II. CALCULUS OF VARIATIONS

1. Euler-Lagrange equations.

We begin with an example.

Example: Let $D \subset \mathbb{R}^n$ be a bounded domain. Let $q(x)$ and $g(x)$ be given smooth functions on D and ∂D respectively. Among all functions $u(x)$ on D (which are sufficiently smooth – say, twice continuously differentiable) and satisfy $u(x) = g(x)$ for $x \in \partial D$, find the one which minimizes the "functional"

$$I(u) := \int_D \left(\frac{1}{2} |\nabla u(x)|^2 + q(x)u(x) \right) dx.$$

Well, suppose $I(u)$ is minimized by a function u^* (with $u^* = g$ on ∂D). Consider now a *one-parameter family* of functions nearby u^*:

$$u_\epsilon(x) := u^*(x) + \epsilon w(x)$$

where w is some fixed function. In order for u_ϵ to be in our allowed class of functions (twice continuously differentiable and equal to g on the boundary), we require that w also be twice continuously differentiable, and $w = 0$ on ∂D. Key observation:

$$u^* \text{ minimizes } I(u) \quad \Longrightarrow \quad I(u_\epsilon) \text{ is minimized at } \epsilon = 0.$$

Since $I(u_\epsilon)$ is a function of just the single variable ϵ, we are in the realm of simple calculus, and we know that at a minimum, the derivative is zero:

$$\begin{aligned}
0 = \frac{d}{d\epsilon} I(u_\epsilon)\Big|_{\epsilon=0} &= \frac{d}{d\epsilon} \int_D \left(\frac{1}{2} |\nabla u_\epsilon(x)|^2 + q(x) u_\epsilon(x) \right) dx\Big|_{\epsilon=0} \\
&= \int_D \left(\nabla u_\epsilon(x) \cdot \nabla \frac{\partial}{\partial \epsilon} u_\epsilon(x) + q(x) \frac{\partial}{\partial \epsilon} u_\epsilon(x) \right) dx\Big|_{\epsilon=0} \\
&= \int_D \left(\nabla u^*(x) \cdot \nabla w(x) + q(x) w(x) \right) dx \\
&= \int_D \left(\nabla \cdot [w(x) \nabla u^*(x)] + w(x)[-\Delta u^*(x) + q(x)] \right) dx \\
&= \int_D w(x) \left[-\Delta u^*(x) + q(x) \right] dx + \int_{\partial D} w(x) \frac{\partial}{\partial n} u^*(x) dS(x) \\
&= \int_D w(x)[-\Delta u^*(x) + q(x)] dx
\end{aligned}$$

where we used the divergence theorem toward the end, and the fact that $w = 0$ on ∂D (so the boundary term disappears).

To summarize: in order for u^* to minimize $I(u)$ (among sufficiently smooth functions with fixed boundary values $g(x)$), the integral above must be zero for **any** (sufficiently smooth) function $w(x)$ vanishing on the boundary ∂D. It turns out that the only way this can happen, is if the function multiplying w in the integral is zero – that is, $-\Delta u^*(x) + q(x) = 0$ – as the following lemma shows.

Lemma: if $g(x)$ is a continuous function on D, and if $\int_D g(x)w(x)dx = 0$ for *all* smooth functions w on D vanishing on the boundary, then $g(x) \equiv 0$ on D.

Proof: suppose, for some $x_0 \in D$, $g(x_0) \neq 0$. Since g is continuous, in some ball B around x_0, we have either $g(x) > 0$ or $g(x) < 0$. So let $w(x)$ be a little "bump function" at x_0: that is, a smooth, non-negative function, with $w(x_0) > 0$, and vanishing outside B. Then $0 = \int_D g(x)w(x)dx$ is contradicted. Hence we must have $g \equiv 0$ in D. $\square$

So, to conclude:

$$u^* \text{ minimizes } I(u) \text{ with fixed boundary data} \quad \Longrightarrow \quad \Delta u^*(x) = q(x) \text{ in } D.$$

That is, the PDE $\Delta u^* = q$ is a necessary condition – known as the **Euler-Lagrange equation** for I – for u^* to be a minimizer of I.

Remarks:

1. Notice that we have not touched the question of whether or not a minimizer u^* exsists, but merely written down a necessary condition that such a minimizer would have to satisfy. (This example, however, is simple enough that we can be sure of the existence of a minimizer – it is the solution of the Poisson equation $\Delta u^* = q$ in D, $u^* = g$ on ∂D.) The situation is completely analogous to finding the minimum of a function of several variables. There, the necessary condition is that the gradient vanish, which gives an algebraic equation to solve to find the "critical points". The question of which (if any) of these minimize the function is then addressed separately. The difference is that in the calculus of variations, we are trying to minimize a function defined on a (infinite-dimensional) space of functions, rather than (finite-dimensional) $\mathbb{R}^n$ – and as our example showed, the necessary condition may be not just an algebraic equation, but in fact a differential equation.

2. Indeed, the key point here is the relation between "variational problems" (minimizing functionals) and the (partial) differential equations which arise as their Euler-Lagrange equations. This idea is ubiquitous in physics (and many other applications) – think, for example, of the "principle of minimal action" in mechanics, and its relation to the equations of motion (Newton's equations). In practice, the relationship works both ways: we may try to solve a variational problem by solving the corresponding Euler-Lagrange equation; or, we may try

to solve a PDE by recognizing it as the E-L equation of some variational problem, which we then try to solve directly (we got a hint of this latter approach when we discussed variational problems for eigenvalues — more on this to come).

