# EXISTENCE OF GLOBAL MINIMUM FOR CALCULUS OF VARIATION

LONG CHEN

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

## CONTENTS

1. Introduction 1
2. Dirichlet's Principle 2
2.1. Minimizing sequence 3
2.2. Dirichlet's principle 3
2.3. Examples 4
3. Coercivity and Convexity 5
3.1. Coercivity 5
3.2. Weak compactness 6
3.3. Lower semi-continuity 9
3.4. Lower semi-continuity of convex functions 10
3.5. Lower semi-continuity for nonnegative integrands 10
3.6. Convexity and lower weak semi-continuity 12
3.7. Application to Calculus of Variations 13
Appendix: Several Convergence and Littlewood's Three Principles 15
References 16

## 1. INTRODUCTION

We consider the calculus of variations problem defined by

$$\inf_{u\in\mathcal{M}} I(u) = \int_\Omega L(x,u(x),\nabla u(x))\,\mathrm{d}x, \tag{1}$$

where

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

In the notation $L(x,u,p)$, the variables $(x,u,p)$ are treated as independent. In the expression $L(x,u(x),\nabla u(x))$, we substitute $p = \nabla u(x)$, so the second and third variables depend on $x$. After integration, $I(u)$ maps each admissible function to a real number. The goal is to find an admissible function that minimizes this functional.

*Date*: April 22, 2026.

As an analogue of the calculus of variations problem, consider the optimization problem

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

where $M \subset \mathbb{R}^d$. Two conditions guarantee the existence of a solution to (2):

1. The function $f$ is continuous on $M$.
2. The set $M$ is compact which is equivalent to closed and bounded.

Finding a minimizer for (1) is more subtle. The admissible set $\mathcal{M}$ lies in an infinite-dimensional function space, while $M$ in (2) is finite-dimensional. This change forces us to reconsider key ideas such as *compactness and continuity*.

A main difference is the notion of compactness. In finite-dimensional spaces, every closed and bounded set is compact. In infinite-dimensional spaces, this is no longer true. For example, the sequence $\{e_n\} \subset \ell^2$, where $e_n = (0, 0, \dots, 1, 0, \dots)$ with 1 in the $n$-th position, satisfies $\|e_n - e_m\| = \sqrt{2}$ for $n \neq m$. Hence it has no convergent subsequence.

To recover compactness, we introduce weaker topologies. A topology is a collection of subsets, called open sets, such that the empty set and the whole space are open, arbitrary unions of open sets are open, and finite intersections of open sets are open. A set is compact if every open cover admits a finite subcover. A weaker topology has fewer open sets, so it is easier for this condition to hold.

A related notion is sequential compactness: every bounded sequence has a convergent subsequence whose limit lies in the set. Compactness implies sequential compactness. In norm-induced topologies, the two are equivalent. Although sequential compactness is often what we need, the language of open sets is often more useful.

Now consider a map $T : X \to Y$ between topological spaces. It is continuous if the preimage of every open set in $Y$ is open in $X$. When $X$ carries fewer open sets, that is, a weaker topology, continuity is harder to obtain. In particular, $I(u)$ may fail to be continuous in the weak topology. We therefore relax continuity to lower weak semicontinuity (l.w.s.c.), which can be characterized geometrically by the closedness of the epigraph.

There is a trade-off between compactness and continuity. From a topological point of view, this can be summarized as follows:

**Weak vs Strong Topology**

| | |
|---|---|
| **Weak topology:** | fewer open sets, better compactness, weaker continuity. |
| **Strong topology:** | more open sets, weaker compactness, better continuity. |

To ensure the existence of a minimizer, we need conditions on $L$ that guarantee:

- *coercivity* to ensure the boundedness of a minimizing sequence $\{u_k\}$;
- *convexity* to ensure the lower weak semi-continuity of $I(\cdot)$.

By the scaling argument, the dominant contribution to the functional comes from the gradient $\nabla u$. Therefore, we impose structural conditions: coercivity and convexity on the Lagrangian $L(x, u, p)$ only with respect to the variable $p$.

## 2. Dirichlet's Principle

To begin with, we need to assume $I(\cdot)$ is bounded below. Otherwise, $\inf I(u) = -\infty$.

2.1. **Minimizing sequence.** A minimizing sequence $\{u_k\} \subset \mathcal{M}$ is defined by

$$\lim_{k\to\infty} I(u_k) = \inf_{u\in\mathcal{M}} I(u).$$

One standard way to construct such a sequence is to use a family of nested finite-dimensional subspaces or subsets $\{\mathcal{M}_k\}$ that approximate the admissible set $\mathcal{M}$. That is, we assume

$$\mathcal{M}_1 \subset \mathcal{M}_2 \subset \cdots, \quad \text{with} \quad \bigcup_{k=1}^{\infty} \mathcal{M}_k \text{ dense in } \mathcal{M}.$$

We then restrict the minimization problem (1) to each $\mathcal{M}_k$. Since each $\mathcal{M}_k$ lies in a finite-dimensional space, finding

$$u_k = \arg\inf_{u\in\mathcal{M}_k} I(u)$$

is usually easier.

Due to the nested structure, we typically have

$$I(u_1) \geq \cdots \geq I(u_k) \geq I(u_{k+1}) \geq \cdots \geq \inf_{u\in\mathcal{M}} I(u) > -\infty,$$

so $\{I(u_k)\}$ is monotone decreasing and bounded below. It follows that $\lim_{k\to\infty} I(u_k)$ exists. Because $\bigcup_k \mathcal{M}_k$ is dense in $\mathcal{M}$, we obtain

$$\lim_{k\to\infty} I(u_k) = \inf_{u\in\mathcal{M}} I(u),$$

and thus $\{u_k\}$ is a minimizing sequence.

An important question remains: can we conclude that the sequence $\{u_k\}$, or a subsequence, converges to some limit $u \in \mathcal{M}$, and that this limit $u$ is a global minimizer?

