THE CALCULUS OF VARIATIONS

AND THE BRACHISTOCHRONE PROBLEM

Introduction and Rationale

The Calculus of Variations is a field of calculus which is concerned with the maxima and minima of mathematical objects called *functionals*, which are basically functions mapping from a set of functions to the set of say, real numbers.

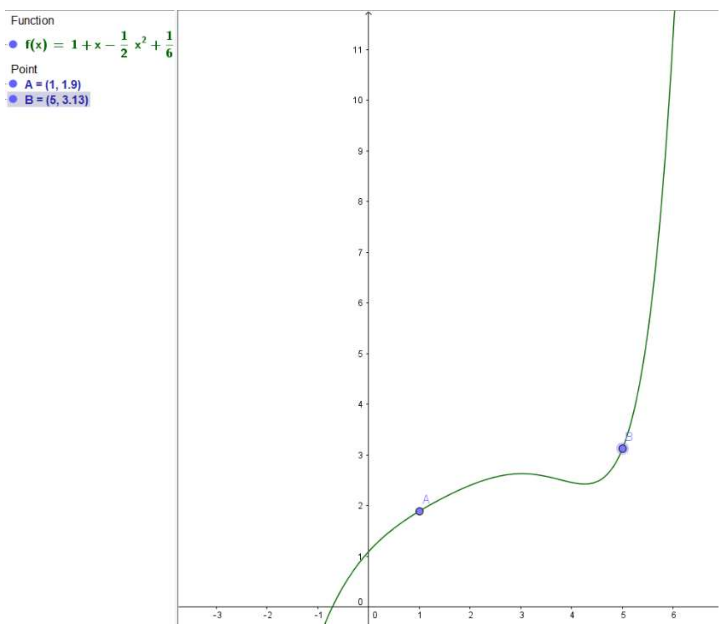
This tool has a great number of applications, especially in physics, where the majority of modern physics is based on the Calculus of Variations. As I am interested in fundamental physics, I naturally found this field intriguing and worthy of study.

The Calculus of Variations also apparently has applications in econometrics, which I have not yet understood.

Example 1: Length of a curve

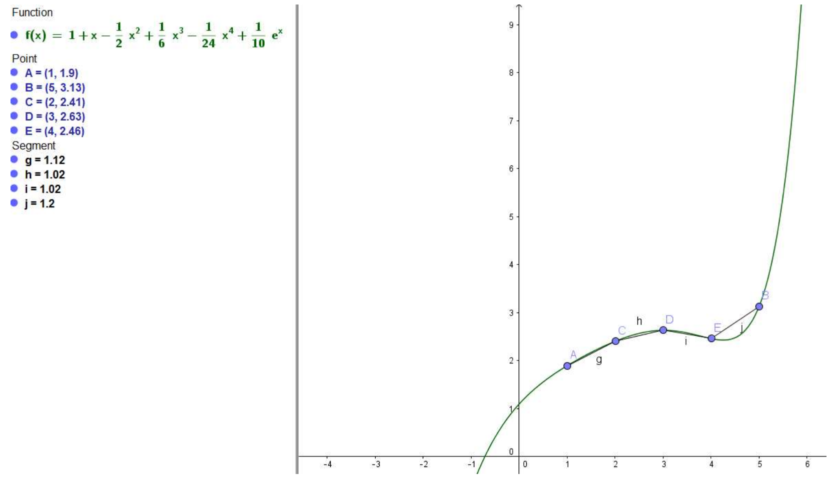
Before formally introducing and defining a functional, we will discuss the simple example of one specific functional: the length of a curve.

Take some function $y = f(x)$ and plot it on the *xy*-plane, and say we want to find the length of the resulting curve plotted between two points $x = a$ and $x = b$. Note that when $x = a$, $y = f(a)$ and when $x = b$, $y = f(b)$, and thus we could say that our two endpoints are $A = \left(a, f\left(a\right)\right)$ and $B = \left(b, f\left(b\right)\right)$.

