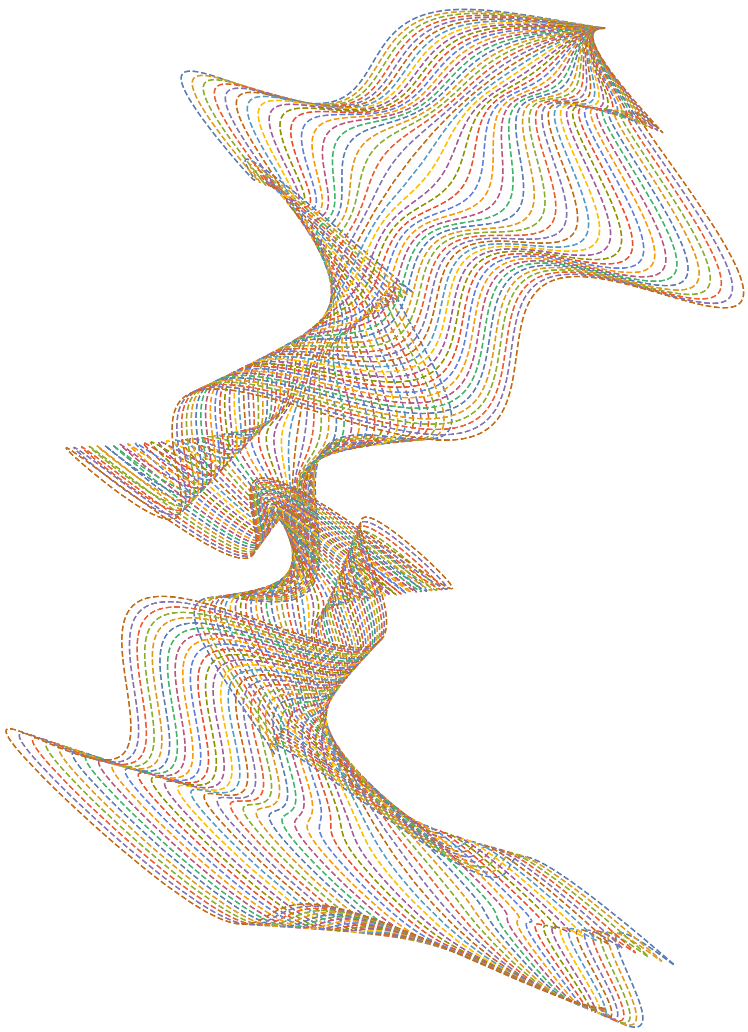
Calculus of Variations
and
Partial Differential Equations

Huan Q. Bui

May 18, 2019



Contents

1 Introduction

From an applied viewpoint, calculus of variations poses a fascinating indirect approach to solving optimization problems: considering small deviations from a solution (variations) and finding solutions based on what makes non-solutions are sub-optimal. In a way, the idea of calculus of variations is much like the scientific process where the necessary conditions for the truth - a solution to a problem, or a theory of a natural phenomenon - is often obtained by trial and error. 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 and the Standard Model.

Calculus of variations and differential equations go almost hand-in-hand in physics. Physical laws are often written as differential equations. And while many theories, such as the Schrödinger's equation or the Eintein's field equations, were postulated, it has been shown from time to time that these laws could be obtained from variational methods, which are often purely mathematical.

This paper attempts to provide an overview of the connection between calculus of variations and (partial) differential equations. Starting with physically-motivated examples, the Euler-Lagrange equations and how physical laws (differential equations) can be obtained from variational methods. Through a general initial boundary value problem of an inhomogeneous Laplace's equation with Dirichlet boundary condition, the paper will also show how this problem in partial differential equation can be written as a minimization problem solvable by variational methods.

What else here?

1.1 Euler-Lagrange's Equation

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 $\bar{y}(x)$. Thus any deviation from the correct path associated with $\eta(x)$ can be written as

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

where η 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 admissible. Since we require that S is minimized,

$dS(\epsilon)/d\epsilon = 0$. Thus it is necessary that

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

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

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

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

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

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

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

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

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

Thus,

$$
\begin{aligned}
\frac{\partial f}{\partial y} &= 0 \qquad (8) \\
\frac{\partial f}{\partial y'} &= \frac{1}{2}\frac{2y'}{\sqrt{1+y'^2}} = \frac{y'}{\sqrt{1+y'^2}}. \qquad (9)
\end{aligned}
$$

By the Euler-Lagrange equation,

$$
0 - \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}} \right) = 0, \qquad (10)
$$

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

$$
\begin{aligned}
y' &= C\sqrt{1+y'^2} \\
y'^2 &= C^2(1+y'^2) \\
(1-C^2)y'^2 &= C^2. \qquad (11)
\end{aligned}
$$

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

$$y(x) = ax + b \tag{12}$$

for some constants a, b. This is nothing but an equation for a line on a plane as expected.

