STARTING POINT FOR THE CALCULUS OF VARIATIONS

Jan Coufal

Abstract

The text is about the start of the calculus of variations. Optimal control deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved. Optimal control theory, an extension of the calculus of variations, is a mathematical optimization method for deriving control policies. The calculus of variations is concerned with the maxima or minima of functionals, which are collectively called extrema. The brachistochrone problem, a priori a simple game for mathematicians, turns out in the end to be a considerable problem. Indeed, the different approaches tried out in its solution may be considered, in a more or less direct way, as the starting point for new theories. While the true "mathematical" demonstration involves what we now call the calculus of variations, a theory for which Euler and then Lagrange established the foundations, the solution which Johann Bernoulli originally produced, obtained with the help analogy with the law of refraction on optics, was empirical. A similar analogy between optics and mechanics reappears when Hamilton applied the principle of least action in mechanics which Maupertuis justified in the first instance, on the basis of the laws of optics.

Key words: calculus of variations, brachistochrone problem, principle of least action, Johann Bernoulli, Jacob Bernoulli, P. L. Moreau de Maupertuis

JEL Code: A12, B23

Introduction

This paper continues and extends the articles (Coufal, 2013) and (Chabert, 1997). Our intention here is to write the history of the brachistrone and its remarkable consequences. In the contemporary socio-cultural context, the question would essentially be formulated in the following text: what shape should we make slides in children's playgrounds so that the time of descent should be minimized? The considerable importance of this question is well understood when we consider how children behave, and they want to obtain the best performance, but the question is also important in a more general way, and a great number of scholars have attempted to solve this problem.

Unfortunately the problem appears to be particularly tricky, and it depends upon a number of parameters, including the variable value of the friction between the clothes of the child and the surface of the slide. We shall not attempt to solve that particular problem here, but content ourselves with theory of the idealized problem, simplifying the situation sufficiently in order to be able to find a solution. In fact we shall replace the child by a perfectly smooth marble, and we assume that it rolls down a smooth surface, thus assuming that friction forces are negligible with respect to gravity.

Now, we are simply confronted with the *problem of brachistrone* as Johann Bernoulli expressed it in the *Acta Eruditorum* published in Leipzig in June 1696 ((Bernoulli, 1742), vol. 1, p. 161): *Datis in plano vertacali duobus punctis A & B, assignare Mobili M viam AMB, per quam gravitate sua descenden, & moveri incipiens a puncto A, brevissimo tempore perveniat ad alterum punctum B*. The expression *brevissimo tempore* is the latin translation of the greek term *brachistochrone* (brachys is brief, brachisto is quickest, chronos is time and brachistochrone is the shortes time). In a modern style: Given two points A and B in a vertical plane, what is the curve traced out by a point subject only to the force gravity, starting from rest at A, such that it arrives at B in the shortest time?

Common sense suggests that this curve is necessarily situated in the vertical plane containing the points A and B. Common sense also leads us to think that the quickest route is the shortest, and is given by the line segment joining the points A and B. But this is not the case. We know, for example that a longer journey on a motorway be faster than going a shorter distance on an ordinary road. Here, in order to try to solve the problem of brachistochrone, it is necessary to consider all the curves joining points A and B and compare all the corresponding times of travel. Taking everything into account, even under these restrictions, the problem turns out to be a subtle one.

1 Falling bodies, reflection and refraction

In 1638, well before the problem had been explicitly stated, Galileo gave his solution to the brachistochrone problem in the course of the Third Day of his (Galileo, 1638). It is here that he studied uniform acceleration – Galileo called it “natural acceleration” – comparing it with uniform motion, and showed that a body falling in space traverses a distance proportional to the square of the time of descent (Theorem II in (Galileo, 1638). With regard to bodies moving on inclined planes he deduced (Galileo, 1638)):

Theorem V. The times of descent along planes of different length, slope and height bear to one another a ratio which is equal to the product of the ratio of the lengths by square root of inverse ratio of their heights.

We interpret the proportionality to be: a body travels a distance *L* and descends a height *H* in time *t* such that:

$$t = \frac{k \cdot L}{\sqrt{H}}.$$

Galileo then proves the following neat result (Galileo, 1638):

Theorem VI. If from the highest or lowest point in a vertical circle there be drawn any inclined planes meeting the circumference, the times of descent along these chords are each equal to the other.

