Mastermath

Radboud University

Spring 2023

Calculus of Variations

Riccardo Cristoferi and Stefanie Sonner

Radboud University, Nijmegen
IMAPP – Mathematics

Preface

These lecture notes are written for the Mastermath course "Calculus of Variations" at Radboud University in Spring 2023. Calculus of variations is an active area of research with important applications in science and technology, e.g. in physics, material science or image processing. Moreover, variational methods play an important role in many other disciplines of mathematics such as the theory of partial differential equations, optimization, geometry and probability theory.

This course provides an introduction to different facets of this interesting field, which is concerned with the minimization (or maximization) of functionals. Further details, applications and many additional topics can be found in, e.g. the monographs by B. Dacorogna [5], H. Kielhöfer [11], J. Jost and X. Li-Jost [10] and F. Rindler [12]. To follow the course a solid understanding of real analysis, functional analysis and measure theory is required.

Contents

Chapter 1

Introduction

The **Calculus of Variations (CoV)** is concerned with the optimization of shapes, states or processes. The property to optimize is given in terms of a *functional*,

$$I(u) = \int_\Omega f(x, u(x), Du(x))dx, \qquad \Omega \subset \mathbb{R}^m, \tag{1.1}$$

involving an unknown function $u : \Omega \to \mathbb{R}^n$. The function $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{n \times m} \to \mathbb{R}$ is given and $\Omega \subset \mathbb{R}^m$ is bounded. Our aim is to find a *minimizer* (or *maximizer*) u of I within a suitable class of functions.

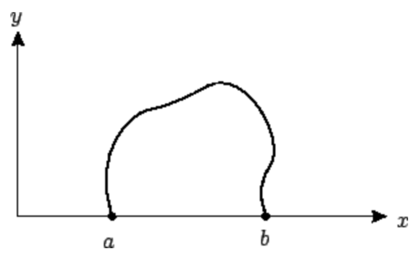
Different from the problem of finding a minimum (or maximum) of a real-valued function $g : \mathbb{R}^n \to \mathbb{R}$, we look for an unknown function u in an admissible function space which is typically infinite dimensional. Hence, we are concerned with *infinite dimensional minimization problems*.

The Calculus of Variations is a classical branch of mathematics and has diverse applications in physics, engineering and economics. The functional I can represent, e.g. a surface area, path length, an action, energy or cost. The research field Calculus of Variations is also closely related to other mathematical disciplines such as geometry, partial differential equations (PDEs), functional analysis and optimal control.

Historically, the Calculus of Variations emerged from concrete problems in geometry and physics and its origins date back to ancient times. The field had a large impact on the development of analysis, in particular, on functional analysis and the theory of PDEs. We discuss several problems and achievements that played an important role for the development of this field.

- The oldest problem in the Calculus of Variations is **Dido's problem** or "isoperimetric problem" (in Greek, *iso* means equal and *perimetron* circumference).

Dido Purchases Land for the Foundation of Carthage. Engraving by Matthäus Merian the Elder, in *Historische Chronica*, Frankfurt a.M., 1630. Dido's people cut the hide of an ox into thin strips and try to enclose a maximal domain.



The question is which curve of a given length encloses the largest area. Dido was the legendary founder of the Phoenician city-state of Carthage. When she arrived in the 9th century BC on the coast of Tunisia, she asked for a piece of land. The king offered her as much land as she could encompass by an ox-hide. She cut the ox-hide into a long thin strip and encircled the land, which became Carthage.

A rigorous mathematical proof of the isoperimetric problem was only given in the 19th century. For an overview and the history of the problem we refer to [2].

- **Fermat's principle** in geometrical optics is another important example of a variational problem. The question is along which path a light ray travels. Pierre de Fermat (1662) claimed that a path taken between two points by a light ray is the path that can be traversed in the shortest time. This leads to the problem of finding the function u that minimizes the travel time T,
$$T = \frac{1}{c}\int_a^b n(x, u(x), u'(x))\sqrt{1 + u'^2(x)}dx,$$
where n is the *index of refraction*. It can depend on the position $(x, u(x))$ and its direction $u'(x)$.

- The birth year of the Calculus of Variations is considered 1696 when Johann Bernoulli challenged his colleagues with the **Brachistochrone problem** (in Greek, *brachystos* means shortest and *chronos* time) of finding the fastest slide. More precisely, if two points A and B in a vertical plane are given, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time?

The Brachistochrone problem was solved shortly after by Johann and Jacob Bernoulli, Gottfried Wilhelm Leibniz and Isaac Newton.

- In the 18th century various variational problems were formulated by Johann and Jakob Bernoulli, Leonhard Euler, Joseph Lagrange and Adrien-Marie Legendre. Leonhard Euler introduced in 1744 the notion "Calculus of Variations". A systematic approach to study variational problems was developed by Leonhard Euler and Joseph Lagrange around 1755. In particular, they postulated that the minimizer satisfies the so-called **Euler-Lagrange equations**. At that time, the existence of minimizers was taken for granted. However, it turns out that the Euler-Lagrange equations are only necessary conditions for a minimizer.

  These so-called "*classical methods in the Calculus of Variations*" developed at that time aim at deriving methods to determine minimizers and to investigate qualitative properties while the existence of minimizers is taken for granted.

- Very significant for the development of the Calculus of Variations was **Dirichlet's principle** in the 19th century. It allowed to solve the Laplace equation $\Delta u = 0$ by reformulation it as a variational problem. However, Carl Weierstrass presented in 1860 a counterexample for the existence of minimizers showing that functionals that are bounded from below do not necessarily posses minimizers. Hence, Dirichlet's solution of the Laplace equation required an existence theory for minimizers.

  In beginning of the 20th century, this gap in the proof of the Dirichlet problem was finally solved by David Hilbert, in particular, he proved the existence of a minimizer. The problem was very important for development of functional analysis, distribution theory, Sobolev spaces and PDEs. The existence theory for minimizers is nowadays known as the "*direct method in the Calculus of Variations*". It required to introduce new function spaces (Sobolev spaces) and weaken the notion of classical derivatives.

To systematically study variational problems, we need to determine a suitable functional I as in (1.1) and to specify the admissible class of functions. The admissible class is either determined by properties resulting from a particular application or by the minimal requirements for the well-posedness of the functional I in (1.1).

In this course, we will first address the classical theory in one dimension, i.e. $\Omega \subset \mathbb{R}$ in (1.1). We take the existence of minimizers for granted, derive the Euler-Lagrange equations and aim to compute and/or derive properties of minimizers. Then, we introduce Sobolev spaces that are needed to study the existence of minimizers and variational problems in higher dimensions, i.e. $\Omega \subset \mathbb{R}^m, m \geq 2$. We derive the Euler-Lagrange equations for higher dimensional variational problems and consider direct methods that allow to prove the existence of minimizers. Finally, we discuss modern methods and applications including relaxation and Γ-convergence.

Chapter 2

Classical theory in one dimension

In this chapter, we consider one-dimensional variational problems and address the classical theory, i.e. we assume minimizers exists and aim to compute them and/or to investigate qualitative properties of the minimizers.

2.1 Fundamental problem and examples

The fundamental problem of the Calculus of Variations is to minimize functionals of the form

$$I(u) = \int_a^b f(x, u(x), u'(x))dx \tag{2.1}$$

in a class Φ of admissible functions $u : [a, b] \to \mathbb{R}^n$, $b > a$. Here, $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z)$, is a given function which at least has the following properties:

$$\begin{aligned} f(\cdot, y, z) &\quad \text{is measurable for all } \; y, z \in \mathbb{R}^n, \\ f(x, \cdot, \cdot) &\quad \text{is continuous for all } \; x \in [a, b]. \end{aligned}$$

We remark that most of the time, we assume f to be continuous or even more regular.

Under these assumptions, the functional I in (2.1) is certainly well-defined if $\Phi \subset C^1([a, b]; \mathbb{R}^n)$. However, this class is often too restrictive in applications. In this chapter, we will mainly consider the following classes of continuous functions.

Definition 2.1.

- The class of continuous functions on $[a, b]$ we denote by

$$C([a, b]; \mathbb{R}^n) = \{u : [a, b] \to \mathbb{R}^n : u \text{ continuous}\}.$$

- The class of continuously differentiable functions on $[a, b]$ we denote by

$$C^1([a, b]; \mathbb{R}^n) = \{u \in C([a, b]; \mathbb{R}^n) : u \text{ differentiable and } u' \in C([a, b]; \mathbb{R}^n)\}.$$

  We remark that in the endpoints, the one-sided derivatives $u'_+(a), u'_-(b)$ are taken.

- The **piecewise continuously differentiable** functions we denote by

$$D^1([a, b]; \mathbb{R}^n) = \Big\{u \in C([a, b]; \mathbb{R}^n) : \exists a = x_0 < x_1 < \cdots, x_N = b : \\ u|_{[x_{i-1}, x_i]} \in C^1([x_{i-1}, x_i]; \mathbb{R}^n), i = 1, \ldots, N\Big\}.$$

- Furthermore, we frequently consider functions that vanish in the endpoints,

$$C_0^1([a,b];\mathbb{R}^n) = \{u \in C^1([a,b];\mathbb{R}^n) : u(a) = u(b) = 0\},$$
$$C_c^1([a,b];\mathbb{R}^n) = \{u \in C^1([a,b];\mathbb{R}^n) : \operatorname{supp}(u) \subset (a,b) \text{ is compact}\}.$$

  Here, the support of u is defined as $\operatorname{supp}(u) = \overline{\{x \in (a,b) : u(x) \neq 0\}}$.

  Analogously, we define the classes of functions $D_0^1([a,b];\mathbb{R}^n), D_c^1([a,b];\mathbb{R}^n), C_0([a,b];\mathbb{R}^n)$ and $C_c([a,b];\mathbb{R}^n)$.

For these spaces we consider the following norms. For $u \in C([a,b];\mathbb{R}^n)$ we denote the maximum norm by

$$\|u\|_{C^0([a,b])} = \|u\|_{C^0} = \|u\|_\infty = \max_{x\in[a,b]} |u(x)|,$$

for $u \in C^1([a,b];\mathbb{R}^n)$ we consider the norm

$$\|u\|_{C^1([a,b])} = \|u\|_{C^1} = \|u\|_\infty + \|u'\|_\infty,$$

and for $u \in D^1([a,b];|R^n)$ we define

$$\|u\|_{D^1([a,b])} = \|u\|_{D^1} = \|u\|_\infty + \max_{i=1,\dots,N} \left\{\|u'\|_{C^0([x_{i-1},x_i])}\right\}$$

We remark that the convergence in $C^0([a,b];\mathbb{R}^n)$ with respect to $\|\cdot\|_{C^0}$ is equivalent to uniform convergence. Moreover, the spaces $C^0([a,b];\mathbb{R}^n)$ and $C^1([a,b];\mathbb{R}^n)$ are complete and hence, they are Banach spaces. However, this is *not* the case for $D^1([a,b];\mathbb{R}^n)$.

Remark 2.2. From a theoretical point of view one may want to minimize I in the largest possible class of functions Φ for which (2.1) is well-defined. This is the class of **absolutely continuous functions** $AC([a,b];\mathbb{R}^n)$. We recall that a function $u \in C([a,b];\mathbb{R}^n)$ is absolutely continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every finite sequence of pairwise disjoint subintervals, $a \le x_1 < y_1 \le x_2 < y_2 \le \cdots \le x_N < y_N \le b, N \in \mathbb{N}$, with $\sum_{i=1}^N (y_i - x_i) < \delta$, it follows that

$$\sum_{i=1}^N (u(y_i) - u(x_i)) < \varepsilon.$$

If $u \in AC([a,b];\mathbb{R}^n)$, one can show that its derivative u' exists a.e. in $[a,b]$, $u' \in L^1((a,b);\mathbb{R}^n)$ and the *fundamental theorem of calculus* holds, i.e.

$$u(x) = u(a) + \int_a^x u'(t)dt \qquad \forall x \in [a,b].$$

Our assumptions on f then ensure the measurability of $x \mapsto f(x, u(x), u'(x))$.

Finally, we remark that the space $AC([a,b];\mathbb{R}^n)$ can be identified with the Sobolev Space $W^{1,1}((a,b);\mathbb{R}^n)$. Sobolev spaces will be introduced in Chapter 3.

In the sequel, we only consider the function classes $C^1([a,b];\mathbb{R}^n)$ and $D^1([a,b];\mathbb{R}^n)$, but most of the results in this chapter can be generalized for absolutely continuous functions.

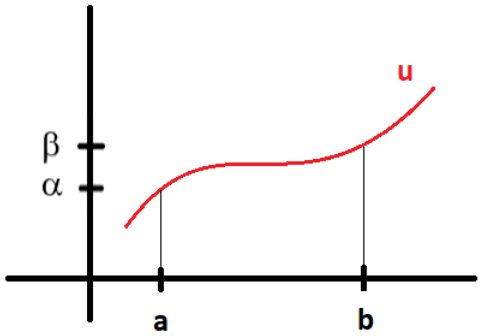
In many applications, $u \in \Phi$ additionally has to satisfy certain *boundary conditions*, e.g.

$$u(a) = \alpha, \qquad u(b) = \beta, \qquad \alpha, \beta \in \mathbb{R}^n \text{ given.}$$

If we prescribe the values of the function u in the endpoints of the interval, we call these conditions **Dirichlet boundary conditions.**

Example 2.3. We discuss several classical examples for minimization problems.

(i) We aim to minimize the **length of the graph** of a continuously differentiable function $u : [a,b] \to \mathbb{R}$, $a < b$, with prescribed boundary conditions, $u(a) = \alpha$, $u(b) = \beta$, where $\alpha, \beta \in \mathbb{R}$ are given.



Hence, we look for a minimizer of the functional

$$I(u) = \int_a^b \sqrt{1 + (u'(x))^2}dx$$

within the class $\Phi = \{u \in C^1([a,b];\mathbb{R}) : u(a) = \alpha, u(b) = \beta\}$.

(ii) **Fermat's principle** states that a light ray travels from one point $A = (a, \alpha) \in \mathbb{R}^2$ to another point $B = (b, \beta) \in \mathbb{R}^2$ along the path that can be traversed in the shortest time.

Assume that the light ray travels along the graph of a function $u \in D^1([a,b];\mathbb{R})$ with $u(a) = \alpha$, $u(b) = \beta$ in time $T > 0$. Then, the distance traveled at time $t < T$ is given by

$$s(t) = \int_0^{x(t)} \sqrt{1 + (u'(z))^2}dz. \tag{2.2}$$

The speed of light in a medium is given by $\frac{c}{n}$, where c is the speed of light in vacuum and n the index of refraction. In general, the index of refraction depends on the position, $(x, u(x))$, and the travel direction, i.e. $u'(x)$. Consequently, using (2.2) we obtain

$$\frac{c}{n(x, u(x), u'(x))} = \frac{ds}{dt} = \sqrt{1 + (u'(x))^2}\frac{dx}{dt},$$

and this implies that

$$\begin{aligned} T &= \int_0^T dt = \frac{1}{c}\int_0^T n(x(t), u(x(t)), u'(x(t)))\sqrt{1 + (u'(x(t)))^2}\frac{dx(t)}{dt}dt \\ &= \frac{1}{c}\int_a^b n(x, u(x), u'(x))\sqrt{1 + (u'(x))^2}dx = I(u). \end{aligned}$$

Hence, to determine the path of the light ray we aim to find a minimizer of the functional I within the class

$$\Phi = \left\{u \in D^1([a,b];\mathbb{R}) : u(a) = \alpha, u(b) = \beta\right\}.$$

(iii) **Brachistochrone**: We aim to construct the fastest slide that starts at the point $A = (0,0) \in \mathbb{R}^2$ and ends at the point $B = (b, \beta)$ with given $b > 0$ and $\beta < 0$. Assume that in time $T > 0$ a mass point moves without friction in the gravity field from A to B along the graph of a function u with $u(0) = 0$ and $u(b) = \beta$.

Energy conservation implies that

$$\text{gain in kinetic energy} = \text{loss of potential energy},$$
$$\frac{1}{2}mv^2 = -mgu,$$

where g is the gravitational constant. Hence, we obtain, as in the case of Fermat's principle,

$$v = \sqrt{-2gu} = \frac{ds}{dt} = \sqrt{1 + (u'(x))^2}\frac{dx}{dt},$$

where s is the distance traveled at time $t > 0$. To find the shape of the slide we need to minimize the travel time

$$T = \int_0^T dt = \int_0^b \sqrt{\frac{1 + (u'(x))^2}{-2gu(x)}}dx$$

within the class $\Phi = \{u \in AC([0,b];\mathbb{R}) : u(0) = 0, u(b) = \beta\}$. We remark that the classes $C^1([a,b];\mathbb{R})$ and $D^1([a,b];\mathbb{R})$ would be too restrictive in this case, since this excludes the possibility that the slope of u becomes infinitely steep in $x = 0$.

2.2 First variation and Euler-Lagrange equations

Let $a < b$. We consider the functional

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

where $f \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n;\mathbb{R})$, $(x,y,z) \mapsto f(x,y,z)$, and assume that there exists a minimizer u in the class

$$\Phi = \{u \in D^1([a,b];\mathbb{R}^n) : u(a) = \alpha, u(b) = \beta\},$$

where $\alpha, \beta \in \mathbb{R}^n$ are given. Then, for all $\eta \in C_c^\infty([a,b];\mathbb{R}^n)$ and $s \in (-\varepsilon, \varepsilon)$ we have $u + s\eta \in \Phi$ and $I(u + s\eta) \geq I(u)$, as u is a minimizer of I. The real-valued function $s \mapsto I(u + s\eta)$, $s \in (-\varepsilon, \varepsilon)$, is continuously differentiable and has a minimum at $s = 0$. Therefore, we conclude that

$$\begin{aligned} 0 &= \frac{d}{ds}(I(u + s\eta))\Big|_{s=0} \\ &= \int_a^b f_y(x, u(x), u'(x)) \cdot \eta(x) + f_z(x, u(x), u'(x)) \cdot \eta'(x) dx, \end{aligned} \tag{2.3}$$

where we interchanged differentiation and integration. Here, $f_y = (f_{y_1}, \dots f_{y_n})$ and $f_z = (f_{z_1}, \dots f_{z_n})$ denote the partial derivatives of f with respect to y and z, and $\cdot$ the inner product in $\mathbb{R}^n$.

Definition 2.4. Let $u \in D^1([a,b];\mathbb{R}^n)$. For $\eta \in D^1([a,b];\mathbb{R}^n)$,

$$\delta I(u, \eta) = \int_a^b f_y(x, u(x), u'(x)) \cdot \eta(x) + f_z(x, u(x), u'(x)) \cdot \eta'(x) dx$$

is called the **first variation of I at u in direction of η.**

If $\delta I(u, \eta) = 0$ for all $\eta \in C_c^\infty([a,b];\mathbb{R}^n)$, then u is called a **weak extremal of I.**

The calculations above show that if $u \in \Phi$ is a minimizer, then u is a weak extremal of I. We observe that this also holds for local minimizers. A function $u \in \Phi$ is a **local minimizer of** I, if there exists $\varepsilon > 0$ such that for all $v \in \Phi$ with $\|u - v\|_{D^1([a,b])} < \varepsilon$, it follows that $I(u) \leq I(v)$.

If, in addition, we assume that $f_z \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ and $u \in C^2([a,b];\mathbb{R}^n)$, we can apply integration by parts in (2.3) and obtain the following theorem.

Theorem 2.5. *If $u \in C^2([a,b];\mathbb{R}^n)$ is a weak extremal of I and $f, f_z \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$, then u satisfies the* ***Euler-Lagrange equations***

$$\frac{d}{dx}\left(f_z(x, u(x), u'(x))\right) - f_y(x, u(x), u'(x)) = 0. \tag{2.4}$$

Proof. The theorem follows from applying integration by parts in (2.3) and the Fundamental Lemma of the Calculus of Variations (Lemma 2.6). □

Lemma 2.6 (Fundamental Lemma of the Calculus of Variations)**.** *If a function $u \in C([a,b];\mathbb{R}^n)$ satisfies*

$$\int_a^b u(x) \cdot \eta(x) dx = 0 \qquad \forall \eta \in C_c^\infty([a,b];\mathbb{R}^n),$$

then $u \equiv 0$ on $[a,b]$.

Proof. By contradiction, we assume that $u \not\equiv 0$ in $[a,b]$. Then, there exists $x_0 \in (a,b)$ and $i \in \{1, \dots, n\}$ such that $u_i(x_0) \neq 0$. Since u is continuous, there exists $\delta > 0$ such that $a < x_0 - \delta < x_0 + \delta < b$ and

$$u_i(x) > \frac{1}{2}u_i(x_0) > 0 \qquad \text{or} \qquad u_i(x) < \frac{1}{2}u_i(x_0) < 0 \qquad \forall x \in (x_0 - \delta, x_0 + \delta).$$

We now choose a function $\psi \in C_c^\infty([a,b];\mathbb{R}^n)$ with

$$\operatorname{supp}(\psi_i) \subset (x_0 - \delta, x_0 + \delta), \qquad \psi_i(x_0) > 0, \qquad \psi_i \geq 0 \quad \text{in } (x_0 - \delta, x_0 + \delta)$$

and $\psi_j \equiv 0$ for all $j \neq i$. It then follows that

$$\int_a^b u(x) \cdot \psi(x) dx = \int_{x_0-\delta}^{x_0+\delta} u_i(x)\psi_i(x) dx \neq 0,$$

which is a contraction. □

Later, we will show that Lemma 2.6 remains valid for functions that are locally integrable, i.e. $u \in L^1_{loc}((a,b))$, and generalize it for higher dimensions.

Remark 2.7.

- Writing out the Euler-Lagrange equations (2.4) yields

$$\begin{aligned} f_{zz}(x, u(x), u'(x))u''(x) &+ f_{zy}(x, u(x), u'(x))u'(x) \\ &+ f_{zx}(x, u(x), u'(x)) - f_y(x, u(x), u'(x)) = 0. \end{aligned}$$

  Here we use the notation $f_{yy} = \left(f_{y_i y_j}\right)_{1 \le i,j \le n}$, $f_{zz} = \left(f_{z_i z_j}\right)_{1 \le i,j \le n}$ and $f_{zx} = (f_{z_1 x}, \ldots, f_{z_n x})$. This is a system of second order ordinary differential equations (ODEs) that linearly depend on u''. If the matrix f_{zz} is invertible, we obtain an explicit system of second order ODEs.

- In the particular case that f does not explicitly depend on x, i.e. $f(x, y, z) = f(y, z)$, we observe that

$$\begin{aligned} \frac{d}{dx}\left(f(u, u') - u' \cdot f_z(u, u')\right) &= f_y(u, u') \cdot u' + f_z(u, u') \cdot u'' - u'' \cdot f_z(u, u') - u' \cdot \frac{d}{dx}\left(f_z(u, u')\right) \\ &= u' \cdot \left(f_y(u, u') - \frac{d}{dx}(f_z(u, u'))\right) = 0, \end{aligned}$$

  since u is a solution of the Euler-Lagrange equation. Hence, we have a *first integral*,

$$f(u, u') - u' \cdot f_z(u, u') \equiv c, \tag{2.5}$$

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

Example 2.8. We consider the examples discussed in Example 2.3.

(i) **Graph length:** For the problem of minimizing the length of the graph of a function, we have $f(x, y, z) = \sqrt{1 + z^2}$. The corresponding Euler-Lagrange equation is

$$\begin{aligned} 0 = \frac{d}{dx}\left(\frac{u'(x)}{\sqrt{1 + (u'(x))^2}}\right) &= \frac{u''(x)}{\sqrt{1 + (u'(x))^2}} - \frac{(u'(x))^2 u''(x)}{\left(\sqrt{1 + (u'(x))^2}\right)^3} \\ &= \frac{u''(x)}{\left(\sqrt{1 + (u'(x))^2}\right)^3}, \end{aligned}$$

and therefore, $u''(x) = 0$. This implies that the graph is a straight line.

(ii) **Fermat's principle:** In this case we have $f(x, y, z) = \frac{1}{c} n(x, y, z)\sqrt{1 + z^2}$. In an *isotropic medium* the velocity is independent of the direction, i.e. $n(x, y, z) = n(x, y)$. If we assume that $n \in C^1$, then

$$f_z(x, y, z) = \frac{1}{c} n(x, y) \frac{z}{\sqrt{1 + z^2}}, \qquad f_y(x, y, z) = \frac{1}{c} n_y(x, y)\sqrt{1 + z^2}.$$

The corresponding Euler-Lagrange equation is

$$\frac{d}{dx}\left(n(x,u(x))\frac{u'(x)}{\sqrt{1+(u'(x))^2}}\right) - n_y(x,u(x))\sqrt{1+(u'(x))^2} = 0,$$

which implies that

$$\begin{aligned} n_x(\cdot,u)\frac{u'}{\sqrt{1+u'^2}} + n_y(\cdot,u)\frac{u'^2}{\sqrt{1+u'^2}} + n(\cdot,u)\frac{u''}{\sqrt{1+u'^2}} \\ - n(\cdot,u)\frac{u'^2u''}{\left(\sqrt{1+u'^2}\right)^3} - n_y(\cdot,u)\sqrt{1+u'^2} = 0. \end{aligned}$$

We can rewrite the equation as

$$n_x(\cdot,u)u'\left(1+u'^2\right) - n_y(\cdot,u)\left(1+u'^2\right) + n(\cdot,u)u'' = 0,$$

and if $n \neq 0$, we obtain

$$u'' = -\frac{n_x(\cdot,u)}{n(\cdot,u)}u'\left(1+u'^2\right) + \frac{n_y(\cdot,u)}{n(\cdot,u)}\left(1+u'^2\right). \tag{2.6}$$

(iii) **Brachistochrone problem:** For this problem we have $f(x,y,z) = \frac{1}{\sqrt{-2gy}}\sqrt{1+z^2}$. Hence, the Euler-Lagrange equation (2.6) with $\frac{1}{c}n(x,y) = \frac{1}{\sqrt{-2gy}}$ implies that

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

We remark that the equation holds only formally. Due to the singularity at $u = 0$, Theorem 2.5 cannot be applied.

