Nonlinear analysis and applications
Survey course

© John Ball

November 11, 2015

1 Variational problems for nonlinear elasticity and liquid crystals

For more information on these topics see `http://people.maths.ox.ac.uk/ball/teaching.shtml` (e.g. the TCC course on Mathematical Foundations of Elasticity Theory and the 2015 Harbin short course on the Mathematics of Solid and Liquid Crystals).

1.1 Nonlinear elasticity

Nonlinear elasticity is the central model of solid mechanics. Rubber, metals (and alloys), rock, wood, bone ... can all be modelled as elastic materials, even though their chemical compositions are very different.

For example, metals and alloys are crystalline, with grains consisting of regular arrays of atoms. Polymers (such as rubber) consist of long chain molecules that are wriggling in thermal motion, often joined to each other by chemical bonds called *crosslinks*. Wood and bone have a cellular structure ...

A brief history.

1678 Hooke's Law

1705 Jacob Bernoulli

1742 Daniel Bernoulli

1744 L. Euler elastica (elastic rod)

1821 Navier, special case of linear elasticity via molecular model (Dalton's atomic theory was 1807)

1822 Cauchy, stress, nonlinear and linear elasticity

For a long time the nonlinear theory was ignored/forgotten.

1927 A.E.H. Love, Treatise on linear elasticity

1950's R. Rivlin, Exact solutions in incompressible nonlinear elasticity (rubber)

1960 – 80 Nonlinear theory clarified by J. L. Ericksen, C. Truesdell

1980 – Mathematical developments, applications to materials, biology

1.1.1 Kinematics (statics)

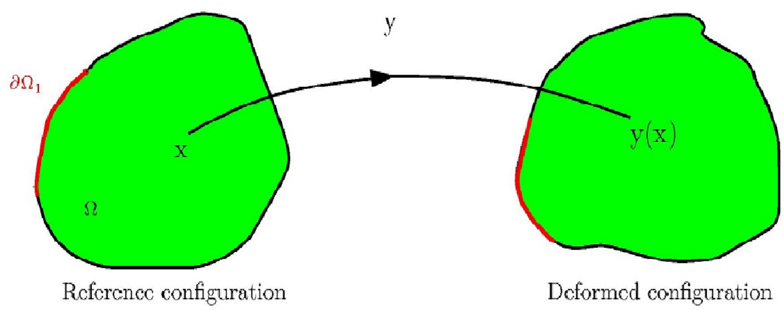

Figure 1: Deformation of an elastic body

We adopt the *Lagrangian* or *material* description, labelling the material points of a body by their positions in a reference configuration, in which the body occupies a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ with boundary $\partial\Omega$. A typical deformation is described by a mapping $\mathbf{y} : \Omega \to \mathbb{R}^3$, in terms of which the *deformation gradient* is given by

$$\mathbf{F} = \nabla \mathbf{y}(x), \quad F_{i\alpha} = \frac{\partial y_i(x)}{\partial x_\alpha}. \tag{1.1}$$

To avoid interpenetration of matter, we require that $\mathbf{y} : \Omega \to \mathbb{R}^3$ is invertible. Note that it would not be reasonable to suppose that $\mathbf{y} : \bar{\Omega} \to \mathbb{R}^3$ is invertible (see Fig. 2). We also suppose that $\mathbf{y}$ is orientation-preserving, so that

$$J = \det \nabla \mathbf{y}(x) > 0 \ \text{ for } x \in \Omega. \tag{1.2}$$

(Note that if $\mathbf{y} \in C^1$ then (1.2) implies that $\mathbf{y}$ is *locally* invertible.)

Notation.
$M^{3\times 3} = \{\text{real } 3 \times 3 \text{ matrices}\}$
$M^{3\times 3}_+ = \{\mathbf{F} \in M^{3\times 3} : \det \mathbf{F} > 0\}$
$SO(3) = \{\mathbf{R} \in M^{3\times 3} : \mathbf{R}^T\mathbf{R} = \mathbf{1}, \det \mathbf{R} = 1\} = \{\text{rotations}\}$.

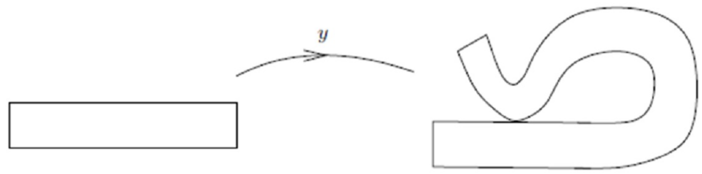

Figure 2: Deformation $\mathbf{y}$ invertible on Ω but not on $\bar{\Omega}$.

Theorem 1 (Polar decomposition)**.** *Let $\mathbf{F} \in M_+^{3\times 3}$. Then there exist positive definite symmetric $\mathbf{U}, \mathbf{V} \in M_+^{3\times 3}$ and $\mathbf{R} \in SO(3)$ such that*

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}.$$

These representations (right and left respectively) are unique.

1.1.2 Variational formulation of nonlinear elastostatics

Find $\mathbf{y} : \Omega \to \mathbb{R}^3$ (invertible) minimizing

$$I(\mathbf{y}) = \int_\Omega W(\nabla \mathbf{y}(x))\, dx$$

subject to suitable boundary conditions, e.g.

$$\mathbf{y}|_{\partial\Omega_1} = \bar{\mathbf{y}},$$

for given $\bar{\mathbf{y}}$, where $\partial\Omega_1 \subset \partial\Omega$. Here we have assumed (i) that the body is *homogeneous* (same material response at each point x) (ii) no body forces (such as gravity).

$W : M_+^{3\times 3} \to \mathbb{R}$ is the *stored-energy function* of the material (sometimes called the *strain-energy* or *free-energy* function).

Properties of W.

(H1) $W : M_+^{3\times 3} \to [0, \infty)$ is C^1.
(H2) $W(\mathbf{F}) \to \infty$ as $\det \mathbf{F} \to 0+$.
(H3) (Frame-indifference)

$$W(\mathbf{RF}) = W(\mathbf{F}) \text{ for all } \mathbf{R} \in SO(3), \mathbf{F} \in M_+^{3\times 3}.$$

It follows from (H3) and the polar decomposition theorem that

$$W(\mathbf{F}) = W(\mathbf{RU}) = W(\mathbf{U}) = \tilde{W}(\mathbf{C}),$$

where $\mathbf{C} = \mathbf{F}^T\mathbf{F}, \mathbf{U} = \mathbf{C}^{\frac{1}{2}}$.

(H4) (Material symmetry)

$$W(\mathbf{FH}) = W(\mathbf{F}) \text{ for all } \mathbf{H} \in \mathcal{S},$$

where the *symmetry group* $\mathcal{S}$ is a subgroup of the group of unimodular matrices (that is those matrices $\mathbf{H} \in M^{3\times 3}$ with $\det \mathbf{H} = 1$). If $\mathcal{S} \subset SO(3)$ then W is said to be *isotropic*.

Exercise 1. $\mathbf{T}_R(\mathbf{F}) = DW(\mathbf{F})$, $T_{Ri\alpha}(\mathbf{F}) = \frac{\partial W}{\partial F_{i\alpha}}$ is the *Piola-Kirchhoff stress tensor*, while

$$\mathbf{T}(\mathbf{F}) = J^{-1}\mathbf{T}_R\mathbf{F}^T$$

is the *Cauchy stress tensor*. Prove that (H1)+(H3) imply that $\mathbf{T}$ is symmetric. (*Hint.* choose $\mathbf{R} = e^{\mathbf{K}t}$, where $\mathbf{K}$ is skew.)

Exercise 2. Show that if there were an $\mathbf{H} \in M^{3\times 3}_+$ with $\det \mathbf{H} \neq 1$ and

$$W(\mathbf{F}) = W(\mathbf{FH}) \text{ for all } \mathbf{F} \in M^{3\times 3}_+$$

then this would contradict (H2).