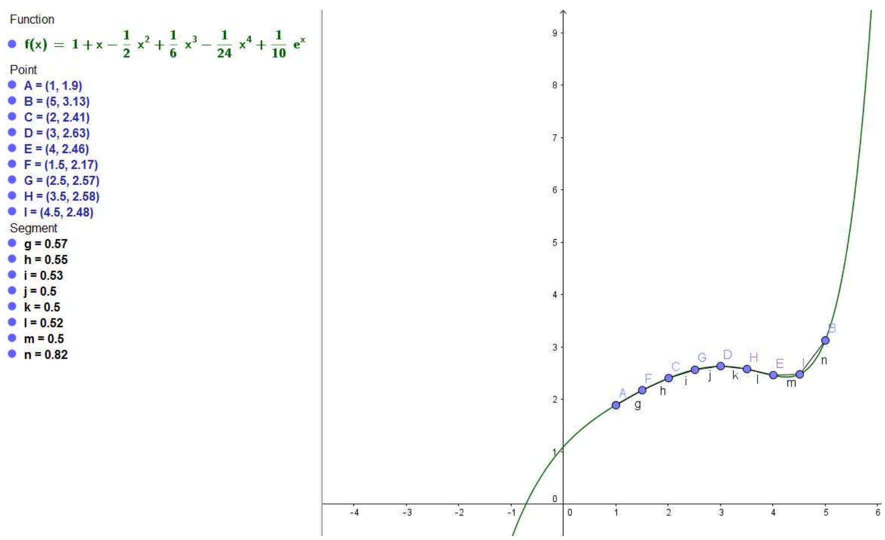


To estimate the length of this curve segment, we break the curve into small pieces whose length we *can* calculate. For example, we can divide the curve into the segments AC, CD, DE, and EB as follows:

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey



Then we can say that the length of the curve segment/arc AB can be approximated as $AB \approx AC + CD + DE + EB$. Notice that we have taken intervals whose length has a horizontal component of 1 here, but we can take any real number between 0 and *b* – *a* here, and the smaller value we take, the more accurate our estimate becomes. For example, if we take the horizontal component to be ½ instead, we have $AB \approx AF + FC + CG + GD + DH + HE + EI + IB$, as shown in the following graph:



In general, we can say that the length *S* can be approximated as:

$$S \approx \sum_{x=a}^{b} \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

Where the step of the summation index is taken as Δx, as opposed to 1 (i.e. instead of summing over *x* = *a*, *x* = *a* + 1, *x* = *a* + 2,… *x* = *b* we sum over $x = a,\ x = a + \Delta x,\ x = a + 2\Delta x...,\ \ x = b$), and $b / \Delta x$ must be an integer so there is an integer number of line segments to take.

This approximation becomes more and more accurate as Δx decreases, and becomes perfectly accurate at the limit that it becomes zero.

Therefore we can say that:

$$S = \lim_{\Delta x \to 0} \sum_{x=a}^{b} \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

But the limit of a sum (as the increment in the index of the summation approaches zero) is just a definite integral, and the limit of Δx approaching zero is the infinitesimal, *dx*. Hence

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey

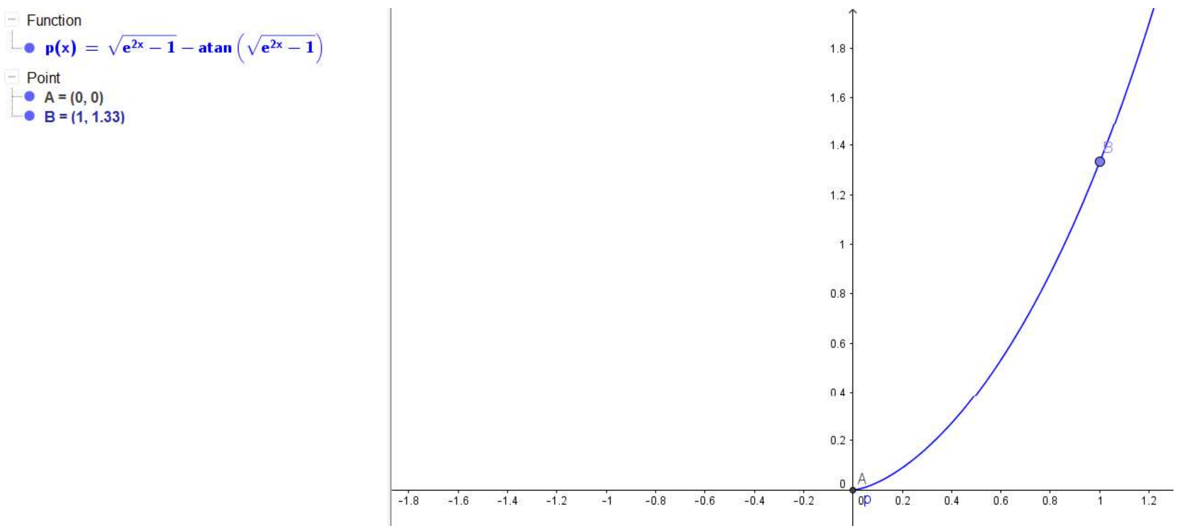
$$S=\int_{x=a}^{x=b}\sqrt{(dx)^2+(dy)^2}$$

