Calculus of Variations

Valeriy Slastikov & Georgy Kitavtsev **Spring, 2017**

1 1D Calculus of Variations.

We are going to study the following general problem:

- Minimize functional
$$I(u) = \int_a^b f(x, u(x), u'(x))\, dx$$
subject to boundary conditions $u(a) = \alpha,\ u(b) = \beta$.

In order to correctly set up this problem we have to assume certain properties of $f(x, u, \xi)$ and a function $u(x)$. Since we know how to integrate continuous functions it is reasonable to assume that function $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous. In fact in this section we will need stronger assumption that $f \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$, meaning that f, f' and f'' are continuous functions.

Concerning function $u : [a, b] \to \mathbb{R}$ we may assume that $u \in C^1([a, b])$ and $u(a) = \alpha$, $u(b) = \beta$. It seems reasonable since minimization problem involves $u'(x)$ (that needs to be continuous since we know how to integrate continuous functions) and boundary conditions. In fact in this section we will use a bit stronger regularity assumption $u \in C^2([a, b])$.

Note: we can change boundary conditions by removing constraints at one or both ends. In general, one can not prescribe u' at the end points (this will be explained later).

We are ready to set up the problem with mathematical precision:
Define
$$I(u) \equiv \int_a^b f(x, u(x), u'(x))\, dx, \tag{1}$$
where $f \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$, $u \in C^1([a, b])$.

Problem:
$$\inf_{u \in X} I(u), \tag{2}$$
where $X = \{u \in C^1([a, b]) : u(a) = \alpha,\ u(b) = \beta\}$.

We want to find a function $u_0(x)$ such that $u_0 \in X$ and u_0 delivers minimum to the functional $I(u)$ (we call such function a *minimizer*). In order to do this we need to understand what it means to minimize the functional $I(u)$.

Definition 1.1 *Function $u_0 \in X$ is a minimizer (or global minimizer) of the functional $I(u)$ on a set X if*
$$I(u_0) \le I(u) \quad \text{for all } u \in X.$$

Definition 1.2 *Function $u_0 \in X$ is a local minimizer of the functional $I(u)$ if the following holds for some open set $\mathcal{N} \ni u_0$*
$$I(u_0) \le I(u) \quad \text{for all } u \in \mathcal{N} \subset X.$$

In general it is very difficult to solve (2), however one can try to characterize minimizers. Let's consider a simple example from Calculus.

Example 1.1 *Let $g \in C^1([a,b])$ and we want to find a minimizer of g. By well known result of Analysis we know that any continuous function on a compact interval (closed and bounded) achieves its minimum (and maximum). Obviously the minimizer is just some point $x_0 \in [a,b]$. In order to find it we have to use the definition, i.e.*

$$g(x_0) \leq g(x) \quad \text{for all } x \in [a,b].$$

Assume that $x_0 \in (a,b)$ then for any $x \in (a,b)$ we have

$$\frac{f(x)-f(x_0)}{x-x_0} \geq 0, \quad \text{if } x > x_0$$

and

$$\frac{f(x)-f(x_0)}{x-x_0} \leq 0, \quad \text{if } x < x_0.$$

Taking a limit $x \to x_0$ we obtain $f'(x_0) \geq 0$ and $f'(x_0) \leq 0$. Therefore we have $f'(x_0) = 0$. **Note**: we must assume $x_0 \in (a,b)$ in order to obtain $f'(x_0) = 0$, explain it.
In order to find a minimizer we must find all points where $f'(x) = 0$ and compare values at these points with values at the ends of the interval ($f(a)$, $f(b)$).

In order to characterize minimizers of $I(u)$ we proceed in the similar way. We know that if u_0 is a minimizer then $I(u_0) \leq I(u)$ for all $u \in X$. Take any function $h \in C_0^1([a,b])$, obviously $u_0 + th \in X$ and therefore

$$I(u_0) \leq I(u_0 + th) \quad \text{for all } h \in C_0^1([a,b]).$$

We can define a function $F(t) = I(u_0 + th)$ and notice that since u_0 is a minimizer of $I(u)$ then $t = 0$ is a minimizer of $F(t)$ on $\mathbb{R}$. Therefore we must have $F'(0) = 0$ (see example).
Note: we a priori assumed existence of minimizer $u_0 \in X$ for $I(u)$ and just trying to characterize it **if it exists**. *In the example above for 1D function existence was guaranteed by a simple result of Analysis (look it up).*

We characterize minimizers in the following theorem

Theorem 1.3 *Let $f \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$ ($f \equiv f(x,u,\xi)$) and*

$$\inf_{u \in X} \left\{ I(u) \equiv \int_a^b f(x, u(x), u'(x))\, dx \right\} = m, \tag{3}$$

where $X = \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\ u(b) = \beta\}$. Then

1. *If (3) admits a minimizer $u_0 \in X \cap C^2([a,b])$ then*

$$\frac{d}{dx}\left(f_\xi(x, u_0(x), u_0'(x))\right) = f_u(x, u_0(x), u_0'(x)) \quad x \in (a,b) \tag{4}$$

$$u_0(a) = \alpha,\ u_0(b) = \beta.$$

2. *If some function $u_* \in X \cap C^2([a,b])$ satisfies (4) and if $(u,\xi) \to f(x,u,\xi)$ is convex (as a function of two variables) for every $x \in [a,b]$ then u_* is a minimizer of (3).*

Proof. 1) As we discussed before, if $u_0 \in X$ is a minimizer of $I(u)$ then $F'(0) = 0$ (see the definition of F above). Let's find $F'(0)$: by definition of derivative

$$F'(0) = \lim_{t\to 0} \frac{F(t) - F(0)}{t} = \lim_{t\to 0} \frac{I(u_0 + th) - I(u_0)}{t}.$$

It's not difficult to see that

$$\frac{I(u_0 + th) - I(u_0)}{t} = \int_a^b \frac{f(x, u_0(x) + th(x), u_0'(x) + th'(x)) - f(x, u_0(x), u_0'(x))}{t}\, dx. \tag{5}$$

Passing to the limit and using standard Analysis results to interchange integral and limit (for instance Dominated Convergence Theorem) we obtain

$$F'(0) = \int_a^b \frac{d}{dt} f(x, u_0(x) + th(x), u_0'(x) + th'(x))\big|_{t=0}\, dx.$$

Using standard rules of differentiation we have

$$\frac{d}{dt} f(x, u_0(x) + th(x), u_0'(x) + th'(x))\big|_{t=0} = f_u(x, u_0(x), u_0'(x))h(x) + f_\xi(x, u_0(x), u_0'(x))h'(x). \tag{6}$$

Therefore if $u_0 \in X$ is a minimizer of $I(u)$ we must have

$$\int_a^b f_u(x, u_0(x), u_0'(x))h(x) + f_\xi(x, u_0(x), u_0'(x))h'(x)\, dx = 0$$

for all $h \in C_0^1([a, b])$. This is a weak form of Euler-Lagrange equation. It is not clear how to deal with this integral identity and therefore we want to obtain a pointwise identity. We can use integration by parts to obtain

$$\int_a^b f_\xi(x, u_0(x), u_0'(x))h'(x)\, dx$$
$$= -\int_a^b \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right) h(x)\, dx + f_\xi(x, u_0(x), u_0'(x))h(x)\big|_a^b. \tag{7}$$

Recalling that $h(a) = h(b) = 0$ we have

$$\int_a^b \left[f_u(x, u_0(x), u_0'(x)) - \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right) \right] h(x)\, dx = 0$$

for all $h \in C_0^1([a, b])$. Using Fundamental Lemma of Calculus of Variations (proved after this theorem) we obtain

$$f_u(x, u_0(x), u_0'(x)) - \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right) = 0 \quad x \in (a, b).$$

Conditions $u_0(a) = \alpha$, $u_0(b) = \beta$ follow from the fact $u_0 \in X$.

