Euler Euler Everywhere

Using the Euler-Lagrange Equation to Solve Calculus of Variation Problems

Jeanine Smallwood
Principles of Analysis
Professor Flaschka

May 12, 1998

1. Introduction

Calculus of variations is a branch of the more general theory of calculus of functionals which deals specifically with optimizing functionals. In the late 1600s, John Bernoulli posed the brachistochrone problem, which marks the beginnings of calculus of variations. Although such types of problems had been considered by the ancient Greeks, the problems had only been solved geometrically, rather than analytically.

The question that John Bernoulli introduced in 1696 was as follows: "Given two points A and B in a vertical plane, to find the path AMB down which a movable point M, by virtue of its weight, proceed from A to B in the shortest possible time."[1] The problem was solved by John Bernoulli, James Bernoulli, Newton, L'Hospital and Leibniz. Besides the brachistochrone, other standard problems include geodesics, minimizing surfaces of revolution, and isoperimetric problems. All of these problems, including the brachistochrone, employ a common technique for finding the optimal solution. Each can be modelled by $J[y] = \int f(x, y, y')\, dx$ where $J : C^1 \longrightarrow \mathbf{R}$ and $y \in C^1$ is sought to optimize J. The problem can then be translated into a differential equation, called the Euler-Lagrange equation, and the problem reduces to one which can be solved using techniques of ordinary differential equations.

2. Theoretical Background

In the presentation of some introductory problems of calculus of variations, we will be considering functionals from C^1 to $\mathbf{R}$ of the form $J[y] = \int f(x, y, y')\, dx$. In order for these functionals to be continuous, we need to use the Sobolev norm on the space, which is defined to be:

$$\|y\| = \max_{a \le x \le b} |y(x)| + \max_{a \le x \le b} |y'(x)|.$$

We use this norm so that two functions will be "close" if both the functions *and* their derivatives are close in the standard sup-norm. Using the norm $\|y\| = \max_{a \le x \le b} |y(x)|$ does not satisfy the requirement for continuity for functionals of this form.

[1] Goldstine p. 30

To begin talking about necessary conditions for when a functional has a maximum or minimum, we need to have a definition of a derivative for a functional.

Definition 1: F is *differentiable* at u if there exists a bounded linear map $d_u F : C^1 \longrightarrow \mathbf{R}$ such that $\|F(u+h) - F(u) - d_u F(h)\| = \circ(\|h\|)$. [2]

Definition 2: Let $J[y]$ be a functional defined on a normed linear space and let

$$\Delta J[h] = J[y+h] - J[y]$$

be its increment. Suppose

$$\Delta J[h] = \phi[h] + \varepsilon \|h\|$$

where $\phi[h]$ is a linear functional and $\varepsilon \to 0$ as $\|h\| \to 0$. Then the principal linear part of the increment $\Delta J[h]$ is called the *variation* of $J[h]$ and is denoted by $\delta J[h]$.[3]

Theorem 1: A necessary condition for the differentiable functional $J[y]$ to have an extremum at $y = \hat{y}$ is that its variation vanish for $y = \hat{y}$, i.e., that

$$\delta J[h] = 0$$

for $y = \hat{y}$ and all h satisfying the constraints of the variational problem.

Proof: By contradiction. Suppose $J[y]$ has a minimum at $\hat{y}$, but there exists h_0 such that $\delta J[h_0] \neq 0$. We have that $\Delta J[h] = \delta J[h] + \varepsilon \|h\|$ where $\varepsilon \to 0$ as $\|h\| \to 0$. Since $\varepsilon > 0$, and $\|h\| > 0$, for sufficiently small $\|h\|$, the sign of $\Delta J[h]$ will be the same as $\delta J[h]$. Since $\hat{y}$ is a minimum, this implies that $\Delta J[h] \geq 0$ for all $\|h\|$ sufficiently small. So suppose without loss of generality that $\delta J[h_0] > 0$. Then for any $\alpha > 0$, since δJ is linear, $\delta J[-\alpha h_0] = -\alpha \delta J[h_0] < 0$ which implies that $\Delta J[-\alpha h] < 0$ which is a contradiction. Since $\hat{y}$ is a minimum, $\Delta J[h] \geq 0$ for all sufficiently small $\|h\|$.[4]

---

[2]Lega
[3]Gelfand p. 11
[4]Gelfand p. 13

This leads us to the question, given that there exists a continuous, twice-differentiable function which minimizes the functional $J[h] = \int_a^b f(x, y, y')dx$, what is the differential equation which this function must satisfy? This differential equation is known as the Euler-Lagrange equation.