1.2 Nematic liquid crystals

Nematic liquid crystals consist of short (2-3nm) rod-like molecules (see Fig. 3) that are usually more or less well aligned in the direction of some unit vector $\mathbf{n}(x)$ (see Fig. 4).

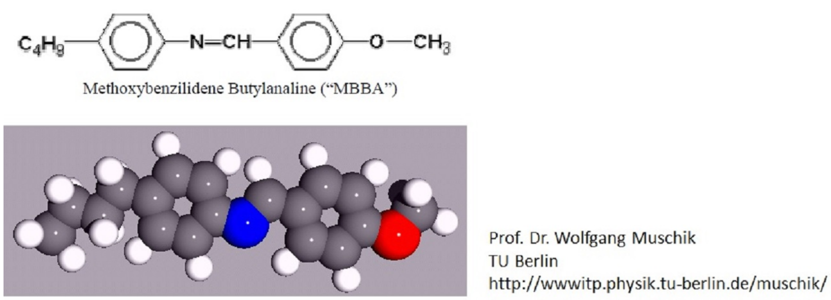

Figure 3: A typical liquid crystal molecule: MBBA

The oldest, and widely used, theory is that of *Oseen-Frank* with corresponding energy functional

$$I(\mathbf{n}) = \int_\Omega W(\mathbf{n}, \nabla\mathbf{n})\, dx, \tag{1.3}$$

which we seek to minimize subject to the constraint

$$|\mathbf{n}(x)| = 1, \quad x \in \Omega,$$

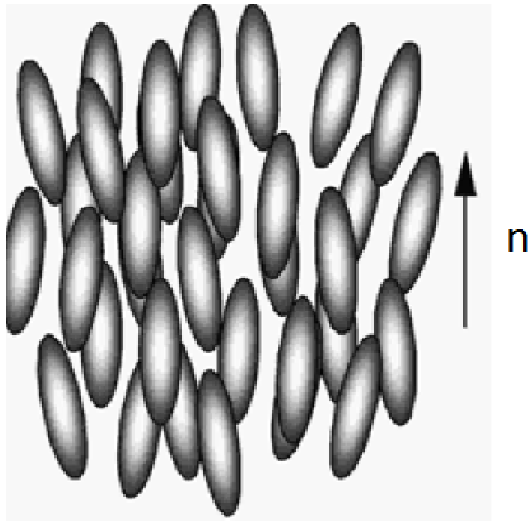

Figure 4: Director $\mathbf{n}$ gives the mean orientation of the rod-like molecules.

and suitable boundary conditions, such as

$$\mathbf{n}|_{\partial\Omega} = \bar{\mathbf{n}}.$$

Here Ω is a container filled with liquid crystal.

For this theory, frame-indifference is the condition that

$$W(\mathbf{Rn}, \mathbf{RAR}^T) = W(\mathbf{n}, \mathbf{A}) \text{ for all } \mathbf{R} \in SO(3), \mathbf{n} \in S^2, \mathbf{A} \in M^{3\times 3}. \tag{1.4}$$

We also assume the symmetry conditions

$$W(-\mathbf{n}, \mathbf{A}) = W(-\mathbf{n}, -\mathbf{A}) = W(\mathbf{n}, \mathbf{A}) \text{ for all } \mathbf{n}, \mathbf{A}. \tag{1.5}$$

If $W(\mathbf{n}, \mathbf{A})$ is quadratic in $\mathbf{A}$ then (1.4), (1.5) imply that W has the form

$$\begin{aligned} W(\mathbf{n}, \nabla\mathbf{n}) = {} & K_1(\operatorname{div}\mathbf{n})^2 + K_2(\mathbf{n}\cdot\operatorname{curl}\mathbf{n})^2 + K_3|\mathbf{n}\times\operatorname{curl}\mathbf{n}|^2 \\ & + (K_2+K_4)(\operatorname{tr}(\nabla\mathbf{n})^2 - (\operatorname{div}\mathbf{n})^2), \end{aligned} \tag{1.6}$$

for constants $K_1, \ldots, K_4$. Necessary and sufficient conditions for

$$W(\mathbf{n}, \nabla\mathbf{n}) \geq C|\nabla\mathbf{n}|^2 \text{ for some } C > 0,$$

are the Ericksen inequalities

$$2K_1 > K_2 + K_4, \quad K_2 > |K_4|, \quad K_3 > 0 \tag{1.7}$$

Exercise 3. Show that if $K_1 = K_2 = K_3 = K > 0, K_4 = 0$ then

$$I(\mathbf{n}) = K\int_\Omega |\nabla\mathbf{n}|^2\,dx$$

for which the corresponding minimizers are *harmonic maps* $\mathbf{n} : \Omega \to S^2$.

There is another more sophisticated theory, popular with physicists, called the *Landau - de Gennes theory.* It is an approximation to even more detailed theories in which the ordering of the molecules is represented by a probability distribution $\rho(x, \mathbf{p})$ on S^2, so that $\rho(x, \mathbf{p}) \geq 0$, $\int_{S^2} \rho(x, \mathbf{p})\, d\mathbf{p} = 1$, and $\rho(x, \mathbf{p}) = \rho(x, -\mathbf{p})$ (reflecting the statistical head-to-tail symmetry of the molecules). The first moment of $\rho(x, \cdot)$

$$\int_{S^2} \mathbf{p}\, \rho(x, \mathbf{p})\, d\mathbf{p} = 0.$$

The second moment is

$$\mathbf{M}(x) = \int_{S^2} \mathbf{p} \otimes \mathbf{p}\, \rho(x, \mathbf{p})\, d\mathbf{p},$$

where $(\mathbf{p} \otimes \mathbf{p})_{ij} = p_i p_j$. Note that $\mathbf{M} = \mathbf{M}^T \geq 0$ and $\operatorname{tr} \mathbf{M} = 1$. Define the de Gennes $\mathbf{Q}$ tensor by

$$\mathbf{Q} = \mathbf{M} - \frac{1}{3}\mathbf{1},$$

so that

$$\mathbf{Q} = \mathbf{Q}^T, \;\; \mathbf{Q} \geq -\frac{1}{3}\mathbf{1}, \;\; \operatorname{tr} \mathbf{Q} = 0.$$

(Note that $\mathbf{M}_0 = \frac{1}{3}\mathbf{1}$ is the value of $\mathbf{M}$ corresponding to the isotropic distribution of molecules $\rho(x, \mathbf{p}) = \frac{1}{4\pi}$, so that $\mathbf{Q}$ measures the deviation of $\mathbf{M}$ from its isotropic value.)

In the Landau - de Gennes theory the energy functional is given by

$$I_{\mathrm{LdG}}(\mathbf{Q}) = \int_\Omega \psi(\mathbf{Q}, \nabla \mathbf{Q})\, dx,$$

which needs to be minimized subject to suitable boundary conditions such as

$$\mathbf{Q}|_{\partial\Omega} = \bar{\mathbf{Q}},$$

where $\bar{\mathbf{Q}}$ is given. Often we write

$$\begin{aligned}
\psi(\mathbf{Q}, \nabla\mathbf{Q}) &= \psi(\mathbf{Q}, 0) + (\psi(\mathbf{Q}, \nabla\mathbf{Q}) - \psi(\mathbf{Q}, 0)) \\
&= \psi_B(\mathbf{Q}) + \psi_E(\mathbf{Q}, \nabla\mathbf{Q}) \\
&= \text{bulk energy } + \text{ elastic energy.}
\end{aligned}$$

Relation between the theories.