The sequence $\{I(u_k)\}$ is a sequence of real numbers in $\mathbb{R}$, while $\{u_k\}$ lies in an infinite-dimensional space, whose topology is much richer than that of $\mathbb{R}$.

2.2. **Dirichlet's principle.** Historically, the existence of minimizers was often justified by Dirichlet's principle, which asserts:

**Dirichlet's principle**

If $I(\cdot)$ is bounded below, then there exists a minimizer $u \in \mathcal{M}$ such that

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

A classic 'proof' of Dirichlet's principle, attributed to Riemann [3], proceeds as follows:

> Choose a minimizing sequence $\{u_k\} \subset \mathcal{M}$ such that $I(u_k) \to \inf_{u\in\mathcal{M}} I(u)$.
> Since $\{u_k\}$ is bounded, there exists a convergent subsequence $u_{k_j} \to u_0$.
> Then $u_0$ is the desired minimizer, that is,
>
> $$I(u_0) = \inf_{u\in\mathcal{M}} I(u).$$

However, this argument is flawed for several reasons:

1. Boundedness of $\{I(u_k)\}$ does not imply boundedness of $\{\|u_k\|_{H^1(\Omega)}\}$.
2. Even if $\{\|u_k\|_{H^1(\Omega)}\}$ is bounded, this does not guarantee the existence of a strongly convergent subsequence in $H^1(\Omega)$.

To address these issues, we introduce two crucial notions:

- **Coercivity**: to ensure that boundedness of the functional implies boundedness in the function space norm, thus addressing issue 1.
- **Weak convergence** $u_k \rightharpoonup u_0$: to replace strong convergence in non-compact infinite-dimensional settings, thus addressing issue 2.

However, passing to the weak limit introduces a new obstacle:

3. Weak convergence $u_k \rightharpoonup u_0$ does not imply $I(u_k) \to I(u_0)$. That is, the functional $I(\cdot)$ may fail to be weakly continuous.

Recall that there is a trade-off between compactness and continuity. Since the weak topology has fewer open sets, a functional that is continuous in the strong topology may fail to be continuous in the weak topology.

Thus, the existence of a minimizer cannot be deduced from a minimizing sequence without additional assumptions. We first explore several examples that show how a minimum may fail to exist. We begin with examples from finite-dimensional calculus and then move to examples from the calculus of variations.

2.3. **Examples.** We first show that the functional can be bounded below while a minimizing sequence remains unbounded.

**Example 2.1** (Minimizing sequence is unbounded)**.** Consider $f(x) = e^{-x}$ on $M = \mathbb{R}_+$. The function $f$ is bounded below: $f \geq 0$ and $\inf_{x\in M} f(x) = 0$. We can take $M_k = [0, k]$, for which the minimizer is $x_k = k$. However, there is no global minimizer on $M$. The minimizing sequence $\{x_k\}$ is unbounded, even though $f(x_k) \to 0$. Here $f$ is strictly convex but not coercive.

**Example 2.2** (Minimizing sequence is unbounded)**.** Consider

$$I(u) = \int_{-1}^{1} x^2 (u')^2 \,\mathrm{d}x, \quad \mathcal{M} = \{u \in H^1(-1,1) : u(-1) = -1,\ u(1) = 1\}.$$

The Lagrangian $L(x, u, p)$ is convex in $p$, but it is not coercive because of the degeneracy at $x = 0$.

Construct a piecewise linear function $u_h$ that changes from $-1$ to $1$ on the interval $(-h, h)$, so that its slope there is $1/h$. See Fig. 1. Then $I(u_h) \to 0$ as $h \to 0$.

If $u_0$ were a minimizer of $I$ on $\mathcal{M}$, then we would have $I(u_0) = 0$, which implies $u_0' = 0$ a.e. on $(-1, 1)$. Hence $u_0$ would be constant, which cannot satisfy the boundary conditions. Therefore no minimizer exists. Note that the minimizing sequence $\{u_h\}$ is bounded in $L^2(-1, 1)$ but not in $H^1(-1, 1)$.

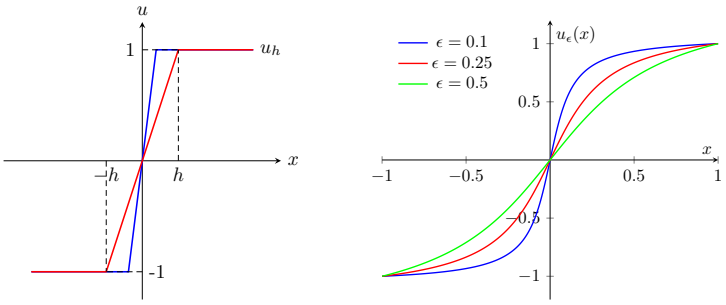



FIGURE 1. A minimizing sequence with unbounded $H^1$-norm.

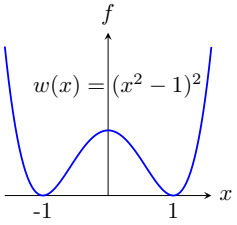
In the original example constructed by Weierstrass, it is a smooth function

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

**Example 2.3** (Loss of continuity under weak convergence)**.** Consider the functional

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

with boundary conditions $u(0) = u(1) = 0$. The function $\left(x^2-1\right)^2$ is called a double-well potential. We can construct a zig-zag piecewise linear function $u_h$ such that $|u_h(x)| \le h$ and $|u_h'(x)| = 1$. See Fig. 2.



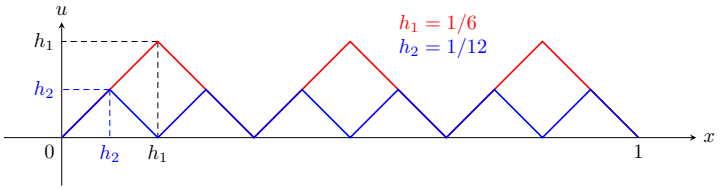




FIGURE 2. A minimizing sequence $\{u_h\}$ with $u_h \to 0$ in $L^2$, while $I(\cdot)$ is not continuous in the $L^2$ topology.

