right| dx \tag{III.269}$$

$$\leq ||\partial_k u^\gamma_\nu||_{L^{p'}} ||\partial_j u^\beta_\nu - \partial_j u^\beta||_{L^p} + ||\partial_j u^\beta||_{L^{p'}} ||\partial_k u^\gamma_\nu - \partial_k u^\gamma||_{L^p} \tag{III.270}$$

$$\leq (\alpha K + ||\partial_j u^\beta||_{L^{p'}}) ||u_\nu - u||_{W^{1,p}} (\text{cf. theorems (II.1) and (II.2)}) \tag{III.271}$$

hence $u \mapsto \partial_j u^\beta \partial_k u^\gamma$ is continuous as required. Now, as noted above, (III.261) holds for $u \in C^\infty(\Omega)$. So, suppose that in (III.261) $u \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$. Multiply both sides by a fixed $\varphi \in \mathcal{D}(\Omega)$, integrate over Ω and apply Green's formula to obtain:

$$\int_\Omega (\operatorname{adj}(\nabla u)_{i\alpha} \varphi dx = -\frac{(-1)^{i+\alpha}}{2} \epsilon_{\alpha\beta\gamma} \epsilon_{ijk} \int_\Omega u^\beta \partial_k \partial_j \varphi dx \tag{III.272}$$

This is well defined since $p \geq 2 \Rightarrow \operatorname{adj}(\nabla u)_{i\alpha}, u^\beta \in L^1(\Omega)$. Now since $C^\infty(\Omega) \cap W^{1,p}(\Omega)$ is dense in $W^{1,p}(\Omega)$ we may use the continuity of both sides of (III.272) to extend the formula to arbitrary $u \in W^{1,p}(\Omega)$ and since $\varphi \in \mathcal{D}(\Omega)$ was arbitrary we conclude that:

$$(\operatorname{adj}(\nabla u))_{i\alpha} = \frac{(-1)^{i+\alpha}}{2} \epsilon_{\alpha\beta\gamma} \epsilon_{ijk} \partial_j (u^\beta \partial_k u^\gamma) \tag{III.273}$$

in $\mathcal{D}^{'}(\Omega)$ for all $u \in W^{1,p}(\Omega;\mathbb{R}^3)$ with $p \geq 2$.

We have that $\det(\nabla u) = \partial_j u_1 (\operatorname{adj}(\nabla u))_{1j}$. If $\partial_j u_1 \in L^p(\Omega)$ and $\operatorname{adj}(\nabla u))_{1j} \in L^q(\Omega)$ (with q and s as described in the statement of the lemma) then by applying Hölder's inequality we have

$$||\partial_j u (\operatorname{adj}(\nabla u))_{1j}||_{L^s} \leq ||\partial_j u||_{L^p} || \operatorname{adj}(\nabla u))_{1j}||_{L^q} \tag{III.274}$$

Hence $\partial_j u (\operatorname{adj}(\nabla u))_{1j} \in L^s(\Omega)$ $s \geq 1$ and the map:

$$W^{1,p}(\Omega) \times L^p(\Omega) \to L^s(\Omega) \tag{III.275}$$

$$(u, \operatorname{adj}(\nabla u)) \mapsto \partial_j u (\operatorname{adj}(\nabla u))_{1j} \tag{III.276}$$

is well-defined and continuous. Now suppose $u \in C^\infty(\Omega)$, then

$$\partial_i (\mathrm{adj}(\nabla u))_{1i} = \frac{(-1)^{1+i}}{2} \epsilon_{1\beta\gamma}\epsilon_{ijk}\partial_i\partial_j u_\beta \partial_k u_\gamma = 0 \tag{III.277}$$

since ϵ_{ijk} is antisymmetric in i, j and $\partial_i\partial_j u_\beta \partial_k u_\gamma$ is symmetric in i, j. So

$$\det(\nabla u) = \partial_i u_1 (\mathrm{adj}(\nabla u))_{1i} = \partial_i \left[u_1(\mathrm{adj}(\nabla u)_{1i}\right] \quad \text{for } u \in C^\infty(\Omega) \tag{III.278}$$

Returning to (III.277) we have for $u \in C^\infty(\Omega)$, $\varphi \in \mathcal{D}(\Omega)$

$$0 = \int_\Omega \partial_i (\mathrm{adj}(\nabla u))_{1i}\varphi dx = -\int_\Omega (\mathrm{adj}(\nabla u))_{1i}\partial_i\varphi dx \tag{III.279}$$

