Calculus of Variations - Lecture 7 - 10/28/09

[Note: Lectures 5+6 were based on my typed notes "A brief introduction to optimal control".]

Today's topic: nonconvex variational problems, emphasizing nonlinear elasticity (where we expect existence) and examples from optimal design + phase transformation (where we don't).

Let's start with nonlinear elasticity. (Places to read more: Appendix A of Dacorogna's book, on reserve; or Ball's 1977 article in Arch Rat Mech Anal.)

Ω = undeformed body (assumed stress-free)

$x : \Omega \to \mathbb{R}^3$ deformation, takes $X \in \Omega$ to $x(X)$ = deformed position

$F = \frac{\partial x}{\partial X}$ $n \times n$ matrix ("deformation gradient")

elastic energy $= \int_\Omega W(F)\, dx$ where $W(RF) = W(F)$ for all $R \in SO(3)$,

$$= \int_\Omega \Phi\left((F^TF)^{1/2}\right) dx$$ where $(F^TF)^{1/2}$ is the "pos def symm" part of F

using polar decomposition $F = R(F^TF)^{1/2}$
rotn pos def symm

for any orientation-preserving F.

Condition $W(F) = W(RF)$ for $R \in SO(3)$ is "prin of frame indifference": rotn does no work. Further conds on W:

a) $F = I$ is a local min (so undeformed state is stable as well as stress-free)

b) $W \to \infty$ as $\det F \to 0$ (compression should be difficult) and also as $|F| \to \infty$ (extension should also be difficult)

c) if material is isotropic then $W(F) = W(FR)$ for all rotns R; corresp to $\Phi((F^TF)^{1/2})$ being a symmetric fn of the eigenvalues of $(F^TF)^{1/2}$ ("principal stretches").

Key pt for us: W cannot be a convex fn of F, since SO(3) is not convex.

In fact, expect W to be min on both

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad + \quad R_0 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

but $W = \infty$ on $\frac{1}{2} I + \frac{1}{2} R_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$!

So, for problems like this we need a larger class of fnls for which direct method of the calculus of variations applies.

For simplicity, let's think first in 2D (so $x(X) : \mathbb{R}^2 \to \mathbb{R}^2$ and F is 2×2). Everything said above still makes sense. Ball observed that the class of polyconvex fns is big enough to include reasonable "elastic energies" W(F), + still permits direct method to be used.

In 2D, a polyconvex fn is one of the form

$$W(F) = \varphi(F, \det F)$$ where $F = \frac{\partial x}{\partial X}$ is 2×2 + φ is a convex fn of its 5 arguments.

Obvious questions:

a) is this class really big enough to include reasonable elastic energies?

b) why is $E = \int W(F)$ lsc when W has this form (with suitable growth conds)?

c) how does this extend to 3D?

Short answer to (a): a simple isotropic law could use, in 2D,

$$W(F) = A(v_1^\alpha + v_2^\alpha - c) + g(\det F)$$

where $v_1 + v_2$ are eigenvalues of $(F^TF)^{1/2}$ (so $\det F = v_1 v_2$) and g is convex (value of c can be chosen so that $W = \min$ at $v_1 = v_2 = 1$). This W is polyconvex, due to

<u>Lemma</u>: a symmetric fn of $v_1 + v_2$ determines a <u>convex</u> fn of F if it is convex + nondecreasing in each v_i

(Pf of Lemma is nontrivial; see eg Ball's 1977 paper.)

In 3D, polyconvexity means $W(F) =$ convex fn of F, $\det F$, + 2x2 minors of F. Typical 3D isotropic polyconvex law: $W = A(v_1^\alpha + v_2^\alpha + v_3^\alpha - c) + B((v_1v_2)^\beta + (v_2v_3)^\beta + (v_3v_1)^\beta - d) + g(\det F)$.

