Calculus of Variations Chapter Notes

Adapted from *Calculus of Variations by I.M. Gelfand and S.V. Fomin*

Navid Shamszadeh

Contents

1 Preliminaries

Lemma 1.1. *If $\alpha(x)$ is continuous in $[a,b]$ and if*

$$\int_a^b \alpha(x)h(x)dx = 0$$

for every $h(x) \in \mathcal{C}[a,b]$ such that $h(a) = h(b) = 0$, then $\alpha(x) = 0$ for every $x \in [a,b]$

Proof. Suppose $\alpha(x)$ is nonzero, say positive at some point in $[a,b]$. Then since $\alpha(x)$ is continuous in $[a,b]$, $\alpha(x)$ is positive in some interval $[x_1, x_2] \subset [a,b]$. If we set

$$h(x) = (x - x_1)(x - x_2)$$

for $x \in [x_1, x_2]$ and 0 everywhere else, then $h(x)$ satisfies the conditions for the lemma. However,

$$\int_a^b \alpha(x)h(x)dx = \int_a^b \alpha(x)(x - x_1)(x - x_2)dx < 0$$

Since the integrand is negative (except at x_1 and x_2), we are done. □

Lemma 1.2. *If $\alpha(x)$ is continuous in $[a,b]$ and if*

$$\int_a^b \alpha(x)h'(x)dx = 0$$

for every $h(x) \in \mathcal{C}^1[a,b]$ such that $h(a) = h(b) = 0$, then $\alpha(x) = c$ for all $x \in [a,b]$, where c is a constant.

Proof. Let c be the constant defined by the condition

$$\int_a^b [\alpha(x) - c]dx = 0$$

that is,

$$\int_a^b \alpha(x)dx - c\int_a^b dx = 0$$

$$\iff \int_a^b \alpha(x)dx = c\int_a^b dx \iff \frac{1}{b-a}\int_a^b \alpha(x)dx = c$$

and let

$$h(x) = \int_a^x [\alpha(\xi) - c]d\xi$$

so that $h(x)$ automatically belongs to $\mathcal{C}^1[a,b]$ and satisfies the conditions $h(a) = h(b) = 0$. Then by the fundamental theorem of calculus,

$$h'(x) = \alpha(x) - c$$

and on one hand,

$$\int_a^b [\alpha(x) - c]h'(x)dx = \int_a^b \alpha(x)h'(x)dx - c\int_a^b h'(x)dx = 0 - c[h(b) - h(a)] = 0$$

and on the other hand,

$$\int_a^b [\alpha(x) - c]h'(x)dx = \int_a^b [\alpha(x) - c]^2 dx$$

Since $[\alpha(x) - c]^2$ is nonnegative for all values of x,

$$\int_a^b [\alpha(x) - c]^2 dx = 0 \implies \alpha(x) - c = 0 \implies \alpha(x) = c$$

□

Lemma 1.3. *If $\alpha(x)$ is continuous in $[a, b]$ and if*

$$\int_a^b \alpha(x)h''(x)dx = 0$$

for all $h(x) \in \mathcal{C}^2[a, b]$ with $h(a) = h(b) = 0$ and $h'(a) = h'(b) = 0$, then $\alpha(x) = c_0 + c_1 x$ where c_0, c_1 are constants.

Proof. Let c_0, c_1 be defined by the conditions

$$\int_a^b [\alpha(x) - c_0 - c_1 x]dx = 0$$

$$\int_a^b dx \int_a^x [\alpha(\xi) - c_0 - c_1 \xi]d\xi = 0$$

and let

$$h(x) = \int_a^x d\xi \int_a^\xi [\alpha(t) - c_0 - c_1 t]dt$$

so that $h(x)$ automatically belongs to $\mathcal{C}^2[a, b]$ and satisfies $h(a) = h(b) = h'(a) = h'(b) = 0$. Then on one hand

$$\int_a^b [\alpha(x) - c_0 - c_1 x]h''(x)dx$$

$$= \int_a^b \alpha(x)h''(x)dx - c_0 \int_a^b h''(x)dx - c_1 \int_a^b xh''(x)dx$$

$$= \int_a^b \alpha(x)h''(x)dx - c_0[h'(b) - h'(a)] - c_1 \int_a^b xh''(x)dx$$

$$= -c_1 \int_a^b xh''(x)dx = -c_1 \left(xh'(x)\right)_a^b + c_1 \int_a^b h'(x)dx$$

$$= c_1 \int_a^b h'(x)dx = c_1[h(b) - h(a)] = 0$$

while on the other hand,

$$\int_a^b [\alpha(x) - c_0 - c_1 x]h''(x)dx = \int_a^b [\alpha(x) - c_0 - c1x]^2 dx = 0$$

It follows that $\alpha(x) - c_0 - c_1 x = 0$, that is, $\alpha(x) = c_0 + c_1 x$ □

1.1 Question

I'm not sure how

$$\int_a^b [\alpha(x) - c_0 - c_1 x] h''(x) dx = \int_a^b [\alpha(x) - c_0 - c1x]^2 dx$$

I tried differentiating $h(x)$ and also integrating $h''(x)$ to reach $h(x)$ but I'm not sure how to proceed.

Lemma 1.4. *If $\alpha(x)$ and $\beta(x)$ are continuous in $[a, b]$ and if*

$$\int_a^b \alpha(x)h(x) + \beta(x)h'(x)dx = 0 \tag{1}$$

for every $h(x) \in \mathcal{C}^1[a, b]$ such that $h(a) = h(b) = 0$, then $\beta(x)$ is differentiable and $\beta'(x) = \alpha(x)$ for all $x \in [a, b]$.

Proof. Setting

$$A(x) := \int_a^x \alpha(\xi)d\xi$$

We integrate by parts:

$$\int_a^b \alpha(x)h(x)dx = \left(h(x)\int_a^x \alpha(\xi)d\xi\right)_a^b - \int_a^b \left(\int_a^x \alpha(\xi)d\xi\right) h'(x)dx = -\int_a^b A(x)h'(x)dx$$

We write (1) in terms of $A(x)$:

$$\begin{aligned}\int_a^b \alpha(x)h(x)dx + \int_a^b \beta(x)h'(x)dx &= -\int_a^b A(x)h'(x) + \int_a^b \beta(x)h'(x) \\ &= \int_a^b [\beta(x) - A(x)]\, h'(x)dx \qquad (2)\end{aligned}$$

Now since $\alpha(x)$ is continuous in $[a, b]$, $A(x)$ is continuous in $[a, b]$. And since $\beta(x)$ is continuous in $[a, b]$, $\beta(x) - A(x)$ is continuous in $[a, b]$. So we can apply Lemma 1.2 to (2) which gives $\beta(x) - A(x) = c$ where c is a constant. Hence $\beta(x) = A(x) + c$ and we can differentiate to get $\beta'(x) = \alpha(x)$ for all $x \in [a, b]$. □

Definition 1.5. Let $J[y]$ be a functional defined on some normed linear space. Let

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

be an arbitrary increment of $J[y]$ by a function $h(x)$. Now suppose that

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

where $\varphi[h]$ is a linear functional and $\varepsilon \to 0$ as $||h|| \to 0$. Then we say $J[y]$ is *differentiable,* and the *principal linear part* of the increment $\Delta J[h]$, that is, $\varphi[h]$, which differs from $\Delta J[h]$ by an infinitesimal of order higher than 1 relative to $||h||$, is called the *variation (or differential) of $J[y]$ and is denoted $\delta J[y]$.*

Recall. (See E.I. Gordon, A.G. Kusraev, and S.S. Kutateladze. Infinitesimal Analysis) We call ε an infinitesimal if

$$|\varepsilon| < r \quad \forall r \in \mathbb{R}^+$$

0 is the only real valued infinitesimal by this definition. Furthermore, if ε, δ are infinitesimals, then so are

1. $\varepsilon + \delta$
2. $\varepsilon \cdot \delta$
3. $s \cdot \varepsilon \quad \forall s \in \mathbb{R}$

Theorem 1.6. *The differential of a differentiable functional is unique.*

Proof. We first claim that if $\varphi[h]$ is a linear functional and if

$$\frac{\varphi[h]}{||h||} \to 0$$

as $||h|| \to 0$, then $\varphi[h] \equiv 0$ for all h. To prove this, suppose for contradiction that $\varphi[h_0] \neq 0$ for some $h_0 \neq 0$. Then, setting

$$h_n = \frac{h_0}{n}, \qquad \lambda = \frac{\varphi[h_0]}{||h_0||} \neq 0,$$

It is easy to see that $||h_n|| \to 0$ as $n \to \infty$, but

$$\lim_{n\to\infty} \frac{\varphi[h_n]}{||h_n||} = \lim_{n\to\infty} \frac{\varphi[\frac{h_0}{n}]}{||\frac{h_0}{n}||} = \frac{\varphi[h_0]}{||h_0||} = \lambda \neq 0$$

Since $||h_n|| \to 0$ as $n \to \infty$, $\frac{\varphi[h_n]}{||h_n||}$ must approach 0 by hypothesis; however, by assuption, $\lambda \neq 0$, yielding a contradiction. Therefore $\varphi[h_0] = 0$.
Now, to prove uniqueness, suppose

$$\Delta J[h] = \varphi_1[h] + \varepsilon_1 ||h||$$

$$\Delta J[h] = \varphi_2[h] + \varepsilon_2 ||h||$$

It is easy to see that $\varphi_1[h] - \varphi_2[h] = \varepsilon_2||h|| - \varepsilon_1||h||$. Recall, we call δ an infinitesimal if and only if $\delta < \frac{1}{n}$ for any $n \in \mathbb{N}$. The laws of algebra hold for infinitesimals. Therefore, $\varphi_1[h] - \varphi_2[h] = \varepsilon_2||h|| - \varepsilon_1||h|| = (\varepsilon_2 - \varepsilon_1)||h||$ is an infinitesimal of order higher than 1 with relative to $||h||$. Recall from our definition of differentiable functional, $\varepsilon \to 0$ as $|h|| \to 0$, hence from our first claim, by setting $\varphi[h] := \varphi_1[h] - \varphi_2[h]$ we see that

$$\varphi[h] = (\varepsilon_2 - \varepsilon_1)||h|| \iff \frac{\varphi[h]}{||h||} = \varepsilon_2 - \varepsilon_1$$

So as $||h|| \to 0$, $\varepsilon_2 - \varepsilon_1 \to 0$ and

$$\frac{\varphi[h]}{||h||} \to 0 \implies \varphi[h] = 0 \implies \varphi_1[h] = \varphi_2[h]$$

as desired. $\square$

Recall. Let $F(x_1, \ldots, x_n)$ be a differentiable function of n variables. Then $F(x_1 \ldots, x_n)$ is said to have a (relative) extremum at $(\hat{x_1}, \ldots, \hat{x_n})$ if $\Delta F = F(x_1, \ldots, x_n) - F(\hat{x_1}, \ldots, \hat{x_n})$ has the same sign for all points $(x_1, \ldots, x_n)$ belonging to some neighborhood of $(\hat{x_1}, \ldots, \hat{x_n})$. The extremum is a minimum if $\Delta F \geq 0$ and a maximum if $\Delta F \leq 0$.