We already have that the map $u \mapsto (\mathrm{adj}(\nabla u))_{1i}$ is $||\cdot||_{W^{1,p}}$ continuous, so again by the density of $C^\infty(\Omega) \cap W^{1,p}(\Omega)$ in $W^{1,p}(\Omega)$ we deduce that (III.279) holds for all $u \in W^{1,p}(\Omega)$.
Now for arbitrary $v \in W^{1,p}(\Omega)$, suppose $w \in L^{p'}(\Omega, \mathbb{R}^n)$ has the property that

$$\int_\Omega w_j \partial_j \varphi dx = 0 \quad \forall \varphi \in \mathcal{D}(\Omega) \tag{III.280}$$

Then if in addition $v \in C^\infty(\Omega)$, $v\varphi \in \mathcal{D}(\Omega)$ and so

$$0 = \int_\Omega w_j \partial_j [v\varphi] dx = \int_\Omega \partial_j v w_j \varphi dx + \int_\Omega v w_j \partial_j \varphi dx \tag{III.281}$$

$$\Rightarrow \int_\Omega \partial_j v w_j \varphi dx = -\int_\Omega v w_j \partial_j \varphi dx \tag{III.282}$$

Now for fixed φ and w, both sides of (III.283) are linear in v and bounded with respect to $||\cdot||_{W^{1,p}}$ so, again by the density of $C^\infty(\Omega)$ in $W^{1,p}(\Omega)$, (III.283) holds for arbitrary $v \in W^{1,p}(\Omega)$. Setting $u_1 = v$ and $(\mathrm{adj}(\nabla u))_{1,j} = w_j$ we have

$$\int_\Omega \partial_j u_1 (\mathrm{adj}(\nabla u))_{1,j}\varphi dx = -\int_\Omega u_1 (\mathrm{adj}(\nabla u))_{1,j}\partial_j\varphi dx \quad \forall\, \varphi \in \mathcal{D}(\Omega) \tag{III.283}$$

and so

$$\partial_j u_1 (\mathrm{adj}(\nabla u))_{1,j} = \partial_j [u_1 (\mathrm{adj}(\nabla u))_{1,j}] \quad \text{in } \mathcal{D}'(\Omega) \tag{III.284}$$

$\square$

Lemma III.16. Suppose that $\Omega \subset \mathbb{R}^3$ is bounded and open, $1 \le p \le \infty$ and $u_\nu \rightharpoonup u$ in $W^{1,p}(\Omega, \mathbb{R}^3)$ ($u_\nu \rightharpoonup^* u$ if $p = \infty$)

1. If $p \ge 2$ then

$$\mathrm{adj}(\nabla u_\nu) \rightharpoonup \mathrm{adj}(\nabla u)) \quad \text{in } (\mathcal{D}'(\Omega))^9 \tag{III.285}$$

2. If $p \ge 2$ and $\mathrm{adj}(\nabla u_\nu) \rightharpoonup \mathrm{adj}(\nabla u))$ in $(L^q(\Omega))^9$ with $1/p + 1/q \le 1$ then

$$\det(\nabla u_\nu) \rightharpoonup \det(\nabla u) \quad \text{in } \mathcal{D}'(\Omega) \tag{III.286}$$

Proof. By lemma III.15 to prove the first assertion it will suffice to show that

$$\int_\Omega u_\nu^\beta \partial_k u_\nu^\gamma \partial_j \varphi dx \to \int_\Omega u^\beta \partial_k u^\gamma \partial_j \varphi dx \quad \forall\, \varphi \in \mathcal{D}(\Omega) \tag{III.287}$$

Consider the bilinear map

$$L^r(\Omega) \times W^{1,p}(\Omega) \to \mathbb{R} \tag{III.288}$$

$$(\xi, \chi) \mapsto \int_\omega \xi \partial_k \chi \partial_j \varphi dx \tag{III.289}$$

where $\varphi \in \mathcal{D}(\Omega)$ is fixed. If $p^{-1} + r^{-1} \leq 1$ then by an application of Hölder's inequality the map is bounded and hence continuous. Now by assumption

$$u_\nu^\beta \rightharpoonup u^\beta \quad \text{in } W^{1,p}(\Omega) \tag{III.290}$$

In addition, by the Rellich-Koňdrasov embedding theorem we know that

$$u_\nu^\beta \to u^\beta \quad \text{in } L^r(\Omega)\ 1 \leq r \leq p^* = \begin{cases} \frac{3p}{3-p} & p < 3 \\ \infty & p \geq 3 \end{cases} \tag{III.291}$$