Theorem 2: Let $J[y]$ be a functional of the form

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

defined on the set of functions $y(x)$ which have continuous first derivatives in [a,b] and satisfy the boundary conditions $y(a) = A$ and $y(b) = B$. Then a necessary condition for $J[y]$ to have an extremum for a given function $y = y(x)$ is that y satisfy the Euler-Lagrange equation:

$$F_y - \frac{d}{dx} F_{y'} = 0$$

Proof: Let $y(x)$ be the function for which $J[y] = \int_a^b F(x, y, y')\ dx$ has a weak extremum. We increment $y(x)$ by $h(x)$, where $h(a) = h(b) = 0$ in order for $y(x) + h(x)$ to satisfy the given boundary conditions.
Then define

$$\begin{aligned}
\Delta J &\equiv J[y + h] - J[y] \\
&= \int_a^b F(x, y + h, y' + h')\ dx - \int_a^b F(x, y, y')\ dx \\
&= \int_a^b [F(x, y + h, y' + h') - F(x, y, y')]\ dx.
\end{aligned}$$

Using Taylor's theorem we have:

$$F(x, y + h, y' + h') = F(x, y, y') + hF_y(x, y, y') + h'F_{y'}(x, y, y') + h.o.t$$

from which it follows that

$$\begin{aligned}
\Delta J &= \int_a^b [F(x, y, y') + hF_y(x, y, y') + h'F_{y'}(x, y, y') - F(x, y, y')]\ dx \\
&= \int_a^b [hF_y(x, y, y') + h'F_{y'}(x, y, y')]\ dx.
\end{aligned}$$

We now define

$$\Phi[h] \equiv \int_a^b [hF_y(x,y,y') + h'F_{y'}(x,y,y')]\ dx$$

and the higher order terms as $\varepsilon\|h\|$, so $\Delta J = \Phi[h] + \varepsilon\|h\|$. Since $\varepsilon\|h\| \to 0$ as $h \to 0$, this will give us $\Phi[h] = \int_a^b [hF_y(x,y,y') + h'F_{y'}(x,y,y')]\ dx$ as the variation of J. Denote the variation as

$$\delta J = \int_a^b [hF_y(x,y,y') + h'F - y'(x,y,y')]\ dx.$$

By Theorem 1, a necessary condition for $J[y]$ to have an extremum for $y = y(x)$ is that

$$\delta J = \int_a^b (hF_y + h'F_{y'})\ dx = 0$$

Integrating by parts, we have

$$\begin{aligned} \delta J &= \int_a^b (hF_y + h'F_{y'})\ dx = 0 \\ &= \int_a^b hF_y\ dx + F_{y'}h|_a^b - \int_a^b \frac{d}{dx}(F_{y'})\ dx = 0 \end{aligned}$$

But $h(a) = h(b) = 0$

$$\Longrightarrow \delta J = \int_a^b h(F_y - \frac{d}{dx}(f_{y'}))\ dx = 0$$

Since this is true for all $h \Rightarrow F_y - \frac{d}{dx}(F_{y'}) = 0$, which is the Euler-Lagrange equation.[5]

The solution of the Euler-Lagrange equation applied to any of the above-mentioned introductory problems is the twice-differentiable function that is a minimum of the integral in question. The Euler-Lagrange equation can be generalized to functionals depending on higher-order derivatives as well as to functionals that have more than one independent variable.

[5]Gelfand p.14

3. Background on some Introductory Problems

There are four classical types of problems that are solved using the calculus of variations. While the brachistochrone marked the beginning of interest in the subject, the other three classical problems were variations on the original brachistochrone problem.

