The Australian Journal of Mathematical Analysis and Applications

http://ajmaa.org

**Volume 9**, Issue 2, Article 8, pp. 1-6, 2012

# A ONE-LINE DERIVATION OF THE EULER AND OSTROGRADSKI EQUATIONS

OLIVIER DE LA GRANDVILLE

*Received 3 April, 2012; accepted 27 June, 2012; published 28 December, 2012.*

STANFORD UNIVERSITY, DEPARTMENT OF MANAGEMENT SCIENCE AND ENGINEERING,
STANFORD, CA 94305
ola@stanford.edu

ABSTRACT. At the very heart of major results of classical physics, the Euler and Ostrogradski equations have apparently no intuitive interpretation. In this paper we show that this is not so. Relying on Euler's initial geometric approach, we show that they can be obtained through a direct reasoning that does not imply any calculation. The intuitive approach we suggest offers two benefits: it gives immediate significance to these fundamental second-order non-linear differential equations; and second, it allows to obtain a property of the calculus of variations that does not seem to have been uncovered until now: the Euler and Ostrogradski equations can be derived not necessarily by giving a variation to the optimal function – as is always done; one could equally well start by giving a variation to their derivative(s).

*Key words and phrases:* Calculus of variations; Euler; Ostrogradski equations.

2000 *Mathematics Subject Classification.* 49.

ISSN (electronic): 1449-5910

## 1. INTRODUCTION.

One of the most fascinating areas of mathematics, the calculus of variations, was born when Pierre de Fermat, Isaac Newton and Johann Bernoulli raised and solved problems of a physical nature. Fermat showed that the kinked path of light, moving accross media of different densities, minimised travel time. Newton investigated the optimal shape of a ship's prow that would incur minimal resistance; and Bernoulli, in June 1696, challenged the scientific community with the famous brachystochrone problem, whose solution had eluded even Galileo: find the path between two points A and B in a vertical plane such that a bead slide from A to B in minimum time. The problem was solved by Johann Bernoulli himself as well as by his brother Jakob, L'Hospital, Leibniz, Newton and Tschirnhaus.

All solutions to these problems relied on careful geometrical and physical considerations. But solving the general problem of optimizing functionals of simple types such as $\int_a^b F(x, y, y')\, dx$ or $\int\int_R F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dxdy$ eluded mathematicians for quite some time, and it is only about three decades later that Leonhard Euler discovered the second-order differential equations that governed the solution leading to an extremum of functionals such as these. (To illustrate the trickiness of the calculus of variations, it may be useful to remember that correct sufficient conditions for extrema of functionals still escaped the sagacity of the best mathematicans at the turn of the eighteenth century, until they appeared in the work of Jacobi and Weierstrass – on this, see Herman Goldstine, *A History of the Calculus of Variation*s, 1980).

A first order condition for an extremum of $\int_a^b F(x, y, y')\, dx$ is

$$\frac{\partial F}{\partial y}(x, y, y') - \frac{d}{dx}\frac{\partial F}{\partial y'}(x, y, y') = 0, \tag{1.1}$$

the celebrated equation published by Euler in 1744, but which he most certainly had discovered in the 1730's already (see Goldstine, op.cit.). Written in full, it reads

$$\frac{\partial F}{\partial y}(x, y, y') - \frac{\partial^2 F}{\partial y'\partial x}(x, y, y') - \frac{\partial^2 F}{\partial y'\partial y}(x, y, y')y' - \frac{\partial^2 F}{\partial y'^2}(x, y, y')y'' = 0. \tag{1.2}$$