If $p \geq 2$ then $p^* \geq 6$ so we may find an r such that $1/p + 1/r \leq 1$ and $1 \leq r \leq p^*$ hence by lemma III.14

$$\int_\Omega u_\nu^\beta \partial_k u_\nu^\gamma \partial_j \varphi dx \to \int_\Omega u^\beta \partial_k u^\gamma \partial_j \varphi dx \tag{III.292}$$

as required.
By the second half of lemma (III.15) to prove the second assertion it will suffice to show that

$$\int_\Omega u_\nu^1 (\mathrm{adj}(\nabla u_\nu))_{1j} \partial_j \varphi dx \to \int_\Omega u^1 (\mathrm{adj}(\nabla u))_{1j} \partial_j \varphi dx \tag{III.293}$$

Since $\mathrm{adj}(\nabla u_\nu)_{1j} \rightharpoonup \mathrm{adj}(\nabla u)_{1j}$ in $L^q(\Omega)$ by identical reasoning to the proof of the first assertion we need to show that

$$u_\nu^1 \to u^1 \text{ in } L^r(\Omega) \text{ with } \frac{1}{r} + \frac{1}{q} \leq 1 \tag{III.294}$$

since by assumption $1/q + 1/p \leq 1$ the above condition on r translates to

$$\frac{1}{r} \leq 1 - \frac{1}{q} \leq 1 - (1 - \frac{1}{p}) \leq \frac{1}{p} \tag{III.295}$$

$$\Rightarrow r \geq p \tag{III.296}$$

Again by the Rellich-Koňdrasov theorem

$$u_\nu^\beta \to u^\beta \quad \text{in } L^r(\Omega)\ 1 \leq r \leq p^* = \begin{cases} \frac{3p}{3-p} & p < 3 \\ \infty & p \geq 3 \end{cases} \tag{III.297}$$

and if $p \geq 2$, $p^* \geq 6$ so we may choose r as required. □

IV. THE EXISTENCE THEOREMS OF BALL AND MORREY

We shall now use the results of the previous section to obtain the promised existence theorems. Our first result concerns integrands which are quasiconvex and finite everywhere and is originally due to Morrey ((Morrey, 1952), (Morrey, 1966)) and further refined by Meyers ((Meyers, 1965)).

Theorem IV.1. *Let $\Omega \subset \mathbb{R}^n$ be bounded and open, and $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \to \mathbb{R}$ be continuous, quasiconvex and satisfying*

$$\alpha |A|^p \leq f(x,u,A) \leq \gamma(1 + |u|^p + |A|^p) \tag{IV.1}$$

$$|f(x,u,A) - f(x,v,B)| \leq \beta \left(1 + |u|^{p-1} + |v|^{p-1} + |A|^{p-1} + |B|^{p-1}\right)(|u-v| + |A-B|) \tag{IV.2}$$

$$|f(x,u,A) - f(y,u,A)| \leq \eta(|x-y|)(1 + |u|^p + |A|^p) \tag{IV.3}$$

(IV.4)

with $p > 1$, $\alpha, \beta\gamma > 0$ and η a continuous increasing function with $\eta(0) = 0$. Then

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

admits at least one solution.

Proof. (IV.1) implies that the infimum in (P) is strictly less than infinity since consider $u_0 \in u_0 + W_0^{1,p}(\Omega, \mathbb{R}^m)$:

$$I(u_0) \leq \gamma \int_\Omega (1 + |u_0(x)|^p + |\nabla u_0(x)|^p)\, dx \tag{IV.5}$$

$$= \mu(\Omega) + ||u_0||_{W^{1,p}} \tag{IV.6}$$

$$< \infty \tag{IV.7}$$

In addition, (IV.1), together with Poincare's inequality, implies that I is coercive over $W^{1,p}(\Omega, \mathbb{R}^n)$. Finally by, (IV.2),(IV.3) and the quasiconvexity of f we have by theorem (III.12) that I is s.w.l.s.c., and so by theorem (III.1) (P) has at least one solution. □

The above theorem however, has several shortcomings, specifically with regard to problems of the form (.1) where f is the stored energy function of a hyperelastic body. Firstly, one cannot define quasiconvexity for a function taking the value of ∞. This sort of singular behaviour is useful in elasticity as, if u represents a deformation of Ω, the condition

$$f(x, \nabla u(x)) \to \infty \quad \text{when } \det(\nabla u) \to 0 \tag{IV.8}$$

encodes the fact that we cannot compress our body to a point. Secondly, the growth conditions and coercivity conditions (equations (IV.1) to (IV.3)) are too stringent and rule out many cases of interest. In light of this, Ball was motivated to study the case where the quasiconvexity of f is replaced by polyconvexity, and the growth conditions on f are weakened.