We note that f does not explicitly depend on x. Hence, we can apply (2.5) in Remark 2.7 and conclude that there exists $c \in \mathbb{R}$ such that

$$\sqrt{\frac{1+u'^2}{-2gu}} - \frac{u'^2}{\sqrt{-2gu}}\frac{1}{\sqrt{1+u'^2}} = c,$$

i.e.

$$u(1+u'^2) = \frac{1}{-2gc^2}.$$

Remark 2.9. The Euler-Lagrange equations are only a *necessary condition* for minimizers, but *not a sufficient condition* for the existence of a minimizer. Not all functionals that are bounded from below posses a minimizer. Furthermore, we remark that minimizers are not always of class C^2. For counterexamples we refer to the tutorials.

Later, we address *direct methods* that allow to prove the existence of minimizers in suitable (weaker) function spaces.

To derive the Euler-Lagrange equations under weaker regularity assumptions, we need the following lemma.

Lemma 2.10 (Du Bois-Reymond)**.** *Let $u : [a,b] \to \mathbb{R}^n$ be a piecewise continuous function. If*

$$\int_a^b u(x) \cdot \eta'(x) dx = 0 \qquad \forall \eta \in C_c^\infty([a,b];\mathbb{R}^n),$$

then, there exists $c \in \mathbb{R}$ such that $u \equiv c$.

Proof. By an approximation argument one can show that

$$\int_a^b u(x) \cdot \eta'(x) dx = 0 \qquad \forall \eta \in D_0^1([a,b];\mathbb{R}^n).$$

For the proof of this statement we refer to the tutorials. It suffices to show that for every $\psi \in D_0^1([a,b];\mathbb{R})$ there exists a sequence of functions $\psi_m \in C_c^\infty((a,b);\mathbb{R})$ such that

$$\lim_{m\to\infty} \int_a^b |\psi_n'(x) - \psi'(x)| dx = 0.$$

To this end we approximate $\psi \in D_0^1([a,b];\mathbb{R})$ first with a function $\tilde{\psi} \in D_c^1([a,b];\mathbb{R})$ and then construct the sequence ψ_m by convolution of $\tilde{\psi}$ with mollifiers (see Chapter 3).

Let now $c = \frac{1}{b-a} \int_a^b u(x) dx$ and $\psi(x) = \int_a^x (u(s) - c) ds$. Then, $\psi \in D_0^1([a,b];\mathbb{R}^n)$ and we conclude that

$$0 = \int_a^b u(x) \cdot \psi'(x) dx = \int_a^b u(x) \cdot (u(x) - c) dx = \int_a^b (u(x) - c) \cdot (u(x) - c) dx,$$

where we used that $\psi(b) = \int_a^b (u(x) - c) dx = 0$. Consequently, $u \equiv c$ in $[a,b]$, which concludes the proof. □

Theorem 2.11. *Let the function $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, (x,y,z) \mapsto f(x,y,z)$, be continuous with respect to x and continuously differentiable with respect to y and z. If $u \in D^1([a,b];\mathbb{R}^n)$ is a weak extremal of I, then $f_z(\cdot, u, u') \in D^1([a,b];\mathbb{R}^n)$ and u satisfies*

$$\frac{d}{dx} \left(f_z(x, u(x), u'(x)) \right) - f_y(x, u(x), u'(x)) = 0 \qquad \textit{for a.e. } x \in [a,b]. \tag{2.7}$$

Proof. Let $\eta \in C_c^\infty([a,b];\mathbb{R}^n)$ and u be a weak extremal of I. Integration by parts implies that

$$\begin{aligned} \int_a^b f_y(x, u(x), u'(x)) \cdot \eta(x) dx &= \int_a^b \left(\frac{d}{dx} \int_a^x f_y(s, u(s), u'(s)) ds \right) \cdot \eta(x) dx \\ &= - \int_a^b \int_a^x f_y(s, u(s), u'(s)) ds \cdot \eta'(x) dx. \end{aligned}$$

Since u is a weak extremal, $\delta I(u, \eta) = 0$. Using this and the equality above we conclude that

$$0 = \int_a^b \left(- \int_a^x f_y(s, u(s), u'(s)) ds + f_z(x, u(x), u'(x)) \right) \cdot \eta'(x) dx.$$

Lemma 2.10 now implies that there exists $c \in \mathbb{R}$ such that

$$f_z(x, u(x), u'(x)) = \int_a^x f_y(s, u(s), u'(s)) ds + c, \tag{2.8}$$

i.e. $f_z(\cdot, u, u') \in D^1([a,b];\mathbb{R}^n)$. Moreover, differentiating the equation implies that u satisfies (2.7). □

The equation (2.8) in the proof of Theorem 2.11 is called the *integrated form of the Euler-Lagrange equations*.

We remark that Lemma 2.10 remains valid for functions $u \in L^1_{loc}((a, b); \mathbb{R}^n)$, which implies that Theorem 2.11 also holds for functions $u \in AC([a, b]; \mathbb{R}^n)$.

Regularity of minimizers

Under suitable assumptions on f one can show that weak extremals are of class C^2 and hence, they are solutions of the Euler-Lagrange equations (2.4).

Theorem 2.12. *Let $f, f_z \in C^1([a, b] \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ and $u \in C^1([a, b]; \mathbb{R}^n)$ be a weak extremal of I. If*

$$\det(f_{zz}(x, u(x), u'(x))) = \det\left((f_{z_i z_j}(x, u(x), u'(x))_{1\le i,j\le n}\right) \neq 0 \qquad \forall x \in [a, b],$$

then, $u \in C^2([a, b]; \mathbb{R}^n)$.

Proof. We define $p(x) := \int_a^x f_y(s, u(s), u'(s))ds + c$, which is the right hand side of (2.8), and consider the function

$$\psi : \mathbb{R}^n \times [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (w, x, y, z) \mapsto f_z(x, y, z) - w,$$

which is continuously differentiable. For arbitrary $x_0 \in [a, b]$ we set $y_0 = u(x_0), z_0 = u'(x_0)$ and $w_0 = p(x_0)$. Then, (2.8) implies that $\psi(w_0, x_0, y_0, z_0) = 0$.

By assumption, $\det(\psi_z(w_0, x_0, y_0, z_0)) = \det(f_{zz}(x_0, y_0, z_0)) \neq 0$, and therefore, the Implicit Function Theorem implies that there exists a neighborhood U of (w_0, x_0, y_0) and V of z_0 and a function $\varphi : U \to V$ of class C^1 such that

$$\psi(w, x, y, z) = 0 \qquad \Longleftrightarrow \qquad z = \varphi(w, x, y) \qquad \forall (w, x, y, z) \in U \times V. \tag{2.9}$$

Equation (2.8) implies that

$$\psi(p(x), x, u(x), u'(x)) = 0 \qquad \forall x \in [a, b],$$

and by continuity, we can conclude that $(p(x), x, u(x), u'(x)) \in U \times V$ for all x that are sufficiently close to x_0. Therefore, from (2.9) it follows that

$$u'(x) = \varphi(p(x), x, u(x)),$$

which shows that u' is continuously differentiable in a neighborhood of x_0. Since x_0 was arbitrary, the theorem follows. □

Next, we prove regularity for weak extremals $u \in D^1([a, b]; \mathbb{R}^n)$. In this case, stronger hypotheses are needed and the arguments are more involved.

Theorem 2.13. *Let $f, f_z \in C^1([a, b] \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ and $u \in D^1([a, b]; \mathbb{R}^n)$ be a weak extremal of I. Moreover, we assume that*

$$f_{zz}(x, u(x), u'(x)) = \left(f_{z_i z_j}(x, u(x), u'(x)\right)_{1\le i,j\le n}$$

is positive (or negative) definite on $\Omega \times \mathbb{R}^n$, where $\Omega \subset \mathbb{R}^{n+1}$ is a subset that contains the set $\{(x, u(x)) : x \in [a, b]\}$. Then, $u \in C^2([a, b]; \mathbb{R}^n)$.

Proof. We repeat the arguments in the proof of Theorem 2.12 until we arrive at (2.9). However, if $u \in D^1([a,b];\mathbb{R}^n)$ we cannot conclude that $(p(x), x, u(x), u'(x)) \in U \times V$ for all x that are sufficiently close to x_0, since the derivative u' might be discontinuous.

Let $x \in [a,b]$ be such that $(p(x), x, u(x)) \in U \cap (\mathbb{R}^n \times \Omega)$. Then, we conclude that

$$\psi(p(x), x, u(x), \varphi(p(x), x, u(x))) = 0 = \psi(p(x), x, u(x), u'(x)).$$

By the definition of ψ it follows that

$$\begin{aligned} 0 &= f_z(x, u(x), \varphi(p(x), x, u(x))) - f_z(x, u(x), u'(x)) \\ &= \int_0^1 \frac{d}{ds} \left(f_z(x, u(x), u'(x) + s\left(\varphi(p(x), x, u(x)) - u'(x)\right))\right) ds \\ &= \int_0^1 f_{zz}\left(x, u(x), u'(x) + s\left(\varphi(p(x), x, u(x)) - u'(x)\right)\right) ds \cdot \left(\varphi(p(x), x, u(x)) - u'(x)\right). \end{aligned}$$

The integral (matrix) is positive (or negative) definite, since f_{zz} is positive (or negative) definite, and therefore, it is invertible. We conclude that

$$\varphi(p(x), x, u(x)) - u'(x) = 0,$$

which implies that $u' = \varphi(p, \cdot, u)$ is piecewise continuously differentiable, since $p \in D^1([a,b];\mathbb{R}^n)$.

It follows that u' is continuous and hence, that $u \in C^1([a,b];\mathbb{R}^n)$. We can now apply Theorem 2.12 to conclude that $u \in C^2([a,b];\mathbb{R}^n)$. □

If $f_y(\cdot, u, u')$ and $f_z(\cdot, u, u')$ are integrable, one can show that Theorem 2.13 remains valid for functions $u \in AC([a,b];\mathbb{R}^n)$.

Corollary 2.14. *Under the assumptions of Theorem 2.12 or Theorem 2.13, the weak extremal satisfies the Euler-Lagrange equations* (2.4).

Proof. This is an immediate consequence of Theorem 2.12 or Theorem 2.13 and Theorem 2.5. □

For smoother functions f one can show that the weak extremal u is even more regular.

Theorem 2.15. *Let the assumptions of Theorem 2.12 or Theorem 2.13 be satisfied. If* $f, f_z \in C^k([a,b] \times \mathbb{R}^n \times \mathbb{R}^n;\mathbb{R})$, $k \in \mathbb{N}$, *then the weak extremal* u *is in* $C^{k+1}([a,b];\mathbb{R}^n)$.

Proof. We assume that $k \geq 2$. By Theorems 2.12, 2.13 and Theorem 2.5, it follows that $u \in C^2([a,b];\mathbb{R}^n)$ and u satisfies the Euler-Lagrange equations, i.e.

$$f_{zz}(\cdot, u, u')u'' + f_{zy}(\cdot, u, u') \cdot u' + f_{zx}(\cdot, u, u') - f_y(\cdot, u, u') = 0.$$

Since f_{zz} is invertible, we can rewrite the equation as

$$u'' = (f_{zz}(\cdot, u, u'))^{-1}\left(-f_{zy}(\cdot, u, u') \cdot u' - f_{zx}(\cdot, u, u') + f_y(\cdot, u, u')\right).$$

We observe that the right hand side is continuously differentiable, since by assumption $u \in C^2([a,b];\mathbb{R}^n)$ and f, f_z are of class C^2. Therefore, $u'' \in C^1([a,b];\mathbb{R}^n)$ which implies that $u \in C^3([a,b];\mathbb{R}^n)$. The statement of the theorem now follows by iteration. Indeed, if $u \in C^j([a,b];\mathbb{R}^n)$ and $f, f_z \in C^j([a,b] \times \mathbb{R}^n \times \mathbb{R}^n;\mathbb{R})$, $j \leq k$, then the right hand side of the equation is of class $C^{j-1}([a,b];\mathbb{R}^n)$, which implies that $u'' \in C^{j-1}([a,b];\mathbb{R}^n)$, i.e. $u \in C^{j+1}([a,b];\mathbb{R}^n)$. □

We remark that if the invertibility conditions on f_{zz} are not satisfied, then the regularity results may not hold (for examples we refer to the tutorials). On the other hand, if the invertibility assumptions hold, the Euler-Lagrange equations allow to express u'' in terms of u and u'.

Natural boundary conditions

So far, we considered variational problems with Dirichlet boundary conditions, i.e. the functions satisfy $u(a) = \alpha, u(b) = \beta$, for given $\alpha, \beta \in \mathbb{R}^n$. If no such conditions are imposed in one (or both) endpoints of the interval $[a, b]$, we have variational problems with so-called *free boundary values.*

Assume that both ends are free and that u is a (local) minimizer of the functional

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

in $\Phi = D^1([a, b]; \mathbb{R}^n)$. Then, it follows by repeating the arguments in the beginning of Section 2.2, that

$$\delta I(u, \eta) = 0 \qquad \forall \eta \in C^\infty([a, b]; \mathbb{R}^n).$$

However, note that this condition not only holds for all $\eta \in C_c^\infty([a, b]; \mathbb{R}^n)$, but for all $\eta \in C^\infty([a, b]; \mathbb{R}^n)$. We now show that this enforces certain boundary conditions on u.

Theorem 2.16. *Let $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z)$, be continuous in x and continuously differentiable with respect to y and z. If $u \in D^1([a, b]; \mathbb{R}^n)$ and*

$$\delta I(u, \eta) = 0 \qquad \forall \eta \in C^\infty([a, b]; \mathbb{R}^n),$$

then, u satisfies the boundary conditions

$$f_z(b, u(b), u'(b)) = f_z(a, u(a), u'(a)) = 0.$$

These boundary conditions are called ***natural boundary conditions****.*

Since $C_c^\infty([a, b]; \mathbb{R}^n) \subset C^\infty([a, b]; \mathbb{R}^n)$, under the hypotheses of Theorem 2.16 it follows that u satisfies the Euler-Lagrange equations a.e. in $[a, b]$ (see Theorem 2.11).

Proof. Let $\eta \in C^\infty([a, b]; \mathbb{R}^n)$. Using integration by parts and the fact that $f_z(\cdot, u, u') \in D^1([a, b]; \mathbb{R}^n)$ by Theorem 2.12, we conclude that

$$\begin{aligned}
0 = \delta I(u, \eta) &= \int_a^b f_y(x, u(x), u'(x)) \cdot \eta(x) + f_z(x, u(x), u'(x)) \cdot \eta'(x)dx \\
&= f_z(x, u(x), u'(x)) \cdot \eta(x)\Big|_{x=a}^{x=b} + \int_a^b \left(f_y(x, u(x), u'(x)) - \frac{d}{dx}(f_z(x, u(x), u'(x)) \right) \cdot \eta(x)dx \\
&= f_z(a, u(a), u'(a)) \cdot \eta(a) - f_z(b, u(b), u'(b)) \cdot \eta(b).
\end{aligned}$$

In the last step we used that u satisfies the Euler-Lagrange equations.

Finally, choosing first the function $\eta(x) = \frac{x-a}{b-a} f_z(b, u(b), u'(b))$, and then the function $\eta(x) = \frac{x-b}{a-b} f_z(a, u(a), u'(a))$, it follows that

$$0 = \left| f_z(b, u(b), u'(b)) \right|^2 = \left| f_z(a, u(a), u'(a)) \right|^2,$$

which proves the theorem. □

The case of one free and one fixed endpoint can be shown accordingly.

Example 2.17. We consider the Brachistochrone problem, where the starting point $A = (0, 0)$ and the horizontal coordinate $b > 0$ of the endpoint B are given, but the value of the vertical coordinate is not prescribed. We then aim to find the shape of the fastest slide from A to a point $B = (x, y)$ with $x = b$. In this case, the minimizer satisfies the boundary conditions

$$u(0) = 0, \qquad u'(b) = 0.$$

2.3 Inner variations and corner conditions

We consider the functional I in (2.1). In this section, we look at *inner variations*, i.e. suitable variations of the variable x and not of the minimizer u as before, and derive additional conditions. Let $\eta \in C_c^\infty([a, b]; \mathbb{R})$ and the function $\psi : [a, b] \times (-s_0, s_0)$ be defined by

$$\psi(x, s) = x + s\eta(x).$$

If $s_0 > 0$ is small enough, then $\partial_x\psi(x, s) > 0$ for all $(x, s) \in [a, b] \times (-s_0, s_0)$, which implies that $\psi(\cdot, s)$ is a family of diffeomorphisms. Moreover, we observe that

$$\begin{aligned} &\psi(a, s) = a \qquad && \psi(b, s) = b \qquad && \forall s \in (-s_0, s_0),\\ &\psi(x, 0) = x && \forall x \in [a, b], \end{aligned}$$

and call ψ an *admissible parameter variation*.

Indeed, if we consider the the class

$$u \in \Phi = \{u \in D^1([a, b]; \mathbb{R}^n) : u(a) = \alpha, u(b) = \beta\},$$

for some given $\alpha, \beta \in \mathbb{R}^n$, then the function

$$v(\cdot, s) = u \circ \psi(\cdot, s) \in \Phi \qquad \forall s \in (-s_0, s_0).$$

Moreover, if $u \in \Phi$ is a local minimizer of the functional

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

then, we conclude that

$$\frac{d}{ds}\left(I(v(\cdot, s))\right)\Big|_{s=0} = 0.$$

Evaluating the right hand side we will obtain additional conditions for the minimizer. To this end we denote the inverse of $\psi(\cdot, s)$ by $\tau(\cdot, s) = \psi^{-1}(\cdot, s)$ and observe that

$$\partial_x v(x, s) = u'(\psi(x, s))\partial_x\psi(x, s) = u'(\psi(x, s))\frac{1}{\partial_x\tau(\psi(x, s))}.$$

This implies that

$$\begin{aligned} I(v(x, s)) &= \int_a^b f\Big(x, v(x, s), \partial_x(v(x, s))\Big)dx\\ &= \int_a^b f\left(\tau(x, s), u(x), \frac{u'(x)}{\partial_x\tau(x, s)}\right)\partial_x\tau(x, s)dx, \end{aligned}$$

where we used the substitution $x \mapsto \tau(x,s)$. Hence, we obtain

$$\begin{aligned}\frac{d}{ds}\left(I(v(x,s))\right)\Big|_{s=0} = \int_a^b \Bigg(& f_x\left(\tau(x,0), u(x), \frac{u'(x)}{\partial_x\tau(x,0)}\right)\partial_s\tau(x,0)\partial_x\tau(x,0) \\ &+ f_z\left(\tau(x,0), u(x), \frac{u'(x)}{\partial_x\tau(x,0)}\right)\cdot u'(x)\frac{(-\partial_s\partial_x\tau(x,0))}{(\partial_x\tau(x,0))^2}\partial_x\tau(x,0) \\ &+ f\left(\tau(x,0), u(x), \frac{u'(x)}{\partial_x\tau(x,0)}\right)\partial_s\partial_x\tau(x,0)\Bigg)dx.\end{aligned} \tag{2.10}$$

To simplify the expression we observe that $\tau(x,0) = x = \psi(x,0)$ and $\partial_x\tau(x,0) = 1$. Moreover, differentiating the equation

$$\tau(\psi(x,s),s) = x$$

with respect to s and evaluating it in $s = 0$ implies that

$$\begin{aligned}0 &= \partial_x\tau(\psi(x,0),0)\partial_s\psi(x,0) + \partial_s\tau(\psi(x,0),0) = \partial_x\tau(x,0)\eta(x) + \partial_s\tau(x,0) \\ &= \eta(x) + \partial_s\tau(x,0).\end{aligned}$$

Consequently, it follows that

$$\partial_s\tau(x,0) = -\eta(x).$$

Similarly, by differentiating the equation $\psi(\tau(x,s),s) = x$ first with respect to s and then with respect to x and evaluating it in $s = 0$ we find

$$\partial_x\partial_s\tau(x,0) = -\eta'(x).$$

Inserting these identities in (2.10) we finally obtain

$$\begin{aligned}\frac{d}{ds}\left(I(v(\cdot,s))\right)\Big|_{s=0} = \int_a^b \Big(& -f_x\left(x,u(x),u'(x)\right)\eta(x) + f_z\left(x,u(x),u'(x)\right)\cdot u'(x)\,\eta'(x) \\ & - f\left(x,u(x),u'(x)\right)\eta'(x)\Big)dx.\end{aligned}$$

This motivates the following definition.

Definition 2.18. Let $u \in D^1([a,b];\mathbb{R}^n)$ and $\eta \in D^1([a,b];\mathbb{R})$. The functional

$$\partial I(u,\eta) = \int_a^b \Big(-f_x\left(\cdot,u,u'\right)\eta + \left(f_z\left(\cdot,u,u'\right)\cdot u' - f\left(\cdot,u,u'\right)\right)\eta'\Big)$$

is called the **first inner variation of** u **in direction** η**.**

Proposition 2.19. Let $f \in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n;\mathbb{R})$ and $u \in \Phi$ be a local minimizer of I, then

$$\partial I(u,\eta) = 0 \qquad \forall \eta \in C_c^\infty([a,b];\mathbb{R}).$$

Proof. This immediately follows from the derivation above. □

We now state the main result in this section which provides additional conditions for minimizers u in point where the derivative u' is discontinuous.

Theorem 2.20. *Let $f \in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n;\mathbb{R})$ and $u \in D^1([a,b];\mathbb{R}^n)$ be such that*

$$\partial I(u,\eta) = 0 \qquad \forall \eta \in C_c^\infty([a,b];\mathbb{R}).$$

Then, there exists $c \in \mathbb{R}$ such that

$$f(x,u(x),u'(x)) - u'(x)\cdot f_z(x,u(x),u'(x)) = c + \int_0^x f_x(t,u(t),u'(t))dt.$$

Proof. By assumption, for all $\eta \in C_c^\infty([a,b];\mathbb{R})$ we have

$$\begin{aligned} 0 &= \int_a^b \Big(- f_x(\cdot,u,u')\,\eta + (f_z(\cdot,u,u')\cdot u' - f(\cdot,u,u'))\,\eta' \Big) \\ &= \int_a^b \left(\int_a^{\cdot} f_x(\cdot,u,u') + f_z(\cdot,u,u')\cdot u' - f(\cdot,u,u') \right) \eta', \end{aligned}$$

where we used integration by parts. The statement now follows from Lemma 2.10. □

Corollary 2.21. *Let $f \in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n;\mathbb{R})$ and $u \in \Phi$ be a local minimizer of I. Then, the functions*

$$f_z(\cdot,u,u') \qquad \text{and} \qquad f(\cdot,u.u') - u'\cdot f_z(\cdot,u,u')$$

are continuous.

Proof. The continuity of the function $f_z(\cdot,u,u')$ was shown in Theorem 2.11. In fact, we even have $f_z(\cdot,u,u') \in D^1([a,b];\mathbb{R}^n)$. Moreover, the continuity of the function $f(\cdot,u.u') - u'\cdot f_z(\cdot,u,u')$ follows from Theorem 2.20, since the right hand side of the equation is continuous (even piecewise continuously differentiable). □

If the minimizer u is in $D^1([a,b];\mathbb{R}^n)$, these continuity results yield conditions for those points $x \in (a,b)$ where u' is discontinuous, i.e. where u has a "corner". In fact, in such a point $x \in (a,b)$ it follows that

$$f_z(x,u(x),u'_-(x)) = f_z(x,u(x),u_+(x))$$

and

$$\begin{aligned} &f(x,u(x),u'_-(x)) - u'_-(x)\cdot f_z(x,u(x),u'_-(x)) \\ =&f(x,u(x),u'_+(x)) - u'_+(x)\cdot f_z(x,u(x),u'_+(x)). \end{aligned}$$

These conditions are called the **Erdmann-Weierstraß corner conditions**.

Example 2.22. Consider the functional

$$I(u) = \int_0^2 u^2(x)\,(1+(u'(x))^2)dx$$

with $\Phi = \{u \in D^1([a,b];\mathbb{R}^n) : u(0)=0, u(2)=1\}$. One can show that I possesses no minimizers in $C^2([a,b];\mathbb{R}^n)$. Assuming that a minimizer $u \in D^1([a,b];\mathbb{R}^n)$ exists Corollary 2.21 allows to uniquely determine is (see tutorials).

2.4 Second variation

We consider the functional

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

with $f \in C^2([a, b]; \mathbb{R})$ and

$$\Phi = \{u \in D^1([a, b]; \mathbb{R}^n) : u(a) = \alpha, u(b) = \beta\},$$

for given $\alpha, \beta \in \mathbb{R}^n$.

If $u \in \Phi$ is a local minimizer of I and we assume that $u \in C^1([a, b]; \mathbb{R}^n)$ then for every $\eta \in C_0^1([a, b]; \mathbb{R}^n)$ the function

$$s \mapsto I(u + s\eta), \qquad s \in (-s_0, s_0),$$

has a local minimum at $s = 0$. This implies that

$$\frac{d}{ds}(I(u + s\eta))\Big|_{s=0} = 0 \qquad \text{and} \qquad \frac{d^2}{ds^2}(I(u + s\eta))\Big|_{s=0} \geq 0.$$

While the first condition leads to the first variation, evaluating the second condition leads to the so-called second variation which is an additional necessary condition for the minimizer.

