EXISTENCE OF GLOBAL MINIMUM FOR CALCULUS OF VARIATION

LONG CHEN

ABSTRACT. The existence of a global minimum for problems in calculus of variation is briefly reviewed in this notes. The main ingredients are: 1. use coercivity to get the boundedness of a minimizing sequence; 2. use weak compactness to get a weak convergent sub-sequence; 3. use convexity in the derivative to prove the functional is lower weak semi-continuous.

CONTENTS

1. INTRODUCTION

We explore the calculus of variations problem given by the following expression:

$$\inf_{u\in\mathcal{M}} I(u), \tag{1}$$

where $I(u)$ denotes the integral functional defined as:

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))\, dx,$$

where

- $\Omega \subset \mathbb{R}^d$ represents an open and bounded Lipschitz domain;
- $\mathcal{M} = \{u \in W^{1,q}(\Omega) : \operatorname{trace} u = g \text{ on } \partial\Omega\}$ signifies the space of admissible functions;
- $L : \Omega \times \mathbb{R} \times \mathbb{R}^d$ is known as the Lagrangian.

When expressing the Lagrangian as $L(x, u, p)$, the variables (u, p) are independent. However, in the notation $L(x, u(x), \nabla u(x))$, $p = \nabla u$ is substituted, resulting in the second and third variables being related, and both becoming functions of x. After integration, $I(u)$ maps a function to a real value. Our objective is to find an admissible function that minimizes this functional.

Date: May 26, 2023.

Similar to the calculus of variations problem, we can also examine the following calculus problem:

$$\inf_{x\in M} f(x), \tag{2}$$

where M denotes a subset of $\mathbb{R}^d$.

There are two crucial sufficient conditions for the existence of solutions to (2):

(1) The function f exhibits continuity on the set M.
(2) The set M is compact, which is equivalent to M being both bounded and closed.

In summary, when searching for the minimum value of the function f over the set M, we can ensure the existence of a solution provided that f is continuous and M is compact.

We want to find the minimum value for the calculus of variations problem (1). This task is not straightforward, and it is essential to understand the main differences and challenges when moving from the calculus problem (2) to the calculus of variations problem (1).

In the calculus of variations problem, the subset $\mathcal{M}$ is part of an infinite-dimensional function space. On the other hand, the simpler calculus problem has a subset M that belongs to a finite-dimensional space. We need to reexamine many concepts that we know from finite-dimensional spaces.

One crucial difference between finite and infinite-dimensional spaces involves compactness. In finite-dimensional spaces, a set is compact if it is both bounded and closed. However, in infinite-dimensional spaces, having a bounded and closed set is not enough to ensure compactness. For instance, consider the set $\{e_n\} \subset \ell^2$, where $e_n = (0, 0, \ldots, 1, 0, \ldots)$. There is no convergent subsequences as $\|e_n - e_m\| = \sqrt{2}$.

To make sure a unit ball is compact, we introduce weaker topologies. Topologies describe a collection of subsets of a space as open sets. A set is called compact if, for every open covering, a finite open covering exists. Weaker topologies have fewer open sets, increasing the likelihood of satisfying "for every open covering."

A mapping $T : X \to Y$ between two topological spaces is continuous if the preimage of open sets is open. However, having fewer open sets in X is not favorable to continuity. Thus, $I(u)$ may not be continuous in the weak topology. To address this, we relax continuity to lower semi-continuity in the weak topology (w.l.s.c).

It remains to figure out conditions on the Lagrangian L to ensure the boundedness of a minimizing sequence $\{u_k\}$ and the w.l.s.c. of $I(\cdot)$. By the scaling argument, the derivative ∇u is the dominate part and thus conditions are only imposed on the p variable for the Lagrangian $L(x, u, p)$. More specifically, we shall use the following properties of function $p \to L(x, u, p)$

- Coercivity $\to$ boundedness.
- Convexity $\to$ w.l.s.c.

2. Preliminary and Examples

A minimizing sequence $\{u_k\}$ is defined by the property that $\lim_{k\to\infty} I(u_k) = \inf I(u)$. This concept can also be applied to the calculus problem (2). One approach to construct a minimizing sequence involves using nested finite-dimensional subspaces or subsets $\{\mathcal{M}_k\}$ to approximate the admissible function set $\mathcal{M}$. That is, $\mathcal{M}_1 \subset \mathcal{M}_2 \subset \cdots$, and $\cup_{k=1}^{\infty}\mathcal{M}_k$ is dense in $\mathcal{M}$. We then restrict the problem in Equation (1) to $\mathcal{M}_k$ to obtain u_k. Due to the nested structure, $I(u_{k+1}) \le I(u_k)$, and if $I(\cdot)$ is bounded below, we can conclude that a minimizing sequence exists.