3. A natural question follows from the example above: suppose we indeed have a solution u^* of the problem

$$\begin{cases} \Delta u^* = q & \text{in } D \\ u^* = g & \text{on } \partial D \end{cases}.$$

Does it really minimize $I(u)$ among functions u with boundary values g? For this example, in fact, it does, since for any such u, using the divergence theorem,

$$\begin{aligned} I(u) - I(u^*) &= \int_D \left(\frac{1}{2}|\nabla u|^2 - \frac{1}{2}|\nabla u^*|^2 + q(u-u^*) \right) dx \\ &= \int_D \left(\frac{1}{2}|\nabla (u-u^*)|^2 + \nabla u^* \cdot \nabla (u-u^*) + q(u-u^*) \right) dx \\ &= \int_D \left(\frac{1}{2}|\nabla (u-u^*)|^2 + (u-u^*)(-\Delta u^* + q) \right) dx \\ &\quad + \int_{\partial D} (u-u^*) \frac{\partial u^*}{\partial n} dS(x) \\ &= \frac{1}{2} \int_D |\nabla (u-u^*)|^2 dx \geq 0. \end{aligned}$$

4. **Natural BCs**: consider the same minimization problem as above, except now we would like to minimize $I(u)$ among *all* (smooth) functions u (without the condition that $u = g$ on ∂D). What problem should a minimizer u^* solve in this case? Proceeding as above, we conclude that for $u_\varepsilon = u^* + \varepsilon w$, $\frac{d}{d\varepsilon} I(u_\varepsilon)|_{\varepsilon=0} = 0$ where w now is *any* smooth function (it does not have to vanish on ∂D). This leads, as above, to

$$0 = \int_D w(x) \left[-\Delta u^*(x) + q(x) \right] dx + \int_{\partial D} w(x) \frac{\partial}{\partial n} u^*(x) dS(x)$$

(the boundary term remains). In particular, this must hold for all w vanishing on the boundary, and so we conclude, as above, that the E-L equation $\Delta u^* = q$ must hold. Then in addition we require that $\int_{\partial D} w(x) \frac{\partial}{\partial n} u^*(x) dS(x) = 0$, again for *any* function w. This can only hold if $\frac{\partial u^*}{\partial n} = 0$ on ∂D. Thus

$$\begin{cases} \Delta u^* = q & \text{in } D \\ \frac{\partial u^*}{\partial n} = 0 & \text{on } \partial D \end{cases}.$$

These Neumann BCs are sometimes called the **natural boundary conditions** for the minimization problem, since they arise naturally from the minimiation when no BCs are imposed in the problem.

Our next example generalizes our first one.

Example: Find the Euler-Lagrange equation for the problem

$$\min_{u\in H} I(u), \quad I(u) := \int_D F(x,u,\nabla u)\,dx, \quad H = \{u \in C^2(D) \quad u = g \text{ on } \partial D\}$$

(notation: $C^2(D)$ denotes the twice continuously differentiable functions on D).

As above, a necessary condition for u to be a minimizer is, for *any* smooth function w vanishing on ∂D,

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon} I(u+\varepsilon w)\big|_{\varepsilon=0} = \frac{d}{d\varepsilon}\int_D F(x, u+\varepsilon w, \nabla u + \varepsilon \nabla w)dx\big|_{\varepsilon=0} \\ &= \int_D (F_u(x,u,\nabla u)w + F_{\nabla u}(x,u,\nabla u)\cdot\nabla w) \\ &= \int_D (-\nabla\cdot F_{\nabla u}(x,u,\nabla u) + F_u(x,u,\nabla u))\, w\, dx \end{aligned}$$

(where again we used the divergence theorem , and the fact that w vanishes on the boundary). So the Euler-Lagrange equation for this problem is

$$\nabla\cdot F_{\nabla u}(x,u,\nabla u) = F_u(x,u,\nabla u). \tag{39}$$

(Note that $F_{\nabla u}$ is a gradient – that is, it is a vector field.) As a special case, consider the previous example above, which corresponds to $F(x,u,\nabla u) = \frac{1}{2}|\nabla u|^2 + q(x)u$, so that $F_u = q$, $F_{\nabla u} = \nabla u$, $\nabla\cdot F_{\nabla u} = \Delta u$, and the Euler-Lagrange equation is $\Delta u = q$.

Remark: Notice that in general, the Euler-Lagrange equation (39) is a *nonlinear* PDE – whereas so far this course has been concerned almost exclusively with linear problems.

2. Some further examples.

Example: ("Brachistochrone problem") A ball rolls down a curve from a point P to a (lower) point Q. What curve gives the shortest travel time?

