PY501 - Mathematical Physics

Hongwan Liu

Table of Contents

1 Calculus of Variations

References: Stone & Goldbart (SG) Chapter 1; Byron & Fuller (BF) Chapter 2; Arfken, Weber & Harris (AWH) Chapter 22.

Finding where the minimum or maximum value of some quantity occurs is an extremely common task. For example, you might want to know where the highest point on a map is, or when you had the highest heartrate throughout the day. Mathematically, we have a function $f(x)$, and we want to find the value of x which maximizes or minimizes $f(x)$. To do so, for a differentiable function f, we simply take the derivative and set it to zero

$$f'(x) = 0 \tag{1}$$

and solve for x to find the stationary points of the function.

Often though, we run into problems where we want to find the *function* at which the minimum or maximum value of some function occurs. Some examples include:

1. What is the shortest path to take between points A and B?
2. What closed curve of fixed length encloses the maximum possible area?
3. What form does a hanging heavy chain of fixed length take, so as to minimize its potential energy?

To answer these questions mathematically, we need an object called a **functional** $J[y]$, which maps smooth[1] functions y (e.g. a path, a curve) to a real number (e.g. a distance, an area). This is just another map, like a function is. But now, we want to develop the tools required to define a **functional derivative** such that setting

$$\frac{\delta J}{\delta y(x)} = 0 \tag{2}$$

will allow us to find a function $y(x)$ (e.g. a path, a curve) that maximizes $J[y]$ (e.g. a distance, an area).

[1] "Smooth" means that all derivatives of the function exists. We won't ever be interested in subtleties involving continuity and differentiability in this course.

1.1 Functionals

What does a functional look like? For our purposes, we will be dealing with functionals that have the following form:

$$J[y] = \int_{x_1}^{x_2} dx\, f(x, y, y', y'', \cdots, y^{(n)})\,, \tag{3}$$

where f is a function of the real numbers x, y, $y' \cdots$, *independently*.[2] We call these kinds of functionals **local** in x. As you can see, J takes in a function y, performs an integral, and returns a real number, which is exactly what a functional should do.

[2] This is a cause of endless confusion, so pay attention! From the perspective of f, y and y' are *independent variables*.

1.1.1 The functional derivative

Let us work out the functional derivative for the case where

$$J[y] = \int_{x_1}^{x_2} dx\, f(x, y, y')\,. \tag{4}$$

To do this, suppose we make an infinitesimal shift $y(x) \to y(x) + \varepsilon\eta(x)$, where ε is an infinitesimally small constant,[3] and $\eta(x)$ is some arbitrary function. Then

[3] i.e. with ε^2 and higher powers of ε all being zero.

the shift in J is

$$\delta J = J[y + \varepsilon\eta] - J[y] = \int_{x_1}^{x_2} dx\, \left[f(x, y + \varepsilon\eta, y' + \varepsilon\eta') - f(x, y, y') \right] . \quad (5)$$

Since ε is an infinitesimal quantity, we can perform a Taylor expansion up to first order about y and y' to find

$$\delta J = \int_{x_1}^{x_2} dx\, \left[\varepsilon\eta \frac{\partial f}{\partial y} + \varepsilon\eta' \frac{\partial f}{\partial y'} \right] . \quad (6)$$

To make further progress, we integrate the second term by parts, giving

$$\delta J = \left[\varepsilon\eta \frac{\partial f}{\partial y'} \right]_{x_1}^{x_2} + \int_{x_1}^{x_2} dx\, \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right] \varepsilon\eta\, . \quad (7)$$

We are frequently—but not always!—concerned with finding functions of y with fixed endpoints;[4] in that case, $\eta(x_1) = \eta(x_2) = 0$, and the boundary terms in the first term on the right vanishes, leaving

[4] For example, if we are interested in finding the path with the short distance between two fixed points.

$$\delta J = \int_{x_1}^{x_2} dx\, \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right] \varepsilon\eta\, . \quad (8)$$

This can be written suggestively as

$$\delta J = \int_{x_1}^{x_2} dx\, \delta y(x) \left(\frac{\delta J}{\delta y(x)} \right) , \quad (9)$$

where $\delta y(x) \equiv \varepsilon\eta(x)$, and

$$\boxed{\frac{\delta J}{\delta y(x)} \equiv \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \quad (10)}$$

is the **functional derivative** of J with respect to $y(x)$.

