Calculus of Variations, Lecture 1, 1/23/2017

[See syllabus for semester plan, etc]

Today's topic: "the direct method of the calculus of variations"

- review of essential idea, in familiar context of Laplace's eqn
- many examples of things that can go wrong
- proof that it works, under suitable hypotheses
- numerical approximation (minimization in a subspace)

Goal may be either

1) to minimize a functional that has intrinsic meaning (eg geodesics, minimal surfaces, etc)

or

2) to solve a pde by recognizing it as the EL eqn of a suitable var'l pbm
[eg $\Delta u + u^p = 0$ is EL eqn of $\int \frac{1}{2}|\nabla u|^2 - \frac{1}{p+1} u^{p+1}$; since fnl is nonconvex, solns of pde are crit pts, not necessarily local min (and never global min)]

A simple case should be familiar to all: if $\Omega \subset \mathbb{R}^n$ is bounded, then minimizer of

$$\min_{u=\varphi \text{ at } \partial\Omega} \int_\Omega \tfrac{1}{2}|\nabla u|^2 + fu \, dx$$

solves

$$-\Delta u + f = 0 \quad \text{in } \Omega$$
$$u = \varphi \quad \text{at } \partial\Omega$$

Reminder of why: if $E[u] = \int_\Omega \frac{1}{2}|\nabla u|^2 + fu \, dx$ then

$$1^{st} \text{ varn of } E = \frac{d}{dt}\Big|_{t=0} E(u+tv)$$
$$= \int_\Omega \langle \nabla u, \nabla v\rangle + fv \, dx$$

Dir bc requires that we use v st $v|_{\partial\Omega} = 0$.

1^{st} var'n $= 0$ at $E \iff u$ is a "weak solution" of $-\Delta u + f = 0$

To see that u is a classical soln (eg that $u \in C^{2,\alpha}$ if f is C^α) one must use pde regularity theory (the var'l principle doesn't easily give regularity).

Soln is unique due to strict convexity of E: if u solves EL eqn (ie it is a crit pt of E) and $u|_{\partial\Omega} = \varphi$, then for any w st $w|_{\partial\Omega} = \varphi$,

$E[w] = E[u] +$ 1st var'n at u in dir'n $w-u$
+ 2nd var'n term that's positive

which simplifies in this quadratic setting to

$$\int_\Omega \tfrac{1}{2}|\nabla w|^2 + fw = \int_\Omega \tfrac{1}{2}|\nabla u|^2 + fu$$
$$+ \int_\Omega \langle \nabla u, \nabla(w-u)\rangle + f(w-u)$$
$$+ \int_\Omega \tfrac{1}{2}|\nabla u - \nabla w|^2$$

Middle term on RHS vanishes if u is a crit pt (note that $w-u=0$ at $\partial\Omega$). So

$E[w] \geq E[u]$ with equality only when $\int |\nabla u - \nabla w|^2 = 0$, i.e. when $u = w$

(recall that $w-u=0$ at bdry, so $\nabla(u-w)=0 \Rightarrow u-w = \text{const} \Rightarrow u-w=0$).

In a basic pde class we often prove existence of weak sol'ns of 2nd order, linear pde using Lax-Milgram lemma. It has the advantage of applying to a fairly general class of pde's, including many that don't have variational principles. (For Laplace's eq'n you don't need

Lax-Milgram; The Riesz repr Thm is sufft.)
But var'l prin provides an alternative route.
Key advantage: it generalizes straightforwardly to nonlinear pde (non-quadratic var'l problems).

We'll discuss variational pf of existence later in this lecture.