2) Let $u_* \in X \cap C^2([a, b])$ is a solution of (4). Since $f(x, u, \xi)$ is convex in last two variables we have

$$f(x, u(x), u'(x)) \geq f(x, u_*(x), u_*'(x)) + f_u(x, u_*(x), u_*'(x))(u(x) - u_*(x))$$
$$+ f_\xi(x, u_*(x), u_*'(x))(u'(x) - u_*'(x)) \tag{8}$$

for every $u \in X$ (this inequality will be shown later for a convex function of one variable, extend it). We can integrate both parts of this inequality over (a, b) to obtain

$$\int_a^b f(x, u(x), u'(x))\,dx \geq \int_a^b f(x, u_*(x), u_*'(x))\,dx + \int_a^b f_u(x, u_*(x), u_*'(x))(u(x) - u_*(x))\,dx + \int_a^b f_\xi(x, u_*(x), u_*'(x))(u'(x) - u_*'(x))\,dx. \quad (9)$$

Using integration by parts and noting that $u - u_* \in C_0^1([a, b])$ we obtain

$$I(u) \geq I(u_*) + \int_a^b \left[f_u(x, u_*(x), u_*'(x)) - \frac{d}{dx}\left(f_\xi(x, u_*(x), u_*'(x))\right)\right](u(x) - u_*(x))\,dx. \quad (10)$$

Since u_* satisfies (4) we have $I(u) \leq I(u_*)$ for all $u \in X$.
Theorem is proved.

Remark. Note that we did not prove existence of minimizer in this theorem. In the first statement we showed that if u_0 is a minimizer then it solves Euler-Lagrange equation. In the second statement we showed that if u_* solves Euler-Lagrange equation and **the function f is convex in last two variables** then u_* is a minimizer. But we did not prove existence of solution for Euler-Lagrange equation and therefore even the case of convex f we don't know if there is a minimizer.

Now we prove Fundamental Lemma of Calculus of Variations (FLCV)

Lemma 1.1 *Let $g \in C([a, b])$ and*

$$\int_a^b g(x)\eta(x)\,dx = 0 \quad \text{for all } \eta \in C_0^\infty([a, b]),$$

then $g(x) = 0$ on $[a, b]$.

Proof. Suppose there exists a point $c \in (a, b)$ such that $g(c) \neq 0$. We may assume without loss of generality that $g(c) > 0$. By continuity of function $g(x)$ there exists an open interval $(s, t) \subset (a, b)$ containing point c such that $g(x) > 0$ on (s, t). Taking $\eta(x)$ to be positive on (s, t) with $\eta(s) = \eta(t) = 0$ we obtain the contradiction with $\int_a^b g(x)\eta(x)\,dx = 0$. Therefore $g(x) = 0$ for all $x \in (a, b)$. By continuity of g we obtain $g(x) = 0$ on $[a, b]$.
Lemma is proved.

For convenience we also prove a simple fact that for $g \in C^1([a, b])$ convexity is equivalent to

$$g(x) \geq g(y) + g'(y)(x - y) \quad \text{for all } x, y \in (a, b).$$

Proof. Let g be convex then by definition

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y)$$

for all $\alpha \in (0, 1)$ and x, y in (a, b). After simple rearrangement we have

$$\frac{g(y + \alpha(x - y)) - g(y)}{\alpha} \leq g(x) - g(y).$$

Taking a limit as $\alpha \to 0$ we obtain

$$g'(y)(x-y) \leq g(x) - g(y).$$

We proved it in one direction. Now we assume

$$g(x) \geq g(y) + g'(y)(x-y) \quad \text{for all } x, y \in (a,b).$$

Using this we have

$$g(x) \geq g(\alpha x + (1-\alpha)y) + (1-\alpha)g'(\alpha x + (1-\alpha)y)(x-y)$$

and

$$g(y) \geq g(\alpha x + (1-\alpha)y) - \alpha g'(\alpha x + (1-\alpha)y)(x-y).$$

Multiplying first inequality by α, second inequality by $1-\alpha$ and adding them we obtain

$$\alpha g(x) + (1-\alpha)g(y) \geq g(\alpha x + (1-\alpha)y).$$

Result is proved.

Note that this fact is true also for functions defined on $\mathbb{R}^n$.

Let's consider several examples.

Example 1.2 (Brachistochrone problem) *In this example we have*

$$f(u,\xi) = \sqrt{\frac{1+\xi^2}{u}},$$

and $X = \{u \in C^1([0,1]) \, : \, u(0) = 0, \, u(1) = \beta, \, u(x) > 0 \text{ for } x \in (0,1)\}$. We would like to find solutions of Euler-Lagrange equation.

Using theorem 1.3 we can easily find Euler-Lagrange equations

$$\left(\frac{u'}{\sqrt{u[1+(u')^2]}}\right)' = -\sqrt{\frac{1+(u')^2}{4u^3}}.$$

We can multiply both parts by $\frac{u'}{\sqrt{u[1+(u')^2]}}$ and obtain

$$\left(\frac{u'}{\sqrt{u[1+(u')^2]}}\right)' \frac{u'}{\sqrt{u[1+(u')^2]}} = -\frac{u'}{2u^2}.$$

It is clear that this is equivalent to

$$\left(\frac{(u')^2}{u[1+(u')^2]}\right)' = \left(\frac{1}{u}\right)'.$$

Now we obtain

$$\frac{1}{u} = \frac{(u')^2}{u[1+(u')^2]} + C$$

and simplifying this expression we obtain

$$u[1+(u')^2] = 2\mu,$$

where $\mu > 0$ is some constant that we can find from boundary conditions. Separating variables we obtain

$$dx = \sqrt{\frac{u}{2\mu - u}}du.$$

Now we integrate both parts

$$x = \int_0^u \sqrt{\frac{y}{2\mu - y}}dy.$$

We can make a natural substitution $y = \mu(1-\cos\theta)$ to obtain $x = \mu(\theta - \sin\theta)$. Therefore solution to this problem can be represented in the following parametric form:

$$u(\theta) = \mu(1-\cos\theta),\ x(\theta) = \mu(\theta - \sin\theta).$$

It is clear that $u(0) = 0$ so that first boundary condition is satisfied and we can find μ by applying second boundary condition.
Note that we don't know if this solution is a minimizer of the problem. For this we have to prove either existence of minimizer or convexity of $f(u,\xi)$. However from physical model we can see that there must be a solution here (therefore minimizer should exist) and since solution to the problem is unique it must be a minimizer (I guess that was the reasoning in 17-th century).

Example 1.3 (Minimal surface of revolution) *In this example we have*

$$f(u,\xi) = 2\pi u\sqrt{1+\xi^2},$$

and $X = \{u \in C^1([0,1]) : u(0) = \alpha, u(1) = \beta, u(x) > 0\}$. We would like to find solutions of Euler-Lagrange equation.

Again we can use the theorem 1.3 to obtain

$$\left(\frac{uu'}{\sqrt{1+(u')^2}}\right)' = \sqrt{1+(u')^2}.$$

Multiplying both parts by $\frac{uu'}{\sqrt{1+(u')^2}}$ and integrating we obtain

$$\frac{u^2}{1+(u')^2} = C,\ \text{or}\ (u')^2 = \frac{u^2}{a^2} - 1$$

for some constant $a > 0$. We can separate variables to obtain the solution here but there is a better way: we search for a solution in the following form

$$u(x) = a\cosh\frac{f(x)}{a}.$$

Plugging this into equation we obtain $[f'(x)]^2 = 1$ and therefore either $f(x) = x+\mu$ or $f(x) = -x+\mu$. Since $\cosh x$ *is even function and μ is any constant we have*

$$u(x) = a\cosh\frac{x+\mu}{a}.$$

Continue to find the solutions of this problem assuming $u(0) = u(1) = \alpha > 0$. The number of solutions will depend on α.