Minimizers $\mathbf{Q}$ of $\psi_B(\mathbf{Q})$ are typically of the uniaxial form

$$\mathbf{Q}(\mathbf{n}) = s\left(\mathbf{n} \otimes \mathbf{n} - \frac{1}{3}\mathbf{1}\right) \quad \text{for some } \mathbf{n} \in S^2,$$

where $s > 0$ is constant. Then, under suitable assumptions on $\psi(\mathbf{Q}, \nabla\mathbf{Q})$ (frame indifference, quadratic dependence on $\nabla\mathbf{Q}$ ...) we have that

$$W(\mathbf{n}, \nabla\mathbf{n}) = \psi\left(\mathbf{Q}(\mathbf{n}), \nabla\mathbf{Q}(\mathbf{n})\right).$$

1.3 Common features and questions

So we see that the variational problems for nonlinear elasticity and nematic liquid crystals are very similar, and have the form

$$\text{Minimize } \; I(\mathbf{u}) = \int_\Omega f(x, \mathbf{u}(x), \nabla \mathbf{u}(x))\, dx$$

for some f, where $\mathbf{u} : \Omega \to \mathbb{R}^m$, subject to constraints and boundary conditions.

What are the generic questions to ask about such energy functionals?

1. Why minimize energy?
(This is a very deep question, to which a very rough answer might be that the Second Law of Thermodynamics implies that corresponding dynamic equations have a Lyapunov function similar to $I(\mathbf{u})$.)

2. What function space to use?
(*This is part of the model*, but what should govern its choice?)

3. Do (global, local) minimizers exist? If not, what happens to minimizing sequences?
(For example, the existence of minimizers for commonly used models of rubber can be proved, but for W corresponding to elastic crystals the minimum is in general not attained, leading to infinitely fine microstructures.)

4. Are minimizers smooth in x, or do they have singularities (defects)? What is the physical interpretation of any singularities?

5. How many solutions are there, and how do they depend on parameters and boundary conditions?
(Can we use bifurcation theory to illuminate this?) What other properties do solutions have, such as symmetries, topological properties, self-similarity ... ?

6. Do solutions to appropriate dynamic equations approach equilibrium solutions/ minimizers as time $t \to \infty$?

2 Review of weak convergence

(For the functional analysis aspects see Dunford & Schwartz [2].)

Let $(X, \|\cdot\|)$ be a real Banach space (i.e. a complete real normed linear space). The dual space X^* of X is the space of continuous linear maps $f : X \to \mathbb{R}$. With the norm

$$\|f\|_{X^*} = \sup_{\|u\| \le 1} |f(u)|$$

X^* is a Banach space.

Given $u \in X$, we can define $u^{**} \in X^{**} = (X^*)^*$ by

$$u^{**}(f) = f(u) \text{ for all } f \in X^*.$$

Now recall the following consequence of the Hahn-Banach theorem.

Proposition 2. *For every $u \neq 0$ in X there exists $f \in X^*$ with $\|f\|_{X^*} = 1$ and $f(u) = \|u\|$.*

Hence

$$\|u^{**}\|_{X^{**}} = \sup_{\|f\|_{X^*} \leq 1} |f(u)| = \|u\|.$$

Hence $T : u \mapsto u^{**}$ is an isometric isomorphism between X and its image $T(X) \subset X^{**}$.

X is to be *reflexive* if $T(X) = X^{**}$ (so that X^{**} can be identified with X).

Definition 1. *A sequence $u^{(j)}$ in X converges weakly to $u \in X$ (written $u^{(j)} \rightharpoonup u$) if $f(u^{(j)}) \to f(u)$ for all $f \in X^*$.*

A sequence $f^{(j)}$ in X^* converges weak to $f \in X^*$ (written $f^{(j)} \overset{*}{\rightharpoonup} f$) if $f^{(j)}(u) \to f(u)$ for all $u \in X$.*

Weak* limits are obviously unique, while weak limits are unique by Proposition 2. If X is reflexive then weak and weak* convergence coincide.

Example 1. $X = L^p(\Omega)$, $\Omega \subset \mathbb{R}^n$ open (or Lebesgue measurable), with $1 \leq p < \infty$.

If $1 \leq p < \infty$ then $L^p(\Omega)^* \cong L^{p'}(\Omega)$, $\frac{1}{p} + \frac{1}{p'} = 1$, so that given $f \in L^p(\Omega)^*$ there exists a unique $g \in L^{p'}(\Omega)$ with

$$f(u) = \int_\Omega ug\,dx \text{ for all } u \in L^p(\Omega).$$

(Thus $L^p(\Omega)$ is reflexive iff $1 < p < \infty$.)

Hence $u^{(j)} \rightharpoonup u$ in $L^p(\Omega)$, $1 \leq p < \infty$, iff

$$\int_\Omega u^{(j)} g\,dx \to \int_\Omega ug\,dx \text{ for all } g \in L^{p'}(\Omega).$$

Similarly, $u^{(j)} \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega)$ iff

$$\int_\Omega u^{(j)} g\,dx \to \int_\Omega ug\,dx \text{ for all } g \in L^1(\Omega).$$

2.1 Useful results concerning weak convergence

1. $u^{(j)} \to u$ strongly in X implies $u^{(j)} \rightharpoonup u$ in X.
(The converse is false: e.g. take $u^{(j)} = \omega_j$, $\{\omega_j\}$ an orthonormal basis of a Hilbert space H. Then for any $v \in H$ we have $\|v\|_H^2 = \sum_{j=1}^\infty (v, \omega_j)^2$ and so $\omega_j \rightharpoonup 0$ in H. But $\|\omega_j\| = 1$, and so $\omega_j \not\to 0$ strongly.)

2. If $u^{(j)} \rightharpoonup u$ in X then $\|u^{(j)}\|$ is bounded. Similarly, $f^{(j)} \overset{*}{\rightharpoonup} f$ in X^* implies $\|f\|_{X^*}$ bounded.
(Follows from uniform boundedness theorem.)

3. If $u^{(j)} \rightharpoonup u$ in X then $\|u\| \leq \liminf_{j\to\infty} \|u^{(j)}\|$.
(*Proof.* By Proposition 2 there exists $f \in X^*$ with $\|f\|_{X^*} = 1$ and $f(u) = \|u\|$. Therefore $\|u\| = \lim_{j\to\infty} f(u^{(j)}) \leq \liminf_{j\to\infty} \|f\|_{X^*}\|u^{(j)}\|$.)

4. (Mazur's theorem.) If $u^{(j)} \rightharpoonup u$ then there exists a sequence of finite convex combinations

$$v^{(k)} = \sum_{j=k}^{N_k} \lambda_j^{(k)} u^{(j)}, \;\; (\lambda_j^{(k)} \geq 0, \; \sum_{j=k}^{N_k} \lambda_j^{(k)} = 1)$$

converging strongly to u in X.

5. Let X be separable (that is X has a countable dense subset). Then if $\|f^{(j)}\|_{X^*} \leq M < \infty$ there exists a subsequence $f^{(j_k)} \overset{*}{\rightharpoonup} f$ for some $f \in X^*$.
(In particular if X is also reflexive then any bounded sequence in X has a weakly convergent subsequence.)

The following lemma will be useful.

Lemma 3. *Let $\Omega \subset \mathbb{R}^n$ be open, let $1 < p < \infty$, and suppose that $f^{(j)} \rightharpoonup f$ in $L^p(\Omega)$, $g^{(j)} \to g$ strongly in $L^{p'}(\Omega)$. Then $f^{(j)} g^{(j)} \rightharpoonup fg$ in $L^1(\Omega)$.*

Proof. Let $\psi \in L^\infty(\Omega)$. Then

$$\begin{aligned} \left| \int_\Omega (f^{(j)} g^{(j)} - fg)\psi \, dx \right| &= \left| \int_\Omega [(g^{(j)} - g) f^{(j)} + g(f^{(j)} - f)]\psi \, dx \right| \\ &\leq C\|g^{(j)} - g\|_{p'} \cdot \|f^{(j)}\|_p + \left| \int_\Omega (f^{(j)} - f) g\psi \, dx \right|, \end{aligned}$$

which tends to zero as $j \to \infty$. □

Exercise 4. Let $1 < p < \infty$ and $u^{(j)}, u \in L^p(\Omega)$. Show that $u^{(j)} \rightharpoonup u$ in $L^p(\Omega)$ iff

(i) $\|u^{(j)}\|_p$ is bounded, and

(ii) $⨍_B u^{(j)} \, dx \to ⨍_B u \, dx$ for all balls $B \subset \Omega$,

where $⨍_B v \, dx = \frac{1}{\mathcal{L}^n(B)} \int_B v \, dx$.

