MATH 354: Calculus of Variations

Brief lecture notes

Rashid M. Talha

School of Natural Sciences, NUST

(Date: March 22, 2024)

Textbook: Calculus of Variations, I. M. Gelfand and S. V. Fomin

1 Euler-Lagrange Equation

Theorem 1.1. *Consider the functional*

$$J[y] = \int_a^b F(x, y, y') \, dx, \tag{1}$$

defined on the set of functions $y \in C^1[a, b]$, with $y(a) = A$ and $y(b) = B$. Then, a necessary condition for $J[y]$ to have an extremal is that y satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0. \tag{2}$$

Proof. Let y be an extremal of $J[y]$. Let $h \in C^1[a, b]$ be an arbitrary function with $h(a) = h(b) = 0$, and consider the variation $y \to y + h$. If we assume this variation to be small then we can use Taylor expansion to obtain

$$\delta J = \int_a^b F(x, y + h, y' + h') \, dx - \int_a^b F(x, y, y') \, dx \tag{3}$$

$$= \int_a^b h \partial F \partial y(x, y, y') + h' \partial F \partial y'(x, y, y') \, dx \tag{4}$$

$$= \int_a^b h\left(\frac{\partial F}{\partial y}(x, y, y') + \frac{\partial F}{\partial y'}(x, y, y')\right) dx - h(x)F(x, y, y')\Big|_a^b \tag{5}$$

$$= \int_a^b h\left(\frac{\partial F}{\partial y}(x, y, y') + \frac{d}{dx}\left(\frac{\partial F}{\partial y'}(x, y, y')\right)\right) dx. \tag{6}$$

Here, the boundary term vanishes due to $h(a) = h(b) = 0$.

Now, requiring $\delta J = 0$ (the condition for y being an extremal of $J[y]$) and using the fundamental theorem of calculus of variations, we get

$$\frac{\partial F}{\partial y}(x, y, y') + \frac{d}{dx}\left(\frac{\partial F}{\partial y'}(x, y, y')\right) = 0 \tag{7}$$

as required. $\square$

This is a second order ODE in y

$$\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial x \partial y'} - y'\frac{\partial^2 F}{\partial y \partial y'} - y''\frac{\partial^2 F}{\partial y' \partial y'} = 0 \tag{8}$$

We shall later see that if the highest order derivative contained in the integrand is $y^{(k)}$, then the resulting Euler-Lagrange equation is a $2k$ order ODE. Similarly, if the integrand contains multiple independent variables then we obtain a PDE instead.

The integral curves of equation (2) are called extremals (or critical points) of J.

2   First Integrals and Special Cases

For some integrands $F(x, y, y')$ the Euler-Lagrange equation (2) can be reduced into a first order ODE, called the first integral. In some other cases, it can be rearranged into a more practically useful form.

1. $F = F(x, y')$
   Here, $\frac{\partial F}{\partial y} = 0$. As a result, $\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0$, if y satisfies the Euler-Lagrange equation. Therefore, y satisfies the Euler-Lagrange equation if and only if

$$\frac{\partial F}{\partial y'} = c \tag{9}$$

   for some constant $c \in \mathbb{R}$.

2. $F = F(y, y')$
   It is easy to show that, in this case, y satisfies the Euler-Lagrange equation if and only if

$$F - y'\frac{\partial F}{\partial y'} = c, \tag{10}$$

   for some constant $c \in \mathbb{R}$. This is because

$$\frac{d}{dx}\left(F - y'\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial x} + y'\frac{\partial F}{\partial y} + y''\frac{\partial F}{\partial y'} - y''\frac{\partial F}{\partial y'} - y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \tag{11}$$

$$= y'\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right). \tag{12}$$

   So, if the Euler-Lagrange equation (2) is satisfied then the LHS is zero. Conversely, if the LHS is zero, then the Euler-Lagrange equation (2) is satisfied. (The case $y' = 0$, when the LHS vanishes, leads to $F = $ constant, which again implies that the Euler-Lagrange equation is satisfied.)

3. $F = F(x, y)$
   In this case, the Euler-Lagrange equation reduces to $\frac{\partial F}{\partial y}(x, y) = 0$. This is an algebraic equation; not an ODE.

An important class of functionals that is often encountered is

$$J[y] = \int_a^b g(x, y)\sqrt{1 + y'^2}dx. \tag{13}$$

Using the Euler-Lagrange equation (2) we obtain (after several lines of careful differentiation and algebraic manipulations)

$$g_y - yg_x - \frac{y''}{1 + y'^2}g = 0. \tag{14}$$

For example, if $g(x, y) \equiv 1$, then we immediately obtain $y'' = 0$, which leads to $y = ax + b$ for $a, b \in \mathbb{R}$. This is the solution to the length functional variational problem on $\mathbb{R}^2$.

Exercise 2.1.

(i) Find the extremum of $J[y] = \int_1^2 \frac{1}{x}\sqrt{1 + y'^2}\,dx$ such that $y(1) = 0$ and $y(2) = 1$.

