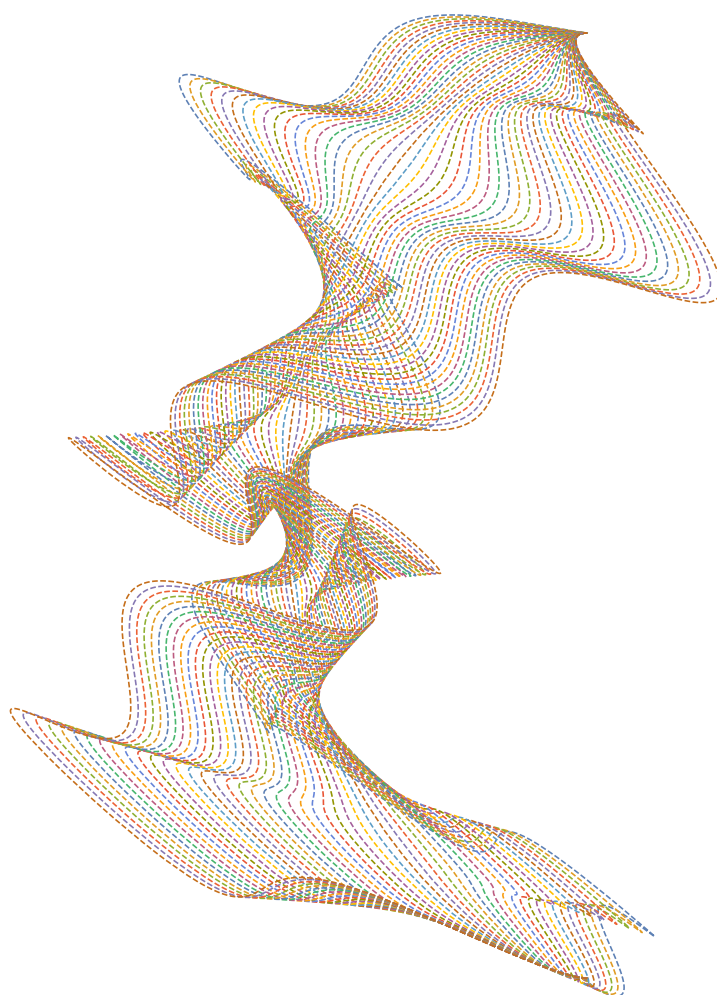


# Calculus of Variations & Partial Differential Equations

Huan Q. Bui

May 18, 2019



Spot the Cycloid!

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Calculus of Variations</b>	<b>1</b>
2.1	Euler-Lagrange Equation . . . . .	1
2.2	The Brachistochrone Problem . . . . .	3
2.3	Remark: Necessary and (almost) sufficient conditions . . . . .	4
<b>3</b>	<b>Euler-Lagrange equations as PDE's</b>	<b>4</b>
3.1	Solutions as Equations of Motion . . . . .	5
3.2	Example: Klein-Gordon equation for photons (the wave equation) . . . . .	5
<b>4</b>	<b>PDE's as Minimization Problems</b>	<b>7</b>
4.1	A Dirichlet-type Initial Boundary Value Problem . . . . .	7
4.2	Example: Poisson's Equation in Electrostatics . . . . .	9
<b>5</b>	<b>Above &amp; Beyond</b>	<b>10</b>

# 1 Introduction

From an applied viewpoint, calculus of variations poses a fascinating alternative approach to solving optimization problems: by considering small deviations from a solution (variations) and finding solutions based on what makes non-solutions sub-optimal. In a way, the idea of calculus of variations is much like the scientific process where the necessary conditions for a solution to a problem or a scientific theory of a natural phenomenon is often obtained by trial (deviations from the true solution) and error (how a deviation from the solution is penalized). It is thus not surprising that calculus of variation gives a *natural* and appealing way for finding physical laws. Calculus of variations appears in almost all corners of physics, often under the name of the Principle of Least Action: from Lagrangian's formalism of classical mechanics to the foundations of classical and quantum field theory, the Standard Model of physics and its extension (SME).