Thus weak convergence of bounded sequences in L^p is equivalent to the convergence of averages. It can also be thought of as convergence of measurements (since, for example, a probe designed to measure temperature at a point will in practice measure an average).

Weak convergence is of central importance for the study of nonlinear PDE because it is often possible to show that a sequence of approximate solutions $u^{(j)} : \Omega \to \mathbb{R}^m$ is bounded in some Banach space X (e.g. from energy estimates).

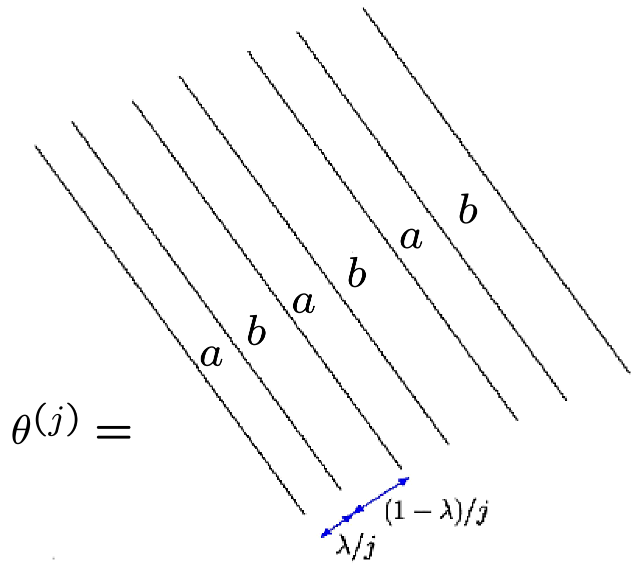

Figure 5: Laminate defined by Rademacher functions

Then if X is reflexive, say, we can extract a *weakly* convergent subsequence $u^{(j_k)} \rightharpoonup u$, whose weak limit u is a candidate solution. However to prove this we need to pass to the weak limit in the approximate equations (or energy functional).

2.2 Weak continuity and semicontinuity

Example 2. *Rademacher functions.*
Let $a, b \in \mathbb{R}^m$, $0 < \lambda < 1$, and define $\theta : \mathbb{R} \to \mathbb{R}^m$ by

$$\theta(t) = \begin{cases} a, & 0 \leq t < \lambda \\ b, & \lambda < t \leq 1 \end{cases}$$

extended to the whole of $\mathbb{R}$ as a function of period 1. For $N \in \mathbb{R}^n$, $|N| = 1$, let (see Fig. 5)

$$\theta^{(j)}(x) = \theta(jx \cdot N), \quad x \in \mathbb{R}^m. \tag{2.1}$$

Then (Exercise) $\theta^{(j)} \overset{*}{\rightharpoonup} \lambda a + (1-\lambda)b$ in $L^\infty(\mathbb{R}^n; \mathbb{R}^m)$.

Proposition 4. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be continuous, $\Omega \subset \mathbb{R}^n$ bounded open.*
(i) *The functional*

$$I(u) = \int_\Omega f(u)\, dx$$

is sequentially weak lower semicontinuous (sw*lsc) on* $L^\infty(\Omega;\mathbb{R}^m)$ *(i.e.* $u^{(j)} \overset{*}{\rightharpoonup} u$ *in* $L^\infty(\Omega;\mathbb{R}^m)$ *implies* $I(u) \leq \liminf_{j\to\infty} I(u^{(j)})$*) if and only if* f *is convex.*

(ii) The map $u \mapsto f(u)$ *is sequentially weak* continuous (i.e.* $u^{(j)} \overset{*}{\rightharpoonup} u$ *in* $L^\infty(\Omega;\mathbb{R}^m)$ *implies* $f(u^{(j)}) \overset{*}{\rightharpoonup} f(u)$ *in* $L^\infty(\Omega)$*) if and only if* f *is affine, i.e.* $f(u) = c\cdot u + d$ *for some* $c \in \mathbb{R}^m,\ d \in \mathbb{R}$*.*

Proof. (i) *Necessity.* We choose $u^{(j)} = \theta^{(j)}$. Then $u^{(j)} \overset{*}{\rightharpoonup} u = \lambda a + (1-\lambda)b$ in $L^\infty(\mathbb{R}^n;\mathbb{R}^m)$ (and hence in $L^\infty(\Omega;\mathbb{R}^m)$), and similarly $f(u^{(j)}) \overset{*}{\rightharpoonup} \lambda f(a) + (1-\lambda)f(b)$ in $L^\infty(\Omega)$. Therefore

$$f(\lambda a + (1-\lambda)b) = \frac{1}{\mathcal{L}^n(\Omega)} I(u) \leq \frac{1}{\mathcal{L}^n(\Omega)} \liminf_{j\to\infty} I(u^{(j)}) = \lambda f(a) + (1-\lambda)f(b).$$

Hence f is convex.

Sufficiency. If $f \in C^1$ this is easy, since

$$f(u^{(j)}) \geq f(u) + Df(u)\cdot(u^{(j)} - u),$$

and $\lim_{j\to\infty}\int_\Omega Df(u)\cdot(u^{(j)}-u)\,dx = 0$.

If f is just continuous, for $\varepsilon > 0$ we can let $f_\varepsilon(v) = \int_{\mathbb{R}^m}\rho_\varepsilon(v-z)f(z)\,dz$, where ρ_ε is a mollifier (i.e. $\rho_\varepsilon \geq 0$ is smooth with $\operatorname{supp}\rho_\varepsilon \subset B(0,\varepsilon)$ and $\int_{\mathbb{R}^m}\rho_\varepsilon(v)\,dv = 1$). Then f_ε is convex, since

$$\begin{aligned} f_\varepsilon(\lambda u_1 + (1-\lambda)u_2) &= \int_\Omega \rho_\varepsilon(z) f(\lambda u_1 + (1-\lambda)u_2 - z)\,dz \\ &\leq \lambda\int_\Omega \rho_\varepsilon(z)f(u_1 - z)\,dz + (1-\lambda)\int_\Omega \rho_\varepsilon(z)f(u_2-z)\,dz \\ &= \lambda f_\varepsilon(u_1) + (1-\lambda)f_\varepsilon(u_2), \end{aligned}$$

and $f_\varepsilon(v) \to f(v)$ as $\varepsilon \to 0$ uniformly for v in compact subsets of $\mathbb{R}^m$. Thus

$$\begin{aligned} \int_\Omega f_\varepsilon(u)\,dx &\leq \liminf_{j\to\infty}\int_\Omega f_\varepsilon(u^{(j)})\,dx \\ &\leq \liminf_{j\to\infty}\int_\Omega f(u^{(j)})\,dx + \sup_j \int_\Omega |f_\varepsilon(u^{(j)}) - f(u^{(j)})|\,dx, \end{aligned}$$

so that letting $\varepsilon \to 0$ we get $I(u) \leq \liminf_{j\to\infty} I(u^{(j)})$ by bounded convergence, using the fact that $|u^{(j)}(x)| \leq M < \infty$ for some M.

(Exercise. Give another proof using Mazur's theorem.)