Direct method looks easy, but there are subtleties. Let's identify some of them by giving examples where things go wrong

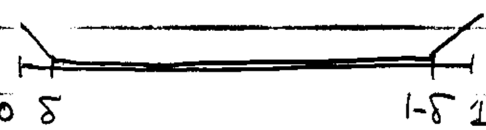
(A) In view of preceding discn abt var'l prin for $-\Delta u + f = 0$ with Dir bc, one might be tempted to solve same eqn with Neumann bc (say, $\partial u/\partial n = g$ at $\partial\Omega$) by

$$\min_{\frac{\partial u}{\partial n} = g \text{ at } \partial\Omega} \int_\Omega \tfrac{1}{2}|\nabla u|^2 + fu \, dx \qquad \boxed{\text{WRONG}}$$

but this doesn't work. Consider for example the 1D case with $g = 1$ and $f = 0$

$$\min_{\substack{u_x(0) = -1 \\ u_x(1) = +1}} \int_0^1 \tfrac{1}{2} u_x^2 \, dx$$

The min value is zero, and a minimizing sequence looks like



Limit of this seq is constant, so it solves $u_{xx}=0$ but boundary condition is lost. (Exercise: what happens when $f \neq 0$? What about space dim ≥ 2? Why can a Neumann cond be lost, but a Dir cond cannot be lost?)

Correct var'l prin for Neumann plm recovers bc as part of EL eqn

$$\min_{u} \int_{\Omega} \left(\tfrac{1}{2}|\nabla u|^2 + fu\right) dx - \int_{\partial\Omega} ug\, ds$$

since crit pt now satisfies

$$\int_{\Omega} \langle \nabla u, \nabla v \rangle + fv\, dx - \int_{\partial\Omega} gv\, ds = 0$$

for all $v \in H^1(\Omega)$. If u is smooth enough this gives

$$\int_{\partial\Omega} \left(\frac{\partial u}{\partial n} - g\right) v\, ds + \int_{\Omega} (-\Delta u + f)\, v\, dx = 0$$

for all v. Since v is <u>unrestricted</u> at $\partial\Omega$, we get both the pde and <u>the Neumann</u> bc.

(B) We must <u>take care that</u> min is not $-\infty$. For Neumann pblm discussed just above, we expect something to go wrong if f + g don't satisfy <u>consistency</u> condn

$$\int_{\partial\Omega} g\, ds = \int_\Omega f\, dx$$

(because $\Delta u = f$ in Ω, $\frac{\partial u}{\partial n} = g$ at $\partial\Omega$ $\Rightarrow \int_\Omega f = \int_\Omega \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial n} = \int_{\partial\Omega} g$)

In fact, if consistency fails then

$$\min_u \int_\Omega \left(\tfrac{1}{2}|\nabla u|^2 + fu\right) \underline{dx} - \int_{\partial\Omega} ug\, \underline{ds} = -\infty,$$

since we can take u = c = constant (use $c \to \infty$ or $-\infty$).

Pf that min is not $-\infty$ when data is consistent uses the inequality

(a) $\displaystyle\int_\Omega |\nabla u|^2 \geq C \int_\Omega u^2$ if $\displaystyle\int_\Omega u = 0$

(Note: best C is 1st nonzero eigenvalue of Neumann Laplacian on Ω.) We also need inequality

(b) $\|u\|_{L^2(\partial\Omega)} \le C \|u\|_{H^1(\Omega)}$

(Sketch of pf, assuming $\partial\Omega$ is smooth enough: suff't to show $\int_{\partial\Omega} u^2 ds \le C \int_\Omega u^2 + |\nabla u|^2$ when u is smooth. Let $\nu(x)$ be a smooth extension of the unit normal to $\partial\Omega$ (ν could vanish away from a nbhd of $\partial\Omega$). Then $\int_{\partial\Omega} u^2 = \int_{\partial\Omega} u^2(\nu \cdot n) = \int_\Omega \operatorname{div}(u^2 \nu) \le C \int_\Omega u^2 + |\nabla u|^2$.)

(Note: (b) says the map $H^1(\Omega) \to L^2(\partial\Omega)$ taking a fn to its "boundary trace" is cont's; actually we used it implicitly earlier, to know that $\int_{\partial\Omega} ug\, ds$ made sense.)

