CALCULUS OF VARIATION: BIRTH AND RISE OF A NEW DISCIPLINE

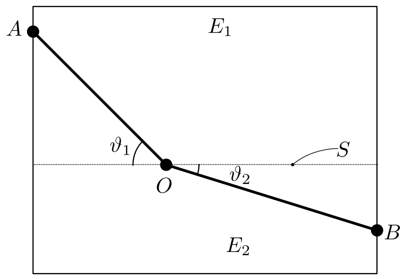
Escola de verão de matematica

MARCO CAROCCIA*

June 20 - June 22, 2018

CONTENTS

* Faculdade de Ciencias da Universidade de Lisboa - email: mcaroccia@fc.ul.pt



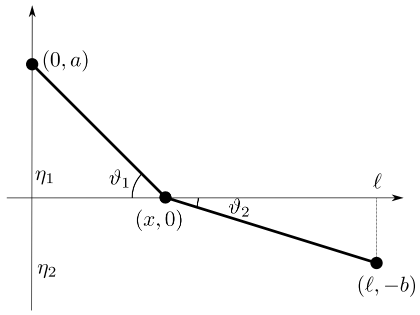


Figure 1: The light ray enter E_2 from the point O and is deflected from ϑ_1 to ϑ_2. On the right the situation modeled in the Cartesian plane.

1 FERMAT'S LEAST TIME PRINCIPLE

"Nature operates by means and ways that are 'easiest and fastest"
Pierre de Fermat - The analysis of refractions (1662).

1.1 Do light knows calculus?

Imagine to have two bodies E_1 and E_2 made by different materials that share a flat interface S. Imagine now that a light ray is shot through E_1 it passes through S and then exit from E_2 (typical example is *air-water*). What is the angle of deviation to which the ray is subject (see Figure 1)? The experience tells us that if the materials are different some sort of refraction is expected. In particular, if v_1 is the speed of light in the medium E_1 and v_2 is the speed of light in the medium E_2 then the Snell-Descartes Law (1637) reads as

$$\frac{\cos(\vartheta_1)}{\cos(\vartheta_2)} = \frac{v_1}{v_2}.$$

In his two papers "The analysis of refractions" and "The synthesis of refractions" Pierre De Fermat try to re-read the above law as a consequence of a very general principle that, according to Goldstine [1], defines the clear first milestone of what will become famous as *Calculus of Variation*. He states "Nature operates by means and ways that are 'easiest and fastest" and that "nature always acts along shortest paths". Thus the light ray will choose the point O so that the trajectory between A and B is achieved in the least amount of time. To see how powerful this principle is, let us go through the same computation (in modern language) that Fermat went through.

We can imagine to put a reference Frame as in Figure 1 and set the *indexes of refraction* $\eta_1, \eta_2 > 1$. The index of refraction represents how much the medium is slowing down the speed of light compared to the void. In particular, if c is the speed of light in the void we have that the speed of light across E_i is given by $v_i = \frac{c}{\eta_i}$. Notice that, the time that is required to the light to go from A to O is

$$\frac{|O-A|}{v_1} = \frac{\eta_1}{c}\sqrt{a^2+x^2}.$$

In the same way the amount of time needed to reach B starting from O in the second medium is

$$\frac{|O-B|}{v_2} = \frac{\eta_2}{c}\sqrt{b^2+(x-\ell)^2}.$$

Then the total amount of time needed to reach B starting from A and passing through the point O is thus

$$T(O) = T(x) = \frac{\eta_1}{c}\sqrt{a^2+x^2} + \frac{\eta_2}{c}\sqrt{b^2+(x-\ell)^2}.$$

So, according to Fermat's principle of least time, the point O chosen by the light is the point that minimizes the above time. Let's check it. We just need to search for x_0 such that

$$T'(x_0) = 0.$$

Notice

$$T'(x) = c^{-1}\left(\frac{\eta_1 x}{\sqrt{x^2+a^2}} - \frac{\eta_2(\ell - x)}{\sqrt{b^2+(x-\ell)^2}}\right).$$

It is quite hard to re-read the Snell-Descartes relation in the above derivative. But, by looking in the above picture we can see that

$$x = \sqrt{a^2+x^2}\cos(\vartheta_1(x)), \quad (\ell - x) = \sqrt{b^2+(\ell-x)^2}\cos(\vartheta_2(x)).$$

Thus we see that

$$T'(x) = c^{-1}\left(\eta_1\cos(\vartheta_1(x)) - \eta_2\cos(\vartheta_2(x))\right).$$

Then, a critical point must satisfies

$$c^{-1}\eta_1\cos(\vartheta_1(x_0)) = c^{-2}\eta_2\cos(\vartheta_2(x_0)),$$
$$\frac{\cos(\vartheta_1)}{v_1} = \frac{\cos(\vartheta_2)}{v_2}$$

yielding precisely the Snell-Descartes law!