Example 1.4 (Convexity) *In this example we assume that* $f(\xi)$ *is convex on* $\mathbb{R}$ *and* $X = \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\, u(b) = \beta,\, \}$. *We will prove existence of minimizer for*

$$I(u) = \int_a^b f(u'(x))\,dx$$

in this case and find it explicitly. By convexity of f *we know that*

$$f(x) \geq f(y) + f'(y)(x-y).$$

We can define $u_0(x) = \frac{\beta-\alpha}{b-a}(x-a) + \alpha$. *Using above inequality we see that for any* $u \in X$

$$f(u'(x)) \geq f(u_0'(x)) + f'(u_0'(x))(u'(x) - u_0'(x)).$$

It is clear that $u_0'(x) \equiv \frac{\beta-\alpha}{b-a}$ *and therefore integrating both parts over* (a,b) *we obtain*

$$\int_a^b f(u'(x))\,dx \geq \int_a^b f(u_0'(x))\,dx.$$

Therefore u_0 *is a minimizer of* $I(u)$.

Now we give several examples of nonexistence.

Example 1.5 (Non-convexity implies no minimizer) *We consider* $f(\xi) = e^{-\xi^2}$ *and* $X = \{u \in C^1([0,1]) \,:\, u(0) = 0,\, u(1) = 0,\, \}$ *and want to show that for the problem*

$$\inf_X \int_0^1 f(u'(x))\,dx$$

there is no solution.
We see that Euler-Lagrange equation is

$$\frac{d}{dx} f'(u') = 0.$$

And therefore $u' = const$ *that implies in this case (using boundary conditions) that* $u(x) \equiv 0$. *It is clear that this is a maximizer of the problem, not a minimizer. Moreover, if we take*

$$u_n(x) = n(x - \frac{1}{2})^2 - \frac{n}{4}$$

we have $u_n \in X$ *for any* $n \in \mathbb{N}$. *Calculating* $I(u_n)$ *we obtain*

$$I(u_n) = \int_0^1 e^{-4n^2\left(x-\frac{1}{2}\right)^2}\,dx = \frac{1}{2n}\int_{-n}^{n} e^{-x^2}\,dx \to 0,$$

as $n \to \infty$. *Therefore* $\inf_X I(u) = 0$ *but obviously no function can deliver this infimum and so minimizer does not exist.*

Example 1.6 (Not smooth minimizers) *Consider* $f(\xi) = (\xi^2 - 1)^2$, $X = \{u \in C^1([0,1]) \,:\, u(0) = 0,\, u(1) = 0,\, \}$. *Show that there are no minimizers for this problem. However there are infinitely many minimizers that are piecewise smooth.*

1.1 Problems with constraints - Lagrange multipliers

In many problems arising in applications there exist some integral constraints on the set of functions where we look for a minimizer. One basic example is Catenary problem where we have to find the shape of the hanging chain of fixed length (see example below). In order to characterize the minimizer of the problem (2) with integral constraint

$$G(u) \equiv \int_a^b g(x, u(x), u'(x))\, dx = 0 \tag{11}$$

we can not simply take a variation of the functional $I(u)$ with respect to **any** function $h \in C_0^1([a, b])$ since we can violate this integral constraint. Therefore we have to invent a method that allows us to derive Euler-Lagrange equation in this situation. This method is called Method of Lagrange Multipliers. Let's state the problem that we are solving here with mathematical precision.

We investigate minimizers of the following problem

$$\inf_{u \in Y} \int_a^b f(x, u(x), u'(x))\, dx, \tag{12}$$

where

$$Y = X \cap \{u \in C^1([a, b]) \,:\, G(u) \equiv \int_a^b g(x, u(x), u'(x))\, dx = 0\}.$$

We can prove the following theorem

Theorem 1.4 *Let $f, g \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$ and u_0 be a minimizer of the problem (12) above such that*

$$\int_a^b g_\xi(x, u_0(x), u_0'(x))w'(x) + g_u(x, u_0(x), u_0'(x))w(x)\, dx \neq 0 \tag{13}$$

for some $w \in C_0^1([a, b])$. Then u_0 satisfies Euler-Lagrange equations corresponding to the following problem

$$\inf_{(u,\lambda) \in X \times \mathbb{R}} I(u) - \lambda G(u). \tag{14}$$

Proof. Let u_0 be a minimizer of (12). We would like to take a variation o $I(u)$ preserving constraint $G(u) = 0$. By (13) (rescaling if needed) we know that there exists a function $w \in C_0^1([a, b])$ such that

$$\int_a^b g_\xi(x, u_0(x), u_0'(x))w'(x) + g_u(x, u_0(x), u_0'(x))w(x)\, dx = 1. \tag{15}$$

Let $v \in C_0^1([a, b])$ be any and w be as above, we define eeewsddq4

$$F(\varepsilon, h) = I(u_0 + \varepsilon v + hw) = \int_a^b f(x, u_0(x) + \varepsilon v(x) + hw(x), u_0'(x) + \varepsilon v'(x) + hw'(x))\, dx \tag{16}$$

and

$$G(\varepsilon, h) = G(u_0 + \varepsilon v + hw) = \int_a^b g(x, u_0(x) + \varepsilon v(x) + hw(x), u_0'(x) + \varepsilon v'(x) + hw'(x))\, dx. \tag{17}$$

We know that $G(0,0) = G(u_0) = 0$ (since u_0 satisfies constraint (11)) and $G_h(0,0) = 1$ (since u_0 satisfies (15)). Using Implicit Function Theorem we can deduce that there exists ε_0 and a continuously differentiable function $t(\varepsilon)$ such that $t(0) = 0$ and

$$G(\varepsilon, t(\varepsilon)) = 0$$

for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. Therefore we know that for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ we have

$$u_0 + \varepsilon v + t(\varepsilon) w \in Y.$$

Now we can take a variation of $I(u)$. We consider $F(\varepsilon, t(\varepsilon)) = I(u_0 + \varepsilon v + t(\varepsilon)w)$, by the same arguments as before since u_0 is a minimizer of $I(u)$ and since $F(0,0) = I(u_0)$ we know that 0 is a minimizer of function $F(\varepsilon, t(\varepsilon))$ and therefore

$$\frac{d}{d\varepsilon} F(\varepsilon, t(\varepsilon))\Big|_{\varepsilon=0} = 0.$$

Computing this derivative we obtain

$$F_\varepsilon(0,0) + F_h(0,0)t'(0) = 0.$$

It's not difficult to see that

$$F_h(0,0) = \int_a^b f_\xi(x, u_0(x), u_0'(x))w'(x) + f_u(x, u_0(x), u_0'(x))w(x)\, dx,$$

and since u_0 and w are fixed functions $F_h(0,0)$ is just a constant that we denote as $\lambda \equiv F_h(0,0)$. We have to understand what is $t'(0)$ and in order to do this we use the fact that $G(\varepsilon, t(\varepsilon)) = 0$ for all $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ and compute

$$0 = \frac{d}{d\varepsilon} G(\varepsilon, t(\varepsilon))\Big|_{\varepsilon=0} = G_\varepsilon(0,0) + G_h(0,0)t'(0).$$

Since $G_h(0,0) = 1$ then $t'(0) = -G_\varepsilon(0,0)$. Therefore we have that if u_0 is a minimizer of $I(u)$ then

$$F_\varepsilon(0,0) - \lambda G_\varepsilon(0,0) = 0.$$

We may rewrite this as

$$\int_a^b f_u(x, u_0(x), u_0'(x))v(x) + f_\xi(x, u_0(x), u_0'(x))v'(x)\, dx - \lambda \int_a^b g_\xi(x, u_0(x), u_0'(x))v'(x) + g_u(x, u_0(x), u_0'(x))v(x)\, dx = 0 \quad (18)$$

for all $v \in C_0^1([a,b])$. Using integration by parts, FLCV and recalling that $G(u_0) = 0$ we obtain

$$\frac{d}{dx}\left(f_\xi(x, u_0(x), u_0'(x))\right) - f_u(x, u_0(x), u_0'(x)) = \lambda\left(\frac{d}{dx}\left(g_\xi(x, u_0(x), u_0'(x))\right) - g_u(x, u_0(x), u_0'(x))\right) \quad x \in (a,b) \quad (19)$$

$$\int_a^b g(x, u_0(x), u_0'(x))\, dx = 0$$

This is exactly Euler-Lagrange equation for the problem (14).
Theorem is proved.

