Nonlinear Variational Problems

I. Convex Problems

- An important class of nonlinear elliptic PDEs arise from problems in the calculus of variations.
- We will focus on a particularly vital set of representatives of this class: problems describing the equilibrium of **elastic** materials.
- In discussing this class of problems we will be able to examine not only many of the classical techniques for elliptic problems, but also more important techniques which have been developed over the last twenty years.

- We begin our brief discussion of nonlinear elasticity by describing the **kinematics** of geometry of deformation of three-dimensional bodies.
- We let a domain $\Omega \subset \mathbb{R}^3$ represent the **reference configuration** of a material body.

– We assume that Ω is bounded with Lipschitz boundary.
(It is usually convenient to think of the reference configuration Ω as the "rest" or "unstressed" configuration of the body, but we will not restrict ourselves to this case.)
– A **deformation** of the body is simply a mapping of the form

$$\mathbb{R}^3 \supset \Omega \ni \mathbf{x} \mapsto \mathbf{u}(\mathbf{x}) \in \mathbb{R}^3. \tag{1}$$

– It will ease our computations somewhat to spell out all of our vectors and matrices in terms of components. Thus, we let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ constitute a fixed orthonormal basis of $\mathbb{R}^3$, and for a given vector $\mathbf{v}$, we define $v_i = \mathbf{v} \cdot \mathbf{e}_i$..

- We will assume a certain amount of smoothness of the deformation: We will either make the classical assumption that $\mathbf{u} \in C^1_b(\overline{\Omega})$ or use a Sobolev space $\mathbf{u} \in W^{1,p}(\Omega)$.

– In either case, we are able to define the **deformation gradient**

$$\mathbf{F}(\mathbf{x}) := \frac{\partial \mathbf{u}}{\partial \mathbf{x}}(\mathbf{x}). \tag{2}$$

In terms of components we have

$$F_{ij}(\mathbf{x}) := \frac{\partial u_i}{\partial x_j}(\mathbf{x}). \tag{3}$$

- We will assume that the material **preserve orientation**.

– If $\mathbf{u} \in C^1_b(\overline{\Omega})$ this constraint can be expressed by the pointwise inequality

$$\det \mathbf{F}(\mathbf{x}) > 0. \tag{4}$$

– If we are taking $\mathbf{u} \in W^{1,p}(\Omega)$, then we will assume that the inequality (4) is satisfied almost everywhere.

- We will specify **displacement boundary conditions** on our deformations.

Typeset by $\mathcal{AMS}$-TEX

- The most natural way to state such conditions is to specify a continuous function $\mathbf{b} : \partial\Omega \to \mathbb{R}^3$ and to require that

$$\mathbf{u}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega. \tag{5}$$

- If we wish to consider $\mathbf{u} \in W^{1,p}(\Omega)$, we have to require that (5) holds in the sense of trace.
- However, when working in Sobolev spaces, it is more convenient to enforce boundary conditions by specifying a function $\mathbf{g} : \Omega \to \mathbb{R}^3$ with $\mathbf{g} \in W^{1,p}(\Omega)$ and requiring

$$\mathbf{p}_0 = \mathbf{u} - \mathbf{g} \in W_0^{1,p}(\Omega). \tag{6}$$

  This ensures that $\mathbf{u}$ and $\mathbf{g}$ have the same trace on $\partial\Omega$.
- In addition, we will be able to assume such conditions as

$$\det\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right) > 0 \quad \text{a.e.} \tag{7}$$

- To sum up, we assume that $\mathbf{g} \in W^{1,p}(\Omega)$ satisfying (7) exists and take the domain of our elasticity boundary problem to be

$$\mathcal{D}_\varepsilon := \{\mathbf{u} \in W^{1,p}(\Omega) | \det \nabla \mathbf{u} > 0 \ \text{ a.e. and } \ \mathbf{u} - \mathbf{g} \in W_0^{1,p}(\Omega)\}. \tag{8}$$

- We now pose a mathematical problem whose solutions will describe the equilibrium configurations of the body.
  - A material whose equilibrium are described by such a problem is said to be **elastic**.
  - We begin by defining an **elastic functional:**

$$\mathcal{E}(\mathbf{u}) := \int_\Omega \mathcal{W}(\mathbf{x}, \nabla \mathbf{u}(\mathbf{x}))\, d\mathbf{x} + \int_\Omega \Psi(\mathbf{u}(\mathbf{x}))\, d\mathbf{x}. \tag{9}$$

  Here $\mathbf{F} := \nabla \mathbf{u}$ and $\mathcal{W} : \Omega \times Q \to \mathbb{R}$ is the **stored energy density**, where Q is the set of 3×3 matrices with positive determinant.
  - The function $\Psi : \mathbb{R}^3 \to \mathbb{R}$ is the **potential energy density**.
  - The first term in the energy is called the **stored energy functional**, and describes the energy stored from mechanical deformation within the material.
  - The second term in the energy is called the **potential energy functional**, and describes the energy from exterior forces (assumed to be conservative).

- We make the following assumptions about the density functionals.
  (A1) We assume that $\mathcal{W} \in C^2(\Omega \times Q)$ and $\Psi \in C^1(\mathbb{R}^3)$.
  (A2) For some $p > 1$, $\exists$ a constant $k > 0$ and a function $w \in C_b(\Omega)$ such that

$$\mathcal{W}(\mathbf{x}, \nabla \mathbf{u}) \geq \omega(\mathbf{x}) + k|\nabla u|^2. \tag{10}$$

  Note: From now on this p is to be used in the definition of $\mathcal{D}_\varepsilon$.

(A3) There exists a constant C such that for every $\mathbf{u} \in \mathcal{D}_\varepsilon$ we have

$$\int_\Omega \Psi(\mathbf{u}(\mathbf{x})) \geq C. \tag{11}$$

(A4) There exists $\mathbf{u} \in \mathcal{D}_\varepsilon$ such that $\mathcal{E}(\nabla \mathbf{u}) < \infty$.
(A5) For every $\mathbf{x} \in \Omega$, we have $\mathcal{W}(\mathbf{x}, \nabla \mathbf{u}) \to \infty$ as $\det(\nabla \mathbf{u}) \to 0$.