In physics, calculus of variations and differential equations go almost hand-in-hand. Physical laws are often written as differential equations. For example, Newton's law of gravitation relates the spatial gradient of a force field to the rate of change of the velocity of a particle in the field. On the other hand, many theories such as the Schrödinger's equation or the Einstein's field equations, while originally postulated, has been shown from time to time to be obtainable from variational methods. For instance, the Einstein's field equations can be obtained through variational methods applied to the metric tensor  $g_{\mu\nu}$ .

This paper attempts to provide an overview of the connection between calculus of variations and (partial) differential equations. Starting with some simple physically-motivated examples, this paper will derive the Euler-Lagrange equations and show how physical laws, expressed as a differential equation, can be generally obtained from variational methods in Lagrangian mechanics and beyond. Through an initial boundary value problem with Laplace's equation and inhomogeneous boundary condition, the paper will also show how this problem in partial differential equation can be written as a minimization problem that is solvable by variational methods. A similar problem in electrostatics will be considered to further illustrate this idea.

The final section of this paper shows a glimpse into the much more profound mathematical connection between partial differential equation and calculus of variations. In particular, the paper will briefly discuss how some of the questions in calculus of variations and partial differential equations are answered in functional analysis.

## 2 Calculus of Variations

### 2.1 Euler-Lagrange Equation

In this section, we will look at how calculus of variations is applied to a simple problem, and how the Euler-Lagrange equation arises as a necessary condition for a solution to an optimization problem.

Suppose we want to find the shortest arc joining two points on a plane. We are certain that the arc is a straight line. But how is a straight line joining two points is the shortest arc? The idea of calculus of variations is to consider a solution  $\bar{y}(x)$  that minimizes the distance  $S$  between the two points, say  $A(x_1, y_1)$  and  $B(x_2, y_2)$ , and some deviation  $\eta(x)$  from this correct curve  $\bar{y}(x)$ . Any deviation from the correct path associated with  $\eta(x)$  can thus be written as

$$y(x) = \bar{y}(x) + \epsilon\eta(x) \tag{1}$$

where  $\epsilon$  is a constant parameter controlling for the magnitude of the deviation  $\eta(x)$ . Since we want the end points of any general path to be the same as the correct path, it is required that  $\eta(x_1) = \eta(x_2) = 0$ .

The distance  $S(\epsilon)$  between  $A$  and  $B$  for any given  $\eta(x)$  can be found by a little bit of calculus:

$$S(\epsilon) = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx. \tag{2}$$

Before solving this problem, let us think about a more general case. Let the integrand be  $f = f(y', y, x) = f(\bar{y}' + \epsilon\eta', y + \epsilon\eta, x)$ , where  $x$  is the independent variable and  $f$  is sufficiently differentiable. Since we require that  $S(\epsilon)$  is minimized,  $dS(\epsilon)/d\epsilon = 0$ . Thus it is necessary that

$$\begin{aligned}
0 &= \frac{dS}{d\epsilon} \\
&= \frac{d}{d\epsilon} \int_{x_1}^{x_2} f(\bar{y}' + \epsilon\eta', y + \epsilon\eta, x) dx \\
&= \int_{x_1}^{x_2} \frac{d}{d\epsilon} f(\bar{y}' + \epsilon\eta', y + \epsilon\eta, x) dx \\
&= \int_{x_1}^{x_2} \eta' \frac{\partial f}{\partial y'} + \eta \frac{\partial f}{\partial y} dx
\end{aligned} \tag{3}$$

Consider the first term in the integrand. Integrating by parts gives.

$$\begin{aligned}
\int_{x_1}^{x_2} \eta' \frac{\partial f}{\partial y'} dx &= \eta \frac{\partial f}{\partial y'} \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx} \frac{\partial f}{\partial y} dx \\
&= - \int_{x_1}^{x_2} \frac{d}{dx} \frac{\partial f}{\partial y} dx,
\end{aligned} \tag{4}$$

where the boundary term vanishes due to the constraint  $\eta(x_1) = \eta(x_2) = 0$ . From (3) and (4), we have

$$0 = \int_{x_1}^{x_2} \eta \left( \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right) dx. \tag{5}$$