Then $\{u_h\}$ is a minimizing sequence with $I(u_h) \to 0$ as $h \to 0$. What is the limit of $u_h$? The sequence $\{u_h\}$ converges to zero in $L^2$, since

$$\|u_h\|_{L^2(0,1)} \le \|u_h\|_{L^\infty(0,1)} \le h,$$

but it does not converge in $H^1$, because $\|u_h'\|_{L^2(0,1)} = 1$ does not tend to zero. Thus, the minimizing sequence does not converge strongly in $H^1$, but it does converge in the weaker $L^2$ topology.

However, the functional $I$ is not continuous in the $L^2$ topology. Although $u_h \to 0$ in $L^2$, we have

$$I(0) = \int_0^1 0 + (0-1)^2 \, \mathrm{d}x = 1 \neq 0 = \lim_{h\to 0} I(u_h).$$

This shows that a minimum may fail to be attained even when a minimizing sequence converges. The convergence may occur only in a weaker topology, and the lack of continuity in that topology is the main obstruction in this example.

## 3. Coercivity and Convexity

In this section, we give an existence proof that corrects the flawed argument based on Dirichlet's principle. The presentation follows [1, Chapter 9] and [4, Chapter 8].

3.1. **Coercivity.** By scaling, we see that boundedness often comes from control of $\nabla u$. This is why most structural conditions are imposed on the $p$ variable in $L(x,u,p)$.

The Lagrangian $L(x,u,p)$ is called *coercive* with respect to $p$ if there exist constants $\alpha > 0$, $\beta \ge 0$, and $1 < q < \infty$ such that

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

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

$$I(w) := \int_\Omega L(x, w(x), \nabla w(x))\,\mathrm{d}x$$

satisfies the lower bound

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

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

In particular, if $I(w) < \infty$, then $\|\nabla w\|_{L^q(\Omega)} < \infty$. Now fix a reference function $u_g \in \mathcal{M}$ such that $I(u_g) < \infty$. Then for any admissible $w \in \mathcal{M}$, we have $w - u_g \in W^{1,q}_0(\Omega)$. By the Poincaré inequality,

$$\|w - u_g\|_{L^q(\Omega)} \leq C\|\nabla(w - u_g)\|_{L^q(\Omega)},$$

and hence

$$\|w\|_{L^q(\Omega)} \leq \|w - u_g\|_{L^q(\Omega)} + \|u_g\|_{L^q(\Omega)} < \infty.$$

In summary, the coercivity assumption (3) on $p$ gives boundedness of minimizing sequences in $W^{1,q}(\Omega)$ and also ensures that the functional $I(\cdot)$ is bounded below.

3.2. **Weak compactness.** Let $X$ be a Banach space, that is, a complete normed space, and let $X^*$ denote its dual space, consisting of all linear and continuous functionals on $X$. A sequence $\{u_k\} \subset X$ is said to converge weakly to $u \in X$, denoted by

$$u_k \rightharpoonup u,$$

if

$$\langle f, u_k\rangle \to \langle f, u\rangle \quad \text{for all } f \in X^*,$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $X^*$ and $X$.

In contrast, strong convergence is defined by

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

By the definition of a continuous linear functional, strong convergence implies weak convergence:

$$|\langle f, u - u_k\rangle| \leq \|f\|_{X^*}\|u - u_k\|_X \to 0.$$

However, weak convergence does not imply strong convergence. A counterexample is given in the exercise below.

**Exercise 3.1** (Exercise 1, Chapter 8 of [4])**.** *This problem shows that a weakly convergent sequence can behave quite badly.*

1. *Show that $u_k(x) = \sin(kx)$ converges weakly to zero in $L^2(0,1)$ as $k \to \infty$.*
2. *Fix constants $a, b \in \mathbb{R}$ and $0 < \lambda < 1$. Define a sequence of piecewise constant functions:*

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

*Prove that $u_k \rightharpoonup \lambda a + (1-\lambda)b$ in $L^2(0,1)$.*

The duality pairing

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

can be used to define a natural embedding $X \hookrightarrow X^{**} := (X^*)^*$ by mapping each $u \in X$ to the element $T_u \in X^{**}$ defined by

$$T_u(f) := \langle f, u\rangle \quad \text{for all } f \in X^*.$$

If this mapping is surjective, that is, if every element of $X^{**}$ arises from some $u \in X$, then we say that $X$ is *reflexive*.

**Exercise 3.2.** *The space $L^p(\Omega)$ is reflexive for $1 < p < +\infty$, but not for the endpoints $p = 1$ and $p = +\infty$. We do have*

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

*Find $(L^\infty(\Omega))^*$.*

In the dual space $X^*$, weak convergence is defined by

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

If we restrict the class of test functions $u$ to the subspace $X \subseteq X^{**}$ via the natural embedding, we obtain *weak-* convergence*:

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

If $X$ is reflexive,then weak convergence is equivalent to weak-* convergence.

If $u_k \rightharpoonup u$, then in general we cannot conclude that $\|u_k - u\|_X \to 0$. However, we can conclude that the norms are uniformly bounded:

$$u_k \rightharpoonup u \quad \Rightarrow \quad \sup_k \|u_k\|_X < \infty.$$

Indeed, if $u_k \rightharpoonup u$ in $X$, then for each $f \in X^*$, the sequence $\{\langle f, u_k\rangle\}_k$ is bounded. Then $\{T_k := T_{u_k}(f)\}_k$ is bounded for every $f \in X^*$, and the Uniform Boundedness Principle yields

$$\sup_k \|T_k\| = \sup_k \|u_k\|_X < \infty.$$

In finite-dimensional spaces, every bounded sequence has a convergent subsequence by the Bolzano–Weierstrass theorem. In infinite-dimensional Banach spaces, this is no longer true in general for strong convergence; recall the example $\{e_n\} \subset \ell^2$. We therefore need replace strong convergence by a weaker notion of convergence.

