Calculus of Variations, Lecture 11, 4/17/2017

Continuation of discn abt relaxation + quasiconvexity: main goals are to

a) tell whether, for a given W, it is in need of relaxation (ie: is $QW = W$ or not?)

b) when $QW < W$, understand the character of the (optimal) oscillatory gradients in defn of QW

c) identify QW in some cases of practical interest

Sketch of answers to be presented below:

<u>about (a)</u>: most convenient (+ very powerful) test is provided by the "layering constrn": if W is lsc (ie $QW = W$) then W must be "rank-one convex"

$$W(\theta F_1 + (1-\theta) F_2) \le \theta W(F_1) + (1-\theta) W(F_2)$$
whenever $F_2 - F_1$ is rank-one.

This is <u>not</u> equivalent to quasiconvexity

(Sverak gave an example of a rank-one-convex fn that's not quasiconvex, Proc Roy Soc Edinburgh Sect A 120 (1992) 185-189.)

<u>about (b)</u>: layering or "multiscale layering" provides a convenient framework, though often (eg in the scalar case $u: \mathbb{R}^n \to \mathbb{R}$) there are natural constrns with less complexity. Key task, after guessing the answer, is to prove your guess is right. This amounts, more or less, to showing that the "proposed" QW is quasiconvex. Typically achieved by showing convexity or polyconvexity (though there are also other tricks)

<u>about (c)</u>: the simplest case is when $u: \mathbb{R}^n \to \mathbb{R}$ (scalar valued). Then QW is the "convexification of W".

Let's start with the "layering constrn":

<u>Layering lemma</u>: Suppose $F = \theta F_1 + (1-\theta) F_2$ with $0<\theta<1$ & $F_2 - F_1 = a \otimes n$ (rank one). Then $\exists\, u^\varepsilon$ with

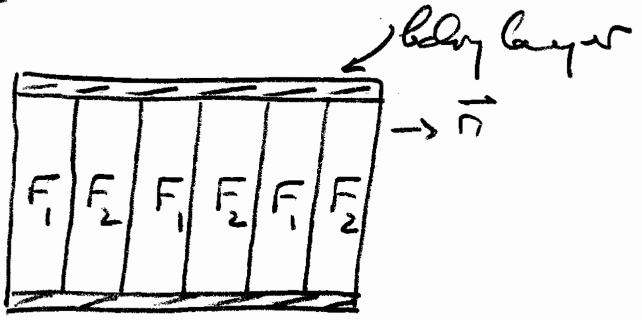
$$Du^\varepsilon = F_1 \text{ or } F_2 \quad \text{except on a set of measure} \to 0$$

with vol fr $\approx \theta$ of F_1, $(1-\theta)$ of F_2, and st

$$|Du^\varepsilon| \text{ unif bdd (indep of } \varepsilon)$$

$$u^\varepsilon = F \cdot x \quad \text{at bdry}$$

Proof: use layering with normal $\vec{n}$ + length scale $\varepsilon \to 0$, combined with a bdry layer of thickness about ε. Constrn can be done eg with u^ε piecewise linear



(Details left as exercise.)

Note that in scalar-valued setting F_1 + F_2 are vectors (so difference is automatically rank-one) and layering normal is parallel to $F_2 - F_1$.

Consequences of layering lemma:

1) $QW(\theta F_1 + (1-\theta)F_2) \le \theta W(F_1) + (1-\theta)W(F_2)$
if $F_2 - F_1$ is rank one

pf: obvious from defn of QW

2) $QW(\theta F_1 + (1-\theta)F_2) \le \theta QW(F_1) + (1-\theta)QW(F_2)$

1st pf: clear from fact that QW is lsc

2nd pf: use layering constrn, then replace piecewise linear test fn in each layer by a more oscillatory one, using defn of $QW(F_1)$ + $QW(F_2)$.