As to the first order condition for $\int\int_R F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dxdy$ to go through an extremum, it is given by

$$\frac{\partial F}{\partial z}(x, y, z, p, q) - \frac{\partial}{\partial x}\frac{\partial F}{\partial p}(x, y, z, p, q) - \frac{\partial}{\partial y}\frac{\partial F}{\partial q}(x, y, z, p, q) = 0, \tag{1.3}$$

where $p \equiv \partial z/\partial x$ and $q \equiv \partial z/\partial y$. This is the Ostrogradski, or Euler-Ostrogradski equation where the last two terms on the left-hand side are the total derivatives of $\partial F/\partial p$ and $\partial F/\partial q$ with respect to $x$ and $y$ respectively. Hence it is a second order partial differential equation whose fully written expression is the following; note that for brevity we have omitted the dependence of the partial derivatives of $F$ on their arguments $(x, y, z, p, q)$ :

$$\begin{aligned} &\frac{\partial F}{\partial z} - \frac{\partial^2 F}{\partial p\partial x} - \frac{\partial^2 F}{\partial p\partial z}\frac{\partial z}{\partial x} - \frac{\partial^2 F}{\partial p^2}\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 F}{\partial p\partial q}\frac{\partial^2 z}{\partial x\partial y} \\ &- \frac{\partial^2 F}{\partial q\partial y} - \frac{\partial^2 F}{\partial q\partial z}\frac{\partial z}{\partial y} - \frac{\partial^2 F}{\partial q\partial p}\frac{\partial^2 z}{\partial x\partial y} - \frac{\partial^2 F}{\partial q^2}\frac{\partial^2 z}{\partial y^2} = 0. \end{aligned} \tag{1.4}$$

The non-linear nature of these second-order equations implies that most often their solution cannot be expressed analytically, thus requiring to be determined through numerical or direct methods.

In this paper, we will show that these equations, intricate as they look and are, can nevertheless be immediately derived through economic reasoning. This derivation does not imply any calculation. Also it will allow us to obtain a new property of the calculus of variations.

## 2. AN IMMEDIATE DERIVATION OF THE EULER EQUATION THROUGH ECONOMIC REASONING: THE CASE OF AN EXTREMUM FOR $\int_a^b F(x, y, y')\, dx$.

Suppose we want to find $y(x)$ over an interval $[x_0, x_n]$ such that the functional $I[y(x)] = \int_a^b F(x, y, y')\, dx$ passes through an extremum. The problem is constrained by the fact that $y(x_0) = y_0$ and $y(x_n) = y_n$; the function $y(x)$ is supposed to be differentiable twice, and so is $F(x, y, y')$. Our first step is to rely on Euler's geometrical insight: using the fact that $y(x)$ is differentiable, we partition the interval $[x_0, x_n]$ into $n$ intervals of equal length $\Delta x = (x_n - x_0)/n$; it is then sufficient to determine the optimal values of $y$ at the end of the first $n-1$ intervals; the resulting polygonal line will tend toward the solution when $n$ tends to $\infty$. Although Euler did not investigate the validity of this limiting process, he was fully vindicated at the end of the 19th and at the beginning of the 20th century, in particular by those mathematicians who developed in the same vein direct methods, used today in computing - see for instance I. Gelfand and S. Fomin, *Calculus of variations*, 1963.

Our reasoning will be the following: if a solution exists, at any point of the optimal curve a change in its ordinate should impart an additional gain to the functional exactly equal to any possible additional cost, otherwise we would not have reached optimality. At all points of the polygonal line, and then at all points of the curve, any small change in the optimal value $y$ must entail the equality

$$\textbf{Additional gain to the functional} = \textbf{Additional cost} \tag{2.1}$$

We will now put the corresponding algebraic symbols on this equality. Consider any value $y_x$ on the polygonal line, and give it a *unit* increase (see Figure 2.1).

Without any loss in generality, suppose that both derivatives $\partial F/\partial y$ and $\partial F/\partial y'$ are positive (our reasoning would follow analogous lines if one of these derivatives, or both, were negative). Such a unit increase given to $y_x$ generates a positive increase of the functional, equal in linear approximation to $(\partial F/\partial y_x)\Delta x$; we call this quantity an *additional gain.*