Definition 1.7. Similarly, we say that a functional $J[y]$ has a (relative) extremum for $y = \hat{y}$ if $J[y] - J[\hat{y}]$ does not change its sign in some neighborhood of the curve $y = \hat{y}(x)$. We will limit our arguments to functions belonging to $\mathcal{C}^0[a,b]$ and $\mathcal{C}^1[a,b]$.
We say a functional $J[y]$ has a *weak extremum* for $y = \hat{y}$ if there exists an $\varepsilon > 0$ such that the sign of $J[y] - J[\hat{y}]$ is the same for all y contained in $||y - \hat{y}||_1 < \varepsilon$ where

$$|| \cdot ||_1 \colon \mathcal{C}^1[a,b] \longrightarrow \mathbb{R}, \quad y \longmapsto \max_{a \leq x \leq b} |y(x)| + \max_{a \leq x \leq b} |y'(x)|$$

We say $J[y]$ has a strong extremum at $y = \hat{y}$ if there exists an $\varepsilon > 0$ such that the sign of $J[y] - J[\hat{y}]$ is the same for all y contained in $||y - \hat{y}||_0 < \varepsilon$ where

$$|| \cdot ||_0 \colon \mathcal{C}^1[a,b] \longrightarrow \mathbb{R}, \quad y \longmapsto \max_{a \leq x \leq b} |y(x)|$$

Remark. Note that by our definition if $\hat{y}$ is a strong extremum it is also a weak extremum; however, the converse is not necessarily true.

Theorem 1.8. *A necessary condition for the differential functional $J[y]$ to have an extremum for $y = \hat{y}$ is that its variation vanishes for $\hat{y}$, i.e.*

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

for $y = \hat{y}$ and all admissible h.

Proof. Without loss of generality, suppose $J[\hat{y}]$ is a minimum. So there exists an $\varepsilon > 0$ such that $J[y] - J[\hat{y}]$ is positive where y is contained in $||y - \hat{y}|| < \varepsilon$. By definition we have

$$\Delta J[h] = J[y+h] - J[y] = \varphi[h] + \varepsilon ||h||$$

where $\varepsilon \to 0$ as $||h|| \to 0$ and $\varphi[h] = \delta J[h]$ is linear. Then for sufficiently small $||h||$, $sign(\Delta J[h]) = sign(\delta J[h])$. Now suppose for contradiction that $\delta J[h_0] \neq 0$ for some admissible h_0. Then for any $\alpha > 0$, no matter how small, we have by linearity

Since h_0 is admissible (continuous with zero as endpoints), $-\alpha h_0$ is also admissible. Hence for sufficiently small $||h||$ we can make $\delta J[h] + \varepsilon ||h||$ have either sign, a contradiction since for $y = \hat{y}$ we know

$$\Delta J[h] = \varphi[h] + \varepsilon ||h|| = J[\hat{y} + h] - J[\hat{y}] \geq 0$$

□

2 The Euler-Lagrange Equation

Theorem 2.1. *Let $J[y]$ be a functional of the form*

$$\int_a^b L(x, y, y')dx \tag{3}$$

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

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

Proof. Suppose $y(x) \in \mathcal{C}^1[a,b]$ is the function which admits an extremum of (3). Then if we increment $y(x)$ by some function $\varepsilon h(x)$ where ε is small and

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

we still satisfy our boundary conditions. The corresponding increment of our functional equals

$$\Delta J = J[y + \varepsilon h] - J[y] = \int_a^b \left[L(x, y + \varepsilon h, y' + \varepsilon h') - L(x, y, y') \right] dx. \tag{4}$$

Recall, Taylor's theorem for an n times differentiable function $f(x,y)$ of two variables asserts the following expansion at $(a,b) \in \mathbb{R}^2$

$$f(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x - a + \frac{\partial f}{\partial y}(a,b)(y-b) + G(x,y)$$

where $G(x,y)$ denotes the higher order $n-2$ terms. Therefore, by fixing x in (4) and expanding at the point $(y + \varepsilon h, y' + \varepsilon h')$, Taylor's theorem asserts that

$$\Delta J = \int_a^b \left[\frac{\partial L}{\partial y} \varepsilon h + \frac{\partial L}{\partial y'} \varepsilon h' + \varepsilon^2 G(x) \right] dx$$

Where $G(x)$ denotes higher order terms, scaled by ε^2 because the terms are quadratic with respect to εh and $\varepsilon h'$. Now if $\varepsilon > 0$, we have

$$\int_a^b \left[\frac{\partial L}{\partial y} \varepsilon h + \frac{\partial L}{\partial y'} \varepsilon h' + \varepsilon^2 G(x) \right] dx = \varepsilon \int_a^b \left[\frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y'} h' \right] dx + \varepsilon^2 \int_a^b G(x) dx$$

Where the first integral on the right hand side is the principal linear part of the increment ΔJ and the second integral denotes the vanishing terms. By Theorem 1.8, we must have

$$\int_a^b \left[\frac{\partial L}{\partial y} h(x) + \frac{\partial L}{\partial y'} h'(x) \right] dx = \int_a^b \frac{\partial L}{\partial y} h(x) dx + \int_a^b \frac{\partial L}{\partial y'} h'(x) = 0 \tag{5}$$

We integrate by parts the second integral in (5):

$$\int_a^b \frac{\partial L}{\partial y'} h'(x) dx = \left[\frac{\partial L}{\partial y'} h(x) \right]_a^b - \int_a^b \frac{dh}{dx} \frac{\partial L}{\partial y'} dx = - \int_a^b \frac{dh}{dx} \frac{\partial L}{\partial y'} dx \tag{6}$$

Combining (6) with (5) gives

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

Since $h(x)$ is arbitrary, we must have

$$\frac{\partial L}{\partial y} - \frac{d}{dx} \frac{\partial L}{\partial y'} = 0 \tag{7}$$

□

Theorem 2.2. *(Bernstein) If the functions L, L_y, and $L_{y'}$ are continuous at every finite point (x, y) for any finite y', and if a constant $k > 0$ and functions*

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

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

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

then one and only one integral curve of $y'' = L(x, y, y')$ passes through any two points (a, A) and (b, B) with different abscissas (different first coordinates $a \neq b$)

Remark. Bernstein's theorem gives conditions for the existence and uniqueness for the solution to $y'' = L(x, y, y')$, a second order differential equation. Note that the Euler-Lagrange equation is of this form. The next theorem gives conditions which guarantee that a solution to the Euler-Lagrange equation has a second derivative.

Theorem 2.3. *Suppose $y = y(x)$ has a continuous first derivative and satisfies*

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0.$$

Then, if $L(x, y, y')$ has continuous first and second derivatives with respect to all its arguments, $y(x)$ has a continuous second derivative at all points (x, y) where

$$\frac{\partial^2 L}{\partial y'^2}(x, y(x), y'(x)) \neq 0$$

Proof. Consider the difference

$$\Delta L_{y'} = L_{y'}(x + \Delta x, y + \Delta y, y' + \Delta y') - L(x, y, y') = \Delta x \bar{L}_{y'x} + \Delta y \bar{L}_{y'y} + \Delta y' \bar{L}_{y'y'}$$

Where the bar indicates that the derivatives are evaluated intermediate curves. Dividing by Δx and taking the limit as $\Delta x \to 0$ gives us

$$\lim_{\Delta x \to 0} \frac{\Delta L_{y'}}{\Delta x} = \lim_{\Delta x \to 0} \bar{L}_{y'x} + \frac{\Delta y}{\Delta x}\bar{L}_{y'y} + \frac{\Delta y'}{\Delta x}\bar{L}_{y'y'}$$

This limit exists because the Euler-Lagrange equation tells us the derivative of $L_{y'}$ with respect to x is L_y. Since the second derivatives of L are continuous, we have as $\Delta x \to 0$, $\bar{L}_{y'x} \to L_{y'x}$, that is,

$$\bar{L}_{y'x} \to \frac{\partial^2 L}{\partial y' \partial x}(x, y(x), y'(x))$$

Where (x, y) is the aforementioned point in space. It follows from the existence of y' and the continuity of all second derivatives of L that $\frac{\Delta y}{\Delta x}\bar{L}_{y'y}$ also exists as $\Delta x \to 0$. Therefore the third term must also have a limit since the limit of the sum of the three terms exists. As $\Delta x \to 0$ $\bar{L}_{y'y'} \to L_{y'y'} \neq 0$, and hence

$$\lim_{\Delta x \to 0} \frac{\Delta y'}{\Delta x} = y''(x)$$

must exist. Finally, from the Euler-Lagrange equation,

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

we can find an expression for y'', from which it is clear that y'' is continuous wherever $L_{y'y'} \neq 0$.

□

2.1 Question

The proof argues that the third term's limit must exist because the first two limits exist and the limit of the entire expression exists, but why can't this argument be used in the first place to say all three terms must have a limit?

Remark. Reducing the Euler-Lagrange equation to several cases simplifies the process of finding extremal curves.

Case 1. *Suppose the integrand does not depend on y, that is,*

$$J[y] = \int_a^b L(x, y')dx$$

Then the Euler-Lagrange equation states

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = 0 \iff \frac{\partial L}{\partial y'} = C$$

Where C is some constant. This is a first order differential equation of the form

$$y' = f(x, C)$$

Case 2. *Suppose the integrand does not depend on x, so*

$$J[y] = \int_a^b L(y, y')dx$$

We use the chain rule to compute $\frac{d}{dx}\frac{\partial L}{\partial y'}$,

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = \frac{\partial^2 L}{\partial y' \partial y} y' + \frac{\partial^2 L}{\partial y'^2} y''$$

Note that the second derivative with respect to x is zero. The Euler-Lagrange equation becomes

$$\frac{\partial L}{\partial y} - \frac{\partial^2 L}{\partial y' \partial y} y' - \frac{\partial^2 L}{\partial y'^2} y'' = 0$$

Multiplying by y' gives

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

It can be verified using the chain rule that (8) is equivalent to

$$\frac{d}{dx}\left(L - \frac{\partial L}{\partial y'} y'\right) = 0$$

Which means

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

Where C is some constant.

Case 3. *Suppose the integrand does not depend on y', so*

$$J[y] = \int_a^b L(x, y)dx$$

In this case, the Euler-Lagrange equation states that

$$\frac{\partial L}{\partial y} = 0,$$

which is not a differential equation. The solution consists of one or more curves $y = y(x)$.

Case 4. *One often encounters functionals of the form*

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

representing an integral of a function $f(x,y)$ with respect to the arc length s ($ds = \sqrt{1+y'^2}dx$)

We use the chain rule and product rule to compute the Euler-Lagrange equation,

$$\frac{\partial L}{\partial y} = \frac{\partial f}{\partial y}\sqrt{1+y'^2},$$

$$\frac{\partial L}{\partial y'} = f(x,y)\frac{y'}{\sqrt{1+y'^2}},$$

$$\frac{d}{dx}\left(f(x,y)\frac{y'}{\sqrt{1+y'^2}}\right) = \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}y'\right)\frac{y'}{\sqrt{1+y'^2}} + f(x,y)\frac{y''}{(1+y'^2)^{\frac{3}{2}}}$$

Therefore, Euler-Lagrange equation becomes

$$\frac{\partial f}{\partial y}\sqrt{1+y'^2} - \frac{\partial f}{\partial x}\frac{y'^2}{\sqrt{1+y'^2}} - f(x,y)\frac{y''}{(1+y'^2)^{\frac{3}{2}}} = 0$$

Which can be simplified to

$$\frac{1}{\sqrt{1+y'^2}}\left(\frac{\partial f}{\partial y} - \frac{\partial f}{\partial x}y' - f(x,y)\frac{y''}{1+y'^2}\right) = 0$$

i.e.,

$$\frac{\partial f}{\partial y} - \frac{\partial f}{\partial x}y' - f(x,y)\frac{y''}{1+y'^2} = 0$$

2.2 Problem

$$J[y] = \int_1^2 \frac{\sqrt{1+y'^2}}{x}dx, \qquad y(1) = 0,\ y(2) = 1$$

The integrand does not contain y, so $\frac{\partial L}{\partial y'} = C$. We have that

$$\frac{\partial L}{\partial y'} = \frac{\partial}{\partial y'}\left(\frac{\sqrt{1+y'^2}}{x}\right) = \frac{y'}{x\sqrt{1+y'^2}} = C$$

$$y' = Cx\sqrt{1+y'^2}$$
$$y'^2 = C^2x^2 + C^2x^2y'^2$$
$$y'^2(1 - C^2x^2) = C^2x^2$$

$$y'^2 = \frac{C^2x^2}{1 - C^1x^2}$$

$$y' = \frac{Cx}{\sqrt{1 - C^2x^2}}$$

$$y = \int \frac{Cx}{\sqrt{1 - C^2x^2}}dx = \frac{1}{C}\sqrt{1 - C^2x^2} + C_1$$

$$(y - C_1)^2 + x^2 = \frac{1}{C^2}$$

By plugging in the initial conditions we get $C = \frac{1}{\sqrt{5}},\ C_1 = 2$ for a final solution of

$$(y - 2)^2 + x^2 = 5.$$

3 The Case of Several Variables

Functionals which take functions of one independent variable as arguments map curves to scalar values. We can generalize this notion to functionals which take as argument functions of multiple independent variables. For example, a functional which takes as input a two argument function maps a surface to a scalar value. For notational purposes we will limit our examples to functions of two independent variables, i.e.,

$$J[z] = \iint_R L(x, y, z, z_x, z_y)dxdy \tag{9}$$

where R is a closed region and $z_x,\ z_y$ are the partial derivatives of $z = z(x, y)$, but our results can all be generalize to functions of n independent variables.
Suppose we are looking for a function $z = z(x, y)$ such that

1. $z(x, y)$ and its first and second derivatives are continuous on R;
2. $z(x, y)$ takes given values on the boundary ∂R;
3. (9) has an extremum for $z = z(x, y)$.