We're ready to prove that in the scalar-valued case, QW = convexn of W

Thm: when $u: \mathbb{R}^n \to \mathbb{R}$,

$$QW = \text{largest convex fn} \le W$$

Pf: Point 2 just above shows that QW is convex. But we can easily show that

$$\Phi \text{ convex} + \Phi \le W \Rightarrow \Phi \le QW$$

as follows: observe that $u = F\cdot x$ at ∂U $\Rightarrow$
$\frac{1}{|U|}\int_U Du = F$, so by Jensen's ineq

$$\Phi(F) = \inf_{u = F\cdot x \text{ at } \partial U} \frac{1}{|U|}\int_U \Phi(Du). \quad [\text{convexity}]$$

$$\leq \inf_{u=F\cdot x \text{ at } \partial U} \frac{1}{U}\int_U W(Du). \quad [\Phi \leq W].$$

$$= QW(F)$$

The proof just completed is simple, but it hides the assoc oscillations in ∇u.

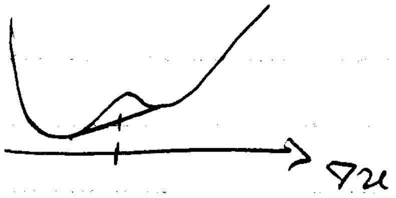
Let's make these more evident, ie let's discuss how given $F \in \mathbb{R}^n$, we can create u st

$$u\Big|_{\text{bdry}} = F\cdot x, \qquad \frac{1}{|U|}\int_U W(\nabla u) \approx \mathcal{C}W(F)$$

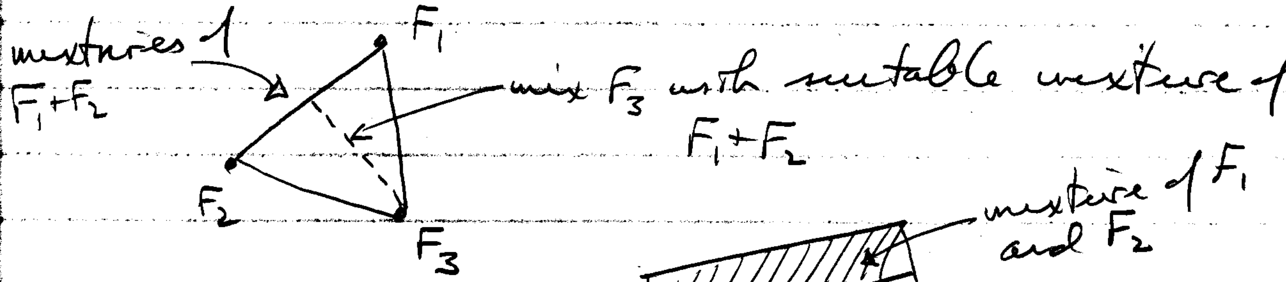
where $\mathcal{C}W$ = convexn of W.

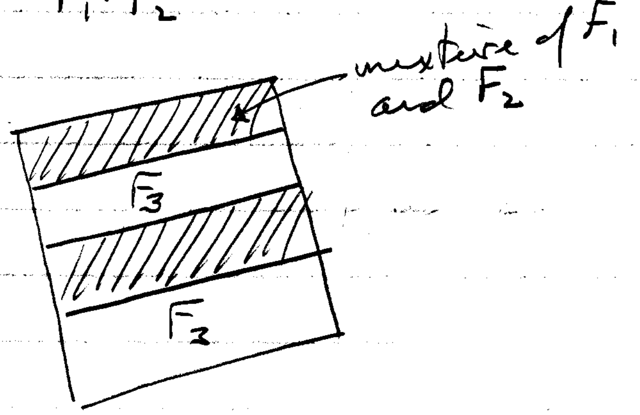
Start with observation that supergraph of $\mathcal{C}W$ is a convex set in $\mathbb{R}^{n+1}$ $\Rightarrow$ any pt on graph of $\mathcal{C}W$ is a convex hull of (finitely many!) extreme pts. Moreover the extreme pts are where $W = \mathcal{C}W$.

When this convex hull involves just 2 extreme pts, use layering constrn



When convex hull involves 3 extreme pts, use layering lemma twice

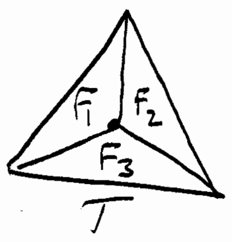




This can of course be iterated to handle any # of extreme pts.]

<u>Note:</u> when we want to mix just 2 gradients, layering is more or less the <u>only</u> constrn.