**Exercise 3.3.** *Let $\ell^2$ denote the space of square-summable sequences. Suppose $\{x_n\}_{n=1}^\infty \subset \ell^2$ is a bounded sequence, that is,*

$$\sup_n \|x_n\|_{\ell^2} < C.$$

*Prove that there exists a subsequence $\{x_{n_k}\}$ and an element $x \in \ell^2$ such that for each coordinate $j \in \mathbb{N}$,*

$$(x_{n_k})_j \to x_j \quad \textit{as } k \to \infty.$$

*That is, a bounded sequence in $\ell^2$ admits a coordinate-wise convergent subsequence.*

We now present a sequential compactness result under the additional assumption that $X$ is separable. Recall that a normed space $X$ is called *separable* if it contains a countable dense subset. The Banach–Alaoglu theorem states that the closed unit ball of $X^*$ is compact in the weak-$*$ topology. For our purposes, it is convenient to prove this sequential form directly by a diagonal argument.

**Theorem 3.4** (Sequential weak-$*$ compactness in the separable case)**.** *Let $X$ be a separable normed linear space, and let $\{f_n\}_{n=1}^\infty \subset X^*$ be a norm-bounded sequence:*

$$\sup_n \|f_n\| < \infty.$$

*Then there exist a subsequence $\{f_{m_j}\}$ and a functional $f \in X^*$ such that*

$$f_{m_j} \rightharpoonup^* f \qquad \textit{in } X^*.$$

*Proof.* Since $X$ is separable, there exists a countable dense subset $\{u_k\}_{k=1}^{\infty} \subset X$.

We construct a diagonal subsequence. For $u_1$, the scalar sequence $\{\langle f_n, u_1\rangle\}_{n=1}^{\infty}$ is bounded, hence it has a convergent subsequence $\{f_{n_j^1}\}$. Next, for $u_2$, the scalar sequence $\{\langle f_{n_j^1}, u_2\rangle\}_{j=1}^{\infty}$ is again bounded, so it has a convergent subsequence $\{f_{n_j^2}\} \subset \{f_{n_j^1}\}$. Continuing inductively, for each $k$ we choose a subsequence $\{f_{n_j^k}\}$ of $\{f_{n_j^{k-1}}\}$ such that

$$\langle f_{n_j^k}, u_k\rangle$$

converges as $j \to \infty$.

Now define the diagonal subsequence by

$$f_{m_j} := f_{n_j^j}.$$

Then for each fixed $k$, the sequence $\langle f_{m_j}, u_k\rangle$ converges as $j \to \infty$.

We claim that for every $u \in X$, the limit $\lim_{j\to\infty}\langle f_{m_j}, u\rangle$ exists. Indeed, since $\{u_k\}$ is dense in $X$ and $\sup_j \|f_{m_j}\| \le C$, for any $\varepsilon > 0$ we may choose $u_k$ with $\|u - u_k\| < \varepsilon/(3C)$; then the convergence of $\langle f_{m_j}, u_k\rangle$ and the triangle inequality show that $\{\langle f_{m_j}, u\rangle\}$ is Cauchy, hence convergent.

Define

$$f(u) := \lim_{j\to\infty} \langle f_{m_j}, u\rangle, \qquad u \in X.$$

It is easy to check that $f$ is linear. Moreover,

$$|f(u)| = \lim_{j\to\infty} |\langle f_{m_j}, u\rangle| \le C\|u\|,$$

so $f \in X^*$.

Therefore,

$$\langle f_{m_j}, u\rangle \to \langle f, u\rangle \qquad \text{for all } u \in X,$$

that is,

$$f_{m_j} \rightharpoonup^* f.$$

□

When $X$ is reflexive, the canonical embedding identifies $X$ with $X^{**}$, and hence weak convergence and weak-$*$ convergence in $X^*$ coincide.

**Corollary 3.5.** *Let $X$ be a reflexive and separable Banach space, and let $\{u_k\} \subset X$ be a bounded sequence, that is,*

$$\sup_k \|u_k\|_X < \infty.$$

*Then there exist a subsequence $\{u_{k_j}\} \subset \{u_k\}$ and an element $u \in X$ such that*

$$u_{k_j} \rightharpoonup u \qquad \text{in } X.$$

For weak compactness of bounded sequences in $X$, the assumption "reflexive and separable" may be replaced by "reflexive." The separability assumption is needed only to obtain a subsequence result by a diagonal argument. If one is interested only in topological compactness, rather than sequential compactness, then neither reflexivity nor separability is needed. This is precisely the content of the Banach–Alaoglu theorem: if $X$ is a normed space, then the closed unit ball of $X^*$ is compact in the weak-$*$ topology.

**Remark 3.6.** In Sobolev spaces $W^{1,q}(\Omega)$, we exclude the endpoint case $q = 1$, since it lies outside the reflexive setting. Because $W^{1,1}(\Omega)$ is not reflexive, bounded sequences in $W^{1,1}(\Omega)$ need not admit weakly convergent subsequences in $W^{1,1}(\Omega)$. For this reason, existence theory for functionals with linear growth is usually developed in the larger space

$BV(\Omega)$. In that setting, bounded sequences admit weak-$*$ convergent subsequences, which replace the weak compactness argument used for $1 < q < \infty$.

3.3. **Lower semi-continuity.** A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is called *lower semi-continuous* (l.s.c.) if

$$f(u) \leq \liminf_{k\to\infty} f(u_k) \qquad \text{whenever } u_k \to u.$$

It is called *lower weakly semi-continuous* (l.w.s.c.) if

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

We allow the value $+\infty$ in order to include singular behavior and constraints. Define

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

The function $f$ is called *proper* if $\mathrm{dom}(f) \neq \varnothing$.

The *graph* of $f$ is

$$G(f) := \{(x, f(x)) : x \in \mathrm{dom}(f)\} \subset X \times \mathbb{R}.$$

The *epigraph* of $f$ is

$$\mathrm{epi}(f) := \{(x, t) \in X \times \mathbb{R} : t \geq f(x)\}.$$

Geometrically, for each $x \in \mathrm{dom}(f)$, the set $\{x\} \times [f(x), \infty)$ is a vertical ray starting at $(x, f(x))$ and extending upward.