The proof of Theorem 1.8 does not depend on the form of the functional J, hence for the case of several variables, the variation must still vanish at extrema.

Lemma 3.1. *If $\alpha(x, y)$ is a fixed function which is continuous on R and if*

$$\iint_R \alpha(x, y)h(x, y)dxdy = 0$$

for every continuous $h(x, y)$ with continuous first and second derivatives and zero boundary values, i.e. $h(x, y) = 0\ \forall(x, y) \in \partial R$, then $\alpha(x, y)$ vanishes everywhere in R.

Proof. Suppose for contradiction that $\alpha(x, y) > 0$ at (x_0, y_0). Then from continuity we know that α is positive in some circle $S = (x - x_0)^2 + (y - y_0)^2 \leq \varepsilon^2$. Define $h(x, y)$ by

$$h(x, y) = \begin{cases} 0 & if\ h(x, y) \notin S \\ [(x - x_0)^2 + (y - y_0)^2 - \varepsilon^2]^3 & if\ h(x, y) \in S \end{cases}$$

Then $h(x, y)$ satisfies the conditions of the lemma, but

$$\iint_R \alpha(x,y)h(x,y)dxdy = \iint_S \alpha(x,y)h(x,y)dxdy$$

which must be nonzero as the function values within S are nonzero. This contradiction proves the lemma. $\square$

Remark. One should compare this lemma to Lemma 1.1, where this lemma is its the two variable analog.

3.1 Problem

$$J[z] = \iint_R \sqrt{1 + z_x^2 + z_y^2}\ dxdy$$

The Euler-Lagrange equation for two variable functions states that

$$L_z - \frac{\partial}{\partial x} L_{z_x} - \frac{\partial}{\partial y} L_{z_y} = 0$$

Where L is the Lagrangian. Using the chain rule and implicit differentiation we calculate the given derivatives. Since the Lagrangian does not depend on z, $L_z = 0$. For $\frac{\partial}{\partial x} L_{z_x}$ we have

$$L_{z_x} = \frac{\partial}{\partial z_x}\left(\sqrt{1 + z_x^2 + z_y^2}\,\right) = \frac{z_x}{\sqrt{1 + z_x^2 + z_y^2}}$$

$$\frac{\partial}{\partial x} L_{z_x} = \frac{\partial}{\partial x}\left(\frac{z_x}{\sqrt{1 + z_x^2 + z_y^2}}\right) = \frac{z_{xx}\sqrt{1 + z_x^2 + z_y^2} - z_x\left(\frac{\partial}{\partial x}\sqrt{1 + z_x^2 + z_y^2}\right)}{1 + z_x^2 + z_y^2}$$

$$= \frac{z_{xx}\sqrt{1 + z_x^2 + z_y^2} - z_x\left(\frac{z_x z_{xx} + z_y z_{xy}}{\sqrt{1+z_x^2+z_y^2}}\right)}{1 + z_x^2 + z_y^2} \tag{10}$$

For $\frac{\partial}{\partial y} L_{z_y}$ we have

$$L_{z_y} = \frac{\partial}{\partial z_y}\sqrt{1 + z_x^2 + z_y^2} = \frac{z_y}{\sqrt{1 + z_x^2 + z_y^2}}$$

$$\frac{\partial}{\partial y}\left(\frac{z_y}{\sqrt{1 + z_x^2 + z_y^2}}\right) = \frac{z_{yy}\sqrt{1 + z_x^2 + z_y^2} - z_y\left(\frac{\partial}{\partial y}\sqrt{1 + z_x^2 + z_y^2}\right)}{1 + z_x^2 + z_y^2}$$

$$= \frac{z_{yy}\sqrt{1 + z_x^2 + z_y^2} - z_y\left(\frac{z_y z_{yy} + z_x z_{xy}}{\sqrt{1+z_x^2+z_y^2}}\right)}{1 + z_x^2 + z_y^2} \tag{11}$$

Combining (10) and (11) the Euler-Lagrange equation states that

$$\frac{z_x\left(\frac{z_x z_{xx} + z_y z_{xy}}{\sqrt{1+z_x^2+z_y^2}}\right) + z_y\left(\frac{z_y z_{yy} + z_x z_{xy}}{\sqrt{1+z_x^2+z_y^2}}\right) - (z_{xx} + z_{yy})\sqrt{1 + z_x^2 + z_y^2}}{1 + z_x^2 + z_y^2} = 0$$

Since this implies the numerator must equal zero, we ignore the denominator and multiply the equation by $\sqrt{1+z_x^2+z_y^2}$ to obtain

$$z_x(z_x z_{xx} + z_y z_{xy}) + z_y(z_y z_{yy} + z_x z_{xy}) - (z_{xx} + z_{yy})\,(1 + z_x^2 + z_y^2) = 0$$

With further simplification we get

$$\begin{aligned} z_x^2 z_{xx} + z_x z_y z_{xy} + z_y^2 z_{yy} + z_x z_y z_{xy} - z_x x - z_x^2 z_{xx} - z_y^2 z_{xx} - z_{yy} - z_x^2 z_{yy} - z_y^2 z_{xx} \\ = z_{xx}(1 + z_y^2) - 2 z_{xy} z_x z_y + z_{yy}(1 + z_x^2) = 0 \end{aligned} \tag{12}$$

4 A Variable Endpoint Problem

A particular case of variable endpoints is finding the curve for which the functional

$$J[y] = \int_a^b L(x, y, y')dx$$

has an extremum, where the class of curves has points which lie on two given vertical lines $x = a$ and $x = b$. A more general case is where the endpoints lie on two given curves $y = \varphi(x)$ and $y = \psi(x)$, but we will consider this case later. To begin, we calculate the increment

$$\Delta J = J[y+h] - J[y] = \int_a^b \big[L(x, y+h, y'+h') - L(x, y, y')\big]\,dx$$

The Taylor expansion of the Lagrangian gives

$$\Delta J = \int_a^b (L_y h + L_{y'} h')dx + \cdots$$

Where the dots denote terms of order higher than 1 relative to h and h'. Hence

$$\delta J = \int_a^b (L_y h + L_{y'} h')dx \tag{13}$$

We integrate (13) by parts:

$$\int_a^b (L_y h + L_{y'} h')dx = \int_a^b L_y h dx + \int_a^b L_{y'} h' dx = \int_a^b L_y h dx - \int_a^b \frac{d}{dx} L_{y'} h dx + \big[L_{y'} h(x)\big]_{x=a}^{x=b}$$

If $h(a) = h(b) = 0$, which it can, then we are reduced to the original case where

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

However since $h(x)$ is arbitrary, the cases in which the endpoints do not vanish leave us with

$$\delta J = 0 \implies \left[\frac{\partial L}{\partial y}\right]_{x=b} h(b) - \left[\frac{\partial L}{\partial y'}\right]_{x=a} h(a) = 0$$

It follows that

$$\left[\frac{\partial L}{\partial y}\right]_{x=b} = 0, \qquad \left[\frac{\partial L}{\partial y'}\right]_{x=a} = 0$$

4.1 Problem (Incomplete)

Starting from $P = (a, A)$, a heavy particle slides down a curve in the vertical plane. Find the curve such that the particle reaches the vertical line $x = b \neq a$ in the shortest time.
This is a variant of the Brachistochrone problem. We know velocity is change in position with respect to time, i.e.

$$v = \frac{ds}{dt} = \frac{dx}{dt}\sqrt{1 + y'^2}$$

Recall that in Newtonian mechanics $v = \sqrt{2gy}$, where g is the acceleration of gravity. Therefore, we have

$$dt = \frac{\sqrt{1 + y'^2}}{v}dx = \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}dx$$

So total time is the integral

$$T = \int \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}dx$$

The Lagrangian does not depend on x, so our problem reduces to the case

$$\frac{d}{dx}\left[L - \frac{\partial L}{\partial y'}y'\right] = 0 \implies L - \frac{\partial L}{\partial y'}y' = C$$

that is,

$$\frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} - \frac{y'^2}{\sqrt{2gy}\sqrt{1 + y'^2}} = C$$

We can simplify this equation to

$$\frac{1}{\sqrt{2gy}\sqrt{1 + y'^2}} = C \tag{14}$$

Squaring (14) gives

$$\frac{1}{2gy(1 + y'^2)} = C^2$$

The solution in the book states that the general solution to the Euler-Lagrange equation is a family of cycloids,

$$x = r(\theta - \sin\theta) + c, \qquad y = r(1 - \cos\theta)$$

But the book didn't explain why or how to get to that point.

4.2 Problem

A set $\mathcal{M}$ in a normed linear space $\mathscr{R}$ is said to convex if $\mathcal{M}$ contains all elements of the form $\alpha x + \beta y$, where $\alpha + \beta = 1$, provided that $\mathcal{M}$ contains x and y. Prove that the set, A, of all elements $x \in \mathscr{R}$ satisfying $||x - x_0|| \leq c$, where x_0 is fixed and $c > 0$, is convex.