(ii) Find the extremum of $J[y] = \int_{x_1}^{x_2} 2\pi y \sqrt{1 + y'^2}\,dx$ subject to $y(x_1) = y_1$ and $y(x_2) = y_2$.

(iii) Find the extremum of $J[y] = \int_{x_1}^{x_2}(x - y)^2 dx$.

3 Existence and Uniqueness of Solutions to Variational Problems

The variational problems are usually posed as boundary value problems, where the solution is required to exist over the entire domain. Unlike the initial value ODE problems, there are no general results for the existence of uniqueness of solutions for boundary value problems. Moreover, the existence and uniqueness theorems for IVPs only discuss the solution in a neighbourhood of the initial point, and this might not extend to cover the entire domain.

A certain class of variational (boundary value) problems, however, does have an associated existence and uniqueness theorem.

Theorem 3.1 (Bernstein)**.** *Consider the ODE $y'' = G(x, y, y')$. If the functions F, F_y and $F_{y'}$ are continuous at every finite point (x, y) for any finite y', and if a constant $k > 0$ and function*

$$\alpha = \alpha(x, y) \geq 0, \qquad \beta = \beta(x, y) \geq 0$$

(which are bounder in every finite region of the plane) can be found such that

$$F_y(x, y, y') > k, \qquad |F(x, y, y')| \leq \alpha y'^2 + \beta$$

then one and only one integral curve of the stated ODE passes through any two point (a, A) and (b, B) where $a \neq b$.

4 Multiple Independent Variables

Suppose the functional J takes as its input a function of multiple variables $z = z(x, y)$. Here, x and y are both independent variables, and $J = J[z]$. (For simplicity we restrict to two independent variables. All of the following results can easily be extended to more independent variables.) In that case, we can write

$$J[z] = \int_R F(x, y, z(x, y), z_x(x, y), z_y(x, y))\,dx\,dy. \tag{15}$$

Here, R is some suitable closed region in the plane. The goal now is to find a function $z = z(x, y)$ such that $z \in C^2$ and z satisfies the prescribed boundary conditions on ∂R, and that it extremises the functional $J[z]$.

Lemma 4.1. *If $\alpha(x, y)$ is a fixed function which is continuous in a closed region R, and if the integral*

$$\int_R \alpha(x, y)h(x, y)\,dx\,dy \tag{16}$$

vanishes for every function h that is twice continuously differentiable on R and equals zero in the boundary ∂R, then $\alpha(x, y) \equiv 0$.

The corresponding Euler-Lagrange equation is

$$\frac{\partial F}{\partial z} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial z_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial z_y}\right) = 0 \tag{17}$$

Example 4.2. Consider the functional

$$J[z] = \int_R \left(z_x^2 + z_y^2 + 2zf(x,y)\right) dx\,dy. \tag{18}$$

Its Euler-Lagrange equation is the Poisson equation

$$z_{xx} + z_{yy} = f(x,y) \tag{19}$$

Example 4.3. Finding the surface $z = z(x,y)$ of least area, bounded by a given contour can be represented as a variational problem with the functional

$$J[z] = \int_R \sqrt{1 + z_x^2 + z_y^2}\,dx\,dy. \tag{20}$$

The Euler-Lagrange equation for this functional becomes

$$r(1 + q^2) - 2spq + t(1 + p^2) = 0 \tag{21}$$

where $p = z_x$, $q = z_y$, $r = z_{xx}$, $s = z_{xy}$ and $t = z_{yy}$.

This is a complicated non-linear PDE. However, a result from differential geometry states that for a surface given by $z = z(x,y)$, the mean curvature M takes the form

$$M = \frac{r(1 + q^2) - 2spq + t(1 + p^2)}{\sqrt{1 + p^2 + q^2}} \tag{22}$$

Therefore, the required surfaces are precisely those whose mean curvature vanishes. Surfaces whose mean curvature is zero are called minimal surfaces.

5 Case of Variable End Point

Suppose we need to find the function $y = y(x)$ that extremises the functional

$$J[y] = \int_{x_1}^{x_2} F(x, y, y')\,dx \tag{23}$$

subject to the the condition $y(x_1) = y_1$, but whose second endpoint $y(x_2)$ is not fixed.

In this case the variation δJ with $y \to y + h$ requires $h(x_1) = 0$ but $h(x_2)$ is not fixed. From the usual procedure of Taylor expansion we obtain

$$\delta J = \int_{x_1}^{x_2} h(x)\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right) dx \; - h(x_1)\left.\frac{\partial F}{\partial y'}\right|_{x_1} - h(x_2)\left.\frac{\partial F}{\partial y'}\right|_{x_2} \tag{24}$$

$$= \int_{x_1}^{x_2} h(x)\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right) dx \; - h(x_2)\left.\frac{\partial F}{\partial y'}\right|_{x_2} \tag{25}$$