At the end of the Third Day, Galileo shows that it is also possible to improve on this descent (Galileo, 1638):

Theorem XXII. If from the lowest point of a vertical circle, a chord is drawn subtending an arc not greater than a quadrant, and if from the two ends of this chord two other chords be drawn to any point on the arc, the time of descent along the two later chords will be shorter than along the first, and shorter also, by the same amount, than along the lower of these two latter chords.

This result is false, since arguing the case from two to three segments is based on a faulty intuition from arguing from one to two segments. The brachistrochrone problem is considerably more subtle than the one of the research into optimum inclination of planes, which is a simple problem of the extremum for a function of single variable.

The demonstration by Johann Bernoulli (Bernoulli, 1742) also derives from an intuitive approach. This approach, an analogy with the law of refraction, leads to the curve solution which one cannot find a priori, without an arsenal of sufficiently sophisticated techniques. Let us begin by recalling the first laws of Optics, which are in fact consequences of the principles of optimization.

Experience tells us that light travels in straight lines. This phenomenon is stated as a principle: light chooses the shortest path. This formulation led to a real theoretical advance since it allowed Hero of Alexandria in the first century AD to explain the law of reflection, namely, the equality of the angles of incidence and reflection. In the case of reflection, the speed remains constant. It is not so for refraction, where the speed of light varies as a function of the index n of the medium traversed. However, the principle stated above could have been stated in the following form as the Fermat's Principle: light chooses the fastest route, which in a homogenous medium where its speed is constant, is equivalent to the previous principle.

So, to go from *A* to *B*, passing from a medium of index n_1 to medium of index n_2, the trajectory of the light will not be the line segment *AB*, but broken line *AIB* such that the trajectory *AIB* will have the shortest time of all trajectories from *A* to *B*. Using the initial conditions we calculate that the angle of incidence *i* and the angle of refraction *r* are related to the respective speeds v_i and v_r by the formula:

$$\frac{\sin i}{v_i} = \frac{\sin r}{v_r}, \tag{1}$$

or using the indices n_i and n_r we have the sine formula

$$n_i \cdot \sin i = n_r \cdot \sin r.$$

This formula, discovered by the Dutch scientist Snell in 1621, received its correct interpretation with Fermat. In a letter of the 1st of January 1662 to M De la Chambre, Fermat explains ((Fermat, 1894), vol. II, p. 457-463): *As I said in my previous letter, M. Descartes has never demonstrated his principle; because not only do the comparisons hardly serve as a foundation for the demonstrations, but he uses them in the opposite sense and supposes that the passage of light is more easy in dense bodies than in rare bodies, which is clearly false. I will not say anything to you about the shortcomings of the demonstration itself …*

Fermat puts his principle to work, and proves the sine formula using his method 'de maximis et minimis' (Fermat (1894)). Another example of a non homogeneous medium where the shortest trajectory is not the quickest occurs in mechanics, where the effect of gravity is in the vertical direction. And this is the context for Johann Bernoulli brachistochrone problem. Johann Bernoulli in the *Acta Eruditorum* of May 1697 ((Bernoulli, 1742), vol. 1, pp. 187-193). His method typically corresponds to what we now call a discretisation of the problem. He images space carved into small lamina, sufficiently fine so that within each one it is possible to imagine that the speed is constant. Within each strip the trajectory becomes the shortest route, and necessarily a segment. The complete trajectory appears as a sequence of segments. But how we move from one strip to another? We must always optimize the time of travel. As in refraction of light, this is done by using Fermat's principle. Thus, if v_i is the speed in a given band and v_r in the band immediately below, the angle *i* is the angle made with the vertical by segment of the trajectory in the first band, and the angle *r* in the neighboring band, then they are connected by the rule of sines (1). If we now imagine that the horizontal strips become progressively thinner, and their number increases indefinitely, the line of segments tends towards a curve. The tangents at each point of this curve approach the sequence of segments. The angle *u* which the tangent makes with the vertical is then connected to the speed *v* by the relation:

$$\frac{\sin u}{v} = const.$$

Here, the speed *v* of a particle is know; it is result of the action of gravity and, as we know from Galileo, it is a function of the distance fallen *y*, according to the formula

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

And so the rule of sines leads to the equation:

$$\frac{\sin u}{\sqrt{y}} = const.$$

In particular, for $y = 0$, the tangent is vertical.