- We now **define the equilibrium configurations** of the body to be those that minimize the energy $\mathcal{E}$; i.e., we say that $\widetilde{\mathbf{u}}$ is the **equilibrium state** or the **minimizer** if

$$\mathcal{E}(\widetilde{\mathbf{u}}) \leq \mathcal{E}(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{D}_\varepsilon. \tag{12}$$

- The primary goal of this section is to examine the question of existence of solutions of the minimization problem (12).

– Bur before doing so, we wish to explore the connection of the problem to elliptic PDEs. We do this in the following result.

Theorem 1 (Euler-Lagrange equation). *Suppose $\overline{\mathbf{u}} \in \mathcal{D}_\varepsilon \in C^2(\overline{\Omega})$ solves (12) and that (4) holds almost everywhere in Ω. Then at every $\mathbf{x} \in \Omega$, $\overline{\mathbf{u}}$ must satisfy the* **Euler-Lagrange equations**

$$-\sum_j \frac{\partial}{\partial x_j} \mathcal{A}_{ij} + \frac{\partial \Psi}{\partial u_j}(\mathbf{u}(\mathbf{x})) = 0, \quad i = 1, 2, 3. \tag{13}$$

where

$$\mathcal{A}_{ij}(\mathbf{x}) := \frac{\partial \overline{\mathcal{W}}}{\partial F_{ij}}(\mathbf{x}, \nabla \mathbf{u}(\mathbf{x})), \tag{14}$$

with $F_{ij}(\mathbf{x})$ defined in (3). In a homogeneous material $\mathcal{W} = \mathcal{W}(\nabla \mathbf{u})$, we have

$$\sum_j \frac{\partial}{\partial x_j} \mathcal{A}_{ij} = \sum_{j,k,\ell} \frac{\partial^2 u_k}{\partial x_j \partial x_\ell}, \tag{15}$$

where

$$a_{ijk\ell}(\mathbf{x}) = \frac{\partial^2 \mathcal{W}}{\partial F_{ij} \partial F_{k\ell}}(\nabla \mathbf{u}(\mathbf{x})).$$

Proof. Since $\overline{\mathbf{u}} \in \mathcal{D}_\varepsilon \cap C^2(\overline{\Omega})$, we see that for any $\phi = \sum \phi_i \mathbf{e}_i \in [C_0^\infty(\Omega)]^3$ we have

$$\overline{\mathbf{u}} + \varepsilon \phi \in \mathcal{D}_\varepsilon \quad \text{for } \varepsilon \text{ sufficiently small.} \tag{17}$$

Thus, for any ϕ, the real-valued function

$$f(\varepsilon) := \mathcal{E}(\overline{\mathbf{u}} + \varepsilon \phi)$$

is well defined in an interval about $\varepsilon = 0$. Furthermore, f is minimized at $\varepsilon = 0$ since by hypothesis $\mathcal{E}$ is minimized at $\mathbf{u}$. Now, using standard results on uniform convergence to take the derivative under the integral, we see that f is differentiable. We get

$$f'(\varepsilon) = \int_\Omega \left[\sum_{ij} \frac{\partial \mathcal{W}}{\partial F_{ij}}(\mathbf{x}, \nabla \overline{\mathbf{u}}(\mathbf{x}) + \varepsilon \nabla \phi(\mathbf{x})) \frac{\partial \phi_i}{\partial x_j} + \frac{\partial \Psi}{\partial u_j}(\overline{\mathbf{u}} + \varepsilon \phi)\phi_i \right] d\mathbf{x}. \tag{19}$$

Setting the derivative equal to zero at $\varepsilon = 0$ and integrating by parts gives us

$$\int_\Omega \left[-\sum_{ij} \frac{\partial}{\partial x_j} \frac{\partial \mathcal{W}}{\partial F_{ij}}(\mathbf{x}, \nabla \overline{\mathbf{u}}(\mathbf{x})) + \frac{\partial \Psi}{\partial u_j}(\overline{\mathbf{u}}) \right] \phi_i(\mathbf{x})\, d\mathbf{x} = 0. \tag{20}$$

Since ϕ is arbitrary, this implies that (13) is satisfied in the sense of distributions. Since each term in the equation is continuous, it must, in fact, be satisfied pointwise. □

- We now return to the question of existence.
- We will use what is called a **direct method** in the calculus of variations. (We will try to minimize the energy directly rather than solve, for instance, the Euler-Lagrange equations.)
- The following lemma provides a first important step.

Lemma 2. *There exists a minimizing sequence* $\{\mathbf{u}_k\} \in \mathcal{D}_\varepsilon$ *such that*

$$\lim_{k\to\infty} \mathcal{E}(\mathbf{u}_k) \le \mathcal{E}(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{D}_\varepsilon \tag{21}$$

and with the additional property that $\mathbf{u}_k$ *is weakly convergent; i.e. there exists* $\widetilde{\mathbf{u}} \in W^{1,p}(\Omega)$ *such that*

$$\mathbf{u}_k \rightharpoonup \widetilde{\mathbf{u}} \in W^{1,p}(\Omega). \tag{22}$$