Remark. Think about why we need condition with w.

Example 1.7 (Catenary problem) *Solve this problem.*

2 Second Variation and Stability

In the previous section we introduced a minimization problem and learned a method to find critical points of the functional $I(u)$ using Euler-Lagrange equations. However we still don't know anything about how to solve the original minimization problem. In particular we don't understand if the critical points that we find by solving Euler-Lagrange equations are minimisers or not, i.e. if they "stable" or "unstable" (in a certain sense). This section is devoted to investigating stability of the critical point of $I(u)$ using its second variation.

2.1 Various Derivatives

In this section we start with introduction for the functional (1) an analog of Taylor expansion (for a function of two variables) and subsequently of two different notions of functional derivatives, namely *Gateaux and Frechet*.

Let us assume that $f \in C^3([a,b] \times \bar{\mathbb{R}} \times \bar{\mathbb{R}})$ (in particular, its third order partial derivatives $f_{uuu},\ f_{uu\xi},\ f_{u\xi\xi},\ f_{\xi\xi\xi}$ are continuous and bounded functions on $[a,b] \times \mathbb{R} \times \mathbb{R}$) and recall Peano's form of it's Taylor expansion at a point (x,u,ξ):

$$\begin{aligned} f(x, u+\phi, \xi+h) &= f(x,u,\xi) + (f_u\phi + f_\xi h) + \frac{1}{2}(f_{uu}\phi^2 + 2f_{uh}\phi h + f_{\xi\xi}h^2) \\ &+ \frac{1}{6}(\bar{f}_{uuu}\phi^3 + 3\bar{f}_{uuh}\phi^2 h + 3\bar{f}_{uhh}\phi h^2 + \bar{f}_{\xi\xi\xi}h^3). \end{aligned} \tag{20}$$

In the last expression derivatives without bars are evaluated at the point (x,u,ξ), and derivatives with bars are evaluated at an intermediate point $(x, u+\theta_1\phi, \xi+\theta_2 h)$ for some numbers $0 \le \theta_1,\ \theta_2 \le 1$.

Substitution of expansion (20) into the functional $I(u)$ in (1) and taking

$$\xi(x) = u'(x),\ h(x) = \phi'(x) \quad \text{for} \quad x \in [a,b] \quad \text{with} \quad \phi(a) = \phi(b) = 0$$

results in the following expansion of functional $I(u)$ at point u:

$$\begin{aligned} I(u+\phi) - I(u) &= \int_a^b f_u\phi + f_\xi\phi'\,dx + \frac{1}{2}\int_a^b f_{uu}\phi^2 + 2f_{u\xi}\phi\phi' + f_{\xi\xi}|\phi'|^2\,dx \\ &+ \frac{1}{6}\int_a^b (\bar{f}_{uuu}\phi^3 + 3\bar{f}_{uuh}\phi^2\phi' + 3\bar{f}_{uhh}\varphi|\phi'|^2 + \bar{f}_{\xi\xi\xi}(\phi')^3)\,dx \\ &=: \delta I(u)[\phi] + \frac{1}{2}\delta^2 I(u)[\phi] + R(u,\phi), \end{aligned} \tag{21}$$

Above we have introduced three new functionals, which depend on point u:

$$\begin{aligned} \delta I(u)[\phi] &:= \int_a^b f_u\phi + f_\xi\phi'\,dx, \\ \delta^2 I(u)[\phi] &:= \int_a^b (f_{uu}\phi^2 + 2f_{u\xi}\phi\phi' + f_{\xi\xi}|\phi'|^2)\,dx, \\ R(u,\phi) &:= \frac{1}{6}\int_a^b (\bar{f}_{uuu}\phi^3 + 3\bar{f}_{uuh}\phi^2\phi' + 3\bar{f}_{uhh}\phi|\phi'|^2 + \bar{f}_{\xi\xi\xi}(\phi')^3)\,dx. \end{aligned} \tag{22}$$

Next, it is easy to get the following estimates:

$$
\begin{aligned}
|\delta^2 I(u)[\phi]| &\leq \frac{1}{2}M\int_a^b (\phi^2 + 2|\phi\phi'| + \phi'^2)\,dx = \frac{1}{2}M\int_a^b (|\phi| + |\phi'|)^2\,dx \\
&\leq \frac{b-a}{2}M||\phi||^2_{C^1[a,b]},
\end{aligned} \tag{23}
$$

where $M := \max\{f_{uu}, f_{u\xi}, f_{\xi\xi}\}$ in $[a,b]\times\mathbb{R}\times\mathbb{R}$ (by assumptions on f the maximum exists) and as usually

$$||\phi||_{C^1[a,b]} := \max_{x\in[a,b]} \left(|\phi(x)| + |\phi'(x)|\right).$$

Analogously, it can be shown that

$$|R(u,\phi)| \leq \frac{b-a}{3}M_1||\phi||^3_{C^1[a,b]}, \tag{24}$$

where $M_1 := \max\{f_{uuu}, f_{uu\xi}, f_{u\xi\xi}, f_{\xi\xi\xi}\}$ in $[a,b]\times\mathbb{R}\times\mathbb{R}$.

Using expansion (21) of the functional $I(u)$ we can give a meaning to the following definition.

Definition 2.1 *Let $u \in C^1([a,b])$, $\phi \in C_0^1([a,b])$ and $I : C^1([a,b]) \to \mathbb{R}$ be a functional (see for instance (1)). We say that I is Frechet differentiable at u if there exists a linear functional $\delta I_F(u) : C_0^1([a,b]) \to \mathbb{R}$ (called a first Frechet variation/derivative of $I(u)$ at the point u) such that*

$$\lim_{||\phi||_{C^1([a,b])}\to 0} \frac{|I(u+\phi) - I(u) - \delta I_F(u)[\phi]|}{||\phi||_{C^1([a,b])}} = 0.$$

If the above limit exists for all $u \in C^1([a,b])$, we say that I is Frechet differentiable in $C^1([a,b])$.

Note that in the above definition you can change $C^1([a,b])$ and $\mathbb{R}$ to any Banach spaces X, Y. Note, that expansion (21) and estimates (23)–(24) imply together for functional (1) that its Frechet derivative $\delta I(u)[\phi]$ is indeed given by the first formula in (22).

Analogously, we introduce a definition of the second variation as follows.

Definition 2.2 *Functional $F(\phi,h)$ is called to be bilinear on $C^1[a,b]\times C^1[a,b])$ if $F(\cdot,h)$ is linear functional of ϕ for any fixed $h \in C^1[a,b])$ and also $F(\phi,\cdot)$ is linear functional of h for any fixed $\phi \in C^1[a,b])$. Given a bilinear functional $F(\phi,h)$ we can always define a quadratic functional $E[\phi]$ by a restriction:*

$$E(\phi) := F(\phi,\phi), \quad \textit{for all} \quad \phi \in C^1[a,b].$$

Note that according to the last definition functional $\delta^2 I(u)[\phi]$ in (22) is quadratic with the corresponding bilinear functional $F(u)$ given by

$$F(u)[\phi,h] = \int_a^b (f_{uu}\phi h + 2f_{u\xi}\phi\phi' + f_{\xi\xi}\phi h)\,dx.$$

Definition 2.3 *Let $u \in C^1([a,b])$, $\phi \in C_0^1([a,b])$ and $I : C^1([a,b]) \to \mathbb{R}$ be a functional (see for instance (1)). We say that I is twice Frechet differentiable at u if there exists $\delta I(u) : C^1([a,b]) \to \mathbb{R}$ from Definition 2.1 and* **a quadratic functional** *$\delta^2 I(u) : C_0^1([a,b]) \to \mathbb{R}$ (called as second Frechet variation/derivative of $I(u)$ at the point u) such that*

$$\lim_{||\phi||_{C^1([a,b])}\to 0} \frac{|I(u+\phi) - I(u) - \delta I_F(u)[\phi] - \frac{1}{2}\delta^2 I_F(u)[\phi]|}{||\phi||^2_{C_0^1([a,b])}} = 0.$$

If the above limit exists for all $u \in C^1([a,b])$ we say that I is twice Frechet differentiable in $C^1([a,b])$.

