An Introduction to The Calculus of Variations

Caio C.S.

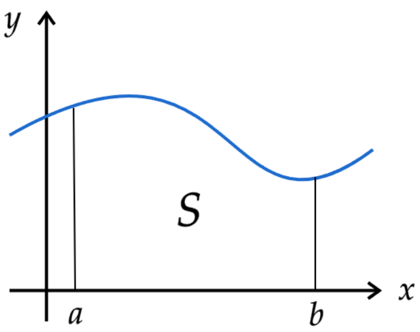
Contents

1 Introduction & First Steps

Usually in physics and mathematics, one wishes to extremize *stuff*. That happens for the first time when studying differential calculus of a single variable, and also when taking introductory physics courses, in which the trajectory of a cannonball is maximized, the time that an object takes to climb down an inclined plane is minimized, and energy is a quantity that also is minimized. In physics, especially in classical mechanics, things tend to be always minimized, energy, the path between two points, and so forth. That is intrinsically related to the so-called 'Principle of Least Action', first stated by William Rowan Hamilton, it states that the extremization of a quantity called 'action', establishes the dynamics of a physical system[1]. The Principle of Least Action is the core of physical theories of modern physics, such as General Relativity, with the Einstein-Hilbert action, Quantum Mechanics and Quantum Field Theory with the path integral formulation, which invokes the action principle. However, it is formulated in the language of functionals, mathematical objects studied by the realm of mathematics called Functional Analysis.

Definition 1.1 (Linear Functional)**.** Let $\exists$S; S is a linear functional, then, S is a linear mapping from the set of all smooth scalar functions $\mathfrak{F} \in \mathrm{C}^{\infty}$, to the real line, i.e, $S : \mathfrak{F} \to \mathrm{R}$.

Somehow, although not formally, functionals have been used without the knowledge that they were indeed functionals. For instance,



The Riemann integral is a functional on the smooth functions defined on $[a, b]$ into the real line. So as long as the definition 1.1 is satisfied, functionals can assume the structure of many mathematical objects.

Example 1:

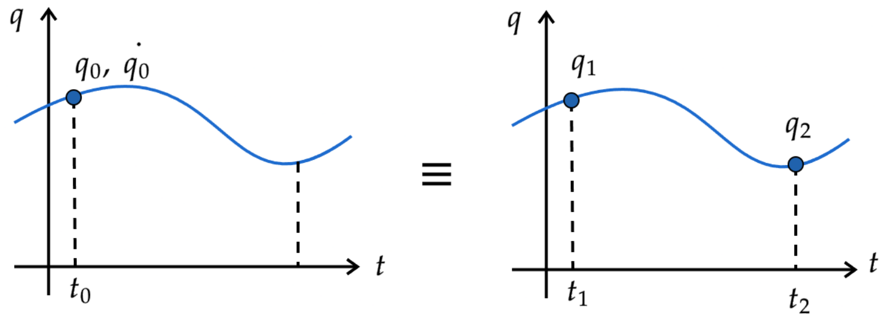
$$S[y] = \int_0^1 dx e^x = (e - 1) \in \mathrm{R}$$

Example 2:

$$S[y] = \left.\frac{\mathrm{d}\sin x}{\mathrm{d}x}\right|_{x=\pi} = \cos x|_{x=\pi} = -1 \in \mathrm{R}$$

[1]Notice how this principle is very general and does not restrict itself to classical mechanical systems.

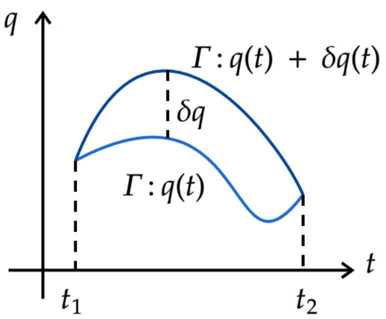
So one can say that two examples of functionals are integral-like functionals and derivative-like functionals[2]. Furthermore, in mathematics and physics, problems are usually formulated in terms of initial value problems, even though, sometimes they can be formulated as boundary condition problems, the language of calculus of variations makes this transition from initial value $\rightarrow$ boundary condition[3] as natural as possible, since fundamentally one is simply solving second order differential equations.