1.2 Do dogs know calculus?

In 2003 Professor Timothy Pennings was playing with his dog down to the shore of lake Michigan. The game was the classical game you can play with your dogs, namely he was throwing the ball and waiting for Elvis to fetch it and bringing it back. Timothy was standing on the sand and he was throwing the ball in the lake. During the whole game he was thinking "there is something odd the way Elvis run to fetch the ball". Indeed his dog seemed to behaved similarly to the refraction law explained by Snell-Descartes in some sense and thus he decided to come back with all the requirement needed to do measurements. After all, is quicker to run on the sand than to swim, hence the problem of finding the right place in where entering the lake is not a trivial task (especially for a dog!). What he discovered, surprisingly, was that Elvis was always making the good choice (up to approximation that he explain very good in his paper titled exactly *do dogs know calculus?*).

> *...we confess that although he made good choices, Elvis does not know calculus. In fact, he has trouble differentiating even simple polynomials. More seriously, although he does not do the calculations, Elvis's behavior is an example of the uncanny way in which nature (or Nature) often finds optimal solutions. Consider how soap bubbles minimize surface area, for example. It is fascinating that this optimizing ability seems to extend even to animal behavior. (It could be a consequence of natural selection, which gives a slight but consequential advantage to those animals that exhibit better judgment.) Finally, for those intrigued by this general study, there are further experiments that are available, other than using your own favorite dog. One might do a similar experiment with a dog running in deep snow versus a cleared sidewalk. Even more interesting, one might test to determine whether the optimal path is found by six-year-old children, junior high aged pupils, or college students. For the sake of their pride, it might be best not to include professors in the study.*

2 THE BRACHISTOCHRONE PROBLEM

"I, Johann Bernoulli, address the most brilliant mathematicians in the world.
Nothing is more attractive to intelligent people than an honest,
challenging problem, whose possible solution will
bestow fame and remain as a lasting monument.
Following the example set by Pascal, Fermat, etc.,
I hope to gain the gratitude of the whole scientific community
by placing before the finest mathematicians
of our time a problem which will
test their methods and the strength of their intellect.
If someone communicates to me the solution of the proposed problem,
I shall publicly declare him worthy of praise"
Johan Bernoulli - Acta Eruditorum (1696).

Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time?

Bernoulli challenged the mathematical community with this problem in May 1696. Galileo has already studied such problem in 1632, lately proposed as a mathematical challenge by Bernoulli. Galileo's clever approach to the problem was the following: "Let's simplify the problem and let's ask who is the straight line from A to a point B on a vertical line which it would reach the quickest" (see Figure 2). So the target is variable but we restrict ourselves to straight lines! Even if he adopted this nice point of view his answer was wrong. But he proved that is always more convenient to roll on a circular arc than on a straight line.

In May 1697 four solutions to the brachistochrone problem were published: Leibniz's solution, Johann Bernoulli's solution, Jacob Bernoulli's and a Latin translation of Newton's solution who has submitted an anonymous manuscript. Even if Newton's solution was submitted anonimously, Bernoulli identify the hand of Sir Isaac Newton behind the proof and lately declared "tanquam ex ungue leonem" (we recognize the lion by his claw).

2.1 Bernoulli's solution

Johann Bernoulli applied the Fermat's principle of least time and argued in the following way: imagine that we are dealing with a light ray starting from A and reaching B through several medium $S_1, \ldots, S_k$ each one that has (increasing) index of refraction η_i (see Figures 3, 4). Then, according to Fermat's principle of least time, the trajectory pursued by the light has to be the fastest one. He noticed that the Snell-Descartes law tells that, if ν_i is the speed in the medium S_i then

$$\frac{\cos(\vartheta_i)}{\nu_i} = \frac{\cos(\vartheta_{i+1})}{\nu_{i+1}}.$$

In particular

$$\frac{\cos(\vartheta_i)}{\nu_i} = C \quad \text{for all } i.$$

Thus, since (see Figure 4) $\Delta y_i = \cos(vartheta_i)\Delta z_i$ we achieve

$$\Delta y_i = C\Delta z_i \nu_i \quad \Rightarrow \quad \Delta y_i^2 = C^2\nu_i^2(\Delta y_i^2 + \Delta x_i^2) \quad \Rightarrow \quad \Delta y_i^2 = \frac{C^2\nu_i^2}{(1 - C^2\nu_i^2)}\Delta x_i^2.$$

He knew that $\nu_i = \kappa\sqrt{x_i}$ for a constant κ and hence he deduced

$$\Delta y_i^2 = \frac{x_i}{a - x_i}\Delta x_i^2 \quad \Rightarrow \quad \frac{\Delta y_i}{\Delta x_i} = \sqrt{\frac{x_i}{a - x_i}}.$$