Is it possible to conclude that the limit of u_k will be the global minimizer? We will first present several examples below. Typically, we begin with a calculus example and then proceed to the calculus of variations.

Example 2.1 (No lower bound)**.** Consider $f(x) = x(x^2 - 1)$ and $M = \mathbb{R}$, there is no global minimizer as $\lim_{x\to-\infty} f(x) = -\infty$. The function is not bounded below.

Example 2.2 (Minimizing sequence is unbounded)**.** Consider $f(x) = e^{-x}$ and $M = \mathbb{R}_+$. Function f is bounded below, i.e. $f \geq 0$ and $\inf f = 0$. But no global minimizer. The minimizing sequence $\{x_k\}$ is not bounded. Here f is strictly convex but not μ-strongly convex for any $\mu > 0$.

We then move to examples in calculus of variation.

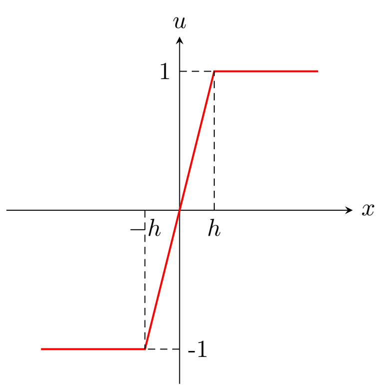
Example 2.3 (Not bounded below)**.** Consider $I(u) = \int_0^1 (1+u)(u')^2 \,\mathrm{d}x$ with boundary condition $u(0) = 0, u(1) = 3$. Solving Euler-Lagrange equation, we can find a local minimum $u(x) = (7x+1)^{2/3} - 1$. But there is no global minimum as $\inf I(u) = -\infty$ by constructing $1 + u < 0$ and $(u')^2$ becomes large.

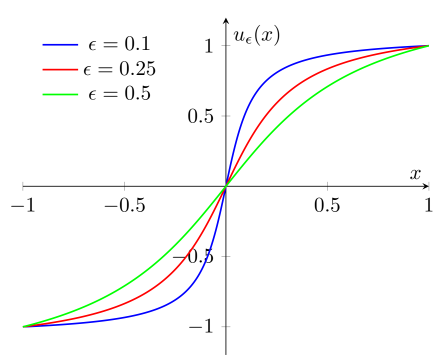
Take a point $h \in (0,1)$. Construct a piecewise linear function u_h such $(1+u_h)(h) = -4$, then the integral from $[h,1]$ and $[0, 2h/5]$ are zero by the symmetry. Overall $I(u_h) = O(-1/h)$. As $h \to 0$, $I(u_h)$ goes to $-\infty$.

Example 2.4 (Minimizing sequence is unbounded)**.** Consider $I(u) = \int_{-1}^1 x^2(u')^2 \,\mathrm{d}x$ with $\mathcal{M} = \{u \in H^1(-1,1), u(-1) = -1, u(1) = 1\}$. The Lagrangian is convex but not at $x = 0$. Construct a piecewise linear function with slop $1/h$ from $(-h,h)$ to connect -1 and 1. Then $I(u_h) \to 0$. If u_0 is a minimum of I on M, then $u_0' = 0$ and u_0 is constant which cannot satisfy the boundary condition. Notice that the minimizing sequence $\{u_h\}$ is bounded in $L^2(0,1)$ but not in $H^1(0,1)$

In the original example constructed by Weierstrass, it is a smooth function

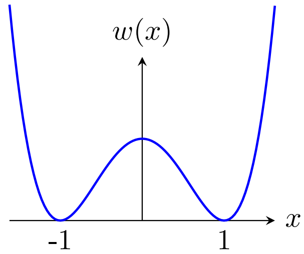
$$u_\epsilon = \frac{\arctan(x/\epsilon)}{\arctan(1/\epsilon)}.$$





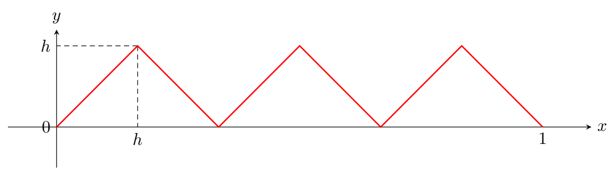
Example 2.5 (Weak convergence)**.** Consider

$$I(u) = \int_0^1 u^2 + \left[(u')^2 - 1\right]^2 \,\mathrm{d}x$$



with boundary condition $u(0) = u(1) = 0$. The potential function $w(x) = (x^2 - 1)^2$ is called double well potential and illustrated below.

