

Lecture 11: Calculus of Variations

Ethan Ebbighausen

July 20, 2026

1 Introduction

Leonhard Euler and Joseph-Louis Lagrange worked in collaboration in the 1750s on the tautochrone problem. They sought to determine a curve along which a weighted particle (a ball, but with no rotational analysis), would fall from any given height to a fixed point in a fixed amount of time (so two balls at different starting points arrive at the bottom at the same moment). Lagrange solved the problem in 1755 and sent his solution to Euler. The method he used was the precursor to Lagrangian mechanics, and their collaboration led to the Calculus of variations (a term coined by Euler) where problems in classical mechanics are expressed as minimization of an action functional.

The action functional, in the original setting, was the integral of a *Lagrangian function*, or kinetic energy minus potential energy.

While the tautochrone problem is interesting, it is a little opaque to why we make some choices on first viewing. Let us look instead at a related but slightly simpler problem called the Brachistochrone problem, where we are trying to find a curve $(x, y(x))$ in $\mathbb{R}^2$ along which a point mass sliding frictionlessly moves from point A to point B the fastest.

The time it takes to move along the curve is the distance divided by the velocity, so

$$T = \int_{x_A}^{x_B} dt = \int_{x_A}^{x_B} \frac{1}{v(x, y)} \sqrt{1 + y'(x)^2} dx$$

From energy conservation, we may determine the speed as

$$v = \sqrt{2g(y_A - y)}$$

and so

$$T = \int_{x_A}^{x_B} \sqrt{\frac{1 + y'(x)^2}{2g(y_A - y)}} dx$$

We may view T as the output of a map $y \mapsto T(y)$ and finding the curve amounts to minimizing this function. To do so, we try to look for critical points:

$$T(y) = \int_{x_A}^{x_B} F(y, y') dx$$
$$\frac{d}{ds} T(y + sv) = \int_{x_A}^{x_B} \frac{d}{ds} (F(y + sv, y' + sv')) dx$$

so that this is a critical point when it is 0 for all v . Computing this carefully gives a solution that y is a cycloid.

Learning Objectives:

- Write weak PDE solutions as minimizers of integral formulas
- Identify special cases where minimizers exist
- Use the Rayleigh Coefficient to understand the spectrum
- Approximate spectrum values using minimax principles
- Identify this general structure as Euler-Lagrange Equations

2 The Poisson Equation and Dirichlet's Principle

We can twist this physical situation for our benefit in the case of viewing the Laplace equation as an energy minimizer (time invariant/equilibrium solution) for the wave equation energy.

We start with the *Dirichlet energy* for $w \in C^2(\overline{\Omega})$, defined by

$$\mathcal{E}[w] = \frac{1}{2} \int_{\Omega} |\nabla w|^2 dx$$

We view $\mathcal{E}$ as a function $\mathcal{E} : C^2(\Omega) \rightarrow \mathbb{R}$, and so we emphasize that its domain is a function space by calling it a *functional*. Let us suppose u is a minimizer of this energy, so for all $\phi \in C_c^\infty(\overline{\Omega}, \mathbb{R})$,

$$\mathcal{E}[u] \leq \mathcal{E}[u + \phi]$$

Therefore, for each such ϕ ,

$$\frac{d}{dt} \mathcal{E}[u + t\phi]|_{t=0} = 0$$

If we differentiate under the integral,

$$\begin{aligned} \frac{d}{dt} \mathcal{E}[u + t\phi]|_{t=0} &= \frac{1}{2} \frac{d}{dt} \int_{\Omega} (|\nabla u|^2 + 2t \nabla u \cdot \nabla \phi + t^2 |\nabla \phi|^2) dx|_{t=0} \\ &= \int_{\Omega} \nabla u \cdot \nabla \phi dx = - \int_{\Omega} \phi \Delta u dx \end{aligned}$$

Since this holds for all test functions ϕ ,

$$\int_{\Omega} \phi \Delta u dx = 0$$

implies $\Delta u = 0$. Minimizers are then Laplace solutions, and Laplace solutions are critical points.

Let us add some generality. Consider the PDE

$$\begin{aligned} -\Delta u &= f \\ u|_{\partial\Omega} &= 0 \end{aligned}$$

to solve this via minimization, it is helpful to use the weak formulation instead, where we take $f \in L^2(\Omega)$ and try to find $u \in H_0^1(\Omega)$ such that

$$\int_{\Omega} [\nabla u \cdot \nabla \psi - f\psi] dx = 0$$

for every $\psi \in C_c^\infty(\Omega)$.

To match this formulation, we define the functional

$$\mathcal{D}_f[w] = \int_{\Omega} \left[\frac{1}{2} \nabla w \cdot \nabla w - fw \right] dx$$

[2.1] Dirichlet's Principle Suppose $\Omega \subset \mathbb{R}^n$ is a bounded domain and $f \in L^2(\Omega, \mathbb{R})$. If $u \in H_0^1(\Omega, \mathbb{R})$ satisfies

$$\mathcal{D}_f[u] \leq \mathcal{D}_f[w]$$

for all $w \in H_0^1(\Omega, \mathbb{R})$, then u is a weak solution of the Poisson equation.

Proof:

Since $C_c^\infty \subset H_0^1$,

$$\mathcal{D}_f[u] \leq \mathcal{D}_f[u + t\psi]$$

for test functions ψ and $t \in \mathbb{R}$. Therefore,

$$0 = \frac{d}{dt} \mathcal{D}_f[u, t\psi]|_{t=0} = \int_{\Omega} \nabla u \cdot \nabla \psi - f\psi dx$$

Q.E.D.

The term “calculus of variations” refers to the fact that we are differentiating across the family of variations $u + t\psi$. We have now reduced solving the PDE to finding a minimum, and next show that a minimum exists.