To aid our understanding, it can be helpful to think discretely. We can discretize x between x_1 and x_2 into N discrete steps, so that the function y takes up values $y_i = y(x_i)$, where $i = 1, 2 \cdots , N$. A choice of the function y corresponds in this discrete picture to a choice of $\{y_i\}$, which is a single point in an N-dimensional space. Fig. 1 has a visualization of this. At every point in this N-dimensional space, we can assign a value to J. The small variation $\varepsilon\eta$ can likewise be discretized, so that $\delta y_i = \varepsilon\eta_i$, which can be thought of as a step in a particular direction in the same N-dimensional space. In this discrete picture,

Figure 1: A discretized visualization of varying over functions. (**HL**: To be completed, but not difficult to imagine!)

$$\delta J = \sum_{i=1}^{N} \frac{\partial J}{\partial y_i} \delta y_i\, , \quad (11)$$

just as one might expect for a function J defined in the N-dimensional space indexed by y_i. In the continuous limit, we need to trade the summation over discrete i to an integral over the continuous label x, leading to Eq. (9).[5]

[5] In this picture, a choice of the function y is a point in an uncountably infinite dimensional space, and $\epsilon\eta$ is a step in some arbitrary direction, and J is function that returns a real number at every point in this space.

1.1.2 The Euler-Lagrange equation

Now, to find the **stationary points**—maxima, minima or saddle points—of J, we want to set $\delta J = 0$ for any arbitrary variation $\varepsilon\eta$, just like for a function g

on $\mathbb{R}^n$, we want δg to be zero for a step in any direction at a stationary point. Referring to Eq. (8), we require

$$\int_{x_1}^{x_2} dx \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] \varepsilon\eta = 0 \,. \tag{12}$$

Since this applies for *any* $\eta(x)$, the term in the square brackets $[\cdots]$ must vanish.[6] we must have

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0} \tag{13}$$

[6] We can prove that this is true rigorously, given various conditions on η and f. This is often known as the **fundamental lemma of the calculus of variations**. For further discussion, see S&G 1.2.2 and Wikipedia.

for a stationary point for J. This is the famous **Euler-Lagrange equation**.

Through derivations similar to what we saw above, we can get generalized Euler-Lagrange equations for more complicated versions of J. If J depends on more than one function y_i, for example, the stationary points are given by

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) = 0 \,, \tag{14}$$

which is one equation for each variable y_i. If on the other hand, f depends on higher derivatives y'', y''' and so on, then the generalized Euler-Lagrange equation we get is

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial y''}\right) - \frac{d^3}{dx^3}\left(\frac{\partial f}{\partial y'''}\right) + \cdots = 0 \,. \tag{15}$$

1.1.3 Applications

Time to apply what we've learnt! We'll apply the Euler-Lagrange equations to two examples.

1.1.3.1 Soap film supported by a pair of coaxial rings Consider Fig. 2, where a pair of co-axial rings support a soap film. The energy associated with the configuration is directly proportional to the area, and hence the soap film tends to minimize this energy by minimizing its area, subject to the constraint that the soap film has to end on the rings at either end. The area of associated with a segment of the film of width dx is

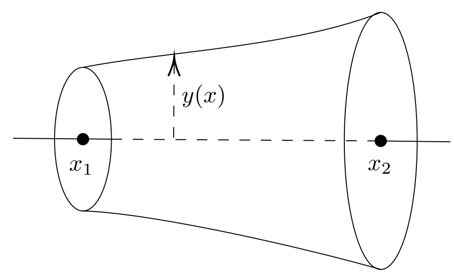

Figure 2: Soap film between two rings, centered at x_1 and x_2, with radii $y(x_1)$ and $y(x_2)$.