Proof. Given $x, y \in A$ we want to show $\alpha x + \beta y \in A$ where $\alpha + \beta = 1$. This is equivalent to showing $\alpha x + (1 - \alpha)y \in A$, where $0 \leq \alpha \leq 1$. We know that $||x - x_0|| < c$ and $||y - x_0|| < c$. Therefore, using the triangle inequality, we see that

$$\alpha||x - x_0|| + (1 - \alpha)||y - x_0|| < \alpha c + (1 - \alpha)c = c$$

$$\iff ||\alpha(x - x_0)|| + ||(1 - \alpha)(y - x_0)|| < c$$

$$\iff ||\alpha x + (1-\alpha)y - \alpha x_0 - (1-\alpha)x_0|| < c$$

Since $\alpha x_0 + (1-\alpha)x_0 = x_0$, we have that

$$||\alpha x + (1-\alpha)y - x_0|| < c$$

□

5 The Variational Derivative

We do an informal derivation of the variational (or functional) derivative. We consider functionals of the type

$$J[y] = \int_a^b L(x, y, y')dx, \qquad y(a) = A, \; y(b) = B \tag{15}$$

We divide $[a, b]$ into $n+1$ subintervals of equal length:

$$x_0 = a, \; x_1, \; x_2, \ldots, x_n, \; x_{n+1} = b \qquad \Delta x := x_{i+1} - x_i$$

Next we approximate the curve $y(x)$ by the polygonal line with vertices given by the points

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n), (x_{n+1}, y_{n+1})$$

Where $y_i \equiv y(x_i)$. We can approximate (15) by

$$J[y] \approx J(y_1, \ldots, y_n) = \sum_{i=0}^{n} L\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right)\Delta x \tag{16}$$

To optimize this approximation, we must take the partial derivatives

$$\frac{\partial J}{\partial y_k}, \qquad k \in \{1, \ldots, n\} \tag{17}$$

Note that y_0 and y_{n+1} are both fixed. For each k, y_k appears in the k and $k-1$ terms of the summation. Therefore, the partial derivative in (17) is equal to

$$\frac{\partial}{\partial y_k}\left[L\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) + L\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right)\right]\Delta x \tag{18}$$

Observe in (18) that y_k appears in y' parameter of the first Lagrangian and the y parameter of the second Lagrangian, and $-y_k$ appears in the y' parameter of the second Lagrangian. Therefore we must add the partial derivatives of each Lagrangian with respect to those parameters, multiplying by constants when necessary. This gives us

$$\frac{\partial J}{\partial y_k} = \frac{\partial L}{\partial y}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right)\Delta x + \frac{\partial L}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) - \frac{\partial L}{\partial y'}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) \tag{19}$$

Dividing (19) by Δx gives us

$$\frac{\partial J}{\partial y_k \Delta x} = \frac{\partial L}{\partial y}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right) + \frac{1}{\Delta x}\left[\frac{\partial L}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) - \frac{\partial L}{\partial y'}\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right)\right]$$

Taking the limit as $\Delta x \to 0$ gives

$$\lim_{\Delta x \to 0} \frac{\partial J}{\partial y_k \Delta x} = \lim_{\Delta x \to 0} \frac{\partial L}{\partial y}\left(x_k, y_k, \frac{y_{k+1}-y_k}{\Delta x}\right) + \frac{1}{\Delta x}\left[\frac{\partial L}{\partial y'}\left(x_{k-1}, y_{k-1}, \frac{y_k - y_{k-1}}{\Delta x}\right) - \frac{\partial L}{\partial y'}\left(x_k, y_k, \frac{y_{k+1}-y_k}{\Delta x}\right)\right]$$

Which converges to the limit

$$\frac{\delta J}{\delta y} \equiv \frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'}$$

which is the Euler-Lagrange equation. To summarize, we took a functional $J[y]$ and approximated it by an n variable function $J(y_1, \ldots, y_n)$ where the curve $y(x)$ is approximated by a polygonal line of vertives $(x_i, y(x_i))$. Taking the limit as $n \to \infty$ is equivalent to taking the limit as $\Delta x \to 0$, which gives the functional $J[y]$. Taking the partial derivatives of our approximation gave an approximation to the variational derivative, which converges to the Euler-Lagrange equation as the approximation converges to the functional.

Remark. This is not a rigorous derivation as the limits were not proven to converge to their given values.

5.1 Problem (Incomplete)

Use Euler's method of finite differences to find the shortest plane curve joining two points A and B.
We know the formula for length of a curve between a and b is

$$J[y] = \int_a^b \sqrt{1 + y'^2}\ dx$$

We approximate this functional with

$$J[y] \approx \hat{J}(y_1, \ldots, y_n) = \sum_{i=1}^{n} \sqrt{1 + \left(\frac{y_{i+1}-y_i}{\Delta x}\right)^2} \Delta x$$

The derivative, which should vanish at extremal points, of $\hat{J}$ with respect to y_i is:

$$\frac{y_i - y_{i-1}}{\Delta x \sqrt{1 + \left(\frac{y_i - y_{i-1}}{\Delta x}\right)^2}} + \frac{y_i - y_{i-1}}{\Delta x \sqrt{1 + \left(\frac{y_{i+1} - y_i}{\Delta x}\right)^2}} = 0 \quad \forall i \in \{1, \ldots, n\}$$

6 Variance of the Euler-Lagrange Equation

Here we explain the general case of changing coordinate systems when optimizing functionals. We introduce curvilinear coordinates u and v such that

$$x = x(u,v), \quad y = y(u,v), \quad \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \neq 0 \tag{20}$$

Then $y(x)$ in the xy-plane corresponds to $v(u)$ in the uv-plane, and the functional

$$J[y] = \int_a^b L(x, y, y')dx$$

goes into the functional

$$J_1[v] = \int_{a_1}^{b_1} L\left(x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v')du$$

Let

$$L_1(u,v,v') := L\left(x(u,v), y(u,v), \frac{y_u + y_v v'}{x_u + x_v v'}\right)(x_u + x_v v')$$

We show that if $y(x)$ satisfies

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

Then $v(u)$ satisfies

$$\frac{\partial L_1}{\partial v} - \frac{d}{du}\frac{\partial L_1}{\partial v'} = 0$$

7 Fixed Endpoints for n Unknown Functions

Let $\vec{\mathbf{y}} = (y_1, y_2, \ldots, y_n)$ and $\vec{\mathbf{y'}} = (y_1', \ldots, y_n')$. Where $y_i \in \mathcal{C}^\infty[a,b]$, for $i = 1, \ldots, n$, and $y_i(a) = A_i,\ y_i(b) = B_i$ are fixed. We wish to optimize the functional

$$J[\vec{\mathbf{y}}] = \int_a^b L(x, \vec{\mathbf{y}}, \vec{\mathbf{y'}})dx \tag{21}$$

Where $L \in \mathcal{C}^2[a,b]$.

Theorem 7.1. *A necessary condition for the curve $y_i(x),\ i \in \{1, \ldots, n\}$ to be an extremal of* (1) *is that $y_i(x)$ satisfies*

$$\frac{\partial L}{\partial y_i} - \frac{d}{dx}\frac{\partial L}{\partial y_i'} = 0$$

Proof. Suppose $\vec{\mathbf{y}}$ locally minimizes (1). Let ε be small, let $\vec{\mathbf{h}}(x) = (h_1(x), \ldots, h_n(x))$ be a vector of smooth functions on $[a,b]$ satisfying $h_i(a) = h_i(b) = 0$ for $i \in \{1, \ldots, n\}$, and let $\vec{\mathbf{h'}}(x) = (h_1'(x), \ldots, h_n'(x)$. We apply an increment to J,

$$J[\vec{\mathbf{y}} + \varepsilon\vec{\mathbf{h}}] - J[\vec{\mathbf{y}}] = \int_a^b L(x, \vec{\mathbf{y}} + \varepsilon\vec{\mathbf{h}}, \vec{\mathbf{y'}} + \varepsilon\vec{\mathbf{h'}}) - L(x, \vec{\mathbf{y}}, \vec{\mathbf{y'}})dx \tag{22}$$

By Taylor's theorem we know (2) is equivalent to

$$\int_a^b \left[\nabla_y L \cdot \varepsilon\vec{\mathbf{h}} + \nabla_{y'} L \cdot \varepsilon\vec{\mathbf{h'}} + \varepsilon^2 G(x)\right] dx \tag{23}$$

Where $\nabla_y = (\frac{\partial}{\partial y_1}, \ldots \frac{\partial}{\partial y_n})$ and similar notation is used for $\nabla_{y'}$. $G(x)$ is the function representing the higher order terms of the Taylor expansion. For $\vec{\mathbf{y}}$ to be minimal, (3) must vanish as $\varepsilon \to 0$. Therefore as $\varepsilon \to 0$,

$$\int_a^b \left[\nabla_y L \cdot \varepsilon\vec{\mathbf{h}} + \nabla_{y'} L \cdot \varepsilon\vec{\mathbf{h'}} + \varepsilon^2 G(x)\right] dx = \varepsilon \int_a^b \left[\nabla_y L \cdot \vec{\mathbf{h}} + \nabla_{y'} L \cdot \vec{\mathbf{h'}} + \varepsilon G(x)\right] dx = 0$$

$$\implies \int_a^b \left[\nabla_y L \cdot \vec{\mathbf{h}} + \nabla_{y'} L \cdot \vec{\mathbf{h'}} + \varepsilon G(x)\right] dx = 0$$

$G(x)$ must vanish as $\varepsilon \to 0$, which leaves us with

$$\int_a^b \left[\nabla_y L \cdot \vec{\mathbf{h}} + \nabla_{y'} L \cdot \vec{\mathbf{h}'}\right] dx = \int_a^b \nabla_y L \cdot \vec{\mathbf{h}}\; dx + \int_a^b \nabla_{y'} L \cdot \vec{\mathbf{h}'}\; dx = 0$$

Integrating the second integral by parts gives us

$$\int_a^b \left[\nabla_y L - \frac{d}{dx}\nabla_{y'} L\right] \vec{\mathbf{h}}\; dx = 0$$

$$\iff \int_a^b \vec{\mathbf{h}}(x) \sum_{i=1}^n \left(\frac{\partial L}{\partial y_i} - \frac{d}{dx}\frac{\partial L}{\partial y_i'}\right) dx = 0$$

$$\iff \sum_{i=1}^n \int_a^b \vec{\mathbf{h}}(x) \left(\frac{\partial L}{\partial y_i} - \frac{d}{dx}\frac{\partial L}{\partial y_i'}\right) dx = 0$$

Since $\vec{\mathbf{h}}(x)$ is arbitrary, this implies

$$\frac{\partial L}{\partial y_i} - \frac{d}{dx}\frac{\partial L}{\partial y_i'} = 0 \qquad \forall i \in \{1, \ldots, n\}$$

□

Remark. Different integrands can lead to the same system of Euler-Lagrange equations. Let

$$\Phi = \Phi(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'})$$

be any twice differentiable function and let

$$\Psi(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'}) = \frac{\partial \Phi}{\partial x} + \sum_{i=1}^n \frac{\partial \Phi}{\partial y_i} y_i' = \frac{d\Phi}{dx}$$

Then by direct calculation we find that

$$\frac{\partial}{\partial y_i} - \frac{d}{dx}\frac{\partial \Psi}{\partial y_i'} = 0 \tag{24}$$

Hence the functionals

$$\int_a^b L(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'}) dx \tag{25}$$

and

$$\int_a^b \left[L(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'}) + \Psi(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'})\right] dx \tag{26}$$

lead to the same system of Euler-Lagrange equations. In fact, since $\Psi(x, \vec{\mathbf{y}}, \vec{\mathbf{y}'}) = \frac{d\Phi}{dx}$, (5) and (6) only differ by a constant.

Remark. We say two functionals are equivalent if they have the same extremals. Therefore, two functionals of the form (1) are equivalent if their integrands differ by a constant or a constant factor $c \neq 0$.

7.1 Question

I tried calculating (4) but I was confused by what $\frac{\partial}{\partial y_i}$ meant, and after assuming $\frac{\partial}{\partial y_i} \equiv \frac{\partial \Psi}{\partial y_i}$ I still could not get the equation to equal zero. I understand intuitively why functionals differing by a constant admit the same system of Euler-Lagrange equations, but I would still like to go over the calculations.

8 Variational Problems in Parametric Form

Theorem 8.1. *A necessary and sufficient condition for the functional*

$$\int_{t_0}^{t_1} \Phi(t, x, y, \dot{x}, \dot{y}) dt \tag{27}$$

to depend only on the curve in the xy-plane defined by the parametric equations $x = x(t)$, $y = y(t)$ and not on the choice of the parametric representation of the curve, is that the integrand Φ should not involve t explicitly and should be a positive-homogeneous function of degree 1 in $\dot{x}$ and $\dot{y}$.