This variation needs to be zero for all suitable h. In particular, it must vanish for the case where $h(x_2) = 0$. The second boundary term vanishes for this choice of h, and, using the fundamental theorem of calculus of variations, we get

$$\delta J = \int_{x_1}^{x_2} h(x)\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right) dx = 0 \implies \frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0. \tag{26}$$

This shows that if y is a critical point of J then it must satisfy the Euler-Lagrange equation. However, note that this Euler-Lagrange equation is independent of h. Therefore, it must remain zero even when a different choice of h is selected. In particular, if we revert to the case that h is arbitrary (subject to $h(x_1) = 0$) then the variation δJ reduces to

$$\delta J = \int_{x_1}^{x_2} h(x)\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right) dx \; - h(x_2)\left.\frac{\partial F}{\partial y'}\right|_{x_2} = -h(x_2)\left.\frac{\partial F}{\partial y'}\right|_{x_2}. \tag{27}$$

Therefore, $\delta J = 0$ gives

$$\left.\frac{\partial F}{\partial y'}\right|_{x_2} = 0. \tag{28}$$

Overall, a function $y \in C^2[x_1, x_2]$ is a critical point of the given functional $J[y]$ subject to the condition $y(x_1) = y_1$ only if it satisfies the *system* of equations

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 \quad \text{and} \quad \frac{\partial F}{\partial y'}(x_2, y(x_2), y'(x_2)) = 0. \tag{29}$$

Similarly, if the other point is fixed (while the first varies) then the condition becomes

$$\frac{\partial F}{\partial y'}(x_1, y(x_1), y'(x_1)) = 0. \tag{30}$$

Likewise, if both end points are variable, then the resulting set of equations is

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0,$$
$$\frac{\partial F}{\partial y'}(x_1, y(x_1), y'(x_1)) = 0 \quad \text{and} \quad \frac{\partial F}{\partial y'}(x_2, y(x_2), y'(x_2)) = 0. \tag{31}$$

Example 5.1. Consider a massive particle in the xy-plane. It slide down under the influence of gravity. Find the trajectory $y = y(x)$ such that this particle reaches the vertical line $x = b$ in the shortest time under the influence of gravity. Without loss of generality, we take the starting point to be the origin $(0,0)$ and assume that $b > 0$.

This corresponds to minimising the functional

$$T[y] = \int_a^b \sqrt{\frac{1 + y'^2}{2gy}}\, dx. \tag{32}$$

subject to $y(0) = 0$ (and $y(b)$ is not fixed).

This functional is independent of x so the associated first integral gives

$$F - y'\frac{\partial F}{\partial y'} = \text{constant} \implies \frac{1}{\sqrt{y(1 + y'^2)}} = k \implies y'^2 = \frac{k}{y} - 1 \tag{33}$$

where k is an arbitrary constant.

We solve this by introducing a parameter θ, such that $y' = \cot\left(\frac{\theta}{2}\right)$. Consequently,

$$y'^2 = \frac{k}{y} - 1 \implies 1 + \cot^2 \frac{\theta}{2} = \frac{k}{y} \implies y = k \sin^2 \frac{\theta}{2} \implies y = \frac{k}{2}(1 - \cos\theta) \quad (34)$$

and

$$\frac{dy}{dx} = \cot\frac{\theta}{2} \implies dx = \tan\frac{\theta}{2}\, dy \implies x = k \int \tan\frac{\theta}{2} \sin\frac{\theta}{2} \cos\frac{\theta}{2}\, d\theta. \quad (35)$$

Here, we used $y = k \sin^2 \frac{\theta}{2}$ to get $dy = k \sin\frac{\theta}{2}\cos\frac{\theta}{2}\, d\theta$. This integral can be evaluated to obtain the corresponding parametric expression for x by using elementary trigonometric relations

$$x = k \int \sin^2\left(\frac{\theta}{2}\right) d\theta = \frac{k}{2} \int (1 - \cos\theta)\, d\theta = c + \frac{k}{2}(\theta - \sin\theta) \quad (36)$$

Overall,

$$x = c + \frac{k}{2}(\theta - \sin\theta), \quad y = \frac{k}{2}(1 - \cos\theta) \quad (37)$$

At one end point we have $y = 0$ when $x = 0$. While the variable end point condition

$$\left.\frac{\partial F}{\partial y'}\right|_{x=b} = 0 \implies \left.\frac{y'}{\sqrt{2g}\sqrt{y(1 + y'^2)}}\right|_{x=b} = 0 \quad (38)$$

gives $y' = 0$ when $x = b$.

Now, $y = 0$ when $x = 0$ means that we need to simultaneously solve

$$c + \frac{k}{2}(\theta - \sin\theta) = 0 \quad \text{and} \quad \frac{k}{2}(1 - \cos\theta) = 0 \quad (39)$$

The second equation gives $\theta = 0$, because $k = 0$ would result in the trivial solution $y \equiv 0$. Using this value of θ in the first equation gives $c = 0$.