Again expansion (21) and estimate (24) imply together for functional (1) that thus defined $\delta^2 I(u)[\phi]$ is indeed given by the second formula in (22).

We end this section with introducing an alternative definition of the first and second variations which are easier to use in practice.

Definition 2.4 *Let $u \in C^1([a,b])$, $\phi \in C_0^1([a,b])$ and $I : C^1([a,b]) \to \mathbb{R}$ be a functional (see for instance* (1)*). We define Gateaux derivative of I at the point u in the direction ϕ as*

$$\delta I_G(u)[\phi] = \lim_{t \to 0} \frac{I(u+t\phi) - I(u)}{t}.$$

If the above limit exists for all $\phi \in C_0^1([a,b])$ we say that I is Gateaux differentiable in $C^1([a,b])$.

We already implicitly used Gateaux derivative in the previous section when we were taking first variation of the functional $I(u)$.

Definition 2.5 *Assuming $f \in C^2$ we define the second Gateaux variation of I at the point $u \in C^1([a,b])$ in the direction $\phi \in C_0^1([a,b])$ as*

$$\delta^2 I_G(u)(\phi) = \frac{d^2}{dt^2}\Big|_{t=0} I(u+t\phi).$$

Replacement of ϕ by $t\phi$ in the expansion (21) and differentiation with respect to t (once or twice) along with the subsequent limit $t \to 0$ give again the formulas for $\delta I(u)[\phi]$ and $\delta^2 I(u)[\phi]$ stated in (22). Therefore, we may conclude with the following statement.

Proposition 2.6 *Let $f \in C^3([a,b] \times \bar{\mathbb{R}} \times \bar{\mathbb{R}})$ then the first and the second variations of functional* (1) *in the sense of Gateaux and Frechet coincide and are given by*

$$\delta I(u)[\phi] := \int_a^b f_u \phi + f_\xi \phi' \, dx, \tag{25}$$

$$\delta^2 I(u)[\phi] := \int_a^b (f_{uu}\phi^2 + 2f_{u\xi}\phi\phi' + f_{\xi\xi}|\phi'|^2)\, dx. \tag{26}$$

Remark. Note that there is a difference between Gateaux and Frechet derivatives. In fact it's clear that existence of Frechet derivative implies existence of Gateaux derivative. The opposite, in general, is not true. We explain it in the next section.

2.2 Examples showing difference between Frechet and Gateaux derivatives

Definitions of Frechet and Gateaux derivatives can be easily transferred to the case of functions of two variables $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, in which case Frechet derivative coincides with the full differential of f, while Gateaux derivative coincides with derivatives in the direction of vector $\bar{\nu} \in \mathbb{R}$. In this case we know from Analysis course that:

Existence of Frechet derivative $\Longrightarrow$ Existence of Gateaux derivative $\Longrightarrow$ Existence of partial derivatives.

The following examples show that inverse implications are not true.

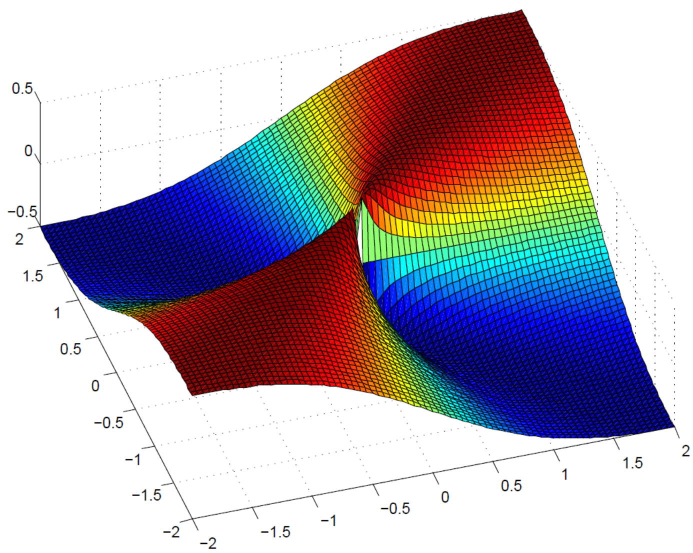

Figure 1: Function $f(x, y)$ from Example 1.

Example 1. Consider a function with its plot shown on Fig. 1

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

It is easy to check that partial derivatives f_x and f_y exist at $(0, 0)$, e.g.

$$|f_x| = \left| \frac{y}{x^2+y^2} - \frac{2x^2y}{(x^2+y^2)^2} \right| = \left| \frac{y(y^2-x^2)}{x^2+y^2} \right| \leq |y| \to 0 \quad \text{as} \quad \sqrt{x^2+y^2} \to 0.$$

But all other directional derivatives don't exist because $f(x, y)$ is discontinuous at $(0, 0)$. Therefore, both Gateaux and Frechet derivatives don't exist for $f(x, y)$.

Example 2. Consider function with its plot shown on Fig. 2

$$g(x,y) = \begin{cases} \frac{x^2y}{x^2+y^2}, & xy > 0 \\ 0 & xy \leq 0. \end{cases}$$

It is easy to check that $g(x, y)$ is continuous. Directional derivatives in the directions given by vectors $\nu_1 = [1, 1]$ and $\nu_2 = [1, -1]$ at $(0, 0)$ are

$$\begin{aligned} D_{\nu_1} g(0,0) &= \lim_{t \to 0} \left(\frac{t^3}{t^2+t^2} \right) / t = 1, \\ D_{\nu_2} g(0,0) &= \lim_{t \to 0} \frac{0}{t} = 0. \end{aligned}$$

Therefore, Gateaux derivative of g exists at $(0, 0)$ but it is no linear operator, because e.g.

$$D_{\nu_3} g(0,0) = 0 \neq D_{\nu_1} g(0,0) + D_{\nu_2} g(0,0),$$

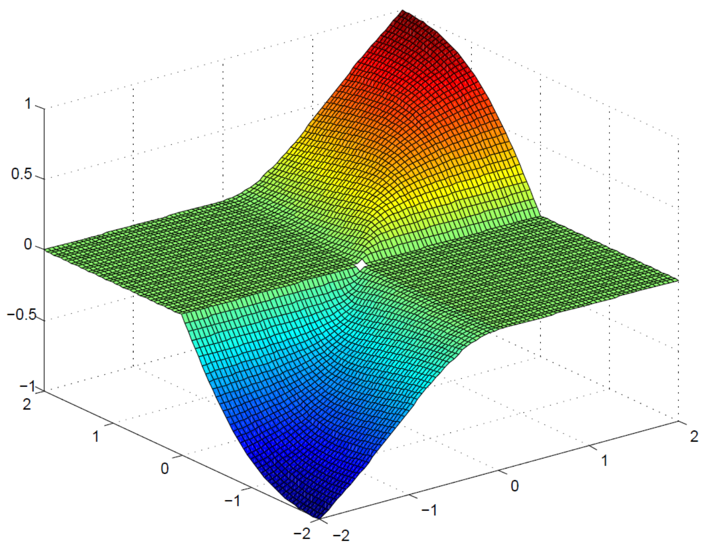

Figure 2: Function $g(x, y)$ from Example 2.

where $\nu_3 = \nu_1 + \nu_2 = [2, 0]$. Hence, Frechet derivative of g doesn't exist.