Proof. We first note that $\mathcal{E}$ is bounded below on $\mathcal{D}_\varepsilon$; i.e. for any $\mathbf{u} \in \mathcal{D}_\varepsilon$, we can use (10) and (11) to obtain

$$\mathcal{E}(\mathbf{u}) \ge \int_\Omega [\omega(\mathbf{x}) + \Psi(\mathbf{u}(\mathbf{x}))]\, d\mathbf{x} \ge C - \|\omega\|_\infty |\Omega|, \tag{23}$$

where $|\Omega|$ is the volume of Ω. Since $\mathcal{E}(\mathbf{u})$ is bounded below, it must have a greatest lower bound L, and hence there must be a sequence $\mathbf{u}_k \in \mathcal{D}_\varepsilon$ such that

$$\lim_{k\to\infty} \mathcal{E}(\mathbf{u}_k) = L. \tag{24}$$

– Since $\mathcal{E}(\mathbf{u}_k)$ is a convergent sequence in $\mathbb{R}$ it must be bounded, say, by a constant K. Using this and (10) we obtain

$$K \ge |\mathcal{E}(\mathbf{u}_k)| \ge k \int_\Omega |\nabla \mathbf{u}_k(\mathbf{x})|^p\, d\mathbf{x} - \|\omega\|_\infty |\Omega| - C. \tag{25}$$

Rearranging and combining this with Poincaré's inequality on $\mathbf{u}_k - \mathbf{g}$ gives us

$$\|\mathbf{u}_k\| \le \widehat{K}, \tag{26}$$

for some constant $\widehat{K}$ independent of n.

– Since $\mathbf{u}_k$ is bounded, it has a weakly convergent subsequence, which we also label $\mathbf{u}_k$. Defining $\widetilde{\mathbf{u}}$ to be the weak limit of this subsequence gives us (22).
– Since the original sequence $\mathcal{E}(\mathbf{u}_k)$ converges to L, so does any subsequence; this gives us (21) and completes the proof. □

- Of course, since $\mathbf{u}_n$ is a minimizing sequence, our first guess is that its "limit" $\widetilde{\mathbf{u}}$ is a solution of our problem.

– However, two questions remain.

(i) Is $\widetilde{\mathbf{u}} \in \mathcal{D}_\varepsilon$; ie., the constraint $\det \nabla\widetilde{\mathbf{u}}$ is satisfied?
(ii) If so, is $\widetilde{\mathbf{u}}$ actually a minimizer; i.e., is it true that

$$\mathcal{E}(\widetilde{\mathbf{u}}) = \lim_{k\to\infty} \mathcal{E}(\mathbf{u}_k) = L \le \mathcal{E}(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{D}_\varepsilon? \tag{27}$$

⊙ In order to answer the first question, we extend the domain of definition of $\mathcal{E}$ to functions which do not satisfy the constraint (4).

– We define the function

$$\Omega \times \mathcal{M}^{3\times 3} \ni (\mathbf{x}, \mathbf{f}) \mapsto \overline{\mathcal{W}}(\mathbf{x}, \nabla\mathbf{u}) \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} \tag{28}$$

by

$$\overline{\mathcal{W}}(\mathbf{x}, \nabla\mathbf{u}) = \begin{cases} \mathcal{W}(\mathbf{x}, \nabla\mathbf{u}), & \det \nabla\mathbf{u} > 0 \\ \infty, & \det \nabla\mathbf{u} \le 0. \end{cases} \tag{29}$$

We can extend the domain of definition of our total energy to

$$W^{1,p}_{\mathbf{g}}(\Omega) := \{\mathbf{u} \in W^{1,p}(\Omega) | \mathbf{u} - \mathbf{g} \in W^{1,p}_0(\Omega)\},$$

by defining

$$\overline{\mathcal{E}}(\mathbf{u}) = \int_\Omega \overline{\mathcal{W}}(\mathbf{x}, \nabla\mathbf{u}(\mathbf{x}))\, d\mathbf{x} + \int_\Omega \Psi(\mathbf{u}(\mathbf{x}))\, d\mathbf{x}. \tag{31}$$

The following result is then immediate.

Lemma 3. *If $\overline{E}(\mathbf{u}) < \infty$, then $\mathbf{u} \in \mathcal{D}_\varepsilon$.*

- This means that if we can indeed answer the second question in the affirmative; i.e. if $\widetilde{\mathbf{u}}$ is a minimizer, then $\widetilde{\mathbf{u}} \in \mathcal{D}_\varepsilon$.

⊙ Thus, we focus on the second question and identify conditions that will ensure the weak limit is indeed a minimizer.

Definition. *Let X be a Banach space. A sequence $x_n \in X$* **converges weakly** *to x if $f(x_n)$ converges to $f(x)$ for every $f \in X^*$. A sequence $f_n \in X^*$* **converges weakly-*** *to f if $f_n(x)$ converges to $f(x)$ for every $x \in X$.*

A weakly convergent sequence in X^* also converges weakly-*. The converse is false in general, but is true if X is reflexive.

Definition. *Let X be a Banach space.*

⊙ *We say that a nonlinear subspace $\mathcal{F} : X \to \mathbb{R}$ is* **sequentially weakly lower semicontinuous (wlsc)** *if whenever $u_n \rightharpoonup \overline{v}$ in X, we have*

$$\mathcal{F}(\overline{v}) \le \liminf_{n\to\infty} \mathcal{F}(v_n). \tag{33}$$

⊙ *We say that a nonlinear subspace $\mathcal{F} : X \to \mathbb{R}$ is* **sequentially weak-star lower semicontinuous (wslsc)** *if whenever $u_n \overset{*}{\rightharpoonup} \overline{v}$ in X, it follows that (33) holds.*

⊙ *A mapping $\mathcal{F}$ is* **sequentially weakly continuous** *if*

$$\mathcal{F}(\overline{v}) = \lim_{n\to\infty} \mathcal{F}(v_n) \tag{35}$$

whenever $v_n \rightharpoonup \overline{v}$ in X.

⊙ *A mapping $\mathcal{F}$ is* **sequentially weak-star continuous** *if (35) holds whenever $v_n \overset{*}{\rightharpoonup} \overline{v}$ in X.*

- We state a theorem (usually attributed to Tonelli) which is the fundamental result on weak lower semicontinuity.

– However, before doing this we give a definition of a convex function whose domain is in general Banach space X and whose range is the exterior real line $\overline{\mathbb{R}}$.