Next, $y' = 0$ at $x = b$ corresponds to the system

$$\frac{k}{2}(\theta - \sin\theta) = b \quad \text{and} \quad \cot\frac{\theta}{2} = 0. \quad (40)$$

Again, the second equation gives $\theta = n\pi$, but we can't take $\theta = 0$ otherwise the first equation will reduce to the contradiction $0 = b > 0$. Taking $\theta = \pi$ gives $k = 2b/\pi$. Therefore, the solution to this variational problem is a cycloid

$$x = \frac{b}{\pi}(\theta - \sin\theta), \quad y = \frac{b}{\pi}(1 - \cos\theta). \quad (41)$$

6   Variational Derivative

Let f be a real-valued function of n variables. We study its differential and partial derivatives in real analysis. Analogous concepts also exist for real-valued functional. Among these, the concept of partial derivatives is generalised to the concept of a variational derivative. The procedure is carried out as follows:

We approximate the continuous function $y = y(x)$ as a polygonal line

$$(x_0, y_0), \ldots, (x_{n+1}, y_{n+1}). \tag{42}$$

This discretises the functional $J[y] = \int_a^b F(x, y, y') \, dx$ subject to $y(a) = A$ and $y(b) = B$ into a function of n variables

$$J(y_1, \ldots, y_n) = \sum_{i=0}^{n} F\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right) \Delta x. \tag{43}$$

(This is not a function of $n + 2$ variables because $y_0 = A$ and $y_{n+1} = B$ are fixed.)

We can now calculate $\partial J / \partial y_k$ as the usual partial derivative to obtain

$$\frac{\partial J}{\partial y_k}(y_1, \ldots, y_n) = \frac{\partial F}{\partial y}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \Delta x$$
$$+ \frac{\partial F}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) \Delta x - \frac{\partial F}{\partial y'}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \Delta x.$$

Dividing by Δx and taking the limit $\Delta x \to 0$ reduces this expression to

$$\frac{\partial J}{\partial y_k \Delta x} = \frac{\partial F}{\partial y}(x, y, y') - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}(x, y, y')\right), \tag{44}$$

where

$$\frac{\delta J}{\delta y} \equiv \lim_{\Delta x \to 0} \frac{\partial J}{\partial y_k \Delta x} \tag{45}$$

is called the variational derivative of the functional $J[y]$. We note that the variational derivative is precisely the LHS of the Euler-Lagrange equation.

7 Invariance of the Euler-Lagrange Equation

So far we have studied the variation of functional

$$J[y] = \int_a^b F(x, y, y') \, dx \tag{46}$$

using the Cartesian variables $y = y(x)$, leading to the Euler-Lagrange equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0. \tag{47}$$

Now we consider a change of variables $x = x(u, v)$, $y = y(u, v)$ with a non-vanishing Jacobian, and the assumption that the required solution curves have the form $v = v(u)$. Simple algebra shows that this transform the function J into another functional

$$J_1[v] = \int_\alpha^\beta F\left(x(u, v), y(u, v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v') \, du = \int_\alpha^\beta F_1(u, v, v') \, du \tag{48}$$

with

$$F_1(u, v, v') = F\left(x(u, v), y(u, v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v'). \tag{49}$$

It can be shown that if $y = y(x)$ satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0, \tag{50}$$

then $v = v(u)$ satisfies the Euler-Lagrange equation

$$\frac{\partial F_1}{\partial v} - \frac{d}{du}\left(\frac{\partial F_1}{\partial v'}\right) = 0, \tag{51}$$

corresponding to J_1. We note that the form of the Euler-Lagrange equation remains unchanged even when we transform to new variables; the form of the functional and the integrand does not (in general) remain the same. In other words, whether or not a curve is an extremal is a property that is independent of the choice of coordinates.

This result is known as the invariance of the Euler-Lagrange equation.

One way to prove this result is through the use of the variational derivative.

Example 7.1. Consider the functional

$$J[r] = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + r'^2}\, d\theta \tag{52}$$

where $r = r(\theta)$. Its associated Euler-Lagrange equation is

$$\frac{r}{\sqrt{r^2 + r'^2}} - \frac{d}{d\theta}\left(\frac{r'}{\sqrt{r^2 + r'^2}}\right) = 0. \tag{53}$$

On the other hand, if we transform to new variables $x = r\cos\theta$ and $y = r\sin\theta$ and seek a solution of the form $y = y(x)$, then the functional becomes

$$\int_{\alpha}^{\beta} \sqrt{1 + y'^2}\, dx, \tag{54}$$

leading to the Euler-Lagrange equation $y'' = 0$. This is significantly easier to solve. The general solution is $y = ax + b$. Or in the original coordinates

$$r\sin\theta = ar\cos\theta + b. \tag{55}$$

8 Generalisation to Many Dependent Variables

Suppose we are given a functional of the form

$$J[y_1, \ldots, y_n] = \int_a^b F(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx \tag{56}$$

where $y_i \in C^2[a, b]$ for all $i = 1, \ldots, n$ and $y_i(a) = A_i$ and $y_i(b) = B_i$. Previously we only had a single dependent variable $y = y(x)$ and obtained a single second-order ODE as the Euler-Lagrange equation corresponding to $\delta J/\delta y = 0$. In the current generalisation, this necessary conditions for the extremal will become $\delta J/\delta y_i = 0$ for each i, and we will obtain a system of n second-order Euler-Lagrange equations