Now, because we require that  $\bar{y}(x)$  minimizes  $S$  and that this holds for any deviation  $\eta(x)$  from  $\bar{y}(x)$ , the integrand in (5) must be zero. Thus, we obtain the **Euler-Lagrange equation**:

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0} \tag{6}$$

Back to our original problem with finding the shortest arc joining two points on a plane. We have that

$$f(y', y, x) = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{7}$$

Thus,

$$\begin{cases} \frac{\partial f}{\partial y} = 0 \\ \frac{\partial f}{\partial y'} = \frac{1}{2} \frac{2y'}{\sqrt{1+y'^2}} = \frac{y'}{\sqrt{1+y'^2}}. \end{cases} \tag{8}$$

By the Euler-Lagrange equation,

$$0 - \frac{d}{dx} \left( \frac{y'}{\sqrt{1+y'^2}} \right) = 0, \tag{9}$$

i.e.,  $(y')\sqrt{1+y'^2}$  is some constant  $C$ , i.e.,

$$\begin{aligned}
y' &= C\sqrt{1+y'^2} \\
y'^2 &= C^2(1+y'^2) \\
(1-C^2)y'^2 &= C^2.
\end{aligned} \tag{10}$$

This says  $y' = dy/dx$  is a constant, which means

$$y(x) = ax + b \quad (11)$$

for some constants  $a, b$ . This is nothing but an equation for a line on a plane, as expected. Of course, this is a motivating example for that we can see how the Euler-Lagrange equation arises. The problem of minimizing the distance between two points on a plane certainly does not require calculus of variations, but we shall see in the following example how the Euler-Lagrange equation can come in handy.

## 2.2 The Brachistochrone Problem

Perhaps one of the most well-known examples of the superiority of Lagrangian mechanics over conventional methods in Newtonian mechanics is the problem of finding the frictionless path for an object to slide down with the shortest (*brachistos*) amount of time (*chronos*). This problem was originally posed by John Bernoulli in 1696 and attracted the attention of many eminent mathematicians and physicists at the end time including Isaac Newton, Leibniz, L'Hopital, and Johann Bernoulli, John Bernoulli's brother.

Here we are minimizing time, so we must first find an express for time. Assuming that the object travels from initial height  $c$  to final height  $d$ . With a bit of calculus and basic mechanics, we have

$$T = \int_0^L \frac{ds}{v} = \int_c^d \frac{\sqrt{1+x'^2}}{v} dy, \quad (12)$$

where  $v$  is the speed of the object. To express the speed  $v$  in terms of height  $y$ , we use conservation of energy. Assuming the total energy is zero and is conserved, we get

$$[\text{Potential Energy}] + [\text{Kinetic energy}] = -mgy + \frac{1}{2}mv^2 = 0. \quad (13)$$

Note that the sign convention is so that potential energy is negative. Thus,

$$v = \sqrt{2gy}. \quad (14)$$

Putting this into (12), we get

$$T = \frac{1}{\sqrt{2g}} \int_c^d \sqrt{\frac{1+x'^2}{y}} dy. \quad (15)$$

Let the integrand be  $f[x', x, y]$ , by the Euler-Lagrange equation we have

$$\frac{\partial f}{\partial x} - \frac{d}{dy} \frac{\partial f}{\partial x'} = 0 - \frac{d}{dy} \frac{1}{\sqrt{y}} \cdot \frac{x'}{\sqrt{1+x'^2}} = 0. \quad (16)$$

Thus,

$$\frac{x'}{\sqrt{y(1+x'^2)}} = \frac{1}{2a} \quad (17)$$

where  $a$  is some non-zero constant. It follows that

$$\begin{aligned} 0 &= (2ax')^2 - y(1+x'^2) \\ &= x'^2(2a-y) - y. \end{aligned} \quad (18)$$