Proof. Here $\dot{x} \equiv \frac{dx}{dt}$ and similarly for $\dot{y}$. To begin, suppose (7) depends only on the curve in the xy-plane and not on the choice of its parametric representation. Then (7) is equivalent to

$$\int_{t_0}^{t_1} F\left[x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)}\right] \dot{x}(t) dt = \int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y}) dt$$

Note that $\frac{\dot{y}(t)}{\dot{x}(t)} \equiv \frac{dy}{dx}$. Let $\lambda \in \mathbb{R} \setminus \{0\}$. Then

$$\int_{t_0}^{t_1} \Phi(x, y, \lambda\dot{x}, \lambda\dot{y}) dt = \int_{t_0}^{t_1} F\left[x(t), y(t), \frac{\lambda\dot{y}(t)}{\lambda\dot{x}(t)}\right] \lambda\dot{x}(t) dt$$

$$= \int_{t_0}^{t_1} \lambda F\left[[x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)}\right] \dot{x}(t) dt = \int_{t_0}^{t_1} \lambda\Phi(x, y, \dot{x}, \dot{y}) dt$$

Conversely, let

$$\int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y}) dt$$

be a functional whose integrand does not depend on t explicitly and is positive-homogeneous of degree 1 in $\dot{x}$ and $\dot{y}$. If we go from t to a new parameter τ by setting $t = t(\tau)$, where $\frac{dt}{d\tau} > 0$ and $[t_0, t_1]$ goes into $[\tau_0, \tau_1]$, then

$$\int_{\tau_0}^{\tau_1} \Phi\left[x, y, \frac{dx}{d\tau}, \frac{dy}{d\tau}\right] d\tau = \int_{\tau_0}^{\tau_1} \Phi\left[x, y, \dot{x}\frac{dt}{d\tau}, \dot{y}\frac{dt}{d\tau}\right] d\tau$$

$$= \int_{\tau_0}^{\tau_1} \Phi\left(x, y, \dot{x}, \dot{y}\right) \frac{dt}{d\tau} d\tau = \int_{t_0}^{t_1} \Phi\left(x, y, \dot{x}, \dot{y}\right) dt$$

which completes the proof. □

8.1 Question

In the above proof I'm not sure why the integrand must be positive-homogeneous and not simply homogeneous. The book gives an example with the arc length functional where homogeneity does not hold for $\lambda < 0$, but the book did not give any general proof.

Remark. Suppose some parametrization of $y = y(x)$ reduces the functional

$$\int_{x_0}^{x_1} L(x, y, y')$$

to the form

$$\int_{t_0}^{t_1} L\left[x, y, \frac{\dot{y}}{\dot{x}}\right] \dot{x} dt = \int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y}) dt$$

The variational problem for the right hand side leads to the pair of Euler-Lagrange equations

$$\Phi_x - \frac{d}{dt}\Phi_{\dot{x}} = 0, \qquad \Phi_y - \frac{d}{dt}\Phi_{\dot{y}} = 0 \tag{28}$$

which must be equal to the single Euler-Lagrange equation

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

Hence, the two equations in (8) cannot be independent, and it is easily seen that they are related:

$$\dot{x}\left(\Phi_x - \frac{d}{dt}\Phi_{\dot{x}}\right) = 0, \qquad \dot{y}\left(\Phi_y - \frac{d}{dt}\Phi_{\dot{y}}\right) = 0$$

$$\implies \dot{x}\left(\Phi_x - \frac{d}{dt}\Phi_{\dot{x}}\right) + \dot{y}\left(\Phi_y - \frac{d}{dt}\Phi_{\dot{y}}\right) = 0$$

The book will discuss this point further in a later chapter.

9 Functionals Depending on Higher-Order Derivatives

Consider a functional of the form

$$J[y] = \int_a^b L(x, y, y', \ldots, y^{(n)}) dx \tag{29}$$

where $L(x, y, \ldots, y^{(n)})$ has continuous first and second derivatives with respect to all its variables. We consider the case of fixed endpoints, i.e.

$$y(a) = A_0,\ y'(a) = A_1,\ \ldots,\ y^{(n-1)}(a) = A_{n-1}$$

$$y(b) = B_0,\ y'(b) = B_1,\ \ldots,\ y^{(n-1)}(b) = B_{n-1}$$

with $y \in \mathcal{C}^n[a, b]$, and wish to find the curve $y = y(x)$ which gives an extremum for (9).
We proceed as usual, by supposing $y = y(x)$ is a minimum and giving it an increment $y + \varepsilon h$, where $\varepsilon \to 0$ as $||h|| \to 0$ and

$$h^{(i)}(a) = h^{(i)}(b) = 0 \qquad \forall i \in \{0, \ldots, n-1\}$$

Then as $\varepsilon \to 0$, the difference

$$\Delta J = J[y + \varepsilon h] - J[y] \to 0$$

and by Taylor's theorem, this difference is equal to

$$\varepsilon \int_a^b \left[L_y h + L_{y'} h' + \ldots + L_{y^{(n)}} h^{(n)} + \varepsilon G(x) \right] dx \tag{30}$$

where $G(x)$ denotes the higher order terms of the expansion with respect to x. If ε were approaching 0 as a positive number, since y is a minimum, we have the inequality (10) ≥ 0, which implies

$$\int_a^b \left[L_y h + L_{y'} h' + \ldots + L_{y^{(n)}} h^{(n)} + \varepsilon G(x) \right] dx \geq 0$$

As $\varepsilon \to 0$, $G(x)$ vanishes, giving us

$$\int_a^b \left[L_y h + L_{y'} h' + \ldots + L_{y^{(n)}} h^{(n)} \right] dx \geq 0 \tag{31}$$

Now if ε were approaching 0 as a negative number, since y is assumed to be a minimum we still have the inequality (10) ≥ 0. It follows that we have the inequality

$$\int_a^b \left[L_y h + L_{y'} h' + \ldots + L_{y^{(n)}} h^{(n)} \right] dx \leq 0 \tag{32}$$

Since (11) and (12) must be true we deduce that

$$\int_a^b \left[L_y h + L_{y'} h' + \ldots + L_{y^{(n)}} h^{(n)} \right] dx = 0 \tag{33}$$

Repeatedly integrating (13) by parts and using our given boundary conditions gives

$$\int_a^b \left[\sum_{i=0}^{n} (-1)^i \frac{d^i}{dx^i} \frac{\partial L}{\partial y^{(i)}} h(x) \right] dx = 0 \tag{34}$$

We attempt to provide a rigorous argument for why this is true. First note that the existence of higher order derivatives

$$\frac{d^i}{dx^i} \frac{\partial L}{\partial y^{(i)}}$$

is verified from lemma 1.4 of the previous chapter notes. The existence of $\frac{d}{dx}\frac{\partial L}{\partial y'}$ follows directly, and the existence of higher order derivatives follows by substituting $h(x)$ with $h'(x)$ and $h'(x)$ with $h''(x)$. From there we substitute $\beta(x)$ with $\frac{d}{dx}\frac{\partial L}{\partial y'}$ and $\alpha(x)$ with $\frac{d^2}{dx^2}\frac{\partial L}{\partial y''}$ and so on.

Lemma 9.1. *Assuming all the conditions previously asserted or proven in the section, we have that*

$$\int_a^b \frac{d}{dx} \frac{\partial L}{\partial y^{(n+1)}} h^{(n)}(x) dx = \int_a^b (-1)^n \frac{d^{n+1}}{dx^{n+1}} \frac{\partial L}{\partial y^{(n+1)}} h(x) dx$$

Proof. We proceed with induction.
Base Case n = 1. Then

$$\int_a^b \frac{d}{dx} \frac{\partial L}{\partial y''} h'(x) dx = \left[\frac{d}{dx} \frac{\partial L}{\partial y''} h(x) \right]_a^b - \int_a^b \frac{d^2}{dx^2} \frac{\partial L}{\partial y''} h(x) dx$$

$$= \int_a^b (-1) \frac{d^2}{dx^2} \frac{\partial L}{\partial y''} h(x) dx$$

Now suppose

$$\int_a^b \frac{d}{dx} \frac{\partial L}{\partial y^{(n+1)}} h^{(n)}(x) dx = \int_a^b (-1)^n \frac{d^{n+1}}{dx^{n+1}} \frac{\partial L}{\partial y^{(n+1)}} h(x) dx$$

holds true for some n. Then for $n+1$, we integrate by parts to get

$$\int_a^b \frac{d}{dx} \frac{\partial L}{\partial y^{(n+2)}} h^{(n+1)}(x) dx = -\int_a^b \frac{d^2}{dx^2} \frac{\partial L}{\partial y^{(n+2)}} h^{(n)}(x) dx$$

$$= \int_a^b (-1)^{n+1} \frac{d^{n+1}}{dx^{n+1}} \frac{\partial L}{\partial y^{(n+2)}} h(x) dx$$

Where the last equality follows from the induction hypothesis and the fact that

$$\frac{d^2}{dx^2} \frac{\partial L}{\partial y^{(n+2)}} h^{(n)}(x) = \frac{d}{dx} \left(\frac{d}{dx} \frac{\partial L}{\partial y^{(n+2)}} \right) h^{(n)}(x)$$

□

Lemma 9.2. *Assuming all the conditions previously asserted or proven in the section, we have that*

$$\int_a^b \frac{\partial L}{\partial y^{(n)}} h^{(n)}(x) dx = \int_a^b (-1)^n \frac{d^n}{dx^n} \frac{\partial L}{\partial y^{(n)}} h(x) dx$$

Proof. We proceed with induction.
Base Case n = 1. We have

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

$$= \int_a^b (-1) \frac{d}{dx} \frac{\partial L}{\partial y'} h(x) dx$$

Now suppose for some n,

$$\int_a^b \frac{\partial L}{\partial y^{(n)}} h^{(n)}(x) dx = \int_a^b (-1)^n \frac{d^n}{dx^n} \frac{\partial L}{\partial y^{(n)}} h(x) dx$$

Then for $n+1$,

$$\int_a^b \frac{\partial L}{\partial y^{(n+1)}} h^{(n+1)}(x) dx = \left[\frac{\partial L}{\partial y^{(n+1)}} h^{(n)} \right]_a^b - \int_a^b \frac{d}{dx} \frac{\partial L}{\partial y^{(n+1)}} h^{(n)}(x) dx$$

$$= -\int_a^b \frac{d}{dx} \frac{\partial L}{\partial y^{(n+1)}} h^{(n)}(x) dx \tag{35}$$

From lemma 3.1, we know (15) is equal to

$$\int_a^b (-1)^{n+1} \frac{d^{n+1}}{dx^{n+1}} \frac{\partial L}{\partial y^{(n+1)}} h(x) dx$$

Which completes the induction proof. □

Corollary 9.3.

$$\int_a^b \sum_{i=0}^{n} \frac{\partial L}{\partial y^{(i)}} h^{(i)}(x)dx \equiv \int_a^b \sum_{i=0}^{n} (-1)^i \frac{d^i}{dx^i} \frac{\partial L}{\partial y^{(i)}} h(x)dx$$

From this we can deduce that the Euler-Lagrange equation for a functional depending on n higher order derivatives is

$$\frac{\partial L}{\partial y} + \sum_{i=1}^{n} (-1)^i \frac{d^i}{dx^i} \frac{\partial L}{\partial y^{(i)}} = 0$$

10 Variational Problems with Subsidiary Conditions

10.1 The Isoperimetric Problem

10.1.1 The Variational Derivative Revisited

Previously the variational derivative was derived in a way so that given a functional

$$J[y] = \int_a^b L(x, y, y')dx$$

with continuous first and second partial derivatives with respect to y and y', we have that

$$\left.\frac{\delta J}{\delta y}\right|_{x=x_0} \equiv \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'}\right]_{x=x_0}$$

Here we consider a more general case. Let $J[y]$ be a functional depending on $y = y(x)$. We give $y(x)$ an increment $\delta y(x)$ and say $\delta y(x) \neq 0$ in a neighborhood around some $x_0 \in [a, b]$. Let $\Delta\sigma$ be the area between $\delta y(x)$ and the x-axis, i.e.