Definition 2.23. For $u, \eta \in D^1([a, b]; \mathbb{R}^n)$ the functional

$$\delta^2 I(u, \eta) = \int_a^b \Big(f_{yy}(\cdot, u, u')[\eta, \eta] + 2 f_{yz}(\cdot, u, u')[\eta, \eta'] + f_{zz}(\cdot, u, u')[\eta', \eta'] \Big)$$

is called the **second variation of I at u in the direction η.**

Here, to shorten notations we use the abbreviation

$$f_{zz}(x, y, z)[\varphi, \zeta] = \sum_{i,j=1}^n f_{z_i z_j}(x, y, z)\varphi_i \zeta_j \qquad \text{for } \zeta, \varphi \in \mathbb{R}^n.$$

Theorem 2.24. *If $u \in \Phi$ is a local minimizer of the functional I, then*

$$\delta^2 I(u, \eta) \geq 0 \qquad \forall \eta \in D_0^1([a, b]; \mathbb{R}^n).$$

Proof. The proof is left as an exercise. The condition can be derived from the fact that the function $s \mapsto I(u + s\eta)$, $s \in (-s_0, s_0)$, $\eta \in C_c^\infty([a, b]; \mathbb{R}^n)$, has a local minimum in $s = 0$, which implies that

$$\frac{d^2}{ds^2}(I(u + s\eta))\Big|_{s=0} \geq 0.$$

Computing the second order derivative and evaluating it at $s = 0$ we obtain the second variation. One can then show that the condition remains valid for all $\eta \in D_0^1([a, b]; \mathbb{R}^n)$. □

Theorem 2.25. *If $u \in \Phi$ is a local minimizer of the functional I, then u satisfies the* ***Legendre condition****, i.e.*

$$f_{zz}(x, u(x), u'_\pm(x)) \qquad \textit{is positive definite} \qquad \forall x \in [a, b].$$

Proof. Let $x_0 \in (a,b)$ and u' be continuous in x_0. Moreover, let $\zeta \in \mathbb{R}^n$ be arbitrary and $\varepsilon > 0$ be such that $(x_0 - \varepsilon, x_0 + \varepsilon) \subset (a,b)$. Then, the piecewise linear and continuous function η, where η is such that $\eta(x_0) = \varepsilon\zeta$ and

$$\begin{aligned} &\eta \equiv 0 && \text{in } [a, x_0 - \varepsilon] \cup [x_0 + \varepsilon, b], \\ &\eta \text{ affine} && \text{in } [x_0 - \varepsilon, x_0 + \varepsilon], \end{aligned}$$

is in $D^1([a,b];\mathbb{R}^n)$.

We observe that $\eta_{x_i} = \pm\zeta$ in $[x_0 - \varepsilon, x_0 + \varepsilon]$, $i = 1, \dots, n$. Theorem 2.24 now implies that

$$\begin{aligned} 0 \leq \frac{\delta^2 I(u,\eta)}{2\varepsilon} &= \frac{1}{2\varepsilon} \int_{x_0-\varepsilon}^{x_0+\varepsilon} \left(f_{zz}(x, u(x), u'(x))[\zeta,\zeta] + O(\varepsilon)\right) dx \\ &= \frac{1}{2\varepsilon} \int_{x_0-\varepsilon}^{x_0+\varepsilon} \left(f_{zz}(x, u(x), u'(x))[\zeta,\zeta]\right) dx + O(\varepsilon) \\ &\xrightarrow[\varepsilon \to 0]{} f_{zz}(x_0, u(x_0), u'(x_0))[\zeta,\zeta]. \end{aligned}$$

If x_0 is a point in $[a,b]$ where u' is discontinuous, or if $x_0 = a$ or $x_0 = b$, then we consider a sequence $(x_n)_{n\in\mathbb{N}_0} \subset (a,b)$ such that $x_n \nearrow x$ or $x_n \searrow x$ as $n \to \infty$. Then, we apply the argument above to the sequence $(x_n)_{n\in\mathbb{N}_0}$, and taking the limit, it follows that

$$f_{zz}(x_0, u(x_0), u'(x_0))[\zeta,\zeta] \geq 0.$$

$\square$

Under suitable assumptions on the second variation, the existence of a minimizer can be shown.

Theorem 2.26. *Let $u \in \Phi$ be a weak extremal of the functional I, where*

$$\Phi = \{u \in C^1([a,b];\mathbb{R}^n) \mid u(a) = \alpha, u(b) = \beta\}$$

and $\alpha, \beta \in \mathbb{R}^n$ are given. If there exists $\lambda > 0$ such that

$$\delta^2 I(u,\eta) \geq \lambda \int_a^b \left(|\eta(x)|^2 + |\eta'(x)|^2\right) dx \qquad \forall \eta \in C_0^1([a,b];\mathbb{R}^n),$$

then, u is a strict local minimizer of I.

Proof. For the proof we refer to the tutorials. $\square$

Chapter 3

Function spaces and tools from functional analysis

In this chapter, we recall properties of Lebesgue and Sobolev spaces and several basic tools and results from functional analysis. If not stated otherwise, $\Omega \subset \mathbb{R}^n$ always denotes an open set, where $\Omega = \mathbb{R}^n$ or $\Omega \subsetneq \mathbb{R}^n$.

- We denote the space of real-valued continuous functions on Ω by $C(\Omega) = C(\Omega; \mathbb{R})$. For $k \in \mathbb{N}$ the space of real-valued k-times continuously differentiable functions on Ω is denoted by $C^k(\Omega) = C^k(\Omega; \mathbb{R})$ and $C^\infty(\Omega) = C^\infty(\Omega; \mathbb{R}) = \bigcap_{k=0}^\infty C^k(\Omega)$.
- Moreover, $C^k(\overline{\Omega}), k \in \mathbb{N}_0$, denotes the subspace of functions in $C^k(\Omega)$ such that the function and its derivatives up to order k can be continuously extended to the boundary $\partial\Omega$, and $C^\infty(\overline{\Omega}) = \bigcap_{k=0}^\infty C^k(\overline{\Omega})$.
- We denote by
$$C_c^k(\Omega) = \{u \in C^k(\Omega) : \operatorname{supp}(u) \subset \Omega \text{ compact}\}$$
the subspace of functions in $C^k(\Omega)$ with compact support. Here, $\operatorname{supp}(u) = \overline{\{x \in \Omega : u(x) \neq 0\}}$ denotes the support of u.

We consider $C^k(\overline{\Omega}), k \in \mathbb{N}_0$, with the norm

$$\|u\|_{C^k(\overline{\Omega})} = \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^\infty}, \qquad u \in C^k(\overline{\Omega}),$$

where $\|u\|_{L^\infty} = \max_{x \in \overline{\Omega}}\{|u(x)|\}$ is the maximum norm of $u : \overline{\Omega} \to \mathbb{R}$. Moreover, we use the multi-index notation to denote partial derivatives,

$$\partial^\alpha u = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u, \qquad \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n,$$

where the order of α is defined as $|\alpha| = \sum_{i=1}^n \alpha_i$.

3.1 L^p-Spaces

In this section, we summarize several important properties of Lebesgue spaces, for proofs and further details we refer to [5] or [7].

Definition 3.1. Let $1 \leq p \leq \infty$. A Lebesgue measurable function $u : \Omega \to \mathbb{R}$ **belongs to $\mathbf{L^p(\Omega)}$** if $\|u\|_{L^p} < \infty$, where

$$\|u\|_{L^p} = \begin{cases} \left(\int_\Omega |u|^p\right)^{\frac{1}{p}} & 1 \leq p < \infty, \\ \operatorname{ess\,sup}_\Omega\{|u|\} & p = \infty. \end{cases}$$

Recall that

$$\operatorname*{ess\,sup}_{\Omega}\{u\} = \inf\left\{\mu \in \mathbb{R} : |u| \leq \mu \text{ } a.e. \text{ in } \Omega\right\}.$$

Moreover, $u \in \mathbf{L^p_{loc}(\Omega)}$, if $u : \Omega \to \mathbb{R}$ is measurable and $u|_V$ belongs to $L^p(V)$ for every open subset V that is compactly contained in Ω, i.e. $\overline{V}$ is compact and $\overline{V} \subset \Omega$. For compactly contained subsets we use the notation $V \subset\subset \Omega$.

The abbreviation *a.e.* means that a property holds almost everywhere, i.e. for almost all $x \in \Omega$. Two functions that coincide a.e. are identified in $L^p(\Omega)$ and hence, the spaces $L^p(\Omega)$ consist of *equivalence classes of functions*.

Example 3.2. Consider the function $u(x) = \frac{1}{|x|}$. Then, $u \notin L^1(\mathbb{R})$ and $u \notin L^1_{loc}(\mathbb{R})$, as

$$\int_0^1 \frac{1}{|x|} dx = \infty.$$

However, for $v(x) = \frac{1}{|x|^r}$, $r > 0$, we have

$$\int_0^1 \frac{1}{|x|^r} dx = \begin{cases} \infty & r \geq 1, \\ \frac{1}{1-r} & r < 1. \end{cases}$$

This implies that $v \in L^p_{loc}(\mathbb{R})$ if and only if $pr < 1$, i.e. $p < \frac{1}{r}$.

Theorem 3.3 (Important inequalities)**.** *Let $1 \leq p \leq \infty$.*

- ***Minkowski inequality**: If $u, v \in L^p(\Omega)$, then*

$$\|u + v\|_{L^p} \leq \|u\|_{L^p} + \|v\|_{L^p}.$$

- ***Hölder inequality**: If $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$, then $uv \in L^1(\Omega)$ and*

$$\|uv\|_{L^1} \leq \|u\|_{L^p}\|v\|_{L^q},$$

  with the convention that $p = 1$ if $q = \infty$ and $p = \infty$ if $q = 1$.

- ***General Hölder inequality**: If $u_j \in L^{p_j}(\Omega)$, $1 \leq p_j \leq \infty$ for $j = 1, \ldots, m$, and $\frac{1}{r} = \sum_{j=1}^m \frac{1}{p_j}$, then $\prod_{j=1}^m u_j \in L^r(\Omega)$ and*

$$\left\| \prod_{j=1}^m u_j \right\|_{L^r} \leq \prod_{j=1}^m \|u_j\|_{L^{p_j}}.$$

Remark 3.4. If Ω is bounded, an important consequence of Hölder's inequality is the *continuous embedding*

$$L^p(\Omega) \hookrightarrow L^q(\Omega) \qquad \forall 1 \le q \le p \le \infty,$$

i.e. $L^p(\Omega) \subset L^q(\Omega)$ and there exists a constant $c > 0$ such that $\|u\|_{L^q} \le c\|u\|_{L^p}$ for all $u \in L^p(\Omega)$.

Definition 3.5. We say that a sequence $(u_m)_{m\in\mathbb{N}} \subset L^p(\Omega)$ **converges (strongly) to** u **in** $L^p(\Omega)$, if $u \in L^p(\Omega)$ and

$$\lim_{m\to\infty} \|u_m - u\|_{L^p} = 0.$$

We then use the notation $u_m \to u$ in $L^p(\Omega)$.

Moreover, a sequence $(u_m)_{m\in\mathbb{N}} \subset L^p_{loc}(\Omega)$ **converges to** u **in** $L^p_{loc}(\Omega)$ if $u_m \to u$ in $L^p(V)$ for every compact subset $V \subset \Omega$.

Theorem 3.6. *Lebesgue spaces have the following important properties:*

- *For $1 \le p \le \infty$, $\|\cdot\|_{L^p}$ is a norm and $L^p(\Omega)$ equipped with this norm is a* ***Banach space****, i.e. a complete normed vector space.*

  For $p = 2$, the space $L^2(\Omega)$ with the inner product

  $$\langle u, v\rangle = \int_\Omega uv, \qquad u, v \in L^2(\Omega),$$

  is a ***Hilbert space****, i.e. a complete inner product space.*

- *Let $1 \le p \le \infty$. If $(u_m)_{m\in\mathbb{N}} \subset L^p(\Omega)$ is a sequence that converges to u in $L^p(\Omega)$, then there exists a subsequence $(u_{m_k})_{k\in\mathbb{N}}$ that converges to u a.e. in Ω.*

- *If $1 \le p < \infty$, then the step functions as well as $C_c^\infty(\Omega)$ are dense in $L^p(\Omega)$.*

  Consequently, $L^p(\Omega)$ is ***separable****, i.e. it contains a countable dense subset.*

Finally, we recall properties of *mollifiers*. For $\varepsilon > 0$ we write $\Omega_\varepsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\}$. Let $\varphi \in C^\infty(\mathbb{R}^n)$ be defined as

$$\varphi(x) = \begin{cases} ce^{\frac{1}{|x|^2-1}}, & |x| < 1, \\ 0 & |x| \ge 1, \end{cases}$$

where the constant $c > 0$ is chosen such that $\int_{\mathbb{R}^n} \varphi = 1$. The function φ is called the **standard mollifier**.

Moreover, for $\varepsilon > 0$ we set

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi\left(\frac{x}{\varepsilon}\right), \qquad x \in \mathbb{R}^n.$$

We observe that φ_ε satisfies

$$\varphi_\varepsilon \in C_c^\infty(\mathbb{R}^n), \quad \varphi_\varepsilon \ge 0, \quad \int_{\mathbb{R}^n} \varphi_\varepsilon = 1, \quad \operatorname{supp}(\varphi_\varepsilon) \subset B_\varepsilon(0),$$

where $B_\varepsilon(0) = \{x \in \mathbb{R}^n : |x| < \varepsilon\}$ denotes the open ball of radius $\varepsilon > 0$ and center 0 in $\mathbb{R}^n$.

Definition 3.7. Let $f \in L^1_{loc}(\Omega)$. Then, its **mollification** is defined as the convolution of f with φ_ε,

$$f_\varepsilon = \varphi_\varepsilon * f \qquad \text{in } \Omega_\varepsilon,$$

i.e.

$$f_\varepsilon(x) = \int_\Omega \varphi_\varepsilon(x-y)f(y)dy = \int_{B_\varepsilon(0)} \varphi_\varepsilon(y)f(x-y)dy, \qquad x \in \Omega_\varepsilon.$$

Mollifiers allow to construct smooth approximations for general (rough) functions. Several properties (e.g., certain identities or inequalities) are much easier to prove if the functions involved are smooth. Hence, typically one first proves a certain statement for smooth functions and then uses smooth approximations to extend it for less regular functions.

Theorem 3.8. *Let $f \in L^1_{loc}(\Omega)$. Then, the mollification f_ε has the following properties:*

(i) $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$.

(ii) $f_\varepsilon \to f$ *almost everywhere as* $\varepsilon \to 0$.

(iii) If f is continuous on Ω, then $f_\varepsilon \to f$ as $\varepsilon \to 0$ uniformly on compact subsets of Ω.

(iv) If $1 \le p < \infty$ and $f \in L^p_{loc}(\Omega)$, then $f_\varepsilon \to f$ as $\varepsilon \to 0$ in $L^p_{loc}(\Omega)$.

Proof. See, e.g. Appendix C in [7]. □

3.2 Sobolev spaces

Definition and basic properties

We call

$$C_c^\infty(\Omega) = \{u \in C^\infty(\Omega) : \text{supp}(u) \subset \Omega \text{ compact}\}$$

the space of *test functions*.

If $u \in C^1(\Omega)$, then integration by parts implies that

$$\int_\Omega u \partial_{x_i}\varphi = -\int_\Omega (\partial_{x_i} u)\varphi \qquad \forall \varphi \in C_c^\infty(\Omega),$$

$i = 1, \ldots, n$. This identity allows to generalize the notion of derivatives. In fact, the integral on the left hand side is well-defined if u is locally integrable. This motivates the following definition.

Definition 3.9. Let $u, v \in L^1_{loc}(\Omega)$ and $i \in \{1, \ldots, n\}$. Then, v is the **weak partial derivative of** u **w.r.t.** x_i, if

$$\int_\Omega u \partial_{x_i}\varphi = -\int_\Omega v\varphi \qquad \forall \varphi \in C_c^\infty(\Omega).$$

In this case, we use the notation $v = \partial_{x_i} u$.

Moreover, we call u **weakly differentiable** if the weak partial derivatives $\partial_{x_1} u, \ldots, \partial_{x_n} u$ exist.

If the weak partial derivative exists, it is unique (up to sets of measure zero) by the following generalization of the Fundamental Lemma of the Calculus of Variations:

Lemma 3.10. *Let $u \in L^1_{loc}(\Omega)$ be such that*

$$\int_\Omega u\varphi = 0 \qquad \forall \varphi \in C_c^\infty(\Omega).$$

Then, $u \equiv 0$ a.e. in Ω.

Proof. We note that the integral is well-defined by Hölder's inequality.

Let φ_ε be the standard mollifier and $u_\varepsilon = \varphi_\varepsilon * u$. If $x \in \Omega$ and $\varepsilon > 0$ is small enough, then $\operatorname{supp}(\varphi_\varepsilon(x - \cdot)) \subset \Omega$ is compact. Consequently, $\varphi_\varepsilon(x - \cdot)) \in C_c^\infty(\Omega)$ and we conclude that

$$u_\varepsilon(x) = \int_\Omega \varphi_\varepsilon(x-y)u(y)dy = 0.$$

Finally, since $u_\varepsilon \to u$ a.e. in Ω as $\varepsilon \to 0$, it follows that

$$u(x) = \lim_{\varepsilon\to 0} u_\varepsilon(x) = 0 \text{ a.e.}.$$

□

Remark.

- Weak derivatives are unique (up to a set of measure zero).
- If $u \in C^1(\Omega)$, then the classical derivative and the weak derivative coincide.
- The notion of weak derivatives generalizes the classical notion of derivatives, however, note that not every function in $L^1_{loc}(\Omega)$ is weakly differentiable. An example is given below.

Example 3.11. Let $\Omega = (-1, 1) \subset \mathbb{R}$.

(a) Consider the function $u(x) = |x|$. Then, u is weakly differentiable and

$$u'(x) = v(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0. \end{cases}$$

Indeed, let $\varphi \in C_c^\infty(\Omega)$. Then, we obtain

$$\begin{aligned}\int_{-1}^1 u(x)\varphi'(x)dx &= \int_{-1}^0 (-x)\varphi'(x)dx + \int_0^1 x\varphi'(x)dx \\ &= \int_{-1}^0 \varphi(x)dx - \int_0^1 \varphi(x)dx = -\int_{-1}^1 v(x)\varphi(x)dx,\end{aligned}$$

where we used integration by parts and the fact that φ has compact support in $(-1, 1)$.

(b) Consider the function

$$u(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0. \end{cases}$$

Then, u is not weakly differentiable. Indeed, assume that the weak derivative exists, $u' = v \in L^1_{loc}(\Omega)$. Let $\varphi \in C_c^\infty((0, 1))$ and extend it by zero on $(-1, 0]$. Using that v is the weak derivative of u we obtain

$$\int_{-1}^1 v(x)\varphi(x)dx = -\int_0^1 u(x)\varphi'(x)dx = -\int_0^1 \varphi'(x)dx = 0,$$

since $\varphi(0) = \varphi(1) = 0$. Consequently,

$$\int_0^1 v(x)\varphi(x)dx = 0 \qquad \forall \varphi \in C_c^\infty((0,1)),$$

and Lemma 3.10 implies that $v = 0$ a.e. in $(0,1)$. Similarly, we conclude that $v = 0$ a.e. in $(-1,0)$.

Therefore, for any $\varphi \in C_c^\infty((-1,1))$ it now follows that

$$0 = \int_{-1}^1 v(x)\varphi(x)dx = -\int_{-1}^1 u(x)\varphi'(x)dx = -\int_0^1 \varphi'(x)dx = \varphi(0),$$

i.e. $\varphi(0) = 0$ for all $\varphi \in C_c^\infty((-1,1))$, which is a contradiction.

Similarly, one can define weak derivatives of higher order.

Definition 3.12. Let $u, v \in L^1_{loc}(\Omega)$ and $\alpha \in \mathbb{N}_0^n$. Then, v is the αth **weak partial derivative of** u, if

$$\int_\Omega u\partial^\alpha\varphi = (-1)^{|\alpha|}\int_\Omega v\varphi \qquad \forall \varphi \in C_c^\infty(\Omega).$$

In this case, we write $v = \partial^\alpha u$.

Using these generalized notions of derivatives we now introduce spaces of weakly differentiable functions.

Definition 3.13. Let $1 \leq p \leq \infty$.

- The **Sobolev space** $W^{1,p}(\Omega)$ is defined as

$$\begin{aligned} W^{1,p}(\Omega) = \{u \in L^p(\Omega) : &\text{ the weak partial derivative } \partial_{x_i}u \\ &\text{exists and } \partial_{x_i}u \in L^p(\Omega), i = 1, \dots, n\}, \end{aligned}$$

  endowed with the norm

$$\begin{aligned} \|u\|_{W^{1,p}} &= \left(\|u\|_{L^p}^p + \sum_{i=1}^n \|\partial_{x_i}u\|_{L^p}^p\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty, \\ \|u\|_{W^{1,\infty}} &= \|u\|_{L^\infty} + \sum_{i=1}^n \|\partial_{x_i}u\|_{L^\infty}. \end{aligned}$$

- The **Sobolev space** $W^{k,p}(\Omega)$, $k \in \mathbb{N}$, is defined as

$$\begin{aligned} W^{k,p}(\Omega) = \{u \in L^p(\Omega) : &\text{ the weak partial derivative } \partial^\alpha u \\ &\text{exists and } \partial^\alpha u \in L^p(\Omega), \text{ for all multiindices } |\alpha| \leq k\}, \end{aligned}$$

  endowed with the norm

$$\begin{aligned} \|u\|_{W^{k,p}} &= \left(\sum_{0\leq|\alpha|\leq k} \|\partial^\alpha u\|_{L^p}^p\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty, \\ \|u\|_{W^{k,\infty}} &= \sum_{0\leq|\alpha|\leq k} \|\partial^\alpha u\|_{L^\infty}. \end{aligned}$$

- We say that a sequence $(u_m)_{m\in\mathbb{N}} \subset W^{1,p}(\Omega)$ **converges to** u **in** $W^{1,p}(\Omega)$, and use the notation $u_m \to u$ in $W^{1,p}(\Omega)$, if $u \in W^{1,p}(\Omega)$ and

$$\lim_{m\to\infty} \|u_m - u\|_{W^{1,p}} = 0.$$

  Moreover, we say that $u_m \to u$ in $W^{1,p}_{loc}(\Omega)$ if $u_m \to u$ in $W^{1,p}(V)$ for every open subset V such that $V \subset\subset \Omega$.

Remark 3.14.
- For $p = 2$ it is customary to write $W^{1,2}(\Omega) = H^1(\Omega)$. In this case, $H^1(\Omega)$ is a Hilbert space with inner product

$$\langle u, v\rangle_{H^1} = \int_\Omega (uv + \nabla u \cdot \nabla v), \qquad u, v \in H^1(\Omega).$$

- For $n = 1$ and $\Omega \subset \mathbb{R}$ an open interval, one can show that $u \in W^{1,p}(\Omega)$ if and only is $u = \tilde{u}$ a.e. in Ω, where $\tilde{u}$ is absolutely continuous and its derivative $\tilde{u}' \in L^p(\Omega)$ (which exists a.e.).

- If Ω is bounded, we have the following inclusions

$$C^1(\overline{\Omega}) \subsetneq W^{1,\infty}(\Omega) \subsetneq W^{1,p}(\Omega) \subsetneq L^p(\Omega) \qquad \forall 1 \le p < \infty.$$

Example 3.15. Let $\Omega = B_1(0) = \{x \in \mathbb{R}^n : |x| < 1\}$. We aim to determine for which values of $s > 0, n$ and p the function

$$u(x) = |x|^{-s}$$

belongs to $W^{1,p}(\Omega)$.

First, we observe that if the weak derivative $\partial_{x_i} u$ exists, it must be given by

$$\partial_{x_i} u(x) = -s\frac{x_i}{|x|^{s+2}} = v_i(x), \qquad i = 1, \dots n.$$

Indeed, u is smooth in $B_1(0) \setminus \{0\}$ and the claim follows by considering test functions that vanish in a neighborhood of the origin.

Let $\varphi \in C_c^\infty(\Omega)$ be an arbitrary test function and $0 < \varepsilon < 1$. Then,

$$\begin{aligned}\int_\Omega u\partial_{x_i}\varphi &= \int_{B_\varepsilon(0)} u\partial_{x_i}\varphi + \int_{\Omega\setminus B_\varepsilon(0)} u\partial_{x_i}\varphi \\ &= \int_{B_\varepsilon(0)} u\partial_{x_i}\varphi + \int_{\partial B_\varepsilon(0)} u\varphi\nu_i dS - \int_{\Omega\setminus B_\varepsilon(0)} (\partial_{x_i} u)\varphi =: I_1 + I_2 - I_3,\end{aligned}$$

where ν denotes the inward pointing unit normal vector on $\partial B_\varepsilon(0)$.

We note that $u \in L^1(\Omega)$ if and only if $s < n$ and $|\nabla u| \in L^1(\Omega)$ if and only if $s+1 < n$. Moreover, if $s + 1 < n$, then by dominated convergence it follows that

$$I_1 \to 0 \qquad \text{and} \qquad I_3 \to \int_\Omega (\partial_{x_i} u)\varphi \qquad \text{as } \varepsilon \to 0.$$

Since $u(x) = \varepsilon^{-s}$ for $x \in \partial B_\varepsilon(0)$, we obtain

$$|I_2| \le \|\varphi\|_{L^\infty} \int_{\partial B_\varepsilon(0)} \varepsilon^{-s} dS(x) \le c\varepsilon^{-s+n-1} \to 0 \quad \text{as } \varepsilon \to 0,$$

for some constant $c \geq 0$. Consequently, taking the limit $\varepsilon \to 0$ in the equation above, we conclude that

$$\int_\Omega u \partial_{x_i} \varphi = -\int_\Omega (\partial_{x_i} u)\varphi,$$

which shows that v_i is the weak partial derivative $\partial_{x_i} u$ of u if $s + 1 < n$.

Finally, $|\nabla u(x)| = \frac{s}{|x|^{s+1}}$ lies in $L^p(\Omega)$ if and only if $p(s+1) < n$, which implies that $u \in W^{1,p}(\Omega)$ if and only if $s < \frac{n-p}{p}$.

Using the Minkowski inequality and completeness of $L^p(\Omega)$ one can show that Sobolev spaces are Banach spaces.