The following characterization is immediate from the definition.

**Theorem 3.7.** *For any function $f : X \to \mathbb{R} \cup \{+\infty\}$, the following are equivalent:*

1. *$f$ is lower semi-continuous;*
2. *$\mathrm{epi}(f)$ is closed in $X \times \mathbb{R}$.*

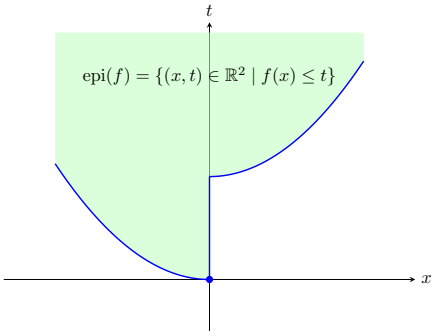



FIGURE 3. A lower semi-continuous function and its closed epigraph.

Thus, lower semicontinuity can be interpreted geometrically as the closedness of the epigraph. This viewpoint is especially useful in variational problems, where closedness properties are often easier to verify than continuity itself. Closedness is a geometric notion, whereas continuity, such as a vertical jump, depends on the choice of coordinates. For the vertical jump in Fig. 3, the boundary is locally Lipschitz after a change of coordinates.

3.4. **Lower semi-continuity of convex functions.** A subset $K$ of a vector space $V$ is convex if

$$x, y \in K, \lambda \in (0,1) \quad \Longrightarrow \quad (1-\lambda)x + \lambda y \in K.$$

Recall for a linear subspace $S$, it requires $\alpha x + \beta y \in S$ for any $\alpha, \beta \in \mathbb{R}$. The convexity restricts the coefficients $\alpha + \beta = 1, \alpha, \beta \geq 0$, which is called convex combination. Geometrically, convexity implies the line segment connecting $x$ and $y$ is inside the set while the linear subspace implies the plane spanned by vectors $x$ and $y$ is contained in the subspace.

A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is called *convex* if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \qquad \forall\, x, y \in X,\ \forall\, \alpha \in [0,1]. \tag{5}$$

It is called *strictly convex* if the inequality is strict whenever $x \neq y$ and $\alpha \in (0,1)$.

The convexity of a function can be characterized geometrically through its epigraph. The following theorem follows directly from the definition.

**Theorem 3.8.** *A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if its epigraph $\operatorname{epi}(f)$ is convex.*

A useful consequence of convexity is the equivalence between norm closedness and weak closedness for convex sets. More precisely, if $K \subset X$ is convex, then $K$ is closed in the norm topology if and only if it is closed in the weak topology.

One way to understand this fact is through Mazur's theorem: if $u_k \rightharpoonup u$ in $X$, then there exists a sequence of convex combinations of the tail $\{u_k, u_{k+1}, \dots\}$, for example

$$v_k = \sum_{n=k}^{N(k)} \alpha_n^{(k)} u_n, \quad \sum_n \alpha_n^{(k)} = 1, \alpha_n^{(k)} \geq 0,$$

such that $v_k \to u$ strongly in $X$. Thus, for convex sets, weak convergence can be upgraded to strong convergence after passing to suitable convex combinations, rather than merely to a subsequence.

An outline of the proof of Mazur's theorem is as follows: for each tail

$$C_m = \operatorname{conv}\{u_m, u_{m+1}, \dots\},$$

the weak limit $u$ must belong to the norm closure $\overline{C_m}$; otherwise, the Hahn–Banach theorem gives a functional $f \in X^*$ that strictly separates $u$ from $\overline{C_m}$, which contradicts the convergence $f(u_k) \to f(u)$. Choosing $v_m \in C_m$ such that $\|v_m - u\| < 1/m$ then gives a sequence of convex combinations converging strongly to $u$.

**Theorem 3.9.** *Let $X$ be a Banach space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex. Then $f$ is lower semi-continuous if and only if it is lower weakly semi-continuous.*

*Proof.* By the epigraph characterization, $f$ is lower semi-continuous if and only if $\operatorname{epi}(f)$ is closed in the norm topology of $X \times \mathbb{R}$. By Theorem 3.8, the set $\operatorname{epi}(f)$ is convex. Since a convex set is norm closed if and only if it is weakly closed, it follows that $\operatorname{epi}(f)$ is closed in the norm topology if and only if it is closed in the weak topology. Therefore, $f$ is lower semi-continuous if and only if it is lower weakly semi-continuous. □

3.5. **Lower semi-continuity for nonnegative integrands.** Suppose $u_k \to u$ strongly in $L^1(\Omega)$, and let the Lagrangian $L(x, u, p)$ be continuous in the variables $(x, u, p)$. A central question in real analysis is: Can we exchange the limit and the integral to conclude

$$\lim_{k\to\infty} \int_\Omega L(u_k(x))\ \mathrm{d}x = \int_\Omega \lim_{k\to\infty} L(u_k(x))\ \mathrm{d}x = \int_\Omega L(u(x))\ \mathrm{d}x?$$

Here we suppress the variables $(x, p)$, since the convergence is only in $u$.

The answer is no. Continuity of $L$ alone does not provide a uniform bound or an integrable dominating function for $L(u_k(x))$. In particular, one cannot in general find $g \in L^1(\Omega)$ such that

$$|L(u_k(x))| \leq g(x) \qquad \text{for all } k \text{ and a.e. } x \in \Omega.$$

Without such a bound, the Dominated Convergence Theorem (DCT) cannot be applied.