Now this unit change in $y_x$ has two other consequences on the polygonal line and hence on the value of the functional: first, the slope of the segment between $y_{x-\Delta x}$ and $y_x$, initially equal to $s_{x-\Delta x} = (y_x - y_{x-\Delta x})/\Delta x$, is increased by $1/\Delta x$; secondly, since $y_{x+\Delta x}$ is fixed, the slope of the adjacent segment, $s_x = (y_{x+\Delta x} - y_x)/\Delta x$, is *reduced* by $1/\Delta x$; it is precisely this reduction in the slope of the polygonal line that generates a cost: the value of the functional will change, in linear approximation, by the negative value $(\partial F/\partial s_x)(\partial s_x/\partial y_x)\Delta x = (\partial F/\partial s_x)(-1/\Delta x)\Delta x = -\partial F/\partial s_x$. This is an additional *cost* equal to $\partial F/\partial s_x$, which will be reduced by the very fact that the slope of the segment linking $y_{x-\Delta x}$ to $y_x$ has increased, generating a gain for the functional $(\partial F/\partial s_{x-\Delta x})(1/\Delta x)\Delta x = \partial F/\partial s_{x-\Delta x}$. Overall, the *net additional cost* generated by a unit change in $y$ is equal to $\partial F/\partial s_x - \partial F/\partial s_{x-\Delta x}$. Thus our equation (2.1) reads

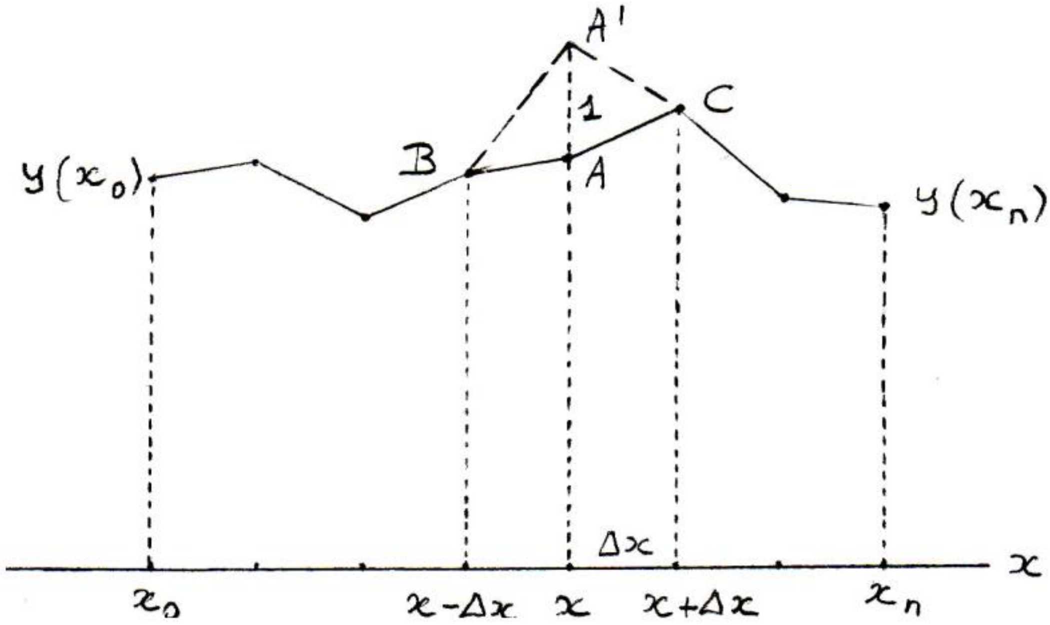


*Figure 2.1: Increasing the polygonal line by one unit generates a gain for the functional through two channels: the increase of the function at point A, and the increase of the slope of the line segment BA. At the same time, a cost is generated by the decrease of the slope of AC.*

$$\frac{\partial F}{\partial y_x}\Delta x = \frac{\partial F}{\partial s_x} - \frac{\partial F}{\partial s_{x-\Delta x}}. \tag{2.2}$$

Dividing by $\Delta x$ and taking the limit of (2.2) when $\Delta x \to 0$ and $n \to \infty$ yields the Euler equation (1.1).

## 3. THE CASE OF AN EXTREMUM FOR $\int\int_R F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dxdy$

We now extend this reasoning to the search of a twice differentiable function $z(x, y)$ leading to an extremum of $\int\int_R F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) dxdy$, where $F$ is twice differentiable; $z(x, y)$ is assumed to take fixed values at all points of the closed curve defining the boundary of $R$. Consider any point A$(x, y)$ on the optimal surface $z$, where $(x, y)$ is strictly inside the region $R$; points B, C, D and E, defined in Figure 3.1 in an obvious way also belong to the optimal surface $z$, and their horizontal coordinates belong to $R$.