Theorem 3.16. *Let $1 \leq p \leq \infty$ and $k \in \mathbb{N}$. Then, $W^{k,p}(\Omega)$ is a Banach space.*

Proof. Let $u, v \in W^{1,p}(\Omega)$. Then, the Minkowski inequality implies that

$$\begin{aligned}
\|u+v\|_{W^{1,p}} &= \left(\|u+v\|_{L^p}^p + \sum_{i=1}^n \|\partial_{x_i} u + \partial_{x_i} v\|_{L^p}^p \right)^{\frac{1}{p}} \\
&\leq \left((\|u\|_{L^p} + \|v\|_{L^p})^p + \sum_{i=1}^n (\|\partial_{x_i} u\|_{L^p} + \|\partial_{x_i} v\|_{L^p})^p \right)^{\frac{1}{p}} \\
&\leq \left(\|u\|_{L^p}^p + \sum_{i=1}^n \|\partial_{x_i} u\|_{L^p}^p \right)^{\frac{1}{p}} + \left(\|v\|_{L^p}^p + \sum_{i=1}^n \|\partial_{x_i} v\|_{L^p}^p \right)^{\frac{1}{p}} \\
&= \|u\|_{W^{1,p}} + \|v\|_{W^{1,p}},
\end{aligned}$$

where we used the triangle inequality for the p-norm in $\mathbb{R}^{n+1}$ in the last inequality. The other properties of a norm and that $W^{1,p}(\Omega)$ is a linear space are clear.

To prove completeness, let $(u_m)_{m\in\mathbb{N}}$ be a Cauchy sequence in $W^{1,p}(\Omega)$. Then, $(u_m)_{m\in\mathbb{N}}$ and $(\partial_{x_i} u_m)_{m\in\mathbb{N}}, i = 1, \ldots n$, are Cauchy sequences in $L^p(\Omega)$. Since $L^p(\Omega)$ is complete, there exist $v_0, v_1, \ldots, v_n \in L^p(\Omega)$ such that

$$u_m \to v_0, \quad \partial_{x_i} u_m \to v_i \quad \text{in } L^p(\Omega), \qquad i = 1, \ldots, n.$$

It remains to show that $u_m \to v_0$ in $W^{1,p}(\Omega)$, i.e. we need to show that $\partial_{x_i} v_0 = v_i, i = 1, \ldots, n$. To this end let $\varphi \in C_c^\infty(\Omega)$. Then, we obtain

$$\int_\Omega v_0 \partial_{x_i} \varphi = \lim_{m\to\infty} \int_\Omega u_m \partial_{x_i} \varphi = -\lim_{m\to\infty} \int_\Omega (\partial_{x_i} u_m)\varphi = -\int_\Omega v_i \varphi,$$

where we used that $\partial_{x_i} u_m$ is the weak partial derivative of v_m, $i = 1, \ldots, n$. This shows that the weak partial derivatives $\partial_{x_i} v_0$ of v_0 exist and $\partial_{x_i} v_0 = v_i \in L^p(\Omega)$.

We remark that the limits in the integral equation above can be justified by Hölder's inequality. Indeed, if $f_m \to f$ in $L^p(\Omega)$ and $g \in L^q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\left| \int_\Omega f_m g - \int_\Omega f g \right| \leq \int_\Omega |(f_m - f)g| \leq \|f_m - f\|_{L^p} \|g\|_{L^q} \to 0 \quad \text{as } m \to \infty.$$

□

Approximations, extensions and traces

Theorem 3.17 (Global approximation)**.** *Let $1 \le p < \infty$ and $u \in W^{1,p}(\Omega)$. Then, there exists a sequence $u_m \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$ such that*

$$u_m \xrightarrow[m\to\infty]{} u \quad \text{in } W^{1,p}(\Omega).$$

Note that the approximation theorem is not valid for $p = \infty$. If the boundary $\partial\Omega$ is sufficiently regular, functions in Sobolev spaces can even be approximated by functions in $C^\infty(\overline{\Omega})$, and not only in $C^\infty(\Omega)$. In fact, the following statement holds:

Let Ω be open and bounded with $\partial\Omega$ of class C^1 and $1 \le p < \infty$. If $u \in W^{1,p}$, then there exists a sequence $u_m \in C^\infty(\overline{\Omega})$ such that

$$u_m \xrightarrow[m\to\infty]{} u \quad \text{in } W^{1,p}(\Omega).$$

Remark 3.18. Theorem 3.17 was proven by N. G. Meyers and J. Serrin in 1964. Before that, the statement of Theorem 3.17 was believed to be wrong without smoothness assumption on $\partial\Omega$. At that time it was customary to use the notation

$$H^{k,p}(\Omega) = \overline{C^\infty(\Omega) \cap W^{k,p}(\Omega)}^{\|\cdot\|_{W^{k,p}}}.$$

Hence, Theorem 3.17 shows that $W^{k,p}(\Omega) = H^{k,p}(\Omega)$.

We aim to extend functions in $W^{1,p}(\Omega)$, $\Omega \subset \mathbb{R}^n$ open, to functions $u \in W^{1,p}(\mathbb{R}^n)$. This can be subtle, since extending a function $u \in W^{1,p}(\Omega)$ by zero in $\mathbb{R}^n \setminus \Omega$ may lead to a function that no longer has a weak derivative (see Example 3.11). The following theorem provides sufficient conditions for the existence of an extension operator.

Theorem 3.19. *Let $1 \le p \le \infty$ and $\Omega \subset \mathbb{R}^n$ be bounded with Lipschitz boundary $\partial\Omega$. Let $V \subset \mathbb{R}^n$ be open and bounded such that $\Omega \subset\subset V$. Then, there exists a bounded linear operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ such that the following properties hold for all $u \in W^{1,p}(\Omega)$:*

- *$Eu = u$ a.e. in Ω,*
- *supp$(Eu) \subset V$,*
- *$\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \le c\|u\|_{W^{1,p}(\Omega)}$, for some constant $c > 0$ depending on p, Ω and V.*

In this case, E is called the ***extension operator*** *and Eu the extension of u.*

We recall that a bounded set $\Omega \subset \mathbb{R}^n$ has a *Lipschitz boundary* if for every $x \in \partial\Omega$ there exists a ball $B_\rho(x) = \{y \in \mathbb{R}^n : \|x - y\| < \rho\} \subset \mathbb{R}^n$, a neighborhood $U \subset \mathbb{R}^n$ of the origin and a bijective mapping $\psi : U \to B_\rho(x)$ such that $\psi : \overline{U} \to \overline{B_\rho(x)}$ and $\psi^{-1} : \overline{U} \to \overline{B_\rho(x)}$ are Lipschitz continuous and $\psi(U_+) = B_\rho(x) \cap \Omega$, $\psi(U_0) = B_\rho(x) \cap \partial\Omega$, where

$$U_+ = \{x \in U : x_n > 0\}, \qquad U_0 = \{x \in U : x_n = 0\}.$$

Next, we aim to assign boundary values along $\partial\Omega$ to a function $u \in W^{1,p}(\Omega)$, which can be done via traces. Note that, since $\partial\Omega$ is a set of measure zero, there is no direct meaning to the restriction of a function $u \in W^{1,p}(\Omega)$ to $\partial\Omega$, as u is only defined a.e. in Ω.

Theorem 3.20. *Let $\Omega \subset \mathbb{R}^n$ be bounded with Lipschitz boundary $\partial\Omega$ and $1 \leq p \leq \infty$. Then, there exists a bounded linear operator $T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ such that*

- $Tu = u|_{\partial\Omega}$, *if* $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$
- $\|Tu\|_{L^p(\partial\Omega)} \leq c\|u\|_{W^{1,p}(\Omega)} \quad \forall u \in W^{1,p}(\Omega)$, *for some constant* $c > 0$ *only depending on* p *and* Ω.

In this case, T is called the ***trace operator*** *and Tu the trace of u on $\partial\Omega$.*

Definition 3.21. We define $W_0^{1,p}(\Omega)$ as the closure of $C_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$,

$$W_0^{1,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{W^{1,p}}}$$

Corollary 3.22. *Let $\Omega \subset \mathbb{R}^n$ be bounded with Lipschitz boundary $\partial\Omega$ and $1 \leq p \leq \infty$. Then, the following holds:*

$$u \in W_0^{1,p}(\Omega) \qquad \Longleftrightarrow \qquad Tu = 0.$$

Sobolev embeddings and Poincaré inequality

Next, we investigate whether functions in Sobolev spaces $W^{1,p}(\Omega)$ lie in "nicer" functions spaces. This turns out to be the case, but it depends on whether $p < n$, $p > n$ or $p = n$.

Theorem 3.23 (Sobolev embeddings)**.** *Let Ω be be bounded with Lipschitz boundary $\partial\Omega$.*

(i) If $1 \leq p < n$ then $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all $1 \leq q \leq p^*$, where $p^* = \frac{np}{n-p} \in (p, \infty)$, i.e. $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ and there exists $c > 0$ depending on Ω and p such that

$$\|u\|_{L^{p^*}(\Omega)} \leq c\|u\|_{W^{1,p}(\Omega)} \qquad \forall u \in W^{1,p}(\Omega).$$

(ii) If $p = n$ then $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for any $1 \leq q < \infty$, i.e. $W^{1,p}(\Omega) \subset L^q(\Omega)$ and there exists $c > 0$ depending on Ω, p and q such that

$$\|u\|_{L^q(\Omega)} \leq c\|u\|_{W^{1,p}(\Omega)} \qquad \forall u \in W^{1,p}(\Omega).$$

(iii) If $p > n$ then $W^{1,p}(\Omega) \hookrightarrow C(\overline{\Omega})$, i.e. $W^{1,p}(\Omega) \subset C(\overline{\Omega})$ and there exists $c > 0$ depending on Ω and p such that

$$\|u\|_{C(\overline{\Omega})} \leq c\|u\|_{W^{1,p}(\Omega)} \qquad \forall u \in W^{1,p}(\Omega).$$

Remark 3.24. Taking the limit $p \to n$ in Theorem 3.23 (i) in the case $p < n$, we would expect that $W^{1,n}(\Omega) \hookrightarrow L^\infty(\Omega)$, but this is not true for $n \geq 2$. A counterexample is given by the function

$$u(x) = \ln\left(\ln\left(1 + \frac{1}{|x|}\right)\right), \qquad x \in B_1(0) \subset \mathbb{R}^n,$$

(see exercises). On the other hand, by Theorem 3.23, we conclude

$$W^{1,n}(\Omega) \subset \bigcap_{1 \leq p < \infty} L^p(\Omega).$$

For bounded subsets $\Omega \subset \mathbb{R}^n$ one can even show that embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is *compact*, i.e. the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is continuous and every bounded sequence in $W^{1,p}(\Omega)$ has a subsequence that converges in $L^p(\Omega)$.

If a space V is compactly embedded into a space W we use the notation $V \hookrightarrow\hookrightarrow W$.

Theorem 3.25. *Let $\Omega \subset \mathbb{R}^n$ be be bounded with Lipschitz boundary $\partial\Omega$. Then, the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact for all $1 \leq p \leq \infty$.*

Next, we prove the Poincaré inequality, which allows to bound the L^p-norm of functions in $W_0^{1,p}(\Omega)$ in terms of the L^p-norm of their gradient.

Theorem 3.26 (Poincaré inequality)**.** *Let $1 \leq p \leq \infty$ and $\Omega \subset \mathbb{R}^n$ be open and bounded with respect to one coordinate direction. Without loss of generality we assume that*

$$\Omega \subset \{x \in \mathbb{R}^n : 0 \leq x_n \leq d\}, \qquad \textit{for some } d > 0.$$

Then, there exists a constant $c > 0$ such that

$$\|u\|_{L^p(\Omega)} \leq c \||\nabla u|\|_{L^p(\Omega)} \qquad \forall u \in W_0^{1,p}(\Omega).$$

Proof. Let us first assume that $u \in C_c^\infty(\Omega)$. Then, using that $u(0) = 0$ it follows that

$$u(x) = \int_0^{x_n} \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) dt.$$

Consequently, Hölder's inequality implies that

$$\begin{aligned}
|u(x)|^p &\leq \left(\int_0^{x_n} \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right| dt \right)^p \\
&\leq \left(\int_0^{x_n} 1 dt \right)^{\frac{p}{q}} \int_0^{x_n} \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right|^p dt \\
&= x_n^{p-1} \int_0^{x_n} \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right|^p dt.
\end{aligned}$$

Integrating over Ω we obtain

$$\begin{aligned}
\|u\|_{L^p(\Omega)}^p &\leq \int_\Omega x_n^{p-1} \int_0^{x_n} \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right|^p dt dx \\
&\leq \int_{\mathbb{R}^{n-1}} \int_0^d x_n^{p-1} \int_0^d \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right|^p dt dx_n dx_1 \cdots dx_{n-1} \\
&\leq \frac{d^p}{p} \int_{\mathbb{R}^{n-1}} \int_0^d \left| \frac{d}{dt} u(x_1, \ldots, x_{n-1}, t) \right|^p dt dx_1 \cdots dx_{n-1} \leq \frac{d^p}{p} \||\nabla u|\|_{L^p(\Omega)},
\end{aligned}$$

where we used that Ω is bounded with respect to x_n, Fubini's Theorem and that u has compact support in Ω.

Finally, if $u \in W_0^{1,p}(\Omega)$, then there exists a sequence $(u_m)_{m \in \mathbb{N}}$ in $C_c^\infty(\Omega)$ such that $u_m \xrightarrow[m \to \infty]{} u$ in $W^{1,p}(\Omega)$. For $u_m \in C_c^\infty(\Omega)$, $m \in \mathbb{N}$, the Poincaré inequality holds. Since $\|u_m\|_{L^p(\Omega)} \xrightarrow[m \to \infty]{} \|u\|_{L^p(\Omega)}$ and $\||\nabla u_m|\|_{L^p(\Omega)} \xrightarrow[m \to \infty]{} \||\nabla u|\|_{L^p(\Omega)}$, we can pass to the limit $m \to \infty$ in the inequality for u_m and conclude that the Poincaré inequality remains valid for functions in $W_0^{1,p}(\Omega)$. $\square$

Theorem 3.27 (Poincaré inequality; mean-value version)**.** *Let $1 \le p \le \infty$ and $\Omega \subset \mathbb{R}^n$ be bounded and connected with Lipschitz boundary $\partial\Omega$. Then, there exists a constant $c > 0$ depending on Ω and p such that*

$$\|u - (u)_\Omega\|_{L^p(\Omega)} \le c\||\nabla u|\|_{L^p(\Omega)} \qquad \forall u \in W^{1,p}(\Omega),$$

where we use the notation $(u)_\Omega = \frac{1}{|\Omega|}\int_\Omega u$.

Proof. By contradiction we assume that there exists a sequence $(u_m)_{m\in\mathbb{N}} \subset W^{1,p}(\Omega)$ such that

$$\|u_m - (u_m)_\Omega\|_{L^p} > m\||\nabla u_m|\|_{L^p}.$$

We define

$$v_m = \frac{u_m - (u_m)_\Omega}{\|u_m - (u_m)_\Omega\|_{L^p}}, \qquad m \in \mathbb{N},$$

and observe that $\|v_m\|_{L^p} = 1, (v_m)_\Omega = 0$ and $\||\nabla v_m|\|_{L^p} < \frac{1}{m.}$ Consequently, $(v_m)_{m\in\mathbb{N}}$ is a bounded sequence in $W^{1,p}(\Omega)$ and Theorem 3.25 implies that there exists a subsequence $(v_{m_k})_{k\in\mathbb{N}}$ such that $v_{m_k} \to v$ in $L^p(\Omega)$. We conclude that $\|v\|_{L^p} = 1$ and $(v)_\Omega = 0$.

Moreover, we also have $\nabla v = 0$ a.e. in Ω. Indeed, for all $\varphi \in C_c^\infty(\Omega)$ we observe that

$$\begin{aligned}\left|\int_\Omega v\partial_{x_i}\varphi\right| &= \lim_{k\to\infty}\left|\int_\Omega v_{m_k}\partial_{x_i}\varphi\right| = \lim_{k\to\infty}\left|\int_\Omega (\partial_{x_i}v_{m_k})\varphi\right| \\ &\le \lim_{k\to\infty}\|\partial_{x_i}v_{m_k}\|_{L^p}\|\varphi\|_{L^q} = 0,\end{aligned}$$

where we used Hölder'e inequality. This implies that $v \in W^{1,p}(\Omega)$ and $\nabla v = 0$ a.e. in Ω. Since Ω is connected, Lemma 3.28 implies that $v = d$ for some constant $d \in \mathbb{R}$. We conclude that $(v)_\Omega = 0$ and therefore, $v = 0$ a.e. in Ω. This is a contradiction to the fact that $\|v\|_{L^p} = 1$. □

Lemma 3.28. *Let Ω be connected and $u \in W^{1,p}(\Omega)$ such that $\nabla u = 0$ a.e. in Ω. Then, $u = c$ a.e. in Ω for some constant $c \in \mathbb{R}$.*

Proof. The proof is left as an exercise. □

3.3 Weak convergence and weak compactness

In this section we recall results about weak convergence and weak compactness that we need to prove the existence of minimizers. In fact, when we apply the *direct method* we consider minimizing sequences. We then try to extract a subsequence that converges and show that the limit of this subsequence is the minimizer of the functional.

Let X be a Banach space with norm $\|\cdot\|_X$ and X^* be its dual space, i.e. the space of all bounded linear functionals $l : X \to \mathbb{R}$. The dual space endowed with the norm

$$\|l\|_{X^*} = \sup_{\|u\|_X\le 1} |l(u)|, \qquad l \in X^*,$$

is a Banach space.

We say that a sequence $(u_m)_{m\in\mathbb{N}}$ in a Banach space converges (or *converges strongly*) if there exists $u \in X$ and $\lim_{m\to\infty}\|u_m - u\|_X = 0$.

Definition 3.29. We say that a sequence $(u_m)_{m\in\mathbb{N}} \subset X$ **converges weakly** to u in X if

$$l(u_m) \xrightarrow[m\to\infty]{} l(u) \qquad \forall l \in X^*.$$

We use the notation $u_m \rightharpoonup u$ in X.

We say that a sequence $(l_m)_{m\in\mathbb{N}} \subset X^*$ **converges weakly*** to l in X^* if

$$l_m(u) \xrightarrow[m\to\infty]{} l(u) \qquad \forall u \in X.$$

We use the notation $u_m \rightharpoonup^* u$ in X^*.

Example 3.30.
- Let $1 \leq p < \infty$. The dual space of $L^p(\Omega)$ can be identified with $L^q(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. We have the dual pairing

  $$\langle u, v\rangle_{L^p,L^q} = \int_\Omega uv, \qquad u \in L^p(\Omega), v \in L^q(\Omega),$$

  which is well-defined due to Hölder's inequality.

  This implies that $u_m \rightharpoonup u$ in $L^p(\Omega)$ if and only if

  $$\int_\Omega u_m v \xrightarrow[m\to\infty]{} \int_\Omega uv \qquad \forall v \in L^q(\Omega).$$

- If $1 \leq p < \infty$ and $l \in \left(W^{1,p}(\Omega)\right)^*$, then there exist $v_0, v_1, \ldots, v_n \in L^q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, such that

  $$l(u) = \int_\Omega \left(v_0 u + \sum_{i=1}^{m} v_i \partial_{x_i} u \right) \qquad \forall u \in W^{1,p}(\Omega).$$

  Moreover, $u_m \rightharpoonup u$ in $W^{1,p}(\Omega)$ if and only if $u_m \rightharpoonup u$ in $L^p(\Omega)$ and $\partial_{x_i} u_m \rightharpoonup \partial_{x_i} u$ in $L^p(\Omega)$ for $i = 1, \ldots, n$.

Remark. Let X be a Banach space and X^* be its dual.

- Every strongly convergent sequence in X converges weakly, but not every weakly convergent sequence converges strongly.
- If $\dim X < \infty$ then weak convergence is equivalent to strong convergence.

Proposition 3.31 (Properties of weakly and weakly* convergent sequences)**.** Let X be a Banach space and X^* be its dual.

- Every sequence in X that weakly converges is bounded in X.

  Every sequence in X^* that weakly* converges is bounded in X^*.

- Let $(u_m)_{m\in\mathbb{N}}$ be a sequence in X and $(l_m)_{m\in\mathbb{N}}$ be a sequence in X^*.

  If $u_m \rightharpoonup u$ in X and $l_m \to l$ in X^*, then $l_m(u_m) \to l(u)$.

  If $u_m \to u$ in X and $l_m \rightharpoonup^* l$ in X^*, then $l_m(u_m) \to l(u)$.

Corollary 3.32. *Let $1 \le p \le \infty$, $\Omega \subset \mathbb{R}^n$ is bounded with Lipschitz boundary $\partial\Omega$ and $(u_m)_{m\in\mathbb{N}}$ be a sequence in $L^p(\Omega)$ such that $u_m \rightharpoonup u$ in $W^{1,p}(\Omega)$. Then, there exists a subsequence (u_{m_k}) such that $u_{m_k} \to u$ in $L^p(\Omega)$.*

Proof. By Proposition3.31, the sequence $(u_m)_{m\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega)$. Therefore, the statement follows from the compact embedding $W^{1,p}(\Omega) \hookrightarrow\hookrightarrow L^p(\Omega)$, see Theorem 3.25. □

A consequence of the Hahn-Banach theorem is that convex subsets are weakly closed.

Theorem 3.33. *Let X be a Banach space and $V \subset X$ be closed and convex. If $(v_m)_{m\in\mathbb{N}} \subset V$ is a sequence and $v \in X$ such that $v_m \rightharpoonup v$ in X, then $u \in V$.*

Finally, we recall an important property of *reflexive* Banach spaces, i.e. $(X^*)^* \simeq X$, which is a consequence of the Banach-Alaoglu Theorem.

Proposition 3.34. Let X be a reflexive Banach space. Then, every bounded sequence has a weakly convergent subsequence.

Example 3.35. If $1 < p < \infty$, then the spaces $L^p(\Omega)$ and $W^{1,p}(\Omega)$ are reflexive. However, this is not the case for $L^1(\Omega)$, $L^\infty(\Omega)$, $W^{1,1}(\Omega)$ and $W^{1,\infty}(\Omega)$.

Consequently, if $(u_m)_{m\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega)$ or in $W^{1,p}(\Omega)$ and $1 < p < \infty$, then it possesses a subsequence that converges weakly.

Chapter 4

Euler-Lagrange Equations in higher dimensions

In this chapter, we consider variational problems, where the integrals are not defined on an interval as in Chapter 2, but on higher dimensional domains $\Omega \subset \mathbb{R}^n, n \geq 2$. We first address the classical theory and show how the theory of elliptic partial differential equations (PDEs) relates to the field Calculus of Variations. Then, we extend the class of admissible functions and introduce the setting for variational problems in Sobolev spaces. The latter setting is needed to apply direct methods which allow to prove the existence of minimizers.

4.1 Classical theory

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with C^1-boundary $\partial\Omega$. We consider the functional

$$I(u) = \int_\Omega f\,(x, u(x), \nabla u(x))\,dx, \tag{4.1}$$

where $f : \overline{\Omega}\times\mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z,)$ is assumed to be twice continuously differentiable. As in Chapter 2, we first aim to find a minimizer of I within the class

$$\Phi = \left\{u \in C^1(\overline{\Omega}) : u = g \text{ on } \partial\Omega\right\},$$

where $g : \partial\Omega \to \mathbb{R}$ is a given C^1-function.

Assume that $u \in \Phi$ is a local minimizer of I. If $\varphi \in C_c^\infty(\Omega)$ and $s \in (-s_0, s_0)$, $s_0 > 0$, then $u + s\varphi \in \Phi$ and we conclude that

$$\begin{aligned} 0 &= \frac{d}{ds}\,(I(u + s\varphi))\Big|_{s=0} = \frac{d}{ds}\left(\int_\Omega f(\cdot, u + s\varphi, \nabla(u + s\varphi))\right)\Big|_{s=0} \\ &= \int_\Omega \left(f_y(\cdot, u, \nabla u))\varphi + \sum_{i=1}^n \partial_{z_i} f(\cdot, u, \nabla u)\partial_{x_i}\varphi\right) =: \delta I(u, \varphi). \end{aligned} \tag{4.2}$$

Note that we can interchange differentiation and integration as the integrand is sufficiently regular and the set $\overline{\Omega}$ is compact. As in Chapter 2 in case of one-dimensional problems, we call $\delta I(u, \varphi)$ the **first variation of I at u in the direction of φ**.

We conclude that weak extremals satisfy the Euler Lagrange equation, which is stated in the following theorem.

Theorem 4.1. *Let* $u \in C^2(\overline{\Omega}) \cap \Phi$ *satisfy*

$$\delta I(u, \varphi) = 0 \qquad \forall \varphi \in C_c^\infty(\Omega).$$

Then, u satisfies the ***Euler-Lagrange equation***

$$-\operatorname{div}\left(\nabla_z f(x, u(x), \nabla u(x))\right) + \partial_y f(x, u(x), \nabla u(x)) = 0, \qquad x \in \Omega.$$

Proof. Applying integration by parts in the second term of $\delta I(u, \varphi)$, see (4.2), we conclude that

$$\int_\Omega \Big(f_y(\cdot, u, \nabla u)) - \operatorname{div}\left(\nabla_z f(\cdot, u, \nabla u)\right)\Big)\varphi = 0 \qquad \text{for all } \varphi \in C_c^\infty(\Omega).$$

The Euler-Lagrange equations now follow from the Fundamental Lemma of the Calculus of Variations (Lemma 3.10). □

Remark. For one-dimensional variational problems the Euler-Lagrange equations are ODEs, while higher dimensional variational problems ($n > 1$) lead to PDEs. In general, the Euler-Lagrange equations are *quasilinear elliptic PDEs of second order*.

To simplify notations we only consider scalar functions $u : \Omega \to \mathbb{R}$, but the results in this chapter remain valid for vector-valued functions $u : \Omega \to \mathbb{R}^m$, $m \in \mathbb{N}$. In this case, the Euler-Lagrange equations are *systems* of quasilinear elliptic PDEs.

Example 4.2.

- The probably most important example is the *Dirichlet integral* which corresponds to the function $f(x, y, z) = \frac{1}{2}|z|^2$. It leads to the functional

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u|^2$$

  and the corresponding Euler-Lagrange equation is the **Laplace equation**

$$\Delta u = 0 \qquad \text{in } \Omega.$$

  It is the a linear elliptic PDE of second order.