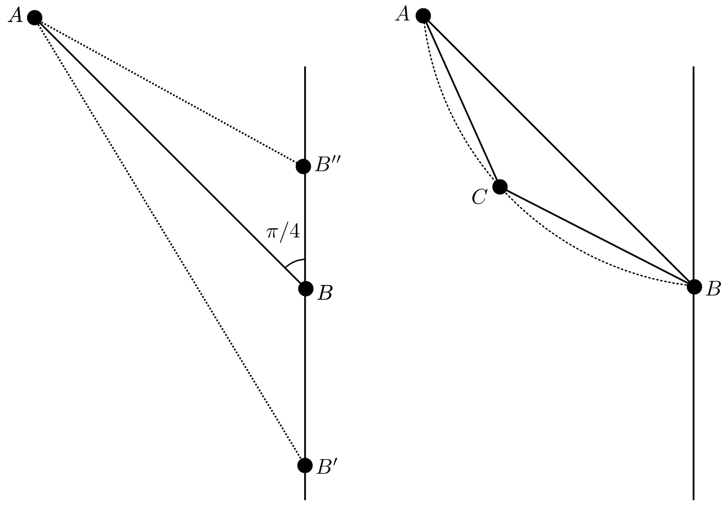


Figure 2: The brachistochrone problem approached by Galileo: he first proved that among all possible straight line the one with angle $\pi/4$ achieved the fastest. Then he realized that, if a point passes through a spot C placed on a circular arc lying on the chord over AB the time can be decreased.

By taking the limit as $\Delta x_i \to 0$ the differential relation

$$y'(x) = \sqrt{\frac{x}{a-x}}$$

is achieved. The above can be integrated with the following trick

$$\begin{aligned} y(x) - y(0) &= \int_0^x \sqrt{\frac{x}{a-x}}\,dx = \frac{1}{2}\int_0^x \frac{a}{\sqrt{ax-x^2}}\,dx - \frac{1}{2}\int_0^x \frac{a-2x}{\sqrt{ax-x^2}}\,dx \\ &= a \arcsin\left(\frac{x}{a}\right) - \sqrt{ax-x^2}. \end{aligned}$$

To recognize the above figure we need to do the following change of variable $x(\theta) = \frac{a}{2}(1-\cos(\vartheta))$ yielding $y(\vartheta) = \frac{a}{2}(\vartheta - \sin(\vartheta))$. In particular we have that the curve

$$\frac{a}{2}(1-\cos(\vartheta), \vartheta - \sin(\vartheta))$$

is a *cycloid*, namely the curve drawn by a point linked to a rolling circumference of radius $\frac{a}{2}$ (see Figure 5).

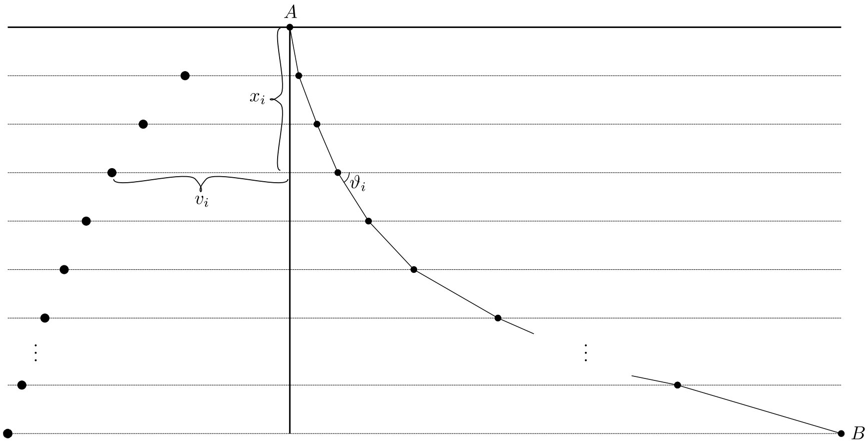


Figure 3: The brachistochrone problem approached by Bernoulli: he imagine that a set of velocities v_i have been assigned and that we need to find the path of light ray from A to B through different mediums. He was figuratively representing the velocities v_i in the sector S_i as the length of a segment at height x_i.