Definition. *Let $K \subset X$ be a convex set. Then we say a mapping $G : K \to \overline{\mathbb{R}}$ is convex if for every $u, v \in K$, we have*

$$G(\lambda u + (1-\lambda)v) \le \lambda G(u) + (1-\lambda)G(v) \quad \forall \lambda \in [0,1]. \tag{36}$$

whenever the right-hand side of the inequality is well defined.

Theorem 4 (Tonelli). *For functions $\mathbf{v} : \Omega \to \mathbb{R}^n$, define the nonlinear function*

$$\mathcal{F}(\mathbf{v}) := \int_\Omega f(\mathbf{v}(\mathbf{x}))d\mathbf{x} \tag{37}$$

where $f : \overline{\mathbb{R}^n} \to \overline{\mathbb{R}}$ is continuous. Then the function $\mathcal{F}$ is

(i) sequentially weakly lower semicontinuous on $L^p(\Omega)$ for $1 < p < \infty$ and

(ii) weak-star lower semicontinuous on $L^\infty(\Omega)$

if and only if $\mathbb{R}^n \ni \mathbf{v} \mapsto f(\mathbf{v}) \in \overline{\mathbb{R}}$ is convex.

Sketch of the proof. We begin by showing that **weak-star lower semicontinuous of $\mathcal{F}$ in $L^\infty(\Omega)$ implies that f is convex**.

⊙ We first prove a lemma below which highlights one of the most important types of weak convergence: **a wildly oscillating sequence which converges to its average value**.

⊙ Indeed, assume that $\mathcal{F}$ is weak-star lower semicontinuous.
Let $C \subset \Omega \subset \mathbb{R}^n$ be a hypercube.
For any $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^n$, $\theta \in (0,1)$, the lemma below says that there exists a sequence of oscillating functions $\mathbf{v}_k$ with support on C such that

$$\mathbf{v}_k \overset{*}{\rightharpoonup} \overline{\mathbf{v}} := \theta\mathbf{a} + (1-\theta)\mathbf{b} \ \ \text{(a constant function)} \quad \text{in } L^\infty(\Omega).$$

The key observation here is that the sequence of composite functions $f(\mathbf{v}_k)$ oscillate between the values $f(\mathbf{a})$ and $f(\mathbf{b})$ with volume functions θ and $1-\theta$, respectively.

Lemma 5. *For any* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *and* $\theta \in (0,1)$*, define* $\mathbf{v} : \mathbb{R} \to \mathbb{R}^n$ *to be a function of period one such that*

$$\mathbf{v}(x) = \begin{cases} \mathbf{a}, & x \in [0,\theta) \\ \mathbf{b}, & x \in [\theta, 1). \end{cases}$$

Let the sequence of functions $\mathbf{v}_k : [0,1] \to \mathbb{R}^n$ *be given by*

$$\mathbf{v}_k(x) := \mathbf{v}(kx), \quad x \in [0,1].$$

Let $\overline{\mathbf{v}} = \theta\mathbf{a} + (1-\theta)\mathbf{b}$ *be a constant function. Then*

$$\mathbf{v}_k \rightharpoonup \overline{\mathbf{v}} \quad \text{in } L^p(0,1) \text{ for } 1 < p < \infty$$

and

$$\mathbf{v}_k \overset{*}{\rightharpoonup} \overline{\mathbf{v}} \quad \text{in } L^\infty(0,1)$$

where

Proof. Here as well, we just give a sketch of the proof. We wish to show that

$$\lim_{k\to\infty} \int_0^1 \mathbf{v}_k(x)\phi(x)\,dx = [\theta\mathbf{a} + (1-\theta)\mathbf{b}] \int_0^1 \phi(x)\,dx, \tag{42}$$

for all $\phi \in L^p(0,1)$, $1 \le p \le \infty$.

Step 1. The assertion is easy to prove for functions of the form

$$\phi(x) = \begin{cases} c, & x \in I, \\ 0, & x \notin I, \end{cases} \quad \text{where } I \subset [0,1] \text{ is an interval..} \tag{43}$$

(For large k the interval I will contain a large integral number of periods of $\mathbf{v}_k$ plus some small "slop" at the ends.)

Step 2. One can use the previous observation to show that (42) holds if ϕ is a *simple function*; i.e. if ϕ is piecewise constant with a finite number of jump discontinuities.

Step 3. One can then show that for any $\phi \in L^p(0,1)$, $p \in [1,\infty)$, and any $\varepsilon > 0$, we can find a simple function ϕ_s such that

$$\|\phi - \phi_s\|_1 < \varepsilon. \qquad \square \tag{44}$$

Remark. The choice of the interval $[0,1]$ was arbitrary. With only trivial modifications of the proof, we can show that

$$\mathbf{v}_k \to \overline{\mathbf{v}} \quad \text{in } L^p(I),$$

for any compact interval $I \subset \mathbb{R}$.

- Furthermore, we proved this result for functions whose domain is a subset of $\mathbb{R}$ only for clarity; an analogous construction can be created using functions $\mathbf{v}_k$ defined on the domain $\Omega \subset \mathbb{R}^n$ by letting the function oscillate in a single coordinate direction.

Return to the Proof of Tonelli's Theorem.