We can construct a zig-zag piecewise linear function s.t. $|u_h| \leq h$ but $|u_h'| = 1$. Then u_h is a minimizing sequence and $I(u_h) \to 0$ but what is the limit of u_h? The minimizing sequence $\{u_h\}$ will converge to 0 in L^2 but not in H^1 as $\|u_h'\| = 1$ not zero. The functional is not continuous in L^2: $I(0) = 1 \neq 0 = \lim_{h\to 0} I(u_h)$.



3. Dirichlet's Principle and Fix

In this section, we present an existence proof, which corrects a flawed proof based on Dirichlet's principle. The following presentation closely follows the works in [1, Chapter 9] and [4, Chapter 8].

3.1. **Dirichlet's principle.**

Historically, the existence proof relied on Dirichlet's principle, which states: "If $I(\cdot)$ is bounded below, then there exists a minimizer u such that $I(u) = \inf_{v\in\mathcal{M}} I(v)$."

Here is a 'proof' of the Dirichlet's principle (provided by Riemann [3]): "Choose a minimizing sequence $\{u_k\} \subset \mathcal{M}$ such that $I(u_k) \to \inf_{u\in\mathcal{M}} I(u)$. Since the sequence $\{u_k\}$ is bounded, there exists a convergent subsequence $u_{k_j} \to u_0$. This u_0 is then the desired solution, i.e., $I(u_0) = \inf_{u\in\mathcal{M}} I(u)$."

The flaw in the 'proof' consists of

(1) $\{I(u_k)\}$ is bounded $\not\Longrightarrow$ $\left\{\|u_k\|_{H^1(\Omega)}\right\}$ is bounded. (Example 2.4)

(2) $\left\{\|u_k\|_{H^1(\Omega)}\right\}$ is bounded $\not\Longrightarrow$ existence of a convergent subsequence $u_{k_j} \to u_0$ in $H^1(\Omega)$. (Example 2.5)

We will introduce the assumption *coercivity* to fix (1) and *weak convergence* to fix (2). Then the third difficulty arises.

(3) $u_k \rightharpoonup u_0 \not\Longrightarrow I(u_0) = \lim_{k\to\infty} I(u_k)$. (Example 2.5)

In the weak topology, we have less open sets and thus the functional $I(\cdot)$ may not be continuous or *lower semi-continuous* in the weak topology.

Let us examine these flaws and address them one by one. Using the scaling argument, we find that the derivative ∇u is dominated, which implies that most assumptions are imposed on the p variable in the Lagrangian $L(x, u, p)$.

3.2. **Coercivity.** The Lagrangian $L(x, u, p)$ is said to be coercive with respect to the variable p if there exist constants $\alpha > 0$, $\beta \geq 0$, and $1 < q < \infty$ such that

$$L(x, u, p) \geq \alpha |p|^q - \beta, \quad \forall p \in \mathbb{R}^n, u \in \mathbb{R}, x \in \Omega. \tag{3}$$

As a result, for any $w \in W^{1,q}(\Omega)$,

$$I(w) \geq \delta \|\nabla w\|^q_{L^q(\Omega)} - \gamma,$$

where $\gamma := \beta|\Omega|$ and some constant $\delta > 0$.

If $I(w) < \infty$, then $\|\nabla w\|_{L^q} < \infty$. Fix a specific function $u_g \in \mathcal{M}$. Then, $w - u_g \in W_0^{1,q}(\Omega)$. We apply the Poincaré inequality to bound $\|w - u_g\|_{L^q} \lesssim \|\nabla(w - u_g)\|_{L^q}$ and use the triangle inequality to conclude $\|w\|_{L^q} < \infty$.

In summary, by assuming the Lagrangian is coercive, we obtain the boundedness of a minimizing sequence.