Put point P at the origin of an xy-plane (with the positive y direction pointing down, for convenience), and describe possible curves C from P to Q by functions $y = f(x)$, with Q at $(x_1, f(x_1))$ (draw a picture!). The speed of the ball along the curve is $v = ds/dt$ where s denotes the arc length along the curve (and t denotes time, of course), so the travel time is

$$T = \int_C dt = \int_C \frac{ds}{v}.$$

The arc length differential is given by

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{dx^2 + (f'(x)dx)^2} = \sqrt{1 + (f'(x))^2}dx.$$

We can determine the velocity by conservation of energy:

$$\text{energy} \;=\; \text{kinetic} \;+\; \text{potential} \;=\frac{1}{2}mv^2 - mgy = \;\text{constant}\; = 0,$$

where m is the mass of the ball, and g is the gravitational acceleration, so

$$v = \sqrt{2gy} = \sqrt{2gf(x)}.$$

So now we have an expression for the travel time

$$T = \int_C \frac{ds}{v} = \frac{1}{\sqrt{2g}} \int_0^{x_1} \sqrt{\frac{1 + (f'(x))^2}{f(x)}}dx.$$

So now our variational problem is clear: among smooth functions $f(x)$ defined on $[0, x_1]$ with boundary conditions $f(0) = 0$, $f(x_1) = y_1$, minimize

$$I(f) := \int_0^{x_1} F(f(x), f'(x))dx, \qquad F(f, f') := \sqrt{\frac{1 + (f')^2}{f}}.$$

We can read the Euler-Lagrange equation for this problem directly off (39):

$$\begin{aligned}
0 = F_f - \frac{d}{dx}F_{f'} &= \sqrt{1 + (f')^2}(-1/2)f^{-3/2} - \frac{d}{dx}\left[f^{-1/2}(1 + (f')^2)^{-1/2}f'\right] \\
&= \frac{1}{2(1 + (f')^2)^{3/2}f^{3/2}}\left(-(1 + (f')^2)^2 + (f')^2(1 + (f')^2) + 2f(f')^2f'' - 2f(1 + (f')^2)f''\right) \\
&= -\frac{1}{2(1 + (f')^2)^{3/2}f^{3/2}}\left((1 + (f')^2) + 2ff''\right).
\end{aligned}$$

So if there is a minimizing function $f(x)$, it should satisfy the nonlinear ODE

$$2ff'' + (f')^2 = -1.$$

It turns out that the *cycloid* (the curve traced out by a point on a circle rolling along a line), which is given parametrically by

$$x(\theta) = R(\theta - \sin\theta), \quad y(\theta) = R(1-\cos\theta),$$

solves this ODE, as we now verify:

$$f' = \frac{dy/d\theta}{dx/d\theta} = \frac{R\sin\theta}{R(1-\cos\theta)} = \frac{\sin\theta}{1-\cos\theta}$$

$$f'' = \frac{d}{dx}f' = \frac{[\sin\theta/(1-\cos\theta)]_\theta}{R(1-\cos\theta)} = \frac{(1-\cos\theta)\cos\theta - \sin^2\theta}{R(1-\cos\theta)^3} = -\frac{1}{R(1-\cos\theta)^2}$$

$$2ff'' + (f')^2 + 1 = \frac{-2(1-\cos\theta) + (1-\cos\theta)^2 + \sin^2\theta}{(1-\cos\theta)^2} = 0.$$

So the cycloid does indeed solve the E-L equation. Choosing R and θ_1 so that $(R(\theta_1 - \sin\theta_1), R(1-\cos\theta_1) = (x_1, y_1)$, we can also satisfy the boundary conditions. The question of whether or not the cycloid really minimizes the travel time is a more difficult one to answer.

Example: (Higher derivatives). Let $p(x)$ be a given function on a bounded domain D. Find the E-L equation for the minimization problem

$$min_{u\in C^4,\, u=\frac{\partial u}{\partial n}=0 \text{ on } \partial D} \int_D \left(\frac{1}{2}(\Delta u(x))^2 - p(x)u(x)\right) dx,$$

which arises in elasticity (where $u(x)$ represents the deflection of a plate $D \subset \mathbb{R}^2$ as a result of a load $p(x)$).

If u^* is a minimizer, and $w(x)$ is any smooth function with $w = \frac{\partial w}{\partial n} = 0$ on ∂D, then

$$0 = \frac{d}{d\varepsilon}\int_D \left(\frac{1}{2}(\Delta(u^* + \varepsilon w))^2 - p(u^* + \varepsilon w)\right) dx \,\Big|_{\varepsilon=0} = \int_D (\Delta u^* \Delta w - pw)\, dx$$

$$= \int_D (\Delta\Delta u^* - p) w dx + \int_{\partial D}\left(\Delta u^* \frac{\partial w}{\partial n} - \frac{\partial}{\partial n}\Delta u^*\, w\right) dS = \int_D (\Delta\Delta u^* - p) w dx$$

using the divergence theorem twice (or, if you prefer, Green's second identity) and the bounday conditions on w. Employing the same argument we used above, we conclude that the E-L equation for u^* is $\Delta^2 u^* = \Delta\Delta u^* = p$, sometimes known as the *biharmonic equation*. That is, the problem for a minimizer u^* is

$$\begin{cases} \Delta\Delta u^*(x) = p(x) & D \\ u^* = \frac{\partial u^*}{\partial n} = 0 & \partial D \end{cases}.$$

3. Variational problems with constraints.

Here we consider the **constrained variational problem**

$$\min_{u \in H,\, M(u)=C} I(u) \tag{40}$$

where

$$H = \{u \in C^2(D) \mid u \equiv 0 \text{ on } \partial D\}$$

is the class of functions we work in,

$$I(u) = \int_D F(x, u, \nabla u) dx$$

is the functional we are minimizing, and

$$M(u) = \int_D G(x, u, \nabla u) dx$$

is another functional, which gives the constraint.