Assume that $\mathcal{F}$ is weak-star lower semicontinuous. Let $C \subset \Omega \subset \mathbb{R}^n$ be a hypercube. For any $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^n$, $\theta \in (0,1)$, the lemma provides us with a sequence of oscillating functions $\mathbf{v}_k$ with support on C such that

$$\mathbf{v}_k \overset{*}{\rightharpoonup} \overline{\mathbf{v}} := \theta\mathbf{a} + (1-\theta)\mathbf{b} \ \ \text{(a constant function)} \quad \text{in } L^\infty(\Omega).$$

The key observation here is that the sequence of composite functions $f(\mathbf{v}_k)$ oscillate between the values $f(\mathbf{a})$ and $f(\mathbf{b})$ with volume functions θ and $1-\theta$, respectively. Hence, it follows from the arguments in the proof of the lemma above that

$$f(\mathbf{v}_k) \overset{*}{\rightharpoonup} \theta f(\mathbf{a}) + (1-\theta) f(\mathbf{b}) \quad \text{in } L^\infty(\Omega).$$

Combining this with the weak-* lower semicontinuity of $\mathcal{F}$, we have

$$\begin{aligned} |\Omega| f(\theta\mathbf{a} + (1-\theta)\mathbf{b}) =& \mathcal{F}(\overline{\mathbf{v}}) \\ \leq& \liminf_{n\to\infty} \mathcal{F}(\mathbf{v}_k) \\ =& \lim_{n\to\infty} \int_\Omega f(\mathbf{v}_k) \\ =& |\Omega| \{\theta f(\mathbf{a}) + (1-\theta) f(\mathbf{b})\}. \end{aligned}$$

It follows that f is convex.

⊙ We now **assume that f is convex** and **show that $\mathcal{F}$ is weak-* lower semicontinuous**. We let $\mathbf{u}_k \to \overline{\mathbf{u}}$ be an arbitrary weakly convergent sequence and let

(48) $$L = \liminf \mathcal{F}(\mathbf{u}_k.)$$

By taking a subsequence, we can actually assume that

(49) $$L = \lim \mathcal{F}(\mathbf{u}_k.)$$

To complete the proof of the theorem we will use Mazur's lemma which we state without proof.

Lemma 6 (Mazur). *Let X be a Banach space and suppose*

(50) $$u_k \rightharpoonup \overline{u} \quad \text{in } X \text{ weakly.}$$

Then there exists a function $N : \mathbb{N} \to \mathbb{N}$, a sequence of sets of real numbers $\{\alpha(m)_k\}_{k=m}^{N(m)}$ such that $\alpha(m)_k \geq 0$ and $\sum_{k=m}^{N(m)} \alpha(m)_k = 1$ such that the sequence

(51) $$v_m := \sum_{k=m}^{N(m)} \alpha(m)_k u_k$$

converges **strongly** *to $\overline{u}$ in X.*

Remark. We say that v_m as defined above is a **convex combination** of elements of the set $\{u_k\}_{k=m}^{N(m)}$. The set of all possible convex combinations of elements of a set S is called the **convex hull** of S.

We can use Mazur's lemma to construct a sequence such that

$$\mathbf{v}_m \to \overline{\mathbf{u}} \quad \text{strongly in } L^p(\Omega) \quad \forall p \in [1, \infty).$$

Thus, at least for a subsequence we have

$$\mathbf{v}_m \to \overline{\mathbf{u}} \quad \text{a.e.}$$

There are two steps to the proof of Tonelli's theorem:

(i) Since $f(\mathbf{v}_m) \to f(\overline{\mathbf{u}})$ almost everywhere, it follows from Fatou's lemma and the convexity of f that

$$\mathcal{F}(\overline{\mathbf{u}}) \leq \liminf_{n\to\infty} \mathcal{F}(\mathbf{v}_m).$$

(If f is bounded below, as it is in most applications, we can use Fatou's lemma without using convexity. Otherwise, one can use the fact that any convex function is bounded below by an affine function.)

(ii) We now let $\varepsilon > 0$ be given and choose N sufficiently large so that

$$\mathcal{F}(\overline{\mathbf{u}}) \leq \mathcal{F}(\mathbf{v}_k) + \varepsilon/2 \quad \text{and} \quad \mathcal{F}(\mathbf{u}_k) \leq L + \varepsilon/2, \quad \text{for } k \geq N.$$

Then, for $m \geq N$, we have

$$\begin{aligned}\mathcal{F}(\overline{\mathbf{u}}) \leq& \mathcal{F}(\mathbf{v}_m) + \varepsilon/2 \\ =& \int_\Omega f\Big(\sum_{k=m}^{N(m)} \alpha(m)_k \mathbf{u}_k\Big)\, d\mathbf{x} + \varepsilon/2 \\ \leq& \sum_{k=m}^{N(m)} \alpha(m)_k \int_\Omega f(\mathbf{u}_k)\, d\mathbf{x} + \varepsilon/2 \\ \leq& L + \varepsilon.\end{aligned}$$

Here we have used the fact that $\alpha(m)_k \geq 0$ and $\sum_{k=m}^{N(m)} \alpha(m)_k = 1$. Since ε is arbitrary, the proof is complete. □

Corollary 7. *The mapping $\mathcal{F}$ is weakly or weak-star continuous iff $\mathbf{u} \mapsto f(\mathbf{u})$ is affine; i.e.*

$$f(\mathbf{u}) = \alpha + \mathbf{b} \cdot \mathbf{u}, \quad \textit{for some } \alpha \in \mathbb{R} \textit{ and } \mathbf{b} \in \mathbb{R}^n.$$

Proof. Apply the previous theorem to $\mathcal{F}$ and $-\mathcal{F}$. Notice that if both f and $-f$ are convex, then f is affine. □

Corollary 8. *If the function $\nabla\mathbf{u} \mapsto \overline{\mathcal{W}(\mathbf{x}, \nabla\mathbf{u})}$ is convex, then $\overline{\mathcal{E}}$ is weakly lower semicontinuous. Furthermore, there exists an equilibrium state $\widetilde{\mathbf{u}}$.*

Proof. Tonelli's theorem applies directly to the minimizing sequence and the stored energy. To take care of the potential energy term, we use compact imbedding to get $\mathbf{u}_k \to \overline{\mathbf{u}}$ strongly to $L^p(\Omega)$. Then, using Fatou's Lemma once again, we get

$$\int_\Omega \Psi(\overline{\mathbf{u}})\, d\mathbf{x} \leq \liminf_{k\to\infty} \int_\Omega \Psi(\mathbf{u}_k)\, d\mathbf{x}.$$

This completes the proof. □

- The result above is useful in some situations (e.g., linear elasticity).