The straight line segments BA, AC and DA, DE are approximations of the curves on the surface $z(x, y)$ joining points B, A, C and D, A, E, respectively. The slopes of segments BA, AC are $s_{x-\Delta x,y} = (z_{x,y} - z_{x-\Delta x,y})/\Delta x$ and $s_{x,y} = (z_{x+\Delta x,y} - z_{x,y})/\Delta x$ ; those of DA and AE are $\sigma_{x,y-\Delta y} = (z_{x,y} - z_{x,y-\Delta y})/\Delta y$ and $\sigma_{x,y} = (z_{x,y+\Delta y} - z_{x,y})/\Delta y$.

Let us us now give a *unit* increase to surface $z(x, y)$ at point A. Point A becomes A$'$. Asssume without any loss in generality that $\partial F/\partial z$, $\partial F/\partial p$ and $\partial F/\partial q$ are positive. Such a unit increase in $z$ first generates an additional gain for the functional equal to $(\partial F/\partial z)\Delta x\Delta y$, in linear approximation. On the other hand, it will entail two additional costs due to the fact that the slopes of segments AC and AE are *reduced* by amounts $1/\Delta x$ and $1/\Delta y$, respectively.

These translate into additional costs equal to

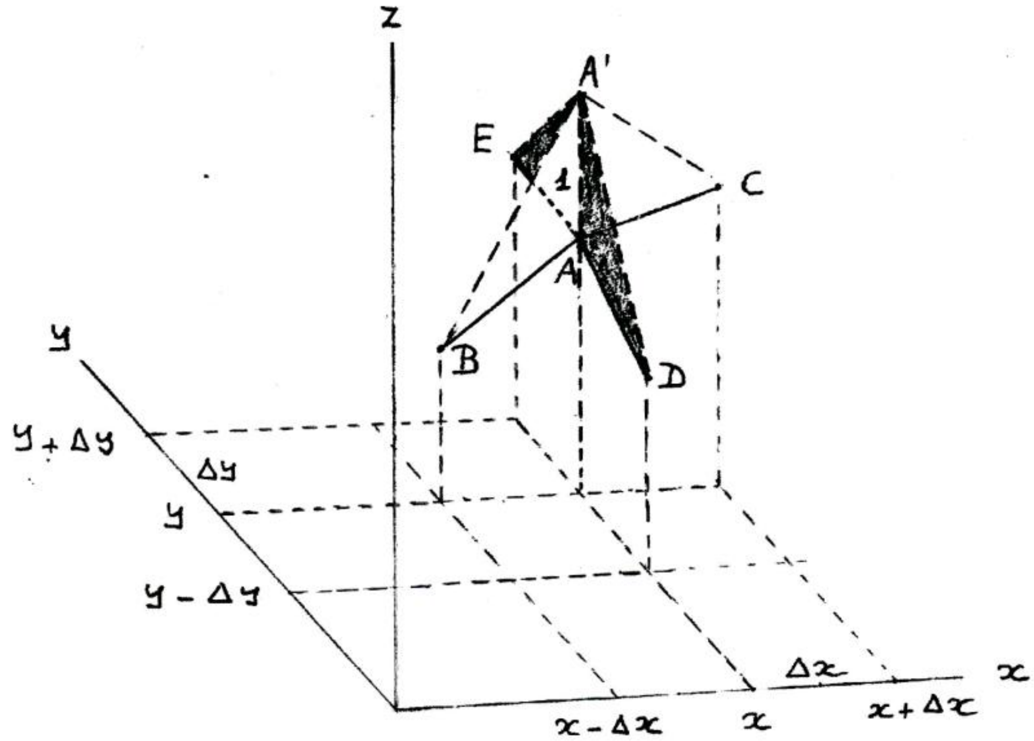


*Figure 3.1: Three factors contribute to an additional gain for the functional: the increase of the function from A to A′, and the increases in the slopes of BA and DA. Two additional costs are generated by the decreases in the slopes of AC and AE.*

$$(\partial F/\partial s_{x,y})(\partial s_{x,y}/\partial z)\Delta x\Delta y = (\partial F/\partial s_{x,y})(1/\Delta x)\Delta x\Delta y = (\partial F/\partial s_{x,y})\Delta y$$

and

$$(\partial F/\partial \sigma_{x,y})(\partial \sigma_{x,y}/\partial z)\Delta x\Delta y = (\partial F/\partial \sigma_{x,y})(1/\Delta y)\Delta x\Delta y = (\partial F/\partial \sigma_{x,y})\Delta x.$$