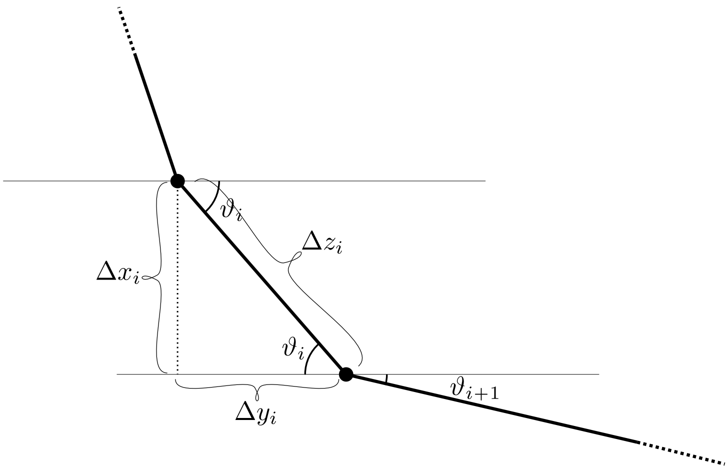


Figure 4: We clearly see that: $\Delta y_i = \Delta z_i \cos(\vartheta_i)$ and $\Delta z_i = \sqrt{\Delta y_i^2 + \Delta x_i^2}$.

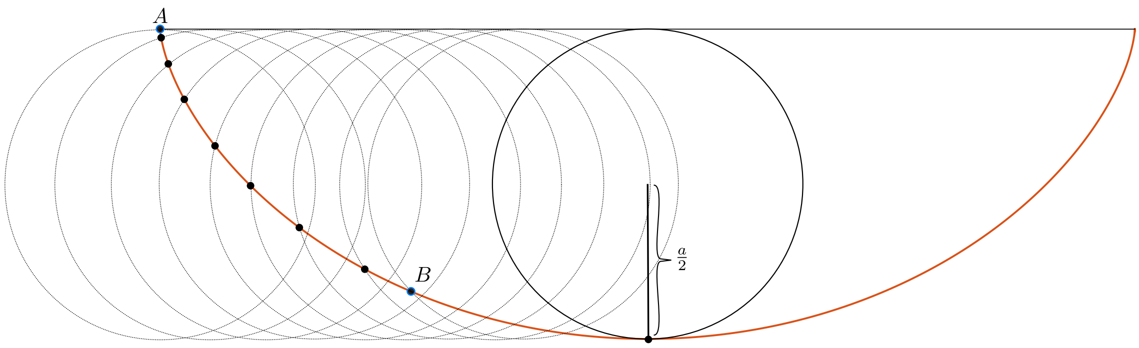


Figure 5: The brachistochrone given by a circle of radius $\frac{a}{2}$

3 MATHEMATICAL FRAMEWORK: LAGRANGE RULES!

"On ne trouvera point de Figures dans cet Ouvrage.
Les méthodes que j'y expose ne demandent ni constructions,
ni raisonnements géométriques ou mécaniques,
mais seulement des opérations algébriques,
assujetties à une marche régulière et uniforme."
Joseph Louis Lagrance - Mécanique analytique, (1788)

Given two instant of time t_1 and t_2 we can imagine that a system determined by the n coordinates $\mathbf{q}(t) := (q_1(t), \ldots, q_n(t))$ is endowed with a *Lagrangian*

$$\mathcal{L} : [t_1, t_2] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+.$$

We define the action $\mathcal{S} : X \to \mathbb{R}_+$ on $\mathbf{q} = \{\mathbf{q}(t)\}_{t \in [t_1, t_2]} \in X$ as

$$\mathcal{S}(\mathbf{q}) := \int_{t_1}^{t_2} \mathcal{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) dt.$$

for all curve $\mathbf{q} \in X$ where X denotes the *family of admissible state* namely the set of all possible states in which we allow the system to be. It is well posed then the problem

$$\ell := \inf\{\mathcal{S}(\mathbf{q}) \mid \mathbf{q} \in X\}.$$

We say that $\mathbf{q}_\nu$ provides a *minimum for* $\mathcal{S}$ if $\mathbf{q}_\nu \in X$ and

$$\mathcal{S}(\mathbf{q}_\nu) := \min\{\mathcal{S}(\mathbf{q}) \mid \mathbf{q} \in X\} \tag{1}$$

Remark 1. The existence of minimum may fail!

Theorem 1. *Any minimum $\mathbf{q}_\nu$ (with C^2 regularity) satisfies the Euler-Lagrange equation*

$$\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right) = 0 \tag{2}$$

Proof. Let $\mathbf{q}_\nu$ be a minimum. Consider a C^1 curve $\varphi : [t_1, t_2] \to \mathbb{R}_+$ with

$$\varphi(t_1) = \varphi(t_2) = 0 \tag{3}$$

By definition of minimum we observe that

$$f(\varepsilon) := \mathcal{S}(\mathbf{q}_\nu + \varepsilon\varphi) - \mathcal{S}(\mathbf{q}_\nu) \geqslant 0 \quad \text{(up to warning 1)}.$$

Warning 1. *We are exploiting the following property of the set of state X: for any $\mathbf{q} \in X$ and for any C^1 curve φ*

$$\mathbf{q} + \varepsilon\varphi \in X \ \textit{for all} \ \ \varepsilon \geqslant 0 \ \ \textit{small enough}. \tag{4}$$

We are also assuming that the constraint (3) on φ does not violates the above requirement. If X changes, the constraint on φ might change as well in order to fit with (4), yielding different kinds of equations in place of (2). To the aim of this short notes we will consider admissible states where the boundary condition have been fixed and thus a curve φ which has no effect at t_1 and at t_2 will always safeguard (4).