Theorem IV.2. *Suppose that $\Omega \subset \mathbb{R}^3$ is bounded and open and that $f : \Omega \times \mathbb{R}^3 \times \mathbb{R}^3 \to \bar{\mathbb{R}}$ is such that there exists a Carathé odory function $F : \Omega \times \mathbb{R}^{19} \to \bar{\mathbb{R}}$ with*

1. *$f(x, u, A) = F(x, u, T(A))\ \forall\ A \in \mathbb{R}^{3\times3}$ where $T(A) = (A, \operatorname{adj}(A), \det(A))$*

2. *$F(x, u, \cdot)$ convex for every $u \in \mathbb{R}^n$ and almost every $x \in \Omega$.*

3. $$f(x, u, A) = F(x, u, T(A)) \geq \alpha(x) + \beta_1|A|^{p_1} + \beta_2|\operatorname{adj}(A)|^{p_2} + \beta_3|\det(A)|^{p_3} \tag{IV.9}$$

   with $p_1 \geq 2$, $p_2 \geq \frac{p_1}{p_1 - 1}$ and $p_3 > 1$

Consider again the minimisation problem:

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

If there exists a $\tilde{u} \in u_0 + W_0^{1,p}(\Omega, \mathbb{R}^n)$ such that $U(\tilde{u}) < \infty$ then (P) admits at least one solution.

Proof. By assumption, the infimum in (P) is finite, so if u_ν is an infimising sequence we have that

$$\int_\Omega \alpha(x)dx + \beta_1 \int_\Omega |\nabla u(x)|^{p_1} dx + \beta_2 \int_\Omega |\operatorname{adj}(\nabla u(x))|^{p_2} dx + \beta_3 \int_\Omega |\det(\nabla u(x))|^{p_3} dx \leq I(u_\nu) \leq K \tag{IV.10}$$

By Poincare's inequality

$$||u||_{W^{1,p_1}} \leq ||\nabla u||_{L^{p_1}} \tag{IV.11}$$

so (IV.10) implies that

$$||u_\nu||_{W^{1,p_1}} + ||\operatorname{adj}(\nabla u_\nu)||_{L^{p_2}} + ||\det(\nabla u_\nu)||_{L^{p_3}} \leq K^{'} \tag{IV.12}$$

for some fixed $K^{'} > 0$ and so by the Banach-Alaoglu theorem we may extract a subsequence (still denoted as u_ν) such that

$$\begin{aligned}
&u_\nu \rightharpoonup u \text{ in } \; W^{1,p_1}(\Omega,\mathbb{R}^m) \Rightarrow u_\nu \to u \text{ in } L^{p_1}(\Omega,\mathbb{R}^m) && \text{(IV.13)}\\
&\nabla u_\nu \rightharpoonup \nabla u_\nu \text{ in } \; L^{p_1}(\Omega;\mathbb{R}^m) && \text{(IV.14)}\\
&\operatorname{adj}(\nabla u_\nu) \rightharpoonup H \text{ in } \; L^{p_2}(\Omega;\mathbb{R}^m) && \text{(IV.15)}\\
&\det(\nabla u_\nu) \rightharpoonup \delta \text{ in } \; L^{p_3}(\Omega;\mathbb{R}) && \text{(IV.16)}\\
\Rightarrow &(\nabla u_\nu, adj(\nabla u_\nu), \det(\nabla u_\nu)) \rightharpoonup (\nabla u, H, \delta) \text{ in } Ł^q(\Omega,\mathbb{R}^{19}) \text{ with } q \geq 1 && \text{(IV.17)}
\end{aligned}$$

and so we may apply theorem (III.9) to obtain that

$$\liminf_{\nu\to\infty} I(u_\nu) = \liminf_{\nu\to\infty} \int_\Omega f(x,u_\nu(x),\nabla u_\nu(x))dx \tag{IV.18}$$

$$= \liminf_{\nu\to\infty} \int_\omega F(x,u_\nu,\nabla u_\nu,\operatorname{adj}(\nabla u_\nu),\det(\nabla u_\nu))dx \tag{IV.19}$$

$$\geq \int_\Omega F(x,u,\nabla u,H,\delta)dx \tag{IV.20}$$

In order to establish the sequential weak lower semicontinuity of I it remains only to show that $H = \operatorname{adj}(\nabla u)$ and $\delta = \det(\nabla u)$.
But, since $p_1 \geq 2$, from the first assertion in lemma (III.16):

$$\operatorname{adj}(\nabla u_\nu) \rightharpoonup \operatorname{adj}(\nabla u) \quad \text{in } (\mathcal{D}^{'}(\Omega))^9 \tag{IV.21}$$