Rearranging and integrating both sides, we get

$$x = \int_c^d \sqrt{\frac{y}{2a-y}} dy. \quad (19)$$

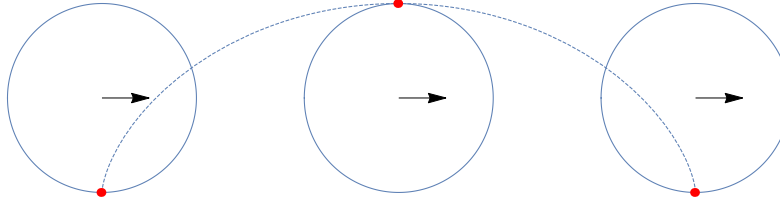
Now, we notice that  $0 \leq y \leq 2a$ , so we can make the substitution  $y = a(1 - \cos \theta)$ . Then,  $dy = a \sin \theta d\theta$ . And so,

$$\begin{aligned}
 x &= \int \sqrt{\frac{1 - \cos \theta}{1 + \cos \theta}} a \sin \theta d\theta \\
 &= a \int \sqrt{\frac{(1 - \cos \theta)^2}{1 - \cos^2 \theta}} \sin \theta d\theta \\
 &= a \int \frac{1 - \cos \theta}{\sin \theta} \sin \theta d\theta \\
 &= a \int 1 - \cos \theta d\theta \\
 &= a(\theta - \sin \theta) + C.
 \end{aligned} \tag{20}$$

We can define locations such that initially, the object is at  $(x, y) = (0, 0)$ . This gives  $C = 0$ . And so,

$$\begin{cases} x = a(\theta - \sin \theta) \\ y = a(1 - \cos \theta). \end{cases} \tag{21}$$

This is the parameterization for a **cycloid**, the curved traced by a point on a circle as it rolls without slipping in a straight line.



It is now clear that variation method can give us an elegant way to find a solution describing a physical system that is not very obvious to us from the start. One might find that attempting this problem using only Newtonian mechanics is much more challenging.

### 2.3 Remark: Necessary and (almost) sufficient conditions

It is very important to keep in mind that the Euler-Lagrange equation gives a *necessary condition* for some function to be a solution; i.e., if some function is a solution to the optimization problem in consideration then the function must satisfy the Euler-Lagrange equation. The Euler-Lagrange equation is not a sufficient condition [2]. By considering the Euler-Lagrange equation, we are only concerned with the first-order change (or first variation) in the action. To obtain sufficient conditions, we need to look for second-order change (or second-variation) in the action. It turns out that in most cases, especially in physical systems, a first variation in the Lagrangian is sufficient for finding a minimizing solution. [3]

## 3 Euler-Lagrange equations as PDE's

The Euler-Lagrange equation is a partial differential equation. In practice, however, we often don't solve the Euler-Lagrange equation but instead solve the differential equation that is *obtained from* the Euler-Lagrange equation for a given Lagrangian  $\mathcal{L}$ . For physical systems, the differential equation obtained from the Euler-Lagrange equation is often referred to as the "equations of motion," while solutions to equations of motion are often the trajectory (either

in time, space, or both) of a particle. In this section, we will look at two examples: one in classical mechanics and one in classical field theory to show how this process works in detail. The point of these examples is to demonstrate how the Euler-Lagrange equation is used in physics, and how powerful a method it is for finding physical laws/equation of motion when the solution is not obvious.

### 3.1 Solutions as Equations of Motion

“Mr. Bader told me the following: Suppose you have a particle (in a gravitational field, for instance) which starts somewhere and moves to some other point by free motion—you throw it, and it goes up and comes down. It goes from the original place to the final place in a certain amount of time. Now, you try a different motion. Suppose that to get from here to there, but got there in just the same amount of time. Then he said this: If you calculate the kinetic energy at every moment on the path, take away the potential energy, and integrate it over the time during the whole path, you’ll find that the number you’ll get is *bigger* than that for the actual motion. In other words, the laws of Newton could be stated not in the form  $\vec{F} = m\vec{a}$  but in the form: the average kinetic energy less the average potential energy is as little as possible for the path of an object going from one point to another.”