$$dA = 2\pi y(x)\sqrt{dx^2 + dy^2} = 2\pi y(x)\sqrt{1 + y'^2}\, dx \,, \tag{16}$$

and so the functional that we want to minimize is

$$J[y] = \int_{x_1}^{x_2} dx\, f(y, y') \,, \quad f(y, y') \equiv y\sqrt{1 + y'^2} \,, \tag{17}$$

with the endpoint values fixed at $y(x_1)$ and $y(x_2)$. The minimum for this functional can therefore be found by applying the Euler-Lagrange equations. The partial derivatives that we need for the Euler-Lagrange equation are

$$\frac{\partial f}{\partial y} = \sqrt{1 + y'^2} \,, \quad \frac{\partial f}{\partial y'} = \frac{yy'}{\sqrt{1 + y'^2}} \,, \tag{18}$$

and so the Euler-Lagrange equation says that the minimal surface area profile $y(x)$ must satisfy

$$
\begin{aligned}
\sqrt{1+y'^2} - \frac{d}{dx}\left(\frac{yy'}{\sqrt{1+y'^2}}\right) &= 0 \\
\implies \sqrt{1+y'^2} - \frac{y'^2}{\sqrt{1+y'^2}} - \frac{yy''}{\sqrt{1+y'^2}} + \frac{yy'^2y''}{(1+y'^2)^{3/2}} &= 0 \\
\implies \frac{1}{\sqrt{1+y'^2}} - \frac{yy''}{(1+y'^2)^{3/2}} &= 0\,. \qquad (19)
\end{aligned}
$$

This differential equation looks difficult to solve, but fortunately there's a neat little trick to do so. Multiplying by y' on both sides gives

$$0 = \frac{y'}{\sqrt{1+y'^2}} - \frac{yy'y''}{(1+y'^2)^{3/2}} = \frac{d}{dx}\left(\frac{y}{\sqrt{1+y'^2}}\right)\,. \qquad (20)$$

We'll return to how we knew this trick would work later on. In the mean time, the solution is

$$\frac{y}{\sqrt{1+y'^2}} = \kappa \qquad (21)$$

for some constant κ. Rewriting this as

$$\frac{dy}{dx} = \sqrt{\frac{y^2}{\kappa^2} - 1} \implies \int \frac{dy}{\sqrt{y^2/\kappa^2 - 1}} = \int dx\,, \qquad (22)$$

we can integrate this first-order ordinary differential equation by substituting $y = \kappa \cosh t$ and $dy = \kappa \sinh t$ to find

$$\kappa \int dt = \int dx \implies \kappa t = x + C \implies y = \kappa \cosh\left(\frac{x+C}{\kappa}\right) \qquad (23)$$

for some constants κ and C. These can be determined by enforcing the two boundary conditions—the radii of the two rings, $y(x_1)$ and $y(x_2)$.

1.1.3.2 The brachistochrone The next problem we will consider is a famous one, posed by Johann Bernoulli in 1696. What shape should a wire with endpoints $(0,0)$ and (a,b) take, in order that a frictionless bead will slide from rest down the wire in the shortest possible time?

First, the total time T taken down a given path can be written as

$$T = \int_0^T dt = \int_0^L \frac{ds}{v}\,, \qquad (24)$$

where v is the speed of the bead, and s is the distance along the path, with a total length L. However, we can once again write $ds^2 = dx^2 + dy^2$ so that $ds = \sqrt{1+y'^2}\,dx$, and apply conservation of energy to find $v = \sqrt{2gy}$. Thus, we can define a functional T

$$T[y] = \int_0^a dx \sqrt{\frac{1+y'^2}{2gy}} \qquad (25)$$

that we want to minimize with respect to y, again with fixed end points. We can therefore apply the Euler-Lagrange equation, which gives after some algebra

$$yy'' + \frac{1}{2}(1+y'^2) = 0\,. \qquad (26)$$

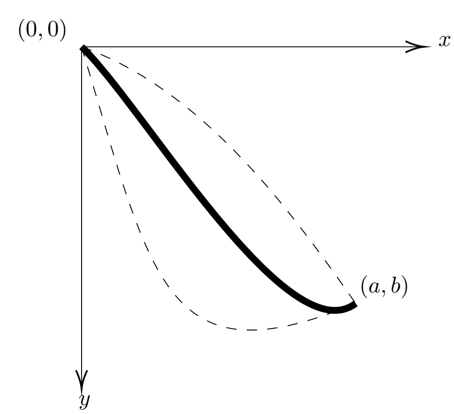

Figure 3: Possible shapes of a wire for a frictionless bead to travel from the origin $(0,0)$ to a point (a,b).