A natural question to make that transition then is, "What is a variation?".

Consider a curve Γ in R^d defined by the parameter t, from $t = t_1$ to $t = t_2$. Then, for a functional S defined on this curve, the equation that extremizes Γ is given by the extremals of the S functional.

Definition 1.2 (A Variation)**.** Given a curve Γ defined by a parameter $t; t \in [t_1, t_2]$, the δ operator acting on q is said to be a variation of the curve Γ if and only if δq doesn't have any relation to any t along the curve.



Notice that $\delta q(t_1) = 0 = \delta q(t_2)$ by the seeking of extremum. The definition 1.2 implies that $\delta q \neq \mathrm{d}q$, since one can associate dq with a dt. Also, the variation δ has some properties that are exactly like the total differentiation variation, with the exception of the one mentioned previously, but one unique property is that $\frac{\mathrm{d}}{\mathrm{d}q}\delta = \delta\frac{\mathrm{d}}{\mathrm{d}q}$.

[2]Although they are not the only type of functionals, a common tool in Quantum Field Theory in Curved Spacetime is the Functional Renormalization Group.

[3]Boundary conditions are of much higher interest in physics.

2 The Euler-Lagrange Equations

Furthermore, now the ground basis has been established to answer the main question on extremization of a functional defined on a curve in a given interval. Thus,

$$\delta S = \delta S[q(t) + \delta q(t)] - S[q(t)] = 0 \tag{1}$$

Then, taking S as an integral functional of the arbitrary function L, that depends on the coordinates $q(t)$, the first derivative of the coordinate $\dot{q}(t)$, and the parameter t along the curve.

$$\delta S = \delta \int_{t_1}^{t_2} dt L(q_i, \dot{q}_i, t) = \int_{t_1}^{t_2} dt \delta L(q_i, \dot{q}_i, t) = 0 \tag{2}$$

$$\iff \int_{t_1}^{t_2} dt \left\{ \frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \frac{\partial L}{\partial t}\delta t \right\} = \int_{t_1}^{t_2} dt \left\{ \frac{\partial L}{\partial q_i}\delta q + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i \right\} = 0$$

Separating the integral terms,

$$\because \int_{t_1}^{t_2} dt \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i = \left[\frac{\partial L}{\partial \dot{q}_i}\delta q \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\frac{\partial L}{\partial \dot{q}_i} \right)\delta q_i = - \int_{t_1}^{t_2} dt \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\frac{\partial L}{\partial \dot{q}_i} \right)\delta q_i \tag{3}$$

$$\implies \int_{t_1}^{t_2} dt \left\{ \frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\frac{\partial L}{\partial \dot{q}_i} \right) \right\}\delta q_i = 0$$

Which can only be zero, if the integrating term is zero,

$$\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0; \;\; i = 1, 2, ..., d \tag{4}$$

These are the famously so-called Euler-Lagrange equations. Given a function L, the Euler-Lagrange equations, which are second-order differential equations, have as a solution, the function that extremizes the Γ curve.

In physics, the function L is called the *Lagrangian*. The L function has some properties that go as the following,

i) $L = L(q, \dot{q}, t) \implies L \in C^2$.

ii) $L = L^* \implies L$ is a real scalar function

iii) In physics, $L = T - V$

iv) L is not unique, $L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{\mathrm{d}\lambda(q,t)}{\mathrm{d}t}$; gives the same Euler-Lagrange equations.

The last property is actually a theorem and it can be easily proven.

Theorem 2.1 (Non-uniqueness of L)**.** Let $\exists L \in C^2$, such that L is a real scalar function that obeys the Euler-Lagrange equations, then, $\exists L'$; $L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{\mathrm{d}\lambda(q,t)}{\mathrm{d}t}$, in a way that the Euler-Lagrange equations are preserved.