- Consider the function $f(x, y, z) = \frac{1}{2}|z|^2 + H(y)$, where H is continuously differentiable with $H'(y) = h(y)$. This leads to the functional

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + H(u)\right),$$

  and the corresponding Euler-Lagrange equation is the semilinear **Poisson equation**

$$-\Delta u + h(u) = 0 \qquad \text{in } \Omega.$$

- Finally, we consider the function $f(x, y, z) = \sqrt{1 + |z|^2}$. Then, the functional

  $$I(u) = \int_\Omega \sqrt{1 + |\nabla u|^2}$$

  describes the surface area of the graph of u, and the corresponding Euler-Lagrange equation is the quasilinear **minimal surface equation**

  $$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0.$$

As we observed for one-dimensional minimization problem, the class of continuously differentiable functions is too restrictive. We now generalize the setting to allow for larger classes of admissible functions Φ. In particular, we consider less regular minimizers and introduce the variational setting in Sobolev Spaces. This is essential for proving the existence of minimizers based on *direct methods*.

4.2 Variational integrals on Sobolev spaces

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary $\partial\Omega$. We consider the functional

$$I(u) = \int_\Omega f\left(x, u(x), \nabla u(x)\right) dx, \tag{4.3}$$

where the function $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z)$, is of class C^1. Moreover, we assume that

$$|f(x, y, z)| \leq c\left(1 + |y|^p + |z|^p\right) \qquad \forall (x, y, z) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n, \tag{4.4}$$

for some $1 < p < \infty$ and some constant $c > 0$. This growth condition implies that I is well-defined for $u \in W^{1,p}(\Omega)$, in particular, $|I(u)| < \infty$.

We aim to find minimizers within the class

$$\begin{aligned} \Phi &= \left\{ u \in W^{1,p}(\Omega) : u = g \text{ on } \partial\Omega \text{ in the sense of traces} \right\} \\ &= \left\{ u \in W^{1,p}(\Omega) : u - g \in W_0^{1,p}(\Omega) \right\}, \end{aligned}$$

where $g \in W^{1,p}(\Omega)$ is given. Note that the second equality holds by Corollary 3.22.

Remark. The condition $u - g \in W_0^{1,p}(\Omega)$ allows us to formulate the boundary conditions $u = g$ on $\partial\Omega$ in a weak sense. Indeed, by Theorem 3.20 the trace operator T is a bounded linear operator from $W^{1,p}(\Omega)$ to $L^p(\partial\Omega)$.

For the first variation of $\delta I(u, \varphi)$ to be well-defined for $u \in W^{1,p}(\Omega)$ we need stronger assumptions than (4.4). In particular, we need the following growth conditions for the partial derivatives of f,

$$\begin{aligned} \left|\partial_y f(x, y, z)\right| &\leq c\left(1 + |y|^{p-1} + |z|^{p-1}\right), \\ |\partial_z f(x, y, z)| &\leq c\left(1 + |y|^{p-1} + |z|^{p-1}\right) \end{aligned} \qquad \forall (x, y, z) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n, \tag{4.5}$$

for some constant $c > 0$.

We show that (4.5) implies that we can interchange differentiation and integration and derive the first variation of I at u in the direction of φ for functions $u \in \Phi$ and $\varphi \in W_0^{1,p}(\Omega)$. Let $s \in (-s_0, s_0)$, $s_0 > 0$, then we observe that

$$\begin{aligned}\frac{d}{ds}\left(f(\cdot, u + s\varphi, \nabla u + s\nabla\varphi)\right) &\leq \left|\partial_y f(\cdot, u + s\varphi, \nabla u + s\nabla\varphi)\varphi + \nabla_z f(\cdot, u + s\varphi, \nabla u + s\nabla\varphi) \cdot \nabla\varphi\right| \\ &\leq c_1 \left(1 + |u + s\varphi|^{p-1} + |\nabla u + s\nabla\varphi|^{p-1}\right)(|\varphi| + |\nabla\varphi|) \\ &\leq c_2 \left(1 + |u + s\varphi|^p + |\nabla u + s\nabla\varphi|^p + |\varphi|^p + |\nabla\varphi|^p\right) \\ &\leq c_3 \left(1 + |u|^p + |\nabla u|^p + |\varphi|^p + |\nabla\varphi|^p\right),\end{aligned} \tag{4.6}$$

for some constants $c_1, c_2, c_3 > 0$. Here, we used the growth restriction (4.5) in the second step and **Young's inequality**, i.e.

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad \frac{1}{p} + \frac{1}{q} = 1 \qquad \forall a, b \in [0, \infty),$$

in the second step. Moreover, in the last estimate we used that the function $y \mapsto y^p, p > 1$, is convex on $[0, \infty)$ and consequently, $|a + b|^p \leq c(|a|^p + |b|^p)$ for some $c > 0$ and all $a, b \in [0, \infty)$.

The upper bound in (4.6) is uniform with respect to $s \in (-s_0, s_0)$ and belongs to $L^1(\Omega)$, since $u, \varphi \in W^{1,p}(\Omega)$. Therefore, by dominated convergence we can interchange differentiation and integration and conclude that

$$\begin{aligned}\delta I(u, \varphi) &= \frac{d}{ds}\left(I(u + s\varphi)\right)\Big|_{s=0} \\ &= \int_\Omega \left(\partial_y f(\cdot, u, \nabla u)\varphi + \sum_{i=1}^n \partial_{z_i} f(\cdot, u, \nabla u)\partial_{x_i}\varphi\right).\end{aligned}$$

We obtain the following theorem.

Theorem 4.3. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary $\partial\Omega$ and let $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}, (x, y, z) \mapsto f(x, y, z)$, be a continuously differentiable function that satisfies the growth conditions* (4.5). *If $u \in \Phi$ is a minimizer of I, then u satisfies*

$$\int_\Omega \left(\partial_y f(\cdot, u, \nabla u)\varphi + \nabla_z f(\cdot, u, \nabla u) \cdot \nabla\varphi\right) = 0 \qquad \forall \varphi \in W_0^{1,p}(\Omega). \tag{4.7}$$

The equation (4.7) is called the **weak form of the Euler-Lagrange equation**.

Proof. We observe that

$$f(x, y, z) = f(x, 0, 0) + \int_0^1 \frac{d}{dt}(f(x, ty, tz))dt \qquad \forall (x, y, z) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n.$$

Therefore, by following the same steps as in the estimate (4.6) we can show that (4.5) implies that

$$|f(x, y, z)| \leq c(1 + |y|^p + |z|^p),$$

for some constant $c > 0$. This shows that if $u \in W^{1,p}(\Omega)$, then $|I(u)| < \infty$. The theorem now follows from the derivation of the weak form of the Euler-Lagrange equation above. □

Definition 4.4. We call $u \in \Phi$ a **weak solution of the boundary value problem**

$$\begin{aligned} -\mathrm{div}\,(\nabla_z f(x, u(x), \nabla u(x))) + \partial_y f(x, u(x), \nabla u(x)) &= 0 \qquad && x \in \Omega, \\ u(x) &= g(x) && x \in \partial\Omega, \end{aligned} \tag{4.8}$$

if (4.7) holds for all $\varphi \in W_0^{1,p}(\Omega)$.

Remark. We remark that the PDE in (4.8) is of second order, however, the definition of weak solutions only requires that the solution possesses weak derivatives of first order.

On the other hand, if a weak solution of the boundary value problem (4.8) is of class C^2, then it is a solution of the boundary value problem in the classical sense, i.e. it satisfies the PDE and boundary values pointwise. This can be shown by integration by parts and applying the Fundamental Lemma of the Calculus of Variations.

Example 4.5. In case of the Dirichlet integral,

$$I(u) = \frac{1}{2} \int_\Omega |\nabla u|^2,$$

we have $f(z) = \frac{1}{2}z^2$, which satisfies (4.5) with $p = 2$. The weak form of the Euler-Lagrange equation is

$$\int_\Omega \nabla u \cdot \nabla \varphi = 0 \qquad \forall \varphi \in W_0^{1,2}(\Omega).$$

Remark 4.6. We remark that the results of this chapter remain valid in the vector-valued case, i.e. for functions $u : \Omega \to \mathbb{R}^m$, $m \geq 2, \Omega \subset \mathbb{R}^n$. Then, the function $f : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m \times n} \to \mathbb{R}$,

$$(x, y, z) = (x_1, \ldots, x_n, y_1, \ldots, y_m, z_{11}, \ldots, z_{mn}) \mapsto f(x, y, z),$$

is a assumed to be of class C^1 and I is given by

$$I(u) = \int_\Omega f(x, u(x), Du(x))dx, \qquad u \in W^{1,p}(\Omega; \mathbb{R}^m),$$

where $Du(x)$ denotes the Jacobian matrix. Under suitable growth assumptions on f the first variation exists and the corresponding Euler-Lagrange equations are a *system of quasilinear elliptic PDEs*,

$$-\mathrm{div}\,(D_z f(x, u(x), Du(x))) + D_y f(x, u(x), Du(x)) = 0, \qquad x \in \Omega,$$

where $D_z f$ is an $\mathbb{R}^{m \times n}$-matrix with entries $\partial_{z_{ij}} f$. The system can be written componentwise as

$$-\sum_{j=1}^n \partial_{x_j} \partial_{z_{ij}} f(x, u(x), Du(x)) + \partial_{y_i} f(x, u(x), Du(x)) = 0, \qquad i = 1, \ldots, m.$$

Chapter 5

Direct methods

Throughout this chapter $\Omega \subset \mathbb{R}^n$ denotes an open and bounded subset with Lipschitz boundary $\partial\Omega$. We aim to prove the existence of minimizers for functionals of the form

$$I(u) = \int_\Omega f(\cdot, u, \nabla u) \tag{5.1}$$

in the class $\Phi = \{u \in W^{1,p}(\Omega) : u - g \in W_0^{1,p}(\Omega)\}$, where $g \in W^{1,p}(\Omega)$ is a given function and $1 < p < \infty$. The minimal requirement for the function $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is that it is a *Carathéodory function*, i.e. $x \mapsto f(x, y, z)$ is measurable for every $(y, z) \in \mathbb{R} \times \mathbb{R}^n$ and $(y, z) \mapsto f(x, y, z)$ is continuous for almost every $x \in \Omega$. However, to simplify the presentation we typically assume that f is more regular. We use the *direct method* that allows us to prove the existence of minimizers without the detour of solving differential equations (Euler-Lagrange equations). We first explain the key ideas of the method in an abstract setting and then apply it to functionals of the above form.

The origins of the direct methods go back to David Hilbert, Henri Lebesgue and Leonida Tonelli who proved the existence of a minimizer for the Dirichlet integral at the beginning of the 20th century. However, Sobolev spaces did not exist at that time and the result was phrased differently. The problem of the existence of minimizers for the Dirichlet integral was important in the development of analysis in general, in particular, for functional analysis, measure theory, the theory of Sobolev spaces and PDE theory.

5.1 General strategy

Let X be a topological space (e.g. a Banach space with the norm topology or a closed and convex subset of a Banach space with the weak or weak* topology) and $J : X \to \mathbb{R} \cup \{\pm\infty\}$. We aim to solve the abstract minimization problem

$$\inf_{u \in X} J(u), \tag{5.2}$$

i.e. we aim to find $u^* \in X$ such that $J(u^*) = \min_{u \in X} J(u)$.

Definition 5.1. We call J **sequentially coercive** if for all $\alpha \in \mathbb{R}$ the sublevel set $\{u \in X : J(u) \leq \alpha\}$ is sequentially precompact, i.e. for every sequence $(u_m)_{m\in\mathbb{N}} \subset X$ such that $J(u_m)_{m\in\mathbb{N}} \leq \alpha$ for some $\alpha \in \mathbb{R}$ possesses a convergent subsequence.

We call J **sequentially lower semicontinuous** if for all sequences $(u_m)_{m\in\mathbb{N}} \subset X$ with $u_m \to u$ in X as $m \to \infty$ it holds that

$$\liminf_{m\to\infty} J(u_m) \geq J(u).$$

The direct method provides sufficient conditions for the existence of a minimizer and can be phrased as follows.

Theorem 5.2 (Direct Method)**.** *Let X be a topological space $J : X \to \mathbb{R} \cup \{\pm\infty\}$ be sequentially coercive and sequentially lower semicontinuous. Then, there exists at least one solution of* (5.2)*, i.e. there exists $u^* \in X$ such that $J(u^*) = \min_{u\in X} J(u)$.*

Proof. Assume that there exists $u \in X$ such that $J(u) < \infty$. Otherwise, every $u \in X$ is a minimizer of the (degenerate) minimization problem. We consider a minimizing sequence $(u_m)_{m\in\mathbb{N}} \subset X$,

$$\lim_{m\to\infty} J(u_m) = \alpha := \inf\{J(u) : u \in X\} < \infty.$$

Either the sequence $(J(u_m))_{m\in\mathbb{N}}$ diverges to $-\infty$ or the sequence converges. Hence, there exists $\mu \in \mathbb{R}$ such that $J(u_m) \leq \mu$ for all $m \in \mathbb{N}$. By the sequential coercivity of J, the sequence $(u_m)_{m\in\mathbb{N}}$ possesses a convergent subsequence, which we denote again by $(u_m)_{m\in\mathbb{N}}$, such that $u_m \to u^*$ in X.

By the sequential lower semicontinuity of J is follows that

$$\alpha \leq J(u^*) \leq \liminf_{m\to\infty} J(u_m) = \alpha.$$

This implies that $\alpha = J(u^*)$, i.e. u^* is a minimizer. □

The direct method reduces the existence proof for minimizers to establishing sequential coercivity and sequential lower semicontinuity of the functional. We will see that for integral functionals such as (5.1), sequential lower semicontinuity follows from convexity properties of the integrand and the coercivity from certain lower bounds. We remark that both properties depend on the chosen topology in X. In fact, for stronger topologies (i.e. topologies for which fewer sequences converge) the sequential lower semicontinuity of J becomes a weaker requirement, but it becomes harder to show that J is sequentially coercive. For weaker topologies the opposite is true.

The functional J and the set of admissible functions X are typically given in a concrete problem, but we need to choose a topology in which both sequential coercivity and sequential lower semicontinuity of the functional hold. We always consider infinite-dimensional Banach spaces and hence, can use either strong or weak convergence. In most cases, sequential coercivity with respect to the strong convergence does not hold, whereas sequential coercivity with respect to the weak convergence is true under reasonable assumptions. Therefore, we consider weak convergence and use the direct method in the following version:

Corollary 5.3. *Let X be a Banach space and let $J : X \to \mathbb{R} \cup \{\pm\infty\}$ be a functional with the following properties:*

- ***Sequential weak coercivity:*** *For all $\alpha \in \mathbb{R}$, the sublevel set $\{u \in X : J(u) \leq \alpha\}$ is sequentially weakly relatively compact, i.e. for every sequence $(u_m)_{m\in\mathbb{N}} \subset X$ such that $J(u_m) \leq \alpha$ for all $m \in \mathbb{N}$ and some $\alpha \in \mathbb{R}$ possesses a weakly convergent subsequence.*

- ***Sequential weak lower semicontinuity:*** *For every sequence $(u_m)_{m\in\mathbb{N}} \subset X$ such that $u_m \rightharpoonup u$ (weak convergence), we have*

$$J(u) \leq \liminf_{m\to\infty} J(u_m).$$

Then, there exists $u^* \in X$ such that $J(u^*) = \min_{u\in X} J(u)$.

In *reflexive Banach spaces X* we have the following sufficient condition for sequential weak coercivity.

Proposition 5.4. Let X be a reflexive Banach space and $J : X \to \mathbb{R}$ be a functional such that

$$\forall (u_m)_{m\in\mathbb{N}} \subset X \text{ such that } \lim_{m\to\infty} \|u_m\| = \infty \quad \text{it follows that} \quad \lim_{m\to\infty} J(u_m) = \infty.$$

Then, J is sequentially weakly coercive.

Proof. For a proof we refer to the tutorials. □

In the next sections, we formalize these ideas for Sobolev spaces and functionals of the form (5.1). Before we state sufficient conditions for weak coercivity and weak lower semicountinuity for general functionals we consider the Dirichlet integral as a model case.

5.2 Model case: Dirichlet integral

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary $\partial\Omega$. We consider the functional

$$I(u) = \int_\Omega |\nabla u(x)|^2 \, dx, \tag{5.3}$$

and aim to find a minimizer of I within the class

$$\Phi = \left\{ u \in W^{1,2}(\Omega) : u - g \in W_0^{1,2}(\Omega) \right\},$$

for a given $g \in W^{1,2}(\Omega)$.

Theorem 5.5. *There exists a unique minimizer $u \in \Phi$ of* (5.3) *and u satisfies the weak form of the corresponding Euler-Lagrange equations*

$$\int_\Omega \nabla u \cdot \nabla \varphi = 0 \qquad \forall \varphi \in W_0^{1,2}(\Omega). \tag{5.4}$$

Conversely, if $u \in \Phi$ satisfies (5.4)*, then it is a minimizer of* (5.3).

Proof. The proof is divided into several steps. We first show the existence of the minimizer.

Step 1: We observe that $g \in \Phi$ and this implies that

$$0 \leq m := \inf\{I(u) : u \in \Phi\} \leq I(g) < \infty.$$

Let $(u_k)_{k\in\mathbb{N}} \in \Phi, k \in \mathbb{N}$, be a minimizing sequence, i.e.

$$I(u_k) \xrightarrow[k\to\infty]{} m.$$

We aim to show that $(u_k)_{k\in\mathbb{N}}$ is bounded. To this end, we estimate the functional $I(u_k)$ as follows,

$$\begin{aligned}
I(u_k) &= \frac{1}{2}\int_\Omega |\nabla u_k|^2 + \frac{1}{2}\int_\Omega |\nabla(u_k - g) + \nabla g|^2 \\
&= \frac{1}{2}\int_\Omega |\nabla u_k|^2 + \frac{1}{2}\int_\Omega \Big(|\nabla(u_k - g)|^2 + |\nabla g|^2 + 2\nabla(u_k - g)\cdot\nabla g\Big) \\
&\geq \frac{1}{2}\int_\Omega |\nabla u_k|^2 + \frac{c}{2}\int_\Omega |u_k - g|^2 + \frac{1}{2}\int_\Omega \Big(2\nabla u_k\cdot\nabla g - |\nabla g|^2\Big) \\
&= \frac{1}{2}\int_\Omega |\nabla u_k|^2 + \frac{c}{2}\int_\Omega \Big(|u_k|^2 + |g|^2 - 2u_k g\Big) + \frac{1}{2}\int_\Omega \Big(2\nabla u_k\cdot\nabla g - |\nabla g|^2\Big),
\end{aligned}$$

for some constant $c > 0$, where we used Poincaré's inequality (Theorem 3.27) and $u_k - g \in W_0^{1,p}(\Omega)$. We now use Young's inequality, $ab \leq \frac{1}{\varepsilon}a^2 + \varepsilon b^2$, for all $a, b \in \mathbb{R}$ and $\varepsilon > 0$, to conclude that

$$2\nabla u_k\cdot\nabla g \geq -\frac{1}{2}|\nabla u_k|^2 - 8|\nabla g|^2, \qquad -2u_k g \geq -\frac{1}{2}|u_k|^2 - 8|g|^2.$$

Inserting these estimates in the inequality above it follows that

$$\begin{aligned}
I(u_k) &\geq \frac{1}{4}\int_\Omega |\nabla u_k|^2 + \frac{c}{4}\int_\Omega \left(\frac{1}{2}|u_k|^2 - 7|g|^2\right) - \frac{9}{2}\int_\Omega |\nabla g|^2 \\
&\geq c_1\|u_k\|^2_{W^{1,2}(\Omega)} - c_2\|g\|^2_{W^{1,2}(\Omega)},
\end{aligned}$$

for some constants $c_1, c_2 > 0$. Since $g \in W^{1,2}(\Omega)$, $I(u_k) \xrightarrow[k\to\infty]{} m$ and $m < \infty$, this implies that $(u_k)_{k\in\mathbb{N}}$ is a bounded sequence in $W^{1,2}(\Omega)$.

By Proposition 3.34, there exists a subsequence $(u_{k_l})_{l\in\mathbb{N}}$ and $u \in W^{1,2}(\Omega)$ such that

$$u_{k_l} \rightharpoonup u \quad \text{in } W^{1,2}(\Omega).$$

Moreover, by Theorem 3.33 we conclude that $u - g \in W_0^{1,2}(\Omega)$, since Φ is closed (the trace operator is continuous) and convex. This shows that $I : \Phi \to \mathbb{R}$ is weakly coercive.

Step 2: Next, we show the weak lower semicontinuity, i.e. if $(u_k)_{k\in\mathbb{N}}$ is a sequence in $W^{1,2}(\Omega)$ such that $u_k \rightharpoonup u$ in $W^{1,2}(\Omega)$, then $\liminf_{k\to\infty} I(u_k) \geq I(u)$.

In fact, using the estimate

$$\begin{aligned}
|\nabla u_k|^2 &= |\nabla(u_k - u) + \nabla u|^2 = |\nabla(u_k - u)|^2 + |\nabla u|^2 + 2\nabla(u_k - u)\cdot\nabla u \\
&\geq |\nabla u|^2 + 2\nabla(u_k - u)\cdot\nabla u,
\end{aligned}$$

it follows that

$$I(u_k) \geq I(u) + 2\int_\Omega \nabla u\cdot\nabla(u_k - u).$$

The last term converges to zero, since $u \in W^{1,2}(\Omega)$ and $u_k \rightharpoonup u$ in $W^{1,2}(\Omega)$. Consequently, we have

$$\liminf_{k\to\infty} I(u_k) \geq I(u),$$

which proves the claim.

Step 3: By Step 1 I is weakly coercive and by Step 2, I is weakly lower semicontinuous. Hence, Theorem 5.3 implies that $I(u) = m$.

Step 4: To prove uniqueness, let $u, v \in \Phi$ be such that $I(u) = I(v) = m$. Then, $w = \frac{u+v}{2} \in \Phi$, and the convexity of the function $x \mapsto x^2$ implies that

$$m \leq I(w) \leq \frac{1}{2}I(u) + \frac{1}{2}I(v) = m.$$

Consequently, equality must hold and it follows that

$$\frac{1}{2}I(u) + \frac{1}{2}I(v) - I(w) = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla v|^2 - \frac{|\nabla u + \nabla v|^2}{2} \right) = 0.$$

However, by the convexity of $x \mapsto x^2$, the integrand is non-negative, and we conclude that

$$\frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla v|^2 - \frac{|\nabla u + \nabla v|^2}{2} = 0 \qquad \text{a.e. in } \Omega.$$

Since $x \mapsto x^2$ is strictly convex, it follows that $\nabla u = \nabla v$ a.e. in Ω.

Finally, Poincaré's inequality and using that $\nabla u = \nabla v$ a.e. implies that

$$\int_\Omega |u - v|^2 \leq c \int_\Omega |\nabla u - \nabla v|^2 = 0,$$

which proves that $u = v$ a.e. in Ω.

Step 5: The weak form of the Euler-Lagrange equation follows from Theorem 4.3. On the other hand, if $\bar{u} \in \Phi$ satisfies (5.4), then $w = u - \bar{u} \in W_0^{1,2}(\Omega)$ for all $u \in \Phi$. Moreover,

$$\begin{aligned} I(u) = I(\bar{u} + w) &= \int_\Omega |\nabla(\bar{u} + w)|^2 = I(\bar{u}) + \int_\Omega \nabla \bar{u} \cdot \nabla w + I(w) \\ &= I(\bar{u}) + I(w) \geq I(\bar{u}), \end{aligned}$$

where we used that $\bar{u}$ satisfies the weak form of the Euler-Lagrange equations and $I(w) \geq 0$. This shows that $\bar{u}$ is a minimizer of I. □

5.3 Integral functionals with convex integrands

We now address more general functionals and extend the existence (and uniqueness) theory for minimizers that we developed for the Dirichlet integral. Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary $\partial\Omega$. We consider the functional (5.1),

$$I(u) = \int_\Omega f(\cdot, u, \nabla u),$$

where $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. Moreover, the admissible class of functions is

$$\Phi = \{u \in W^{1,p}(\Omega) : u - g \in W_0^{1,p}(\Omega)\},$$

for some $p \in (1, \infty)$ and a given function $g \in W^{1,p}(\Omega)$.

The following lemma provides sufficient conditions for weak coercivity.

Lemma 5.6. *Assume that there exist constants $c_1 > 0$ and $c_2 \geq 0$ such that*

$$f(x,y,z) \geq c_1|z|^p - c_2 \qquad \forall (x,y,z) \in \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n. \tag{5.5}$$

Then, the functional I in (5.1) *is sequentially weakly coercive.*

Remark. Allowing that $I(u) = \infty$, the functional I is well-defined under assumption (5.5).

Proof. Let $u \in \Phi$ with $I(u) \leq c_0$. Then, (5.5) implies that

$$c_1 \int_\Omega |\nabla u|^p \leq \int_\Omega (f(\cdot, u, \nabla u) + c_2) = I(u) + c_2|\Omega| \leq c_0 + c_2|\Omega|. \tag{5.6}$$

Moreover, using the Poincaré inequality (Theorem 3.26) we can estimate the L^p-norm of u by

$$\begin{aligned} \|u\|_{L^p} &\leq \|u - g\|_{L^p} + \|g\|_{L^p} \leq c\||\nabla(u-g)|\|_{L^p} + \|g\|_{L^p} \\ &\leq c\,(\||\nabla u|\|_{L^p} + \||\nabla g|\|_{L^p}) + \|g\|_{L^p}. \end{aligned} \tag{5.7}$$

Let now $(u_k)_{k\in\mathbb{N}} \subset \Phi$ be a sequence such that $I(u_k) \leq c_0$. Then, combining the estimates (5.6) and (5.7) implies that $(u_k)_{k\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega)$. Therefore, by Proposition 3.34 there exists $\bar{u} \in \Phi$ and a subsequence $(u_{k_l})_{l\in\mathbb{N}}$ such that $u_{k_l} \rightharpoonup \bar{u}$ in $W^{1,p}(\Omega)$, which shows that I is weakly coercive. □

Next, we address the weak lower semicontinuity of I. For the proof we need the following result from measure theory.