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right) = 0, \quad i = 1, \ldots, n. \tag{57}$$

This conclusion follows from varying the functional $J[y_1, \ldots, y_n]$ as follows. Consider n independent arbitrary functions $h_i \in C^2[a,b]$ such that $h_i(a) = h_i(b) = 0$ for all $i = 1, \ldots, n$. Let y_i denote the extremal of J. Then, varying it as $y_i \to y_i + h_i$ gives

$$\delta J = \int_a^b F\left(x, y_1 + h_1, \ldots, y_n + h_n, y_1' + h_1', \ldots, y_n' + h_n'\right) dx \tag{58}$$

$$- \int_a^b F\left(x, y_1, \ldots, y_n, y_1', \ldots, y_n'\right) dx \tag{59}$$

$$= \int_a^b \sum_{i=1}^n \left(h_i \frac{\partial F}{\partial y_i} + h_i' \frac{\partial F}{\partial y_i'} \right) dx \tag{60}$$

$$= h_i(x) F(x, y_i, y_i') \Big|_a^b + \int_a^b \sum_{i=1}^n h_i \left(\frac{\partial F}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_i'} \right) \right) dx \tag{61}$$

$$= \int_a^b \sum_{i=1}^n h_i \left(\frac{\partial F}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_i'} \right) \right) dx. \tag{62}$$

Now, since h_i are arbitrary (except the regularity and boundary conditions), we can consider the case where all h_i are zero except h_k. This reduces the variation of J to

$$\delta J = \int_a^b h_k \left(\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_k'} \right) \right) dx. \tag{63}$$

Requiring $\delta J = 0$ and using the fundamental theorem of calculus of variations then gives

$$\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_k'} \right) = 0. \tag{64}$$

The same process can be repeated for each h_i to get one Euler-Lagrange equation for each dependent variable.

9 Geodesics

Functionals with multiple dependent variables are particularly suitable for stating the problem of finding geodesics on a surface. A geodesic between two points P and Q on a surface $\mathbf{r} = \mathbf{r}(u, v)$ is a curve of shortest length connecting these two points. In finding a geodesic we, therefore, minimise the length functional $L = \int_\alpha^\beta ds$ where ds is the length element for the given surface. From differential geometry, we know that

$$ds^2 = E(u, v)\, du^2 + 2F(u, v)\, dudv + G(u, v)\, dv^2, \tag{65}$$

with $E = \mathbf{r}_u \cdot \mathbf{r}_u$, $F = \mathbf{r}_u \cdot \mathbf{r}_v$ and $G = \mathbf{r}_v \cdot \mathbf{r}_v$. Equation (65) is also called the first fundamental form for the surface $\mathbf{r} = \mathbf{r}(u, v)$. Consequently, finding geodesics on the given surface is equivalent to finding the extremals $u = u(t), v = v(t)$ of

$$J[u, v] = \int_\alpha^\beta \sqrt{E(u, v)u'^2 + 2F(u, v)u'v' + G(u, v)v'^2}\, dt. \tag{66}$$

The resulting Euler-Lagrange equations are

$$\frac{E_u u'^2 + 2F_u u'v' + G_u v'^2}{2Q} - \frac{d}{dt}\left(\frac{Eu' + Fv'}{Q} \right) = 0, \tag{67}$$

$$\frac{E_v u'^2 + 2F_v u'v' + G_v v'^2}{2Q} - \frac{d}{dt}\left(\frac{Fu' + Gv'}{Q} \right) = 0, \tag{68}$$

where $Q = \sqrt{E(u,v)u'^2 + 2F(u,v)u'v' + G(u,v)v'^2}$.

Example 9.1. To find the equation of a general geodesic on a circular cylinder of radius a and centred on the z-axis, we start by parametrising the surface as

$$\mathbf{r}(\varphi, z) = (a \cos \varphi, a \sin \varphi, z). \tag{69}$$

Then, $E = \mathbf{r}_\varphi \cdot \mathbf{r}_\varphi = a^2$, $F = \mathbf{r}_\varphi \cdot \mathbf{r}_z = 0$ and $G = \mathbf{r}_z \cdot \mathbf{r}_z = 1$. So, the length functional becomes

$$J[\varphi, z] = \int_\alpha^\beta \sqrt{a^2 \dot\varphi^2 + \dot z^2} \, dt, \tag{70}$$

leading to a system of two Euler-Lagrange equations (first integrals)

$$\frac{d\varphi}{dt} = k_1 \sqrt{a^2 \dot\varphi^2 + \dot z^2}, \qquad \frac{dz}{dt} = k_2 \sqrt{a^2 \dot\varphi^2 + \dot z^2}. \tag{71}$$

Here, k_1, k_2 are arbitrary constants. We can solve these by eliminating the parameter t and giving z as a function of φ,

$$\frac{dz/dt}{d\varphi/dt} = \frac{k_2 \sqrt{a^2 \dot\varphi^2 + \dot z^2}}{k_1 \sqrt{a^2 \dot\varphi^2 + \dot z^2}} \implies \frac{dz}{d\varphi} = c_1 \implies z = c_1 \varphi + c_2. \tag{72}$$

The arbitrary constant c_1, c_2 are determined by the points P and Q.