But to mix 3 gradients in $\mathbb{R}^2$, preceding "2-scale-layering" constrn is <u>not</u> the only possibility. For example, one can show that if the shape of the triangle T is chosen just right then $\exists\, u$ st

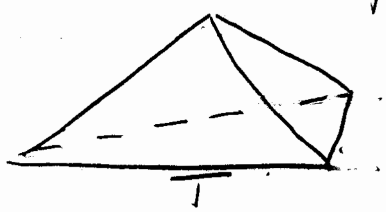


$\nabla u = F_j$ on subtriangles as shown

$u|_{\partial T} = F \cdot x$ where

$F = \text{avg of } \nabla u$
$= \theta_1 F_1 + \theta_2 F_2 + \theta_3 F_3$

(graph of u is a sort of pyramid).

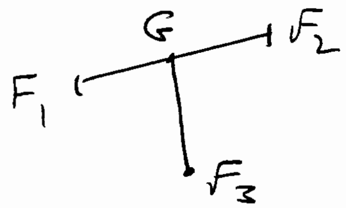


graph of u

By filling any domain with scaled copies of T, we get an example of u st $\nabla u = F_j$ on fraction θ_j whose oscillations have just one scale rather than two.

Another note: 2 scale layering (or higher-scale layering) provides a rich family of candidate oscillations also in the vector-valued case. For example:

if $\begin{cases} F_2 - F_1 \text{ is rank one} \\ G = \theta F_1 + (1-\theta) F_2 \\ F_3 - G \text{ is rank one} \end{cases}$



Then $$QW\left[\theta' F_3 + (1-\theta')(\theta F_1 + (1-\theta)F_2)\right] \le \theta' W(F_3) + (1-\theta')\theta W(F_1) + (1-\theta')(1-\theta) W(F_2).$$

All that remains: techniques for proving <u>lower bds</u> <u>on</u> <u>QW</u> in vector-valued settings. We need these

(a) to show, if it's true, that $QW = W$

(b) to show, when finding the relaxation, that a candidate test fn on defn of QW is in fact optimal.

We have already discussed two of the main tools: Jensen's ineq and polyconvexity. (There are also other techniques, but not many!)

<u>Rmk</u>: lsc of polyconvex fn is easy to prove using equivalence of lsc + quasiconvexity, as follows. Focusing for simplicity on $u: \mathbb{R}^2 \to \mathbb{R}^2$, suppose $W(Du) = f(Du, \det Du)$ with f convex as fn $\mathbb{R}^5 \to \mathbb{R}$. Observe that $u|_{\partial U} = F\cdot x \Rightarrow \frac{1}{|U|}\int_U \det Du = \det F$,

So Jensen's ineq gives, if $u|_{\partial U} = F \cdot x$,

$$\begin{aligned} W(F) &= f(F, \det F) \\ &\le \frac{1}{|U|} \int_U f(Du, \det Du)\, dx \\ &= \frac{1}{|U|} \int_U W(Du)\, dx \end{aligned}$$

Thus $QW = W$ when W is polyconvex.

Here's an alternative viewpt that's sometimes convenient: observe that

$$QW(F) = \min_{\substack{\mu = \text{distribution} \\ \text{measure of } Du \text{ for} \\ \text{some } u \text{ st } u|_{\partial U} = F \cdot x}} \int W(\lambda)\, d\mu(\lambda).$$

(the measure μ is the "Gradient Young measure" assoc to the test fn u) we see that looking for lower bds on $QW \Longleftrightarrow$ looking for possible restrictions on such distrn measures μ. In scalar setting there are <u>no restrns</u> other than the obvious one $\int \lambda\, d\mu(\lambda) = F$, but in vector-valued case we get restrns, eg in 2x2 case $\int \det \lambda \, d\mu = \det F$.

Focusing as usual on $u: \mathbb{R}^2 \to \mathbb{R}^2$, we have

a) $QW \geq$ [largest polyconvex fn $\leq W$]

b) $$QW \geq \min_{\substack{\int \lambda \, d\mu = F \\ \cancel{\int \det \lambda \, d\mu = \det F}}} \int W(\lambda) \, d\mu(\lambda).$$

Pf of (a) is parallel to pf in scalar setting that $QW \geq CW$

Pf of (b) is obvious, since enlarging the class of measures decreases the value of the min.