**Example 3.10.** Let $\Omega = [0,1]$ and define $L(u) = u^2$. Consider

$$u_h(x) := h^{-2/3}\chi_{[0,\,h]}(x), \qquad h \to 0,$$

where $\chi_{[0,\,h]}$ is the characteristic function of $[0,h]$. Then

$$\int_0^1 u_h(x)\ \mathrm{d}x = h^{1/3} \to 0,$$

so $u_h \to u = 0$ strongly in $L^1(\Omega)$. However,

$$\int_0^1 u_h^2(x)\ \mathrm{d}x = \int_0^h h^{-4/3}\ \mathrm{d}x = h^{-1/3} \to \infty \neq 0.$$

Thus,

$$\int_\Omega L(u_h)\ \mathrm{d}x \not\to \int_\Omega L(u)\ \mathrm{d}x.$$

We can still prove lower semi-continuity if $L \geq 0$. In the example above, even though

$$\int_\Omega L(u_k)\ \mathrm{d}x \to +\infty \geq \int_\Omega L(u)\ \mathrm{d}x,$$

this causes no difficulty for lower semi-continuity, since only a lower bound is required.

**Lemma 3.11** (Fatou)**.** *Let $\{f_k \geq 0\}$ be a sequence of nonnegative measurable functions on $\Omega$. Then the integral is lower semi-continuous in the sense that*

$$\int_\Omega \liminf_{k\to\infty} f_k(x)\ \mathrm{d}x \ \leq\ \liminf_{k\to\infty} \int_\Omega f_k(x)\ \mathrm{d}x.$$

**Lemma 3.12** (Lower semicontinuity for nonnegative integrands)**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded, and let $u_k, u \in L^1(\Omega)$ satisfy*

$$u_k \to u \qquad \textit{strongly in } L^1(\Omega).$$

*Assume $L : \mathbb{R} \to \mathbb{R}$ is continuous and $L \geq 0$. Then*

$$\int_\Omega L\big(u(x)\big)\ \mathrm{d}x \ \leq\ \liminf_{k\to\infty} \int_\Omega L\big(u_k(x)\big)\ \mathrm{d}x.$$

*Proof.* Choose a subsequence $\{u_{k_j}\}$ such that

$$\int_\Omega L\big(u_{k_j}(x)\big)\ \mathrm{d}x \to \ell = \liminf_{k\to\infty} \int_\Omega L\big(u_k(x)\big)\ \mathrm{d}x.$$

Since $u_k \to u$ strongly in $L^1(\Omega)$, there exists a subsequence, still denoted by $\{u_{k_j}\}$, such that

$$u_{k_j}(x) \to u(x) \qquad \text{for a.e. } x \in \Omega.$$

By continuity of $L$,

$$L\big(u_{k_j}(x)\big) \to L\big(u(x)\big) \qquad \text{for a.e. } x \in \Omega.$$

Since $L \geq 0$, Fatou's lemma yields

$$\int_\Omega L\big(u(x)\big)\ \mathrm{d}x \leq \liminf_{j\to\infty} \int_\Omega L\big(u_{k_j}(x)\big)\ \mathrm{d}x = \ell.$$

□

**Example 3.13.** Let $L(u) = -u$. Consider $\Omega = [0,1]$ and define

$$u_k(x) = k\chi_{[0,1/k]}(x), \quad u(x) = 0.$$

Then $u_k \to u$ strongly in $L^1(\Omega)$, $\int_\Omega L(u_k)\,\mathrm{d}x = -1$ for all $k$, $\int_\Omega L(u)\,\mathrm{d}x = 0$, and $\liminf_{k\to\infty} \int_\Omega L(u_k)\,\mathrm{d}x = -1 < 0$. Thus the conclusion of Lemma 3.12 fails when $L$ takes negative values.

3.6. **Convexity and lower weak semi-continuity.** If $\nabla u_k \to \nabla u$ strongly and $L(x,u,p)$ is continuous in $p$, then one may apply the lower semi-continuity result from the previous subsection to the sequence $L(x,u_k,\nabla u_k)$. However, boundedness of $\|\nabla u_k\|_{L^q}$ yields only weak convergence, up to a subsequence:

$$\nabla u_k \rightharpoonup \nabla u \qquad \text{in } L^q(\Omega;\mathbb{R}^m).$$

Thus, to recover lower semicontinuity of the functional, one needs an additional structural assumption.

A natural sufficient condition is convexity in the gradient variable $p$. Indeed, convexity provides the supporting-hyperplane inequality, which is the key tool for proving weak lower semicontinuity.

**Lemma 3.14** (Lower weak semicontinuity of convex integral functionals)**.** *Let $\Omega \subset \mathbb{R}^n$ be measurable with finite measure. Suppose $L : \mathbb{R}^m \to \mathbb{R}$ is convex and continuously differentiable. Let $\{p_k\} \subset L^q(\Omega;\mathbb{R}^m)$ for some $1 < q < \infty$, and assume*

- *$p_k \rightharpoonup p$ weakly in $L^q(\Omega;\mathbb{R}^m)$,*
- *$L(p_k) \in L^1(\Omega)$ for all $k$,*
- *$\nabla L(p) \in L^{q^*}(\Omega;\mathbb{R}^m)$, where $1/q + 1/q^* = 1$.*

*Then we have the lower weak semi-continunity*

$$\int_\Omega L(p(x))\ \mathrm{d}x \leq \liminf_{k\to\infty} \int_\Omega L(p_k(x))\ \mathrm{d}x.$$

*Proof.* Since $L : \mathbb{R}^m \to \mathbb{R}$ is convex and differentiable, we have

$$L(p_k(x)) \geq L(p(x)) + \nabla L(p(x)) \cdot \big(p_k(x) - p(x)\big) \qquad \text{for a.e. } x \in \Omega.$$

Integrating over $\Omega$, we obtain

$$\int_\Omega L(p_k(x))\ \mathrm{d}x \geq \int_\Omega L(p(x))\ \mathrm{d}x + \int_\Omega \nabla L(p(x)) \cdot \big(p_k(x) - p(x)\big)\ \mathrm{d}x.$$

Since $p_k \rightharpoonup p$ weakly in $L^q(\Omega;\mathbb{R}^m)$ and $\nabla L(p) \in L^{q^*}(\Omega;\mathbb{R}^m)$, the last term converges to zero. Therefore

$$\liminf_{k\to\infty} \int_\Omega L(p_k(x))\ \mathrm{d}x \geq \int_\Omega L(p(x))\ \mathrm{d}x.$$

□

A variant based on Mazur's theorem shows that one may pass to suitable convex combinations of $\{p_k\}$ that converge strongly in $L^q$, so the assumption $\nabla L(p) \in L^{q^*}(\Omega;\mathbb{R}^m)$ can be avoided.