(ii) If $u \mapsto f(u)$ is sequentially weak* continuous then similarly

$$f(\lambda a + (1-\lambda)b) = \lambda f(a) + (1-\lambda)f(b) \text{ for all } a, b \in \mathbb{R}^m,\ 0 \leq \lambda \leq 1,$$

which implies (Exercise) that f is affine. □

Thus we cannot pass to the weak limit in terms of the form $f(u)$ if f is nonlinear, unless we have further information on the sequence $u^{(j)}$.

3 Young measures

Useful references for the material in this and later sections are [3, 1].

The Young measure is a useful object which characterizes the weak limits of $f(u^j)$ for *all* continuous f for a sequence $u^{(j)} : \Omega \to \mathbb{R}^m$. An intuitive description is the following: Fix j, a point $x \in \Omega$ and a small radius $\delta > 0$. We behave

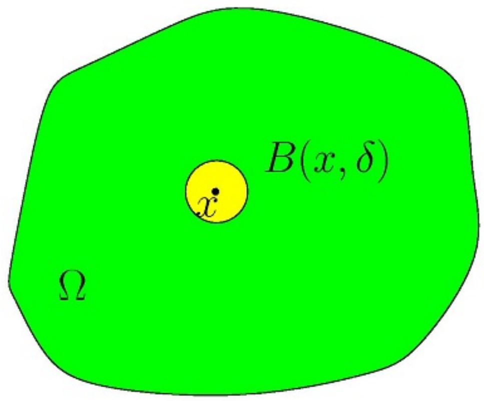

Figure 6: The Young measure $(\nu_x)_{x\in\Omega}$ of a sequence $u^{(j)}$ describes the asymptotic distribution of values of a sequence $u^{(j)}$ in a vanishingly small ball with centre x.

like a microscopist, examining the values of $u^{(j)}(z)$ for $z \in B(x, \delta)$ (see Fig. 6). Given an open subset $E \subset \mathbb{R}^m$ the probability that $u^{(j)}(z) \in E$ if z is chosen at random (uniformly distributed) from $B(x, \delta)$ is

$$\nu^{(j)}_{x,\delta}(E) = \frac{\mathcal{L}^n(\{z \in B(x, \delta) : u^{(j)}(z) \in E\})}{\mathcal{L}^n(B(x, \delta))}.$$

This defines a probability measure $\nu^{(j)}_{x,\delta}$ on $\mathbb{R}^m$. First we let $j \to \infty$, so that any rapid variations in $u^{(j)}$ are smeared out across the ball $B(x, \delta)$, and then $\delta \to 0+$, so that the probability is localized at x. We hope that this leads to a family of probability measures

$$\nu_x = \lim_{\delta\to 0} \lim_{j\to\infty} \nu^{(j)}_{x,\delta} \tag{3.1}$$

on $\mathbb{R}^m$ parametrized by $x \in \Omega$. In fact, given any sequence $u^{(j)}$ satisfying a mild bound, there is a subsequence (again denoted $u^{(j)}$) such that the double limit (3.1) exists for a.e. $x \in \Omega$, the convergence being in the sense of weak* convergence of probability measures. The family of measures $(\nu_x)_{x\in\Omega}$ is called the *Young measure* generated by $u^{(j)}$. A precise statement is the following:

Theorem 5. *Let $\Omega \subset \mathbb{R}^n$ be open (or just measurable), let $K \subset \mathbb{R}^m$ be closed, and let $u^{(j)}$ be a bounded sequence in $L^1(\Omega; \mathbb{R}^m)$ such that $u^{(j)} \to K$ in measure (that is*

$$\lim_{j\to\infty} \mathcal{L}^n(\{x \in \Omega : u^{(j)}(x) \notin U\}) = 0$$

for any open neighbourhood U of K in $\mathbb{R}^m$). Then there exists a subsequence $u^{(\mu)}$ of $u^{(j)}$ and a family of probability measures $(\nu_x)_{x\in\Omega}$, depending measurably on x, and satisfying

$$\operatorname{supp} \nu_x \subset K \quad \textit{for a.e. } x \in \Omega \tag{3.2}$$

such that

$$f(u^{(\mu)}) \rightharpoonup \langle \nu_x, f\rangle = \int_{\mathbb{R}^m} f(z)\, d\nu_x(z) \quad \textit{in } L^1(A) \tag{3.3}$$

whenever $A \subset \Omega$ is $\mathcal{L}^n$ measurable and $f(u^{(\mu)})$ is weakly relatively compact in $L^1(A)$.

Remarks 1. (a) A single mapping $u \in L^1(\Omega; \mathbb{R}^m)$ can be identified with the Young measure $\nu_x = \delta_{u(x)}$ (probability one of finding the value $u(x)$).

(b) The proof of Theorem 5 uses the identification in (a). Let $C_0(\mathbb{R}^m)$ denote the Banach space of continuous functions $f : \mathbb{R}^m \to \mathbb{R}$ such that $\lim_{|u|\to\infty} |f(u)| = 0$ with the norm $\|f\| = \sup_{u\in\mathbb{R}^m} |f(u)|$, whose dual $C_0(\mathbb{R}^m)^*$ can be identified with the space $M(\mathbb{R}^m)$ of bounded Radon measures on $\mathbb{R}^m$. The sequence $\delta_{u^{(j)}(\cdot)}$ is bounded in the space $L^\infty_w(\Omega; M(\mathbb{R}^m))$ of essentially bounded weak* measurable maps from Ω to $M(\mathbb{R}^m)$. Since $L^\infty_w(\Omega; M(\mathbb{R}^m)) = L^1(\Omega; C_0(\mathbb{R}^m))^*$ and $L^1(\Omega; C_0(\mathbb{R}^m))$ is separable, there exists a weak* convergent subsequence with limit $(\nu_x)_{x\in\Omega}$, which can be proved to have the required properties.

(c) If $u^{(j)} \rightharpoonup u$ in $L^1(\Omega; \mathbb{R}^m)$ then

$$u(x) = \bar{\nu}_x := \int_{\mathbb{R}^m} z\, d\nu_x(z) \quad \text{a.e. } x \in \Omega.$$

(d) If $u^{(j)}$ is bounded in $L^p(\Omega; \mathbb{R}^m)$, where Ω is bounded and $1 < p < \infty$, then it follows from Theorem 5 that $f(u^{(\mu)}) \rightharpoonup \langle \nu_x, f\rangle$ in $L^1(\Omega)$ if $|f(v)| \le C(|v|^q + 1)$, where $1 \le q < p$.

(e) Suppose $u^{(j)} \to u$ in $L^1(\Omega; \mathbb{R}^m)$ strongly. Then if $E \subset \Omega$ is bounded, $f(u^{(j)}) \to f(u)$ in $L^1(E)$ for any $f \in C_0(\mathbb{R}^m)$, so that

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

Conversely, if $u^{(j)}$ is bounded in $L^p(\Omega; \mathbb{R}^m)$, where Ω is bounded and $1 < p < \infty$, and if $u^{(j)}$ has Young measure $\nu_x = \delta_{u(x)}$, then $u^{(j)} \rightharpoonup u$ in $L^p(\Omega; \mathbb{R}^m)$, and by (d) if $1 \le q < p$ (taking $f(\cdot) = |\cdot|^q$)

$$\int_\Omega |u^{(j)}|^q dx \to \int_\Omega |u|^p dx,$$

so that $u^{(j)} \to u$ in $L^q(\Omega; \mathbb{R}^m)$.

Example 3. $u^{(j)}(x) = \theta(jx \cdot N)$ (see (2.1)) generates the Young measure

$$\nu_x = \lambda\delta_a + (1-\lambda)\delta_b,$$

since $f(u^{(j)}) \overset{*}{\rightharpoonup} \lambda f(a) + (1-\lambda)f(b) = \langle \lambda\delta_a + (1-\lambda)\delta_b, f\rangle$ in $L^\infty(\Omega;\mathbb{R}^m)$.