Let's prove now that if f + g are consistent then

$$\min_u \int_\Omega \left(\tfrac{1}{2}|\nabla u|^2 + fu\right) dx - \int_{\partial\Omega} ug\, ds \neq -\infty .$$

In fact, consistent data $\Rightarrow$ var'l prin is insensitive to replacement of u by u + const $\Rightarrow$ may restrict att'n to u st $\int_\Omega u = 0$. For such u, inequalities (a) + (b) give

$$\int_\Omega f u \, dx - \int_{\partial\Omega} g u \, ds \le C \|u\|_{H^1(\Omega)}$$
$$\le C' \left(\int_\Omega |\nabla u|^2\right)^{1/2}$$

(assuming $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$). So var'l prin is bounded below by

$$\min_z \tfrac{1}{2} z^2 - C' z$$

which is finite.

(C) In a nonlinear setting, even a Dir bc can be "lost" in the course of minimization, if the "energy" behaves like $|\nabla u|$ near ∞. Consider for example

$$\min_{u=\varphi \text{ at } \partial\Omega} \int_\Omega \left(1+|\nabla u|^2\right)^{1/2} dx$$

ie min (surface area of graph of u, for given bdry data).

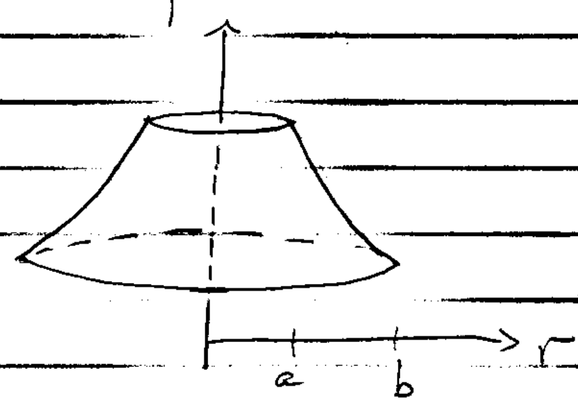
When $\Omega \subset \mathbb{R}^2$ is strictly convex the min is achieved (ie the bc is not lost), but proving this is nontrivial. More relevant today: when Ω is nonconvex the min might not be achieved. A simple example is the case

$$\Omega = \text{annulus } \{a < r < b\} \subset \mathbb{R}^2$$

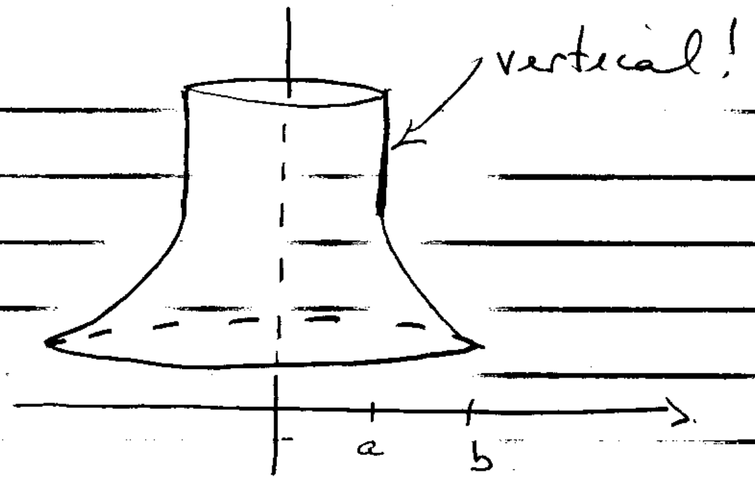
and bc is

$$\varphi = \begin{cases} A & \text{at } r = a \\ B & \text{at } r = b \end{cases}$$

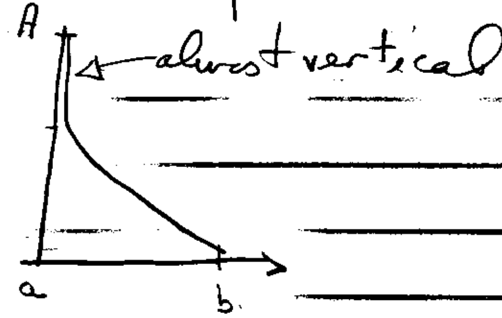
with $A - B$ suff'tly large. In fact: if $A > B$ but $A - B$ is not too large, sol'n is achieved



but when $A - B$ is large enough the area-minimizing surface has a vertical piece at the inner edge:



In latter case a min seq looks like



For a proof of these assertions (and even a formula for the solution) see pp 47-48 of the book by Buttazzo, Giaquinta, + Hildebrandt

Digesting this example a bit: The problem can be "fixed" by considering the "relaxed" problem

$$\min \int_{\Omega} (1+|\nabla u|^2)^{1/2} + \int_{\partial\Omega} |u-g|\, ds$$

which correctly accounts for the "cost" of getting the bc wrong (ie it correctly accounts for the surface area of any vertical parts above $\partial\Omega$). In calling this a "relaxed" problem, I mean

a) its min is achieved

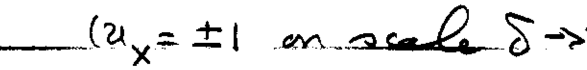
b) any min seq of the orig pblm converges to a minimizer of the relaxed pblm

c) the min values are the same for the original + relaxed pblms

Proving (a) - (c) is not trivial (but it's not so hard if you impose radial symmetry).

(D) If the functional is nonconvex, then a min sequence can easily develop oscillations. Its (weak) limit will not achieve the min. For example:

$$\min_u \int_0^1 (u_x^2-1)^2 + u^2 \, dx \quad = 0$$

but min is not achieved. A possible min sequence is 

($u_x = \pm 1$ on scale $\delta \to 0$).

That's all for examples where the method fails. But here are two additional subtleties:

(E) In nonconvex case we can get a "local min", but it's important to be careful about what "local" means. For example: consider

$$\min \int_0^1 (u_x^2-1)^2 \, dx \qquad \text{subject to } u(0)=u_0, \quad u(1)=u_1$$

and write $W(u_x) = (u_x^2-1)^2$  $W(u_x)$.

EL eqn is $(W'(u_x))_x = 0$. The linear function $u_*(x)$ with the given bc clearly solves the EL eqn. If in addition $W''(u_{*x}) > 0$ then 2nd varn is positive, since

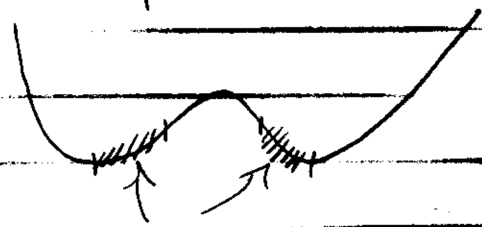
$$\frac{d^2}{dt^2}\Big|_{t=0} \int_0^1 W(u_{*x}+tv_x) = \frac{d}{dt}\Big|_{t=0} \int_0^1 W'(u_{*x}+tv_x)\, v_x$$

$$= \int_0^1 W''(u_{*x})\, v_x^2$$

which is strictly positive for $v(0)=v(1)=0$ (v must vanish at endpts since our pbm has a Dir bc). One can show that in fact, u_* is a <u>local</u>

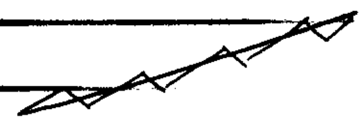
minimum in C^1 [when $u_{*x} = \text{const}$ + $W''(u_{*x}) > 0$].

However it is not a local min in L^∞, since an osc fn clearly does better [if $|u_{*x}| < 1$].



$W''(u_{*x}) > 0$

and $|u_{*x}| < 1$



$u_*(x)$ and an osc fn (with slope ± 1) that does better.

(F) Sometimes it matters what "space" we minimize over, even when we seek a global rather than local min. Example:

$$\min_{\substack{u(0)=0 \\ u(1)=1}} \int_0^1 (u^3 - x)^2 u_x^6 \, dx$$

Obviously, the fnl is ≥ 0 and it vanishes for $u(x) = x^{1/3}$. However, if u is restricted to the class of Lipschitz cont's fns then the min is bounded away from 0. For a proof, see §4.3 of Buttazzo-Giaquinta-Hildebrandt (this example is due to Manià + dates from the 1930's.)