- Feynman Lectures on Physics, *The Principle of Least Action* [1]

One aspect of the Brachistochrone problem that makes it a “classic” apart from its unexpected solution and interesting history is the prescription it provides for finding the equations of motion for certain physical systems. This prescription is as follows:

1. Determine the Lagrangian, which in most cases is the “kinetic energy” pieces minus the “potential energy” pieces. The same principle applies to physics beyond Lagrangian (classical) mechanics.
2. (a) Either apply the Euler-Lagrange equation to the Lagrangian to get a system of relationships among the physical quantities  
(b) Or if it is unclear how to proceed with the Euler-Lagrange equations, vary the action functional with respect to some independent variable and obtain the Euler-Lagrange equation.
3. Simplify and obtain the physical law.

For example, suppose we want to “derive” Hooke’s law from only the energy terms. We first set up the Lagrangian:

$$\mathcal{L} = [\text{Potential energy}] - [\text{Kinetic energy}] = \frac{1}{2}kx^2 - \frac{1}{2}m\dot{x}^2. \quad (22)$$

where  $x$  is the position and  $\dot{x}$  is the velocity of the particle. In independent variable here is time,  $t$ . Applying the Euler-Lagrangian equation:

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} \implies kx = -\frac{d}{dt}(m\dot{x}) = -m\ddot{x}. \quad (23)$$

This is of course a silly example just to demonstrate how this method works. We can only really see how this method becomes useful when the Lagrangian is known but it is not clear what the equation of motion is. Let us look at a slightly more complicated example in the following subsection.

### 3.2 Example: Klein-Gordon equation for photons (the wave equation)

In classical field theory, the independent variable in the Euler-Lagrange equation becomes the field  $\phi$ , which essentially is a function that associates every point in space and time with some number. Here, we will use Einstein’s notation for summation, where  $\mu = 0, 1, 2, 3$  is an index, and  $\partial_\mu$  is defined as

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad (24)$$

where  $\cdot^\mu$  denotes the coordinates  $x^0 = t, x^1 = x, x^2 = y, x^3 = z$ . With this notation and the change in the independence variable, we have that for a given Lagrangian  $\mathcal{L}$ , the Euler-Lagrange equation is

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0. \quad (25)$$

For massless photons, it can be shown from variational methods that the Lagrangian is given by

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi) (\partial^\mu \phi), \quad (26)$$

where this “quadratic” term is the analogue of kinetic energy. Because photons are massless, we don’t have a potential term here ( $E^2 = c^2 p^2 + m^2 c^4 = c^2 p^2$ ). Plugging this Lagrangian into the Euler-Lagrange equation, we get

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \phi} = 0 \\ \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \frac{1}{2} \partial^\mu \phi. \end{cases} \quad (27)$$

So the Euler-Lagrange equation gives:

$$\partial_\mu (\partial^\mu \phi) = 0. \quad (28)$$

Now, the operator

$$\partial_\mu \partial^\mu = \partial^\mu \partial_\mu = \eta_{\mu\nu} \partial^\mu \partial^\nu = \square \quad (29)$$

is called the **d’Alembertian**, and the sign convention is defined by the Minkowski metric

$$[\eta_{\mu\nu}] = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}. \quad (30)$$

If we expand out the terms, with the 0 index associated with the time variable  $t$ , and 1, 2, 3 associated with the spatial components  $x, y, z$ , then we get from (28)

$$\boxed{\frac{\partial^2}{\partial t^2} \phi = \frac{\partial^2}{\partial x^2} \phi + \frac{\partial^2}{\partial y^2} \phi + \frac{\partial^2}{\partial z^2} \phi = \nabla^2 \phi} \quad (31)$$