Outline of answer to (b): in 2x2 setting,

det Dx is weakly cont's

under wk convergence in $W^{1,p}$ for $p>2$. (Note: for maps $\mathbb{R}^2 \to \mathbb{R}^2$, Dx is quadratic in $\partial x_i / \partial X_j$, so $\partial x/\partial X$ bdd in $W^{1,p} \Rightarrow \det Dx$ bdd in $L^{p/2}$. If $p>2$, $p/2>1$, + bdd sets are cpt under wk convergence. So we expect that if $Dx^n \to Dx^\infty$ in $W^{1,p}$ for $p>2$ then $\det Dx^n$ conv. wkly to something. Our assertion is that this wk limit is $\det Dx^\infty$.)

If we grant this, then lscy is clear:

1st pf: just use that convex fnl is lsc under wk convergence: so $F^{(n)} = \partial x^n/\partial X \to F^\infty = \partial x^\infty / \partial X$ and $\det F^n \to \det F^\infty$ (wkly L^p + $L^{p/2}$ resp) implies

$$\liminf_n \int_\Omega \varphi(F^n, \det F^n) \geq \int_\Omega \varphi(F^\infty, \det F^\infty)$$

for any convex φ.

2nd pf: use of Fenchel transform represents $\int \varphi(F, \det F)$ as a sup of linear (hence wkly cont's) fnls:

$$\varphi(F, \det F) = \sup_{G,t} G\cdot F + t\cdot \det F - \varphi^*(G,t)$$

$$\Rightarrow \int \varphi(Dx, \det Dx) = \sup_{\substack{G(x),\, t(x) \\ \text{with cpt spt}}} \int_\Omega G\cdot Dx + t\cdot \det Dx - \varphi^*(G,t)$$

$= \sup$ of wkly cont's fns

$\Rightarrow$ lsc

Wk cont'y of $\det Dx$ comes from the fact that

(*) $\underline{\underline{\det Dx \text{ can be expressed as a divergence}}}$

[How could we have guessed this was so? Well, note that $\int_\Omega \det Dx$ = area of image of Ω, for any map st $\det Dx > 0$; thus the integral depends only on x restricted to $\partial\Omega$. How else could this happen, but by Dx being a divergence? Note corollary to the preceding: EL eqn for $\int_\Omega \det Dx$ viewed as a var'l pblm must be an identity. Therefore one says "$\det Dx$ is a null-Lagrangian".]

Pf of (*) via exterior calculus:

$$d(x_1 \wedge dx_2) = dx_1 \wedge dx_2 = \det Dx \; dX_1 \wedge dX_2$$

Same proof, written out in coordinates:

$$\det Dx = \partial_1 x_1 \partial_2 x_2 - \partial_1 x_2 \partial_2 x_1$$

$$= \partial_1 (x_1 \partial_2 x_2) - \partial_2 (x_1 \partial_1 x_2)$$

To see why (*) implies wk convergence of $\det Dx^n$ to $\det Dx^\infty$ when $x^n \to x^\infty$ wkly $W^{1,p}$ $(p>2)$, it suffices to show wk convergence as distrns. So let φ be smooth with cpt spt in Ω. Then

$$\int_\Omega (\det Dx)\, \varphi \, dX = \int_\Omega -x_1 \partial_2 x_2 \partial_1 \varphi + x_1 \partial_1 x_2 \partial_2 \varphi$$

If $x^n \to x^\infty$ wkly $W^{1,p}$ then

$$\int_\Omega x_1^n \partial_2 x_2^n \partial_1 \varphi \to \int_\Omega x_1^\infty \partial_2 x^\infty \partial_1 \varphi$$

since $x_1^n \to x_1^\infty$ strongly in L^p (by Rellich)
$\partial_2 x_2^n \to \partial_2 x_2^\infty$ wkly in L^p (by hypothesis)

+ similarly for the other term.

Extnsn of above discn to 3×3 case: each 2×2 minor is wkly conts in $L^{p/2}$ under wk conv. in $W^{1,p}$; & 3×3 det is wkly conts in $L^{p/3}$ under wk conv. in $W^{1,p}$ (if $p>3$).