And since $p_2 \geq p_1/(p_1-1) > 1$, (IV.15) implies that

$$\operatorname{adj}(\nabla u_\nu) \rightharpoonup H \quad \text{in } (\mathcal{D}^{'}(\Omega))^9 \tag{IV.22}$$

Finally, since the weak-star topology on $\mathcal{D}^{'}(\Omega)$ is Hausdorff:

$$H = \operatorname{adj}(\nabla u) \tag{IV.23}$$

To establish that $\det(\nabla u) = \delta$, note first that by assumption:

$$\frac{1}{p^2} + \frac{1}{p^1} \leq \frac{p^1-1}{p^1} + \frac{1}{p^1} = 1 \tag{IV.24}$$

and by the first part of the proof:

$$\operatorname{adj}(\nabla u_\nu) \rightharpoonup \operatorname{adj}(\nabla u) \quad \text{in } L^{p^2}(\Omega,\mathbb{R}^n) \tag{IV.25}$$

by the second assertion in lemma (III.16) we have that

$$\det(\nabla u_\nu) \rightharpoonup \det(\nabla u) \quad \text{in } \mathcal{D}^{'}(\Omega) \tag{IV.26}$$

and since $p^3 > 1$

$$\det(\nabla u_\nu) \rightharpoonup \delta \quad \text{in } \mathcal{D}^{'}(\Omega) \tag{IV.27}$$

so again $\delta = \det(\nabla u)$ and returning to (IV.20)

$$\liminf_{\nu\to\infty} I(u_\nu) \geq \int_\Omega F(x,u,\nabla u \operatorname{adj}(\nabla u),\det(\nabla u))dx = I(u) \tag{IV.28}$$

Thus u is a solution to (P). □

V. CONCLUSION AND FURTHER READING

To conclude then, we have seen how we might obtain data on the existence of solutions to variational problems using much milder assumptions on our integrand than we could classically get away with. A question that naturally arises now is how regular is our solution? In the classical case, any solution obtained via the Euler-Lagrange equations will be at least $\mathcal{C}^2$ whereas the weak solution we have shown to exist might not even be continuous. This question is the subject of *regularity theory* where the Rellich-Koňdrasov embedding theorem, amongst other things, is used to estimate the degree of regularity of the solution guaranteed to exist by the direct methods. In the case where we can show our postulated solution is $\mathcal{C}^2$ we have proven the existence of a solution to the Euler-Lagrange equations which is an absolute minimizer. Another pertinent question is how do we find the solution which we now know to exist? Frequently we shall use a numerical scheme, such as the finite element method. The fact that our solution exists in some Sobolev space, combined with the observation that the space of piecewise affine functions is dense in the Sobolev spaces, (provided some mild conditions on the boundary Ω are satisfied cf. (II.8)) implies that our numerical scheme should converge. Finally, it is unfortunate that, due to space constraints, we have been unable to discuss the application of the above theory is to nonlinear elasticity which is a particularly nice example of rigourous mathematical analysis being used to solve real-world problems. For this the reader is referred to (Ball, 1977) or (Ciarlet, 1988).

VI. ACKNOWLEDGEMENTS

I am indebted to my supervisor, Dr Francois Ebobisse, for numerous informative conversations, and his willingness to help out whenever a particular theorem or proof had me stumped.

References

Adams, R. A., 1975, *Sobolev Spaces* (Academic Press).
Ball, J., 1977, Archive of Rational Mechanics and Analysis **63**, 337.
Ciarlet, P., 1988, *Mathematical Elasticity*, volume 20 of *Studies in mathematics and its applications* (Elsevier Science Publishers).
Dacorogna, B., 1989, *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences* (Springer-Verlag).
Ekeland, I., and R. Temam, 1972, *Convex analysis nad Variational problems*, number 1 in Studies in mathematics and its applications (North-Holland).
Kuttler, K., 2006, Topics in analysis, available at http://math.byu.edu/ klkuttle/.
Lee, J., 2003, *An introduction to smooth manifolds*, number 218 in Graduate Texts in Mathematics (Springer-Verlag New York).
Meyers, N., 1965, Trans. A.M.S. **119**, 125.
Morrey, C., 1952, Pacific J. Math. **2**, 25.
Morrey, C., 1966, *Multiple integrals in the Calculus of Variations* (Springer, Berlin).
Tonelli, L., 1921, *Fomdamenti di calcolo delle variazioni* (Zanichelli, Bologna).
Whitley, R., 1967, Mathematische Annalen **172**.