$$\Delta\sigma = \int_a^b \delta y(x)dx$$

One can also interpret $\Delta\sigma$ as the area between the curves $y(x) + \delta y(x)$ and $y(x)$. We divide the increment difference by this area and let $\Delta\sigma \to 0$ in a way so that $||\delta y(x)|| \to 0$ and the length of the interval in which $\delta y(x)$ is nonzero (the neighborhood around x_0) goes to zero as well. If this ratio converges to a limit, i.e. if

$$\lim_{\Delta\sigma \to 0} \frac{J[y + \delta y] - J[y]}{\Delta\sigma} \equiv \lim_{\int_a^b \delta y(x)dx \to 0} \frac{J[y + \delta y] - J[y]}{\int_a^b \delta y(x)dx}$$

exists, then we call this limit the variational derivative of J for the curve $y = y(x)$ (at $x = x_0$), and we write this limit as

$$\left.\frac{\delta J}{\delta y}\right|_{x=x_0}$$

Remark. From this definition we can deduce that $\Delta J \equiv J[y + \delta y] - J[y]$ is equal to

$$\left[\left.\frac{\delta J}{\delta y}\right|_{x=x_0} + \varepsilon\right] \Delta\sigma$$

where $\varepsilon \to 0$ as $||\delta y|| \to 0$ and the length of the interval in which δy is nonzero (the neighborhood around x_0) goes to zero. It follows that for a given curve $y = y(x)$ and abscissa $x = x_0$ that

$$\delta J = \left.\frac{\delta J}{\delta y}\right|_{x=x_0} \Delta\sigma$$

We now return to the isoperimetric problem. Originally, the isoperimetric problem was the following: Among all closed curves of given length l, find the curve enclosing the greatest area. Hence the term isoperimetric (same perimeter). The modern problem is as follows:
Find the curve $y = y(x)$ for which the functional

$$J[y] = \int_a^b L_0(x, y, y')dx$$

has an extremum, where the curves satisfy the boundary conditions

$$y(a) = A, \quad y(b) = B$$

and are such that another functional

$$K[y] = \int_a^b L_1(x, y, y')dx$$

takes a fixed value l. The classic problem is the specific case where $K[y]$ is the arclength functional and $J[y]$ is the area enclosed by the curve $y = y(x)$ with the given the boundary conditions. To attack this problem we first assume L_0 and L_1 have continuous first and second derivatives in $[a, b]$ for y and y'.

Before we begin with the main results, we first need an important theorem.

Recall. The implicit function theorem states, if a function $F(x,y)$ has continuous derivatives F_x and F_y on its domain and if at the point (x_1,y_1) within its domain $F(x_1,y_1)=0$ and $F_y(x_1,y_1)\neq 0$, then we can mark off about the point (x_1,y_1) a rectangle $[x_0,x_2]\times[y_0,y_2]$ such that for every $x\in[x_0,x_2]$ the equation $F(x,y)=0$ determines exactly one $y=f(x)\in[y_0,y_2]$ that satisfies $y_1=f(x_1)$ and for every $x\in[x_0,x_2]$, $F(x,f(x))=0$. $f(x)$ is continuous and differentiable with $y'=f'(x)=-\frac{F_x}{F_y}$.

Theorem 10.1. *Given the functional*

$$J[y]=\int_a^b L_0(x,y,y')dx$$

Let the admissible curves satisfy the conditions

$$y(a)=A,\qquad y(b)=B,\qquad K[y]=\int_a^b L_1(x,y,y')dx=l$$

where l is fixed and $K[y]$ is some other functional. Then if $y=y(x)$ is not an extremum for $K[y]$, there exists a constant λ such that $y=y(x)$ is an extremum for

$$\int_a^b (L_0+\lambda L_1)\,dx$$

i.e. $y=y(x)$ satisfies

$$\frac{\partial L_0}{\partial y}-\frac{d}{dx}\frac{\partial L_0}{\partial y'}+\lambda\left(\frac{\partial L_1}{\partial y}-\frac{d}{dx}\frac{\partial L_1}{\partial y'}\right)=0$$

Remark. We will present a proof using variational derivatives and afterwards a less rigorous but more intuitive derivation adapted from *The Calculus of Variations* by Brunt.

Proof. Let $y=y(x)$ be an extremum of $J[y]$ satisfying the theorem's given conditions. We choose two points x_1 and x_2 in $[a,b]$ and give $y(x)$ an increment $\delta_1y(x)+\delta_2y(x)$ where $\delta_1y(x)$ is nonzero only in a neighborhood around x_1 and similarly for $\delta_2y(x)$ and x_2. Then using variational derivatives, we can write the corresponding increment ΔJ of J in the form

$$\Delta J=\left[\left.\frac{\delta J}{\delta y}\right|_{x=x_1}+\varepsilon_1\right]\Delta\sigma_1+\left[\left.\frac{\delta J}{\delta y}\right|_{x=x_2}+\varepsilon_2\right]\Delta\sigma_2$$

where $\Delta\sigma_1=\int_a^b\delta_1y(x)dx$ and $\Delta\sigma_2=\int_a^b\delta_2y(x)dx$ and $\varepsilon_1,\varepsilon_2\to 0$ as $\Delta\sigma_1,\Delta\sigma_2\to 0$.

10.2 Question

When writing ΔJ in this form, I understand we are considering two neighborhoods around x_1 and x_2 respectively where there is a nonzero difference between $y=y(x)$ and $y=y(x)+\delta_1y(x)$ and $y=y(x)+\delta_2y(x)$. From what I understand, this is why we can write $\Delta J=J[y+\delta_1y+\delta_2y]-J[y]$ as the sum of two variational derivatives times the corresponding areas. I drew some pictures to help me understand this, but the idea is not entirely rigorous, which is why it has taken me so long to understand this proof. I would like to go over this proof in our next meeting.

Now, we require that $\Delta K = K[y + \delta_1 y + \delta_2 y] - K[y] = l - l = 0$. Writing ΔK in a similar form as we wrote ΔJ, we obtain

$$\Delta K = \left[\frac{\delta K}{\delta y}\bigg|_{x=x_1} + \varepsilon_1' \right] \Delta\sigma_1 + \left[\frac{\delta K}{\delta y}\bigg|_{x=x_2} + \varepsilon_2' \right] \Delta\sigma_2 \tag{36}$$

where $\varepsilon_1', \varepsilon_2' \to 0$ as $\Delta\sigma_1, \Delta\sigma_2 \to 0$. By hypothesis, $K[y]$ does not have an extremal for $y = y(x)$, so we can choose x_2 to be a point for which

$$\frac{\delta K}{\delta y}\bigg|_{x=x_2} \neq 0$$

Therefore, we can write (16) as

$$\Delta\sigma_2 = -\left[\frac{\frac{\delta K}{\delta y}\Big|_{x=x_1}}{\frac{\delta K}{\delta y}\Big|_{x=x_2}} + \varepsilon' \right] \Delta\sigma_1$$

where $\varepsilon' \to 0$ as $\Delta\sigma_1 \to 0$. Then we set

$$\lambda := -\frac{\frac{\delta J}{\delta y}\Big|_{x=x_2}}{\frac{\delta K}{\delta y}\Big|_{x=x_2}}$$

From here we get that

$$\Delta J = \left[\frac{\delta J}{\delta y}\bigg|_{x=x_1} + \lambda \frac{\delta K}{\delta y}\bigg|_{x=x_1} \right] \Delta\sigma_1 + \varepsilon \Delta\sigma_1 \tag{37}$$

where $\varepsilon \to 0$ as $\Delta\sigma_1 \to 0$. This expression for ΔJ explicitly involves variational derivatives only at $x = x_1$ since $\delta_2 y$ has been taken to account by the fact that $\Delta K = 0$.

10.3 Question

I tried doing the algebra to get to (17) but I was not able to get rid of all the terms with respect to $x = x_2$.

Therefore, the first term in the right hand side of (17) is the principal linear part of ΔJ, so

$$\delta J = \left[\frac{\delta J}{\delta y}\bigg|_{x=x_1} + \lambda \frac{\delta K}{\delta y}\bigg|_{x=x_1} \right] \Delta\sigma_1$$

We know that $\delta J = 0$ and $\Delta\sigma_1 \neq 0$. Since x_1 is arbitrary it follows that

$$\frac{\partial L_0}{\partial y} - \frac{d}{dx}\frac{\partial L_0}{\partial y'} + \lambda \left(\frac{\partial L_1}{\partial y} - \frac{d}{dx}\frac{\partial L_1}{\partial y'} \right) = 0$$

which completes the proof. $\square$

Remark. The above result generalizes to functionals depending on several functions with several subsidiary conditions. Suppose we are looking for an extremum of the functional

$$J[y_1, \ldots, y_n] = \int_a^b L_0(x, y_1, \ldots, y_n, y_1', \ldots, y_n')dx$$

and subject to the conditions

$$y_i(a) = A_i, \quad y_i(b) = B_i, \quad \forall i \in \{1, \ldots, n\}$$

and

$$\int_a^b L_j(x, y_1, \ldots, y_n, y_1', \ldots, y_n')dx = l_j, \quad \forall j \in \{1, \ldots, k\}$$

In this case a necessary condition for an extremum is that

$$\frac{\partial}{\partial y_i}\left(L_0 + \sum_{j=1}^{k} \lambda_j L_j\right) - \frac{d}{dx}\left[\frac{\partial}{\partial y_i'}\left(L_0 + \sum_{j=1}^{k} \lambda_j L_j\right)\right] = 0 \quad \forall i \in \{1, \ldots, n\}$$

Remark. The constant λ in the theorem is known as a Lagrange multiplier. For some reason my calculus teachers never thought of teaching Lagrange multipliers, so before doing the following derivation we will take a brief detour.

10.4 Lagrange Multipliers

This is adapted from Brunt as well. Let $f : \mathbb{R}^2 \to \mathbb{R}$ and consider the problem of determining the local extremum of f subject to the condition that the values of f are sampled on a curve/surface $\gamma \subset \mathbb{R}^2$, that is, determining the points on γ in which f has a local extremum. Assume f, γ are smooth and let γ be defined parametrically by

$$\vec{\mathbf{r}} : I \to \mathbb{R}^2, \quad I \subset \mathbb{R}$$

where I is an interval and

$$\vec{\mathbf{r}}(t) = \big\langle x(t), y(t)\big\rangle, \quad \forall t \in I$$

Then we can define

$$F : I \to \mathbb{R}, \quad t \mapsto \big\langle f(x(t)), f(y(t))\big\rangle$$

building f into γ. Given that the parametrization is smooth, a necessary condition for a local extremum at $t \in I$ is

$$\frac{d}{dt}F(t) = \frac{\partial f}{\partial x}\dot{x} + \frac{\partial f}{\partial u}\dot{y} = 0$$

Indeed, we can implicitly define a curve by an equation of the form $g(x, y) = 0$. If g is smooth and $\nabla g \neq 0$, by the implicit function theorem we can solve g for one of its variables; however, this method of solving for g is quite painful and generally involves a fair amount of suffering. Furthermore, even if g is smooth, the implicit functions may not be. For example, let $g(x, y) = x^2 + y^2 - 1$. Solving for y gives $y(x) = \sqrt{1 - x^2}$ which is not smooth at $x = \pm 1$.
Instead of addressing this issue head on, we avoid it, because no one wants to deal with that trash. Lagrange multiplies offer an alternative way for solving these optimization problems. Let $g(x, y) = 0$ define the curve γ implicitly and since $\nabla g \neq 0$, we know γ is smooth, so γ has a well

defined unit tangent vector for all its points. This means we can still represent γ by $\vec{\mathbf{r}}$ from above. Since $\nabla g \neq 0$, we have that $\mathbf{r}' \neq \mathbf{0}$ for all $t \in I$. Since $g(x, y) = 0$ we have

$$\frac{d}{dt} g(x, y) = \frac{\partial g}{\partial x} \dot{x} + \frac{\partial g}{\partial y} \dot{y} = 0, \qquad \forall t \in I$$

But because $\nabla g \neq 0$, it follows that $\frac{\partial g}{\partial x} \neq 0$ or $\frac{\partial g}{\partial y} \neq 0$. Suppose $\frac{\partial g}{\partial y} \neq 0$. Then with a bit of machinery it can be shown that

$$\dot{y} = -\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}} \dot{x} = 0$$

And substituting for $\dot{y}$ in $\frac{df}{dt}$ gives

$$\frac{\dot{x}}{\frac{\partial g}{\partial y}} \left[\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \right] = 0$$

Now since $\mathbf{r}' = \langle \dot{x}, \dot{y} \rangle \neq 0$ and $\dot{y} = 0$, we know that $\dot{x} \neq 0$, so we get that

$$\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} = 0$$

But this is just the determinant

$$\begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix}$$

which is the two dimensional cross product $\nabla f \times \nabla g = ||\nabla f|| \cdot ||\nabla g|| \sin \varphi = 0$. Geometrically this means ∇f is parallel to ∇g, so that $\nabla f = \lambda \nabla g$, where λ is some constant. We call this constant the Lagrange multiplier.