This is superficially a difficult problem to solve, but we can algebraically manipulate it to look like this, by taking the *dx* outside:

$$S=\int_a^b\sqrt{1+(dy/dx)^2}\,dx$$

In fact, in many formal areas of mathematics, the length of a curve is *defined* in this way.

Example 1: Find the length of the curve $y=\sqrt{e^{2x}-1}-\arctan\sqrt{e^{2x}-1}$ from $x = 0$ to $x = 1$, i.e. between points *A* and *B* in the following curve.



Solution:

$$\frac{dy}{dx}=\frac{d}{dx}\sqrt{e^{2x}-1}-\frac{d}{dx}\arctan\sqrt{e^{2x}-1}=\frac{1}{2\sqrt{e^{2x}-1}}2e^{2x}-\frac{1}{1+e^{2x}-1}\frac{1}{2\sqrt{e^{2x}-1}}2e^{2x}=\frac{e^{2x}}{\sqrt{e^{2x}-1}}-e^{-2x}\frac{e^{2x}}{\sqrt{e^{2x}-1}}$$
$$=\frac{e^{2x}-1}{\sqrt{e^{2x}-1}}=\sqrt{e^{2x}-1}$$

$$S=\int_0^1\sqrt{1+(dy/dx)^2}\,dx=\int_0^1\sqrt{1+\sqrt{e^{2x}-1}^{\,2}}\,dx=\int_0^1 e^x dx=\left[e^x\right]_{x=0}^{x=1}=e-1$$

Introduction to functionals and variations

The length, which we discussed in the previous section, is an example of a functional. A functional is basically a one-to-one or many-to-one mapping from some set of functions to some set of real numbers. The function can generally have any domain which is a subset of the real numbers. Some examples of functionals are:

- Length
- Area under the curve
- Volume of the solid of revolution formed when the curve is rotated around some axis
- Midpoint of the curve
- Work done along the path by some force
- Time taken for a particle to transverse the path
- Action integral of the path
- Average value of some scalar field along the path, e.g. temperature

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey

Many of these functionals have important applications in other fields, which we will later discuss.

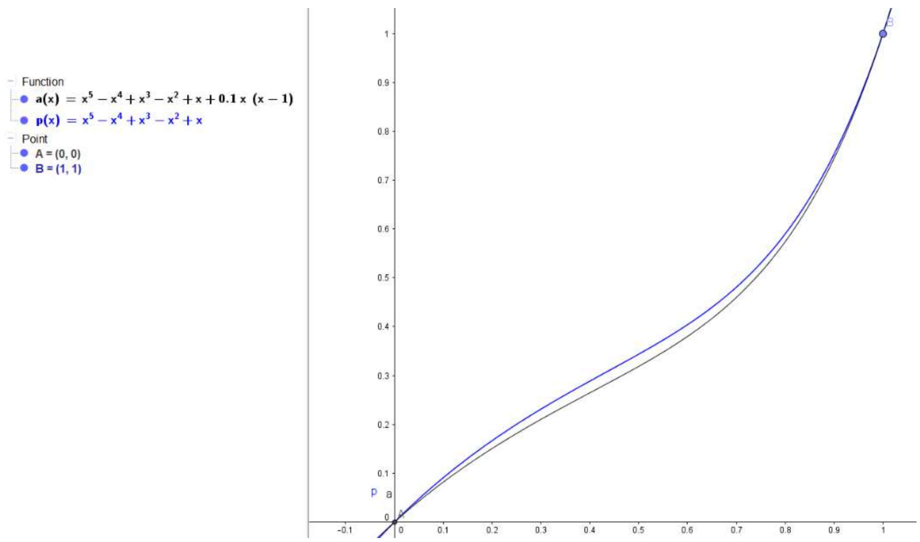
A functional can be written as a definite integral along the domain of some function of the parameter *x*, the *y*-co-ordinate *y*(*x*) and its derivatives, i.e.

$$F[y(x)] = \int_a^b f\left(x,\ y(x),\ y'(x)\right)\, dx$$

Note that the usage of the term "*y*(*x*)" on the left-hand-side refers to the entire curve, or schedule, whereas *y*(*x*) on the right-hand-side refers to the *y*-co-ordinate of the curve at some *x*, i.e. the *value* of *y*(*x*) at some *x*. An appropriate analogy from economics would be demand vs. quantity of demand, or supply vs. quantity of supply.