These additional costs are reduced by the fact that the slopes of BA and DA have increased by $1/\Delta x$ and $1/\Delta y$ respectively, carrying additional gains

$$(\partial F/\partial s_{x-\Delta x,y})(\partial s_{x-\Delta x,y}/\partial z)\Delta x\Delta y = (\partial F/\partial s_{x-\Delta x,y})(1/\Delta x)\Delta x\Delta y = (\partial F/\partial s_{x-\Delta x,y})\Delta y$$

for BA′ and

$$\partial F/\partial \sigma_{x,y-\Delta y})\Delta x$$

for DA′.

In total, the *net* additional costs are

$$(\partial F/\partial s_{x,y} - \partial F/\partial s_{x-\Delta x,y})\Delta y + (\partial F/\partial \sigma_{x,y} - \partial F/\partial \sigma_{x,y-\Delta y})\Delta x.$$

Our fundamental equality (2.1) now reads

$$(\partial F/\partial z)\Delta x\Delta y = (\partial F/\partial s_{x,y} - \partial F/\partial s_{x-\Delta x,y})\Delta y + (\partial F/\partial \sigma_{x,y} - \partial F/\partial \sigma_{x,y-\Delta y})\Delta x. \tag{3.1}$$

Dividing by $\Delta x\Delta y$ and letting $\Delta x$ and $\Delta y$ tend to zero yields the Euler-Ostrogradski equation (1.3). Note the simplicity of this intuitive derivation, made in the spirit of Euler, while an analytical demonstration requires a very clever application of Green's theorem, together with the fundamental lemma of the calculus of variations.

## 4. CONCLUSIONS.

It is a good place to recall how a young italian, citizen of Torino, at the time the capital city of the Kingdom of Sardinia, dramatically shifted the way optimization of functionals was approached. In 1755, nineteen-year old Ludovico de La Grange Tournier sent Euler a letter developing a purely analytic method to obtain first order conditions of such optimization. Herman Goldstine quotes the very words used by Euler to praise this discovery: "Even though the author of this [Euler] had meditated a long time and had revealed to friends his desire, yet the glory of first discovery was reserved to the very penetrating geometer of Turin La Grange who, having used analysis alone, has clearly attained the very same solution which the author had deduced from geometrical considerations" (Goldstine, op. cit., pp. 110-111).

Needless to say, tackling more complicated problems, involving higher dimensions or constraints, definitely required an analytic approach, and there was no going back to geometry. But we feel that Euler's geometric insight should not be sent to oblivion, at least for three reasons. First, for its mere simplicity and beauty – let us not forget that none of the great mathematicians at the turn of the 17th century had managed to find the *general* solution to the optimization of functionals. Second, because it allows to see a property of the calculus of variations which, to the best of our knowledge, has not yet be mentioned: we could have derived the Euler and Ostrogradski equations not necessarily by giving a variation to the optimal functions $y(x)$ or $z(x, y)$ – as is always done; we could have equally well started by giving a variation to their derivative(s). Indeed, increasing any of those derivatives would have generated outcomes exactly equal to those analysed above. And finally, Euler's geometric approach leads to a reasoning in terms of economics rendering those equations, involved as they may be, tremendously meaningful – we would be tempted to say "evident".

## REFERENCES

[1] LEV ELSGOLC, *Calculus of Variations,* International Series of Monographs in Pure and Applied Mathematics, Pergamon Press, Reading, MA, 1962.

[2] I. GELFAND and S. FOMIN, *Calculus of variations*, Englewood Cliffs, NJ, Prentice Hall, 1963.

[3] HERMAN H. GOLDSTINE, *A History of the Calculus of Variations*, Springer, 1980.