2.1 Assumptions for the Existence of a Minimum

Minimizing the functional $\mathcal{D}_f[u + t\psi]$ reduces to minimizing

$$\int_{\Omega} \frac{1}{2} |\nabla u|^2 - fu + t(\nabla u \cdot \nabla \psi - f\psi) + \frac{t^2}{2} |\nabla \psi|^2 dx$$

or that we have a quadratic term, a linear term, and a constant term. The quadratic $ax^2 + bx + c$ has a minimum if and only if $a > 0$, which is of course true for the above form given non-constant ψ . However, to have this across all ψ , we need a better formulation of the lower bound for the quadratic term. The original form of this is a result of Henri Poincaré

[2.2] Poincaré Inequality For a bounded domain $\Omega \subset \mathbb{R}^n$, there is a constant $\kappa > 0$ depending only on Ω such that

$$\|u\|_2^2 \leq \kappa^2 \mathcal{E}[u]$$

for all $u \in H_0^1(\Omega)$. **Remark:** We call the optimal κ above the *Poincaré constant* $\kappa(\Omega)$

Proof:

We first focus on the test functions (which are dense in H_0^1) and deal with the general case after. The basic idea is that we want to compare the size of a function to the size of its derivative. In 1D, for a test function ψ supported in (a, b) , this is obvious due to the Cauchy-Schwarz inequality

$$|\psi(x)|^2 = \left(\int_a^x \psi'(t) dt \right)^2 \leq \|\psi'\|_2^2 (b - a)$$

so that

$$\|\psi\|_2^2 \leq (b - a)^2 \|\psi'\|_2^2 = (b - a)^2 \mathcal{E}[\psi]$$

In higher dimensions, the harder part is defining the right derivative to apply this idea to. If we have a bounded domain $\Omega \subset \mathbb{R}^n$, we pick a symmetrical box bound

$$\Omega \subset [-M, M]^n$$

for some large M . By our previous theorem, smooth and H_0^1 functions may be extended by 0 to this space. Then,

$$\begin{aligned} |\psi(x)|^2 &= \left(\int_{-M}^{x_1} \partial_1 \psi(y, x_2, \dots, x_n) dy \right)^2 \\ &\leq 2M \|\partial_1 \psi(y, x_2, \dots, x_m)\|_{L^2([-M, M])}^2 \end{aligned}$$

so integrating across all x gives

$$\|\psi\|_2^2 \leq 4M^2 \|\partial_1 \psi\|_2^2$$

Taking the average across all x_i ,

$$\|\psi\|_2^2 \leq \frac{4M^2}{n} \sum \|\partial_i \psi(y, x_2, \dots, x_m)\|_2^2$$

Notice that

$$\mathcal{E}[\psi] = \frac{1}{2} \int_{\Omega} \nabla \psi \cdot \nabla \psi dx = \frac{1}{2} \sum \|\partial_i \psi\|_2^2$$

Combining this with the above inequality gives

$$\|\psi\|_2^2 \leq \frac{8M^2}{n} \mathcal{E}[\psi]$$

To extend to $u \in H_0^1(\Omega)$, we take some sequence $\{\psi_k\}$ of test functions which converges to u in H^1 . This immediately implies $\|\psi_k\|_2 \rightarrow \|u\|_2$, so we need only treat the energy on the right hand side of the inequality above to show it converges to the energy of u . Taking the difference,

$$\begin{aligned} \mathcal{E}[u] - \mathcal{E}[\psi_k] &= \int \nabla u \cdot \nabla u - \nabla \psi_k \cdot \nabla \psi_k dx \\ &= \int \nabla u \cdot \nabla(u - \psi_k) + \nabla \psi_k \cdot \nabla(u - \psi_k) dx \\ &\leq \|\nabla(u - \psi_k)\|_2 (\|\nabla u\|_2 + \|\nabla \psi_k\|_2) \end{aligned}$$

so that

$$|\mathcal{E}[u] - \mathcal{E}[\psi_k]| \leq \|u - \psi_k\|_{H^1} (\|u\|_{H^1} + \|\psi_k\|_{H^1})$$

Since the ψ_k are bounded in norm,

$$\lim_{k \rightarrow \infty} \mathcal{E}[\psi_k] = \mathcal{E}[u]$$

so taking the limit of

$$\|\psi_k\|_2^2 \leq \frac{8M^2}{n} \mathcal{E}[\psi_k]$$

as $k \rightarrow \infty$ shows

$$\|u\|_2^2 \leq \frac{8M^2}{n} \mathcal{E}[u]$$

Q.E.D.

We may re-express this theorem in terms of the H^1 -norm using

$$\|u\|_{H^1}^2 = \|u\|_2^2 + \mathcal{E}[u]$$

so that the theorem states

$$\|u\|_{H^1}^2 \leq (\kappa^2 + 1)\mathcal{E}[u]$$

In general, a quadratic operator $L : H \rightarrow \mathbb{R}$ is called *coercive* if $|Lu| \geq C\|u\|^2$ for some $C > 0$. Poincaré's Inequality states that the energy is a coercive operator. A coercive operator is immediately injective (so invertible in some sense) and its inverse is bounded.

We also know

$$\mathcal{E}[u] \leq \|u\|_{H^1}^2$$

so that the energy operator is bounded,.

Using these two properties, we may solve the minimization problem

[2.3] For a bounded domain $\Omega \subset \mathbb{R}^n$ and $f \in L^2(\Omega)$, there is a unique function $u \in H_0^1(\Omega)$ such that

$$\mathcal{D}_f[u] \leq \mathcal{D}_f[w]$$

for all $w \in H_0^1(\Omega)$.