This is not surprisingly the wave equation (or the Klein-Gordon equation for the massless particles). We can in fact verify that the solution to this equation is

$$\phi = e^{-i\mathbf{k} \cdot \mathbf{x}} \quad (32)$$

where

$$\mathbf{k} = K^\mu = (K^0, \vec{K}) \quad (33)$$

$$\mathbf{x} = X^\mu = (X^0, \vec{X}) \quad (34)$$

are the wavenumber vector and the position vector respectively. The first derivative is

$$\begin{aligned} \partial_\mu \phi &= \partial_\mu (e^{-i\mathbf{k} \cdot \mathbf{x}}) \\ &= -i \partial_\mu (\mathbf{k} \cdot \mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} \\ &= -i \partial_\mu (K_\nu X^\nu) e^{-iK_\alpha X^\alpha} \\ &= -i K_\nu \partial_\mu X^\nu \phi \\ &= -i K_\nu \delta_\mu^\nu \phi \\ &= -i K_\mu \phi \end{aligned} \quad (35)$$



Next, we compute the second derivative:

$$\begin{aligned}
 \partial^\mu \partial_\mu \phi &= \eta^{\mu\nu} \partial_\nu \partial_\mu \phi \\
 &= \eta^{\mu\nu} \partial_\nu (-iK_\mu \phi) \\
 &= -iK_\mu \eta^{\mu\nu} (-iK_\nu \phi) \\
 &= (-i)^2 K^\mu K_\mu \phi.
 \end{aligned} \tag{36}$$

If  $K^\mu K_\mu = m^2 = 0$  ( $K^\mu K_\mu$  is associated with the mass of a particle in the de Broglie relation), then

$$\square \phi = 0, \tag{37}$$

which satisfies the Klein-Gordon equation.

This result makes good sense, because photons can be described as propagating electromagnetic waves. While this result might not be too surprising, we can see that if the Lagrangian is somehow modified (which is often the case if, say, we want to describe a different particle with mass and spin, etc), then we should expect a different equation of motion from the Euler-Lagrange equation. It is truly remarkable that by specifying certain characteristics of a particle, we can find an equation describing exactly what the particle does in space and time.

## 4 PDE's as Minimization Problems

In this section we shift our focus to how initial boundary value problems can be expressed as minimization problems and solved using variational methods. First, a Dirichlet-type initial boundary value problem is discussed. Then, an example in electrostatics is given to demonstrate how a PDE can be expressed as a minimization problem and vice versa. We will consider two examples, one from the mathematical theory of PDE, and another from electrostatics. Again, the goal of these examples is to illustrate how the re-formulation of PDE's as minimization problem works and how a solution can be obtained via this reformulation.

### 4.1 A Dirichlet-type Initial Boundary Value Problem

Consider the following initial boundary value problem.

$$(*) \begin{cases} \nabla^2 u = 0 & \text{in } \Omega \\ u = f & \text{on } \partial\Omega. \end{cases} \tag{38}$$

One with some familiarity with partial differential equation may recognize that this is a Laplace's equation with inhomogeneous boundary condition. One can solve this problem by transforming it to get homogeneous boundary condition, but in context of calculus of variations, it turns out that for some admissible function  $u$ ,

$$\boxed{u \text{ solves } (*) \iff u \text{ minimizes } S[w] = \frac{1}{2} |\nabla w|^2 dx} \tag{39}$$

*Sketch of Proof.*

1. ( $\implies$ ) We want to show that for any solution  $u$  of  $(*)$ ,  $S[u] \leq S[w + u]$  for any admissible  $w$ . In context of calculus of variations, the function  $w$  satisfying the boundary conditions is the analogue of the deviation  $\eta$  in the derivation of the Euler-Lagrange equation. Thus, it is natural to consider the perturbed solution  $u' = u + w$  and the action functional associated with it. Let a solution  $u$  to  $(*)$  be given. Then  $\nabla^2 u = 0$ . Let an admissible  $w$  satisfying the boundary condition  $w = f$  in  $\partial\Omega$  be given. Then it is true that

$$\int_{\Omega} w \nabla^2 u dx = 0. \tag{40}$$

Integration by parts gives

$$\begin{aligned} \int_{\Omega} w \nabla^2 u \, dx &= w \nabla u \Big|_{\partial\Omega} - \int_{\Omega} \nabla u \cdot \nabla w \, dx \\ &= - \int_{\Omega} \nabla u \cdot \nabla w \, dx. \end{aligned} \quad (41)$$