Remark (Egoroff's Theorem). Let $u_k : \Omega \to \mathbb{R}$, $k \in \mathbb{N}$, be a sequence of measurable functions such that

$$u_k \to u \qquad \text{a.e. in } \Omega,$$

where $\Omega \subset \mathbb{R}^n$ is measurable and $|\Omega| < \infty$. Then, for every $\varepsilon > 0$ there exists a measurable set $A_\varepsilon \subset \Omega$ such that

$$|\Omega \setminus A_\varepsilon| \leq \varepsilon, \qquad u_k \to u \qquad \text{uniformly in } A_\varepsilon.$$

Theorem 5.7. *Let f be bounded from below and assume that*

$$z \mapsto f(x,y,z) \quad \textit{is convex for all } x \in \Omega, y \in \mathbb{R}. \tag{5.8}$$

Then, I is sequentially weakly lower semicontinuous in $W^{1,p}(\Omega)$.

Proof. Let $(u_k)_{k\in\mathbb{N}}$ be a sequence such that $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$. We need to show that

$$m := \liminf_{k\to\infty} I(u_k) \geq I(u).$$

Upon passing to a subsequence, if necessary, we can assume that

$$\lim_{k\to\infty} I(u_k) = m.$$

Step 1: By Proposition 3.31, $(u_k)_{k\in\mathbb{N}}$ is a bounded sequence in $W^{1,p}(\Omega)$. Moreover, since the embedding $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact (Theorem 3.25), we conclude that, upon passing to a subsequence, $u_k \to u$ in $L^p(\Omega)$. Again, upon passing to a subsequence, if necessary, we can assume that $u_k \to u$ a.e. in Ω by Theorem 3.6.

By Egoroff's theorem, for every $j \in \mathbb{N}$ there exists a measurable set $A_j \subset \Omega$ such that

$$u_k \to u \quad \text{uniformly in } A_j, \qquad |\Omega \setminus A_j| \le \frac{1}{j}, \tag{5.9}$$

and we can choose the sets in such a way that $A_j \subset A_{j+1}$ for all $j \in \mathbb{N}$.

Furthermore, let $B_j := \{x \in \Omega : |u(x)| + |\nabla u(x)| < j\}$, for $j \in \mathbb{N}$, and observe that $|\Omega \setminus B_j| \to 0$ as $j \to \infty$, since the function $|u| + |\nabla u|$ is integrable. Together with (5.9) it follows that

$$\left|\Omega \setminus (A_j \cap B_j)\right| \xrightarrow[j\to\infty]{} 0.$$

Step 2: Without loss of generality we can assume that $f \ge 0$. Otherwise we apply the arguments to $f + \alpha$, for some $\alpha > 0$. We observe that

$$\begin{aligned} I(u_k) = \int_\Omega f(\cdot, u_k, \nabla u_k) &\ge \int_{A_j \cap B_j} f(\cdot, u_k, \nabla u_k) \\ &\ge \int_{A_j \cap B_j} \left(f(\cdot, u_k, \nabla u) + \nabla_z f(\cdot, u_k, \nabla u) \cdot \nabla(u_k - u) \right), \end{aligned}$$

where we used that f is convex with respect to z, see (5.8). Since f is continuously differentiable and $(u_k)_{k\in\mathbb{N}}$ converges uniformly to u in A_j, it follows that

$$\begin{aligned} f(\cdot, u_k, \nabla u) &\xrightarrow[k\to\infty]{} f(\cdot, u, \nabla u), \\ \nabla_z f(\cdot, u_k, \nabla u) &\xrightarrow[k\to\infty]{} \nabla_z f(\cdot, u, \nabla u), \end{aligned}$$

uniformly in $A_j \cap B_j$. Furthermore, $|\nabla u_k| \rightharpoonup |\nabla u|$ in $L^p(\Omega)$ and therefore, we conclude that

$$\begin{aligned} m = \lim_{k\to\infty} I(u_k) &\ge \lim_{k\to\infty} \int_{A_j \cap B_j} \left(f(\cdot, u_k, \nabla u) + \nabla_z f(\cdot, u_k, \nabla u) \cdot \nabla(u_k - u) \right) \\ &\ge \int_{A_j \cap B_j} f(\cdot, u, \nabla u). \end{aligned}$$

For the proof of the convergence of the second term we refer to the tutorials.

Step 3: Finally, by monotone convergence and since $(A_j \cap B_j) \subset (A_{j+1} \cap B_{j+1})$, it follows that

$$m \ge \lim_{j\to\infty} \int_{A_j \cap B_j} = I(u).$$

□

Combining Theorem 5.7 with Lemma 5.6 we obtain an existence theorem for minimizers.

Theorem 5.8. *Let f be continuously differentiable and satisfy the coercivity assumption* (5.5) *and the convexity assumption* (5.8)*. Then, there exists $u \in \Phi$ such that*

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

Proof. Lemma 5.6 implies that I is weakly coercive and by Theorem 5.7 I is weak lower semi-continuous. Hence, the statement is an immediate consequence of Corollary 5.3. $\square$

Remark 5.9.

- If, in addition, f satisfies the hypotheses of Theorem 4.3, then, the minimizer satisfies the weak form of the Euler-Lagrange equations, i.e.

  $$\int_\Omega \left(\partial_y f(\cdot, u, \nabla u)\varphi + \nabla_z f(\cdot, u, \nabla u) \cdot \nabla\varphi\right) = 0 \qquad \forall \varphi \in W_0^{1,p}(\Omega).$$

  Moreover, if the minimizer $u \in C^2(\overline{\Omega})$ and $f : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then one can show that u is a classical solution of the Euler-Lagrange equation, i.e.

  $$\begin{aligned} -\operatorname{div}\left(\nabla_z f(\cdot, u, \nabla u)\right) + \partial_y f(\cdot, u, \nabla u) &= 0 && \text{in } \Omega, \\ u &= g && \text{on } \partial\Omega. \end{aligned}$$

- If u is real-valued, $u : \Omega \to \mathbb{R}$, the hypothesis of Theorem 5.8 are nearly optimal, in the sense that weakening any of them, there exists a counterexample to the existence of a minimizer (see tutorials and [5]).

  The theorem can be generalized for the vectorial case, $u : \Omega \to \mathbb{R}^N, N \geq 2$, however, the assumptions are then far from being optimal.

- Under additional assumptions, one can show that the minimizer is, in fact, more regular, i.e. of class C^1, C^2 or even C^∞. We have shown such regularity results for one-dimensional problems in Chapter 2. For higher dimensional problems, showing additional regularity is much more involved.

Example 5.10.

(i) For the Dirichlet integral, we have

$$f(x, y, z) = f(z) = \frac{1}{2}z^2.$$

Hence, all hypotheses of Theorem 5.8 are satisfies for $p = 2$. In this case, we have even shown stronger results (see Theorem 5.5).

(ii) For the p-Laplacian we have

$$f(x, y, z) = f(z) = \frac{1}{p}z^p, \qquad 1 < p < \infty.$$

All hypotheses of Theorem 5.8 are satisfies and the hypothesis of Theorem 4.3 as well. The corresponding Euler-Lagrange equation is

$$\operatorname{div}\left(|\nabla u|^{p-2}\nabla u\right) = 0 \qquad \text{in } \Omega.$$

(iii) Weierstrass' example: We consider

$$I(u) = \int_0^1 x(u'(x))^2 dx$$

with $\Phi = \{u \in W^{1,2}((0,1)) : u(0) = 1, u(1) = 0\}$. In this case, the coercivity assumption (5.5) is not satisfied and one can show that no minimizer exists (see tutorials).

(iv) Bolza's example: We consider

$$I(u) = \int_0^1 ((u'(x))^2 - 1)^2 + (u(x))^4 dx$$

with $\Phi = W_0^{1,4}((0,1))$. Then, the convexity assumption (5.8) is not satisfied and one can show that no minimizer exists (see tutorials).

Convexity (and even strict convexity) in the variable z is not sufficient for the uniqueness of minimizers. However, we can prove uniqueness, e.g. if f is **strictly convex in** y **and** z, i.e.

$$f(x, \lambda y_1 + (1-\lambda y_2), \lambda z_1 + (1-\lambda)z_2) < \lambda f(x, y_1, z_1) + (1-\lambda) f(x, y_2, z_2) \tag{5.10}$$

for all $x \in \Omega, y_1, y_2 \in \mathbb{R}, z_1, z_2 \in \mathbb{R}^n$ such that $y_1 \neq y_2, z_1 \neq z_2$ and $\lambda \in (0,1)$.

Theorem 5.11. *Let f be strictly convex in y and z, i.e. f satisfies* (5.10). *Then, there exists at most one $u \in \Phi$ such that*

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

Proof. Let $u, v \in \Phi$, $\lambda \in (0,1)$ and consider the function $\eta : \Omega \to \mathbb{R}$,

$$\eta(\cdot) = f(\cdot, \lambda u + (1-\lambda)v, \lambda \nabla u + (1-\lambda)\nabla v) - \lambda f(\cdot, u, \nabla u) - (1-\lambda) f(\cdot, v, \nabla v).$$

We observe that $\eta \leq 0$ in Ω and $\eta(x) < 0$ if $u(x) \neq v(x)$. Consequently, it follows that

$$I(\lambda u + (1-\lambda)v) - \lambda I(u) + (1-\lambda)I(v) = \int_\Omega \eta \leq 0.$$

This inequality is strict, unless $u = v$ a.e. in Ω, which implies that I is strictly convex.

Finally, assume that u and v are both minimizers of I and $u \neq v$. Then, using that $\frac{u+v}{2} \in \Phi$ and the strict convexity of I, we conclude that

$$\inf_{w \in \Phi} I(w) \leq I\left(\frac{u+v}{2}\right) < \frac{1}{2} I(u) + \frac{1}{2} I(v) = \inf_{w \in \Phi} I(w),$$

which is a contradiction. □

Finally, we show that for functionals where the integrand only depends on ∇u the convexity assumption in Theorem 5.7 is also necessary for sequential weak lower semicontinuity.

Theorem 5.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous and satisfy the growth assumption*

$$|f(z)| \leq c(1 + |z|^p) \qquad \forall z \in \mathbb{R}^n,$$

for some $c > 0$ and $1 < p < \infty$. We consider the functional

$$I(u) = \int_\Omega f(\nabla u(x)) dx, \qquad u \in W^{1,p}(\Omega).$$

If I is sequentially weakly lower semicontinuous then f is convex.

Proof. We only present the proof for $n = 1$, i.e. $f : \mathbb{R} \to \mathbb{R}$ and $\Omega \subset \mathbb{R}$ is an open interval. For the general case we refer to [4]. We need to show that f is convex, i.e.

$$f(\lambda z_1 + (1-\lambda)z_2) \le \lambda f(z_1) + (1-\lambda)f(z_2) \qquad \forall z_1, z_2 \in \mathbb{R}, \lambda \in (0,1). \tag{5.11}$$

Step 1: We consider the function $u(x) = zx, x \in \mathbb{R}$, where $z = \lambda z_1 + (1-\lambda)z_2$ which implies that $u' = z$ and hence, $f(u') = f(z)$. We aim to construct a sequence $(u_k)_{k\in\mathbb{N}}$ such that $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$ and such applying the weak lower semicontinuity of I with this sequence implies (5.11).

Let $\tilde{v} : [0,1] \to \mathbb{R}$ be defined as

$$\tilde{v}(x) = \begin{cases} (z_1 - z)x, & x \in [0,\lambda], \\ (z_2 - z)x + z - z_2, & x \in [\lambda, 1], \end{cases}$$

and extend $\tilde{v}$ periodically to a function $v : \mathbb{R} \to \mathbb{R}$, i.e. $v(x) = \tilde{v}(y)$ if $x = m + y, m \in \mathbb{Z}, y \in [0,1]$. Now, we consider the sequence $u_k(x) = \frac{1}{k}v(kx) + u(x), k \in \mathbb{N}$, and show that

$$||u_k - u||_{L^\infty} \le \frac{c}{k} \qquad \forall k \in \mathbb{N}. \tag{5.12}$$

Indeed, we have

$$|u_k(x) - u(x)| = \left|\frac{1}{k}v(kx)\right| = \frac{1}{k}|\tilde{v}(y)| \le \frac{1}{k}\lambda(1-\lambda)(z_1 - z_2) \qquad \forall x \in \Omega,$$

where $kx = l + y$, $l \in \mathbb{Z}, y \in [0,1]$.

Step 2: We show that $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$. Note that by Step 1 it remains to show that $u_k' \rightharpoonup u$ in $L^q(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. Indeed, since Ω is bounded (5.12) implies that $u_k \to u$ in $L^q(\Omega)$ for all $1 \le q \le \infty$.

Let $\phi \in C_c^\infty(\Omega)$, then

$$\int_\Omega u_k'\phi = -\int_\Omega u_k\phi' \xrightarrow[k\to\infty]{} \int_\Omega u\phi' = \int_\Omega u\phi', \tag{5.13}$$

where we that $u_k \rightharpoonup u$ in $L^1(\Omega)$. If $g \in L^q(\Omega), \frac{1}{p} + \frac{1}{q} = 1$, the claim follows by an approximation argument. Indeed, $C_c^\infty(\Omega)$ is dense in $L^q(\Omega)$ and hence, for all $\varepsilon > 0$ there exists $\phi \in C_c^\infty(\Omega)$ such that $||g - \phi|| < \varepsilon$. Using Hölder's inequality we conclude that

$$\begin{aligned} \left|\int_\Omega u_k' g - \int_\Omega u' g\right| &\le \int_\Omega \left|u_k'\phi - u'\phi\right| + \int_\Omega \left|u_k'(g-\phi)\right| + \int_\Omega \left|u'(g-\phi)\right| \\ &\le \int_\Omega \left|(u_k - u)'\phi\right| + \sup_{k\in\mathbb{N}} ||u_k'||_{L^p}||g-\phi||_{L^q} + ||u'||_{L^p}||g-\phi||_{L^q} \\ &\le \int_\Omega \left|(u_k - u)'\phi\right| + c||g-\phi||_{L^q}, \end{aligned}$$

for some constant $c \ge 0$. The first terom on the right hand side converges by (5.13), and the second term can be made arbitrarily small by the denseness of $C_c\infty(\Omega)$ is dense in $L^q(\Omega)$.

Step 3: By assumption, I is weakly lower semicontinuous and $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$. This implies that

$$|\Omega|f(z) = |\Omega|f(\lambda z_1 + (1-\lambda)z_2) \le \liminf_{k\to\infty} \int_\Omega f(u_k').$$

We observe that

$$u_k'(x) = v'(kx) + u'(x) = \begin{cases} z_1 & y \in [0,\lambda], \\ z_2 & y \in [\lambda, 1], \end{cases}$$

where $kx = l + y$, $l \in \mathbb{Z}, y \in [0,1]$.

Finally, if $\Omega = (m,l)$ with $m, l \in \mathbb{Z}$ then

$$\begin{aligned} \left|\{x \in \Omega : u_k'(x) = z_1\}\right| &= \lambda|\Omega|, \\ \left|\{x \in \Omega : u_k'(x) = z_2\}\right| &= (1-\lambda)|\Omega|. \end{aligned}$$

We conclude that

$$\begin{aligned} |\Omega| f(z) = \liminf_{k\to\infty} \int_\Omega f(u_k') &= \liminf_{k\to\infty} \left(\int_{\{u_k'=z_1\}} f(u_k') + \int_{\{u_k'=z_2\}} f(u_k') \right) \\ &= \lambda|\Omega| f(z_1) + (1-\lambda)|\Omega| f(z_2), \end{aligned}$$

which proves (5.11). If Ω is an arbitrary open interval, we can approximate it with a sequence of intervals with integer endpoints and the error will tend to zero as $k \to \infty$. □

From Theorem 5.7 and Theorem 5.12 applied to f and $-f$, we get a characterization of integral functions that are continuous with respect to the weak topology of L^p.

Corollary 5.13. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous and satisfy the growth assumption*

$$|f(z)| \le c(1 + |z|^p) \qquad \forall z \in \mathbb{R}^n,$$

for some $c > 0$ and $1 < p < \infty$. Consider the functional

$$I(u) = \int_\Omega f(\nabla u(x))dx, \qquad u \in W^{1,p}(\Omega).$$

Then, I is sequentially weakly continuous then f is linear.

Chapter 6

Relaxation

We now consider the following situation: suppose that we want to prove the existence of a solution to the minimization problem

$$\min\{ F(x) : x \in X \} \tag{6.1}$$

where $F : X \to (-\infty, +\infty]$. What we have seen in the last chapter is that, if we can find a topology for which F is coercive and lower semicontinuous, then the Direct Method of Calculus of Variations (see Theorem 5.2) ensures the existence of a solution. Note that we need to *choose* a topology!

The two properties required by Direct Methods wants different properties of the topology: (sequential) lower semicontinuity is a condition on weakly converging sequences. The less weakly converging sequences we have, the easier it is for a functional to be weak lower semicontinuous. Thus, lower semicontinuity prefers a topology with a lot of open sets, so it is more difficult for a sequence to converge. On the other hand, coercivity would like as few open sets as possible, since it's a property about compactness. This competition between the two required properties makes the choice of the topology a crucial part of the task to solve a minimization problems. Usually, this choice is dictated by considerations coming from the nature of the phenomenon modeled by F that reflects on the pre-compactness of sequences with equibounded energy, namely considerations on the coercivity: assume that $(x_n)_{n\in\mathbb{N}} \subset X$ is such that

$$\sup_{n\in\mathbb{N}} F(x_n) < +\infty.$$

Is there a topology ensuring that $(x_n)_{n\in\mathbb{N}}$ is pre-compact? This fixed the *natural* topology for the problem[1]. Then, the question is whether or not the functional F is lower semicontinuous or not with respect to such topology. If it is, well, we are in business to get the existence of a solution to the minimization problem. As you can guess, the interesting case is when it is not! This is not simply a technicality for mathematicians, but it is the way the model really captures complex behaviors of what situation F represents. In particular, oscillations, concentration, and jumps. Here, we will focus of the former type. The latter two require the introduction of functional spaces more general than Sobolev spaces (like spaces of function of bounded variations, or Young measures), and it is beyond the scope of the course.

[1] Of course, one can consider different topologies.

Consider the case where $F : X \to (-\infty, +\infty]$ is a functional, and a topology on X. We are not assuming F to be lower semicontinuous, or even coercive with respect to such topology.Thus, the minimization problem (6.1) does not necessarily have a solution. Nevertheless, we wonder if it is possible to understand the behavior of a minimizing sequence $(x_n)_{n\in\mathbb{N}} \subset X$ for F. In particular, we are interested in a *variational* characterization of minimizing sequences for F. Namely, we want to understand if we can we find a functional $\overline{F} : X \to (-\infty, +\infty]$ such that

$$\inf\{\, F(x) : x \in X \,\} = \min\{\, \overline{F}(x) : x \in X \,\}, \tag{6.2}$$

and if $(x_n)_{n\in\mathbb{N}} \subset X$ is a minimizing sequence for F, then, up to a subsequence, it converges (in some topology) to a minimizer of $\overline{F}$. Moreover, it would be also important to have that every minimizer of $\overline{F}$ is the limit of a minimizing sequence for F. In this way we can study minimizers of $\overline{F}$ and, by using the two properties above, be sure that objects that almost minimize the initial functional F are close (with respect to the chosen topology) to them.

We focus on the case of functionals defined on a metric space. This is because the metric structure allows for a *sequential* characterization of the relaxed functional, that coincides with the topological definition. This equivalence is not true in a general topological space. Nevertheless, motivated by applications, we will discuss the case of relaxation with respect to weak topologies.

6.1 Relaxation in metric spaces

In this section, (X, d) will denote a metric space. Moreover, with $\overline{\mathbb{R}}$ we denote $\mathbb{R} \cup \{\pm\infty\}$

The question we want to address is the following: how to find, if possible, such a functional $\overline{F}$? The answer is quite intuitive, if you think about it for a moment, before reading the definition. Consider the function F in Figure 6.1. This function is not lower semicontinuous at the point $\bar{x}$. The lack of lower semicontinuity is because we *defined F at $\bar{x}$ in the wrong way!* Indeed, the *correct* value would have been $F(\bar{x}) = 1$. Why is that? Well, because you can find a sequence $(x_n)_{n\in\mathbb{N}}$ converging to $\bar{x}$ with

$$\lim_{n\to\infty} F(x_n) = 1.$$

Note that the fact that the limit exists is a very nice coincidence. We could have used the liminf instead.

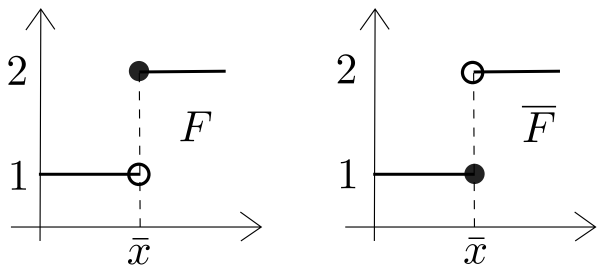


Figure 6.1: The function F and its sequential lower semicontinuous envelope $\overline{F}$.

Thus, given $x \in X$, in order to define $\overline{F}(x)$ we want to look at all possible ways to approximate x with a sequence $(x_n)_{n\in\mathbb{N}}$, and choose that that gives the lowest amount of energy. This gives the heuristics for the following definition.

Definition 6.1. Let $F : X \to \overline{\mathbb{R}}$. We define $\overline{F} : X \to \overline{\mathbb{R}}$, the *sequential lower semicontinuous envelope of* F, as

$$\overline{F}(x) := \inf\left\{ \liminf_{n\to\infty} F(x_n) : x_n \to x \right\},$$

for every $x \in X$.

Remark 6.2. It is important to note the the lower semicontinuous envelope *depends* on the topology we are using to define the notion of convergence of sequences. Different topologies will lead to different lower semicontinuous envelopes for the same functional F.

Remark 6.3. Note that $\overline{F} \leq F$.

Remark 6.4. Note that the sequential lower semicontinuous envelope $\overline{F}$ has a *local* character. Namely, if $G : X \to \overline{\mathbb{R}}$ is a functional such that $G = F$ in a ball $B(x,r)$, for some $r > 0$, than $\overline{F}(y) = \overline{G}(y)$ for all $y \in B(x,r)$.

First of all, we would like to justify the name of the functional $\overline{F}$.

Lemma 6.5. *The functional $\overline{F}$ is sequentially lower semicontinuous.*

Proof. Let $x \in X$, and let $(x_n)_{n\in\mathbb{N}}$ be such that $x_n \to x$. We want to prove that

$$\overline{F}(x) \leq \liminf_{n\to\infty} \overline{F}(x_n).$$

Without loss of generality, we can assume $\overline{F}(x) > -\infty$, otherwise there is nothing to prove. For each $n \in \mathbb{N}$, by definition of $\overline{F}$, we get that there exists $y_n \in X$ such that

$$\mathrm{d}(x_n, y_n) < \frac{1}{n}, \qquad F(y_n) \leq \overline{F}(x_n) + \frac{1}{n}. \tag{6.3}$$

Note that, the second inequality in (6.3), we cannot have $\overline{F}(x_n) = -\infty$, for n large. Indeed, by using our assumption, and the definition of $\overline{F}(x)$, we get

$$-\infty < \overline{F}(x) \leq \liminf_{n\to\infty} F(y_n).$$

By choice of the sequence $(y_n)_{n\in\mathbb{N}}$, we get that $y_n \to x$. Therefore, by definition of $\overline{F}(x)$, and by using 6.3, we obtain that

$$\overline{F}(x) \leq \liminf_{n\to\infty} F(y_n) \leq \liminf_{n\to\infty} \left[\overline{F}(x_n) + \frac{1}{n} \right] = \liminf_{n\to\infty} \overline{F}(x_n).$$

This proves the desired result. $\square$

We now prove a property of the sequential lower semicontinuous envelope that is very useful when .

Lemma 6.6. *Let $F : X \to \mathbb{R}$. Then, for every $x \in X$ there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that*

$$\overline{F}(x) = \lim_{n\to\infty} F(x_n),$$

and with $x_n \to x$.

Proof. First of all, note that if $\overline{F}(x) = +\infty$, then for every sequence $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \to x$, we have

$$\lim_{n\to\infty} F(x_n) = +\infty.$$

Thus, any sequence works. Now, assume $\overline{F}(x) < +\infty$. By definition of $\overline{F}(x)$, for each $n \in \mathbb{N} \setminus \{0\}$, there exists a sequence $(y^n_k)_{k\in\mathbb{N}} \subset X$ with $y^n_k \to x$ as $k \to \infty$, such that

$$\liminf_{k\to\infty} F(y^n_k) \leq \overline{F}(x) + \frac{1}{n},$$

in the case $\overline{F}(x) > -\infty$, and such that

$$\liminf_{k\to\infty} F(y^n_k) \leq -\frac{1}{n},$$

in the case $\overline{F}(x) = -\infty$. Thus, for each $n \in \mathbb{N}$, we can choose $k(n) \in \mathbb{N}$ such that

$$\mathrm{d}(y^n_{k(n)}, x) < \frac{1}{n},$$

and

$$\liminf_{k\to\infty} F(y^n_k) \leq \overline{F}(x) + \frac{2}{n},$$

in the case $\overline{F}(x) > -\infty$, and such that

$$\liminf_{k\to\infty} F(y^n_k) \leq -\frac{1}{2n},$$

in the case $\overline{F}(x) = -\infty$. Define the sequence $(x_n)_{n\in\mathbb{N}}$ as $x_n := y^n_{k(n)}$. This sequence enjoys the desired properties. ◻

In computing the sequential lower semicontinuous envelope, it is useful to focus only on the terms that are not continuous, as the following proposition shows.

Proposition 6.7. Let $F, G : X \to \overline{\mathbb{R}}$. Then,

$$\overline{F + G} \geq \overline{F} + \overline{G},$$

provided that the right-hand side and the left-hand side are well defined. Moreover, the equality

$$\overline{F + G} = \overline{F} + G$$

holds if G is continuous and everywhere finite.

Remark 6.8. The fact that G is continuous and finite every where is essential to have the second part of the above result. The necessity of continuity can be seen by considering the following example: let $F, G : \mathbb{R} \to \mathbb{R}$ be defined as

$$F(t) := \begin{cases} 0 & \text{for } t \neq 0, \\ 1 & \text{for } t = 0, \end{cases} \qquad G(t) := \begin{cases} 1 & \text{for } t \neq 0, \\ 0 & \text{for } t = 0, \end{cases}$$

respectively. Then we have a strict inequality for $t = 0$. Indeed

$$\overline{F}(0) + G(0) = 0 + G(0) < 1 = F(0) + G(0) = \overline{F + G}(0).$$

The necessity of having G finite can be seen by using a similar example.