1.2 The Brachistochrone Problem

Perhaps one of the most well-known examples of the superiority of Lagrangian mechanics over the conventional methods in Newtonian mechanics is the problem of find 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 (or more accurately, *natural philosophers*) at the end time including 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 a to final height b. With a bit of calculus and basic mechanics, we have

$$T = \int_0^L \frac{ds}{v} = \int_a^b \frac{\sqrt{1+x'^2}}{v}\,dy, \tag{13}$$

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

$$[\text{Potential Energy}] - [\text{Kinetic energy}] = mgy - \frac{1}{2}mv^2 = 0. \tag{14}$$

Thus,

$$v = \sqrt{2gy}. \tag{15}$$

Putting this into (13), we get

$$T = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1+x'^2}{y}}\,dy. \tag{16}$$

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

$$\frac{\partial f}{\partial x} - \frac{d}{dy}\frac{\partial f}{\partial x'} = 0 - \frac{d}{dy}\frac{1}{\sqrt{y}} \cdot \frac{x'}{\sqrt{1+x'^2}} = 0. \tag{17}$$

Thus,

$$\frac{x'}{\sqrt{y}\sqrt{1+x'^2}} = \frac{1}{2a} \tag{18}$$

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

$$\begin{aligned} 0 &= (2ax')^2 - y(1+x'^2) \\ &= x'^2(2a-y) - y. \end{aligned} \tag{19}$$

Rearranging and integrating both sides, we get

$$x = \int \sqrt{\frac{y}{2a-y}}\,dy. \tag{20}$$

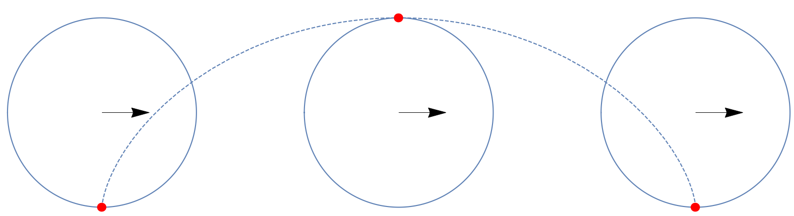
Now, we notice that $0 \leqslant y \leqslant 2a$, so we can make the substitution $y = A(1-\cos\theta)$. Then, $dy = A\sin\theta\,d\theta$. And so,

$$\begin{aligned} x &= \int \sqrt{\frac{1-\cos\theta}{1+\cos\theta}} A\sin\theta\,d\theta \\ &= A\int\sqrt{\frac{(1-\cos\theta)^2}{1-\cos^2\theta}}\sin\theta\,d\theta \\ &= A\int\frac{1-\cos\theta}{\sin\theta}\sin\theta\,d\theta \\ &= A\int 1-\cos\theta\,d\theta \\ &= A(\theta-\sin\theta)+C. \end{aligned} \tag{21}$$

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

$$\begin{cases} x = A(\theta-\sin\theta) \\ y = A(1-\cos\theta). \end{cases} \tag{22}$$

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



2 Euler-Lagrange equations as PDE's

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 "equation of motion" for certain physical systems:

1. 1. Determine the Lagrangian.
2. 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 desired.

For example, suppose we would like to "derive" Hooke's law from only the energy terms. We first set up the Lagrangian:

$$\mathcal{L} = [Potentialenergy] - [Kineticenergy] = \frac{1}{2}kx^2 - \frac{1}{2}m\dot{x}^2. \tag{23}$$

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:

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} \implies kx = -\frac{d}{dt}(m\dot{x}) = -m\ddot{x}. \tag{24}$$

This is of course a silly example just to demonstrate how this method works. This method only becomes useful when the Lagrangian is known but it is not clear what the "equations of motion" are. For example, suppose we would like to know

3 PDE's as Minimization Problems

Consider the following initial boundary value problem.

$$(*) \begin{cases} \nabla^2 u = 0 & \text{in } \Omega \\ u = f & \text{on } \partial\Omega. \end{cases} \tag{25}$$

It turns out that

$$\boxed{u \text{ solves } (*) \iff u \text{ minimizes } S[w] = \frac{1}{2}|\nabla w|^2\, dx} \tag{26}$$

Sketch of Proof.