Now, we look at the action associated with the solution  $u$  perturbed by some amount  $w$ :

$$\begin{aligned} S[u'] &= S[u + w] = \frac{1}{2} \int_{\Omega} |\nabla(u + w)|^2 \, dx \\ &= \frac{1}{2} \int_{\Omega} \nabla(u + w) \cdot \nabla(u + w) \, dx \\ &= \frac{1}{2} \int_{\Omega} (\nabla u \cdot \nabla u + 2\nabla u \cdot \nabla w + \nabla w \cdot \nabla w) \, dx, \quad \text{by linearity of } \nabla \\ &= \frac{1}{2} \int_{\Omega} \nabla u \cdot \nabla u + \nabla w \cdot \nabla w \, dx \\ &= \frac{1}{2} \int_{\Omega} |\nabla u|^2 \, dx + \frac{1}{2} \int_{\Omega} |\nabla w|^2 \, dx \\ &= S[u] + S[w] \\ &\geq S[u]. \end{aligned} \quad (42)$$

Thus,  $S[u] \leq S[u + w]$  for any admissible perturbation  $w$ . This means  $u$  minimizes the action  $S[\cdot]$ .

2. ( $\Leftarrow$ ) Here we want to show that if  $u$  minimizes the action functional  $S[\cdot]$  then  $u$  solves (\*). Here the idea of calculus of variations comes in handy. Suppose that  $u$  satisfies the boundary condition  $u = f$  in  $\partial\Omega$  and minimizes the action  $S[\cdot]$ . Now, consider some perturbation in  $u$ , i.e., we let

$$u \rightarrow u + \epsilon w \quad (43)$$

for some constant  $\epsilon$  and function  $w$  that satisfies the boundary condition of (\*). Once again, we consider the action associated with this perturbed  $u$ :

$$\begin{aligned} S[u + \epsilon w] &= \frac{1}{2} \int_{\Omega} |\nabla(u + \epsilon w)|^2 \, dx \\ &= \frac{1}{2} \int_{\Omega} |\nabla u|^2 + 2\epsilon \nabla u \cdot \nabla w + \epsilon^2 |\nabla w|^2 \, dx \quad \text{by linearity of } \nabla. \end{aligned} \quad (44)$$

Now, because  $u$  minimizes  $S[\cdot]$ ,  $\partial S / \partial \epsilon = 0$  at  $\epsilon = 0$ , so we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial \epsilon} S[u + \epsilon w] \Big|_{\epsilon=0} \\ &= \int_{\Omega} \nabla u \cdot \nabla w \, dx + \int_{\Omega} \epsilon |\nabla w|^2 \, dx \Big|_{\epsilon=0} \\ &= \int_{\Omega} \nabla u \cdot \nabla w \, dx. \end{aligned} \quad (45)$$

But since we have argued that in (41)

$$\int_{\Omega} w \nabla^2 u \, dx = - \int_{\Omega} \nabla u \cdot \nabla w \, dx. \quad (46)$$

Therefore,

$$\int_{\Omega} \nabla u \cdot \nabla w \, dx = 0, \quad (47)$$

which must hold for any perturbation  $w$ . This means  $\nabla^2 u = 0$ . But since  $u$  also satisfies the boundary condition of (\*),  $u$  solves (\*). □

## 4.2 Example: Poisson's Equation in Electrostatics

In this subsection we are looking at home we can obtain Poisson's equation for electrostatics from a minimization problem. Suppose we are given some charge distribution  $\rho$  and want to know what the potential (scalar) field associated with this charge distribution is. One way to solve this problem is to consider solving Poisson's equation:

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0}. \quad (48)$$

Alternatively, it turns out that this PDE can be expressed as the following problem: Find  $\phi$  such that the action

$$J = \frac{\epsilon_0}{2} \int |\nabla \phi|^2 \, dV - \int \rho \phi \, dV \quad (49)$$

is minimum. We shall verify that the solution to this problem is indeed Poisson's equation for electrostatics, via variational methods.