Exercise 5. (i) Let Q be an n-cube in $\mathbb{R}^n$, let $1 \leq p \leq \infty$, and let $u \in L^p_{\text{loc}}(\mathbb{R}^n)$ be Q-periodic. For $j = 1, 2\ldots$ define

$$u^{(j)}(x) = u(jx).$$

Prove that as $j \to \infty$

$$u^{(j)} \rightharpoonup \fint_Q u(y)\,dy \ \ (\overset{*}{\rightharpoonup} \ \text{ if } p = \infty)$$

in $L^p(\Omega)$ for every bounded open subset $\Omega \subset \mathbb{R}^n$.

(ii) Let $u^{(j)}(x) = \sin jx, \quad x \in (0,1)$. Show that $u^{(j)}$ generates the Young measure on $\mathbb{R}$

$$d\nu_x(z) = \chi_{[-1,1]}(z)\frac{dz}{\pi\sqrt{1-z^2}}.$$

4 Quasiconvexity

Let

$$I(u) = \int_\Omega f(x, u, \nabla u)\,dx, \tag{4.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain and f is continuous. We would like to prove the existence of a minimizer of I in

$$\mathcal{A} = \{v \in W^{1,1}(\Omega, \mathbb{R}^m) : v|_{\partial\Omega_1} = \bar{u}\},$$

where $\partial\Omega_1 \subset \partial\Omega$ and $\bar{u} : \partial\Omega_1 \to \mathbb{R}^m$. The idea of the direct method of the calculus of variations is the following. Suppose, for example, that f satisfies the coercivity condition

$$f(x, u, A) \geq c_0|A|^p - c_1 \quad \text{for all } x \in \Omega,\, u \in \mathbb{R}^m,\, A \in M^{m\times n},$$

for some $p > 1$ and constants $c_0 > 0$, c_1. Suppose also that $I(\tilde{v}) < \infty$ for some $\tilde{v} \in \mathcal{A}$. Let

$$l = \inf_{\mathcal{A}} I > -\infty.$$

Let $u^{(j)}$ be a minimizing sequence, i.e. $u^{(j)} \in \mathcal{A}$ for each j and $I(u^{(j)}) \to l$. Then $l < \infty$ and from coercivity we have that

$$\int_\Omega |\nabla u^{(j)}|^p dx \leq M < \infty.$$

Since Ω is connected, if $\mathcal{H}^{n-1}(\partial\Omega_1) > 0$ we have the Poincaré type inequality

$$\int_\Omega |v|^p dx \leq C\left(\int_\Omega |\nabla v|^p dx + \left|\int_{\partial\Omega_1} v\, d\mathcal{H}^{n-1}\right|^p\right)$$

for all $v \in W^{1,p}(\Omega;\mathbb{R}^m)$. Hence if $\bar{u}$ is bounded, say, then

$$\int_\Omega |u^{(j)}|^p dx \leq M_1 < \infty.$$

Therefore there exists a subsequence $u^{(\mu)}$ of $u^{(j)}$ with

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

and $u \in \mathcal{A}$ (since $u^{(\mu)}|_{\partial\Omega} \to u|_{\partial\Omega}$ in $L^p(\partial\Omega;\mathbb{R}^m)$ by trace theory).

We want to prove that u is a minimizer. This follows if I is sequentially weakly lower semicontinuous on $W^{1,p}$, i.e. $u^{(j)} \rightharpoonup u$ in $W^{1,p}$ implies

$$I(u) \leq \liminf_{j\to\infty} I(u^{(j)}).$$

Modulo technical assumptions Morrey ($\sim$ 1950) showed that I is swlsc on $W^{1,p}$ if and only if $f(x,u,\cdot)$ is *quasiconvex* for each x, u.

From now on we consider the case $f = f(\nabla u)$, which contains the essential difficulties.

4.1 Necessary conditions for lower semicontinuity

Let

$$I(u) = \int_\Omega f(\nabla u)\, dx,$$

where $f : M^{m\times n} \to [0,\infty]$ is Borel measurable.

Theorem 6. *If I is swlsc on $W^{1,p}$ (weak* if $p = \infty$) then*
(i) f is lower semicontinuous
(ii) for every n-cube Q

$$f\left(\fint_Q \nabla v(x)\, dx\right) \leq \fint_Q f(\nabla v(x))\, dx \tag{4.2}$$

for all $v \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with ∇v Q-periodic.

Proof. Let $A^{(j)} \to A$ in $M^{m\times n}$, $u^{(j)}(x) = Ax$. Then $u^{(j)} \to u$, $u(x) = Ax$, in $W^{1,p}$, and so

$$f(A) \leq \liminf_{j\to\infty} f(A^{(j)}).$$

(ii) Let $v \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with ∇v Q-periodic. Define

$$u^{(j)}(x) = \frac{1}{j}v(jx), \quad u(x) = \left(\fint_Q \nabla v(y)\, dy\right)x.$$

We claim that $u^{(j)} \rightharpoonup u$ in $W^{1,p}$ (weak* if $p=\infty$). To see this we can take $Q=(0,1)^n$ and let

$$z_i(x) = v(x+e_i) - v(x).$$

Since $\nabla z_i(x) = \nabla v(x+e_i) - \nabla v(x) = 0$, $z_i(x)$ is independent of x. Define $A \in M^{m\times n}$ by $Ae_i = z_i$, and let $w(x) = v(x) - Ax$. Then

$$w(x+e_i) - w(x) = z_i - Ae_i = 0,$$

and so w is Q-periodic, hence $⨍_Q \nabla w\, dx = 0$ (why?) and so $A = ⨍_Q \nabla v\, dx$. But $u^{(j)}(x) = Ax + \frac{1}{j} w(jx)$, $\nabla u^{(j)}(x) = \nabla w(jx)$. Hence $u^{(j)} \to Ax$ in L^p, and by Exercise 5 above $\nabla u^{(j)} \rightharpoonup ⨍_Q \nabla v(y)\, dy$ in L^p as claimed.

Hence

$$I(u) = \int_\Omega f\left(⨍_Q \nabla v\, dy\right) dx \le \liminf_{j\to\infty} \int_\Omega f(\nabla v(jx))\, dx.$$

If $f(\nabla v) \in L^1(Q)$ then again by Exercise 5 the RHS $= \int_\Omega \left(⨍_Q f(\nabla v)\, dy\right) dx$ and (4.2) follows. If $f(\nabla v) \notin L^1(Q)$ then (4.2) holds anyway. $\square$

From now on we assume for simplicity that $f : M^{m\times n} \to [0,\infty)$ is continuous.

Definition 2. *We say that f is* quasiconvex at the matrix A *if for some bounded open set $E \subset \mathbb{R}^n$*

$$f(A) \le ⨍_E f(A + \nabla\varphi(x))\, dx$$

for all $\varphi \in W_0^{1,\infty}(E;\mathbb{R}^m)$ and is quasiconvex *if it is quasiconvex at every A.*

Remark 1. The definitions are independent of E. Suppose that f is quasiconvex at A and that $E' \subset \mathbb{R}^n$ is bounded and open. Pick $a \in \mathbb{R}^n$ and $\lambda > 0$ such that $a + \lambda E' \subset E$. Given $\varphi \in W_0^{1,\infty}(E';\mathbb{R}^m)$ define $\psi : E \to \mathbb{R}^n$ by

$$\psi(x) = \begin{cases} \lambda\varphi(\frac{x-a}{\lambda}) & \text{if } x \in a + \lambda E', \\ 0 & \text{otherwise.} \end{cases}$$

Then $\psi \in W_0^{1,\infty}(E;\mathbb{R}^m)$ and hence

$$\begin{aligned} \mathcal{L}^n(E) f(A) &\le \int_E f(A + \nabla\psi(x))\, dx \\ &= \int_{E\setminus(a+\lambda E')} f(A)\, dx + \int_{a+\lambda E'} f(A + \nabla\varphi\left(\frac{x-a}{\lambda}\right))\, dx \\ &= [\mathcal{L}^n(E) - \lambda^n \mathcal{L}^n(E')] f(A) + \lambda^n \int_{E'} f(A + \nabla\varphi(x))\, dx, \end{aligned}$$

so that

$$f(A) \le ⨍_{E'} f(A + \nabla\varphi(x))\, dx$$

as required.