Dependence of min value on the space where u resides is called "Lavrentiev phenomenon" since issue was 1st studied by Lavrentiev in 1920's. It arises when the minimizer (in a larger space) is singular + cannot be approx (with abt same energy) in the smaller space.

When minimizer is singular, we must also ask: does the minimizer satisfy the EL eqn? Not obvious, since $t \to E[u+tv]$ may not be diff'ble at $t=0$.

OK, after so many subtleties + failures one might ask: can we ever justify using the "direct method"? Answer is yes (with proper hypotheses). Moreover it's important, not only for existence but also for numerics, since minimization in a finite-dim'l space of fns is the most basic numerical scheme.

Concerning existence: lets discuss just the simple case

$$\min_{u=\varphi \text{ at } \partial\Omega} \int_\Omega W(\nabla u) + fu \, dx$$

where W is convex with "p^{th} power growth"

$$C_1(|\xi|^p - 1) \leq W(\xi) \leq C_2(|\xi|^p + 1)$$

where (crucially important!) $p > 1$. (Same ideas can be applied more generally, eg to $W(x, u, \nabla u)$, see eg Evans chapter 8 or Jost / Li-Jost chapter 4.)

Step 0: To be sure the functional is well-defined we must impose some restrictions on φ + f

For φ, let's just assume $\exists\, \Phi \in W^{1,p}(\Omega)$ st $\Phi|_{\partial\Omega} = \varphi$; then our plm can be written as

$$\min_{\tilde{u} = u - \Phi \in W_0^{1,p}(\Omega)} \int_\Omega W(\nabla\Phi + \nabla\tilde{u}) + f \cdot (\Phi + \tilde{u}) \, dx$$

where $W_0^{1,p}(\Omega)$ = closure of cptly sptd fns in $W^{1,p}$ norm

$= \{ u \in W^{1,p}(\Omega) \text{ st } u|_{\partial\Omega} = 0 \}$

For f, we need

$$\int_\Omega f u \le C \|u\|_{W^{1,p}(\Omega)}$$

ie $u \mapsto \int_\Omega f u$ must be a cont's fn'l on $W^{1,p}(\Omega)$.
It's suff't that $f \in L^q(\Omega)$, $\frac{1}{q} + \frac{1}{p} = 1$, since then

$$\int_\Omega f u \le C \|f\|_{L^q} \|u\|_{L^p} \le C \|f\|_{L^q} \|u\|_{W^{1,p}}$$

Step 1: Show the functional is bounded below on $W^{1,p}(\Omega) \cap \{u = \varphi \text{ at } \partial\Omega\}$. By hypothesis

$$\int_\Omega W(\nabla u) \ge C \int_\Omega |\nabla u|^p \, dx - \text{const}$$

But $u = \Phi + \tilde{u}$ with $\tilde{u} \in W_0^{1,p}(\Omega)$. Using the Poincaré-type inequality

$$\int_\Omega |\tilde{u}|^p \le C \int_\Omega |\nabla \tilde{u}|^p \quad \text{for } \tilde{u} \in W_0^{1,p}(\Omega)$$

we get

$$\begin{aligned}
\|u\|_{W^{1,p}} &\le \|\Phi\|_{W^{1,p}} + \|\tilde{u}\|_{W^{1,p}} \\
&\le C\left(\int_\Omega |\nabla \tilde{u}|^p\right)^{1/p} + \text{const} \\
&\le C\left(\int_\Omega |\nabla u|^p\right)^{1/p} + \text{new constant}
\end{aligned}$$

using triangle ineq for last step. So

$$\int_\Omega W(\nabla u) \geq C \, \|u\|_{W^{1,p}}^p - \text{const}$$

(of course my constants change from line to line)

Since $\min_{z \geq 0} c_1 z^p - c_2 z$ is finite, we see that the functional is bounded below. Moreover: for any μ, the set

$$\left\{ u : \int_\Omega W(\nabla u) + fu \, dx \leq \mu, \; u = \varphi \text{ at } \partial\Omega \right\}$$

is a bounded set in the $W^{1,p}(\Omega)$ norm. (This is an easy consequence of the preceding ests.)