That is a characteristic equation of a well-known curve of the time, the cycloid.

We have just seen that the solution to the curve is a cycloid. But how can we construct such a curve, starting from a point *A*, an arriving exactly at a point *B*? Newton gave a simple solution in a letter to Montague on the 30th of January 1697 (see (Newton, 1967), p. 223). In addition the Newton's contribution to the solution of the problem of the brachistochrone, we must also mention Leibniz, and in a lesser role, the Marquis de l'Hospital, and most of all, Jacob Bernoulli, the older brother of Johann ((Bernoulli, 1742), vol. 1, p. 194-204): … *my elder brother made up the fourth of these, that the three great nations, Germany, England, France, have given us each one of their own to unite with myself in such a beautiful search, all finding the same truth*.

The method used by Jacob Bernoulli is laborious, but quite general. Also, Jacob, in wanting to show the singular character of Johann's method, extended the problem by posing new questions. Indeed, Johann's method, founded on an analogy, does not work except in a particular case, and cannot be used for more general problems of this type. In particular, Jacob Bernoulli put the following question to his brother" given a vertical line which of all the cycloids having the same starting point and the same horizontal base, is the one which will allow a heavy body passing along it to arrive at the vertical line the soonest? Such statement reminds us of Galileo's first version, which was about finding the inclined plane through a given point which gave the shortest time to reach a given vertical. Johann Bernoulli ((Bernoulli, 1742), vol. 1, p. 206-213) replied and showed that the cycloid in question is the one which meets the given line horizontally. More generally, the cycloid which allows us to achieve the swiftest possible descent to a given oblique line is the one which meets the line at right angles. This cycloid which, as we have just said, is a brachistochrone curve, was also known to Huygens from 1659 as the tautochrone curve: bodies which fall in an inverted cycloid arrive at the bottom at the same time, no matter from what height they are released. This property was perhaps closer to that observed by Galileo: the equality of the times for the distance on the chords of the same circle. Among the other problems posed by Jacob Bernoulli to Johann are those which are called isoperimetric problems, which together with brachistochrone problem are prototypes of optimization problems. These scientific exchanges between the two brothers were carried out in the form of letters. Here is a sample of Johann's response to same criticisms by Jacob ((Bernoulli, 1742), vol. 1, p. 194-204): *So there it is, his imagination, stronger and more vivid than those claiming to be sorcerers who believe they have found themselves bodily present at a Sabbath, has seduced him; he is carried along by a torrent of vain conjectures; in a word, he is longer ready to give reign to reason …* The resolution of these problems is then the object – reason or excuse? – for a long dispute between the two brothers; a dispute which developed into a major row, but which gave birth to new area in mathematics, the Calculus of Variations.

2 Start of the Calculus of Variations

When we look for boundary values of a function *f* of a variable *x*, i.e. when we look for values of the variable *x* for which the value *f* (*x*) is a maximum or minimum, we look for the points where the graph of *f* has a horizontal tangent, or we say we look for the values where $f'(x)=0$. In the case of a function *f* of two variables *x* and *y*, we have to consider the points where the tangent plane is horizontal to the surface which has the equation $z=f(x,y)$. Alternatively we could say we seek the number pairs [*x*, *y*] for which

$$\frac{\partial f}{\partial x}(x,y)=\frac{\partial f}{\partial y}(x,y)=0.$$

Or we can say we are looking for the points where the function *f* has a stationary value. In the case of a finite number of variables, the difficulties seem surmountable, and the approach to the problem may be effected with the aid of the differential calculus of Newton and Leibniz. Here the object which changes is not a number or a point, but a curve, a function, and the corresponding quantity to maximize or minimize is a number depending on this curve or on this function. It is necessary to conceive an extension of the differential calculus. The new theory which was created is called the calculus of variations, the variations being those of the function. But, in 1696, this theory had not been formulated and our problem becomes *a priori* somewhat subtle. A problem in the calculus of variations can be presented generally in the following fashion: we try to find a curve, being the graphical representation of a function *y* of *x*, which minimizes or maximizes a certain quantity among all the curves constrained by certain conditions[1]. The quantity whose extreme value has to be found[2] is expressed generally in the form of an integral:

$$I(y) = \int_a^b F(x, y, y')\, dx$$

where *y* represented the unknown function, y' its derivative, *x* variable and *F* a particular function.