Once again, we can use the trick of multiplying by y' to find that

$$y'\left(yy'' + \frac{1}{2}(1+y'^2)\right) = \frac{1}{2}\frac{d}{dx}\left(y(1+y'^2)\right) = 0\,, \tag{27}$$

or

$$y(1+y'^2) = 2C \tag{28}$$

for some constant C. From this point, one can check that the following parametrization $(x(t), y(t))$ is indeed a solution to the differential equation above:

$$\begin{aligned} x &= C(\theta - \sin\theta) \\ y &= C(1-\cos\theta)\,, \end{aligned} \tag{29}$$

although it is surprisingly hard to pin down the details regarding this solution.[7] This parametric curve is known as the **cycloid**, which is the curve traced out by a fixed point on the rim of a wheel that is rolling without slipping along a flat surface.

[7] For example, is there a unique solution, and does the solution always occur with $\theta \in [0, 2\pi)$ for every point (a, b)? See for example Ref. [1] for more details.

1.1.4 First integral

In both applications discussed in Sec. 1.1.3, we were able to rephrase the differential equation to be solved as $dI/dx = 0$, implying that I is some constant associated with the problem. This quirk, which somewhat resembles the conservation of quantities like energy, is something we will revisit in much greater depth later: it is far deeper than just a mathematical coincidence. For now, let's just take a quick look at where it comes from. In both cases, the function inside the integral was $f(y, y')$, with no explicit dependence on x, implying that

$$\frac{df}{dx} = \cancelto{0}{\frac{\partial f}{\partial x}} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'}\,. \tag{30}$$

We define the **first integral** of the Euler-Lagrange equation as

$$\boxed{I \equiv f - y'\frac{\partial f}{\partial y'}\,,} \tag{31}$$

from which we can check that

$$\begin{aligned} \frac{dI}{dx} &= y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'} - y''\frac{\partial f}{\partial y'} - y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \\ &= y'\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\,. \end{aligned} \tag{32}$$

Thus, $dI/dx = 0$ if the Euler-Lagrange equation is satisfied.

You can show that if f depends on more than one variable, so that we have a functional of the form

$$J[y_1, y_2, \cdots, y_n] = \int dx\, f(y_1, y_2, \cdots, y_n; y_1', y_2', \cdots, y_n')\,, \tag{33}$$

the first integral is of the form

$$I = f - \sum_i y_i'\frac{\partial f}{\partial y_i'}\,. \tag{34}$$

Note that there is only one first integral, even when there are multiple dependent variables y_i.

1.2 Lagrangian Mechanics

It turns out that classical mechanics can be reformulated as a problem of finding the stationary function of some functional. Given some system, we first define the **Lagrangian** function $L = T - V$, where T and V are the kinetic and potential energy functions of the system. We can make any choice of coordinates we would like to describe T and V; let's say we choose some set of **generalized coordinates** q with components q^i and time derivatives $\dot{q}^i$.[8]

The equations of motion governing the system between times t_i and t_f can then be obtained by finding the stationary function $q(t)$ of the **action functional**,

$$S[q] = \int_{t_i}^{t_f} dt\, L(t, q^i; \dot{q}^i)\,. \tag{35}$$

[8] This can be the usual x, y and z in 3D space, r, θ and ϕ in 3D spherical coordinates, or something even more abstract than these choices, it doesn't matter.

This is known as the **principle of least action**. It is no exaggeration to say that a lot of theoretical physics basically involves finding the appropriate action that describes the system of interest, once the principle of least action is applied.

(**End of Lecture**: Wednesday Sep 4 2024)

References

[1] Philippe G. Ciarlet and Cristinel Mardare. "On the Brachistochrone Problem". In: *Communications in Mathematical Analysis and Applications* 1.1 (2022), pp. 213–240. ISSN: 2790-1939. DOI: `https://doi.org/10.4208/cmaa.2021-0005`. URL: `http://global-sci.org/intro/article_detail/cmaa/20161.html`.