Example 9.2. Suppose that the 3-dimensional space is filled with an optically inhomogeneous medium, such that the velocity of propagation of light at each point is some function $v(x, y, z)$ of the coordinates of the points.

According to Fermat's principle, light travels from point A to point B along the trajectory which takes the least time. Suppose the curve joining A and B is given by the equations $y = y(x)$ and $z = z(x)$. Then, the corresponding time functional is

$$T[y, z] = \int_a^b \frac{\sqrt{1 + y'^2 + z'^2}}{v(x, y, z)} \, dx. \tag{73}$$

The associated Euler-Lagrange equations are

$$\frac{\sqrt{1 + y'^2 + z'^2}}{v^2} v_y + \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2 + z'^2}} \right) = 0, \tag{74}$$

$$\frac{\sqrt{1 + y'^2 + z'^2}}{v^2} v_z + \frac{d}{dx} \left(\frac{z'}{\sqrt{1 + y'^2 + z'^2}} \right) = 0. \tag{75}$$

10 Equivalent Functionals

Let $\varphi = \varphi(x, y_1, \ldots, y_n)$ be a twice continuously differentiable function and consider the function

$$\psi(x, y_1, \ldots, y_n, y_1', \ldots, y_n') = \frac{d\varphi}{dx} \equiv \frac{\partial\varphi}{\partial x} + \sum_{i=1}^n y_i' \frac{\partial\varphi}{\partial y_i}. \tag{76}$$

It is straightforward to check by direct calculation that

$$\frac{\partial\psi}{\partial y_k} - \frac{d}{dx}\left(\frac{\partial\psi}{\partial y_k'}\right) \equiv 0 \tag{77}$$

identically. Consequently, the Euler-Lagrange equations obtained from

$$J = \int_a^b F(x, y_1, \ldots, y_n, y_1', \ldots, y_n') \, dx \tag{78}$$

and

$$J_1 = \int_a^b \left\{ F(x, y_1, \ldots, y_n, y_1', \ldots, y_n') + \psi(x, y_1, \ldots, y_n, y_1', \ldots, y_n') \right\} dx \tag{79}$$

are the same, namely

$$\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_k'} \right) = 0. \tag{80}$$

This can also be seen by noting that J and J_1 differ only by a constant, since ψ is a total derivative of $\varphi(x, y_1, \ldots, y_n)$ and integrates out

$$J_1 = \int_a^b (F + \psi) \, dx = \int_a^b F \, dx + \int_a^b \frac{d\phi}{dx} \, dx = J + \left(\varphi \big|_a^b \right) = J + \text{constant}. \tag{81}$$

Therefore, different integrands (or likewise, different functionals) can lead to the same Euler-Lagrange equations.

Two functionals are said to be equivalent if they lead to the same Euler-Lagrange equations.

In particular, we note that the functionals $J = \int_a^b F \, dx$ and $J = \int_a^b \left(F + \frac{d\phi}{dx} \right) dx$ are equivalent. Similarly, $J = \int_a^b F \, dx$ and $J = \int_a^b cF \, dx$ are equivalent for $c \in \mathbb{R}$.

11 Functionals Depending on Higher-Order Derivatives

Now we consider the generalisation to functionals of the form

$$J[y] = \int_a^b F(x, y, y', \ldots, y^{(n)}) \, dx \tag{82}$$

with $y(a) = A$, $y(b) = B$ and $y^{(i)}(a) = A_i$, $y^{(i)}(b) = B_i$ for $i = 1, \ldots, n$.

As usual we suppose that y extremises this functional and consider its variation $y \to y + h$ where $h(a) = h(b) = 0$ and $h^{(i)}(a) = h^{(i)}(b) = 0$ for $i = 1, \ldots, n$. Then, using Taylor expansion we obtain

$$\delta J = \int_a^b F\left(x, y + h, y' + h', \ldots, y^{(n)} + h^{(n)}\right) dx - \int_a^b F\left(x, y, y', \ldots, y^{(n)}\right) dx \tag{83}$$

$$= \int_a^b \left(h \frac{\partial F}{\partial y} + h' \frac{\partial F}{\partial y'} + \cdots + h^{(n)} \frac{\partial F}{\partial y^{(n)}} \right) dx \tag{84}$$

$$= \int_a^b h(x) \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial F}{\partial y''} + \cdots + (-1)^n \frac{d^n}{dx^n} \frac{\partial F}{\partial y^{(n)}} \right) dx. \tag{85}$$

Here, we have used integration by parts repeatedly in the last step and the boundary terms vanish due to the conditions on h and its derivatives. This last expression is

true for all arbitrary functions h (subject to suitable boundary conditions), therefore, the fundamental theorem of calculus of variations gives

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial F}{\partial y''} + \cdots + (-1)^n \frac{d^n}{dx^n}\frac{\partial F}{\partial y^{(n)}} = 0 \tag{86}$$

when $\delta J = 0$. In other words, we have proved the following theorem.