Since $f(0) = 0$ and $f(\varepsilon) \geqslant f(0)$ for all ε small enough we need to have that $f'(0) \geqslant 0$. Thus

$$f'(0) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{S}(\mathbf{q}_\nu + \varepsilon\varphi) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_{t_1}^{t_2} \mathcal{L}(t, \mathbf{q}_\nu + \varepsilon\varphi, \dot{\mathbf{q}}_\nu + \varepsilon\dot{\varphi})\, dt$$

$$\text{warning } 2 = \int_{t_1}^{t_2} \left[\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\varphi_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\dot{\varphi}_i\right] dt.$$

Warning 2. *We are assuming that we can exchange derivative and integral.*

Notice that we can write

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\dot{\varphi}_i = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\varphi_i\right) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\varphi_i$$

and thus

$$\begin{aligned} \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\dot{\varphi}_i \, dt &= \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\varphi_i\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\varphi_i \, dt \\ \text{due to (3)} &= -\int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\varphi_i \, dt. \end{aligned}$$

Hence

$$f'(0) = \sum_{i=1}^{n} \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\right]\varphi_i \, dt$$

Since $f'(0) \geqslant 0$ we conclude

$$0 \leqslant \sum_{i=1}^{n} \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\right]\varphi_i \, dt$$

By exchanging φ with $-\varphi$ we can reproduce all the computations and reach

$$0 \leqslant -\sum_{i=1}^{n} \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\right]\varphi_i \, dt$$

yielding

$$0 = \sum_{i=1}^{n} \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\right]\varphi_i \, dt$$

and the above equality holds for all $\varphi = (\varphi_1, \ldots, \varphi_n)$ satisfying (3). Thus, by testing the above with function $\varphi = (0, \ldots, 0, \varphi, 0, \ldots, 0)$ with $\varphi(t_1) = \varphi(t_2) = 0$ we achieve the independent relations

$$0 = \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right)\right]\varphi \, dt \quad \text{for all } i = 1, \ldots, n$$

where the above holds for all $\varphi \in C^1([t_1, t_2])$ with $\varphi(t_1) = \varphi(t_2) = 0$.

Warning 3 (Fundamenal Lemma of Calculus of variations)**.** *Notice that if*

$$\int_{t_1}^{t_2} f(t)\varphi(t) \, dt = 0 \quad \textit{for all } \varphi \in C^1([t_1, t_2])$$

then $f = 0$.

By means of warning 3 we can thus conclude that

$$\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}_\nu, \dot{\mathbf{q}}_\nu)\right) = 0 \quad \text{for all } i = 1, \ldots, n.$$

□

Warning 4. *Equation (2) helps to identify a minimum only if we can show that there is a unique stationary point to $\mathcal{S}$. Indeed Theorem 1 tells us that if $\mathbf{q}_v$ is a minimum then it satisfies (2). The converse is, in general, false unless we have a unique point for which (2) holds.*

The curves φ are called *compactly supported perturbation* or *variation* and what the warning 1 requires (in order for this argument to works) is that, if a state $\mathbf{q}$ is an accessible state (belongs to X) then all small perturbations around $\mathbf{q}$, namely $\mathbf{q} + \varepsilon\varphi$, are accessible states as well. In particular we have that any solutions to (2) represents a *stationary point for* $\mathcal{S}$ which might not be a global minimizers (but just a local one, or even a saddle point!).

Warning 5. *The existence of minimizers is a non trivial tasks and actually it represents one of the main ingredient for this all analysis to be consistent. A common mistake is the following chain of deduction*

1) *Assume that a minimum exists;*

2) *Try to deduce some property about such a solution by exploiting its minimality with respect to $\mathcal{S}$ (for example by means of equations (2));*

3) *Try to deduce, starting from the properties obtained at point 2), who is the minimum!*

Such receipt might lead to absurd conclusion as it is well explained by the following example.