Proof:

This proof will be non-constructive. We will show that the Dirichlet operator is bounded below, so an infimum value exists. Then, we will use the existence of a sequence converging to that infimum to find a function that minimizes the value of the Dirichlet operator, which will be our minimizer. Finally, we will show that this minimizer is unique and so this process gave the only possible minimizer.

To show that a finite infimum exists, notice that

$$\mathcal{D}_f[w] \geq \mathcal{E}[w] - |\langle f, w \rangle|$$

so that adding in Poincaré's inequality (Coercivity estimate) and the Cauchy-Schwarz inequality,

$$\begin{aligned} \mathcal{D}_f[w] &\geq \frac{1}{\kappa^2 + 1} \|w\|_{H^1}^2 - \|f\|_2 \|w\|_2 \\ &\geq \frac{1}{\kappa^2 + 1} \|w\|_{H^1}^2 - \|f\|_2 \|w\|_{H^1} \end{aligned}$$

Notice that for $x = \|w\|_{H^1}$, this right side is a quadratic in x with a positive coefficient on x^2 . Therefore, using the vertex of the parabola, we have that

$$\mathcal{D}_f[w] \geq -\frac{\kappa^2 + 1}{4} \|f\|_2^2$$

We have a lower bound and may now take $d_0 = \inf_{w \in H_0^1(\Omega)} \mathcal{D}_f[w]$. We take a sequence $\{w_k\} \subset H_0^1$ so $\mathcal{D}_f[w_k] \rightarrow d_0$. We want to show that this sequence is Cauchy, and so a limit exists because H_0^1 is complete.

We first use a convexity-type inequality.

$$\mathcal{E}\left[\frac{u+v}{2}\right] = \frac{1}{2}\mathcal{E}[u] + \frac{1}{2}\mathcal{E}[v] - \frac{1}{4}\mathcal{E}[u-v]$$

implies

$$d_0 \leq \mathcal{D}_f\left[\frac{w_k + w_m}{2}\right] = \frac{1}{2}\mathcal{D}_f[w_k] + \frac{1}{2}\mathcal{D}_f[w_m] - \frac{1}{4}\mathcal{E}[w_k - w_m]$$

or

$$\mathcal{E}[w_k - w_m] \leq 2\mathcal{D}_f[w_k] + 2\mathcal{D}_f[w_m] - 4d_0$$

As $k, m \rightarrow \infty$, the right hand side of this inequality decreases to 0 (both functional values go to d_0). Using the coercivity estimate $\|w\|_{H^1}^2 \leq C\mathcal{E}[w]$, we then have that

$$\|w_k - w_m\|_{H^1} \rightarrow 0$$

as $k, m \rightarrow \infty$. The sequence is therefore Cauchy, and there is a limit $u \in H_0^1(\Omega)$. By construction, $\mathcal{D}_f[u] = d_0$, so u is a minimizer.

Lastly, for uniqueness, if u_1 and u_2 both have $\mathcal{D}_f[u_i] = d_0$,

$$\|u_1 - u_2\|_{H^1}^2 \leq C\mathcal{E}[u_1 - u_2] \leq C[2\mathcal{D}_f[u_1] + 2\mathcal{D}_f[u_2] - 4d_0] = 0$$

so $u_1 \equiv u_2$.

Q.E.D.

[2.4] For $f \in L^2(\Omega)$, the weak formulation of the Poisson equation admits a unique solution $u \in H_0^1(\Omega)$.

Proof:

By the Dirichlet principle and the theorem above, the existence of a minimizer to $\mathcal{D}_f$ implies the existence of a Poisson solution u . To prove uniqueness, we note that two solutions u_1 and u_2 have

$$\int_{\Omega} \nabla(u_1 - u_2) \cdot \nabla \psi dx = 0$$

for all $\psi \in C_c^\infty(\Omega)$. Since we may take a sequence of test functions ψ_k converging to $u_1 - u_2$ in H^1 , this implies

$$\mathcal{E}[u_1 - u_2] = \lim_{k \rightarrow \infty} \int_{\Omega} \nabla(u_1 - u_2) \cdot \nabla \psi_k dx = 0$$

so that, by the Poincaré inequality, $\|u_1 - u_2\|_2 = 0$ and $u_1 \equiv u_2$.

Q.E.D.

Remark: In practice, we may estimate a solution by looking at finite-dimensional subspace $A \subset H_0^1(\Omega)$. If A has basis $w_1, \dots, w_m$, then

$$D_f[\sum c_i w_i] = c^T E c - c^T F$$

for $E_{ij} = \mathcal{E}[w_i, w_j]$ and $F_i = \langle f, w_i \rangle$. Since E is symmetric, the spectral theorem for matrices allows us to write E as diagonal using an orthogonal change-of-basis, which allows for computation to find c so

$$\sum c_i w_i = \operatorname{argmin}_{w \in A} D_f[w]$$

2.2 Elliptic Regularity

In the classical Poisson equation $-\Delta u = f$ (with true derivatives), where $f \in C^k$, we expect $u \in C^{k+2}$. In the case of a harmonic function, we should then expect it to be smooth. The goal of this section is to derive a similar result for the weak formulation.

As smoothness is a local property, we avoid technicalities with the boundary by working with local Sobolev spaces

$$H_{loc}^m(\Omega) = \{u \in L_{loc}^1(\Omega) \mid u\psi \in H^m(\Omega) \text{ for all } \psi \in C_c^\infty(\Omega)\}$$

which is slightly less restrictive than H_0^1 .

[2.5] Interior Regularity Suppose that $u \in H_{loc}^1(\Omega)$ is a weak solution of

$$-\Delta u = f$$

If $f \in H_{loc}^m(\Omega)$ for $m \geq 0$, then $u \in H_{loc}^{m+2}(\Omega)$.