With the Lagrange multiplier introduced to us, we can derive the results from theorem 4.1 in a more straightforward fashion. Let $J : \mathcal{C}^2[a, b] \to \mathbb{R}$ be defined by $J[y] = \int_a^b L(x, y, y')dx$ where L is smooth with respect to all its variables. Let the admissible curves satisfy the boundary conditions $y(a) = A$ and $y(b) = B$ with the additional constraint that $K[y] = \int_a^b g(x, y, y')dx = l$.

We now present another derivation for Theorem 4.1, adapted from Brunt. Let $J : \mathcal{C}^2[a, b] \to \mathbb{R}$ be defined by

$$J[y] = \int_a^b L(x, y, y')dx$$

where L is smooth with respect to its parameters.

10.5 Finite Subsidiary Conditions

In the isoperimetric problem, the subsidiary conditions which must be satisfied by functions $y_1, \dots, y_n$ are of the form

$$K_i[y_1, \dots, y_n] = \int_a^b L_i(x, y_1, \dots, y_n, y_1', \dots, y_n')dx = l_i, \qquad (i = 1, \dots, k)$$

Now consider the following: Find the functions $y_i(x)$ for which the functional

$$J[y] = \int_a^b L_0(x, y, y')dx$$

has an extremum, where the admissible curves satisfy the boundary conditions

$$y_i(a) = A_i, \quad y_i(b) = B_i, \quad (i = 1, \dots, n) \tag{38}$$

and k finite conditions ($k < n$)

$$g_j(x, y_1, \dots, y_n) = 0, \quad (j = 1, \dots, k) \tag{39}$$

We are not concerned with all curves satisfying the boundary conditions in (18). We are only concerned with the curves which satisfy (18) and lie in the $(n-k)$-dimensional manifold defined in (19). We consider the case $n = 2, \ k = 1$.

Theorem 10.2. *Given the functional*

$$J[y, z] = \int_a^b L(x, y, z, y', z')dx$$

Let the admissible curves lie on the surface

$$g(x, y, z) = 0 \tag{40}$$

and satisfy

$$y(a) = A_1, \quad y(b) = B_1$$
$$z(a) = A_2, \quad z(b) = B_2$$

and let $J[y, z]$ have an extremum for $y = y(x)$ and $z = z(x)$. Then if g_y and g_z do not vanish simultaneously at any point on (20), there exists a function $\lambda(x)$ such that $y(x)$ and $z(x)$ is an extremal of

$$\int_a^b [L + \lambda(x)g] \, dx$$

i.e. y and z satisfy

$$L_y + \lambda g_y - \frac{d}{dx} F_{y'} = 0$$
$$L_z + \lambda g_z - \frac{d}{dx} F_{z'} = 0$$

Proof. The proof is similar to the previous theorem. Let $J[y, z]$ have an extremum for $y = y(x), \ z = z(x)$ subject to the given boundary conditions on the manifold $g(x, y, z) = 0$. Let $x_1 \in [a, b]$ be arbitrary. We give $y(x)$ and $z(x)$ increments $\delta y(x)$ and $\delta z(x)$ respectively, where both increments are nonzero only in a neighborhood $[\alpha, \beta]$ of x_1. The existence of these increments on the manifold defined by $g(x, y, z)$ is verified through the implicit function theorem. Using variational derivatives, we write the corresponding increment ΔJ in the form

$$\Delta J = \left[\frac{\delta L}{\delta y}\bigg|_{x=x_1} + \varepsilon_1 \right] \Delta\sigma_1 + \left[\frac{\delta L}{\delta z}\bigg|_{x=x_1} + \varepsilon_2 \right] \Delta\sigma_2$$

where $\Delta\sigma_1 = \int_a^b \delta y(x)dx$ and $\Delta\sigma_2 = \int_a^b \delta z(x)dx$ and $\varepsilon_1, \varepsilon_2 \to 0$ as $\Delta\sigma_1, \Delta\sigma_2 \to 0$. By hypothesis, the varied curves $\hat{y} = y + \delta y$ and $\hat{z} = z + \delta z$ must satisfy $g(x, \hat{y}, \hat{z}) = 0$. Therefore the integral of the difference must also be zero, i.e.

$$0 = \int_a^b [g(x, \hat{y}, \hat{z}) - g(x, y, z)] \, dx = \int_a^b [\bar{g}_y \delta y + \bar{g}_z \delta z] \, dx$$

$$= \left[g_y|_{x=x_1} + \varepsilon_1' \right] \Delta\sigma_1 + \left[\bar{g}_z|_{x=x_1} + \varepsilon_2' \right] \Delta\sigma_2 \tag{41}$$

where $\varepsilon_1', \varepsilon_2' \to 0$ as $\Delta\sigma_1, \Delta\sigma_2 \to 0$ and the overbar indicates that the corresponding derivatives are evaluated along intermediate curves. By hypothesis either $g_y|_{x=x_1}$ or $g_z|_{x=x_1}$ is nonzero, suppose without loss of generality that $g_z|_{x=x_1} \neq 0$. Then we can rewrite (21) in the form

$$\Delta\sigma_2 = -\left[\frac{g_y|_{x=x_1}}{g_z|_{x=x_1}} + \varepsilon' \right] \Delta\sigma_1$$

where $\varepsilon' \to 0$ as $\Delta\sigma_1 \to 0$. We substitute this into the equation for ΔJ and obtain

$$\Delta J = \left[\frac{\delta L}{\delta y}\bigg|_{x=x_1} - \left(\frac{g_y}{g_z} \frac{\delta L}{\delta z} \right)\bigg|_{x=x_1} \right] \Delta\sigma_1 + \varepsilon \Delta\sigma_1$$

where $\varepsilon \to 0$ as $\Delta\sigma_1 \to 0$. Therefore,

$$\delta J = \left[\frac{\delta L}{\delta y}\bigg|_{x=x_1} - \left(\frac{g_y}{g_z} \frac{\delta L}{\delta z} \right)\bigg|_{x=x_1} \right] \Delta\sigma_1 = 0$$

Since $\Delta\sigma_1$ is generally nonzero we have

$$\frac{\delta L}{\delta y} - \frac{g_y}{g_z}\frac{\delta L}{\delta z} = \frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} - \frac{g_y}{g_z}\left(\frac{\partial L}{\partial z} - \frac{d}{dx}\frac{\partial L}{\partial z'} \right) = 0$$

$$= \frac{L_y - \frac{d}{dx} L_{y'}}{g_y} = \frac{L_z - \frac{d}{dx} L_{z'}}{g_z}$$

which when evaluated along $y = y(x)$ and $z = z(x)$, simplifies to

$$L_y + \lambda g_y - \frac{d}{dx} F_{y'} = 0$$

$$L_z + \lambda g_z - \frac{d}{dx} F_{z'} = 0$$

where $\lambda(x)$ is some function. □

The above proof is provided by the book. What follows is another proof of the same theorem, except we first parametrize all the variables with respect to t. Here our arbitrary constant is also a function with respect to t.

Remark. Using finite subsidiary conditions allows for a way to compute the geodesics of a surface.

11 Additional Problems

Problem . Find the extremals of

$$\int_a^b \left(y^2 + y'^2 - 2y\sin x\right) dx$$

The Euler-Lagrange equation states

$$y - \sin x - y'' = 0 \implies y - y'' = \sin x$$

Which is a second order differential equation. We first find a solution to the homogeneous equation:

$$y - y'' = 0$$

Which has a characteristic equation of

$$-r^2 + 1 = 0 \implies r = \pm 1$$

Which gives the solution

$$y_H(x) = C_1 e^x + C_2 e^{-x}$$

to the homogeneous equation. For the nonhomogeneous equation, we assume the solution of of the form

$$y_P(x) = A\cos x + B\sin x$$

for a second derivative of

$$y_P''(x) = -A\cos x - B\sin x$$

Which gives us

$$y_P - y_P'' \equiv 2A\cos x + 2B\sin x = \sin x \implies A = 0,\ B = \frac{1}{2}$$

For a general solution

$$y(x) = C_1 e^x + C_2 e^{-x} + \frac{1}{2}\sin x$$

Where C_1, C_2 are constants.

Problem . Find the extremals of

$$\int_a^b \frac{y'^2}{x^3} dx$$

The Euler-Lagrange equation states

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0$$

Since the Lagrangian has no y term, the partial derivative with respect to y is zero, leaving

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = 0 \implies \frac{\partial L}{\partial y'} = C$$

Where C is some constant. The corresponding differential equation is

$$2x^{-3}y' = C$$

Which has a general solution of

$$y(x) = c_1 x^4 + c_2$$

Where c_1, c_2 are constants.

Problem . Find the extremals of

$$\int_a^b \left(y^2 - y'^2 - 2y\cosh x\right) dx$$

The Euler-Lagrange equation states

$$2y - 2\cosh x + 2y'' = 0 \implies y'' + y = \cosh x$$

We first solve for the homogeneous equation:

$$y_H'' + y_H = 0.$$

Let $y_H(x) = e^{\alpha x}$, $y_H'(x) = \alpha e^{\alpha x}$, $y_H''(x) = \alpha^2 e^{\alpha x}$. Then

$$\alpha^2 e^{\alpha x} + e^{\alpha x} \implies \alpha^2 = -1 \implies \alpha = \pm i$$

For a homogeneous solution of

$$y_H(x) = c_1 \cos x + c_2 \sin x$$

A particular solution to the nonhomogeneous equation is of the form

$$y_P(x) = a\cosh x + b\sinh x \implies y_P''(x) = a\cosh x + b\sinh x$$

Which yields a solution

$$y_P(x) = \frac{1}{2}\cosh x$$

For a general solution

$$y(x) = c_1 \cos x + c_2 \sin x + \frac{1}{2}\cosh x$$

Problem . Find the extremals of

$$\int_a^b \left(y^2 + y'^2 + 2ye^x\right) dx$$

The Euler-Lagrange equation states

$$2y + 2e^x - 2y'' = 0 \implies y'' - y = e^x$$

The homogeneous solution is

$$y_H(x) = c_1 e^x + c_2 e^{-x}$$

Our particular solution is of the form

$$y_P(x) = axe^x,\ y_P''(x) = 2ae^x + axe^x \implies a = \frac{1}{2}$$

For a general solution $y(x) = c_1 e^x + c_2 e^{-x} + \frac{1}{2}xe^x$.