We now show that the sequential lower semicontinuous envelope is the functionals that we were seeking for in order to characterize the limiting behavior of minimizing sequences for a given functional. Before, we recall the definition of a cluster point.

Definition 6.9. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a topological space. We say that x is a *cluster point*, or *accumulation point*, of the sequence, if there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to x.

Theorem 6.10. *Let $F : X \to \overline{\mathbb{R}}$ be a coercive functional. Then, $\overline{F}$ is coercive and the following holds*

$$\inf\{\, F(x) : x \in X \,\} = \min\{\, \overline{F}(x) : x \in X \,\}. \tag{6.4}$$

Moreover, every cluster point of a minimizing sequence for F us a minimum point for $\overline{F}$. Finally, every minimun point for $\overline{F}$ is the limit of a minimizing sequence for F.

Proof. Note that, without loss of generality, we can assume that $F \not\equiv +\infty$.

Step 1: Coerciveness. Let $\alpha \in \mathbb{R}$. We claim that

$$\left\{\overline{F} \leq \alpha\right\} = \bigcap_{t>\alpha} \overline{\{F \leq t\}}. \tag{6.5}$$

Indeed, let $x \in \left\{\overline{F} \leq \alpha\right\}$, and let $t > \alpha$. Then, by definition of $\overline{F}(\alpha)$, there exists $(x_n)_{n\in\mathbb{N}}$ with $x_n \to x$ such that

$$\overline{F}(\alpha) = \lim_{n\to\infty} F(x_n).$$

Therefore, there exists $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, it holds $F(x_n) \leq t$. Thus, $x \in \overline{\{F \leq t\}}$.

To prove the opposite inclusion, let

$$x \in \bigcap_{t>\alpha} \overline{\{F \leq t\}}.$$

Then, for all $t > \alpha$, there exists $(x_n^t)_{n\in\mathbb{N}} \subset \{F \leq t\}$ such that $x_n^t \to x$. Then,

$$\overline{F}(x) \leq \liminf_{n\to\infty} F(x_n^t) \leq t.$$

Since this holds for all $t > \alpha$, we get that $\overline{F}(x) \leq \alpha$, as desired.

Now, by using the coercivity of F, we have that $\overline{\{F \leq t\}}$ is closed for all $t \in \mathbb{R}$. In particular, from (6.5), we get that $\left\{\overline{F} \leq \alpha\right\}$ is compact for every $\alpha \in \mathbb{R}$, being intersection of compact sets.

Step 2: Validity of (6.4). The functional $\overline{F}$ is sequentially lower semicontinuous (see Lemma 6.5) and sequentially coercive (see Step 1). Therefore, we can apply the Direct Method (see Theorem 5.2) to ensure that the minimization problem for $\overline{F}$ has a solution.

We now prove the equality in (6.4). Note that

$$\inf\{\, F(x) : x \in X\,\} \geq \min\{\, \overline{F}(x) : x \in X\,\}$$

because $F \geq \overline{F}$. To prove the opposite inequality, we argue by contradiction. Assume that

$$\inf\{\, F(x) : x \in X\,\} > \min\{\, \overline{F}(x) : x \in X\,\}, \tag{6.6}$$

and let $\bar{x} \in X$ be a minimum point of $\overline{F}$. Then, by Lemma 6.6 we can find a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \to \bar{x}$ such that

$$\overline{F}(\bar{x}) = \lim_{n\to\infty} F(x_n).$$

This, combined with (6.6), yields

$$\min\{\, \overline{F}(x) : x \in X\,\} = \overline{F}(\bar{x}) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} F(x_n) \geq \inf\{\, F(x) : x \in X\,\} > \min\{\, \overline{F}(x) : x \in X\,\}.$$

This gives the desired contradiction.

Step 3: Cluster points. Let $(x_n)_{n\in\mathbb{N}} \subset X$ be a minimizing sequence for F, and let $\overline{x} \in X$ be a cluster point of it. Then, by the lower semicontinuity of $\overline{F}$, we have that

$$\overline{F}(\overline{x}) \leq \liminf_{n\to\infty} \overline{F}(x_n) \leq \liminf_{n\to\infty} F(x_n) = \inf\{\, F(x) : x \in X\,\} = \min\{\, \overline{F}(x) : x \in X\,\},$$

where in the last step we used (6.4). Thus, $\overline{x}$ is a minimum point for $\overline{F}$.

Step 4: Minimum points of $\overline{F}$. Let $\overline{x} \in X$ be a minimizer of $\overline{F}$. By Lemma 6.6, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ converging to $\overline{x}$ such that

$$\min\{\, \overline{F}(x) : x \in X\,\} = \overline{F}(\overline{x}) = \lim_{n\to\infty} F(x_n).$$

By using (6.4) we then obtain that $(x_n)_{n\in\mathbb{N}}$ is a minimizing sequence for F. □

6.2 Characterizations of the relaxed functional

In this section, we want to provide several characterizations of the sequential lower semicontinuous envelope.

Proposition 6.11. Let $F : X \to \overline{\mathbb{R}}$. Then, $\overline{F}$ is characterized by the following two properties:

(i) (Liminf inequality) For every $x \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ converging to x, it holds

$$\overline{F}(x) \leq \liminf_{n\to\infty} F(x_n);$$

(ii) (Limsup inequality) For every $x \in X$, there exists $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$\overline{F}(x) \geq \limsup_{n\to\infty} F(x_n),$$

and $x_n \to x$.

Proof. Step 1. We first show that $\overline{F}$ satisfies (i) and (ii). Let us start with (i). Let $x \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ converging to x. Then, using the definition of $\overline{F}(x)$ we get

$$\overline{F}(x) = \inf\left\{ \liminf_{n\to\infty} F(y_n) : y_n \to x \right\} \leq \liminf_{n\to\infty} F(x_n).$$

This proves (i). As for (ii), this follows from Lemma 6.6.

Step 2. Let $G : X \to \overline{\mathbb{R}}$ be a functional satisfying the following two properties:

(a) For every $x \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ converging to x, it holds

$$G(x) \leq \liminf_{n\to\infty} F(x_n);$$

(b) For every $x \in X$, there exists $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$G(x) \geq \limsup_{n\to\infty} F(x_n),$$

and $x_n \to x$.

We claim that $G = \overline{F}$. For, we first show that $G \leq \overline{F}$. Fix $x \in X$. Let $(x_n)_{n\in\mathbb{N}} \subset X$ be the sequence provided by Lemma 6.6. Then, $x_n \to x$ and

$$\overline{F}(x) = \lim_{n\to\infty} F(x_n) \geq G(x),$$

where last equality follows from (a). To prove that $G \geq \overline{F}$, fix $x \in X$, and let $(x_n)_{n\in\mathbb{N}} \subset X$ be the sequence given by (b). Then, $x_n \to x$ and

$$G(x) \geq \limsup_{n\to\infty} F(x_n) \geq \overline{F}(x),$$

where last equality follows from (i). This concludes the proof. □

Remark 6.12. The limsup inequality is also called *recovery sequence*. The reason is that, for the sequence $(x_n)_{n\in\mathbb{N}}$ provided by (ii), it holds

$$\overline{F}(x) = \lim_{n\to\infty} F(x_n).$$

This follows from (ii) combined with (i). Note that there can be more than one recovery sequence.

The characterization provided by Proposition 6.11 is how, very often, the relaxed functional is identified. Assume that you are given a functional $F : X \to \overline{\mathbb{R}}$ and you want to identify its sequential lower semicontinuous envelope. What you do is that you have to *guess* a candidate $G : X \to \overline{\mathbb{R}}$. Then:

Step 1: Prove that G is sequentially lower semicontinuous;
Step 2: Prove that $G \leq F$;
Step 3: Prove that the liminf inequality holds;
Step 4: Construct a recovery sequence.

A possible way to guess the candidate G is to work on the liminf inequality, and then prove that the lower bound obtained is optimal, by constructing a recovery sequence. As a first step, though, you need to make sure that the functional G is lower semicontinuous. For, the following properties come in handy.

Proposition 6.13. Let $G : X \to \overline{\mathbb{R}}$. Then, the followings are equivalent:

(i) G is sequentially lower semicontinuous;

(ii) The superlevel set $\{G > t\}$ is open, for all $t \in \mathbb{R}$. Namely, for every $x \in X$ and every $t \in \mathbb{R}$ with $G(x) > t$, there exists $r > 0$ such that $G(y) > t$, for all $y \in B(x, r)$;

(iii) The sublevel set $\{G \leq t\}$ is closed, for all $t \in \mathbb{R}$;

(iv) It holds
$$G(x) = \sup_{r>0} \inf_{y \in B(x,r)} G(y) = \lim_{r \to 0} \inf_{y \in B(x,r)} G(y),$$
for all $x \in X$.

Moreover, in order to prove that a functional is lower semicontinuous, usually one proceeds by building up on the lower semicontinuity of simple functionals. We start by considering the operation of infimum and supremum. Note that we are *(sequentially)* lower semicontinuous, because these results also holds for lower semicontinuous functionals (see Definition 6.23): indeed, the proofs of the results use the topological properties provided by the above characterization.

Proposition 6.14. Let $F_i : X \to \overline{\mathbb{R}}$ be (sequentially) lower semicontinuous for all $i \in I$, where I is a set of indexes (even not countable). Then, the functional $U : X \to \overline{\mathbb{R}}$ defined as
$$U(x) := \sup_{i \in I} F_i(x)$$
is (sequentially) lower semicontinuous. Moreover, for any *finite* set $J \subset I$, it holds that the functional $L : X \to \overline{\mathbb{R}}$ defined as
$$L(x) := \inf_{i \in J} F_i(x)$$
is (sequentially) lower semicontinuous.

Remark 6.15. The infimum of a countable family of lower semicontinuous functions may fail to be lower semicontinuous. As an example, consider the case where $f_n : \mathbb{R} \to \mathbb{R}$ are given by $f_n(x) := ((nx + 1) \vee 0) \wedge 1$. Then
$$\inf_{n \in \mathbb{N}} f_n(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$
which is not lower semicontinuous.

We now consider what happens when considering algebraic operations.

Proposition 6.16. Let $F, G : X \to \overline{\mathbb{R}}$ be (sequentially) lower semicontinuous. Then, λF is (sequentially) lower semicontinuous for all $\lambda > 0$, and, if $F + G$ is well defined (namely there are no cases where we have to sum $\pm\infty$ with $\mp\infty$), then also $F+G$ is (sequentially) lower semicontinuous.

Remark 6.17. Note that, if F is lower semicontinuous, and $\lambda < 0$, then λF is not necessarily lower semicontinuous. Find an example for this!

The proofs of the above two propositions are left as an exercise to the reader.

We now define a topological notion of lower semicontinuous envelope, that has more to do with the word *envelope*. As it usually happens, the sequential and the topological notion are equivalent in metric spaces, but they differ in a general topological space.

Definition 6.18. Let $F : X \to \overline{\mathbb{R}}$. We define $\mathrm{lsc}(F) : X \to \overline{\mathbb{R}}$, the *lower semicontinuous envelope*, as

$$\mathrm{lsc}(F)(x) := \sup\{\, G(x) \;:\; G \le F,\; G : X \to \overline{\mathbb{R}} \text{ is lower semicontinuous}\,\},$$

for every $x \in X$.

Remark 6.19. Using Proposition 6.14, we get that $\mathrm{lsc}(F)$ is the greatest lower semicontinuous functional that is not greater than F.

We first give a local characterization of the lower semicontinuous envelope, in the same spirit as the characterization (v) in Proposition 6.13 of lower semicontinuous functions.

Proposition 6.20. Let $F : X \to \overline{\mathbb{R}}$. Then,

$$\mathrm{lsc}(F)(x) = \sup_{r>0} \inf_{y \in B(x,r)} F(y) = \lim_{r \to 0} \inf_{y \in B(x,r)} F(y),$$

for every $x \in X$.

Proof. Consider the functional

$$H(x) := \sup_{r>0} \inf_{y \in B(x,r)} F(y).$$

Step 1. We claim that $H \le \mathrm{lsc}(F)$. By the definition of $\mathrm{lsc}(F)$, we just need to prove that H is lower semicontinuous and that $H \le F$. The latter property follows directly from the definition of H. As for the former, we use the characterization of lower semicontinuous functions given by Proposition 6.13: we check that the superlevel sets of H are open. Let $t \in \mathbb{R}$, and let $x \in \{H > t\}$. Then, by definition of $H(x)$, there exists $r > 0$ such that

$$\inf_{y \in B(x,r)} F(y) > t.$$

In particular, for each $y \in B(x, r)$, it holds that

$$\inf_{z \in B(y, r - \mathrm{d}(x,y))} F(z) > t,$$

and thus $H(y) > t$ as well. This proves that $B(x, r) \subset \{H > t\}$.

Step 2. To prove the opposite inequality, the lower semicontinuity of H, together with Proposition 6.13, yields that, for every $G : X \to \overline{\mathbb{R}}$ lower semicontinuous and such that $G \leq F$,

$$G(x) = \sup_{r>0} \inf_{y\in B(x,r)} G(y) \leq \sup_{r>0} \inf_{y\in B(x,r)} F(y) = H(x),$$

for every $x \in X$. Thus, $\mathrm{lsc}(F) \leq H$. This concludes the proof. □

We now prove that, in a metric space, the sequential lower semicontinuous envelope coincides with the lower semicontinuous envelope.

Proposition 6.21. Let $F : X \to \overline{\mathbb{R}}$, where (X, d) is a metric space. Then, $\mathrm{lsc}(F) = \overline{F}$.

Proof. Step 1. We first show that $\mathrm{lsc}(F)$ satisfies (i) of Proposition 6.11. Let $\overline{x} \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ converging to $\overline{x}$. By the lower semicontinuity of $\mathrm{lsc}(F)$ we have that

$$\mathrm{lsc}(F)(\overline{x}) \leq \liminf_{n\to\infty} \mathrm{lsc}(F)(x_n) \leq \liminf_{n\to\infty} F(x_n),$$

where in the last step we used the fact that $\mathrm{lsc}(F) \leq F$.

Step 2. We now show that (ii) of Proposition 6.11 holds. Let $\overline{x} \in X$. Assume $\mathrm{lsc}(F)(\overline{x}) < +\infty$, otherwise there is nothing to prove. Let $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be such that

$$\lim_{n\to\infty} t_n = \mathrm{lsc}(F)(\overline{x}), \qquad \mathrm{lsc}(F)(\overline{x}) < t_n. \tag{6.7}$$

Using Proposition 6.20, we get

$$\sup_{k\in\mathbb{N}} \inf_{y\in B(x,1/k)} F(y) = \sup_{r>0} \inf_{y\in B(x,r)} F(y) = \mathrm{lsc}(F)(x) < t_n.$$

Therefore, for every $k \in \mathbb{N} \setminus \{0\}$, we get that

$$\inf_{y\in B(x,1/k)} F(y) < t_n.$$

In particular,

$$\inf_{y\in B(x,1/n)} F(y) < t_n.$$

for all $n \in \mathbb{N} \setminus \{0\}$. Since we have a strict inequality, by using the definition of infimum, we get that for all $n \in \mathbb{N}$ there exists $x_n \in B(\overline{x}, 1/k(n))$ such that

$$F(x_n) < t_n, \tag{6.8}$$

Therefore, the sequence $(x_n)_{n\in\mathbb{N}}$ converges to $\overline{x}$, and, by using (6.7) we get

$$\limsup_{n\to\infty} F(x_n) \leq \lim_{n\to\infty} t_n = \mathrm{lsc}(F)(\overline{x}),$$

where in the last step we used (6.8). Thus, $(x_n)_{n\in\mathbb{N}}$ is a desired recovery sequence. □

6.3 Relaxation with respect to weak and weak* topologies

The previous section was in the framework of a metric space. That was important because some of the results we proved do *not* hold in a general topological space (X, τ). In particular, in proving the sequential characterization of the relaxation (see Proposition 6.21), and the limiting behaviour of minimizing sequences (see Theorem 6.10) we use is the fact that X satisfies the first axiom of countability, namely that every point has a countable neighbourhood basis (for metric spaces, given by balls with rational radius). This is *not* satisfied by weak topologies.

Remark 6.22. Note that the result of Proposition 6.20 still holds in a general topological space, up to replacing the supremum over $r > 0$ with the supremum over all open sets $U \subset X$ such that $x \in U$.

We are interested in such a issue because in the applications that we will see (and in most of the modern applications), weak topologies are those that need to be used in order to identify the lower semicontinuous envelope. In particular, in the next section, we will consider the weak topology of L^p for $p < \infty$ and the weak*-L^∞ for $p = \infty$. These topologies does *not* come from a metric, and therefore the sequential characterization given by Proposition 6.21 is not valid, for a general functional F. We thus wonder what is the object given by the sequential characterization of Proposition 6.21, and what is its relation to the lower semicontinuous envelope.

First of all, we give the topological definition of lower semicontinuity.

Definition 6.23. Let (X, τ) be a topological space, and let $F : X \to \overline{\mathbb{R}}$. We say that F is *lower semicontinuous at the point* $x \in X$ if, for every $t \in \mathbb{R}$ with $F(x) > t$, there exists an open neighborhood U of x such that

$$F(y) > t$$

for all $y \in U$. We say that F is *lower semicontinuous* if it is lower semicontinuous at all points.

Remark 6.24. A lower semicontinuous function is sequentially lower semicontinuous, but the opposite is not true in general. Proposition 6.13 shows that, in a metric space, the two notions coincide.

By using Proposition 6.14 we get the following.

Lemma 6.25. *Let $F : X \to \overline{\mathbb{R}}$. Then,* $\mathrm{lsc}(F)$ *is lower semicontinuous.*

Remark 6.26. Note that the *sequential* lower semicontinuous envelope $\overline{F}$ might fail to be lower semicontinuous in a general topological space.

We now state a similar characterization of the lower semicontinuous envelope in the same spirit as those for the sequential lower semicontinuous envelope.

Lemma 6.27. *Let $F : X \to \overline{\mathbb{R}}$. Then,*

$$\mathrm{lsc}(F)(x) = \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y),$$

for all $x \in X$, where $\mathcal{N}(x)$ denotes the family of neighbourhood of x.

In general, $\overline{F}$ and $\mathrm{lsc}(F)$ are different, but the two are always ordered.

Lemma 6.28. *Let $F : X \to \overline{\mathbb{R}}$. Then, $\overline{F} \geq \mathrm{lsc}(F)$.*

We now state a result ensuring that the lower semicontinuous envelope is, in a general topological space, the right object to consider, if we want to characterize in a variational way the asymptotic behavior of minimizing sequences.

Theorem 6.29. *Let $F : X \to \overline{\mathbb{R}}$ be a coercive functional. Then,* $\mathrm{lsc}(F)$ *is coercive and the following holds*

$$\inf\{ F(x) : x \in X \} = \min\{ \mathrm{lsc}(F)(x) : x \in X \}. \tag{6.9}$$

Moreover, every cluster point of a minimizing sequence for F us a minimum point for $\mathrm{lsc}(F)$*. Finally, if X satisfies the first axiom of countability, then every minimun point for* $\mathrm{lsc}(F)$ *is the limit of a minimizing sequence for F.*

Remark 6.30. Note that the difference with Theorem 6.10 is in the additional requirement needed to ensure that every minimum point of $\mathrm{lsc}(F)$ can be achieved as limit of a minimizing sequence for F.

Remark 6.31. The above result does not hold, in general, if we put $\overline{F}$ in place of $\mathrm{lsc}(F)$. This is the main difficulty in dealing with relaxation in general topological spaces: the object that we need, $\mathrm{lsc}(F)$, is difficult to work with, while the object is easier to identify, $\overline{F}$, is not the one we need.

We now wonder if there are conditions ensuring that the two notions of relaxation are equal. We have hope because (this is a general result in Functional Analysis), in a reflexive and separable Banach space (X, τ), for each *bounded* set $B \subset X$ there exists a metric $\mathrm{d} : B \times B \to [0, \infty)$ that metrizes the weak convergence on B. Therefore,given a functional $F : X \to \overline{\mathbb{R}}$, in the computation of the relaxed functional, we need to be able to restrict our attention to bounded sets. This is possible if F is coercive, as the following result shows.

Proposition 6.32. Let $(X, \|\cdot\|)$ be a normed vector space. Let τ be either

(i) The weak topology on X, and assume X^* separable;

(i) The weak* topology on X, and assume $X \subset Y'$, with Y a separable normed space.

Let $F : X \to \overline{\mathbb{R}}$ be a coercive functional with respect to the topology τ. Then, $\overline{F} = \mathrm{lsc}(F)$.

In the applications we will consider the space $X = L^p(\Omega; \mathbb{R}^M)$ endowed wit the weak topology if $p < \infty$, or the weak* topology if $p = \infty$. We thus specialize the above result to our case of interest.

Corollary 6.33. *Let $p \in (1, \infty]$, and let $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ be a coercive functional with respect to the weak topology of L^p if $p \in (1, \infty)$, or with respect to the weak* topology of L^∞ if $p = \infty$. Then,*

$$\overline{F} = \mathrm{lsc}(F),$$

where both relaxations are with respect to the above topologies.

Remark 6.34. Let $p \in (1, \infty)$, and let $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$. Using Proposition 5.4, we get that F is weakly coercive if and only if

$$F(u) \to \infty \qquad \text{as} \quad \|u\|_{L^p} \to \infty.$$

In the case $p = \infty$, we can ensure weak* coerciveness of F by imposing that

$$F(u) = +\infty \qquad \text{if} \quad \|u\|_{L^\infty} > R,$$

for some $R > 0$.

The above corollary does not hold for the weak topology of L^1, since L^1 is not a reflexive Banach space. Nevertheless, it is still possible to find a growth condition ensuring the equality between the two notions of relaxation.

Proposition 6.35. Let $F : L^1(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$, where $\Omega \subset \mathbb{R}^N$ is a measurable set, be such that

$$F(u) \geq \int_\Omega \varphi(|u|)\, dx,$$

for all $u \in L^1(\Omega; \mathbb{R}^M)$, where $\varphi : [0, \infty] \to [0, \infty]$ is an increasing function such that

$$\lim_{t\to\infty} \frac{\varphi(t)}{t} = \infty.$$

Then, $\overline{F} = \operatorname{lsc}(F)$, where both relaxations are with respect to the weak L^1 topology.

Remark 6.36. The difference between the growth condition needed in Corollary 6.33 for $p > 1$, and that needed in Proposition 6.35 for $p = 1$, is that in the latter we require a *rate of growth* of F at infinity, while in the former we just ask for it to blow up at infinity.

The technical reason for asking for a superlinear growth in Proposition 6.35 is to ensure that the relaxed functional is defined over L^1 functions. If the functional F has only linear growth, then the relaxed functional might required to be defined over measures. An example is the so called *area functional*: $F : W^{1,1}(\Omega) \to [0, +\infty]$ defined as

$$F(u) := \int_\Omega \sqrt{1 + |\nabla u|^2}\, dx.$$

Note that, since F only depends on the gradient, and not on the function u, we can think of it as defined over the subspace of L^p of gradients of functions. This functional has linear growth in ∇u, and it turns out that its relaxation has to be defined over functions whose gradient is a measure. These are the so called *functions of bounded variation*.

For the proofs of Proposition 6.32 and of Proposition 6.35 we refer to [8, Proposition 3.16, Proposition 3.18].

6.4 Laurentiev phenomenon and relaxation

Minimization problems are difficult to solve. This is why being able to find an approximate solution with numerical methods is a fundamental task. Classical methods in approximations (like the finite element method) use Lipschitz functions. Therefore, it is of high interested to answer the following question: consider a functional $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$, where $p \in [1, \infty)$, and let $\mathcal{A} \subset L^p(\Omega; \mathbb{R}^M)$. For instance, $\mathcal{A} = \operatorname{Lip}(\Omega; \mathbb{R}^M)$. Is is true that

$$\inf_{\mathcal{A}} F = \inf_X F?$$

Of course, the inequality

$$\inf_{\mathcal{A}} F \geq \inf_X F$$

holds, and we want to know if it can be strict. We first illustrate it with a finite dimensional example. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Then, it is clear that

$$\inf_{\mathbb{Q}} f > \inf_{\mathbb{R}} f.$$

Indeed, despite the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$, we cannot achieve the value $0 = \inf_{\mathbb{R}} f$ just by using rational points.

In 1926, Laurentiev studied such a phenomenon for integral functional, and in 1934 Maniá found a surprisingly easy example where the strict inequality holds. Let $F : W^{1,1}((0,1)) \to \overline{\mathbb{R}}$ be defined as

$$F(u) := \int_0^1 [u^3(x) - x]^2 (u'(x))^6 \, dx.$$

He proved that

$$\inf_{\mathrm{Lip}((0,1))} F > \min_{W^{1,1}((0,1))} F.$$

What is surprising is that the functional F seems very nice. Yet, it is not possible to approximate its minimum over $W^{1,1}$ with Lipschitz functions.

We now interpret the Lauretiev phenomenon by using the theory of relaxation. There is a natural way to include the constraint in the functional. Define $G : X \to \overline{\mathbb{R}}$ as

$$G(x) := \begin{cases} F(x) & \text{if } x \in \mathcal{A}, \\ +\infty & \text{else.} \end{cases}$$

Then, it holds that

$$\inf_{\mathcal{A}} F = \inf_{X} G.$$

Moreover, $x \in \mathcal{A}$ is a minimizer of F over $\mathcal{A}$ if and only if x is a minimizer of G over X. Thus, we can equivalently study the *unconstrained* minimization of G. What happens if we relax the functional G? Do we get back the original F or not? This is the essence of the Laurentiev phenomenon:it might happen that

$$\overline{G} \geq \overline{F},$$

with strict inequality for $u \in X \setminus \mathcal{A}$ and this is what causes the Laurentiev phenomenon. Some research has been done in order to identify the mismatch, as well as identifying conditions that prevent this phenomenon from happening. Moreover, delicate numerical techniques have been developed to simulate numerically minimizers for functionals exibiting the Laurentiev phenomenon.