3.7. **Application to Calculus of Variations.** We now aim to establish the lower weak semicontinuity of the functional $I(u)$. Since the variable $p$ will later be replaced by $\nabla u$, and the gradient term has the dominant scaling, it is enough to assume convexity of the Lagrangian with respect to $p$ alone.

In one dimension, this difference in scaling can be seen directly by comparing $u$ and $u'$. In higher dimensions, the same idea is reflected by the Rellich–Kondrachov compact embedding theorem. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary, and let $1 \leq q < \infty$. Then the embedding

$$W^{1,q}(\Omega) \hookrightarrow L^q(\Omega)$$

is compact, in the sense that every bounded sequence in $W^{1,q}(\Omega)$ admits a strongly convergent subsequence in $L^q(\Omega)$.

Since weak convergence implies boundedness of the norm, it follows that if

$$u_k \rightharpoonup u \qquad \text{in } W^{1,q}(\Omega),$$

then $\{u_k\}$ is bounded in $W^{1,q}(\Omega)$. By the compact embedding, there exists a subsequence $\{u_{k_j}\}$ such that

$$u_{k_j} \to u \qquad \text{in } L^q(\Omega).$$

Note that the assumptions that $\Omega$ is bounded with Lipschitz boundary and that $q < \infty$ are essential here.

The proof of the Tonelli–Morrey theorem is technical, but the main ideas are already reflected in Littlewood's three principles. For the zeroth-order term $u_k \to u$, we pass to a good compact subset on which the convergence is uniform. For the gradient term $\nabla u_k \rightharpoonup \nabla u$, we use convexity in $p$. The assumption $L \geq 0$ is used to control the contribution from the complement of the good compact set.

**Theorem 3.15** (Tonelli–Morrey)**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded with Lipschitz boundary, and let $1 < q < \infty$. Suppose*

$$L : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$$

*satisfies*

1. *$L \in C^1(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n)$,*
2. *$L \geq 0$,*
3. *for every $(x, u) \in \Omega \times \mathbb{R}$, the map $p \mapsto L(x, u, p)$ is convex.*

*Then*

$$I(u) := \int_\Omega L(x, u(x), \nabla u(x)) \, \mathrm{d}x$$

*is lower weak semi-continuous on $W^{1,q}(\Omega)$.*

*Proof.* Suppose

$$u_j \rightharpoonup u \qquad \text{in } W^{1,q}(\Omega).$$

We must show

$$I(u) \leq \liminf_{j \to \infty} I(u_j).$$

Passing to a subsequence if necessary, we may assume

$$\liminf_{j \to \infty} I(u_j) = \lim_{j \to \infty} I(u_j).$$

1 By the Rellich–Kondrachov theorem, after passing to a subsequence,

$$u_j \to u \qquad \text{strongly in } L^q(\Omega).$$

Passing to a further subsequence, still denoted by $\{u_j\}$, we may also assume

$$u_j(x) \to u(x) \qquad \text{for a.e. } x \in \Omega.$$

2 Fix $\varepsilon > 0$. By Lusin's theorem and Egorov's theorem, there exists a compact set $K \subset \Omega$ such that $|\Omega \setminus K| < \varepsilon$ and

1. $u_j \to u$ uniformly on $K$,
2. $u$ and $\nabla u$ are continuous on $K$.

Since $L(x, u(x), \nabla u(x)) \geq 0$ and is measurable, by absolute continuity of the Lebesgue integral we may choose $K$ so that

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

By convexity in $p$,

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

for a.e. $x \in K$. Integrating over $K$, we obtain

$$\begin{aligned} I(u_j) \geq \int_K L(x, u_j, \nabla u)\ \mathrm{d}x + \int_K L_p(x, u, \nabla u) \cdot (\nabla u_j - \nabla u)\ \mathrm{d}x \\ + \int_K \big(L_p(x, u_j, \nabla u) - L_p(x, u, \nabla u)\big) \cdot (\nabla u_j - \nabla u)\ \mathrm{d}x. \end{aligned}$$

3 Since $u_j \to u$ uniformly on $K$, and $L$ is continuous, we have

$$L(x, u_j(x), \nabla u(x)) \to L(x, u(x), \nabla u(x)) \qquad \text{uniformly on } K.$$

Hence, by the Dominated Convergence Theorem,

$$\int_K L(x, u_j(x), \nabla u(x))\ \mathrm{d}x \to \int_K L(x, u(x), \nabla u(x))\ \mathrm{d}x.$$

4 Since $u$ and $\nabla u$ are continuous on the compact set $K$, and $L \in C^1$, the function

$$x \mapsto L_p(x, u(x), \nabla u(x))$$

belongs to $L^\infty(K; \mathbb{R}^n)$. Extending by zero outside $K$, define

$$\phi(x) := \chi_K(x) L_p(x, u(x), \nabla u(x)).$$

Then

$$\phi \in L^\infty(\Omega; \mathbb{R}^n) \subset L^{q^*}(\Omega; \mathbb{R}^n), \qquad \frac{1}{q} + \frac{1}{q^*} = 1.$$

Since $\nabla u_j \rightharpoonup \nabla u$ weakly in $L^q(\Omega; \mathbb{R}^n)$, it follows that

$$\int_K L_p(x, u(x), \nabla u(x)) \cdot (\nabla u_j - \nabla u)\ \mathrm{d}x = \int_\Omega \phi(x) \cdot (\nabla u_j - \nabla u)\ \mathrm{d}x \to 0.$$