Lemma 7. *f is quasiconvex if and only if*

$$f\left(\fint_Q \nabla v\,dx\right) \leq \fint_Q f(\nabla v)\,dx \tag{4.3}$$

whenever $v \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with ∇v Q-periodic.

Proof. Suppose (4.3) holds, and let $A \in M^{m\times n}$, $\varphi \in W^{1,\infty}_0(Q;\mathbb{R}^m)$. Extend φ as a Q-periodic mapping on $\mathbb{R}^n$. Choosing $v = Ax + \varphi$, we deduce immediately that f is quasiconvex at A.

Conversely, suppose that f is quasiconvex, and let $v \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with ∇v Q-periodic, where without loss of generality we take $Q = (0,1)^n$. For $\varepsilon > 0$ small, let $\theta_\varepsilon : [0,1] \to \mathbb{R}$ be defined by

$$\theta_\varepsilon(t) = \begin{cases} \frac{t}{\varepsilon} & \text{if } t \in [0,\varepsilon], \\ 1 & \text{if } t \in (\varepsilon, 1-\varepsilon), \\ \frac{1-t}{\varepsilon} & \text{if } t \in [1-\varepsilon, 1], \end{cases}$$

and $\varphi_\varepsilon(x) = \Pi_{i=1}^n \theta_\varepsilon(x_i)$, so that $\varphi_\varepsilon \in W^{1,\infty}_0(Q)$ and $|\varphi_{\varepsilon,i}(x)| \leq \frac{1}{\varepsilon}$ for a.e. $x \in Q$. Let $A = \int_Q \nabla v\,dx$. As in the proof of Theorem 6 we have that $v - Ax$ is Q-periodic, so that $v - Ax$ is bounded on $\mathbb{R}^n$. For $x \in Q$ and $j \geq \frac{1}{\varepsilon}$ define

$$v_\varepsilon^{(j)}(x) = Ax + \frac{1}{j}\varphi_\varepsilon(x)(v(jx) - Ajx),$$

so that

$$\nabla v_\varepsilon^{(j)}(x) = A + \frac{1}{j}(v(jx) - Ajx) \otimes \nabla\varphi_\varepsilon(x) + \varphi_\varepsilon(x)(\nabla v(jx) - A) \quad \text{for a.e. } x \in Q.$$

Then $v_\varepsilon^{(j)} - Ax \in W^{1,\infty}_0(Q;\mathbb{R}^m)$ and so

$$\begin{aligned} f(A) &\leq \int_Q f(\nabla v_\varepsilon^{(j)}(x))\,dx \\ &= \int_Q f(\nabla v(jx))\,dx + \int_Q [f(\nabla v_\varepsilon^{(j)}(x)) - f(\nabla v(jx))]\,dx \\ &\leq \frac{1}{j^n}\int_{jQ} f(\nabla v(x'))\,dx' + C(1-(1-2\varepsilon)^n) \\ &= \int_Q f(\nabla v(x))\,dx + C(1-(1-2\varepsilon)^n), \end{aligned}$$

where C is a constant independent of j, ε. Letting $\varepsilon \to 0$ we obtain (4.3) as required. □

5 Rank-one convexity

Definition 3. *f is* rank-one convex *if*

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

for all $A, B \in M^{m\times n}$ with $\operatorname{rank}(A-B) = 1$, $\lambda \in (0,1)$.

(Note that $\operatorname{rank}(A-B)=1$ if and only if $A-B=a\otimes N$ for nonzero vectors $a\in\mathbb{R}^m, N\in\mathbb{R}^n$. The importance of rank-one matrices in this context is that a continuous $y:\mathbb{R}^n\to\mathbb{R}^m$ satisfies

$$\nabla y(x)=\begin{cases} A & \text{if } x\cdot N<k,\\ B & \text{if } x\cdot N>k,\end{cases}$$

for some unit vector $N\in\mathbb{R}^n$ if and only if $A-B=a\otimes N$ for some $a\in\mathbb{R}^m$ (the *Hadamard jump condition*).)

If $f\in C^2$ then f is rank-one convex if and only if

$$\frac{d^2}{dt^2}f(A+ta\otimes N)|_{t=0}=\frac{\partial^2 f(A)}{\partial A_{i\alpha}\partial A_{j\beta}}a_iN_\alpha a_jN_\beta\geq 0$$

for all $A\in M^{m\times n}, a\in\mathbb{R}^m, N\in\mathbb{R}^n$ (the *Legendre-Hadamard condition*).

Theorem 8. *If f is quasiconvex then f is rank-one convex.*

Proof. Let $A-B=a\otimes N$. We can without loss of generality suppose $N=e_1$. Define

$$v(x)=Bx+a\theta(x\cdot e_1),$$

where the derivative $\theta':\mathbb{R}\to\mathbb{R}$ of θ is the 1-periodic function which on $(0,1)$ equals the characteristic function of $(0,\lambda)$. Then $\nabla v=B+a\otimes\theta'(x\cdot e_1)e_1$ is Q-periodic with respect the cube $Q=(0,1)^n$ and thus by Theorem 7

$$f(\lambda A+(1-\lambda)B)=f\left(\int_Q\nabla v\,dx\right)\leq\int_Q f(\nabla v)\,dx=\lambda f(A)+(1-\lambda)f(B),$$

as required. □

Corollary 9. *If $n=1$ or $m=1$ then f is quasiconvex if and only if f is convex.*

Theorem 10 (van Hove)**.** *Let $f(A)=c_{ijkl}A_{ij}A_{kl}$ be quadratic. Then f is quasiconvex if and only if f is rank-one convex.*

Proof. Let f be rank-one convex. Since for any $\varphi\in W_0^{1,\infty}(E;\mathbb{R}^m)$

$$\int_E[f(A+\nabla\varphi)-f(A)]\,dx=\int_E c_{ijkl}\varphi_{i,j}\varphi_{k,l}\,dx$$

we just need to show that the RHS is ≥ 0. Extend φ by zero to the whole of $\mathbb{R}^n$ and take Fourier transforms. By the Plancherel formula

$$\int_E c_{ijkl}\varphi_{i,j}\varphi_{k,l}\,dx \quad=\quad 4\pi^2\int_{\mathbb{R}^n}\operatorname{Re}\left[c_{ijkl}\hat{\varphi}_i\xi_j\hat{\varphi}_k\xi_l\right]d\xi\geq 0.$$

□

6 Null Lagrangians

We ask when equality holds in the quasiconvexity condition, i.e. for what L is

$$\int_E L(A+\nabla\varphi(x))\,dx = \int_E L(A)\,dx$$

for all $\varphi \in W_0^{1,\infty}(E;\mathbb{R}^m)$? We call such L *quasi-affine.*

Theorem 11 (Landers, Morrey, Reshetnyak ...). *If $L : M^{m\times n} \to \mathbb{R}$ is continuous then the following are equivalent.*

(i) L is quasiaffine

(ii) L is a (smooth) *null Lagrangian, i.e. the Euler-Lagrange equations* $\operatorname{div} D_A L(\nabla u) = 0$ *hold for* all *smooth u.*

(iii) $L(A) = c_0 + \sum_{k=1}^{d(m,n)} c_k J_k(A)$,
where $\mathbf{J}(A) = (J_1(A), \dots, J_{d(m,n)}(A))$ is the list of all minors of A and the c_i are constants.
(*e.g.* $m = n = 3 : L(A) = \text{const.} + C\cdot A + D\cdot \operatorname{cof} A + e\det A$)