In the case of length, for instance, $f(x,y,y') = \sqrt{1+(y')^2}$ for any given function *y*(*x*).

Now, a variation is just what it sounds like – it's a variation in the function *y*(*x*), written as replacing $y(x)$ with $y(x)+\epsilon h(x)$, where $\epsilon \to 0$. For example take the following graph:



Here, the curve to the left is *p*(*x*) and the curve to the right is the perturbation *a*(*x*), with $h(x) = x(x-1)$. As ϵ, currently 0.1, approaches 0, what function *h*(*x*) is, ceases to be significant.

Note that the variation shown in the graph has fixed endpoints at (0,0) and (1,1), i.e. the variation vanishes for these endpoints.

Recall that we had earlier stated the point of the calculus of variations to be to study the functions which yield an extremal value of some functional – e.g. minimum length, maximal action, minimum time taken to transverse the path, etc. Recall from elementary calculus that the extremal value of some function is given at the point where its derivative is zero. Similarly, the extremal value of a functional is given at the point where its variation is zero, i.e.

$$\delta F = 0$$

Where δF is the variation in *F*. This can often be expressed with respect to the function itself, as in $\dfrac{\delta F}{\delta y(x)} = 0$ – i.e. the functional is stationary for an infinitesimal variation in the function by some $\epsilon h(x)$. Once again, recall that we're varying the function itself – we aren't moving along the function by incrementing *x* and therefore changing *f*(*x*). Once again, to give an analogy from economics, we're talking about non-price determinants shifting the supply or demand curve as opposed to movements along the supply or demand curve.

δF, of course, can be more formally defined as

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey

$$\delta F = F(x,\, y+\epsilon,\, y'+\epsilon') - F(x,y,y') = \int_a^b \left(f(x,\, y+\epsilon,\, y'+\epsilon') - f(x,\, y,\, y')\right)dx$$

This is analogous to the definition of the derivative.

The Euler-Lagrange equation

The Euler-Lagrange equation gives us a convenient way to solve for the extremum of a functional. Basically it states that for some given functional $F[y(x)] = \int_a^b f\left(x,\ y(x),\ y'(x)\right)\, dx$, if some $y(x)$ is a curve which satisfies $\delta F = 0$, then the following differential equation must be satisfied:

$$\frac{\partial f}{\partial y(x)} = \frac{d}{dx}\frac{\partial f}{\partial y'(x)}$$

This equation is known as the *Euler-Lagrange Equation*. Note: a partial derivative $\partial y / \partial x$ is calculated pretty much in the same way as a standard derivative, except that we only differentiate with respect to the term when it is explicitly present in the expression for the multivariable ufnction.

Although a general proof is difficult, I managed to prove this result given two fixed endpoints, as follows:

$$\begin{aligned} 0 = \delta F &= \int_a^b \left(f(x,\, y+\epsilon,\, y'+\epsilon') - f(x,\, y,\, y')\right)dx \\ &= \int_a^b \left(\epsilon\frac{\partial L}{\partial y(x)} + \frac{\partial L}{\partial y'(x)}\frac{d\epsilon}{dx}\right)dx \\ &= \int_a^b \left(\epsilon\frac{\partial L}{\partial y(x)} + \epsilon'\frac{\partial L}{\partial y'(x)}\right)dx \end{aligned}$$

Note that here, we're using the symbol ϵ to represent what we previously called $\epsilon h(x)$, for the sake of being compact. We used the chain rule using partial derivatives, $\frac{d}{dx}\psi(x,y) = \frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial y}\frac{dy}{dx}$, in the last step above. Now we can apply integration by parts as follows:

$$0 = \left[\epsilon\frac{\partial f}{\partial y'(x)}\right]_a^b - \int_a^b \left(\epsilon\frac{\partial f}{\partial y} - \epsilon\frac{d}{dx}\frac{\partial f}{\partial y'(x)}\right)dx$$

The first term on the right-hand-side is zero due to our boundary conditions, and thus we get $\epsilon\frac{\partial f}{\partial y} - \epsilon\frac{d}{dx}\frac{\partial f}{\partial y'(x)} = 0$, or $\frac{\partial f}{\partial y} = \frac{d}{dx}\frac{\partial f}{\partial y'(x)}$.