Remark: Our Sobolev embedding theorem (lecture 9) gives $H_{loc}^m(\Omega) \subset C^k(\Omega)$ when $k < m - \frac{n}{2}$. Therefore, if $f \in H_{loc}^m$ for $m > \frac{n}{2}$, this theorem says that u is actually a classical solution.

We begin by proving the toral version using Fourier analysis, then extend to the real numbers using an argument on localities.

[2.6] Lemma Suppose $u \in H^1(\mathbb{T}^n)$ solves $-\Delta u = f$ in the weak sense. If $f \in H^m(\mathbb{T}^n)$ for $m \in \mathbb{N}_0$, then $u \in H^{m+2}(\mathbb{T}^n)$.

Proof:

Let u be a solution as in the statement of the theorem, so for all $\psi \in C^\infty(\mathbb{T}^n)$,

$$\int_{\mathbb{T}^n} \nabla u \cdot \nabla \psi - f \psi dx = 0$$

Taking $\psi = e^{-ikx}$ shows

$$\int_{\mathbb{T}^n} [-ik \cdot \nabla u - f(x)] e^{-ikx} dx = 0$$

or

$$\begin{aligned} c_k[f] &= (-ik) \cdot c_k[\nabla u] \\ &= |k|^2 c_k[u] \end{aligned}$$

Since $f \in H^m$,

$$\sum |k|^{2m+4} |c_k[u]|^2 = \sum |k|^{2m} |c_k[f]|^2 < \infty$$

and so $u \in H^{m+2}$ as desired by our Sobolev regularity theorem of lecture 9.

Q.E.D.

Proof:

[Proof of Theorem] Suppose $-\Delta u = f$ weakly, and $u \in H_0^1(\Omega)$, $f \in L_{loc}^2(\Omega)$. Since the Poisson function is invariant under rescaling and translating, we may alter Ω as necessary to assume $\Omega \subset [0, 2\pi]^n$. Let $\chi \in C_c^\infty(\Omega)$, so $u\chi$ is a localization of u . We will extend this function by 0 to $\mathbb{T}^n$, show that it satisfies a modified Poisson equation here, then use the above lemma to complete the argument.

Indeed, for $\psi \in C^\infty(\mathbb{T}^n)$, the product $\psi\chi$ is as well and

$$\begin{aligned} 0 &= \int_{\mathbb{T}^n} [\nabla u \cdot \nabla(\psi\chi) - f(\psi\chi)] dx \\ &= \int_{\mathbb{T}^n} [\nabla(u\chi) \cdot \nabla\psi - u\nabla\chi \cdot \nabla\psi + \psi\nabla u \cdot \nabla\chi - f\chi\psi] dx = (A) \end{aligned}$$

We notice that we can treat some of these middle terms since, by the weak derivative definition,

$$\int_{\Omega} \nabla u \cdot \psi \nabla \chi dx = - \int_{\Omega} u \nabla \psi \cdot \nabla \chi + u \psi \Delta \chi dx$$

Then,

$$0 = (A) = \int_{\mathbb{T}^n} [\nabla(u\chi) \cdot \nabla\psi + (2\nabla u \cdot \nabla\chi + u\Delta\chi - f\chi)\psi] dx$$

In other words,

$$-\Delta u = (2\nabla u \cdot \nabla\chi + u\Delta\chi - f\chi)$$

weakly since the above holds for all smooth ψ . The right-hand-side of this equation is in L^2 since u and f are. Hence, by the above lemma, $u\chi \in H^2(\mathbb{T}^n)$. Unwinding the localization, $u \in H_{loc}^2(\Omega)$.

Now, if we know $u, f \in H_{loc}^q(\Omega)$, the same function above would be $H^{q+2}(\mathbb{T}^n)$, so that $u \in H_{loc}^{q+2}(\Omega)$. Hence, repeating the argument gives the higher-regularity cases.

Q.E.D.

3 Spectral Analysis

As we've been diving into elliptic equations in general settings both in the last lecture with Green's functions and here with variations, we have neglected one piece of information that we used readily in our simpler studies of the Laplacian: eigenvalues. We used the eigenvalues and eigenfunctions to develop the heat series equation and Fourier series, and now we shall apply variational calculus to the eigenvalue problem to try to estimate more general eigenvalues.

3.1 Eigenvalues by Minimization

The goal of this section is to prove the following result.

[3.1] Spectral Theorem for the Dirichlet Laplacian Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. There exists an orthonormal basis $\{\phi_k\}_{k \in \mathbb{N}}$ for $L^2(\Omega)$ such that

$$-\Delta \phi_k = \lambda_k \phi_k$$

with $\phi_k \in H_0^1(\Omega; \mathbb{R}) \cap C^\infty(\Omega)$. Furthermore, $\lambda_k > 0$ for all k and

$$\lim_{k \rightarrow \infty} \lambda_k = \infty$$

In the 1D case, we have already directly computed these eigenvalues. Now, we will exploit the weak formulation of the eigenvalue equation $-\Delta \phi = \lambda \phi$, given by

$$\int_{\Omega} [\nabla \phi \cdot \nabla \psi - \lambda \phi \psi] dx = 0$$

(for all test functions ψ). This looks like a difference of an energy-type inner product and the L^2 inner product, so we will use the notation

$$\mathcal{E}[u, v] = \int_{\Omega} \nabla u \cdot \overline{\nabla v} dx$$

whereby the weak formulation is

$$\mathcal{E}[\phi, \psi] = \lambda \langle \phi, \psi \rangle$$

for all $\psi \in C_c^\infty(\Omega)$. Therefore, we will aim to compute λ by minimizing

$$\mathcal{R}[v] = \frac{\mathcal{E}[v, v]}{\|v\|_{L^2}^2}$$

a function called the *Rayleigh quotient*.