Example 1. Consider the problem of finding

$$n_0 := \sup\{n \mid n \in \mathbb{N}\}.$$

In this example, the action is simplified to be just the value of the number itself with no integration or boundary values. The family of admissible state is $X = \mathbb{N}$. Consider the above chain then:

1) Assume that we have a maximum $n_0 \in \mathbb{N}$ such that

$$n_0 := \max\{n \mid n \in \mathbb{N}\}.$$

2) Let's deduce some property by exploiting its maximality (instead of minimality); We know that $x^2 \geqslant x$ for all $x \in \mathbb{R}$. Hence

$$n_0^2 \geqslant n_0.$$

From the other side n_0 is the maximum and since n_0^2 is an admissible state, by maximality it holds

$$n_0 \geqslant n_0^2.$$

Thus $n_0^2 = n_0$;

3) The only admissible state satisfying the property deduced in 2) is $n_0 = 1$. Hence 1 is the biggest natural number!

The problem with the above example is that we are not allowed to manipulate a minimizer without ensuring its existence. So all the argument is bugged by point 1) *of the receipt. In particular we should treat* max *and* min *problem in term of* sup *and* inf *and then try to show that there exists a minmizer (maximizer).*

A BRIEF TREATMENT OF EXISTENCE OF MINIMUM (MAXIMUM). Do we remember the *Weierstrass's theorem*? Let's recall it, together with its proof

Theorem 2 (Weierstrass's Theorem)**.** *Any continuous function* $f : [a, b] \to \mathbb{R}_+$ *admits a maximum and a minimum on* $[a, b]$.

Proof. Define

$$\ell := \inf_{x\in[a,b]}\{f(x)\}, \quad L := \inf_{x\in[a,b]}\{f(x)\}.$$

We can always find the so called *minimizing* and *maximizing* sequences $\{x_n^\ell\}_{n\in\mathbb{N}}$ and $\{x_n^L\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to+\infty} f\left(x_n^\ell\right) = \ell, \quad \lim_{n\to+\infty} f\left(x_n^L\right) = L.$$

Notice that $[a, b]$ is a compact interval and thus, by means of Bolzano - Weierstrass we can find two subsequences $\{x_{n_k}^\ell\}_{k\in\mathbb{N}} \subset \{x_n^\ell\}_{n\in\mathbb{N}}$ and $\{x_{n_k}^L\}_{k\in\mathbb{N}} \subset \{x_n^L\}_{n\in\mathbb{N}}$ and two limit points x^ℓ, x^L such that

$$\lim_{k\to+\infty} x_{n_k}^\ell = x^\ell, \quad \lim_{k\to+\infty} x_{n_k}^L = x^L.$$

In particular, by continuity of f and by the property of the maximizing and minimizing sequence, we have

$$\ell = \lim_{n\to+\infty} f\left(x_n^\ell\right) = \lim_{k\to+\infty} f\left(x_{n_k}^\ell\right) = f\left(x^\ell\right)$$

$$L = \lim_{n\to+\infty} f\left(x_n^L\right) = \lim_{k\to+\infty} f\left(x_{n_k}^L\right) = f\left(x^L\right).$$

Hence we can write

$$f\left(x^\ell\right) = \min_{x\in[a,b]}\{f(x)\}, \quad f\left(x^L\right) = \max_{x\in[a,b]}\{f(x)\}.$$

□

What we really used in the above proof was the following two facts

a) Any sequence $\{x_n\}_{n\in\mathbb{N}} \subset [a, b]$ it admits a converging subsequence (compactness of the family of admissible states);

b) The function (action) is continuos;.

In particular, in order to guarantee existence of minimum (maximum) we need to argue in an abstract way and ensure that

a) Any minimizing sequence of admissible state $\{\mathbf{q}_n\}_{n\in\mathbb{N}} \subset X$ it admits a converging subsequence (compactness of the family of admissible states with low action);

b) The action $\mathcal{S}(\mathbf{q})$ is continuos (we underline that it is enough to exploit the notion of *lower semicontinuity* of $\mathcal{S}$).

Indeed, define

$$\ell := \inf\{\mathcal{S}(\mathbf{q}) \mid \mathbf{q} \in X\}$$

and as above pick a minimizing sequence $\{\mathbf{q}_n\}_{n\in\mathbb{N}} \in X$, namely such that

$$\mathcal{S}(\mathbf{q}_n) \to \ell.$$

Point a) implies the existence of a converging subsequence $\{\mathbf{q}_{n_k}\}_{k\in\mathbb{N}} \subset \{\mathbf{q}_n\}_{n\in\mathbb{N}}$, $\mathbf{q}_{n_k} \to \mathbf{q}_v \in X$. Then the (lower semi)continuity of $\mathcal{S}$ yields

$$\ell \leqslant \mathcal{S}(\mathbf{q}_v) = (\leqslant) \lim_{k\to\infty} \mathcal{S}(\mathbf{q}_{n_k}) = \ell.$$

This yields the existence of at least one minimum

$$\mathcal{S}(\mathbf{q}_v) = \ell = \min\{\mathcal{S}(\mathbf{q}) \mid \mathbf{q} \in X\}$$

and allows us to run points 2) and 3) in the receipt in warning 5.