Example 2: Use the Euler-Lagrange equation to prove that the shortest path between two points, given no other conditions, is a straight line.

Solution: So we're asked to take the length of the curve as our functional, and prove that $\frac{\partial f}{\partial y} = \frac{d}{dx}\frac{\partial f}{\partial y'(x)}$ implies that $y(x)$ has an equation of the form $y = mx + c$. We know that $f = \sqrt{1+(y')^2}$, which is only dependent on y', not on y, therefore $\frac{d}{dx}\frac{\partial\sqrt{1+(y')^2}}{\partial y'(x)} = 0 \Rightarrow \frac{d}{dx}\frac{y'}{\sqrt{1+(y')^2}} = 0$. Integrating both sides dx, we see that $\frac{y'}{\sqrt{1+(y')^2}} = C_1$. With some algebraic

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey

simplification, we see that $\sqrt{\frac{(y')^2}{1+(y')^2}} = C_1 \Rightarrow y' = \frac{1}{\sqrt{\frac{1}{C_1^2}-1}}$. Integrating both sides *dx*, we obtain $y(x) = \frac{1}{\sqrt{\frac{1}{C_1^2}-1}} x + C_2$.

Defining new arbitrary constants $m = \frac{1}{\sqrt{\frac{1}{C_1^2}-1}}$ and $c = C_2$, we get the required form $y = mx + c$.

Betrami identity

A special case of the Euler-Lagrange equation can be obtained when $\partial f / \partial x = 0$. Due to lack of space and duration, I will not be able to derive it here and will leave it as an exercise to the reader.

$$f - y' \frac{\partial f}{\partial y'} = C$$

This is known as the Betrami identity. An important application of the Betrami identity is in the Brachistochrone problem.

Example 3: The Brachistochrone problem

Find the shape of the curve down which a particle sliding from rest and accelerating due to gravity without friction from one point to another in the least time.

Solution: The duration taken for a particle to transverse a path is also a functional, given by $t = \int_{P_1}^{P_2} ds / v$, where *v* is the speed and *ds* is the infinitesimal unit of distance transversed. From conservation of energy, we know that $\frac{1}{2} mv^2 = mgy \Rightarrow v = \sqrt{2gy}$, and it is basic Pythagorean theorem that tells us that $ds = \sqrt{dx^2 + dy^2} = dx\sqrt{1+(y')^2}$. Hence, $t = \int_{P_1}^{P_2} \sqrt{\frac{1+(y')^2}{2gy}} dx$. Rather than employing the Euler-Lagrange equation at this stage, we can use Betrami's identity as *x* does not explicitly appear in the integrand, hence $\partial f / \partial x = 0$. Since $\partial f / \partial y' = y'\left(1+(y')^2\right)^{-1/2} (2gy)^{-1/2}$, plugging this in to the identity yields $\frac{1}{\sqrt{2gy}\sqrt{1+y}} = C$, and thus $1+(y')^2 = K / y$ for positive *K*. Solving this differential equation is not within the scope of this report, but the solution can in fact be parametrically expressed as follows:

$$x = (\theta - \sin\theta) K / 2, \quad y = (1 - \cos\theta) K / 2$$

Which are the equations for a shape known as a "cycloid".

Conclusion

This report serves as an introduction to the calculus of variations. The field is, of course, much wider, starting with issues like constrained variational problems and complicated boundary conditions which I did not discuss in this report. The field is extremely important in modern physics, where pretty much all fundamental theories, from General Relativity to String Theory, are or can be expressed in terms of what is called the "Principle of Least Action". Another important idea in the Calculus of Variations is the Path Integral, which is also of very significant importance in all physics that is quantum mechanical, which we did not discuss at all. The specific case of the Brachiostrone problem also has important applications in engineering, especially once you generalize it to an environment with friction or air resistance, and to rigid bodies rather than just point particles. This is an extremely important mathematical tool which serves as a stepping stone to most of advanced mathematics.

Gelfand & Fomin (1963), *Calculus of Variations* : Dover, New Jersey