3.3. **Weak compactness.** Let X be a Banach space. A sequence $\{u_k\} \subset X$ is called *weak convergent* to $u \in X$ and denoted by

$$u_k \rightharpoonup u \text{ if } \langle f, u_k \rangle \to \langle f, u \rangle \quad \forall f \in X^*.$$

For comparison, the strong convergence is

$$u_k \to u \text{ if } \lim_{k\to\infty} \|u_k - u\|_X = 0.$$

Exercise 3.1 (Exercise 1 of Chapter 8 in [4]). *This problem illustrates that a weakly convergent sequence can be rather badly behaved.*

(1) Prove $u_k(x) = \sin(kx) \rightharpoonup 0$ *as* $k \to \infty$ *in* $L^2(0, 1)$.

(2) Fix $a, b \in \mathbb{R}, 0 < \lambda < 1$. *Define*

$$u_k(x) := \begin{cases} a & \text{if } j/k \leq x < (j + \lambda)/k \\ b & \text{if } (j + \lambda)/k \leq x < (j + 1)/k \end{cases} \quad (j = 0, \dots, k - 1).$$

Prove $u_k \rightharpoonup \lambda a + (1 - \lambda) b$.

Recall that the duality pair

$$\langle f, u \rangle, \quad f \in X^*, u \in X$$

can be used to define a natural embedding $X \to X^{**} = (X^*)^*$. If this map is surjective, i.e. $X = X^{**}$, we call X is reflexive.

In the dual space X^*, we have the weak convergence:

$$f_k \rightharpoonup f \text{ if } \langle f_k, u \rangle \to \langle f, u \rangle \quad \forall u \in (X^*)^*.$$

If we restrict the test function to a smaller subspace $X \subseteq (X^*)^*$, we obtain the weak-* convergence:

$$f_k \rightharpoonup^* f \text{ if } \langle f_k, u \rangle \to \langle f, u \rangle \quad \forall u \in X \subseteq X^{**}. \tag{4}$$

In finite element dimensional spaces, a bounded sequence will have a convergence subsequence. In infinite dimensional spaces, this may not be true for strong convergence but did hold in the weak convergence provide the space X is reflexive.

Exercise 3.2. *Consider the space* ℓ^2. *Assume* $\{x_n \in \ell^2\}$ *is a bounded sequence, i.e.* $\sup_n \|x_n\|_{\ell^2} < C$. *Prove that there exists a coordinate-wise convergent subsequence.*

We will present a result under a stronger condition: X is separable. Recall that a normed space X is considered separable if there exists a countable subset Y that is dense in X. When X is separable, we can find a convergent subsequence coordinate-wise and apply the diagonal argument. In functional analysis, the Banach-Alaoglu theorem (also known as Alaoglu's theorem) states that the closed unit ball of the dual space of a separable normed vector space is compact in the weak-* topology.

Theorem 3.3 (Banach-Alaoglu)**.** *Let X^* be the dual space of a separable normed linear space X. Suppose $\{f_n \mid n = 1, 2, \ldots\} \subset X^*$ is a norm-bounded sequence: $M = \sup \|f_n\| < \infty$, then it has a weak-* convergent subsequence.*

Proof. Since X is separable, it has a countable dense subset $\{x_k \mid k = 1, 2, \ldots\}$. For x_1, since $|\langle f_n, x_1\rangle|$ is a bounded sequence, it has a subsequence $f_{n_j^1}$ such that $\left\langle f_{n_j^1}, x_1\right\rangle$ converges.

For x_2, since $\left|\left\langle f_{n_j^1}, x_1\right\rangle\right|$ is a bounded sequence, it has a subsequence $f_{n_j^2}$ such that $\left\langle f_{n_j^2}, x_1\right\rangle$ converges.

Continuing in this fashion and applying the diagonal method, we can choose a subsequence $\left\{f_{n_j^j}\right\}$ such that $\left\langle f_{n_j^j}, x_k\right\rangle$ converges, $\forall k = 1, 2, \ldots$

However, since $\{x_k \mid k = 1, 2, \ldots\}$ is dense and $\{f_n \mid n = 1, 2, \ldots\} \subset X^*$ is bounded in norm, for any $x \in X$, the sequence $\left\{\left\langle f_{n_j^j}, x\right\rangle\right\}$ converges. Define

$$f(x) = \lim_{j\to\infty} \left\langle f_{n_j^j}, x\right\rangle$$

It is clear that $f(x)$ is linear and continuous,

$$|f(x)| \leq \sup_j \left\|f_{n_j^j}\right\|\|x\| \leq M\|x\| \quad \forall x \in X.$$

Thus $f \in X^*$, and

$$f_{n_j^j} \rightharpoonup^* f.$$

□

The separable condition can be relaxed to the space is reflective, i.e. $X = X^{**}$ and thus weak convergence is equivalent to weak-* convergence. In a reflexive Banach space, a bounded sequence is weakly sequential compact.