In elasticity we expected existence of minimizers. But there are also good reasons to consider multidim'l problems where existence is not

[go to pg 7]

expected, generalizing our 1D model $\int_0^1 (u_x^2-1)^2 + u^2$.

An example from optimal design: We discussed earlier this semester how to find the smallest t_0 st

$$\exists\ \sigma,\quad |\sigma| \le t_0,\quad \operatorname{div}\sigma = 0 \text{ in } \Omega, \qquad \sigma\cdot n = f \text{ at } \partial\Omega.$$

where $f : \partial\Omega \to \mathbb{R}$ is specified with avg value zero.

Suppose the smallest t_0 is larger than 1 - so there exist σ (probably lots of them) with

$$(*) \qquad |\sigma| \le 1,\quad \operatorname{div}\sigma = 0 \text{ in } \Omega, \qquad \sigma\cdot n = f \text{ at } \partial\Omega.$$

To make this a discn of gradients (not div-free vector fields) we can write $\sigma = (\partial_2 u, -\partial_1 u)$ in $\mathbb{R}^2$, recognize that $\sigma \cdot n = \partial_{\tan} u$ on $\partial\Omega$. So $(*)$ is equiv to

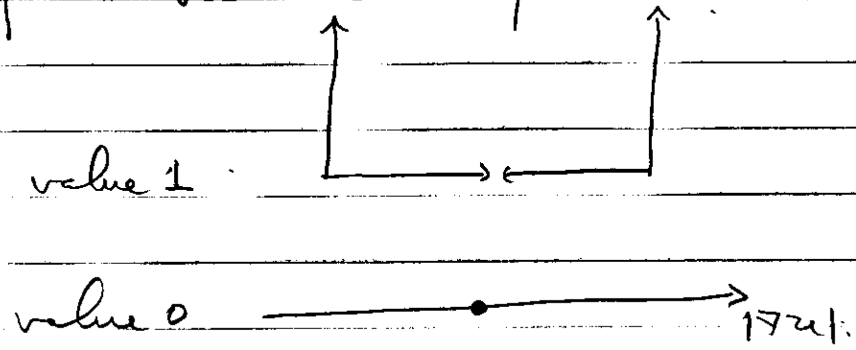
$$|\nabla u| \le 1 \text{ in } \Omega, \qquad u = g \text{ at } \partial\Omega$$

with g = 1st integral of f along $\partial\Omega$.

OK, here's the question: suppose we'd like furthermore to take $\sigma \equiv 0$ on some subset of Ω; how big can this subset be? (if σ represents a flow, we'd like to restrict the flow to a subset of Ω; how much area can we remove?) Answer: maximize set where $\sigma = 0$ $\iff$ min area where $\sigma \neq 0$ $\iff$

$$\min_{\substack{u=g \\ \text{at } \partial\Omega, \\ |\nabla u| \le 1 \text{ in } \Omega}} \int_\Omega 1_{\{\nabla u \neq 0\}}\, dx .$$

The integrand here is very nonconvex:



We'll show (in a week or two) that the answer is obtained by solving the relaxed problem whose integrand is the "convexn" of the preceding one

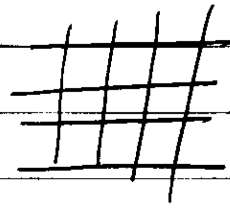
$$\min_{\substack{u=g \text{ at } \partial\Omega \\ |\nabla u| \le 1 \text{ in } \Omega}} \int_\Omega |\nabla u|\, dx$$

(This problem can be solved rather explicitly, by taking advantage of the coarea formula; see "The constrained least gradient problem" by RV Kohn + G. Strang in Knops + Lacey eds, Nonclassical Continuum Mechanics, CUP 1987, 226-243 + later more rigorous work by Sternberg, Williams, + Ziemer, Trans AMS 339 (1993) 403-432.) For a leisurely discn of both "unrelaxed" + "relaxed" problems, see "Fibered structures in optimal design" by R Kohn + G Strang, in Sleeman + Jarvis eds, Theory of Ordinary + Partial Diff'l Eqns, Longman 1986, posted with these notes.