4 APPLICATIONS OF THE VARIATIONAL PRINCIPLE

"In other words, the laws of Newton could be stated not in the form $F = ma$ 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. "
Richard Feynman (1964)

4.1 The Brachistochrone problem revisited

Consider the brachistochrone problem. We set $A = 0$ and $B = (x_b, y_b)$ with $x_b > 0$, $y_b < 0$. All the admissible states can be represented by curves $\mathbf{q}_y(t) = (t, y(t))$, $\mathbf{q} : [a, b] \to \mathbb{R}^2$ with $\mathbf{q}_y(0) = 0$ and $\mathbf{q}_y(b) = \mathbf{b}$. Then

$$X := \{\mathbf{q}_y = (t, y(t)) : [0, x_b] \to \mathbb{R}^2 \mid y \text{ is } C^2 \text{ functions with } y(0) = 0,\ \ y(x_b) = y_b\}.$$

Given a curve $\mathbf{q}_y \in X$ as above we now need to understand how to compute its Lagrangian, relatively to the stated problem. In particular the action should yield back the time spent by the body to go from A to B by moving only on $\mathbf{q}_y$. The speed along the curve is $v(t) = \frac{ds}{dt}$ where ds represents the infinitesimal of length on $\mathbf{q}$. In particular

$$T(\mathbf{q}_y) := \int_{\mathbf{q}_y} \frac{ds(\mathbf{x})}{v(\mathbf{x})}.$$

We also know that the kinetic energy of the body needs to be equal to the potential energy at each time. In this scenario the potential energy at some point $\mathbf{x} \in \mathbb{R}^2$ depends only on the size of vertical component $|\mathbf{x} \cdot \mathbf{e}_2|$ (the "altitude" of the point) we have

$$\frac{1}{2} m v(\mathbf{x})^2 = mg|\mathbf{x} \cdot \mathbf{e}_2|.$$

This yields

$$v(\mathbf{x}) := \sqrt{2g|\mathbf{x} \cdot \mathbf{e}_2|}.$$

Moreover, by differential geometry it is well known that

$$\int_{\mathbf{q}} h(\mathbf{x})\, ds(\mathbf{x}) = \int_a^b h(t, y(t))\sqrt{1 + (y')^2}\, dt$$

and thus

$$\mathcal{S}(\mathbf{q}_y) = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1 + y'(t)^2}{-y(t)}}\, dt$$

Since in our convention $y(t) < 0$. Hence, the Lagrangian $\mathcal{L}(t, y, y') = \mathcal{L}(y, y') = \sqrt{\frac{1}{2g}}\sqrt{\frac{1+y'^2}{-y}}$. In order to run our receipt we need the following Theorem

Theorem 3. *Given the above Lagrangian $\mathcal{L}$ and having defined the above set of admissible states X then there exists a state $\mathbf{q}_v(t)$ that minimizes the action among all the possible states in X. Such state has C^∞ regularity.*

Then we can apply Theorem 1 and deduce that the required curve must satisfy

$$\frac{\partial \mathcal{L}}{\partial q_i}(t, \mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}, \dot{\mathbf{q}})\right) = 0 \quad i = 1, 2.$$

Notice that our Lagrangian is time-independent. We can in particular deduce an interesting thing fromt this fact, through the clever (Hamilton - Jacobi) change of variables

$$p_i := \frac{\partial \mathcal{L}}{\partial \dot{q}_i}(t, \mathbf{q}, \dot{\mathbf{q}}).$$

Thus

$$\dot{p_i} = \frac{\partial \mathcal{L}}{\partial q_i}.$$

Then, having defined

$$H = \sum_{i=1}^{2} \dot{q}_i p_i - \mathcal{L}$$

we have

$$\frac{d}{dt} H = \sum_{i=1}^{2} \ddot{q}_i p_i + q_i \dot{p}_i - \dot{q}_i \frac{\partial \mathcal{L}}{\partial q_i} - \ddot{q}_i p_i = 0.$$

In particular a solution to (2) (for time independent Lagrangian) it has the property that

$$H(\mathbf{q}, \mathbf{p}(\mathbf{q}, \dot{\mathbf{q}})) = \text{cost}$$

for all $t > 0$. Notice that

$$\frac{\partial \mathcal{L}}{\partial y'} = \frac{1}{\sqrt{2g}} \frac{y'}{\sqrt{y(1+y'^2)}}.$$