Theorem 3.4. *Let X be a reflexive Banach space and $\{u_k\} \subset X$ is bounded, i.e. $\sup_k \|u_k\| \leq C$. Then there exists a subsequence $\{u_{k_j}\}$ and an element $u \in X$ s.t. $u_k \rightharpoonup u$.*

Example 3.5. The space $L^p(\Omega), 1 < p < +\infty$, is reflexive. But not including the ending points $p = 1, +\infty$. We do have

$$(L^1(\Omega))' = L^\infty(\Omega), \text{ but } \quad (L^\infty(\Omega))' \supset L^1(\Omega).$$

3.4. **Lower semi-continuity.** A function $f : X \to \mathbb{R}$ is lower weak semi-continuous if:

$$f(u) \leq \liminf_{k\to\infty} f(u_k) \quad \text{ when } u_k \rightharpoonup u.$$

We introduce the extended-valued function $f : X \to \mathbb{R}_{+\infty} = \mathbb{R} \cup \{+\infty\}$, and define the effective domain

$$\operatorname{dom}(f) = \{x \in X : f(x) < +\infty\}.$$

A function is called proper if $\operatorname{dom} f \neq \varnothing$.

The *graph* of f is a surface in space $V \times \mathbb{R}$:

$$G(f) = \{(x, f(x)), x \in \operatorname{dom}(f)\} \subset V \times \mathbb{R}.$$

The *epigraph* (or supergraph) of f is the set

$$\mathrm{epi}(f) = \{(x,t) \in V \times \mathbb{R}, x \in \mathrm{dom}(f), f(x) \le t\}. \tag{5}$$

Theorem 3.6. *For any function $f : X \to \mathbb{R}_{+\infty}$, the following statements are equivalent:*

(1) Function f is lower semi-continuous.
(2) The epigraph $\mathrm{epi}(f)$ is a closed set in $X \times \mathbb{R}_{+\infty}$.

The closeness implies a certain continuity of the boundary of $\mathrm{epi}(f)$. The vertical discontinuity of the graph of f might actually be due to the choice of the coordinate system. By choosing a suitable local coordinate, the function can indeed be Lipschitz continuous.

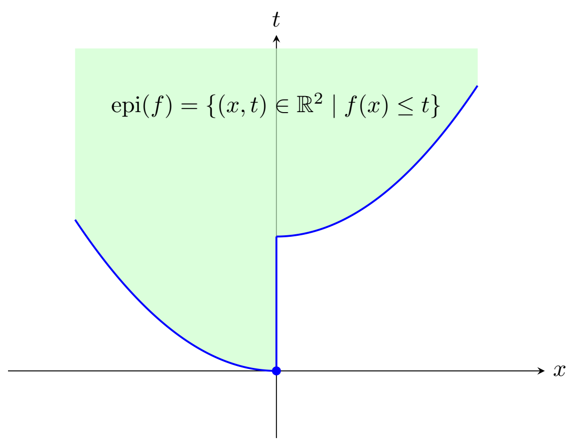


FIGURE 1. A lower semi-continuous function

3.5. **Convexity.** A continuous function $f : X \to \mathbb{R}_{+\infty}$ is *convex* if

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y) \quad \forall\, x, y \in X, \forall\, \alpha \in [0,1], \tag{6}$$

and it is *strictly convex* if

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y) \quad \forall x, y \in X, \forall \alpha \in (0,1).$$

A convex function is called *μ-strongly convex* or *μ-convex* with parameter $\mu > 0$ if

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y) - \frac{\mu}{2}\alpha(1-\alpha)\|x-y\|^2 \quad \forall x, y \in X, \tag{7}$$

for all $\alpha \in [0,1]$. A direct computation shows that (7) is equivalent to $f(x) - \frac{\mu}{2}\|x\|^2$ is convex. Convexity is helpful for the existence of global minimum.

Exercise 3.7 (Strictly convex function)**.** *Show that if f is μ-strongly convex for some $\mu > 0$, then there exists a unique global minimizer of* $\min_{x\in X} f(x)$.

The convexity can be characterized through the geometry set $\mathrm{epi}(f)$. The following equivalence of convexity can be easily verified by definition.