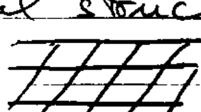
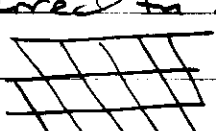
An example involving martensitic phase transfn

A "shape-memory material" is a crystalline solid with several possible crystal structures. Do discn in 2D so it's easy to draw pictures (truth is 3D but idea is similar)

cubic structure (preferred for $T > T_c$)



tetragonal structure (preferred for $T < T_c$)

 or 

sheared rel to cubic phase.

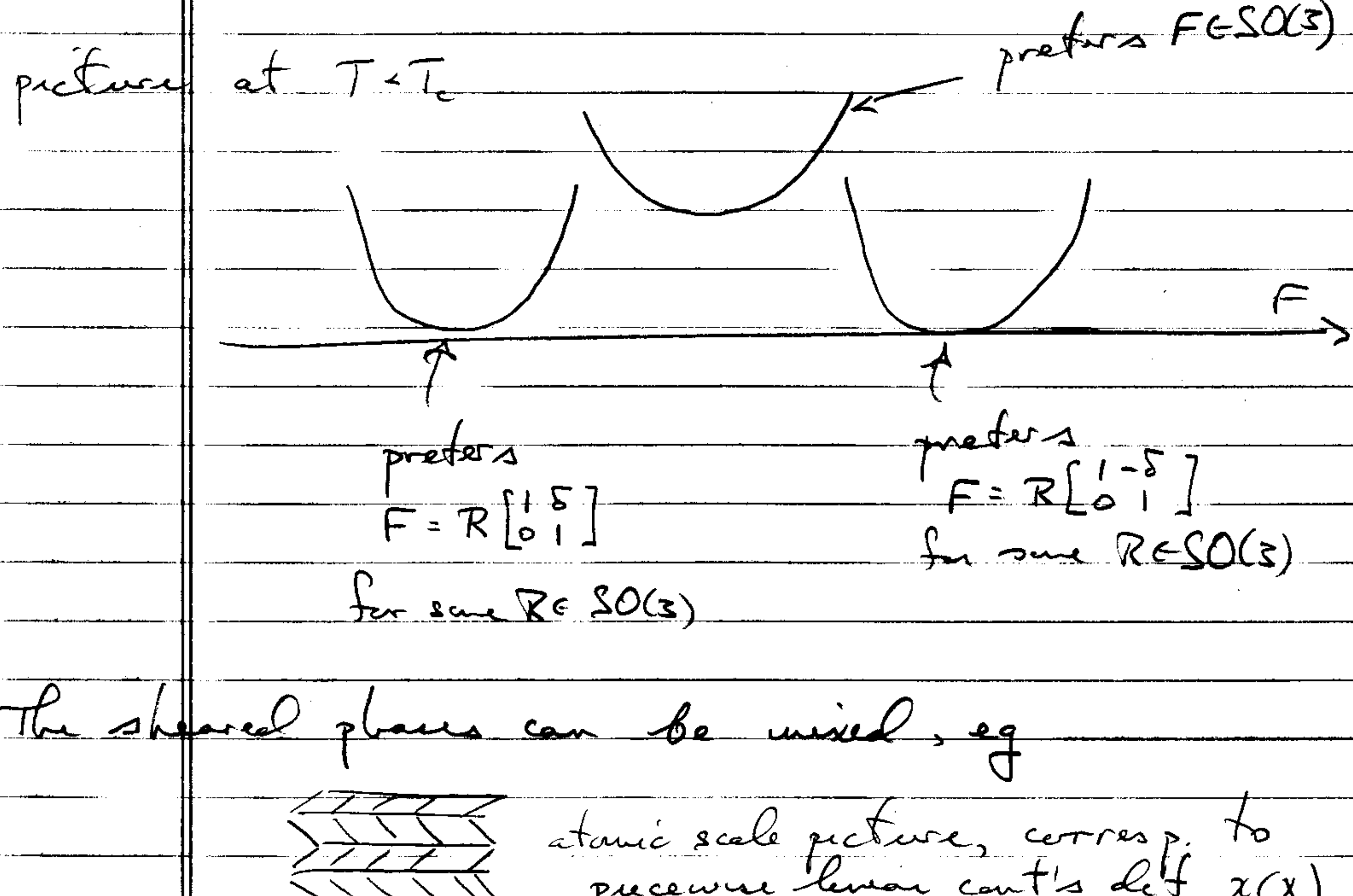
If we ignore issue of nucleation/motion of phase boundaries, we can suppose that material chooses the phase of lowest energy for a given strain. Thus