The study of geodesics arises from the problem of finding the arc with the shortest possible length lying on the surface of a sphere that connects two given points on the surface. In general, the arc of minimum length connecting two points on a given suface is called the *geodesic* for the surface.[6] On the surface of a sphere, the result is the intersection of the sphere with the plane containing the center of the sphere and the two given points – the great circle arc. In finding the geodesic of a circular cylinder $r = (a\cos\phi, a\sin\phi, z)$, the result is $z = c_1\phi + c_2$, which is a two-parameter family of helical cones lying on the cylinder.[7]

Minimum surface of revolution problems can generally be stated as follows: Given two fixed points (x_1, y_1), (x_2, y_2), we would like to pass an arc through them whose rotation about the x-axis generates a surface of revolution whose area included in $x_1 \leq x \leq x_2$ is a minimum. The result in this case is usually a catenary. In solving this problem, we find a family of catenaries that pass through one of the points and satisfies the equations:

$$\begin{aligned} y &= b\cosh\frac{(x-a)}{b} \\ y_1 &= b\cosh\frac{(x_1-a)}{b}. \end{aligned}$$

If the point (x_2, y_2) is outside the envelope of the family, no catenary will satisfy the endpoint conditions and the only solution will be the *Goldschmidt discontinuous solution*, which is defined by[8]

$$\begin{cases} x = x_1 & 0 \leq y \leq y_1 \\ y = 0 & x_1 \leq x \leq x_2 \\ x = x_2 & 0 \leq y \leq y_2. \end{cases}$$

Isoperimetric problems are a class of problems that have added restrictions on the functions besides continuity and end-point conditions. A problem would be stated in a form such as:

[6] Weinstock p. 26
[7] Gelfand p. 38
[8] Weinstock p. 30

> Find the curve $y = y(x)$ for which the functional $J[y] = \int_a^b F(x, y, y')\, dx$ has an extremum, where the admissible curves satisfy the boundary condtions $y(a) = A$, $y(b) = B$ and are such that another functional $K[y] = \int_a^b G(x, y, y')\, dx$ takes a fixed value l.[9]

The best known problem of this class is that of finding the closed curve of a given perimeter for which the area is a maximum, the solution being a circle.

4. The Brachistochrone

We begin some examples of the use of the Euler-Lagrange equation with a classical calculus of variations problem, the brachistrochrone. Several solutions to the brachistochrone problem were put forth, most notably by Newton, Leibniz, L'Hospital, and the Bernoulli brothers. John Bernoulli, who initially posed the problem, gave two solutions, one which used Fermat's principle of least time. The two problems are similar in that Fermat showed that a ray of light moves through a medium on a path to reach a boundary in the shortest time possible. Bernoulli used this idea by dividing up the space between the two fixed endpoints and assigning an index of refraction inversely proportional to the particle's velocity in that layer. Then as the particle moves through each slice, he used Snell's Law to determine how the particle bent as it crossed the interface.[10]

The method used below was not available to Bernoulli at the time when he introduced the brachistochrone problem. Restating the original problem, we have:

> Given two points (x_1, y_1), (x_2, y_2), find the curve that takes the least time to go between these points under the influence of gravity.

We assume that a particle is descending the curve $y(x)$ with an initial velocity v_i. If we let s be the arclength of the curve $y(x)$, then the velocity

[9] Gelfand p. 43

[10] Goldstine p. 39

v of the particle is $\frac{ds}{dt}$. This gives us:

$$\begin{aligned}\frac{ds}{dt} = \frac{ds}{dx} &= \sqrt{1+(y')^2} \\ \Longrightarrow ds &= \sqrt{1+(y')^2}dx\end{aligned}$$

So we have

$$\begin{aligned} v &= \frac{ds}{dt} \\ \Longrightarrow t &= \int_{x_1}^{x_2} \frac{ds}{v} \\ &= \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{v} dx \end{aligned}$$

Since the particle is moving down a frictionless surface, energy is conserved, implying that $\frac{1}{2}mv^2 - \frac{1}{2}mv_i^2 = mg(y-y_1)$ where m is the mass of the particle and g is the acceleration due to gravity. This gives us

$$v^2 = 2g(y - y_1 + \frac{v_i^2}{2g})$$

Letting $k = \frac{v_i^2}{2g}$, we have

$$\begin{aligned} v &= \sqrt{2g(y-y_1+k)} \\ \Longrightarrow t &= \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{y-y_1+k}} dx \end{aligned}$$