Theorem 3.8. *f and $\mathrm{dom}(f)$ is convex if and only if $\mathrm{epi}(f)$ is convex.*

One of the most helpful facts derived from convexity is the equivalence between strong closedness and weak closedness. More precisely, let X be a Banach space and $K \subset X$ be a convex set. Then, K being closed is equivalent to K being weakly closed. This equivalence is due to Mazur's theorem, which states that if $u_k \rightharpoonup u$, then there exists a convex combination of u_k, e.g., $v_k = \sum_{n=k}^{N(k)} \alpha(k)_n u_n$, such that $v_k \to u$ strongly.

Theorem 3.9. *Let X be a Banach space and $f : X \to \mathbb{R}$ is a convex function. Then f is l.s.c. is equivalent to f is l.w.s.c.*

Proof. Using the characterization of l.s.c. as the closedness of epigraph which is convex and thus equivalent to it is l.w.s.c. □

3.6. **Application to Calculus of Variations.** We aim to establish the lower weak semicontinuity (l.w.s.c.) for the functional $I(u)$. A sufficient condition is that $L(x, u, p)$ is convex with respect to (u, p). However, this condition might be too strong. Since p will be substituted by ∇u, which has a dominant scaling, we only need the convexity of the Lagrangian as a function of p.

In a 1-D setting, we can calculate the scaling of u and ∇u to demonstrate the difference in scaling. In an abstract setting, the dominance of the scaling can be argued using the Sobolev embedding: weak convergence in $W^{1,q}$ will imply a strong convergent subsequence.

Theorem 3.10 (Tonelli-Morrey)**.** *Suppose $L : \bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^1$ satisfies*

(1) $L \in C^1\left(\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n\right)$,
(2) $L \geq 0$,
(3) $\forall (x, u) \in \Omega \times \mathbb{R}, p \mapsto L(x, u, p)$ is convex,

then $I(u) = \int_\Omega L(x, u(x), \nabla u(x)) dx$ is weakly sequentially lower semicontinuous in $W^{1,q}(\Omega)\,(1 \leq q < \infty)$.

Proof. Suppose $u_k \rightharpoonup u$ in $W^{1,q}$ and consequently $\sup_k \|u_k\|_{1,q} \leq C$.

The convexity implies that

$$L(x, u_k, \nabla u_k) \geq L(x, u_k, \nabla u) + L_p(x, u_k, \nabla u) \cdot (\nabla u_k - \nabla u).$$

Integrate over Ω to get

$$\begin{aligned} I(u_k) \geq \int_\Omega L(x, u_k, \nabla u) &+ \int_\Omega L_p(x, u, \nabla u) \cdot (\nabla u_k - \nabla u) \\ &+ \int_\Omega (L_p(x, u_k, \nabla u) - L_p(x, u, \nabla u)) \cdot (\nabla u_k - \nabla u). \end{aligned}$$

The third term is where we can use the weak convergence

$$(f, \nabla u_k - \nabla u) \to 0$$

using the weak convergence.

For the first term, using compact embedding theorem, we can find subsequence s.t. $u_k \to u$ strongly. To exchange the integral and limit, we can make it rigorous as follows.

For any $\epsilon > 0$, there exists a subset $K \subset \Omega$ with $|\Omega \backslash K| < \epsilon$ and

(1) $u_j \to u$ uniformly on K (Egorov's theorem),
(2) u and ∇u are continuous on K (Luzin's theorem),
(3) By the continuity of the integral

$$\int_K L(x, u(x), \nabla u(x))\, \mathrm{d}x \geq \int_\Omega L(x, u(x), \nabla u(x))\, \mathrm{d}x - \varepsilon$$

In summary,

$$\liminf_{k\to\infty} I\left(u_k\right) \geq \int_K L(x, u(x), \nabla u(x))\,\mathrm{d}x \geq I(u) - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\liminf_{k\to\infty} I\left(u_k\right) \geq I(u).$$

□

We summarize the existence theorem below.

Theorem 3.11 (Existence of minimizer)**.** *Assume that L satisfies the coercivity inequality and is convex in the variable p. Suppose also the admissible set $\mathcal{M}$ is nonempty. Then there exists at least one function $u \in \mathcal{M}$ solving*

$$I(u) = \min_{w\in\mathcal{M}} I(v).$$

References

[1] K.-C. Chang. *Lecture Notes on Calculus of Variations*, volume 6. World Scientific, 2016. 3

[2] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer Science & Business Media, 2013.

[3] B. Director. Bernhard Riemann's Dirichlet's Principle. *Fidelio, XIII*: 82–88, 2004. 4

[4] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1997. 3, 5