$$E = \int W(F) \qquad F = \frac{\partial x}{\partial X}$$

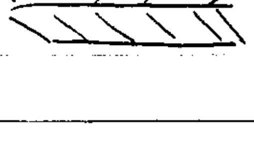
where

$$W = \min_{i=1,2,3} \{ W_i(F) \}.$$

picture at $T < T_c$



The sheared phases can be mixed, eg



atomic scale picture, corresp. to piecewise linear cont's def $x(X)$

st $F = \frac{\partial x}{\partial X}$ takes just 2 values,

$$\begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -\delta \\ 0 & 1 \end{bmatrix}$$

Avg F can be $\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}$ for any $|\lambda| < \delta$. So (as we'll see) the "relaxed" or "macroscopic" energy should be 0 along this line.

Such mixing (but in 3D, with more complex geometry, + in a polycrystal) is the essential mechanism behind the "shape-memory effect".

Key question: which avg defs can be achieved by mixing deformations in the "wells" $SO(2)\begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix} + SO(2)\begin{bmatrix} 1 & -\delta \\ 0 & 1 \end{bmatrix}$? Answer was obtained explicitly by Ball + James in their work "Proposed experimental tests..." Phil Trans: Phys Sci + Eng'g 338, 1992, 389-450. (see also more recent book by K. Bhattacharya for a good exposition)

One more example (mainly to discuss the limitations of this viewpoint): micromagnetics. Here

Ω = magnetic body $\subset \mathbb{R}^3$
$\vec{m}$ = normalized magnetization (a unit vector in $\mathbb{R}^3$)

$$E = \int_\Omega \varphi(m) + \int_\Omega \varepsilon^2 |\nabla m|^2 - \int_\Omega H \cdot m + \int_{\mathbb{R}^3} |\nabla u|^2$$

anisotropy, eg $m_1^2 + m_2^2$ (uniaxial)

exchange energy ($\varepsilon \sim 10$ nanometers)

H = constant (applied field)

nonlocal "magnetostatic" term
involves Helmholtz decompos $m\chi_\Omega = \nabla u +$ divfree,
favors $\operatorname{div} m = 0$ in Ω, $m \cdot \nu = 0$ at $\partial\Omega$

Obvious idea: since ε is small, why not ~~set~~ $\varepsilon = 0$?

$$\min_{|m|=1} \int_\Omega \varphi(m) - H \cdot m + \int_{\mathbb{R}^3} |\nabla u|^2$$

nonconvex due to constraint $|m| = 1$. (and also φ, if present)

nonlocal but convex

Viewpoint of relaxation is OK for finding min energy. But for $\varepsilon > 0$ fcl has lots of local min. Actual material configuration is very history-dependent (energy-min can't capture this)

Recent review that devotes a lot of attn (perhaps too much) to relaxed pbm is: "Recent developments in modeling, analysis, + numerics of ferromagnetism" by Kruzik + Prohl, SIAM Review 48 (2006) 439-483

What to do when pbm is not lsc (and we're interested in the min energy, not in possible local minima and/or effect of small regularizing term)?

Essential answer:

① identify the optimal oscillatory "solutions"

then

② use these to formulate a lsc "relaxed problem" for the "average" gradient.