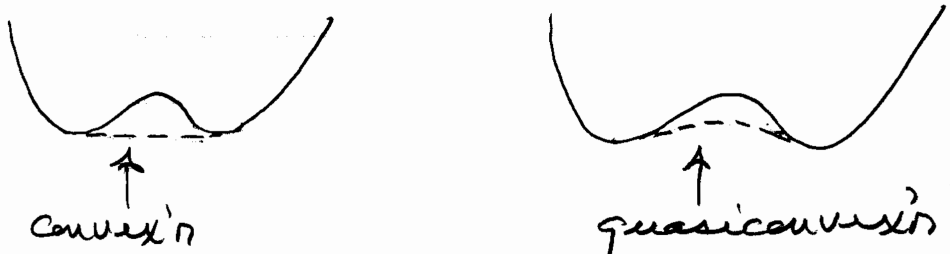
Actually: (a) + (b) are equivalent (ie the RHS of a = RHS of b); see N Firoozye, CPAM 44 (1991) 643-678

Lets use these tools to reach some concrete conclusions in a vector-valued setting

Example 1 Suppose A + B are 2×2 matrices with rank $(B-A)=2$. Then it is not possible that $Du = A$ or B (except perhaps in bdry layer of arb small measure), aside from the trivial cases $u = A\cdot x$ + $u = B\cdot x$

Rigorous version:

Thm: if $W \geq 0$ and $W=0$ at A + B only, then $QW=0$ only at A + B



For proof, use lower bd (b) above:

$$QW \geq \min_{\substack{\int \lambda\, d\mu = F \\ \int \det\lambda\, d\mu = \det F}} \int W(\lambda)\, d\mu .$$

where μ ranges over prob measures on 2×2 matrices. (RHS a linear opt'zn with convex constraints; we are assuming $W \to \infty$ at ∞, so min is achieved at some prob measure.)

Suff't to show $QW(\theta A + (1-\theta)B) > 0$ since at all other pts we have $QW \geq CW > 0$.

So: must show that if μ is admissible for preceding optz'n then it cannot be sptd only at A + B (assuming $0<\theta<1$).

If it were, then $\mu = \theta\delta_A + (1-\theta)\delta_B \Rightarrow$

$$\det(\theta A + (1-\theta)B) = \int \det\lambda \, d\mu$$

$$= \theta \det A + (1-\theta)\det B$$

To see this is possible it's convenient to focus on $A=0$, $B=I$ (the general case can be reduced to this one by change of vars) Then

$$\det(\theta A + (1-\theta)B) = (1-\theta)^2$$
$$\theta \det A + (1-\theta)\det B = 1-\theta$$

and $(1-\theta)^2 = 1-\theta$ only when $\theta = 0$ or 1.

<u>Example 2</u>: Motivated by our 2D discussion of "two martensite wells," let's ask: "what average gradients can be achieved by mixing two martensite phases, using only the

stress-free states (except perhaps for bdry layers)?

<u>Rigorous version</u>: for maps $u: \mathbb{R}^2 \to \mathbb{R}^2$, suppose $W(Du) \geq 0$ with $W = 0$ exactly on $SO(2)U_1 + SO(2)U_2$ with

$$U_1 = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \qquad U_2 = \begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix}$$

Where is $QW = 0$?

(Note: our earlier formulation with $W = 0$ on $SO(2)\begin{pmatrix} 1 & \pm\delta \\ 0 & 1 \end{pmatrix}$ is equivalent to this one, in a rotated coord system, if $\alpha\beta = 1$.)

<u>Answer</u>: $QW(F) = 0 \iff$ ① $\det F = \alpha\beta$

② $F^T F = \begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{pmatrix}$ satisfies

$$D_{11} + D_{22} + 2D_{12} \leq \alpha^2 + \beta^2$$
$$D_{11} + D_{22} - 2D_{12} \leq \alpha^2 + \beta^2$$

Pf that $QW(F) = 0 \Rightarrow$ ① + ②: consider the Young measure μ assoc with a min sequence in the defn of $QW(F)$. It has spt on

$SO(2)U_1 \cup SO(2)U_2$ and $\int \lambda \, d\mu(\lambda) = F$, $\int \det \lambda \, d\mu(\lambda) = \det F$.