⊙ However, in problems in nonlinear elasticity (where we really want to apply this theory in the first place) there are good reasons why the assumption of convexity of the energy density is physically unreasonable.

- If we weaken the convexity assumption on $\mathcal{W}$, Tonelli's theorem would seem to imply that we will not be able to show that the energy $\mathcal{E}$ is weakly lower semicontinuous.
- However, this is not the case.
- A useful result from the point of view of our problem in elasticity is the following theorem on the weak continuity of subdeterminants of gradients.

Theorem 9. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Suppose $n < p < \infty$ and suppose the sequence of functions $\mathbf{u}_k : \Omega \to \mathbb{R}^n$ satisfy*

$$\mathbf{u}_k \to \overline{\mathbf{u}} \quad \text{in } W^{1,p}(\Omega). \tag{61}$$

Let $m \le n$ and let M_k be the sequence of $m \times m$ subdeterminants obtained by taking a fixed m rows and m columns of $\nabla \mathbf{u}_k$, and let $\overline{M}$ be the corresponding subdeterminants of $\nabla \overline{\mathbf{u}}$. Then

$$M_k \rightharpoonup \overline{M} \quad \text{in } L^{p/m}(\Omega). \tag{62}$$

Proof. The proof proceeds by induction.

(i) The first step is to show that for any indices i, j, ℓ, m from 1 to n we have

$$\lim_{k\to\infty} \left(\frac{\partial u_i^k}{\partial x_\ell} \frac{\partial u_j^k}{\partial x_m} - \frac{\partial u_i^k}{\partial x_m} \frac{\partial u_j^k}{\partial x_\ell} \right) = \frac{\partial \overline{u}_i}{\partial x_\ell} \frac{\partial \overline{u}_j}{\partial x_m} - \frac{\partial \overline{u}_i}{\partial x_m} \frac{\partial \overline{u}_j}{\partial x_\ell}.$$

(ii) W.l.o.g., we can complete the proof by showing that

$$\det(\nabla \mathbf{u}_k) \rightharpoonup \det(\nabla \overline{\mathbf{u}}) \quad \text{in } L^{p/n}(\Omega) \tag{63}$$

under the assumption that any $(n-1) \times (n-1)$ subdeterminant M_k satisfies (62) with $m = n-1$.

– However, under this assumption we have

$$\operatorname{cof} \nabla \mathbf{u}_k \rightharpoonup \operatorname{cof} \nabla \overline{\mathbf{u}} \quad \text{in } L^{p/(n-1)}(\Omega); \tag{64}$$

here we have used $\operatorname{cof} \mathbf{A}$ to denote the **cofactor matrix** of $\mathbf{A}$. (The (i,j)-th component of the cofactor matrix is $(-1)^{i+j}$ times the (i,j)-th minor.)

– We now use the fact that for smooth function $\mathbf{u}$

$$\det \nabla \mathbf{u} = \sum_j \frac{\partial u_i}{\partial x_j} (\operatorname{cof} \nabla \mathbf{u}_k)_{ij}, \quad \forall i = 1, \cdots, n. \tag{65}$$

In addition, we use the identity (cf. Evans, PDEs, page 440)

$$\sum_j \frac{\partial}{\partial x_j} (\operatorname{cof} \nabla \mathbf{u})_{ij} = 0, \quad i = 1, \cdots, n. \tag{66}$$

Therefore we can derive the formula

$$\det \nabla \mathbf{u} = \sum_j \frac{\partial}{\partial x_j}(u_i(\operatorname{cof} \nabla \mathbf{u})_{ij}). \tag{67}$$

Thus, after approximating our sequence $\mathbf{u}_k$ by smooth functions, we have for any $\phi \in \mathcal{D}(\Omega)$,

$$\begin{aligned}\int_\Omega \phi \det \nabla \mathbf{u}_k \, d\mathbf{x} &= -\int_\Omega \sum_j \frac{\partial \phi}{\partial x_j}(u_i^k(\operatorname{cof} \nabla \mathbf{u}_k)_{ij}) \, d\mathbf{x} \\ &\to -\int_\Omega \sum_j \frac{\partial \phi}{\partial x_j}\overline{u}_i(\operatorname{cof} \nabla \overline{\mathbf{u}})_{ij}) \, d\mathbf{x} \\ &= \int_\Omega \phi \det \nabla \overline{\mathbf{u}} \, d\mathbf{x}.\end{aligned}$$

In taking the limit above we have used the fact that by compact imbedding $\mathbf{u}_k \to \overline{\mathbf{u}}$ strongly in $L^p(\Omega)$ and hence (since Ω is bounded) $\mathbf{u}_k \to \overline{\mathbf{u}}$ strongly in $L^q(\Omega)$, where $q = p/(1+p-n)$ is the conjugate exponent of $p/(n-1)$.

– Thus, $\det \nabla \mathbf{u}_k \to \det \nabla \overline{\mathbf{u}}$ in $\mathcal{D}'(\Omega)$. To complete the proof, we use the fact that $\det \nabla \mathbf{u}_k$ is bounded in $L^{p/n}(\Omega)$ and a density argument. □

- We can use this result coupled with the following definition in our study of variational problems of nonlinear elasticity.