Problem . Find the extremal of

$$\int_a^b \left(y^2 - y'^2 - 2y\sin x\right) dx$$

The Euler-Lagrange equation states

$$2y - 2\sin x + 2y'' = 0 \implies y + y'' = \sin x$$

Similar to an above problem, the homogeneous solution is

$$y_H(x) = c_1 \cos x + c_2 \sin x$$

The particular solution is of the form

$$y_P(x) = ax\sin x + bx\cos x$$

We multiply the trigonometric functions by x because otherwise the particular solution would be linearly dependent to the homogeneous solution. Solving for the unknown constants gives us a solution $y(x) = c_1\cos x + c_2 \sin x - \frac{1}{2}x\cos x$

Problem . Prove the uniqueness part of Bernstein's theorem.
Hint: Let $\Delta(x) = \varphi_2(x) - \varphi_1(x)$, where φ_2, φ_1 are two solutions of $y'' = F(x, y, y')$, write an expression for $\Delta''(x)$ and use the condition $F_y(x, y, y') > k > 0$.

Proof. Let $\Delta(x) = \varphi_2(x) - \varphi_1(x)$, where φ_2, φ_1 are solutions to $y'' = F(x, y, y')$. Then $\Delta'(x) = \varphi_2'(x) - \varphi_1'(x)$ and $\Delta''(x) = \varphi_2''(x) - \varphi_1''(x) = F(x, \varphi_2, \varphi_2') - F(x, \varphi_1, \varphi_1')$. □

Problem P. rove that one and only one extremal of each of the functionals passes through any two points of the plane with different abscissas.

$$\int e^{-2y^2}(y'^2 - 1)dx$$

Proof. Let $y \in \mathcal{C}^2(\Omega)$, where $\Omega \subset \mathbb{R}$ is a closed interval. Since we are working with extremal curves, we first solve the Euler-Lagrange equation for the given Lagrangian:

$$-4ye^{-2y^2}(y'^2 - 1) - 2y''e^{-2y^2} + 8yy'^2e^{-2y^2} = 0$$

$$\iff -2e^{-2y^2}\left[2y(y'^2 - 1) + y'' - 4yy'^2\right] = 0$$

$$\iff 2yy'^2 - 2y + y'' = 0$$

The resulting second order differential equation is

$$y'' = 2y + 2yy'^2 = F(x, y, y')$$

which is continuous on its domain. The partial derivatives:

$$\frac{\partial F}{\partial y} = 2 + 2y'^2, \quad \frac{\partial F}{\partial y'} = 4yy'$$

are continuous as well. Since $2 > 0$ and $y'^2 \geq 0$, $F_y > 0$, therefore there exists a real number $k > 0$ such that $F_y > k$ for all finite points (x, y). Let $\alpha(x, y) = |2y|$ and $\beta(x, y) = |2y|$. Then by the triangle inequality we know

$$|F(x, y, y')| = |2y + 2yy'^2| \leq |2y| + |2y|y'^2 = \beta + \alpha y'^2.$$

Therefore by Bernstein's theorem only one extremal curve passes through any two points of the plane with different abscissas.

$\square$

$$\int \left(y^2 + y' \tan^{-1} y' - \ln \sqrt{1 + y'^2}\right) dx$$

Proof. Let $y \in \mathcal{C}^2(\Omega)$ where $\Omega \subset \mathbb{R}$ is a closed interval. Again, we are dealing with extremal curves, so we solve for the Euler-Lagrange equation:

$$2y - \frac{y''}{1 + y'^2} = 0 \iff y'' = 2y + 2yy'^2 = F(x, y, y')$$

It is easy to see that F is continuous on its domain. We calculate its partial derivatives:

$$\frac{\partial F}{\partial y} = 2 + 2y'^2 \qquad \frac{\partial F}{\partial y'} = 4yy'$$

The partial derivatives are continuous as well. The resulting second order differential equation is the same as in the first proof, hence the proof is complete.

$\square$

Problem . Find the general solution of the Euler-Lagrange equation corresponding to the functional

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

Since the Lagrangian does not depend on y, $L_y = 0$. We have

$$L_{y'} = \frac{f(x)y'}{\sqrt{1 + y'^2}} = C$$

Where C is a constant. Therefore

$$f(x)y' = C\sqrt{1 + y'^2} \implies y' = \frac{C\sqrt{1 + y'^2}}{f(x)} \implies y'^2 = \frac{C^2 + C^2 y'^2}{f^2(x)}$$

$$\implies y'^2 = \frac{C^2}{f^2(x)} + \frac{C^2}{f^2(x)}y'^2 \implies y'^2\left(1 - \frac{C^2}{f^2(x)}\right) = \frac{C^2}{f^2(x)} \implies y' = \frac{C}{f(x)\sqrt{1 - \frac{C^2}{f^2(x)}}}$$

Let $f(x) = \sqrt{x}$. Then

$$y' = \frac{C}{\sqrt{x - C^2}} \implies y = C\int \frac{1}{\sqrt{x - C^2}} dx \implies y = 2C\sqrt{x - C^2} + D$$

Let $f(x) = x$. Then

$$y' = \frac{C}{\sqrt{x^2 - C^2}} \implies y = C \int \frac{1}{\sqrt{x^2 - C^2}} dx$$

Let $x = C \tan u$ Then our integral becomes

$$\int \frac{C \sec^2 u}{C \sec u}\ du = \int \sec u\ du = \ln|\tan u + \sec u| + D$$

$$= \ln \left| \frac{x + \sqrt{x^2 + C^2}}{C} \right| + D$$

Problem . Find all minimal surfaces whose equations have the form $z = \varphi(x) + \psi(y)$.
We want to minimize the integral

$$\iint_R \sqrt{1 + z_x^2 + z_y^2}\ dxdy$$

The Euler-Lagrange equation is

$$\frac{\partial L}{\partial z} - \frac{\partial}{\partial x}\frac{\partial L}{\partial z_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial z_y} = 0$$

From my notes, we know the corresponding partial differential equation is

$$z_{xx}(1 + z_y^2) + z_{yy}(1 + z_x^2) = 0$$

Therefore $z_{xx} = z_{yy} = 0$ which implies $z_x = C \implies z = Cx + D + \psi(y)$, and $z_{yy} = 0 \implies z = \varphi(x) + Ey + F \implies z = Cx + Ey + G$ where C, E, G are constants.
Another solution arises when we separate the variables,

$$z_{xx}(1 + z_y^2) = -z_{yy}(1 + z_x^2) \implies \frac{z_{xx}}{1 + z_x^2} = \frac{-z_{yy}}{1 + z_y^2} \iff \frac{\varphi''(x)}{1 + \varphi'(x)^2} = \frac{-\psi''(y)}{1 + \psi(y)^2} \tag{42}$$

Given the fact that (1) holds for any admissible point (x, y) on the given surface, we know that (1) is equal to a constant,

$$\frac{\varphi''(x)}{1 + \varphi'(x)^2} = \frac{-\psi''(y)}{1 + \psi(y)^2} = C_1$$

$$\implies \arctan \varphi'(x) + C_x = C_1 x, \qquad -\arctan \psi'(y) + C_y = C_1 y$$

$$\implies \varphi(x) = -\frac{1}{C_1} \ln(\cos(C_1 x - C_x)) + C_2', \qquad \psi(y) = \frac{1}{C_1} \ln(\cos(C_1 y - C_y)) + C_2''$$

Let $C_2 := C_2' + C_2''$, then since $z = \varphi(x) + \psi(y)$, we have

$$z = \frac{1}{C_1}(\ln(\cos(C_1 y - C_y)) - \ln(\cos(C_1 x - C_x))) + C_2$$

$$\implies C_1(z - C_2) = \ln\left(\frac{\cos(C_1 y - C_y)}{\cos(C_1 x - C_x)}\right)$$

$$\implies e^{C_1(z - C_2)} = \frac{\cos(C_1 y - C_y)}{\cos(C_1 x - C_x)}$$

Where all C_i are constants.

Problem . Which curve minimizes

$$\int_0^1 \left(\frac{1}{2}y'^2 + yy' + y' + y\right) dx$$

when the values of y are not specified at the end points?
Evaluating extremals with variable endpoints means solving for the Euler-Lagrange equation and using the natural boundary conditions to solve for the resulting constants. The corresponding Euler-Lagrange equation is

$$1 - y'' = 0 \implies y'' = 1 \implies y(x) = \frac{1}{2}x^2 + c_1 x + c_2$$

The natural boundary conditions are

$$\left[\frac{\partial L}{\partial y'}\right]_{x=1} = \left[\frac{\partial L}{\partial y'}\right]_{x=0} = 0$$

$$\equiv y'(0) + y(0) + 1 = y'(1) + y(1) + 1 = 0 \equiv c_1 + c_2 = 2c_1 + c_2 + \frac{5}{2} = 0$$

Solving the system of linear equations gives us the solution

$$y(x) = \frac{1}{2}x^2 - \frac{3}{2}x + \frac{1}{2}$$

Problem . Calculate the variational derivative at the point x_0 of the quadratic functional

$$J[y] = \int_a^b \int_a^b K(s,t)y(s)y(t)dsdt$$

We have,

$$J[y+\delta y] - J[y] = \iint_{[a,b]^2} K(s,t)\left[y(s)\delta y(t) + y(t)\delta y(s)\right] dsdt$$

$$= \iint_{[a,b]^2} K(s,t)y(s)ds\delta y(t)dt + \iint_{[a,b]^2} K(t,s)y(s)ds\delta y(t)dt$$

$$\implies \frac{\delta J}{\delta y} = \int_a^b \left[K(s,t) + K(t,s)\right] y(s)ds$$

Problem . Find the extremals of the functional

$$\int \sqrt{x^2+y^2}\sqrt{1+y'^2}dx$$

We use polar coordinates to obtain a new integral:

$$\int \sqrt{x^2+y^2}\sqrt{1+y'^2}dx = \int r\sqrt{1+\frac{r'\sin\theta + r\cos\theta}{r'\cos\theta - r\sin\theta}}\left(r'\cos\theta - r\sin\theta\right) d\theta$$

$$= \int r\sqrt{(r'\cos\theta - r\sin\theta)^2 + (r'\sin\theta + r\cos\theta)^2}\ d\theta$$

$$= \int r\sqrt{r'^2\cos^2\theta - 2rr'\cos\theta\sin\theta + r^2\sin\theta + r'^2\sin\theta + 2rr'\cos\theta\sin\theta + r^2\cos\theta}\ d\theta$$

$$= \int r\sqrt{r^2 + r'^2}\ d\theta$$

The newly simplified Lagrangian does not depend on θ, therefore we know

$$L - \frac{\partial L}{\partial r'}r' = C$$

This is equivalent to

$$\frac{r(r^2 + r'^2) - rr'^2}{\sqrt{r^2 + r'^2}} = \frac{r^3}{\sqrt{r^2 + r'^2}} = C$$

Which simplifies to

$$r' = r\sqrt{\left(\frac{r}{\sqrt{C}}\right)^4 - 1}$$

Let $u = \frac{r}{\sqrt{C}}$. Then $\frac{du}{d\theta} = \frac{du}{dr}\frac{dr}{d\theta} = \frac{1}{\sqrt{C}}\sqrt{u^4 - 1} = u\sqrt{u^4 - 1}$. We can integrate,

$$\int \frac{1}{u\sqrt{u^4 - 1}} du = \theta + D$$

The integral evaluates to

$$\frac{1}{2}\arctan\sqrt{u^4 - 1} + D = \frac{1}{2}\arctan\sqrt{\frac{r^4}{C^2} - 1} + D = \theta$$

Where C, D are arbitrary constants. We can continue to evaluate the equation,

$$2\theta - 2D = \arctan\sqrt{\left(\frac{r}{\sqrt{C}}\right)^4 - 1} \implies \tan(2\theta - 2D) = \sqrt{\left(\frac{r}{\sqrt{C}}\right)^4 - 1} \implies \frac{\sin^2 2(\theta - D)}{\cos^2 2(\theta - D)} = \frac{r^4 - C^2}{C^2}$$

$$\implies C^2\sin^2 2(\theta - D) = r^4\cos^2 2(\theta - D) - C^2\cos^2 2(\theta - D)$$

$$\implies C^2\sin^2 2(\theta - D) = (r^4 - C^2)(1 - \sin^2 2(\theta - D))$$