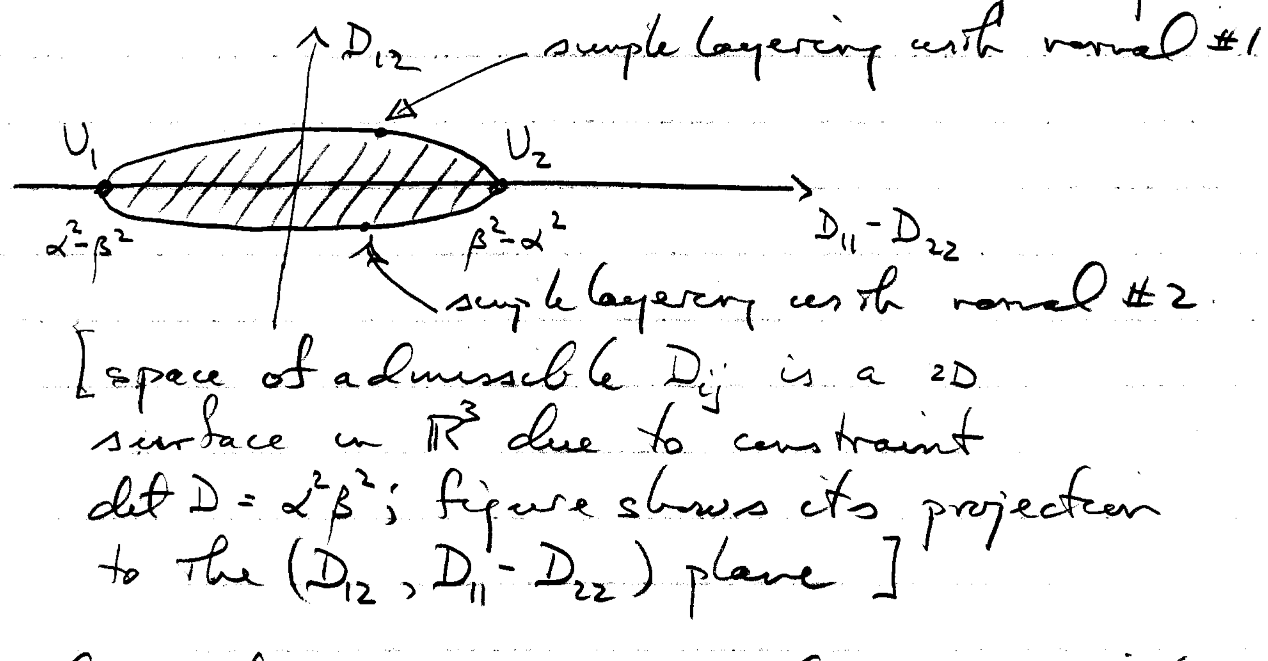
Property ① is clear, since $\det \lambda \equiv \alpha\beta$ on spt μ.

Property ② is elementary: for any vector $\vec{e}$,

$$|Fe|^2 = \langle F^T F e, e \rangle \leq \max_{i=1,2} |U_i e|^2$$

Apply this to $e = (1,1)$ + $e = (1,-1)$ to get the two parts of ②.

Sketch of pf that ① + ② $\Rightarrow$ $QW(F) = 0$: successful constrn uses "rank two layering".



[space of admissible D_{ij} is a 2D surface in $\mathbb{R}^3$ due to constraint $\det D = \alpha^2\beta^2$; figure shows its projection to the $(D_{12}, D_{11} - D_{22})$ plane]

Simple lamination of the 2 wells is possible in two distinct ways (ie $R_1 U_1 - R_2 U_2 = a \otimes n$ has solns with two distinct choices of $\bar{n}$); these laminates give the upper + lower boundaries in the figure (as the vol fractions of the twins vary from 0 to 1).

Corresponding pts on upper + lower bdries are rank-one related, so we can get segment joining them by rank 2 layering.

For more detail see eg Bhattacharya's book. [Why did this work? No idea! In general there is no assurance that using cont'y of det is suff't to prove sharp lower bd, or that multirank layering is suff't to construct optimal microstructure].