Definition. *Let $\mathcal{M}^{m\times n}$ be the set of $m \times n$ matrices. A function $G : \mathcal{M}^{m\times n} \to \mathbb{R}$ is said to be* **polyconvex** *if $\mathbf{A} \mapsto G(\mathbf{A})$ can be represented as a convex function of the subdeterminant of $\mathbf{A}$.*

⊙ In the particular case of three-dimensional elasticity, we say that the stored energy density $\mathcal{W}(\mathbf{x}, \nabla \mathbf{u})$ is **polyconvex** if there exists a function

$$\Omega \times \mathcal{M}^{3\times 3} \times \mathcal{M}^{3\times 3} \times (0,\infty) \ni (\mathbf{x}, \mathbf{A}, \mathbf{B}, d) \mapsto g(\mathbf{x}, \mathbf{A}, \mathbf{B}, d) \in \mathbb{R}, \tag{68}$$

such that for every $\mathbf{x} \in \Omega$, $(\mathbf{A}, \mathbf{B}, d) \mapsto g(\mathbf{x}, \mathbf{A}, \mathbf{B}, d)$ is convex and

$$\mathcal{W}(\mathbf{x}, \nabla \mathbf{u}) = g(\mathbf{x}, \nabla \mathbf{u}, \operatorname{cof} \nabla \mathbf{u}, \det \nabla \mathbf{u}). \tag{69}$$

Remark. Polyconvexity is indeed a weaker assumption than convexity. In particular, such functions as

$$G(\mathbf{F}) = \begin{cases} \frac{1}{\det \mathbf{F}}, & \det \mathbf{F} > 0 \\ \infty, & \det \mathbf{F} \le 0, \end{cases}$$

are polyconvex but not convex.

Remark. In elasticity, where $\mathbf{F}(\mathbf{x}) = \nabla \mathbf{u}(\mathbf{x})$ is the deformation gradient of $\mathbf{x} \in \Omega$, the components of $\mathbf{F}(\mathbf{x})$ reflect local changes under the deformation in the length of curves going through $\mathbf{x}$, the components of $\operatorname{cof} F(\mathbf{x})$ reflect changes in the areas of variously oriented surfaces and $\det \mathbf{F}(\mathbf{x})$ reflects the local change in volume.

Remark. Polyconvexity is widely accepted as an assumption which is both extremely general and physically reasonable.

Theorem 10. *Let $\overline{\mathcal{W}}$ satisfy assumptions* **(A1)-(A5)** *on pages 2-3 with $p > 3$, and, in addition, be polyconvex. Then there exists a minimizer $\widetilde{\mathbf{u}} \in \mathcal{D}_\varepsilon$ of the energy $\overline{\mathcal{E}}$.*

Proof. By Lemma 2, there exists a minimizing sequence $\mathbf{u}_n \in \mathcal{D}_\varepsilon$ and a function $\widetilde{\mathbf{u}} \in W^{1,p}_{\mathbf{g}}(\Omega)$ such that $\mathbf{u}_n \rightharpoonup \widetilde{\mathbf{u}}$ in $W^{1,p}(\Omega)$. Then Theorem 9 gives us

$$\begin{aligned} \operatorname{cof} \nabla \mathbf{u}_n &\rightharpoonup \operatorname{cof} \nabla \widetilde{\mathbf{u}} \quad \text{in } L^{p/2}(\Omega), \\ \det \nabla \mathbf{u}_n &\rightharpoonup \det \nabla \widetilde{\mathbf{u}} \quad \text{in } L^{p/3}(\Omega). \end{aligned}$$

Now, we use Tonelli Theorem to get

$$\begin{aligned} \overline{\mathcal{E}}(\widetilde{\mathbf{u}}) &= \int_\Omega \overline{\mathcal{W}}(\mathbf{x}, \nabla \widetilde{\mathbf{u}}) + \Psi(\widetilde{\mathbf{u}})\, d\mathbf{x} \\ &= \int_\Omega g(\mathbf{x}, \nabla \widetilde{\mathbf{u}}(\mathbf{x}), \operatorname{cof} \nabla \widetilde{\mathbf{u}}(\mathbf{x}), \det \nabla \widetilde{\mathbf{u}}(\mathbf{x})) + \Psi(\widetilde{\mathbf{u}}(\mathbf{x}))\, d\mathbf{x} \\ &\leq \liminf_{n\to\infty} \int_\Omega g(\mathbf{x}, \nabla \mathbf{u}_n(\mathbf{x}), \operatorname{cof} \nabla \mathbf{u}_n(\mathbf{x}), \det \nabla \mathbf{u}_n(\mathbf{x})) + \Psi(\mathbf{u}_n(\mathbf{x}))\, d\mathbf{x} \\ &= \lim_{n\to\infty} \int_\Omega \overline{\mathcal{W}}(\mathbf{x}, \nabla \mathbf{u}_n) + \Psi(\mathbf{u}_n)\, d\mathbf{x} \\ &= \text{the greatest lower bound of } \overline{\mathcal{E}} \text{ over } \mathcal{D}_\varepsilon. \end{aligned}$$

Thus, $\widetilde{\mathbf{u}}$ is indeed a minimizer. □

II. Nonconvex Problems

- The study of nonconvex variational problems has occupied a significant fraction of the PDE community over the last twenty years.
- Such problems arise in PDEs which model phase transitions, e.g. the liquid-gas transition, phase transitions in crystalline solids, ferromagnetism and superconductivity.
- We discuss the main issues which arise in a simple one-dimensional example.

Nonconvex Model Problem.

Let us try to minimize

$$I(u) = \int_0^1 [(u(x))^2 + (u'(x) - 1)^2]\, dx \tag{73}$$

subject to the boundary conditions $u(0) = u(1) = 0$.

⊙ Note that to do so we must meet two conflicting demands. To make it small
  (i)The first term under the integral is a "nice" convex functions of u.
    we must look for functions u that are as close to zero as possible.
  (ii) The second term is nonconvex. It is minimized when $u' = \pm 1$. In between the values 1 and -1 the term achieves a local maximum when $u' = 0$.

⊙ With this in mind minimizing sequence is easily found.

⊙ For each $i = 0, 1, \cdots, n-1$, let

$$u_n(x) = \begin{cases} x - \frac{1}{n}, & \frac{1}{n} \le x \le \frac{1}{n} + \frac{1}{2n}, \\ \frac{i+1}{n} - x & \frac{1}{n} + \frac{1}{2n} \le x \le \frac{i+1}{n}. \end{cases} \tag{74}$$

That is, u_n is a sequence of piecewise linear functions that "zigzag" between derivatives ± 1.