Among the typical problems of the calculus of variations, besides the isoperimetric problems above are investigations of the geodesic lines on surface, i.e. the curves of minimum length joining two points of a surface. Also, the investigation of the shapes of the surfaces of revolution which offer the least resistance to movement, a problem which Newton tackled in 1687 in the *Principia*. The statement of the brachistochrone problem in 1696 could be considered as the definitive origin of the calculus of variations, for it is the problem which generated general methods of investigation which were gradually developed in a competitive context. Johann Bernoulli himself posed the problem of geodetics to Euler. Euler re-worked the ideas of Jacob Bernoulli, simplified them, and finally was the first to formulate the general methods which allowed them to be applied to the principal problems of the calculus variation. He developed these ideas systematically in 1744 in (Euler, 1774). In a way like Jacob Bernoulli, Euler tackles the problem as a problem of limits in an investigation of the ordinary extremum. Euler derived the differential equation:

$$\frac{\partial F}{\partial y}(x, y, y') - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}(x, y, y')\right) = 0 \qquad (2)$$

which satisfies each solution *y*. It is only a necessary condition and the method does not establish the existence of a solution. The equation (2), today called the Euler-Lagrange equation, is a second order differential equation in *y*:

$$\frac{\partial F}{\partial y}(x, y, y') - \frac{\partial^2 F}{\partial y' \partial x}(x, y, y') - y' \frac{\partial^2 F}{\partial y' \partial y}(x, y, y') - y'' \frac{\partial^2 F}{\partial y'^2}(x, y, y') = 0$$

---

[1] For brachistochrone problem – the curve joining two points *A* and *B*.

[2] Here – the time of the journey.

In 1760, Lagrange greatly simplified matters by introducing the differential symbol δ, specifically for the calculus of variations, corresponding to a variation of the complete function. He makes the point of it in the introduction to (Lagrange, 1760-1761): *For as little as we know the principles of the differential calculus, we know the method for determining the largest and smallest ordinates of curves; but there are questions of maxima and minima at a higher level which, although depending on the same method, are not able to be applied so easily. They are those where it is needed to find the curves themselves, in which a given integral expression becomes a maximum or minimum with respect to all the other curves. … Now here is a method which only requires a straightforward use of the principles of the differential and integral calculus; but above all I must give warning that while this method requires that the same quantities vary in two different ways, in order not to mix up these variations, I have introduced into my calculations a new symbol δ. In this way, δZ expressed a difference of Z which is not the same as dZ, but which, however, will be formed by the same rules; such that where we have for any equation dZ=m dx, we can equally have δZ=m δx, and likewise for other cases.*

A century later, Mach was able to write in (Mach, 1883): *In this way, by analogy, Johann Bernoulli accidentally found a solution to the problem. Jacob Bernoulli developed a geometric method for the solution of analogous problems In one stroke, Euler generalized the problem and the geometrical method, Lagrange finally freed it completely from the consideration of diagrams, and provided an analytical method.*

3 The Principle of Least Action from Optics to Mechanics

We shall make a digression, the purpose of which will soon become clear Maupertuis stated his Principle of Least Action in 1744 in (Maupertius, 1744). He explains and justifies his principle from the law of refraction: *In thinking deeply upon this matter, I reflected that light, as it passes from one medium to another, yet not taking the shortest path, which is a straight line, might just as well not take the shortest time. Actually, why should there be a preference here for time over space? Light cannot go at the same time by the shortest path and by the quickest route, so why does it go by one route rather than another? In fact, it does not take either of these; it takes a route that has the greater real advantage*: ***the path taken is the one where the quantity of action is the least****.*

Now I must explain what I mean by the ***quantity of action****. When a body is moved from one place to another, a certain action is needed: this action depends neither on the speed of the* *body and the distance travelled; but it depends on the speed nor the distance taken separately. The quantity of action is moreover greater when the speed of the body is greater and when the path travelled is greater; it is proportional to the sum of the distance multiplied respectively by the speed travelled over each space. … It is quantity of action which is the true expenditure of Nature, and which she uses as sparingly as possible in the motion of light. Let there be two different media, separated by a surface represented by the line CD, such that the speed of light in the medium above is m. and the speed in the medium below is n.*