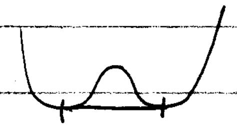
Example 3: We discussed thin elastic sheets as an example where a physically relevant functional may be not-lsc. In this setting $u: \Omega \to \mathbb{R}^3$ where $\Omega \subset \mathbb{R}^2$ is the region occupied by the (flat, stress-free) sheet in the absence of loading. The "membrane energy" $W(Du)$ should prefer isometry. Let's consider the simple example

$$W(Du) = |Du^T Du - I|^2 = (\lambda_1^2 - 1)^2 + (\lambda_2^2 - 1)^2$$

where λ_1, λ_2 are the "principal stretches" (the eigenvalues of $(Du^T Du)^{1/2}$). In this case

$$QW(Du) = (\lambda_1^2-1)_+^2 + (\lambda_2^2-1)_+^2$$

where $(t^2-1)_+^2 = \begin{cases} 0 & |t|<1 \\ (t^2-1)^2 & |t|>1 \end{cases}$



is the convexification of $(t^2-1)^2$.

Sketch: To show $QW(F) \le (\lambda_1^2-1)_+^2 + (\lambda_2^2-1)_+^2$ we must specify, for given F (with prin stretches λ_1, λ_2) a "wrinkling pattern" whose energy is given by the RHS (and whose energy is $F \cdot x$). It suffices to consider $F = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ 0 & 0 \end{pmatrix}$.

- if $\lambda_2 > 1$ but $\lambda_1 < 1$, use the 1D wrinkling picture discussed in Lecture 9
- if $\lambda_1 > 1$ but $\lambda_2 < 1$, same situation.
- if $\lambda_1 < 1$ and $\lambda_2 < 1$, use 2-scale wrinkling to show bc $F \cdot x$ is achievable with almost no membrane energy. (Note: other constrns are possible too, eg a piecewise linear origami-like "Miura-ori" pattern)
- if $\lambda_1 \ge 1$, $\lambda_2 \ge 1$ there's nothing to show.

To show $QW(F) \ge (\lambda_1^2-1)_+^2 + (\lambda_2^2-1)^2$ we must show that $F \to (\lambda_1^2-1)_+^2 + (\lambda_2^2-1)_+^2$ is quasiconvex. In fact it is convex. This can be seen using a more general result (this fnl is symmetric, & convex and increasing in each λ_i). Or it can be proved by using an explicit charzn of convexity for fns of a 3×2 matrix F that are symmetric fns of its prin stretches (see Pipkin, IMA J Appl Math 36, 1986, 85-99)

Suggested problems

(1) Show that

$$\min_{\int \lambda \, d\mu = F} \int W(\lambda)\, d\mu(\lambda)$$

is the largest convex function $\leq W$ (i.e. the "convexification" CW). Here μ ranges over probability measures.

(2) Here is another method for bounding QW from below:

$$QW \geq \mathcal{C}(W-g) + g$$

whenever g is quasiconvex. Prove it.

(3) Let a_1 & a_2 be a pair of 2×2 matrices, and consider the "two quadratic wells" energy

$$W(F) = \min\left\{ \tfrac{1}{2}|F-a_1|^2, \tfrac{1}{2}|F-a_2|^2 \right\}$$

where F is 2×2. Show that

$$(*) \quad QW(F) \geq \min_{0\leq\theta\leq1} \tfrac{1}{2}|F-\bar{a}(\theta)|^2 + \frac{\theta(1-\theta)}{2} h$$

where

$$\bar{a}(\theta) = \theta a_1 + (1-\theta) a_2$$

and

$$h = |a_1-a_2|^2 - \max_{|k|=1} |(a_1-a_2)k|^2$$

by applying pbm 2 with

$$g(F) = \tfrac{1}{2}|F|^2 - c\langle F, a_2-a_1\rangle^2$$

and c approaching

$$c_0 = \frac{1}{2}\left(\max_{|k|=1} |(a_2-a_1)k|^2\right)^{-1}$$

④ Show that the inequality (*) is actually an equality, i.e

$$QW(F) = \min_{0\le\theta\le 1} \frac{1}{2}|F-\bar{a}(\theta)|^2 + \frac{\theta(1-\theta)}{2} h$$

[Hint: an upper bound requires a construction. In this case an application of the layering lemma suffices!]