Let us write a general potential  $\phi$  as  $\phi = \bar{\phi} + \epsilon F$  where  $\bar{\phi}$  is the correct potential and  $F$  is the variation from this correct potential. It follows that

$$|\nabla \phi|^2 = |\nabla \bar{\phi}|^2 + 2\epsilon \nabla \bar{\phi} \cdot \nabla F + \epsilon^2 |\nabla F|^2. \quad (50)$$

Like how we have argued before, because  $\bar{\phi}$  minimizes  $J$ ,  $\partial J / \partial \epsilon = 0$  at  $\epsilon = 0$ , so we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial \epsilon} \left( \frac{\epsilon_0}{2} \int |\nabla \phi|^2 \, dV - \int \rho \phi \, dV \right) \Big|_{\epsilon=0} \\ &= \frac{\partial}{\partial \epsilon} \left( \frac{\epsilon_0}{2} \int |\nabla \bar{\phi}|^2 + 2\epsilon \nabla \bar{\phi} \cdot \nabla F + \epsilon^2 |\nabla F|^2 \, dV - \int \rho(\bar{\phi} + \epsilon F) \, dV \right) \Big|_{\epsilon=0} \\ &= \dots \\ &= \int (\epsilon_0 \nabla \bar{\phi} \cdot \nabla F - \rho F) \, dV. \end{aligned} \quad (51)$$

Using (46), we simplify this to

$$0 = \int -\epsilon_0 F \nabla^2 \bar{\phi} - \rho F \, dV = \int (-\epsilon_0 \nabla^2 \bar{\phi} - \rho) F \, dV, \quad (52)$$

which has to be true for all  $F$ . Therefore,

$$\nabla^2 \bar{\phi} = -\frac{\rho}{\epsilon_0}, \quad (53)$$

which is the Poisson's equation for electrostatics as desired. [1]

## 5 Above & Beyond

So far, we have seen how it is possible for problems in partial differential equations to be formulated as minimization problems and how minimization problems can be written as PDE's. To formulate minimization problems as PDE's, we use variational methods to obtain the Euler-Lagrange equation. For a known Lagrangian, the Euler-Lagrange equation often turns into a second-order ordinary differential equation (in some independent variable), which we often know how to solve. It is not entirely clear, however, *how* one can formulate PDE's as minimization problems. We observe that in both of our examples (the Dirichlet-type initial boundary value problem and Poisson's equation in electrostatics), we were given the quantities which we were trying to minimize. What we did not do was *actually finding* this quantity. So, a natural question is how, in the general theory, one can obtain the to-be-minimized quantity from a given PDE.

Unfortunately, the answer lies in the theory of functional analysis, which is beyond the scope of this paper. Loosely speaking, the idea is not to solve an initial boundary value problem directly and instead formulating it in a way such that **weak solutions** can be found. A weak solution to an initial boundary value problem, as its name suggests, is essentially one that satisfies the problem when *tested* against other admissible candidates, and it is *weak* in the sense that it is not necessarily the absolute solution to the problem. This is exactly the idea of calculus of variations where we obtain some necessary conditions for a solution by introducing some kind of perturbation to a potential solution. For second-order elliptic problems such as ones involving Laplace's and Poisson's equations that we have seen, the general theory tells us that weak solutions exist and are unique in certain circumstances. In these cases, a weak solution minimizes some action functional (which comes up as an inner product in the theory) in the form of (39) or (49). For more information on this topic, refer to Chapter 6: Second-order elliptic equations, in Evans' *Partial Differential Equations*.

## References

- [1] Feynman, R.P. and Leighton, R.B. and Sands, M. and Hafner, EM. 1965. 19. The Principle of Least Action. In. *The Feynman Lectures on Physics*.
- [2] Bliss, G. 1944. *Calculus of Variations*, The Mathematical Association of America.
- [3] Farlow, S.J. 1993. Calculus of Variations (Euler-Lagrange Equations). In. *Partial Differential Equations for Scientists and Engineers*, Dover Publications, Inc. New York.
- [4] Evans, L.C. 1997. *Partial Differential Equations*, American Mathematical Society.