Let a ray of light, starting from point A, reach a point B: to find the point R where the ray changes course, we look for the point where if the ray bends ***the quantity of action is the least****: and I have* $m \cdot AR = n \cdot RB$ *which must be a* ***minimum****. …*

That is to say, ***the sine of the angle of incidence to the sine of the angle of refraction is in inverse proportion to the speed with the light traverses each medium****.*

All the phenomena of refraction now agree with the central principle that ***Nature, in the production of its effects, always tends towards the most simple means****. So this principle follows, that* ***when light passes from one medium to another the sine of the angle of refraction to the sine of the angle of incidence is in inverse ratio to the speed with which the light traverses each medium****.*

And so for Maupertuis, light is propagated so as to minimize $AR \cdot v_1 = RB \cdot v_2$ and not the quantity $\frac{AR}{v_1} = \frac{RB}{v_2}$. For these conclusions to agree with the experimental results of the time, and so that his principle would lead to the sine law. It is true that at that time no one knew how to measure the speed of light and no one could find a way of deciding between the different theories. The experimental proof that light travels faster in air than in water was not established until 1850 Foucault. In 1746, Maupertuis extended his principle from optics to mechanics (Maupertius, 1746): *When a body is carried from one place to another, the action is greater when the mass is heavier, when the speed is faster, when the distance over which it is carried is longer. … Whenever a change in Nature takes place, the quantity of action necessary for this change is the smallest possible*.

With this general principle, Maupertuis established a kind of union between philosophy, physics and mathematics: Nature works in such a way as to minimize its action; the idea of causality is abandoned in favor of the idea achieving an aim, characterized by a harmony between the physical world and rational thought.

Conclusion

It would be right to conclude by revisiting our initial problem of the slides in the playground. We are circumspect, and content ourselves with noticing that in the course of this wander through diverse disciplines, the theme of minimization or maximization briefly the problem of optimalization is ever present, and should not be underestimated during these unhappy times.

References

Bernoulli, J. (1742). *Opera Omnia*. Lausanne and Geneva.

Chabert, J.-L. (1997). The Brachistrone Problem. *History of Mathematics, Histories of Problems*, Elipse, Paris, pp. 183-202.

Coufal, J. (2013). The Origination of the Calculus of Variations. *Mundus symbolicus* [CD], vol. 21, n. 21, pp. 5-12.

Euler, L. (1774). *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes: sive solution problematic isoperimetrici lattissimo sensu accepti*, Lausanne and Geneva (*Œvres*, vol. 24, Berne, Orel Füssli, 1952).

Fermat, P.(1894). *Œvres*, ed. Tannery, P. and Henry, C., Gauthier-Villars, Paris.

Galileo, G. (1638). *Discorsi e dimostrazioni matematiche intorno a duo nuove scienze,* Leyden.

Lagrange, J.-L. (1760-1761). Essai d'une nouvelle méthode pour déterminer les maxima et les minima des formules integrals defines. *Miscellanea Taurinensia*, vol. II (*Œvres*, vol. I, Paris, pp. 333-468).

Mach, E. (1883). *Die Mechanik in ihrer Entwickelung : historisch-kritisch dargestellt*, F. A. Brockhaus, Leipzig,.

Maupertius, P. L. M. de (1744). Accord de différentes lois de la nature qui avaient jusqu'ici paru incompatibles. *Mémoires de l'Acadenie des Sciences de Paris*, pp. 417-426.

Maupertius, P. L. M. de (1746). Les lois du movement et du repos déduites d'un principe métaphysique. *Mémoires de l'Acadenie des Sciences de Berlin*, pp. 267-294.

Newton, I. (1967). *Correspondence*, ed. by Scott, J. F., vol. IV (1694-1709), Cambridge University Press, Cambridge.

Contact

Jan Coufal

University of Economics

Faculty of Statistics and Informatics, Department of Mathematics

Ekonomická 957

148 00 Prague 4

Czech Republic

e-mail: *jan.n.coufal@seznam.cz*