Theorem 11.1. *A necessary condition for $y \in C^{2n}[a, b]$ to be an extremal of*

$$J[y] = \int_a^b F\left(x, y, y^{(1)}, \ldots, y^{(n)}\right) dx \tag{87}$$

subject to $y(a) = A$, $y(b) = B$, $y^{(i)}(a) = A_i$ and $y^{(i)}(b) = B_i$, for $i = 1, \ldots, n$, is given by the (generalised) Euler-Lagrange equation

$$\frac{\partial F}{\partial y} + \sum_{k=1}^{n} (-1)^k \frac{d^k}{dx^k}\left(\frac{\partial F}{\partial y^{(k)}}\right) = 0. \tag{88}$$

Equation (88) is in general an ODE of order $2n$. The regularity conditions on y can be reduced to $y \in C^n$ because it can be shown that under suitable conditions such a function is $y \in C^{(2n)}$.

Here again we can obtain some first integrals when the integrand $F\left(x, y, y^{(1)}, \ldots, y^{(n)}\right)$ takes a specific form.

1. $F = F\left(x, y^{(1)}, \ldots, y^{(n)}\right)$ leads to

$$\frac{\partial F}{\partial y'} + \sum_{k=2}^{n} (-1)^{k-1} \frac{d^{k-1}}{dx^{k-1}}\left(\frac{\partial F}{\partial y^{(k)}}\right) = \text{constant}, \tag{89}$$

which can also be written as

$$\frac{\partial F}{\partial y'} + \sum_{k=1}^{n-1} (-1)^k \frac{d^k}{dx^k}\left(\frac{\partial F}{\partial y^{(k+1)}}\right) = \text{constant}. \tag{90}$$

2. If $F = F\left(x, y^{(s)}, \ldots, y^{(n)}\right)$, i.e. y and its first $s-1$ derivatives don't appear explicitly in the integrand, then the previous result can be extended to obtain

$$\frac{\partial F}{\partial y^{(s)}} + \sum_{k=1}^{n-s} (-1)^k \frac{d^k}{dx^k}\left(\frac{\partial F}{\partial y^{(k+s)}}\right) = \text{constant}. \tag{91}$$

3. If $F = F(y, y', y'')$ then we find that

$$F - y'\left(\frac{\partial F}{\partial y'} - \frac{d}{dx}\left(\frac{\partial F}{\partial y''}\right)\right) - y''\frac{\partial F}{\partial y''} = \text{constant}. \tag{92}$$

(This is the generalisation of the Hamiltonian, H.) For integrands with higher order derivatives, this expression becomes more complicated. In general, when $F = F\left(y, y^{(1)}, \ldots, y^{(n)}\right)$, we get

$$F - \sum_{k=1}^{n} y^{(k)} p_k = \text{constant}, \tag{93}$$

where

$$p_k = \frac{\partial F}{\partial y^{(k)}} + \sum_{j=k}^{n} (-1)^j \frac{d^j}{dx^j}\left(\frac{\partial F}{\partial y^{(j+1)}}\right) \tag{94}$$

12 Variational Problems with Constraints

12.1 Isoperimetric Constraints

Constraints that are given as an integral equation are called isoperimetric or integral constraints. These take the form $K[y] = \int_a^b G(x, y, y') = \ell$. The terminology follows from a famous problem where we find a suitable curve of fixed perimeter, enclosing the largest area.

Theorem 12.1. *Let $y = y(x)$ be an extremal of $J[y] = \int_a^b F(x, y, y')\, dx$ subject to the boundary conditions $y(a) = A$ and $y(b) = B$, and the constraint*

$$K[y] = \int_a^b G(x, y, y') = \ell\, dx. \tag{95}$$

If y is not an extremal of $K[y]$ then there is a constant λ such that y is an extremal of

$$I[y] = \int_a^b (F + \lambda G)\, dx. \tag{96}$$

That is, $y = y(x)$ satisfies the following system of equations

$$\frac{\partial}{\partial y}(F + \lambda G) + \frac{d}{dx}\left(\frac{\partial}{\partial y'}(F + \lambda G)\right) = 0, \tag{97}$$

$$\int_a^b G(x, y, y') = \ell. \tag{98}$$

Example 12.2. Among all the curve of length ℓ in the upper half plane passing through the points $(-a, 0)$ and $(a, 0)$, find the one which together with the interval $[-a, a]$ encloses the largest area. Here, $a > 0$. In other words, maximise $A[y]$ subject to $L[y]$, where

$$A[y] = \int_a^b y\, dx \quad \text{and} \quad L[y] = \int_a^b \sqrt{1 + y'^2}\, dx = \ell. \tag{99}$$

We impose the additional restriction $2a < \ell < a\pi$ on the length of the curve; the shortest distance between the end-points in $2a$.