Just as for functions of several variables, the necessary condition (Euler-Lagrange equation) for a constrained problem invloves **Lagrange mutipliers**:

Theorem: If u^* solves the variational problem (40), then u^* is a critical point (that is, satisfies the Euler-Lagrange equation for) of the functional

$$I(u) + \lambda M(u)$$

for some $\lambda \in \mathbb{R}$ (called a Lagrange multiplier).

Sketch of proof: Suppose $u^*(x)$ is a minimizer of problem (40). That means $M(u^*) = c$, and u^* minimizes $I(u)$ among functions $u \in H$ with $M(u) = c$. In particular, let u_ε be a one-parameter family of functions in H with $M(u_\varepsilon) = c$, and $u_0 = u^*$. Then $I(u_\varepsilon)$ (which is a function of ε) is minimized at $\varepsilon = 0$, and so

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon} I(u_\varepsilon)|_{\varepsilon=0} = \frac{d}{d\varepsilon} \int_D F(x, u_\varepsilon, \nabla u_\varepsilon) dx \,|_{\varepsilon=0} \\ &= \int_D \left[F_u(x, u^*, \nabla u^*) \frac{\partial u_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0} \; + \; F_{\nabla u}(x, u^*, \nabla u^*) \cdot \nabla \frac{\partial u_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0} \right] dx. \end{aligned}$$

Let us denote

$$\xi := \frac{\partial u_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0},$$

notice that $\xi \equiv 0$ on ∂D, and use (as always!) the divergence theorem, to get

$$0 = \int_D \left[F_u(x, u^*, \nabla u^*) - \nabla \cdot F_{\nabla u}(x, u^*, \nabla u^*) \right] \xi(x) dx.$$

Now if $\xi(x)$ could be *any* (smooth) function on D (with zero BCs), we would conclude, as before, that

$$f(x) := F_u(x, u^*, \nabla u^*) - \nabla \cdot F_{\nabla u}(x, u^*, \nabla u^*)$$

is zero, and that would be our E-L equation. However, because of the constraint $M(u_\varepsilon) = 0$, $\xi(x)$ *cannot* be just any function. Indeed, by the same sort of calculation we just did,

$$0 = \frac{d}{d\varepsilon} M(u_\varepsilon)|_{\varepsilon=0} = \int_D g(x)\xi(x)dx, \qquad g(x) := G_u(x, u^*, \nabla u^*) - \nabla \cdot G_{\nabla u}(x, u^*, \nabla u^*).$$

That is, ξ must be orthogonal to the function g. In fact, it turns out that a one-parameter family of functions u_ε as above, with $\frac{\partial u_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0} = \xi$, can be constructed for any (smooth enough) ξ satisfying $\int g\xi = 0$. So, we have

$$\int_D f(x)\xi(x)dx = 0 \;\text{ for any }\; \xi(x) \;\text{ with }\; \int_D g(x)\xi(x)dx = 0.$$

What can we conclude about $f(x)$ from this? The following:

Claim: $f(x) + \lambda g(x) = 0$ for some $\lambda \in \mathbb{R}$.

To see this, consider *any* (smooth) function $\eta(x)$ on D, and write it as

$$\eta(x) = \frac{(g,\ \eta)}{(g,\ g)} g(x) + \tilde{\eta}(x), \qquad \int_D g(x)\tilde{\eta}(x)dx = (g,\ \tilde{\eta}) = 0$$

(i.e., in linear algebra language, $\tilde{\eta}$ is the orthogonal projection onto the subspace perpendicular to g). Since $\tilde{\eta}$ is perpendicular to g, we have

$$\begin{aligned} 0 = \int_D f(x)\tilde{\eta}(x)dx &= \int_D f(x) \left[\eta(x) - g(x)\frac{1}{(g,\ g)} \int_D g(y)\eta(y)dy\right] \\ &= \int_D \left[f(x) - \frac{(g,\ f)}{(g,\ g)} g(x)\right] \eta(x)dx \end{aligned}$$

Finally, since η can be *any* function, we conclude, as in the earlier lemma, that

$$f(x) - \frac{(g,\ f)}{(g,\ g)} g(x) = 0,$$

or, in other words, u^* is a critical point of the functional

$$I(u) + \lambda M(u), \qquad \lambda = -\frac{(g,\ f)}{(g,\ g)}.$$

□

Example: (Eigenvalue problem). Recall that the first (Dirichlet) eigenvalue of $-\Delta$ on domain D is given by

$$\lambda_1 = \min_{u=0 \text{ on } \partial D} \frac{\int_D |\nabla u(x)|^2 dx}{\int_D u(x)^2 dx} = \min_{u=0 \text{ on } \partial D,\ \int_D u^2 = 1} \int_D |\nabla u(x)|^2 dx$$

(the second expression equals the first, since we may normalize any function $u(x)$, by multiplying by a number, to get $\int u^2 = 1$, and this doesn't change the Rayleigh quotient). The second expression is a constrained minimization problem – let's find its Euler-Lagrange equation. The method of Lagrange multipliers tells us to consider the functional

$$\int_D |\nabla u|^2 dx + \ \mu \int_D u^2 dx = \int_D \left[|\nabla u|^2 + \mu\, u^2\right] dx,$$

with Lagrange multiplier μ, whose Euler-Lagrange equation we can find (as usual) by considering, for any (smooth) $\xi(x)$ vanishing on the boundary, $u + \varepsilon\xi$:

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon} \int_D \left[|\nabla u + \varepsilon \nabla \xi|^2 + \mu(u + \varepsilon\xi)^2\right] dx\Big|_{\varepsilon=0} = \int_D 2\left[\nabla u \cdot \nabla \xi + 2\mu u \xi\right] dx \\ &= 2 \int_D \xi(x) \left[-\Delta u(x) + \mu u(x)\right] dx, \end{aligned}$$

and therefore

$$-\Delta u(x) + \mu u(x) = 0,$$

that is, u is an eigenfunction of $-\Delta$ with eigenvalue $-\mu$. And notice, finally, that if we integrate the equation above against u, and use $\int u^2 dx = 1$, we find

$$-\mu = -\mu \int u^2(x) dx = -\int_D u(x)\Delta u(x) dx = \int_D |\nabla u(x)|^2 dx = \lambda_1,$$

so indeed, the minimizer is an eigenfunction with eigenvalue λ_1 (which, in any case, we already knew).

Extensions:

1. (first Neumann eigenvalue). Notice that if we do not impose *any* boundary conditions in the minimization problem, the minimizer will still satisfy the natural boundary conditions – namely Neumann conditions $\frac{\partial u}{\partial n} \equiv 0$ on ∂D. Hence the first *Neumann* eigenvalue of $-\Delta$ on D is given by

$$\lambda_1 = \min_{\int_D u^2 dx = 1} \int_D |\nabla u|^2 dx,$$

and the minimizer is a corresponding eigenfunction.

2. (higher eigenvalues). Recall that the n-th (back to Dirichlet again) eigenvalue is given by

$$\lambda_n = \min_{u=0\ \partial D,\ \int_D u^2=1,\ (u,\phi_1)=\cdots=(u,\phi_{n-1})=0} \int_D |\nabla u|^2 dx$$

where $\phi_1, \ldots, \phi_{n-1}$ are the first $n-1$ eigenfunctions. This is a constrained variational problem with several constraints – in fact n of them. In this case, we need n Lagrange multipliers. That is, we are looking for critical points of

$$\int_D |\nabla u|^2 dx + \mu_0 \int_D u^2 dx + \sum_{j=1}^{n-1} \mu_j \int_D u\phi_j dx$$

for some numbers $\mu_0, \mu_1, \ldots \mu_{n-1}$. The Euler-Lagrange equation (check it!) is

$$-\Delta u + \mu_0 u + \sum_{j=1}^{n-1} 2\mu_j \phi_j = 0.$$

Integrating this equation against ϕ_k for some $k \in \{1, 2, \ldots, n-1\}$, and using

$$(\phi_k,\ u) = 0, \quad (\phi_k,\ \Delta u) = (\Delta\phi_k,\ u) = (-\lambda_k \phi_k,\ u) = -\lambda_k(\phi_k,\ u) = 0,$$

and the orthonormality of the eigenfunctions, we find $\mu_k = 0$ for $k = 1, 2, \ldots, n-1$. And then integrating against u, we find

$$\mu_0 = -\int_D |\nabla u|^2 dx = -\lambda_n.$$

Hence

$$\begin{cases} -\Delta u = \lambda_n u & D \\ u = 0 & \partial D \end{cases}.$$

That is, the minimizer is indeed an eigenfunction corresponding to the eigenvalue λ_n (which, again, we already knew).

Example: (Minimizing the surface area for fixed volume). Let D be a region in $\mathbb{R}^2$, and consider surfaces described by graphs of functions $x_3 = u(x) \geq 0$, $x \in D$, with $u = 0$ on ∂D. The problem is to minimize the surface area

$$A(u) = \int_D \sqrt{1 + |\nabla u(x)|^2} dx$$

subject to the contraint of a fixed volume:

$$V(u) = \int_D u(x) dx = V_0.$$

Thus we consider the functional

$$A + \lambda V = \int_D \left(\sqrt{1 + |\nabla u(x)|^2} + \lambda u(x) \right) dx$$

for a Lagrange mutiplier λ, and find its Euler-Lagrange equation:

$$0 = \frac{d}{d\varepsilon}(A + \lambda V)(u + \varepsilon\xi)\Big|_{\varepsilon=0} = \int_D \left(\frac{1}{2}(1 + |\nabla u(x)|^2)^{-1/2} 2\nabla u(x) \cdot \nabla\xi(x) + \lambda\xi(x) \right) dx$$

$$= \xi(x) \left(-\nabla \cdot \frac{\nabla u(x)}{\sqrt{1 + |\nabla u(x)|^2}} + \lambda \right) dx.$$

Thus a minimizing function should satisfy

$$\nabla \cdot \frac{\nabla u(x)}{\sqrt{1 + |\nabla u(x)|^2}} = \text{ constant } .$$

As a special case, let D be the disk $x_1^2 + x_2^2 \leq r_0^2$, and try

$$u(x_1, x_2) = \sqrt{c_0^2 - x_1^2 - x_2^2} - \sqrt{c_0^2 - r_0^2},$$

for $r_0 \geq c_0$, which is (part of) a sphere, of radius c_0, centred at $(0, 0, -\sqrt{c_0^2 - r_0^2})$. It is a good exercise to check that this satisfies our minimal surface equation above. Notice that the radius c_0 can be chosen to satisfy the volume constraint V_0, provided V_0 is not too large (exercise: find how large, in terms of r_0).