To solve this, we can now use the Euler-Lagrange equation. We have a functional of the form $J[y] = \int_a^b F(y, y')\ dx$. Since our integrand does not depend on x, we can derive a special case of the Euler-Lagrange equation. Since F is only a function of y and y',

$$\Longrightarrow F_y - \frac{d}{dx}F_{y'} = F_y - F_{y'y}y' - F_{y'y'}y'' \ = \ 0.$$

Multiplying through by y', and adding and subtracting $F_{y'}y''$, we have

$$y'F_y + y''F_{y'} - (y')^2F_{y'y} - (y')y''F_{y'y'} - y''F_{y'} = 0$$

$$\Longrightarrow \frac{d}{dx}(F - y'F_{y'}) = 0$$
$$\Longrightarrow F - y'F_{y'} = C$$

In our problem, the integrand is $F(y, y') = \frac{\sqrt{1+(y')^2}}{\sqrt{y-y_1+k}}$, and so the Euler-Lagrange equation will be

$$F - y'F_{y'} = \frac{\sqrt{1+(y')^2}}{\sqrt{(y-y_1+k)}} - \frac{(y')^2}{\sqrt{(y-y_1+k)}\sqrt{1+(y')^2}} \;=\; C$$

If we let $C = 1/\sqrt{2a}$, then we have

$$\begin{aligned}
\frac{1}{\sqrt{(y-y_1+k)}\sqrt{1+(y')^2}} &= \frac{1}{\sqrt{2a}} \\
\Longrightarrow (y')^2 &= \frac{2a-(y-y_1+k)}{y-y_1+k} \\
\Longrightarrow x &= \int \frac{\sqrt{y-y_1+k}}{\sqrt{2a-(y-y_1+k)}}\, dy.
\end{aligned}$$

Letting
$$\begin{aligned}
y - y_1 + k &= 2a\sin^2\tfrac{\theta}{2} \\
dy &= 2a\sin\tfrac{\theta}{2}\cos\tfrac{\theta}{2}\; d\theta
\end{aligned}$$

gives us

$$\begin{aligned}
x &= 2a\int\sqrt{\frac{2a\sin^2\frac{\theta}{2}}{2a\cos^2\frac{\theta}{2}}}\cdot\sin\frac{\theta}{2}\cos\frac{\theta}{2}\; d\theta \\
&= 2a\int sin^2\frac{\theta}{2}\; d\theta \\
\Longrightarrow x &= a\int(1-cos\theta)\; d\theta \\
\Longrightarrow x &= a(\theta - \sin\theta) + C_2
\end{aligned}$$

So we have a parameterized equation for the curve given by

$$\begin{aligned}
y &= (y_1 - k) + a(1-\cos\theta) \\
x &= C_2 + a(\theta - \sin\theta)
\end{aligned}$$

which is the parametrization of a cycloid, generated by the motion of a fixed point on the circumference of a circle of radius a.[11]

5. Isoperimetric Problems

In this section, we will look at another class of problems, called isoperimetric problems, which involve optimizing a functional with an added constraint. Recall that the original isoperimetric problem seeks to find a curve which encloses the largest area given curves of a prescribed length. The problem that we will discuss now is that of the hanging rope. In this problem, the goal is to find the shape of a flexible rope of uniform density which hangs at rest with its endpoints fixed.

Solutions to this, and other isoperimetric problems, can be obtained using the method of Lagrange multipliers. Given a functional $I = \int F(x, y, y')\ dx$ we wish to constrain it by another functional $J = \int G(x, y, y')\ dx$. Define $I^* = I + \lambda J which implies I^* = \int F^* = \int F + \lambda G$. Using the same method to find the Euler-Lagrange equation as before, it can be shown that for the constrained problem, the Euler-Lagrange equation will be

$$F^*_y - \frac{d}{dx}F^*_{y'} = 0.$$

To find our rope of interest, we assume that the rope is hanging in the vertical plane. We let $y = y(x)$ be a rope with the designated boundary conditions and prescribed length, and σ be the constant mass per unit length of the rope. Then the potential energy of a portion of the rope having length ds will be $gy\sigma ds$, where g is the acceleration due to gravity. This implies that the total potential energy of the rope is

$$\int_0^l \sigma gyds = \sigma g \int_{x_1}^{x_2} y\sqrt{1 + (y')^2}dx,$$

where (x_1, y_1), (x_2, y_2) are the fixed endpoints. The constraint in this problem is that $J = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}dx$, i.e. the arclength of the rope has a length l. Using Lagrange multipliers, we are now interested in optimizing

$$\begin{aligned} I^* &= \int \sigma gy\sqrt{1 + (y')^2}dx + \lambda \int \sqrt{1 + (y')^2}dx \\ &= \sigma g \int \sqrt{1 + (y')^2}\left(y + \frac{\lambda}{\sigma g}\right) dx. \end{aligned}$$