Example 3. Consider a function with its plot shown on Fig. 3

$$h(x, y) = \begin{cases} \frac{|x|^3 y}{x^4+y^2}, & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

This function is continuous and all directional derivatives of it at $(0, 0)$ are zero, i.e. its Gateaux derivative at $(0, 0)$ is given by 0 function. Nevertheless, if one compute the limit of the quotient

$$Q(x, y) = \frac{|h(x, y) - h(0, 0)|}{\sqrt{x^2 + y^2}}$$

along the curve $x = t$, $y = t^2$ one gets that $Q(t, t^2) \to 1$. This implies that Frechet derivative doesn't exist, because otherwise it would coincide with zero Gateaux derivative.

The final example shows the difference between notions of Gateaux and Frechet derivatives in infinite dimensional case of functionals $I : C^1[a, b] \to \mathbb{R}$. .

Example 4. Consider functional defined as

$$I(u) = \frac{\int_a^b u(x)\, dx}{u\left(\frac{a+b}{2}\right)}$$

It is easy to see that $I(u)$ is discontinuous at $u(x) \equiv 0$ and naturally its Frechet derivative doesn't exist. Nevertheless, one can formally calculate its Gateaux derivative in the direction ϕ with $\phi((a+b)/2) \neq 0$ as

$$\delta I_G(0)[\phi] = \frac{d}{dt}\left[\frac{\int_a^b t\phi(x)\, dx}{t\phi\left(\frac{a+b}{2}\right)}\right] = \frac{d}{dt}\left[\frac{\int_a^b \phi(x)\, dx}{\phi\left(\frac{a+b}{2}\right)}\right] = 0.$$

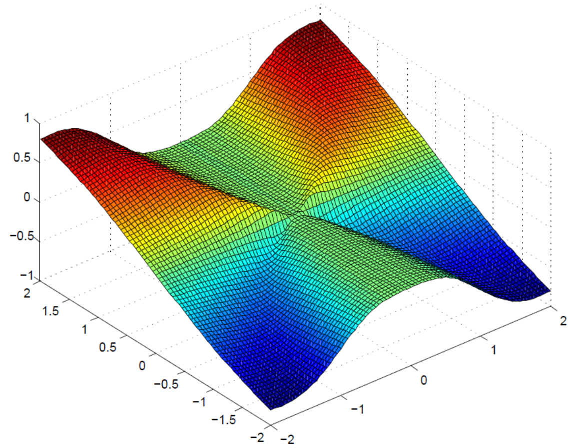

Figure 3: Function $h(x, y)$ from Example 3.

2.3 Computing Second Variation

Now we are ready to compute the second variation of the functional

$$I(u) = \int_a^b f(x, u(x), u'(x))\, dx$$

defined on $X = \{u \in C^1([a, b]), u(a) = \alpha, u(b) = \beta\}$ directly using Definition 2.5.

$$\delta^2 I(u)(\phi) = \int_a^b \left(\phi^2 f_{uu}(x, u, u') + 2\phi\phi' f_{u\xi}(x, u, u') + |\phi'|^2 f_{\xi\xi}\right) dx.$$

We can integrate this expression by parts taking into account $\phi \in C_0^1([a, b])$ to obtain

$$\delta^2 I(u)(\phi) = \int_a^b \left(\phi^2 \left[f_{uu}(x, u, u') - \frac{d}{dx} f_{u\xi}(x, u, u')\right] + |\phi'|^2 f_{\xi\xi}\right) dx.$$

The second variation is extremely important in the investigation of the stability and local minimality of the critical point of functional $I(u)$.

2.4 Second Variation and Stability

We know that if $f\ \mathbb{R} \to \mathbb{R}$ is twice differentiable function and x_0 is a critical point of f then $f'(x_0) = 0$. If we want to understand more about nature of x_0, i.e. if it is local minimum, local maximum or a saddle point we know from basic analysis that we have to investigate the sign of $f''(x_0)$. In particular

- if $f''(x_0) > 0$ then x_0 is local minimum, i.e. there exists $\delta > 0$ such that $f(x_0) \leq f(x)$ for all $x \in (x_0 - \delta, x_0 + \delta)$;
- if $f''(x_0) < 0$ then x_0 is local maximum, i.e. there exists $\delta > 0$ such that $f(x_0) \geq f(x)$ for all $x \in (x_0 - \delta, x_0 + \delta)$;

- if $f''(x_0) = 0$ then nothing definite can be said about properties of x_0 and in order to conclude you need to consider higher order derivatives.

Remark. Recall that we used analogy between 1D problems and infinite dimensional problem in the previous section when finding necessary conditions on the criticality of u_0. Think about how you can prove above statements using convexity (concavity) of f or Taylor expansion of f near x_0. Can you transfer your ideas to the infinite dimensional case?

We are interested in minimization problem and therefore we want to single out local minimizers (if we cannot find global minimizer). We already understand that local minimizers of $I(u)$ have to satisfy Euler-Lagrange equations. Now we want to show that local minimizers have to satisfy more restrictive condition.

Proposition 2.7 *Let* $u_0 \in X = \{u \in C^1([a,b]), u(a) = \alpha, u(b) = \beta\}$ *be a local minimizer of the functional* $I(u)$ *defined in* (1)*, i.e. there exists* $\delta > 0$ *such that* $I(u_0) \le I(u)$ *for all* $u \in X$ *such that* $\|u - u_0\|_{C^1} < \delta$. *Then*

$$\delta^2 I(u_0)(\phi) \ge 0 \quad \text{for all } \phi \in C_0^1([a,b]). \tag{27}$$

Proof Assume that there is $\phi_0 \in C_0^1([a,b])$ such that $\delta^2 I(u_0)(\phi_0) < 0$. Define function $F(t) = I(u_0 + t\phi_0)$. It's clear that $t = 0$ is a critical point of $F(t)$ and $F'(0) = 0$, $F''(0) < 0$. Now you can use 1D arguments to show that there exists $\delta > 0$ such that $F(0) > F(t)$ for all $t \in (t - \delta, t + \delta), t \neq 0$ and this is in contradiction with local minimality of u_0.

Based on this result we introduce several definitions

Definition 2.8 *The critical point* $u_0 \in X$ *of the functional* $I(u)$ *is called locally stable if* $\delta^2 I(u_0)(\phi) \ge 0$ *for all* $\phi \in C_0^\infty([a,b])$*; strictly locally stable if* $\delta^2 I(u_0)(\phi) > 0$ *for all* $\phi \in C_0^\infty([a,b])$*; and uniformly locally stable in space* Y *if* $\delta^2 I(u_0)(\phi) > \mu\|\phi\|_Y^2$ *for all* $\phi \in C_0^\infty([a,b])$.

Remark. Here space Y does not necessarily coincides with X and usually is much weaker than Y. The typical example for Y in our framework would be either $W_0^{1,p}(a,b)$ or $L^p(a,b)$ for some $p > 1$. In general, there is no hope to take $Y = C_0^1([a,b])$ and prove uniform positivity of the second variation for I. That's why you need to study Sobolev and L^p spaces if you want to do research in Calculus of Variations.

There is another necessary condition (Legendre condition) for local stability that directly follows from definition of local stability.

Proposition 2.9 *Let* $I(u)$ *be defined as usually (see* (1)*),* u_0 *is a critical point of* $I(u)$ *and* $\delta^2 I(u_0)(\phi) \ge 0$ *for all* $\phi \in C_0^\infty([a,b])$ *then*

$$f_{\xi\xi}(x, u_0(x), u_0'(x)) \ge 0 \quad \text{for all } x \in [a,b].$$

Proof Proof is a simple exercise. Try to use a mollifier in the neighborhood of a point x to show that $f_{\xi\xi}(x, u_0(x), u_0'(x)) \ge 0$.

2.5 Local Minimality

In the previous section we introduced the notion of local stability. However, in order to show local minimality it is in general not enough to have local stability.

Definition 2.10 *We say that $u_0 \in X$ is a local minimizer of $I(u)$ if there exists $\delta > 0$ such that $I(u_0) \leq I(u)$ for all $u \in X$ such that $\|u - u_0\|_X < \delta$.*

Next example shows that condition (27) is not sufficient for a critical point to be a minimizer.