In settings where lscy <=> convexity our understanding is pretty complete. Otherwise we have examples + a general theory, but no guaranteed-to-succeed algorithm.

Suggested problems:

(1) We showed that if $x^n \to x^\infty$ wkly $W^{1,p}$, $p>2$ (for maps $\mathbb{R}^2 \to \mathbb{R}^2$) then $\det Dx^n \to \det Dx^\infty$ wkly $L^{p/2}$. When $\{x^n\}$ are more singular strange things can happen, due to "cavitation." Explore this by considering the maps $x^n : B_1 \to B_1$ (where B_1 = unit ball in $\mathbb{R}^2$) such that

$$x^n(X) = \begin{cases} \dfrac{1-\rho_n}{\rho_n} X & \text{for } |X| \le \rho_n \\ \left(\dfrac{1-2\rho_n}{1-\rho_n} + \dfrac{\rho_n}{1-\rho_n}|X|\right)\dfrac{X}{|X|} & \text{for } \rho_n \le |X| < 1 \end{cases}$$

so x^n "blows up" B_{ρ_n} to $B_{1-\rho_n}$ and "squashes" $B_1 \setminus B_{\rho_n}$ to $B_1 \setminus B_{1-\rho_n}$. (Here ρ_n can be any sequence $\to 0$.)

a) show that $x^n(X) \to \dfrac{X}{|X|}$ ae

b) show that $\det Dx^n$ converges as a distribution to a pt mass located at 0

c) For which p does $\int_{B_1} |Dx^n|^p \, dX$ stay bounded as $n \to \infty$?

(2) Show that if $\int_\Omega g(\det F)$ is lower semicontinuous then g <u>must</u> be convex. (Hint: from the discussion at the end of the Lecture 4 notes, if $W(F) = g(\det F)$ is lower semicontinuous then it must be rank-one convex, i.e. the function

$$t \to g(\det(F + t a \otimes \nu))$$

should be convex in t, for every $\vec{a} + \vec{\nu} \in \mathbb{R}^n$. Show that this forces $g'' > 0$.)

(3) It can be nontrivial to determine whether a given function $W(F)$ is polyconvex or not. As an example, show that in the 2×2 setting

$$W(F) = \begin{cases} 1 + |F|^2 & \text{if } \rho(F) \geq 1 \\ 2\rho(F) - 2|\det F| & \text{if } \rho(F) \leq 1 \end{cases}$$

is polyconvex, when

$$\rho(F) = \left(|F|^2 + 2|\det F|\right)^{1/2}$$

(this example arose from the "relaxation" of a nonconvex problem, in my work with Strang in the 80's). Hint: show that $W(F) = g(F, \det F)$

where
$$g(F,t) = \max_{\alpha=\pm 1}\left\{ f\left(\left[|F|^2 + 2\alpha \det F\right]^{1/2}\right) - 2\alpha t \right\}.$$

in which
$$f(t) = \begin{cases} 1+t^2 & t \geq 1 \\ 2t & t \leq 1 \end{cases}.$$

Then check that $g(F,t)$ is a convex function of F and t.)

(4) Recall that there exists a cont's u with $Du = F_1$ or F_2 ae iff $F_1 - F_2 = a \otimes \nu$ (and if this happens then the successful construction uses "layers perp to ν," ie u is a function of $\vec{x}\cdot\vec{\nu}$).

In the example involving martensitic transformati[on] this layering is called "twinning," and the normals ($\vec{\nu}$) are readily observable.

In the simplified 2D, "two-well" setting of the notes, $\vec{\nu}$ is a possible twin normal iff $\exists\, R_1, R_2 \in SO(2)$ st.

$$R_1 \begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix} - R_2 \begin{bmatrix} 1 & -\delta \\ 0 & 1 \end{bmatrix} = a \otimes \nu \text{ for some } \vec{a} \in \mathbb{R}^2.$$

Show that (for fixed δ) there are just

7.17 (new)

two possible turning directions, + make them explicit in terms of δ.