To compute this, it is easier to work over H_0^1 , and it suffices to note that the density of C_c^∞ in H_0^1 allows us to take that the weak formulation holds for all such functions. In this case, any eigenfunction has $\mathcal{R}[\phi] = \lambda$. We also know by Poincaré's inequality that

$$\mathcal{R}[v] \geq \frac{1}{\kappa^2} > 0$$

for nonzero $v \in H_0^1$, so that the Rayleigh quotient is bounded below. This suggests a relationship between the constant here and the first eigenvalue.

However, the Rayleigh quotient is not linear, and so we need an extra tool to work with minimization.

[3.2] Rellich's Theorem Suppose $\Omega \subset \mathbb{R}^n$ is a bounded domain and $\{v_k\}$ is a bounded sequence in $H_0^1(\Omega)$. Then, $\{v_k\}$ has a subsequence $\{v_{k_n}\}$ which is convergent in $L^2(\Omega)$.

We will not prove this result (it is proved in Borthwick, section 11.6 if you would like to see it) for the sake of technicality. Results of this form are very important in PDE, but much more of the flavor of higher analysis than we are using for this course.

[3.3] First Eigenvalue There exists $\phi_1 \in H_0^1(\Omega; \mathbb{R}) \cap C^\infty(\Omega)$ satisfying

$$-\Delta \phi_1 = \lambda_1 \phi_1$$

for $\lambda_1 > 0$ such that

$$\lambda_1 \leq \mathcal{R}[v]$$

for all $v \in H_0^1(\Omega)$, $v \neq 0$.

Proof:

Set

$$\lambda_1 = \inf_{v \in H_0^1(\Omega) \setminus \{0\}} \mathcal{R}[v]$$

which is finite and positive by Poincaré's inequality. Pick a sequence of $\{v_k\}$ such that $\mathcal{R}[v_k] \rightarrow \lambda_1$, and rescale the v_k so they have L^2 -norm 1 (i.e. $\mathcal{R}[v_k] = \mathcal{E}[v_k]$).

Since these quotients converge, the energies also converge and are therefore bounded. This implies the sequence is bounded in the H^1 norm, so by passing to a subsequence we may assume that the v_k converge to some function ϕ in L^2 . We want to show that this convergence is actually true in H^1 , and we manipulate the energy functional to do so. We have L^2 convergence already, so we just need to show that $\mathcal{E}[v_k - \phi] \rightarrow 0$ as $k \rightarrow \infty$.

First, from the known inequality

$$\mathcal{R}\left[\frac{v_k + v_m}{2}\right] \geq \lambda_1$$

we have

$$\mathcal{E}\left[\frac{v_k + v_m}{2}\right] \geq \frac{\lambda_1}{4} \|v_k + v_m\|_2^2$$

Combining this with

$$\mathcal{E}\left[\frac{v_k + v_m}{2}\right] = \frac{1}{2} \mathcal{E}[v_k] + \frac{1}{2} \mathcal{E}[v_m] - \frac{1}{4} \mathcal{E}[v_k - v_m]$$

shows that

$$\mathcal{E}[v_k - v_m] \leq 2\mathcal{E}[v_k] + 2\mathcal{E}[v_m] - \lambda_1 \|v_k + v_m\|_2^2$$

The third term of the right-hand side, as we know, converges to $\lambda_1 \|2\phi\|_2^2 = 4\lambda_1$. The two energy terms also converge to this by the infimum sequence assumption, so the RHS converges to 0 and we know

$$\lim_{k, m \rightarrow \infty} \mathcal{E}[v_k - v_m] = 0$$

Since we know the $\{v_k\}$ are Cauchy in L^2 , this gives that the $\{v_k\}$ are Cauchy in H^1 and converge to some $u \in H_0^1$. However, since L^2 convergence is accounted for here, we must have that $u = \phi$. This shows $\phi \in H_0^1$ and also that

$$\mathcal{R}[\phi] = \lambda_1$$

To conclude this proof, we now show that ϕ_1 satisfies the weak solution condition. Suppose $w \in H_0^1(\Omega)$. We may expand the inner product of $\mathcal{E}$ to see

$$\mathcal{E}[\phi + tw] = \lambda_1 + 2t\Re\mathcal{E}[\phi, w] + t^2\mathcal{E}[w]$$

for $t \in \mathbb{R}$. Doing the same to the norm,

$$\|\phi + tw\|_2^2 = 1 + 2t\Re\langle\phi, w\rangle + t^2\|w\|_2^2$$

This gives that

$$\frac{d}{dt}\mathcal{R}[\phi + tw]|_{t=0} = 2\Re\mathcal{E}[\phi, w] - 2\lambda_1\Re\langle\phi, w\rangle$$

Since $\mathcal{R}[\phi]$ is a minimum,

$$\frac{d}{dt}\mathcal{R}[\phi + tw]|_{t=0} = 0$$

which is to say

$$\Re\mathcal{E}[\phi, w] = \lambda_1\Re\langle\phi, w\rangle$$

for all $w \in H_0^1$. The same is true for the imaginary part by replacing w with iw , so

$$\mathcal{E}[\phi, w] = \lambda_1\langle\phi, w\rangle$$

for all $w \in H_0^1$, which is exactly the weak formulation.

Lastly, notice ϕ is smooth by applying elliptic regularity to the eigenvalue equation.

Q.E.D.

To move from the first eigenvalue to the general case, we will simply remove the subspace generated by the first eigenfunction from L^2 . In particular, we will take the orthocomplement

$$A^\perp = \{w \in L^2(\Omega) \mid \langle w, v \rangle = 0 \text{ for all } v \in A\}$$

We notice that $A^\perp \cap H_0^1$ is immediately closed with respect to the H^1 -norm, since the L^2 norm is dominated by it (the L^2 -inner product is continuous with respect to the H^1 -norm).