Thus, by inserting this in the relation

$$y' \frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} = C$$

we obtain the relation

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

with $\frac{1}{2gC^2} = a$. In particular the minimum is a solution to the Cauchy problem

$$\begin{cases} (y')^2 &= \frac{a+y}{-y} \\ y(0) &= 0; \\ y(x_b) &= y_b; \end{cases} \tag{5}$$

with $a \geqslant -y \geqslant 0$. Suppose that $x = x(\vartheta)$, $y(x(\vartheta)) = y(\vartheta)$. Then

$$\frac{dy}{dx}(x(\vartheta)) = \frac{y'(\vartheta)}{x'(\vartheta)}$$

and hence

$$(y'(\vartheta))^2 = (x'(\vartheta))^2 \left(\frac{a - y(\vartheta)}{y(\vartheta)} \right)$$

Consider

$$x(\vartheta) := \frac{a}{2}(\vartheta - \sin(\vartheta)), \quad y(\vartheta) := \frac{a}{2}(\cos(\vartheta) - 1)$$

Then

$$\begin{aligned} x'(\vartheta) &= \frac{a}{2}(1 - \cos(\vartheta)) \\ y'(\vartheta) &= -\frac{a}{2} \sin(\vartheta) \end{aligned}$$

and

$$\frac{y'(\vartheta)^2}{x'(\vartheta)^2} = \frac{\sin(\vartheta)^2}{(1-\cos(\vartheta))^2} = \frac{4\cos\left(\frac{\vartheta}{2}\right)^2 \sin\left(\frac{\vartheta}{2}\right)^2}{4\sin\left(\frac{\vartheta}{2}\right)^4} = \frac{\cos\left(\frac{\vartheta}{2}\right)^2}{\sin\left(\frac{\vartheta}{2}\right)^2}$$

On the other side

$$\frac{a+y}{-y} = \frac{a - \frac{a}{2}(1-\cos(\vartheta))}{\frac{a}{2}(1-\cos(\vartheta))} = \frac{\frac{a}{2}(1+\cos(\vartheta))}{\frac{a}{2}(1-\cos(\vartheta))} = \frac{\cos\left(\frac{\vartheta}{2}\right)^2}{\sin\left(\frac{\vartheta}{2}\right)^2}$$

Thus $\mathbf{q}(\vartheta) = (x(\vartheta), y(\vartheta))$ provide the curves. We need to choose a now. We need, in particular to find, if they exists (ϑ_b, a) such that

$$\begin{cases} x_b &= \frac{a}{2}(\vartheta_b - \sin(\vartheta_b)), \\ y_b &= \frac{a}{2}(1-\cos(\vartheta_b)). \end{cases} \tag{6}$$

4.2 Snell-Descartes law improved

Consider now a medium S with a varying index of refraction $\eta(y)$ (which varies only in the vertical coordinate let's suppose!). Imagine that we have a light ray starting from a and reaching b. At each point $(x, y(x))$ we have a speed of $c/\eta(y(x)) = \nu(x)$ and thus the total time on path γ can be found as

$$T(\gamma) := c^{-1} \int_{x_a}^{x_b} \eta(y(x))\sqrt{1+y'(x)}\, dx$$

In particular the Lagrangian is thus given by

$$\mathcal{L}(y, y') = c^{-1} \frac{\sqrt{1+y'^2}}{\eta(y)}.$$

Notice that the Lagrangian is time independent again and hence we can use the equivalent equations

$$y'\frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} = C.$$

Since

$$\frac{\partial \mathcal{L}}{\partial y'} = c^{-1} \frac{y'\eta(y)}{\sqrt{1+y'^2}}$$

we achieve

$$c^{-1}\frac{y'^2\eta(y))}{\sqrt{1+y'^2}} - c^{-1}\eta(y)\sqrt{1+y'^2} = C$$

$$-\frac{\eta(y)}{\sqrt{1+y'^2}} = C$$

Notice now that $y'(x) = tan(\vartheta(x))$ where $\vartheta(x)$ is the angle between the tangent to graph $y(x)$ ant the horizontal line. This yields

$$\cos(\vartheta(x))\eta(y(x)) = C$$

Remark 2 (Lagrangian as $T - V$). Imagine that you have a potential $V(x)$. Then Newton's second law tells you that a particle of mass m subject to the potential $V(x)$ will have a law of motion given by

$$m\ddot{q} = -V'(q)$$

Notice that this equation comes simply as the Euler-Lagrange equation in the case when $L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q)$. More generally whenever we have a system, in a non-relativistic approximation and evolving in euclidean (not weird relativistic) geometry by choosing the Lagrangian as $\mathcal{L} = T - V$ we will produce the same law of motion given by classical principle and Newton's law. As already observed in the case of the Brachistochrone we have that, when $\mathcal{L}$ is time independent, a constant of motion is given by $\mathcal{H} = T + V$ (energy conservation law!!).

REFERENCES

[1] Herman Heine Goldstine. *A History of the Calculus of Variations from the 17th through the 19th Century*, volume 5. Springer Science & Business Media, 2012.