4. Approximating minimizers: Rayleigh-Ritz method.

Since it is very rare to be able to solve variational problems (indeed, PDE problems in general) explcitly, one usually needs to try and find approximate solutions – whether "by hand", or (more commonly) by computer. The **Rayleigh-Ritz** method is an elementary, classical technique for approximating solutions to variational problems. The idea could not be simpler. Suppose we want to find an approximation to the solution of

$$\min_{u \in H} I(u)$$

where H is some class of functions (eg. with some given BCs), and I is a functional. Let's assume that H is a vector space (i.e. closed under linear combinations), which is often the case. Let

$$v_1(x),\ v_2(x),\ \ldots,\ v_m(x)$$

be m "trial functions", and consider linear combinations of these m functions:

$$u(x) = \sum_{j=1}^{n} c_j v_j(x), \qquad c_j \in \mathbb{R} \text{ for } j = 1, 2, \ldots, m.$$

Such functions lie in H (i.e. satisfy the given BCs), and if we try to minimize the functional I over all such functions (rather than over *all* functions in H), we get a finite-dimensional minimization problem,

$$\min_{c_1,\ldots,c_m} I\left(\sum_{j=1}^{m} c_j v_j\right),$$

which we can solve by standard multi-variable calculus (i.e. set the partial derivatives with respect to each c_j equal to 0).

Example: Find an approximation to the solution of

$$\min_{u=0 \text{ on } \partial D} \int_D \left(\frac{1}{2}|\nabla u|^2 + f(x)u\right) dx$$

where D is the rectangle $[0,1]^2$. (Recall, the minimizer satisfies Poisson's equation $\begin{cases} \Delta u = f & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases}$.)

Convenient trial functions v_j for computation are eigenfunctions of the Laplacian (with zero BCs) – namely, products of sines. That is, let's minimize the given functional among functions of the form

$$\sum_{j=1}^{N} \sum_{k=1}^{N} c_{jk} v_{jk}(x), \qquad v_{jk}(x) = \sin(j\pi x_1)\sin(k\pi x_2)$$

for some positive integer N. Then a nice way to compute the functional is, by the divergence theorem, and the orthogonality of the v_{jk},

$$\frac{1}{2}\int_D |\nabla u|^2 dx = -\frac{1}{2}\int_D u(\Delta u)dx = -\frac{1}{2}\sum_{j,k,j',k'=1}^{N} c_{jk}c_{j'k'}\int_D v_{j'k'}(x)\Delta v_{jk}(x)dx$$

$$= \frac{1}{2}\sum_{j,k,j',k'=1}^{N} c_{jk}c_{j'k'}\int_D v_{j'k'}(x)\pi^2(j^2+k^2)v_{jk}(x) = \frac{\pi^2}{8}\sum_{j,k=1}^{N}(j^2+k^2)c_{jk}^2.$$

And

$$\int_D f(x)udx = \sum_{j,k=1}^{N} c_{jk}\int_D v_{jk}(x)f(x)dx \ =: \ \sum_{j,k=1}^{N} c_{jk}\hat{f}_{n,k}.$$

So our job is to find the choice of the coefficients c_{jk} which minimizes

$$I(u) = \sum_{j,k=1}^{N}\left(\frac{\pi^2}{8}(j^2+k^2)c_{j,k}^2 + \hat{f}_{j,k}c_{j,k}\right).$$

That is, we want to minimize a function of the N^2 variables c_{jk}. So, as we know from calculus, we should take the partial derivatives and set them equal to zero:

$$0 = \frac{\partial}{\partial c_{jk}}\sum_{j,k=1}^{N}\left(\frac{\pi^2}{8}(j^2+k^2)c_{j,k}^2 + \hat{f}_{j,k}c_{j,k}\right) = \frac{\pi^2}{4}(j^2+k^2)c_{j,k} + \hat{f}_{j,k}.$$

So we take

$$c_{j,k} = -\frac{4\hat{f}_{j,k}}{\pi^2(j^2+k^2)},$$

and our approximate minimizer is

$$\tilde{u}(x) = -\frac{4}{\pi^2}\sum_{j,k=1}^{N}\frac{\hat{f}_{j,k}}{j^2+k^2}\sin(j\pi x_1)\sin(k\pi x_2).$$

Remark: In this example, the approximate minimizer turns out to be nothing but the eigenfunction expansion solution of the corresponding Poisson equation, truncated at $j,k \leq N$.

Remark: In general it is somewhat difficult to asess how *good* an approximation the Rayleigh-Ritz method generates. In this particular example, from what we know about the completeness of eigenfunctions, we can at least conclude that our approximation should get better as N gets larger, and, in particular, should approach the true solution as $N \to \infty$.

Example: Let D be the 2D region bounded by the ellipse $x_1^2/a^2 + x_2^2/b^2 = 1$. Approximate the solution u of

$$\min_{u \in C^2(D)} \left\{ I(u) = \int_D \left[\left(\frac{\partial u}{\partial x_1} - x_2 \right)^2 + \left(\frac{\partial u}{\partial x_2} + x_1 \right)^2 \right] dx \right\}.$$

Notice that no boundary conditions are imposed in the problem.