Proof:

[Proof of Spectral Theorem] Let $\phi_1 \in H_0^1(\Omega, \mathbb{R})$ be the eigenvector obtained from the previous theorem, normalized so $\|\phi_1\|_2 = 1$. Set

$$W_1 = \{\phi_1\}^\perp \cap H_0^1(\Omega)$$

which is closed and hence complete as we noticed above. Applying the minimization argument, or *Rayleigh Principle* to W instead of to H_0^1 gives a new minimum λ_2 and new eigenvector ϕ_2 that

is also a solution to the weak Helmholtz equation when looking at other functions in W . By the orthogonal construction, $\langle \phi_1, \phi_2 \rangle = 0$ and, by the weak solution property,

$$\mathcal{E}[\phi_1, \phi_2] = \lambda_1 \langle \phi_1, \phi_2 \rangle = 0$$

We claim ϕ_2 is a full weak solution to the Helmholtz equation. Indeed, take $v \in H_0^1$, and $v = c\phi_1 + w$ for $w \in W$. Therefore,

$$\mathcal{E}[\phi_2, c] - \lambda_2 \langle \phi_2, v \rangle = \mathcal{E}[\phi_2, w] - \lambda_2 \langle \phi_2, w \rangle = 0$$

as desired.

We may find further values by setting

$$W_k = \{\phi_1, \dots, \phi_k\}^\perp \cap H_0^1(\Omega)$$

and applying the same argument to find λ_{k+1} and ϕ_{k+1} . Our next claim to prove is that $\lambda_k \rightarrow \infty$. Assume for contradiction that this sequence is bounded. Then,

$$\|\phi_k\|_{H^1}^2 = 1 + \lambda_k$$

so that the ϕ_k are bounded. By Rellich's theorem, a subsequence converges in L^2 to some Φ . However, the ϕ_k are orthonormal in L^2 , so this is impossible.

Finally, we show that the $\{\phi_k\}$ are a basis for $L^2(\Omega)$. Set

$$W_\infty = \{\phi_1, \phi_2, \dots\}^\perp \cap H_0^1(\Omega)$$

and suppose a nonzero vector ψ is in this space. Then, applying Rayleigh's principle, we have a new eigenvalue λ . Since $W_k \supset W_\infty$, $\lambda > \lambda_k$ for all k . However, this is impossible since the eigenvalues grow to infinity. Hence, $W_\infty = \{0\}$.

Therefore, for any $\psi \in H_0^1$,

$$\sum \langle \psi, \phi_k \rangle \phi_k \rightarrow \psi$$

in the L^2 -norm. Therefore, the ϕ_k are a basis for H_0^1 under the L^2 norm. A vector v orthogonal to the ϕ_k is then orthogonal to H_0^1 , but H_0^1 is dense in L^2 , so the only such vector can be 0.

Q.E.D.

3.2 Estimation of Eigenvalues

The Rayleigh Principle allows us to gain the existence of eigenvalues theoretically, but it also allows us to estimate eigenvalues numerically since we may compute eigenvalues on some domains. Assume the Dirichlet eigenvalues $\{\lambda_k\}$ for a domain Ω are written in increasing order such that

$$\lambda_k = \min_{w \in W_{k-1} \setminus \{0\}} \mathcal{R}[w]$$

as in the previous section, with corresponding eigenfunction ϕ_k .

[3.4] Minimax Principle For a bounded domain $\Omega \subset \mathbb{R}^n$, let Λ_k denote the set of all k -dimensional subspaces of $H_0^1(\Omega)$. Then,

$$\lambda_k = \min_{V \in \Lambda_k} \{ \max_{u \in V \setminus \{0\}} \mathcal{R}[u] \}$$

for each $k \in \mathbb{N}$.

Proof:

Let $\{\phi_k\}$ be the eigenfunction basis as above. We then have

$$\mathcal{E}[\phi_i, \phi_j] = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$$

Set $V_e = \text{span}[\phi_1, \dots, \phi_k] \subset H_0^1(\Omega)$. For $u \in V_e$,

$$u = \sum c_j \phi_j \Rightarrow$$

$$\mathcal{E}[u] = \sum_{j=1}^k \lambda_j |c_j|^2$$

and

$$\mathcal{R}[u] = \frac{\sum_{j=1}^k \lambda_j |c_j|^2}{\sum |c_j|^2}$$

Since the eigenvalues are increasing, $\mathcal{R}[u] \leq \lambda_k$, and of course, $\mathcal{R}[\phi_k] = \lambda_k$ so that the maximum value of the Rayleigh quotient over V_e is attained. This justifies that the minimum is at most λ_k .

Suppose we take a general subspace $V \in \Lambda_k$. Then, by the dimension of the space, there must be some

$$w \in V \cap \text{span}[\phi_1, \dots, \phi_{k-1}]^\perp$$

from which we have, by the definition of our eigenvalues by minimization, $\lambda_k \leq \mathcal{R}[w]$. Therefore,

$$\lambda_k \leq \max_{u \in V \setminus \{0\}} \mathcal{R}[u]$$

which shows that the minimum in the theorem statement is exactly λ_k .

Q.E.D.

Remark: We shall not use or prove it here, but there is also a maxi-min principle that changes these orders:

$$\lambda_k = \max_{S \subset H_0^1(\Omega), \dim(S)=k-1} \left(\min_{u \in S^\perp \setminus \{0\}} \mathcal{R}[u] \right)$$

[3.5] Comparison Principle Consider two bounded domains in $\mathbb{R}^n$ satisfying $\Omega \subset \tilde{\Omega}$. Assuming the Dirichlet eigenvalues are arranged in increasing order,

$$\lambda_k(\tilde{\Omega}) \leq \lambda_k(\Omega)$$

Proof:

Let $\tilde{\Lambda}_k$ denote the k -dimensional subspaces of $H_0^1(\tilde{\Omega})$. Since V_e from the above proof is such a subspace (extending by 0s),

$$\lambda_k(\tilde{\Omega}) = \min_{W \in \tilde{\Lambda}_k} \left\{ \max_{u \in W \setminus \{0\}} \mathcal{R}[u] \right\} \leq \max_{u \in V_e \setminus \{0\}} \mathcal{R}[u] = \lambda_k(\Omega)$$

Q.E.D.