– The maximum height of the function is $\frac{1}{2n}$.
– Since the second term in I is identically zero, we can easily calculate that $I(u_n) = \frac{1}{4n}$. Thus

$$\lim_{n\to\infty} I(u_n) = 0,$$

and since I is nonnegative, the sequence approaches the minimum value of I.

⊙ However, because of the nonconvex term, I is not continuous in the way we might expect. That is even though the sequence of functions u_n converges to zero uniformly, we have

$$I(\lim_{n\to\infty} u_n) = I(0) = 1 \ne \lim_{n\to\infty} I(u_n).$$

⊙ Indeed, it is easy to see that no function can possibly attain the minimum of I. Any function u with $I(u) = 0$ would have to satisfy $u = 0$ and $u' = \pm 1$, but there is no such function.

- The behavior found in this example turns out to be typical of nonconvex variational problems.
- There are two possible strategies to cope with this:
  (i) Identify the problem satisfied by the limit of minimizing sequences; obviously this will be a problem different from the original one.
  (ii) Actually define a "function" which is equal to zero but has derivatives ± 1.
- In this section, we shall give a brief outline of each of these approaches.

Convexification

Definition. *For a function $f(p)$, defined on an interval, we define the lower envelope by*

$$Cf = \{g | g \ \ \text{convex}, \ \ g \le f\}. \tag{75}$$

(a) For variational problems of the form

$$\min \int_a^b f(x, u, u')\, dx,$$

it is possible to show under quite general hypotheses that the limits of minimizing sequences solve the problem

$$\min \int_a^b Cf(x, u, u')\, dx, \tag{77}$$

where the lower convex envelope is w.r.t. the third variables.

– For instance, in the example above, we have

$$f(x,u,u') = ((u')^2-1)^2+u^2, \tag{78}$$

and

$$Cf(x,u,u') = \begin{cases} ((u')^2-1)^2+u^2, & |u'|\geq 1 \\ u^2, & |u'|<1. \end{cases}$$

Obviously, $u=0$ is indeed a minimizer of the modified problem.

(b) In the **multidimensional case**, however, convexity is not the right condition, as we have saw in the previous section.
The appropriate notion is **quasiconvexity**.

Definition. *The function* $f:\mathbb{R}^m\to\mathbb{R}$ *is called* **quasiconvex** *if*

$$\frac{1}{|D|}\int_D f(\mathbf{A}+\nabla\mathbf{u}(\mathbf{x}))\,d\mathbf{x} \geq f(\mathbf{A}) \tag{81}$$

for every bounded domain $D\subset\mathbb{R}^n$*, every* $\mathbf{A}\in\mathbb{R}^{nm}$ *and every* $\mathbf{u}\in(\mathcal{D}(D))^m$*. Here* $|D|$ *denotes the volume of* D*.*

⊙ Quasiconvexity is linked to lower semicontinuity in the multidimensional case in the same way as convexity is in the one-dimensional case.

⊙ If appropriate technical hypotheses are satisfied, it is therefore possible to show that limits of minimizing sequences for the problem

$$\min\int_\Omega f(\mathbf{x},\mathbf{u},\nabla\mathbf{u})\,d\mathbf{x}$$

are minimizers for the functional

$$\min\int_\Omega Qf(\mathbf{x},\mathbf{u},\nabla\mathbf{u})\,d\mathbf{x}$$

where Qf denotes the lower quasiconvex (w.r.t. the third variable) envelope:

$$Qf = \{g|g \ \text{ quasiconvex in } \mathbf{A},\quad g\leq f\}. \tag{83}$$

- Unfortunately, quasiconvexity is not a simple pointwise condition on f as convexity is and is difficult to verify.

– A necessary condition is **rank one convexity**.

Definition. *The function* $f:\mathbb{R}^{mn}\to\mathbb{R}$ *is called* **rank one convex** *if*

$$f(\lambda\mathbf{A}+(1-\lambda)\mathbf{B}) \leq \lambda f(\mathbf{A})+(1-\lambda)f(\mathbf{B})$$

for every $\lambda\in(0,1)$ *and every* $\mathbf{A},\mathbf{B}$ *such that* $\operatorname{rank}(\mathbf{A}-\mathbf{B})=1$.

- We can define a lower polyconvex envelope Pf and a lower rank one convex envelope Rf in an analogous fashion as the lower convex envelope Cf and lower quasiconvex envelope Qf.

⊙ Since it can be shown that

$$\begin{aligned}\text{convexity} &\Rightarrow \text{polyconvexity} \Rightarrow \text{quasiconvexity}\\ &\Rightarrow \text{rank one convexity.}\end{aligned}$$

we have

$$Cf \leq Pf \leq Qf \leq Rf \leq f.$$

Hence we have a characterization of Qf if we can show that $Pf = Rf$.

Generalized Functions

- How can one define a "function" which is equal to zero but has derivative ± 1?

⊙ The solution to this dilemma is to **assign probabilities** rather than values **to the derivative**.

– This allows the function 0 to have derivatives ± 1, as long as both values occur with equal probabilities.

– The development of this theory, originally introduced by L.C. Young, rests on the following result.

Theorem 11. *Let $K \subset \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$ be bounded, open sets, and let $\mathbf{v}_n \in (L^\infty(\Omega))^m$ be such that $\mathbf{v}_n$ has values in K.*
Then there exists a family of probability measures ν_x, $x \in \Omega$, on $\overline{K}$ and a subsequence $\mathbf{v}_{n_k}$ such that, in the sense of weak- convergence in $(L^\infty(\Omega))^m$, we have*

$$\lim_{k\to\infty} f(\mathbf{v}_{n_k}) = \int_K \nu_k(\mathbf{u}) f(\mathbf{u})\, d\mathbf{u}$$

for every continuous function $f : K \to \mathbb{R}$.

Example. If we take v_n to be the derivative of the zigzag function given by (74), we have

$$\lim f(v_n) = \frac{1}{2}(f(1) + f(-1));$$

in particular this yields $\lim v_n = 0$ and $\lim(v_n^2 - 1)^2 = 0$.