Remark: Before we start, notice something: clearly $I(u) \geq 0$. Can we simply get a minimizer with $I(u) = 0$ by setting $u_{x_1} = x_2$ and $u_{x_2} = -x_1$? Well, let's try: the first equation requires $u = x_1 x_2 + f(x_2)$, and then the second requires $x_1 + f'(x_2) = -x_1$ so $f'(x_2) = -2x_1$ which we cannot satisfy. So, no. In vector calculus language, we are trying to find a function $u(x)$ so that $\nabla u = (x_2, -x_1)$. A necessary condition for a vector-field to be a gradient is that its curl vanishes. But $curl(x_2, -x_1) \neq 0$, so we cannot solve the equation.

OK, back to the question of approximating the minimizer. To keep things simple, let's use just *one* trial function ($m = 1$). One computationally nice choice is $v_1(x) = x_1 x_2$ since a multiple of this (v_1 itself) makes the first term in the integral vanish, while a different multiple ($-v_1$) makes the second term vanish. So, we are lead to

$$\min_{\alpha \in \mathbb{R}} I(\alpha x_1 x_2),$$

i.e. just minimizing a function of the single variable α. Let's compute

$$I(\alpha x_1 x_2) = \int_D \left[(\alpha - 1)^2 x_2^2 + (\alpha + 1)^2 x_1^2 \right] dx = (\alpha - 1)^2 \int_D x_2^2 dx + (\alpha + 1)^2 \int_D x_1^2 dx.$$

Before proceeding, let's pause and ask what we expect in the case of a disk: $a = b$. The symmetry suggests that neither positive nor negative α should be favoured, and so $\alpha = 0$ should be the minimizer. We'll keep this idea in mind as a "check" on our computation at the end.

Now

$$\begin{aligned} \int_D x_1^2 dx &= 4 \int_0^a x_1^2 dx_1 \int_0^{b\sqrt{1 - x_1^2/a^2}} dx_2 = 4b \int_0^a x_1^2 \sqrt{1 - x_1^2/a^2} dx_1 \\ &= 4a^3 b \int_0^{\pi/2} \sin^2(\theta) \cos^2(\theta) d\theta = \frac{\pi a^3 b}{4}, \end{aligned}$$

and similarly,

$$\int_D x_2^2 dx = \frac{\pi a b^3}{4}.$$

So our problem is

$$\min_{\alpha} \frac{\pi ab}{4} \left[b^2(\alpha-1)^2 + a^2(\alpha+1)^2\right],$$

which we solve as usual:

$$0 = \frac{d}{d\alpha}\left[b^2(\alpha-1)^2 + a^2(\alpha+1)^2\right] = 2b^2(\alpha-1) + 2a^2(\alpha+1) = 2(a^2+b^2)\alpha + 2(a^2-b^2)$$

hence $\alpha = \frac{b^2-a^2}{a^2+b^2}$ gives the minimum value (it must be a minimum, since the graph of this function of α is an upward parabola), and so our (admittedly crude) approximate minimizer is

$$\tilde{u}(x) = \frac{b^2-a^2}{a^2+b^2} x_1 x_2$$

which gives a value of

$$I(\tilde{u}) = \frac{\pi ab}{4}\left[b^2\left(\frac{-2a^2}{a^2+b^2}\right)^2 + a^2\left(\frac{2b^2}{a^2+b^2}\right)^2\right] = \frac{\pi a^3 b^3}{a^2+b^2}.$$

(And note that our guess that $\alpha = 0$ when $a = b$ is indeed confirmed.)

5. Approximating eigenvalues and eigenfunctions

Recall that the first eigenvalue of the problem

$$\begin{cases} L\phi := -\nabla \cdot [p(x)\nabla\phi] + q(x)\phi = \lambda r(x)\phi \qquad \text{in } D \\ \phi = 0 \quad \text{on } \partial D \end{cases}$$

is given by minimizing the Rayleigh quotient:

$$\lambda_1 = \min_{u=0 \text{ on } \partial D} \frac{\int_D (p|\nabla u|^2 + qu^2)\, dx}{\int_D ru^2 dx},$$

and the minimizing function is a corresponding eigenfunction. So we can use a Rayleigh-Ritz approach to approximate the eigenfunction.

Start with a set of m "trial functions":

$$v_1(x), v_2(x), \ldots, v_m(x)$$

with the correct boundary conditions (in this case Dirichlet: $v_j \equiv 0$ on ∂D), and we will minimize the Rayleigh quotient among linear combinations

$$u(x) = c_1 v_1(x) + c_2 v_2(x) + \ldots + c_m v_m(x).$$

Let's compute the relevant quantities. Using the convention of summation over repeated indices:

$$\int_D ru^2 dx = c_j c_k \int_D r v_j v_k dx = \overrightarrow{c} \cdot B \overrightarrow{c}$$

where $\overrightarrow{c}$ denote the m-vector

$$\overrightarrow{c} = (c_1, c_2, \ldots, c_m),$$

and B is the $m \times m$ matrix

$$B := (B_{jk})_{j,k=1}^m, \qquad B_{jk} = \int_D r v_j v_k dx = (v_j, v_k)_r.$$

Similarly,

$$\int_D \left(p|\nabla u|^2 + qu^2\right) dx = c_j c_k \int_D (p\nabla v_j \cdot \nabla v_k + q v_j v_k)\, dx = \overrightarrow{c} \cdot A \overrightarrow{c}$$