Example 1. Define a functional

$$I(u) = \int_0^1 u^3 \, dx.$$

$u(x) \equiv 0$ is a critical point of this functional and

$$\delta^2 I(0)[\phi] = 0, \quad \text{for all } \phi \in C_0^1([a,b],$$

but $u(x)$ is not a local minimizer. Hint for the proof: calculate the reminder term $R(0, \phi)$ in the expansion (21)-(22) and show that it is negative when $\phi(x) < 0$. In fact, $I(u)$ doesn't have a minimizer and for a sequence $u_n \equiv -n$ one has

$$I(u_n) \to -\infty.$$

Next example shows that in contrast to the case of functions of several variables condition

$$\delta^2 I(u)(\phi) > 0, \quad \text{for all } \phi \in C_0^1([a,b],$$

is still not sufficient for a critical point u to be a local minimizer of functional I.

Example 2. Define a functional

$$I(u) = \int_0^1 \left(\frac{2}{3} u^3 - x^2 u^2 \right) dx.$$

It is easy to show that $u(x) = x^2$ is a critical point of this functional and

$$\delta^2 I(u)[\phi] = 2 \int_0^1 x^2 \phi^2 \, dx > 0 \quad \text{for} \quad \phi \neq 0 \tag{28}$$

We can also calculate the reminder term in the expansion (21)-(22) as

$$R(u, \phi) = 4 \int_0^1 \phi^3 \, dx \tag{29}$$

Let us take a sequence of functions $\{\phi_n\}$ defined as (see also Fig. 4):

$$\phi_n(x) = \begin{cases} -\frac{x}{n}, & 0 \leq x \leq \frac{1}{2n} \\ \frac{1}{n}(x - \frac{1}{n}) & \frac{1}{2n} < x \leq \frac{1}{n} \\ 0 & \frac{1}{n} < x \leq 1. \end{cases}$$

Substituting $\{\phi_n\}$ into the formulas (28) and (29) gives

$$R(u, \phi_n) = -\frac{1}{8n^7} \quad \text{and} \quad \delta^2 I(u)(\phi) = \frac{11}{240n^7}.$$

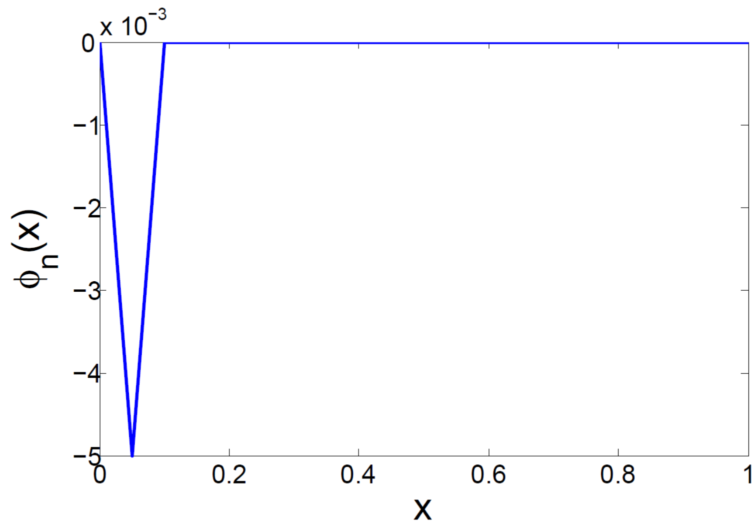

Figure 4: Function $\phi_n(x)$ with $n = 10$ from Example 4.

Therefore, for sufficiently large n one has

$$\delta^2 I(u)(\phi_n) + R(u, \phi_n) < 0 \tag{30}$$

and consequently, expansion (21)-(22) implies

$$I(u) > I(u + \phi_n). \tag{31}$$

The last inequality contradicts to Definition 2.10. Indeed, let us take $X = C[a,b]$ in Definition 2.10 with the corresponding norm

$$||\phi||_{C[a,b]} := \max_{x \in [a,b]} |\phi(x)|.$$

Then for any small $\delta > 0$ and $n > 1/(10\delta)$ one has

$$||\phi_n||_{C[a,b]} < \delta$$

but (31) holds. Therefore, we conclude that $u(x) = x^2$ is not a local minimizer of $I(u)$ in the space $X = C[a,b]$.

We highlight that structure of $\phi_n(x)$ in (30) was chosen in a special way in order to ensure that the reminder dominates the second variation and their sum has a negative sign. This stays in contrast to the case of function of one variable in which

$$f''(u)\phi^2 + f'''(u)\phi^3 > 0$$

holds for all sufficiently small ϕ when $f''(u) > 0$. In infinite dimensional case inequality (30) becomes possible because of the degenerate dependency of $\delta^2 I[u]\phi$ on x. Namely, when x is sufficiently small we can find test functions ϕ such that $x^2|\phi|^2 < 2|\phi|^3$, i.e. the reminder dominates point-wise the second variation.

Proposition 2.11 *Uniform local stability implies local minimality and, therefore, the former is indeed a sufficient condition.*

Proof The proof uses again the expansion (21)-(22). In order to satisfy Definition 2.10 we take $\delta = 3\mu/(2M_1(b-a))$, where M_1 is as in estimate (24) for the reminder. Then for all ϕ such that

$$||\phi||_{C_0^1[a,b]} < \delta \quad \text{and} \quad \phi \neq 0$$

one can estimate

$$I(u+\phi) - I(u) = \delta^2 I(u)[\phi] - R(u,\phi) > \frac{\mu}{2}||\phi||_{C_0^1[a,b]} > 0$$

and hence u is a local minimiser.

3 Important Examples

In this section we study several important examples that shows how to obtain the information on the second variation and how to use it to show local (and sometimes global) minimality.

Example 1. Let's define $I(u)$ in the following way

$$I(u) = \int_0^R \left(\frac{1}{2}|u'|^2 + \frac{1}{4}(u^2-1)^2 \right) dx.$$

We define set $X = \{u \in C^1([0,R]), u(0) = 0, u(R) = 1\}$ and want to solve the following problem

$$\inf_{u \in X} I(u).$$

At this point we know that in order to find a solution to this problem we have to find critical points by solving Euler-Lagrange equations. There is no guarantee that what we find is a minimizer but we will try to investigate its stability later.

Euler-Lagrange equations

It is clear that Euler-Lagrange equations are

$$u'' = u(u^2-1), \quad u(0) = 0, \; u(R) = 1.$$

If you want to find the solution here you can reduce this equation to a first order equation that you might be able to solve. Multiplying both parts of the above equation by u' and integrating from 0 to x we obtain

$$\int_x^R \frac{1}{2}\left(|u'(s)|^2\right)' ds = \int_x^R \frac{1}{4}\left((u^2(s)-1)^2\right)' ds$$

This gives us the following relation

$$\frac{1}{2}|u'(R)|^2 - \frac{1}{2}|u'(x)|^2 = \frac{1}{4}\left((u^2(R)-1)^2\right) - \frac{1}{4}\left((u^2(x)-1)^2\right).$$

Now we can use boundary data to obtain

$$|u'(x)|^2 = |u'(R)|^2 + \frac{1}{2}\left((u^2(x) - 1)^2\right).$$

We now have

$$u'(x) = \pm\sqrt{|u'(R)|^2 + \frac{1}{2}\left((u^2(x) - 1)^2\right)}$$

it's not difficult to show that $u'(x)$ cannot change sign (try to elaborate on it) and therefore

$$u'(x) = \sqrt{|u'(R)|^2 + \frac{1}{2}\left((u^2(x) - 1)^2\right)}.$$

This is a separable equation and one can obtain the solution as

$$x = \int_0^u \frac{du}{\sqrt{K + \frac{1}{2}\left(u^2 - 1\right)^2}},$$

where $K > 0$ should be determined from the boundary condition $u(R) = 1$.

Note, that you have obtained unique solution to the Euler-Lagrange equation and therefore if minimizer exists then it has to be the minimizer. One can show the existence of minimizer in this problem using Sobolev spaces and Direct Methods. Since we don't know anything about Sobolev space and Direct methods yet, we proceed by investigating stability of this solution.