Proof. Consider the S' functional of the primed L function.

$$S' = \int dtL' = \int dtL + \int dt\frac{\mathrm{d}\lambda}{\mathrm{d}t} \tag{5}$$

$$\iff S' = S + \lambda\Big|_{t_1}^{t_2} \tag{6}$$

$$\because \delta S' = \delta S + \delta\lambda\Big|_{t_1}^{t_2} = \delta S' = \delta S + \Big[\frac{\partial\lambda}{\partial q}\delta q\Big]_{t_1}^{t_2} \tag{7}$$

$$\therefore \delta S' = \delta S$$

Preserving the Euler-Lagrange equations. □

Now, from a physics point of view, that's not the most general Lagrangian one could get, for instance, considering a Lagrangian like $L = L(q, \dot{q}, \ddot{q}, t)$, how the seeking extremum of the action functional would change the Euler-Lagrange equations? Would they still be second-order differential equations?

Varying the action functional of this general Lagrangian,

$$\delta S = \delta\int_{t_1}^{t_2} \mathrm{d}tL(q,\dot{q},\ddot{q}) = \int_{t_1}^{t_2} \mathrm{d}t\delta L(q,\dot{q},\ddot{q}) = 0 \tag{8}$$

$$\iff \int_{t_1}^{t_2} \mathrm{d}t\Big\{\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} + \frac{\partial L}{\partial \ddot{q}}\delta\ddot{q}\Big\} = 0$$

The integral of the first two terms is already known,

$$\int_{t_1}^{t_2} \mathrm{d}t\Big\{\Big(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\Big)\delta q + \frac{\partial L}{\partial \ddot{q}}\delta\ddot{q}\Big\} = 0 \tag{9}$$

Integrating the last term by parts,

$$\int_{t_1}^{t_2} \mathrm{d}t\frac{\partial L}{\partial \ddot{q}}\delta\ddot{q} = \Big[\frac{\partial L}{\partial \ddot{q}}\delta\dot{q}\Big]_{t_1}^{t_2} - \int_{t_1}^{t_2} \mathrm{d}t\delta\dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \ddot{q}}\Big) = \Big[\frac{\partial L}{\partial \ddot{q}}\delta\dot{q}\Big]_{t_1}^{t_2} + \Big[\frac{\partial L}{\partial \ddot{q}}\delta q\Big]_{t_1}^{t_2} + \int_{t_1}^{t_2} \mathrm{d}t\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big(\frac{\partial L}{\partial \ddot{q}}\Big)\delta q$$

$$\int_{t_1}^{t_2} \mathrm{d}t\frac{\partial L}{\partial \ddot{q}}\delta\ddot{q} = \int_{t_1}^{t_2} \mathrm{d}t\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big(\frac{\partial L}{\partial \ddot{q}}\Big)\delta q \tag{10}$$

$$\because \int_{t_1}^{t_2} \mathrm{d}t\Big\{\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\Big) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big(\frac{\partial L}{\partial \ddot{q}}\Big)\Big\}\delta q = 0 \tag{11}$$

$$\therefore \frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\Big) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big(\frac{\partial L}{\partial \ddot{q}}\Big) = 0 \tag{12}$$

The equation above represents the fourth-order Euler-Lagrange equations for a Lagrangian that depends not only on the coordinates and its first derivative but also on the second derivative of the coordinates. This type of Lagrangian is of high interest to physics since the Einstein-Hilbert action functional is such that it depends on the Lagrangian density for the gravitational field, $\mathcal{L}_{\mathrm{EH}} = R\sqrt{g}$, such that g is the determinant of the metric tensor and R is the Ricci scalar, a quantity that measures the presence of matter in curved spacetime, but $R = R(g, \partial g, \partial^2 g)$, making the Einstein-Hilbert Lagrangian density such that obeys a differential[4] equation like (12).