where

$$A := (A_{jk})_{j,k=1}^m, \qquad A_{jk} = \int_D (p\nabla v_j \cdot \nabla v_k + q v_j v_k)\, dx = (v_j, L v_k) = (v_k, L v_j) = A_{kj}$$

where we have noted that the matrix A is symmetric, because the operator L is self-adjoint. Notice B is symmetric, too. Hence our problem is:

$$\min_{\vec{c}\in\mathbb{R}^m} \frac{\vec{c}\cdot A\vec{c}}{\vec{c}\cdot B\vec{c}},$$

a simple minimization of a function of m variables. Of course, by calculus, the minimizer will satisfy (going back to summation over repeated indices notation)

$$0 = \frac{\partial}{\partial c_j}\frac{A_{kl}c_k c_l}{B_{nr}c_n c_r} = \frac{\vec{c}\cdot B\vec{c}(A_{jl}+A_{lj})c_l - \vec{c}\cdot A\vec{c}(B_{nj}+B_{jn})c_n}{(\vec{c}\cdot B\vec{c})^2}, \quad j = 1,2,\dots,m$$

which, using the symmetry of A and B, yields

$$(\vec{c}\cdot B\vec{c})A\vec{c} = (\vec{c}\cdot A\vec{c})B\vec{c}$$

or

$$A\vec{c} = \lambda B\vec{c}, \qquad \lambda := \frac{\vec{c}\cdot A\vec{c}}{\vec{c}\cdot B\vec{c}},$$

which is a (generalized) matrix eigenvalue problem.

Example: Approximate the first Dirichlet eiegnfunction of $-\frac{d^2}{dx^2}+\alpha x$ on $[0,1]$ using trial functions $v_1(x) = \sin(\pi x)$, $v_2(x) = \sin(2\pi x)$ (notice that v_1 and v_2 are the first two eigenfunctions of $-\frac{d^2}{dx^2}$ on $[0,1]$ with 0 BCs – that is, of the problem with $\alpha = 0$). Easy computations give

$$(v_1,v_1) = \int_0^1 \sin^2(\pi x)dx = \frac{1}{2}, \quad (v_2,v_2) = \int_0^1 \sin^2(2\pi x)dx = \frac{1}{2},$$
$$(v_1,v_2) = \int_0^1 \sin(\pi x)\sin(2\pi x)dx = 0,$$

so our matrix B is

$$B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Next,

$$v_1' = \pi\cos(\pi x), \quad v_2' = 2\pi\cos(2\pi x),$$

so

$$\int_0^1 (v_1')^2 dx = \frac{\pi^2}{2}, \quad \int_0^1 (v_2')^2 dx = 2\pi^2, \quad \int_0^1 v_1'v_2' dx = 0.$$

Slightly more involved compoutations (using integrations by parts) yield

$$\int_0^1 x v_1^2 dx = \frac{1}{4}, \quad \int_0^1 x v_2^2 dx = \frac{1}{4}, \quad \int_0^1 x v_1 v_2 dx = \frac{1}{2\pi}.$$

Hence our matrix A is

$$A = \begin{pmatrix} \frac{\pi^2}{2} + \frac{\alpha}{4} & \frac{\alpha}{2\pi} \\ \frac{\alpha}{2\pi} & 2\pi^2 + \frac{\alpha}{4} \end{pmatrix}.$$

To solve the eigenvalue problem $(A - \lambda B)\overrightarrow{c} = 0$, we set

$$0 = \det(A - \lambda B) = \begin{vmatrix} \frac{\pi^2}{2} + \frac{\alpha}{4} - \frac{\lambda}{2} & \frac{\alpha}{2\pi} \\ \frac{\alpha}{2\pi} & 2\pi^2 + \frac{\alpha}{4} - \frac{\lambda}{2} \end{vmatrix} = \left(\frac{\pi^2}{2} + \frac{\alpha}{4} - \frac{\lambda}{2}\right)\left(2\pi^2 + \frac{\alpha}{4} - \frac{\lambda}{2}\right) - \frac{\alpha^2}{4\pi^2},$$

and the quadratic formula leads to

$$2\lambda = 5\pi^2 + \alpha \pm \sqrt{(5\pi^2 + \alpha)^2 - 4[(\pi^2 + \alpha/2)(4\pi^2 + \alpha/2) - \frac{\alpha^2}{\pi^2}]}.$$

Since we are trying to approximate the *lowest* eigenvalue/eigenfunction, we take the negative sign here, to obtain the approximate eigenvalue λ. Then one can find an eigenvector $\overrightarrow{c}$ which will produce an approximate eigenfunction $c_1 v_1 + c_2 v_2$. Of course, the expressions are a little messy. One thing we can do, supposing α is small, is expand the approximate eigenvalue in a Taylor series in α to find, for example,

$$\lambda = \pi^2 + \frac{\alpha}{2} + O(\alpha^2).$$

Remark: Notice that this approximate eigenvalue λ is actually an upper-bound for the true first eigenvalue:

$$\lambda_1 \leq \lambda,$$

since we obtain it by minimizing the Rayleigh quotient over a sub-set of functions (the true eigenvalue is the minimizer over *all* admissible functions, hence cannot be larger).