1. ($\implies$) We want to show that for any solution u of $(*)$, $S[u] \leqslant S[w+u]$ for any admissible w. In context of calculus of variations, the function w satisfying the boundary conditions is the analogue of the deviation η in the derivation of the Euler-Lagrange equation. Thus, it is natural to consider the perturbed solution $u' = u + w$ and what 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

$$\int_\Omega w\nabla^2 u\, dx = 0. \tag{27}$$

Integration by parts gives

$$\begin{aligned} \int_\Omega w\nabla^2 u\, dx &= w\nabla u\Big|_{\partial\Omega} - \int_\Omega \nabla u \dot{\nabla} w\, dx \\ &= -\int_\Omega \nabla u \cdot \nabla w\, dx. \end{aligned} \tag{28}$$

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

$$\begin{aligned} S[u'] = S[u+w] &= \frac{1}{2}\int_\Omega |\nabla(u+w)|^2\, dx \\ &= \frac{1}{2}\int_\Omega \nabla(u+w)\cdot\nabla(u+w)\, dx \\ &= \frac{1}{2}\int_\Omega \left(\nabla u\cdot\nabla u + 2\nabla u\cdot\nabla w + \nabla w\cdot\nabla w\right)\, dx, \quad \text{by linearity of } \nabla \\ &= \frac{1}{2}\int_\Omega \nabla u\cdot\nabla u + \nabla w\cdot\nabla w\, dx \\ &= \frac{1}{2}\int_\Omega |\nabla u|^2\, dx + \frac{1}{2}\int_\Omega |\nabla w|^2\, dx \\ &= S[u] + S[w] \\ &\geqslant S[u]. \end{aligned} \tag{29}$$

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

2. ($\Longleftarrow$) Here we want to show that if u minimizes the action functional $S[\,]$ 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 $S[\,]$. Now, consider some perturbation in u, i.e., we let

$$u \to u + \epsilon w \tag{30}$$

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

$$\begin{aligned} S[u+\epsilon w] &= \frac{1}{2}\int_\Omega |\nabla(u+\epsilon w)|^2\,dx \\ &= \frac{1}{2}\int_\Omega |\nabla u|^2 + 2\epsilon\nabla u\nabla w + \epsilon^2|\nabla w|^2\,dx \qquad \text{by linearity of } \nabla. \end{aligned} \tag{31}$$

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

$$\begin{aligned} 0 &= \frac{\partial}{\partial\epsilon} S[u+\epsilon w]\bigg|_{\epsilon=0} \\ &= \int_\Omega \nabla u\cdot\nabla w\,dx + \int_\Omega \epsilon|\nabla w|^2\,dx\bigg|_{\epsilon=0} \\ &= \int_\Omega \nabla u\cdot\nabla w\,dx. \end{aligned} \tag{32}$$

But recall we have argued that

$$\int_\Omega w\nabla^2 u\,dx = -\int_\Omega \nabla u\cdot\nabla w\,dx.$$

Therefore,

$$\int_\Omega \nabla u\cdot\nabla w\,dx = 0, \tag{33}$$

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

$\square$

4 Above and Beyond

References

[1] Huynen, M. A. and Bork, P. 1998. Measuring genome evolution. *Proceedings of the National Academy of Sciences USA* 95:5849–5856.

[2] Caprara, A. 1997. Sorting by reversals is difficult. In: *Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB 97),* New York: ACM. pp. 75-83.

[3] McLysaght, A., Seoighe, C. and Wolfe, K. H. 2000. High frequency of inversions during eukaryote gene order evolution. In Sankoff, D. and Nadeau, J. H., editors, *Comparative Genomics*, Dordrecht, NL: Kluwer Academic Press. pp. 47–58.

[4] Reinelt, G. 1991. *The Traveling Salesman - Computational Solutions for TSP Applications.* Berlin: Springer Verlag.