If for instance, such Lagrangians can be written as $L(q,\dot{q},\ddot{q}) = \tilde{L}(q,\dot{q}) + \frac{\mathrm{d}\hat{L}(q,\dot{q})}{\mathrm{d}t}$

$$\implies S = \int_{t_1}^{t_2} \mathrm{d}t\Big(\tilde{L}(q,\dot{q}) + \frac{\mathrm{d}\hat{L}(q,\dot{q})}{\mathrm{d}t}\Big) \tag{13}$$

$$\implies \delta S = 0 = \int_{t_1}^{t_2} \mathrm{d}t\delta\Big(\tilde{L}(q,\dot{q}) + \frac{\mathrm{d}\hat{L}(q,\dot{q})}{\mathrm{d}t}\Big) = \int_{t_1}^{t_2} \mathrm{d}t\delta\tilde{L}(q,\dot{q}) + \delta\Big[\hat{L}(q,\dot{q})\Big]_{t_1}^{t_2}$$

$$= \int_{t_1}^{t_2} \mathrm{d}t\delta\tilde{L}(q,\dot{q}) + \Big[\frac{\partial\hat{L}}{\partial q}\delta q\Big]_{t_1}^{t_2} + \Big[\frac{\partial\hat{L}}{\partial\dot{q}}\delta\dot{q}\Big]_{t_1}^{t_2} = \int_{t_1}^{t_2} \mathrm{d}t\delta\tilde{L}(q,\dot{q})$$

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t\delta\tilde{L}(q,\dot{q}) \tag{14}$$

Thus, in this special case, the Euler-Lagrange equations become second-order, as they return to the ordinary case already discussed.

3 Noether's Theorem

Even though we are concerned with mathematics, that is not our main concern, what this is all about is *mathematical physics*. Hence all of the discussion made so far has some really interesting applications to physics with the necessary mathematical formalism. Thus, solving the Euler-Lagrange equations for most of the systems[5] can only be done numerically or perturbatively. So what are the conditions for a system to be solved analytically?

Definition 3.1 (Integral of Motion)**.** An integral motion or constant of motion, $I = I(q,\dot{q},t)$ such that $\frac{\mathrm{d}I}{\mathrm{d}t} = 0\ \forall q(t)$, solving equations of motion.

With that definition, the most natural question is, how one knows that a system has such equations of motion? It is due to Noether!

Noether's theorem discusses connections between symmetries and integrals of motion, in a way that given a parameter ε and δt as a generator, the introduction of the ε parameter makes it possible to classify different types of symmetries.

$$\text{Symmetries}\begin{cases}\text{Discrete}(\varepsilon\text{ discrete}) \nrightarrow \text{Conserved quantities}\\ \text{Continuous}(\varepsilon\text{ arbitrariarly small})\end{cases}$$

$$\text{Symmetries}\begin{cases}\text{Global}(\varepsilon\text{ constant}) \implies \text{Conserved Quantities}\\ \text{Local}(\varepsilon = \varepsilon(t)) \implies \text{Gauge Symmetries}\end{cases}$$

Thus, considering transformations like,

$$t \to t' = t + \tilde{\delta}t \tag{15}$$

[4] However, Einstein's equations are not fourth-order differential equations, they are second-order differential equations, this happens because of the conditions imposed onto the geometrical structure of spacetime

[5] They can only be solved exactly for analytically integrable systems

$$q(t) \to q'(t) = q(t) + \tilde{\delta}q(t) \tag{16}$$

Such that

$$\tilde{\delta}t = \varepsilon\Delta t \tag{17}$$

$$\tilde{\delta}q(t) = \varepsilon\Delta q \tag{18}$$

So, equations (17) and (18), define the previously mentioned, generators of symmetries. Hence, the first Noether's theorem is,

Theorem 3.1 (First Noether's Theorem)**.** For every global continuous symmetry of the system, there is a corresponding integral of motion, provided that the equations are satisfied.

That theorem can be proven for a few physical cases in which one has continuous symmetries. For instance,

Corollary 3.1. If $L = L(t) \implies L = L(q, \dot{q})$, then the energy ca be written as $E = \frac{\partial L}{\partial \dot{q}}\dot{q} - L$ is a conserved quantity.