Recall that, for example, we know the eigenvalues on a rectangle $\prod[-a_i, a_i]$ since we wind up with sums

$$\left\{ \sum_{i=1}^n k_i^2 \mid k_i \in \frac{\pi}{a_i} \mathbb{Z} \right\}$$

and we can then put lower bounds on the eigenvalues depending on the size of the domain. We can make this even more precise.

Example: Consider the unit disk $B(0, 1) \subset \mathbb{R}^2$. Consider $w_\alpha(r) = 1 - r^\alpha$. Using the radial definition,

$$\begin{aligned}\mathcal{E}[w_\alpha] &= \int_0^1 \left(\frac{\partial w_\alpha}{\partial r} \right)^2 2\pi r dr \\ &= \pi\alpha\end{aligned}$$

and the L^2 norm is

$$\|w_\alpha\|_2^2 = \int_0^1 (1 - r^\alpha)^2 2\pi r dr = \frac{\pi\alpha^2}{2 + 3\alpha + \alpha^2}$$

so $\lambda_1 \leq \mathcal{R}[\alpha] = \frac{2}{\alpha} + 3 + \alpha$. Picking $\alpha = \sqrt{2}$ minimizes this expression, showing $\lambda_1 = 3 + 2\sqrt{2} \approx 5.828$. The exact value of λ_1 is the first root of the Bessel function squared, or approximately 5.783.

We can estimate higher eigenvalues using a strategy introduced by Walter Ritz in 1909, where we look at subspaces of H_0^1 . Let $A \subset H_0^1(\Omega)$ be a finite-dimensional subspace. Then, the approximate eigenvalues associated to A are

$$\eta_k = \min_{W \in \Lambda_k(A)} \max_{u \in W \setminus \{0\}} \mathcal{R}[u]$$

for $k = 1, \dots, \dim(A)$. In this finite dimensional case, we may rearrange the situation to be a matrix eigenvalue problem. The values η_k are associated the vectors $v_k \in A$ satisfying the weak eigenvalue equation

$$\mathcal{E}[v_k, w] = \eta_k \langle v_k, w \rangle$$

for all $w \in A$. Fix a basis $\{w_i\}$ for A . Then, define

$$E_{ij} = \mathcal{E}[w_i, w_j] \quad F_{ij} = \langle w_i, w_j \rangle$$

If $v = \sum c_j w_j$ in coordinates, then the weak equation is

$$\sum c_i (E_{ij} - \eta_k F_{ij}) = 0$$

In other words, the vector C is in the nullspace of $E - \eta_k F$.

So $\{\eta_1, \dots, \eta_k\} = \{\eta \mid \det(E - \eta F) = 0\}$, or that these are the eigenvalues of EF^{-1} . This is called the *Rayleigh-Ritz Scheme*.

Example:

Set $A = \text{span}\{w_1, w_2, w_3\}$ where $w_j(r) = 1 - r^{2j}$. Then,

$$E = \begin{bmatrix} 2\pi & \frac{8\pi}{3} & 3\pi \\ \frac{8\pi}{3} & 4\pi & \frac{24\pi}{5} \\ 3\pi & \frac{24\pi}{5} & 6\pi \end{bmatrix}$$

$$F = \begin{bmatrix} \frac{\pi}{3} & \frac{5\pi}{12} & \frac{9\pi}{20} \\ \frac{5\pi}{12} & \frac{8\pi}{15} & \frac{7\pi}{12} \\ \frac{9\pi}{20} & \frac{7\pi}{12} & \frac{9\pi}{4} \end{bmatrix}$$

where we see eigenvalues $\eta_1 = 5.783$, $\eta_2 = 30.712$, and $\eta_3 = 113.505$. In fact, $|\lambda_1 - \eta_1| \approx 10^{-6}$. However, $\lambda_2 \approx 14.682$, where we lost accuracy in this case because we focused solely in radial functions (the second

eigenfunction is orthogonal to A). The third eigenvalue, $\lambda_3 \approx 30.471$, is approximated quite well by λ_2 , however.

To prevent the type of issue above, we use the finite-element method to find A which covers $H_0^1(\Omega)$ well. This involves subdividing Ω into small polygons, producing a mesh (or triangulation). Each interior vertex is associated with an element, a piecewise linear function which is positive at this vertex and decays to 0 on the neighboring faces. The subspace A is the span of the elements. This produces an extremely adaptable and accurate method for computing the eigenvalues numerically.