(iv) the map $u \mapsto L(\nabla u)$ is sequentially weakly continuous from $W^{1,p}(E;\mathbb{R}^m) \to L^1(E)$ for any bounded open $E \subset \mathbb{R}^n$ and sufficiently large p ($p > \min(m,n)$ will do)

Proof. (Sketch) $(i) \Rightarrow (iii)$ use L is rank-one affine (by Theorem 8)
$(iii) \Rightarrow (iv)$ Take e.g. $J(\nabla u) = u_{1,1}u_{2,2} - u_{1,2}u_{2,1}$. Then

$$J(\nabla u) = (u_1 u_{2,2})_{,1} - (u_1 u_{2,1})_{,2}$$

if u is smooth. So if $\varphi \in C_0^\infty(E)$

$$\int_E J(\nabla u^{(j)})\cdot\varphi\,dx = \int_E [u_1 u_{2,1}\varphi_{,2} - u_1 u_{2,2}\varphi_{,1}]\,dx,$$

and this is true for $u \in W^{1,2}(E;\mathbb{R}^m)$ by approximation. If $u^{(j)} \rightharpoonup u$ in $W^{1,p}(E;\mathbb{R}^m)$, $p > 2$, then

$$\begin{aligned}\int_E J(\nabla u^{(j)})\varphi\,dx &= \int_E [u_1^{(j)} u_{2,1}^{(j)}\varphi_{,2} - u_1^{(j)} u_{2,2}^{(j)}\varphi_{,1}]\,dx \\ &\to \int_E [u_1 u_{2,1}\varphi_{,2} - u_1 u_{2,2}\varphi_{,1}]\,dx \\ &= \int_E J(\nabla u)\varphi\,dx,\end{aligned}$$

since $u_{2,1}^{(j)} \rightharpoonup u_{2,1}$ in $L^p(E)$, $u_1^{(j)} \to u_1$ strongly in $L^p(E)$, where we have used the compactness of the embedding of $W^{1,p}(E)$ in $L^p(E)$ and Lemma 3. Since $J(\nabla u^{(j)})$ is bounded in $L^{\frac{p}{2}}(E)$, for some subsequence $J(\nabla u^{(j_k)}) \rightharpoonup \chi$ in $L^{\frac{p}{2}}(E)$. Thus

$$\int_E [\xi - J(\nabla u)]\varphi\,dx = 0 \text{ for all } \varphi \in C_0^\infty(E),$$

which implies by the fundamental lemma of the calculus of variations that $\xi = J(\nabla u)$. Since the weak limit is the same for every such subsequence the whole sequence converges.

$(i) \Leftrightarrow (ii)$

$$\frac{d}{dt}\int_E L(A+\nabla\varphi+t\nabla\psi)\,dx|_{t=0} = 0 \text{ for all } \varphi,\psi \in C_0^\infty(E)$$

holds if and only if

$$\int_E DL(A+\nabla\varphi)\cdot\nabla\psi\,dx = 0,$$

and we can set $\nabla u = A+\nabla\varphi$.

$(iv) \Rightarrow (i)$ as in Theorem 6 □

Definition 4. *f is* polyconvex *if there exists a convex function $g : \mathbb{R}^{d(m,n)} \to \mathbb{R}$ such that*

$$f(A) = g(\mathbf{J}(A)) \textit{ for all } A \in M^{m\times n},$$

e.g. $f(A) = g(A, \det A)$ if $m = n = 2$,
$f(A) = g(A, \operatorname{cof} A, \det A)$ if $m = n = 3$,
with g convex.

Theorem 12. *If f is polyconvex then f is quasiconvex.*

Proof.

$$\begin{aligned}
⨍_E f(A+\nabla\varphi) &= ⨍_E g(\mathbf{J}(A+\nabla\varphi(x))\,dx \\
&\geq g\left(⨍_E \mathbf{J}(A+\nabla\varphi)\,dx\right) \text{ by Jensen's inequality} \\
&= g(\mathbf{J}(A)) = f(A).
\end{aligned}$$

□

6.1 A connection with degree theory

A formula for the degree $d(u,p,\Omega)$ of a C^1 map $u : \bar{\Omega} \to \mathbb{R}^n$, where $\Omega \subset \mathbb{R}^n$ is bounded open, and $p \notin u(\partial\Omega)$, is

$$d(u,p,\Omega) = \int_\Omega \rho(u(x)) \det \nabla u(x)\,dx, \tag{6.1}$$

where $\rho \geq 0$ is smooth, $\int_{\mathbb{R}^n} \rho(v)\,dv = 1$, and $\operatorname{supp}\rho$ is a subset of the component of $\mathbb{R}^n \setminus \partial\Omega$ containing p. One important property of the degree is that it depends only on the values of u on $\partial\Omega$. This follows from (6.1) because $L(u, \nabla u) = \rho(u) \det \nabla u$ is a null Lagrangian, since

$$\begin{aligned}
\frac{\partial}{\partial x_\alpha}\left(\frac{\partial L}{\partial u_{i,\alpha}}\right) &= \frac{\partial}{\partial x_\alpha}\left(\rho(u)(\operatorname{cof}\nabla u)_{i\alpha}\right) \\
&= (\operatorname{cof}\nabla u)_{i\alpha}\rho_{,j}(u)u_{j,\alpha} \\
&= (\det \nabla u)\rho_{,i}(u) \\
&= \frac{\partial L}{\partial u_i}.
\end{aligned}$$

7 Examples and counterexamples

We have shown that

$$f \text{ convex } \Rightarrow f \text{ polyconvex } \Rightarrow f \text{ quasiconvex } \Rightarrow f \text{ rank-one convex}$$

The reverse implications are all false if $m > 1, n > 1$ except that *it is not known whether* f rank-one convex $\Leftrightarrow f$ quasiconvex *when* $n \geq m = 2$.

f polyconvex $\not\Rightarrow f$ convex since any 2×2 minor is polyconvex.

f quasiconvex $\not\Rightarrow f$ polyconvex (examples due to Dacorogna & Marcellini, Zhang.

if $m \geq 3, n \geq 2$ then f rank-one convex $\not\Rightarrow f$ quasiconvex (famous example of Šverák 1992).

So is there a tractable characterization of quasiconvexity? This is the main road-block of the subject. Using Šverák's counterexample, Kristensen (1999) proved that for $m \geq 3, n \geq 2$ there is no local condition equivalent to quasiconvexity (for example, no condition involving f and any number of its derivatives at an arbitrary matrix A). This might lead one to think that no characterization is possible. However Kristensen also proved that for $m \geq 2, n \geq 2$ polyconvexity is not a local condition. For example, one might comtemplate a characterization of the type

f quasiconvex $\Leftrightarrow f$ is the supremum of a family of special quasiconvex functions (including null Lagrangians).

8 Lower semicontinuity and existence

A typical lower semicontinuity theorem for quasiconvex integrands is

Theorem 13 (Morrey, Acerbi/Fusco, Marcellini . . .)**.** *Let* $f : M^{m\times n} \to \mathbb{R}$ *be quasiconvex and let* $1 \leq p \leq \infty$*. If* $p < \infty$ *assume that*

$$0 \leq f(A) \leq c(1 + |A|^p \text{ for all } A \in M^{m\times n}.$$

Then

$$I(u) = \int_\Omega f(\nabla u)\, dx$$

is swlsc (weak if $p = \infty$) on $W^{1,p}(\Omega; \mathbb{R}^m)$.*

This leads to existence theorems for minimizers of I subject to suitable boundary conditions. If we make the stronger assumption that f is polyconvex then we can handle the case when f takes the value $+\infty$, and thus prove existence results in nonlinear elasticity under the assumption (H2).

References

[1] B. Dacorogna. *Direct Methods in the Calculus of Variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[2] N. Dunford and J. T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[3] S. Müller. Variational methods for microstructure and phase transitions. In *Calculus of Variations and Geometric Evolution problems*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.