<u>Step 2</u>: From step 1 we can take a minimizing sequence $\{u_k\}_{k=1}^\infty$ and any such sequence stays unif bounded in the $W^{1,p}$ norm.

Since $p>1$, unit ball of $W^{1,p}$ is compact under topology of weak convergence. So $\exists$ subseq $\{u_{k_j}\}$ converging weakly to a limit, call it u_*.

<u>Step 3</u>: We want to show that u_* is a

minimizer. Key point: a <u>closed</u>, <u>convex</u> subset of $W^{1,p}(\Omega)$ is closed under weak convergence. Apply this to

$$\left\{ u \in W^{1,p}(\Omega) \text{ s.t. } u = \varphi \text{ at } \partial\Omega \text{ and } \int_\Omega W(\nabla u) + f \cdot u \, dx \le m \right\}$$

for any $m >$ min value (this set is convex because we assumed W was convex).
Conclude:

$$\int W(\nabla u_*) + f u_* \le \liminf_{k_j} \int W(\nabla u_{k_j}) + f u_{k_j}$$

(ie our functional is <u>lower semicontinuous</u>).
Since u_{k_j} was (by hypoth) a minimizing sequence, u_* must <u>achieve</u> the min:

$$\int_\Omega W(\nabla u_*) + f u_* = \inf_{\substack{u \in W^{1,p}(\Omega) \\ u = \varphi \text{ at } \partial\Omega}} \int_\Omega W(\nabla u) + f \cdot u \, dx.$$

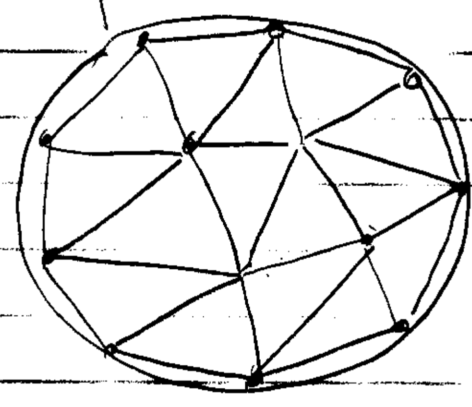
Preceding arg't may seem very abstract. But we saw from examples that its key ingredients ($p > 1$, W convex) are important, in sense that without them the method can fail.

Numerical perspective: We can expect to solve the pde (approximately) by minimizing our functional over a finite-dimensional approxn of $W^{1,p}(\Omega)$.

Obvious question: how well does this work? Let's focus for simplicity on the linear case $-\Delta u + f = 0$ in Ω, $u=0$ at $\partial\Omega$, ie

$$(*)_N \qquad \min_{u \in S_N} \int_\Omega \tfrac{1}{2}|\nabla u|^2 + fu \, dx$$

(Typical choice of S_N: piecewise linear, cont's functions on a triangulation of Ω. If all triangles are inside Ω then we can impose $u=0$ at $\partial\Omega$ by requiring $u=0$ at outer edge of triangulated region.)



eg Ω = circle
u is pl, conts on triangles, vanishes at nodes lying on $\partial\Omega$
(we view u as a fn on Ω via extnsn by 0)

Euler eqn of $(*)_N$ is

$$\int_\Omega \langle \nabla u_N, \nabla v \rangle + f v \, dx = 0$$

for all $v \in S_N$ (vanishing at outer bdry).

How small is $u - u_N$? Most basic result is

$$\int_\Omega |\nabla(u-u_N)|^2 = \min_{v \in S_N} \int_\Omega |\nabla(u-v)|^2$$

ie the approxn error measured in the natural norm is optimal - limited only by how well the pde soln can be approximated in S_N.

Proof: From Euler eqn $(*)_N$ + pde (solved by u) we have

$$\int \langle \nabla(u_N - u), \nabla v \rangle = 0 \quad \text{for } v \in S_N.$$

So

u_N = orthogonal projection of u onto S_N, using inner product

$$(w_1, w_2) = \int \langle \nabla w_1, \nabla w_2 \rangle$$

and we know that the orthog projn is the closest pt in S_N to u.