4 Euler-Lagrange Problems

As we mentioned in the introduction, the development of the Calculus of Variations came from Lagrange and Euler working on problems in classical mechanics and representing solutions in terms of the minimization of an action functional. In particular, this is usually the integral of a *Lagrangian* $L(p, z, x)$ for

$$L : \mathbb{R}^n \times \mathbb{R} \times \bar{\Omega} \rightarrow \mathbb{R}$$

Then, the action functional is

$$I[w] = \int_{\Omega} L(\nabla w(x), w(x), x) dx$$

A critical point of $I[w]$, i.e. w so

$$\frac{d}{dt} I[w + t\psi]|_{t=0} = 0$$

for any test function ψ is a weak solution of an associated PDE called the *Euler-Lagrange Equation* associated to L . Indeed, we simply must compute the derivative to show this. Set $i(t) = I[w + t\psi]$. Then,

$$\begin{aligned} i'(t) &= \int_{\Omega} \sum_i L_{p_i}(\nabla w + t\nabla\psi, w + t\psi, x) \psi_{x_i} + L_z(\nabla w + t\nabla\psi, w + t\psi, x) \psi dx \\ 0 = i'(0) &= \int_{\Omega} \sum_i L_{p_i}(\nabla w, w, x) \psi_{x_i} + L_z(\nabla w, w, x) \psi dx \end{aligned}$$

or

$$0 = \int_{\Omega} \left[- \sum_i ((L_{p_i}(\nabla w, w, x))_{x_i}) + L_z(\nabla w, w, x) \right] \psi dx$$

for all test functions ψ . The associated Euler-Lagrange equation is then

$$- \sum_i ((L_{p_i}(\nabla w, w, x))_{x_i}) + L_z(\nabla w, w, x) = 0$$

Example: Dirichlet's principle relied on the Lagrangian

$$L(p, z, x) = \frac{1}{2} |p|^2$$

which generated equation $-\Delta u = 0$ in Ω .

Example: We may also create Poisson equation $-\Delta u = f(x)$ from

$$L(p, z, x) = \frac{1}{2}|p|^2 - zf(x)$$

and the nonlinear Poisson equation

$$-\Delta u = f(u)$$

from

$$L(p, z, x) = \frac{1}{2}|p|^2 - \int_0^z f(y)dy$$

Example: We may further produce *minimal surfaces* through this approach. Minimal surfaces are defined as being the surface minimizing the surface area while passing through some known regions (bubbles are good examples of creating such surfaces). This problem was first studied by Lagrange, though much work was done by Joseph Plateau who looked precisely at soap films.

The surface area of the graph of a function w is

$$\int_{\Omega} \sqrt{1 + |\nabla w|^2} dx$$

so this looks like the action functional of Lagrangian

$$L(p, z, x) = \sqrt{1 + |p|^2}$$

From our derivation above, this gives the *minimal surface equation*

$$\nabla \cdot \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0$$

which says that the surface has *zero mean curvature*.

In the previous section, we looked at a specific Laplacian and used the properties of coercivity to find a minimum. While the proof may be a bit too technical for us, this may be done with general Lagrangians under further assumptions. In particular, we call the action functional $I[w]$ *coercive* if

$$I[w] \geq \delta \|\nabla w\|_{L^2(\Omega)}^q - \gamma$$

This says roughly that $I[w] \rightarrow \infty$ as $\|\nabla w\|_{L^2} \rightarrow \infty$, which helps us show that we truly achieve a minimizer. We also used a continuity argument that does not hold in all cases. Therefore, we may define an action functional $I[w]$ to be *weakly lower semicontinuous* on $H^1(\Omega)$ provided

$$I[u] \leq \liminf_{k \rightarrow \infty} I[u_k]$$

whenever $\langle u_k, w \rangle_{L^2} \rightarrow \langle u, w \rangle_{L^2}$ for all $w \in L^2(\Omega)$. These two properties are enough to show existence.

[4.1] Assume that $I[w]$ is an action functional generated from Lagrangian L and is coercive and weakly lower semicontinuous. Suppose also that there are functions w so $I[w] < \infty$. Then, there exists at least one w so

$$I[w] = \min_u I[u]$$

If we ascribe some further restrictions to L (in particular, it cannot grow too quickly), then this minimizer w is a weak solution to the associated Euler-Lagrange Equation [See Evans Chapter 8 for a full discussion and Proof].

A helpful tool to arise from all of this theoretical quandry is the *Lagrange Multiplier*. In particular, if we want to restrain $I[w]$ by the condition

$$J[w] = \int_{\Omega} G(w) dx = 0$$

we define $A = \{w \in H_0^1(\Omega) \mid J[w] = 0\}$. When we have some restrictions on G ($|G'(z)| \leq C(|z| + 1)$), minimizers still exist and the following is true.

[4.2] Let $w \in A$ have

$$I[w] = \min_{u \in A} I[u]$$

Then, there exists a real number λ such that

$$\int_{\Omega} Dw \cdot Dv dx = \lambda \int_{\Omega} G'(w)v dx$$

for all $v \in H_0^1(\Omega)$.

For a physicist, the topic of Euler-Lagrange equations is one of the highest-reward topics in PDEs to study due to the sheer breadth of extensions of these results (for example, all of Lagrangian mechanics). For example, optimal control problems may also be reduced to Euler-Lagrange problems (as can several other stochastic minimization problems).

4.1 Finishing the Brastichrone Problem

Recall from the introduction that we had reduced the Brastichrone problem to finding the curve $(x, y(x))$ such that y minimizes

$$T(y) = \frac{1}{\sqrt{2g}} \int_{x_A}^{x_B} \sqrt{\frac{1 + (y')^2}{y}} dx$$

we can now solve this by using the Euler-Lagrange equations. While taking the derivatives is a bit laborious, the equation is, for $F(y, y')$ the integrand of the integral above,

$$0 = \frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = -\frac{1}{2y} - \frac{y''}{1 + (y')^2}$$

We may solve this by multiplying by $2y'$ and integrating,

$$\begin{aligned} 0 &= -\frac{y'}{y} - \frac{2y'y''}{1 + (y')^2} \\ c &= -\ln(y) - \ln(1 + (y')^2) \\ y(1 + (y')^2) &= C \\ y' &= \sqrt{\frac{C}{y} - 1} \end{aligned}$$

Solving this ODE is a bit tricky, but gives the parametric solution

$$\begin{cases} x = \frac{C}{2}(\theta - \sin(\theta)) \\ y = \frac{C}{s}(1 - \cos(\theta)) \end{cases}$$

which is a cycloid based on a circle of diameter C .