Proof. Since, $E = \frac{\partial L}{\partial \dot{q}}\dot{q} - L$, then

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\dot{q} - L\Big) \\
&= \dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\Big) + \frac{\partial L}{\partial \dot{q}}\ddot{q} - \frac{\mathrm{d}L}{\mathrm{d}t} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\dot{q}\Big) + \frac{\partial L}{\partial \ddot{q}}\ddot{q} - \frac{\partial L}{\partial q}\dot{q} - \frac{\partial L}{\partial \dot{q}}\ddot{q} \\
&= \dot{q}\Big[\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\Big) - \frac{\partial L}{\partial q}\Big] = 0
\end{aligned}$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = 0 \tag{19}$$

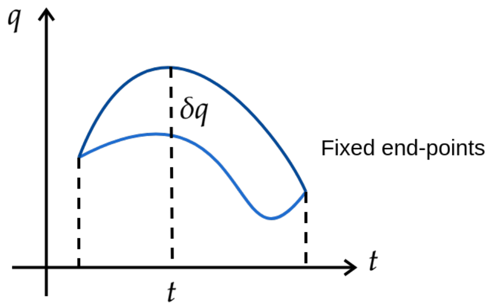
□

Hence, when the Lagrangian of a physical system does not depend on time, the energy is a conserved quantity of that system, and by the first Noether's theorem, there is a global continuous symmetry associated with it, which is **time translation**, because, if $t = t + \delta t$, then the equation (19) still holds.

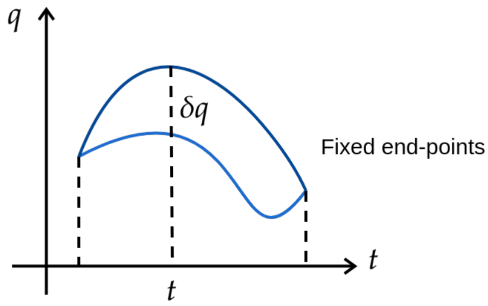
Another immediate example is for cases where $L = L(q_c)$, such that q_c is a cyclic coordinate. Thus, generalised momentum, which is defined as $p = \frac{\partial L}{\partial q_c}$, is an integral of motion, i.e, preserved in time, resulting a global continuous symmetry on spatial translations.

Theorem 3.2 (Second Noether's Theorem)**.** Let $\exists\tilde{\delta}$ be a global continuous symmetry, with any conditions whatsoever imposed, then, $\tilde{\delta}t, \tilde{\delta}q$ such that, $\tilde{\delta}S = 0 \implies I = \frac{\partial L}{\partial \dot{q}}\tilde{\delta}q + (L - \dot{q}\frac{\partial L}{\partial \dot{q}}\Big)\tilde{\delta}t$, is valid provided that I is an integral of motion.

Proof. First of all, $\tilde{\delta}q \neq \delta q$. Then, the best picture to think on the theorem is,



However, when talking about Noether's theorem, there is no consideration of fixed end-points. So if one takes a space translation for instance, then the *fixed end-points* will move! So they are not truly fixed.



So in Noether's theorem, the boundary therms do not go to zero, and they actually have a roll in the conservation game.

$$\tilde{\delta} \neq \delta \tag{20}$$

Since,

$$\tilde{\delta} t = t' - t \implies \tilde{\delta} dt = dt' - dt = d\tilde{\delta} t = \frac{\mathrm{d}\tilde{\delta} t}{\mathrm{d}t} dt \tag{21}$$

Hence, a variation of the coordinates is given as,

$$\tilde{\delta} q(t) = q'(t') - q(t) = q'(t) + \tilde{\delta} t \frac{\mathrm{d}q(t)}{\mathrm{d}t} + ... - q(t) = \delta q(t) + \tilde{\delta} t \frac{\mathrm{d}q(t)}{\mathrm{d}t} \tag{22}$$

$$\iff \tilde{\delta} = \delta + \tilde{\delta} t \frac{\mathrm{d}}{\mathrm{d}t} \, \forall f(q, \dot{q}, t) \tag{23}$$

and thus, the theorem is proven. $\square$