5 Since $u_j \rightharpoonup u$ in $W^{1,q}(\Omega)$, the sequence $\{\nabla u_j\}$ is bounded in $L^q(\Omega)$. Because $K$ has finite measure, this implies

$$\|\nabla u_j - \nabla u\|_{L^1(K)} \leq C$$

for some constant $C$ independent of $j$. Moreover, since $u_j \to u$ uniformly on $K$ and $L_p$ is continuous,

$$\|L_p(\cdot, u_j(\cdot), \nabla u(\cdot)) - L_p(\cdot, u(\cdot), \nabla u(\cdot))\|_{L^\infty(K)} \to 0.$$

Therefore

$$\int_K \big(L_p(x, u_j, \nabla u) - L_p(x, u, \nabla u)\big) \cdot (\nabla u_j - \nabla u)\ \mathrm{d}x \to 0.$$

Combining the three terms, we obtain

$$\liminf_{j\to\infty} I(u_j) \geq \int_K L(x, u(x), \nabla u(x))\ \mathrm{d}x \geq I(u) - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\liminf_{j\to\infty} I(u_j) \geq I(u).$$

□

Put all components together, we obtain the following existence theorem.

**Theorem 3.16** (Existence of a minimizer)**.** *Let $1 < q < \infty$, and let $\mathcal{M} \subset W^{1,q}(\Omega)$ be nonempty and weakly closed. Assume that the Lagrangian $L(x, u, p)$ satisfies the coercivity condition and the assumptions of the Tonelli–Morrey theorem, in particular that $p \mapsto L(x, u, p)$ is convex. Then the functional*

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))\ \mathrm{d}x$$

*attains its minimum on $\mathcal{M}$. That is, there exists $u \in \mathcal{M}$ such that*

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

*Proof.* This theorem may be viewed as a correction of Dirichlet's principle.

Choose a minimizing sequence $\{u_k\} \subset \mathcal{M}$ such that

$$I(u_k) \to \inf_{w\in\mathcal{M}} I(w).$$

By coercivity, the sequence $\{u_k\}$ is bounded in $W^{1,q}(\Omega)$. Since $W^{1,q}(\Omega)$ is reflexive for $1 < q < \infty$, there exists a subsequence, still denoted by $\{u_k\}$, and a function $u_0 \in W^{1,q}(\Omega)$ such that

$$u_k \rightharpoonup u_0 \qquad \text{in } W^{1,q}(\Omega).$$

Because $\mathcal{M}$ is weakly closed, we have $u_0 \in \mathcal{M}$.

By weak lower semicontinuity of $I$,

$$I(u_0) \leq \liminf_{k\to\infty} I(u_k) = \inf_{w\in\mathcal{M}} I(w).$$

Hence $u_0$ is a minimizer. □

## Appendix: Several Convergence and Littlewood's Three Principles

We provide some background on several notions of convergence in real analysis.

- **Weak convergence in $W^{1,q}(\Omega)$:** $u_k \rightharpoonup u$ in $W^{1,q}(\Omega)$ means

$$u_k \rightharpoonup u \quad \text{and} \quad \nabla u_k \rightharpoonup \nabla u \qquad \text{in } L^q(\Omega).$$

  It implies boundedness in $W^{1,q}(\Omega)$.
- **Strong convergence in $L^q(\Omega)$:** $u_k \to u$ in $L^q(\Omega)$ means

$$\|u_k - u\|_{L^q(\Omega)} \to 0.$$

- **Almost everywhere convergence:** $u_k(x) \to u(x)$ a.e. in $\Omega$ is a pointwise statement, not a norm convergence.
- **Uniform convergence on a compact set $K \subset \Omega$:** $u_k \to u$ uniformly on $K$ means

$$\sup_{x\in K} |u_k(x) - u(x)| \to 0.$$

  This is stronger than pointwise convergence on $K$.

We use the relation

$$\begin{aligned} u_k \rightharpoonup u \quad & \text{in } W^{1,q}(\Omega) \\ \Downarrow & \\ u_{k_j} \to u \quad & \text{in } L^q(\Omega) \\ \Downarrow & \\ u_{k_j}(x) \to u(x) \quad & \text{for a.e. } x \in \Omega \\ \Downarrow & \\ u_{k_j} \to u \quad & \text{uniformly on a compact set } K \subset \Omega \end{aligned}$$

We recall *Littlewood's three principles of analysis*, which give useful heuristics in measure theory and real analysis. With negligible modification, we move from real analysis back to the more familiar setting of calculus.

1. **Every measurable set is nearly compact.**
   For any measurable set $E \subset \mathbb{R}^n$ with finite measure and any $\varepsilon > 0$, there exists a compact set $K \subset E$ such that
   $$m(E \setminus K) < \varepsilon.$$
   This is the inner regularity of Lebesgue measure.
2. **Every measurable function is nearly continuous.** *(Lusin's theorem)*
   Let $u : \Omega \to \mathbb{R}$ be a measurable function on a bounded domain $\Omega \subset \mathbb{R}^n$ with finite Lebesgue measure. Then for every $\varepsilon > 0$, there exists a closed set $K \subset \Omega$ such that
   - $|\Omega \setminus K| < \varepsilon$,
   - $u|_K$ is continuous.
3. **Every pointwise convergent sequence of measurable functions is nearly uniformly convergent.** *(Egorov's theorem)*
   Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, and let $\{u_k\}$ be a sequence of measurable functions $u_k : \Omega \to \mathbb{R}$ such that
   $$u_k(x) \to u(x) \qquad \text{for almost every } x \in \Omega.$$
   Then for every $\varepsilon > 0$, there exists a closed set $K \subset \Omega$ such that
   - $|\Omega \setminus K| < \varepsilon$,
   - $u_k \to u$ uniformly on $K$.

## REFERENCES

[1] K.-C. Chang. *Lecture Notes on Calculus of Variations*, volume 6. World Scientific, 2016.
[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] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1997.