Stability

In order to do this we have to study second variation of the functional $I(u)$. Straightforward computation gives us

$$\delta^2 I(u)(\phi) = \int_0^R |\phi'|^2 + (3u^2 - 1)\phi^2,$$

where $\phi \in C_0^\infty(0, R)$. We want to show that $\delta^2 I(u)(\phi) > 0$ for all nontrivial $\phi \in C_0^\infty(0, R)$. If you look at the expression for the second variation you notice that it is not clear if it is positive or negative unless you have extremely good information about behavior of solution u. At this poit we know that $u(x) > 0$ for $x > 0$ (try to understand why this is the case). Any other information about solution comes with a price of solving ODE (that in this particular case one can do) and using obtained information to estimate the sign of the integral. In general, obtaining some information about solution of the nonlinear second order boundary value problem is extremely difficult problem. Therefore we will use a nice trick, called Hardy trick.

The idea is to control nonlinearity through the Euler-Lagrange equations. Since we know that $u > 0$ we can represent any function $\phi \in C_0^\infty(0, R)$ as a product $\phi(x) = u(x)\psi(x)$ for a function $\psi \in C_0^\infty(0, R)$. It's clear that when ψ runs over the whole $C_0^\infty(0, R)$ ϕ also runs over the whole $C_0^\infty(0, R)$ and therefore we can safely take this new representation for ϕ. Now the second variation becomes

$$\delta^2 I(u)(\phi) = \int_0^R |(u\psi)'|^2 + (3u^2 - 1)(u\psi)^2 = \int_0^R |(u\psi)'|^2 + (u^2 - 1)(u\psi)^2 + 2u^4\psi^2,$$

An now we try to control the nonlinearity. We know that $u'' = u(u^2 - 1)$ and therefore we single out this nonlinearity in the equation and substitute u'' instead. The second variation becomes

$$\delta^2 I(u)(\phi) = \int_0^R |u'\psi + \psi' u|^2 + u'' u\psi^2 + 2u^4\psi^2.$$

After integrating by parts second term and basic algebra we obtain

$$\delta^2 I(u)(\phi) = \int_0^R |u'|^2\psi^2 + |\psi'|^2 u^2 + 2u'u\psi'\psi - |u'|^2\psi^2 - 2u'u\psi'\psi + 2u^4\psi^2 \tag{32}$$

$$= \int_0^R |\psi'|^2 u^2 + 2u^4\psi^2 > 0. \tag{33}$$

We have shown that second variation is strictly positive. Therefore we proved that u is strictly locally stable. Now we want to investigate local minimality. In general, to investigate local minimality one has to show uniform local stability in some space. However in this particular case we will be able to prove evenglobal minimality of u by direct calculation.

Global minimality

Let's take two energies $I(v)$ and $I(u)$, where u is our critical point and $v \in X$ is any function. We want to show that $I(v) > I(u)$. We first show that $I(u+\phi) - I(u) > 0$ for any $\phi \in C_0^\infty(0,R)$. Direct computation yeilds

$$I(u+\phi) - I(u) = \frac{1}{2}\int_0^R \left(|\phi'|^2 + (u^2-1)\phi^2\right) + \int_0^R \left(u^2\phi^2 + u\phi^3 + \frac{1}{4}\phi^4\right).$$

It is clear that the first integral is positive due to Hardy trick (supply details here) and the seconf integral is contains full square. Therefore we obtain $I(u+\phi) > I(u)$ for any nontrivial $\phi \in C_0^\infty(0,R)$. Now we can use the density arguments (supply details here, the idea is that C_0^∞ is dense in C_0^1, i.e. for any function w in C_0^1 there exists a sequence of functions $\{w_n\}$ in C_0^∞ such that $\|w_n - w\|_{C^1} \to 0$) to show that $I(u+\phi) > I(u)$ for any nontrivial $\phi \in C_0^1(0,R)$. Since $v - u \in C_0^1(0,R)$ we conclude that u is a global minimizer of I.

Let me stress again: we could have shown that u is a global minimizer by using direct methods and uniqueness of the critical point. However, in "real life problems" you rarely have solvable Euler-Lagrange equations and in most cases you don't know anything about uniqueness of solution for Euler-Lagrange equation. Therefore you usually show existence of minimizer by direct methods but then in order to investigate critical points of the functional $I(u)$ you have to study second variation.

Example 2.

We define $I(u) = \int_0^1 |u'|^2$ and would like to solve the following problem

$$\inf_{u\in X} I(u),$$

where $X = \left\{u \in C_0^1(0,1),\ \int_0^1 u^2 = 1\right\}$. It is clear that we have a problem with constraint and we want to reformulate it using the framework of Lagrange multipliers.

We always start with Euler-Lagrange equations and we know that in this case we should find Euler-Lagrange equations for the following problem

$$\inf_{(u,\lambda)\in C_0^1(0,1)\times\mathbb{R}} \int_0^1 |u'|^2 - \lambda\left(\int_0^1 u^2 - 1\right).$$

A simple calculation yields

$$u'' + \lambda u = 0, \quad u(0) = 0,\ u(1) = 0, \quad \int_0^1 u^2 = 1.$$

Solving differential equation we obtain $\lambda > 0$ (why this is the case?) and the general form of the solution

$$u(x) = A\cos(\sqrt{\lambda}x) + B\sin(\sqrt{\lambda}x)$$

Plugging in the boundary data we obtain

$$u(x) = A\sin(\sqrt{\lambda}x),\ \lambda = \pi^2 k^2,\ k \in \mathbb{N}.$$

In order to find constant A we use constraint $\int_0^1 u^2 = 1$ and it gives us $A = \sqrt{2}$. Now we found all critical points

$$u(x) = \pm\sqrt{2}\sin(\pi k x), \quad k \in \mathbb{N}.$$

We actually want to minimize $\int_0^1 |u'|^2$ and therefore we can plug in the critical points in this expression to obtain

$$\int_0^1 |u'|^2 = \pi^2 k^2.$$

It's clear that this expression is minimizer when $k = 1$ and therefore we found minimizers

$$u(x) = \pm\sqrt{2}\sin(\pi x).$$

Try to use Euler-Lagrange equations to show directly that

$$\lambda = \int_0^1 |u'|^2.$$

Example 3. (Geodesics on a surface)

We define a surface $S(x, y, z) = 0$ and want to find geodesics on this surface. By definition a geodesic curve between the points A and B is a curve that lies on the surface and has the minimal length among all curves connecting points A and B.

Assume the curve in $\mathbb{R}^3$ is given parametrically as $\mathbf{x}(t) = (x(t), y(t), z(t))$ then the distance between points $A = \mathbf{x}(t_0)$ and $B = \mathbf{x}(t_1)$ is given by

$$L(\mathbf{x}) = \int_{t_0}^{t_1} \sqrt{|x'(t)|^2 + |y'(t)|^2 + |z'(t)|^2}\, dt.$$

We also know that this curve has to lie on the surface S and therefore $S(x(t), y(t), z(t)) = 0$. We obtain the following minimization problem

$$\inf_{\mathbf{x} \in X} L(\mathbf{x}),$$

where $X = \{\mathbf{x} \in C^1([t_0, t_1]),\ \mathbf{x}(t_0) = A,\ \mathbf{x}(t_1) = B,\ S(\mathbf{x}(t)) = 0\}$.

It is clear that this is a minimization problem with constraint and therefore unless we can express one of the variables (x, y, z) in terms of the other two using $S(x, y, z) = 0$ we have to use method of Lagrange multipliers. Although here the constraint is pointwise and therefore you will have Lagrange multiplier as a function of t, not as a constant when you have integral constraint (think about it as satisfying $S(x(t), y(t), z(t)) = 0$ at each point t and therefore requiring a different Lagrange multiplier at each point t).

Try to find Euler-Lagrange equations here and solve this problem assuming S is a sphere.