Success of this as a numerical method requires that

(a) u be well-approximated by fns in S_N

(b) finite dim'l pbm be easy to solve

About (a): for piecewise linear finite elements

$$\min_{v \in S_N} \int_\Omega |\nabla(u-v)|^2 \le h^2 \|D^2 u\|_{L^\infty}$$

is nearly obvious, where h = max diameter of triangles. (With some work, the L^∞ norm on RHS can be replaced by L^2)

About (b): we can either solve the var'l pbm (which is a quadratic + linear function of the nodal values of u) or else solve assoc lin system (eg by Gaussian elimination). Key advantage of using a triangulation: the linear system is sparse. (If we used, say, S_N = span of 1st N eigenfns of Dirichlet Laplacian then discrete pbm would be sparse in its quadratic term, though not its lin term. This too is workable.)

For more on such numerics, see a text on the Finite Element Method (My favorite is the one by Strang + Fix, which introduces all the main ideas in the 1st 50 pp.)

Suggested exercises:

(1) I gave two examples where the "direct method" fails because the boundary condition is lost in the limit (examples A + C). Which steps in our "proof" of the direct method fail in each case, and why?

(2) Does the variational problem

$$\min_{u=0 \text{ at } \partial\Omega} \int_\Omega \tfrac{1}{2}|\nabla u|^2 + fu \quad dx$$

achieve its minimum when $f = \delta_{x_0}$ for some $x_0 \in \Omega$ (so that $\int_\Omega fu = u(x_0)$)? Hint: the answer in R^n, $n \geq 2$ is different from dimension 1.

(3) What pde and boundary condition holds at critical points of

$$\int_\Omega \tfrac{1}{2}|\nabla u|^2 + u\cdot f\,dx \pm \beta\int_{\partial\Omega} u^2 \qquad ?$$

Can there be more than one critical point? (Hint: the sign of β matters.)

(4) Suppose W is strictly convex and C^2, and u is a classical solution of

$$\operatorname{div}\left(\frac{\partial W}{\partial \nabla u}(\nabla u)\right) = 0 \quad \text{in } \Omega$$

with bdry condition $u=\varphi$ at $\partial\Omega$. Show that u achieves

$$\min_{\substack{u=\varphi \\ \text{at } \partial\Omega}} \int_\Omega W(\nabla u)\,dx$$

and it is the unique (classical) minimizer.

(5) Consider $W(u_x) = (u_x^2-1)^2$, and suppose

$$W''(b-a) > 0$$

Show that the linear function $u_*(x) = a+(b-a)x$ is a C^1-local-minimizer of

$$\int_0^1 W(u_x)\,dx \quad \text{subject to } u(0)=a,\ u(1)=b,$$

in the sense that for any $v \in C^1(0,1)$ with the same bdry cond. + $\|v-u_*\|_{C^1}$ suffftly small we have

$$\int_0^1 W(v_x)\,dx > \int_0^1 W(u_{*x})\,dx.$$

(Hint: start by showing that the function

$$t \to \int_0^1 W(u_{*x}+t(v_x-u_{*x}))\,dx$$

is convex.)

(6) Let

$$E(u) = \int_\Omega \tfrac{1}{2}|\nabla u|^2 + f u \quad dx$$

and suppose u_* solves

$$\min_{u=0 \text{ at } \partial\Omega} E(u).$$

Show that if $\tilde{u}=0$ at $\partial\Omega$ and $\tilde{u}$ "almost" minimizes E in the sense that

$$E(\tilde{u}) \leq E(u_*) + \delta$$

then
$$\int_\Omega |\nabla u_\delta - \nabla \tilde{u}|^2 \le 2\delta .$$

(Food for thought: can you do something similar for $\int W(\nabla u) + fu$ if $W(\xi)$ strictly convex and
$$\sum_{i,j} \frac{\partial^2 W}{\partial \xi_i \partial \xi_j}(\xi)\, \eta_i \eta_j \ge c|\eta|^2 \quad \text{for all } \xi, \eta \in \mathbb{R}^n$$
for some $c > 0$?)