[11]Weinstock p. 29

The Euler-Lagrange equation for this problem will then be

$$\sigma g\sqrt{1+(y')^2} - \frac{d}{dx}\left(\frac{\sigma g y'}{\sqrt{1+(y')^2}}\left(y+\frac{\lambda}{\sigma g}\right)\right) = 0.$$

Since F^* is independent of x, we can write the Euler-Lagrange equation as

$$\begin{aligned}
y'\left(\frac{\sigma g y'}{\sqrt{1+(y')^2}}\left(y+\frac{\lambda}{\sigma g}\right)\right) - \sigma g\sqrt{1+(y')^2}\left(y+\frac{\lambda}{\sigma g}\right) &= C_1 \\
\Longrightarrow \sigma g\left(y+\frac{\lambda}{\sigma g}\right)\left[\frac{(y')^2}{\sqrt{1+(y')^2}} - \sqrt{1+(y')^2}\right] &= C_1 \\
\Longrightarrow -(\sigma g y+\lambda)\left(\frac{1}{\sqrt{1+(y')^2}}\right) &= C_1
\end{aligned}$$

Solving for y', we have

$$\begin{aligned}
(y')^2 &= \frac{1}{C_1^2}(\sigma g y+\lambda)^2 - 1 \\
\Longrightarrow x &= C_1\int\frac{dy}{\sqrt{(\sigma g y+\lambda)^2-1}} \\
&= \frac{C_1}{\sigma g}\ln\left|\sigma g y+\lambda+\sqrt{(\sigma g y+\lambda)^2-1}\right| + C_2 \\
&= \frac{C_1}{\sigma g}\cosh^{-1}(\sigma g y+\lambda) + C_2 \qquad \sigma g y+\lambda \geq 1
\end{aligned}$$

Writing y in terms of x we have

$$y = \frac{1}{\sigma g}\cosh\left(\frac{\sigma g(x-C_2)}{C_1}\right) - \frac{\lambda}{\sigma g}.$$

We can see that the shape of the hanging rope that we are interested in is a catenary with a vertical axis. By applying any given boundary conditions and knowing the length of the rope we can determine the constants λ, C_1, and C_2.

6. Conclusion

There are many variations to the problems which have been discussed. The brachistochrone problem can be restated as a variable endpoint problem, where we look for a curve where a particle reaches a vertical line in the least amount of time. For the isoperimetric problems, a number of finite subsidary conditions can be added for one specific functional to be extremized.

Although all of the problems presented here have only involved optimizing functionals that result in a function with continuous first and second derivatives, the methods can easily be extended to problems which involve solutions with higher order derivatives. The problems can also be generalized to involve optimizing functionals that depend on n continuously differentiable arguments. In each of these cases, we can derive the appropriate Euler-Lagrange equation for which the extremal function must satisfy.

All of the problems discussed above have smooth solutions as the extremizing function. However, many variational problems have no smooth solutions, and for these problems we must expand the class of admissible functions to piecewise smooth curves. The Euler-Lagrange equation must then be satisfied on each piecewise interval.

The Euler-Lagrange equation is a tough cookie and is central to many of the problems of calculus of variations. It is remarkable that so many problems from so many different settings ultimately boil down to solving this same equation. But then, the unifying power of the Euler-Lagrange equation is probably why mathematicians are still using this equation more than two hundred years after its conception.

Bibliography

[1] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

[2] Herman H. Goldstine. *A History of the Calculus of Variations from the 17th Through the 19th Century*. Springer-Verlag, New York, 1980.

[3] Joceline Lega. Lecture notes, April 1998.

[4] Robert Weinstock. *Calculus of Variations*. Dover Publications, Inc., New York, 1974.