This problem corresponds to solving the following Euler-Lagrange equation

$$1 + \lambda \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + y'^2}}\right) = 0. \tag{100}$$

Here, $F = y$ and $G = \sqrt{1 + y'^2}$. After a few more lines of algebra we get

$$y'^2 = \frac{(x + c)^2}{\lambda^2 - (x + c)^2}, \tag{101}$$

where c is an arbitrary constant. Upon integration this becomes

$$y = \sqrt{\lambda^2 - (x + c)^2} + k \implies (x + c)^2 + (y - k)^2 = \lambda^2 \tag{102}$$

with another arbitrary constant k. The boundary conditions $y(-a) = y(a) = 0$ give

$$k^2 + (c + a)^2 = \lambda^2 \quad \text{and} \quad k^2 + (c - a)^2 = \lambda^2. \tag{103}$$

By subtracting the second equation from the first we obtain $4ac = 0$, which leads to $c = 0$. Moreover, we note that $\lambda^2 = k^2 + a^2$. The reduced solution curve is an arc of a circle of radius $\sqrt{k^2 + a^2}$ centred at $(0, k)$

$$x^2 + (y - k)^2 = k^2 + a^2. \tag{104}$$

To eliminate k from this equation, we need to use the constraint $\int_a^b \sqrt{1 + y'^2}\, dx = \ell$, which leads to a complicated expression,

$$k = a \cot\left(\frac{\ell}{2\sqrt{a^2 + k^2}}\right), \tag{105}$$

with no closed form solution.

In the special case, $\ell = \pi a$, we have $k = 0$ (geometrically).

We can generalise theorem (12.1) to cover the case of several dependent variables and several integral constraints.

Theorem 12.3. *Let $y_1, \ldots, y_n$ be the extremals of*

$$J[y_1, \ldots, y_n] = \int_a^b F(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx \tag{106}$$

subject to the boundary conditions $y_i(a) = A_i$ and $y_i(b) = B_i$ with $i = 1, \ldots, n$ and integral constrains

$$K_j[y_1, \ldots, y_n] = \int_a^b G_j(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx = \ell_j, \quad j = 1, \ldots, k. \tag{107}$$

If $y_1, \ldots, y_n$ are not extremals of $K_j[y_1, \ldots, y_n]$ for all j, then there are constants λ_j such that $y_1, \ldots, y_n$ are the extremals of

$$I[y_1, \ldots, y_n] = \int_a^b \Phi(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx \tag{108}$$

with $\Phi = F + \sum_{j=1}^k \lambda_j G_j$.

In other words, $y_1, \ldots, y_n$ satisfy the system of Euler-Lagrange equations

$$\frac{\partial \Phi}{\partial y_k} - \frac{d}{dx}\left(\frac{\partial \Phi}{\partial y_k'}\right) = 0 \tag{109}$$

along with the boundary conditions and the constraint equations.

12.2 Holonomic Constraints

The constraints of the form $g(x, y_1, \ldots, y_n) = 0$ are called holonomic.

Theorem 12.4. *Let $y_1, \ldots, y_n$ be the extremals of*

$$J[y_1, \ldots, y_n] = \int_a^b F(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx \tag{110}$$

subject to the boundary conditions $y_i(a) = A_i$ and $y_i(b) = B_i$ with $i = 1, \ldots, n$ and holonomic constrains

$$g_j(x, y_1, \ldots, y_n) = 0, \quad j = 1, \ldots, k. \tag{111}$$

If $\partial g_j / \partial y_i$ do not vanish simultaneously for all i for points satisfying $g_j(x, y_1, \ldots, y_n) = 0$, for all j, then there exist functions $\lambda_j(x)$ such that $y_1, \ldots, y_n$ are the extremals of

$$I[y_1, \ldots, y_n] = \int_a^b \Phi(x, y_1, \ldots, y_n, y_1', \ldots, y_n') \, dx \tag{112}$$

with $\Phi = F + \sum_{j=1}^n \lambda_j(x) g_j$.

For the special case $k = 1$ (a single constraint) and $n = 2$ (two dependent variables), we get the following expression for the new functional,

$$I[y_1, y_2] = \int_a^b \left\{ F(x, y_1, y_2, y_1', y_2') + \lambda(x) g(x, y_1, y_2, y_1', y_2') \right\} dx. \tag{113}$$

The associated Euler-Lagrange equations are then

$$\frac{\partial F}{\partial y_1} + \lambda(x) \frac{\partial g}{\partial y_1} - \frac{d}{dx}\left(\frac{\partial F}{\partial y_1'} \right) = 0 \tag{114}$$

$$\frac{\partial F}{\partial y_2} + \lambda(x) \frac{\partial g}{\partial y_2} - \frac{d}{dx}\left(\frac{\partial F}{\partial y_2'} \right) = 0 \tag{115}$$