6.5 Relaxation of integral functionals defined on L^p: strong topology

Strong topology is usually too strong! Namely, when we are given a functional, we need to choose the topology to use in order to either apply the Direct Method, or to relax the functional. In most cases, the functional $F : X \to \overline{\mathbb{R}}$ under consideration

Nevertheless, it is important to understand lower semicontinuity, continuity, and relaxation of integral functionals with respect to strong topologies. We start by investigating lower semicontinuity.

Proposition 6.37. Let $\Omega \subset \mathbb{R}^N$ be a measurable set, and $p \in [1, +\infty]$. Let $f : \Omega \times \mathbb{R}^M \to \mathbb{R}$ be such that

$$z \mapsto f(x, z)$$

is lower semicontinuous for a.e. $x \in \Omega$. Moreover, assume that

- If $p < \infty$, that
$$f(x, z) \geq a(x)|z|^p + b(x),$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $a, b \in L^1(\Omega)$ are non-negative functions;

- If $p = \infty$, that
$$f(x, z) \leq C,$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $C < +\infty$.

Then, the functional $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ defined as

$$F(u) := \int_\Omega f(x, u(x))\, dx$$

is sequentially lower semicontinuous with respect to the strong topology of $L^p(\Omega; \mathbb{R}^M)$.

Proof. Let $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega; \mathbb{R}^M)$ be such that $u_n \to u$, where $u \in L^p(\Omega; \mathbb{R}^M)$. Then, there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ such that

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

Thus, by assumption, we get that

$$f(x, u(x)) \leq \liminf_{k\to\infty} f(x, u_{n_k}(x)) \tag{6.10}$$

for a.e. $x \in \Omega$. Therefore, using Fatous's Lemma, we get

$$F(u) = \int_\Omega f(x, u(x))\, dx \leq \int_\Omega \liminf_{k\to\infty} f(x, u_{n_k}(x)) \leq \liminf_{k\to\infty} \int_\Omega f(x, u_{n_k}(x)) = \liminf_{k\to\infty} F(u_{n_k}).$$

Since the left-hand side is independent of the subsequence $(u_{n_k})_{k\in\mathbb{N}}$, we conclude that

$$F(u) \leq \liminf_{k\to\infty} F(u_n),$$

as desired. $\square$

Remark 6.38. The bounds on the integrand f are necessary in order to get lower semicontinuity of the functional F. Indeed, consider the function $f(x) := x^4$, and let

$$F(u) := \int_0^1 f(u)\, dx.$$

For each $n \in \mathbb{N} \setminus \{0\}$ define $u_n : (0, 1) \to \mathbb{R}$ as

$$u_n := n^{1/3} \mathbb{1}_{(0,1/n)}.$$

Then, $u_n \in L^2$, and $u_n \to u$ in L^2, where $u \equiv 0$, but

$$F(u) = \int_0^1 f(u)\, dx = 0 \neq +\infty = \lim_{n\to\infty} F(u_n) = \int_0^1 f(u_n)\, dx = \lim_{n\to\infty} n^{1/3}.$$

Thus, the functional F is not continuous with respect to the strong topology of L^2. Note though, that it is continuous with respect to the strong topology of L^p, for all $p \geq 4$.

Applying the above result to f and to $-f$ yields the continuity result for integral functionals with respect to the strong topology of L^p.

Proposition 6.39. Let $\Omega \subset \mathbb{R}^N$ be a measurable set, and $p \in [1, +\infty]$. Let $f : \Omega \times \mathbb{R}^M \to \mathbb{R}$ be a Borel function with

- If $p < \infty$, that
$$f(x, z) \geq a(x)|z|^p + b(x),$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $a, b \in L^1(\Omega)$;
- If $p = \infty$, that
$$f(x, z) \geq -C,$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $C > 0$.

Assume that
$$z \mapsto f(x, z)$$
is continuous for a.e. $x \in \Omega$. Then, the functional $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ defined as
$$F(u) := \int_\Omega f(x, u(x))\, dx,$$
is sequentially continuous with respect to the strong topology of $L^p(\Omega; \mathbb{R}^M)$.

As for the case of lower semicontinuity and continuity with respect to the weak topology, we can also characterize strongly lower semicontinuous and strongly continuous integral functionals.

Proposition 6.40. Let $p \in [1, \infty]$, and $\Omega \subset \mathbb{R}^N$ be a measurable set. Let $f : \mathbb{R}^M \to \overline{\mathbb{R}}$ be a Borel function with

- If $p < \infty$, that
$$f(x, z) \geq a(x)|z|^p + b(x),$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $a, b \in L^1(\Omega)$;
- If $p = \infty$, that
$$f(x, z) \geq -C,$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $C > 0$.

Define $F : L^p(\Omega;\mathbb{R}^M) \to \overline{\mathbb{R}}$ as

$$F(u) := \int_\Omega f(x, u(x))\, dx.$$

Then, F is lower semicontinuous with respect to the strong topology of L^p if and only if

$$z \mapsto f(x, z)$$

is lower semicontinuous for a.e. $x \in \Omega$.

Proposition 6.41. Let $p \in [1, \infty]$, and $\Omega \subset \mathbb{R}^N$ be a measurable set. Let $f : \mathbb{R}^M \to \overline{\mathbb{R}}$ be a Borel function with

- If $p < \infty$, that
$$|f(x, z)| \leq a(x)|z|^p + b(x),$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$, where $a, b \in L^1(\Omega)$;
- If $p = \infty$, that
$$|f(x, z)| \leq C,$$
for a.e. $x \in \Omega$, and all $z \in \mathbb{R}^M$.

Define $F : L^p(\Omega;\mathbb{R}^M) \to \overline{\mathbb{R}}$ as

$$F(u) := \int_\Omega f(x, u(x))\, dx.$$

Then, F is continuous with respect to the strong topology of L^p if and only if

$$z \mapsto f(x, z)$$

is continuous for a.e. $x \in \Omega$.

Remark 6.42. Continuity with respect to the weak topology is a more restrictive condition than continuity with respect to the strong topology, since a strongly converging sequence is weakly converging, but the opposite is not true. This reflects on the fact that in Corollary 5.13 the integrand f is required to be linear (in particular, it is continuous), while in Proposition 6.41 we only ask for continuity of the integrand.

6.6 Relaxation of integral functionals defined on L^p: weak topology

We now want to compute explicitly the relaxation of integral functionals of the form

$$F(u) := \int_\Omega f(u)\; dx, \qquad \text{and} \qquad F(u) := \int_\Omega f(\nabla u)\; dx,$$

with respect to the weak topology of $L^p(\Omega;\mathbb{R}^M)$ (and weak*-$L^\infty(\Omega;\mathbb{R}^M)$), where $\Omega \subset \mathbb{R}^N$ is an open set. The reason why we consider both is that, in the latter case, there is a difference in the case $\min\{N, M\} = 1$ and $\min\{N, M\} > 1$. We will only focus, as we did in Theorem 5.12, on the case $N = M = 1$. On the other hand, for the former case, the dimensions N and M play no role, but a deep result on weak convergence in L^p is needed (the so called Riemann-Lebesgue Lemma, see Theorem 6.61).

In particular, we will prove what is called an *integral representation result*: the relaxed functional is still an integral functional. Moreover, we will be able to give an explicit form of the integrand defining the relaxed functional.

Before digging into details, let's make some heuristics: if we have an integral functional $F : L^p(\Omega;\mathbb{R}^M) \to \overline{\mathbb{R}}$ of the form

$$F(u) := \int_\Omega f(\nabla u)\, dx,$$

we know that, if f is convex, then F is sequentially weakly lower semicontinuous in L^p for $p < \infty$, or weakly*-L^∞ for $p = \infty$ (see Theorem 5.7). We also know that the converse is true if $M = 1$ (see Theorem 5.12; this can be extended to the case $\min\{N, M\} = 1$). We might therefore wonder if the relaxation of the functional F, namely finding the greatest lower semicontinuous functional not greater than F has anything to do with the integral functional whose integrand is the greatest convex function not greater than f.

6.6.1 Convex functions

In this section we investigate convex functions and the operation of convexification, namely, given a function f, finding the greatest convex function maximized by f.

Definition 6.43. Let $f : \mathbb{R}^M \to \mathbb{R}$. We say that f is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y),$$

for all $x, y \in \mathbb{R}^N$, and all $\lambda \in [0,1]$.

Remark 6.44. Here we only consider the case of finite valued convex functions. The definition can be extended also to the case of maps in $\mathbb{R}$, and the inequality has to hold when the right-hand side is well-defined.

Convex functions are supremum of affine functions, as the following result shows.

Proposition 6.45. Let $f : \mathbb{R}^M \to \mathbb{R}$. Then, f is convex if and only if f is the supremum of affine functions. Namely, f is convex if and only if

$$f = \sup\left\{ \varphi : \mathbb{R}^M \to \mathbb{R} : \varphi \leq f,\ \varphi \text{ affine} \right\}.$$

Moreover, if f is convex, it is possible to find $(a_i)_{i\in\mathbb{N}} \subset \mathbb{R}^M$ and $(b_i) \subset \mathbb{R}$ such that

$$f(z) = \sup_{i\in\mathbb{N}} [a_i \cdot z + b_i]$$

for all $z \in \mathbb{R}^M$.

The proof is left as an exercise to the reader (and it will be given as an homework).

Remark 6.46. The second part of the previous result follows from a general fact, know as *Lindelöf's Theorem*: let $\Omega \subset \mathbb{R}^M$ be an open set, and consider a family of functions $\mathcal{G} \subset C^0(\Omega)$. Define

$$f := \sup_{g\in\mathcal{G}} g.$$

Then, there exists a *countable* subfamily $\mathcal{G}_1 \subset \mathcal{G}$ such that

$$f = \sup_{g \in \mathcal{G}_1} g.$$

Namely, the supremum over a generic family of continuous functions is the supremum over a countable family of continuous functions.

Finally, note the the countable families $(a_i)_{i\in\mathbb{N}} \subset \mathbb{R}^M$ and $(b_i) \subset \mathbb{R}$ *cannot* be the same for all convex functions $f : \mathbb{R}^M \to \mathbb{R}$. Indeed, if $a \in \mathbb{R}^M$ is such that $a \neq a_i$ for all $i \in \mathbb{N}$, then there is no $i \in \mathbb{N}$ such that $a \cdot z \geq a_i \cdot z + b_i$ for all $z \in \mathbb{R}^M$.

We now prove that convex functions are locally Lipschitz functions.

Definition 6.47. Let $f : \mathbb{R}^N \to \mathbb{R}$. We say that f is *locally Lipschitz* if, for every compact set $K \subset \mathbb{R}^N$, it holds

$$\sup_{x \neq y \in K} \frac{|f(x) - f(y)|}{|x - y|} < +\infty.$$

In such a case, the above supremum is called the *Lipschitz constant of* f, and denoted by $\mathrm{Lip}(f; K)$. Finally, if $\mathrm{Lip}(f; K)$ is uniformly bounded, for all compact subsets $K \subset \mathbb{R}^N$, we say that f is *Lipschitz*.

Proposition 6.48. Let $g : \mathbb{R}^M \to \mathbb{R}$ be convex. Then, g is locally Lipschitz. In particular, for each $R > 0$ we have

$$\frac{|g(x) - g(y)|}{|y - x|} \leq \frac{|M - m|}{R},$$

for all $x, y \in B(\overline{z}, R)$, where $m, M \in \mathbb{R}$ are such that $m \leq g(z) \leq M$ for all $z \in B(\overline{z}, 2R)$.

Proof. Step 1. We first consider the one dimensional case $M = 1$. Fix $\overline{z} \in \mathbb{R}$, $R > 0$, and $y, x \in B(\overline{z}, R)$. Without loss of generality, we can assume $y > x$ and $g(y) \geq g(x)$. Since g is convex, we have that slopes are increasing (actually, this is a characterization of being convex). Namely,

$$\frac{g(x_2) - g(x_1)}{x_2 - x_1} \leq \frac{g(x_4) - g(x_3)}{x_4 - x_3},$$

for all $x_1 < x_2 < x_3 < x_4$. This implies that

$$\frac{g(y) - g(x)}{y - x} \leq \frac{g(\overline{z} + 2R) - g(\overline{z} + R)}{R} \leq \frac{|M - m|}{R},$$

where the last inequality follows from the definition of m and M. This, combined with the other cases (that are treated similarly), gives the result for $M = 1$.

Step 2. Let us now consider the general case $M > 1$. Fix $\overline{z} \in \mathbb{R}^M$, $R > 0$, and $y, x \in B(\overline{z}, R)$. Consider the function $\widetilde{g} : \mathbb{R} \to \mathbb{R}$ given by

$$\widetilde{g}(t) := g(tx + (1 - t)y).$$

Then $\widetilde{g}$ is convex. Let $t_1, t_2 > 0$ be such that

$$P := t_1 x + (1 - t_1)y \in \partial B(\overline{z}, R), \qquad Q := t_2 x + (1 - t_2)y \in \partial B(\overline{z}, 2R).$$

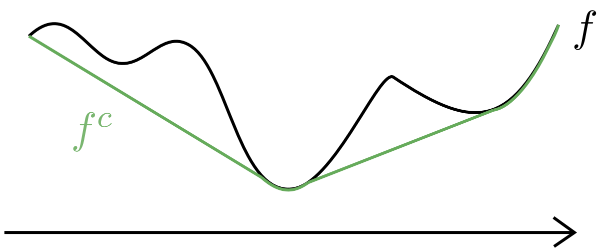


Figure 6.2: The function f and its convex envelope f^c.

By using the definition of $\widetilde{g}$, and of the points t_1 and t_2, from Step 1 we get that

$$|g(y) - g(x)| = |\widetilde{g}(1) - \widetilde{g}(0)| \leq \frac{|M-m|}{t_2 - t_1} \leq \frac{|M-m|}{R}|y-x|,$$

where in the last step we used the fact that $(t_2 - t_1)|y - x| = |P - Q| \geq R$. This gives the desired result. □

Remark 6.49. Two deep results in Analysis ensure that Lipschitz functions are differentiable for almost every point. The case $M = 1$ uses sophisticated techniques in Measure Theory (such as Absolutely Continuous functions and differentiation of measures), and the generalization to the higher dimensional case is known as *Rademacher's Theorem*.

Corollary 6.50. *A convex function is continuous.*

Remark 6.51. Proposition 6.48 can be extended to the case where g is allowed to take the value $+\infty$. In that case, Lipschitzianity holds locally in any open set contained in $\{g < +\infty\}$.

Remark 6.52. The above regularity result is peculiar to finite dimensional vector spaces. Indeed, if X is an infinte dimensional vector space, and $g : X \to \overline{\mathbb{R}}$ is convex, then g could be discontinuous at every point. This comes from the fact that in infinite dimension, we can have a dense subspace where $g = +\infty$.

Remark 6.53. Convex functions enjoy higher order regularity properties, namely they admit a suitable weak notion of second derivative for almost every point. This deep result is known as *Alexandrov's theorem*.

We now investigate the convex envelope of a function.

Definition 6.54. Let $f : \mathbb{R}^M \to \overline{\mathbb{R}}$. We define $f^c : \mathbb{R}^M \to \overline{\mathbb{R}}$, the *convex envelope* of f, as

$$f^c(z) := \sup\{g(z) : g : \mathbb{R}^M \to \overline{\mathbb{R}} \text{ convex }, \ g \leq f\},$$

for $z \in \mathbb{R}^M$.

Remark 6.55. The *convexification* of a function $f : \mathbb{R}^M \to \overline{\mathbb{R}}$, namely taking its convex envelope, is a *nonlocal* operation (see Figure 6.2). Namely, if we change the function f in a ball $B(x, r)$, the convex envelope might change also outside $B(x, r)$. For instance, let $f, g : \mathbb{R} \to \mathbb{R}$ be defined as $f \equiv 1$, and $g := 1 - \mathbb{1}_{[0,1]}$. Then, $f^c = f$, but $g^c \neq g$. This is what makes this operation difficult to compute in practice (other than examples in one dimensions with regular functions, where the convexification can be deduced just by looking at the graph of f).

The following property of the convex envelope follow directly from the definition.

Lemma 6.56. *Let $f : \mathbb{R}^M \to \overline{\mathbb{R}}$. Then, f^c is convex.*

By using the characterization of convex functions as supremum of affine functions (see Proposition 6.45), it is possible to prove the following result.

Lemma 6.57. *Let $f : \mathbb{R}^M \to \overline{\mathbb{R}}$. Then,*

$$f^c(z) = \sup\{g(z) : g : \mathbb{R}^M \to \mathbb{R} \text{ affine }, \, g \leq f\}.$$

In particular, $f^c > -\infty$ if and only if there exist $a \in \mathbb{R}^M$ and $b \in \mathbb{R}$ such that $f(z) \geq a \cdot z + b$, for all $z \in \mathbb{R}^M$.

We now prove a representation formula for the convex envelope, that allows us to recover the value of f^c by using convex combinations of values of f. This will be crucial in the limsup inequality of the proof of the integral representation result.

The heuristic idea is the following: consider the case $N = 1$. Given an affine function $\varphi : \mathbb{R}^M \to \mathbb{R}$ with $\varphi \leq f$, we can *move it up* until it touches the graph of f. This can happen at one point, or more. In the latter case, let $(x, f(x))$ and $(y, f(y))$ be two points of contact. Then, the value of f^c in any point z of the form $\lambda x + (1 - \lambda)y$, with $\lambda \in [0, 1]$ is such that

$$f^c(z) \leq \lambda f(x) + (1 - \lambda) f(y) \leq f(z).$$

By using this idea, and taking into consideration that, in higher dimension, convex combinations of segments are not sufficient, we have the following result.

Proposition 6.58. Let $f : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$. Then,

$$f^c(z) = \inf\left\{ \sum_{i=1}^k \lambda_i f(z_i) \; : \; k \in \mathbb{N}, \, \sum_{i=1}^k \lambda_i = 1, \, \lambda_i \geq 0, \, z_1, \ldots, z_k \in \mathbb{R}^M, \, \sum_{i=1}^k \lambda_i z_i = z \right\},$$

for all $z \in \mathbb{R}^M$

Proof. Without loss of generality, we can assume that there exist $a \in \mathbb{R}^M$ and $b \in \mathbb{R}$ such that $f(z) \geq a \cdot z + b$, for all $z \in \mathbb{R}^M$. In particular, $f^c > -\infty$ (see Lemma 6.57). Let $g : \mathbb{R}^M \to \overline{\mathbb{R}}$ be defined as

$$g(z) := \left\{ \sum_{i=1}^k \lambda_i f(z_i) \; : \; k \in \mathbb{N}, \, \sum_{i=1}^k \lambda_i = 1, \, \lambda_i \geq 0, \, z_1, \ldots, z_k \in \mathbb{R}^M, \, \sum_{i=1}^k \lambda_i z_i = z \right\}.$$

We claim that g is convex.

We first show how to conclude once the claim is proved. By definition of g, we have that $g \leq f$. Thus, g is an admissible competitor in the definition of f^c, and therefore $f^c \geq g$. To prove the opposite inequality, since $f^c \leq f$, for all $z \in \mathbb{R}^M$, we have that

$$\begin{aligned} g(z) &\geq \inf\left\{ \sum_{i=1}^{k} \lambda_i f^c(z_i) \,:\, k \in \mathbb{N},\, \sum_{i=1}^{k} \lambda_i = 1,\, \lambda_i \geq 0,\, z_1, \dots, z_k \in \mathbb{R}^M,\, \sum_{i=1}^{k} \lambda_i z_i = z \right\} \\ &\geq f^c(z), \end{aligned}$$

where in the last step we used the fact that f^c is convex (see Lemma 6.57), and thus the minimum is attained for $k = 1$ and $z_1 = z$ (this is what the inequality defining convexity is about).

We are thus left with proving that g is convex. Let $z_1,\ z_2 \in \mathbb{R}^M$, $t \in (0, 1)$, and set

$$z := tz_1 + (1 - t)z_2. \tag{6.11}$$

Assume without loss of generality that $g(z_1), g(z_2) < \infty$ otherwise there is nothing to prove. Fix $\varepsilon > 0$ and let $k_1, k_2 \in \mathbb{N}$, $y_1^1, \dots, y_{k_1}^1 \in \mathbb{R}^M$, $y_1^2, \dots, y_{k_2}^2 \in \mathbb{R}^M$, $\lambda_1^1, \dots, \lambda_k^1 \geq 0$, $\lambda_1^2, \dots, \lambda_k^2 \geq 0$ with

$$\sum_{i=1}^{k_1} \lambda_i^1 = \sum_{i=1}^{k_2} \lambda_i^2 = 1, \qquad \sum_{i=1}^{k_1} \lambda_i^1 y_i^1 = z_1, \qquad \sum_{i=1}^{k_2} \lambda_i^2 y_i^2 = z_2, \tag{6.12}$$

such that

$$g(z_1) + \varepsilon \geq \sum_{i=1}^{k_1} \lambda_i^1 f(y_i^1), \qquad g(z_2) + \varepsilon \geq \sum_{i=1}^{k_2} \lambda_i^2 f(y_i^2). \tag{6.13}$$

Then, using (6.11), and (6.12), we get

$$z = t \sum_{i=1}^{k_1} \lambda_i^1 y_i^1 + (1 - t) \sum_{i=1}^{k_2} \lambda_i^2 y_i^2 = \sum_{i=1}^{k_1} [t\lambda_i^1] y_i^1 + \sum_{i=1}^{k_2} [(1 - t)\lambda_i^2] y_i^2,$$

and

$$\sum_{i=1}^{k_1} [t\lambda_i^1] + \sum_{i=1}^{k_2} [(1 - t)\lambda_i^2] = 1.$$

Thus, the definition of g yields that

$$\begin{aligned} g(z) &\leq \sum_{i=1}^{k_1} [t\lambda_i^1] f(y_i^1) + \sum_{i=1}^{k_2} [(1 - t)\lambda_i^2] f(y_i^2) \\ &\leq t[g(z_1) + \varepsilon] + (1 - t)[g(z_2) + \varepsilon] \\ &= tg(z_1) + (1 - t)g(z_2) + \varepsilon, \end{aligned}$$

where in the previous to last step we used (6.13). The arbitrariness of $\varepsilon > 0$ allows to conclude. □

Remark 6.59. The infimum in the above representation formula is not always attained. For instance, consider $f(x) := e^{-x^2}$. Then, $f^c \equiv 0$, and the infimum in the above representation formula is never attained.

Remark 6.60. The above result can be refined as follows: if $f : \mathbb{R}^M \to \mathbb{R}$, then

$$f^c(z) = \inf \left\{ \sum_{i=1}^{M+1} \lambda_i f(z_i) \,:\, \sum_{i=1}^{M+1} \lambda_i = 1,\, \lambda_i \geq 0,\, z_1, \ldots, z_{M+1} \in \mathbb{R}^M,\, \sum_{i=1}^{M+1} \lambda_i z_i = z \right\}, \tag{6.14}$$

that is, in dimension M we only need to take convex combinations of at most $M + 1$ points. The proof of (6.14) result is based on the *Carathéodory Theorem* for the convex envelope of a set (stating basically the same result for the convex hull of a set), together with the fact that, for each $z \in \mathbb{R}^M$, it holds

$$f^c(z) = \inf \{ t \in \mathbb{R} : (z, t) \in \mathrm{conv}[\mathrm{epi}(f)] \},$$

where, for a set $E \subset \mathbb{R}^{M+1}$, the *convex hull* $\mathrm{conv}(E)$ of E is the intersection of all of the convex sets containing E.

Note that $M + 1$ is the *optimal number* of points, in the sense that convex combinations of at most M points are not sufficient to get $f^c(z)$. As an example (not trivially even or $M = 1$), take $z_1, z_2, z_3 \in \mathbb{R}^2$ such that $z_2 - z_1$ and $z_3 - z_1$ are linearly independent, and define $f : \mathbb{R}^2 \to \mathbb{R}$ as

$$f(z_1) = f(z_2) = f(z_3) = 0, \qquad f(z) = 1 \text{ for all other points } z \in \mathbb{R}^3.$$

Then $f^c(z) = 0$ if z is a convex combination of z_1, z_2, z_3 (namely in the triangle whose vertexes are z_1, z_2, z_3), and $f^c(z) = 1$ otherwise. But if we were allowed only to take convex combinations of two points, we would get a function $g : \mathbb{R}^2 \to \mathbb{R}$ that is 1 everywhere but on the *segments* joining the points z_1, z_2, and z_3, but not on the interior of the triangle.

6.6.2 Periodic functions and weak convergence in L^p

We now prove an important result that justifies the intuitive fact that the macroscopic behavior of a repeated (periodic) pattern is given by the average of the pattern in a cell. This results was firstly proved in the context of Fourier series, in order to get the convergence of the series, and later extended to the more general situation we consider in here.

Theorem 6.61 (Riemann-Lebesgue's lemma). *Let $p \in [1, \infty]$, $\Omega \subset \mathbb{R}^N$ be an open bounded set, and $Q \subset \mathbb{R}^N$ be an open cube. Let $f \in L^p(Q; \mathbb{R}^M)$. Extend f to the whole $\mathbb{R}^N$ in a Q-periodic way, namely define (with an abuse of notation)*

$$f(z) \coloneqq f(y),$$

where $z \in \mathbb{R}^M$ is written as $z = y + q$ for some $y \in Q$ and $q \in \mathbb{Z}^N$. For $n \in \mathbb{N}$ define $f_n \in L^p(\Omega; \mathbb{R}^M)$ as

$$f_n(x) \coloneqq f(nx).$$

Then,

(i) f_n converges to $\overline{f}$ weakly in $L^p(\Omega; \mathbb{R}^M)$, if $p \in [1, \infty)$;

(ii) f_n converges to $\overline{f}$ weakly in $L^\infty(\Omega; \mathbb{R}^M)$, if $p = \infty$,*

where $\overline{f}$ denote the average of f in Q.

Proof. First of all, note that since Ω is bounded, we have that $f \in L^1(\Omega)$ (see Remark 3.4). We divide the proof in several steps.

Step 0. Without loss of generality, we can assume that

(i) $Q := (-1/2, 1/2)^N$;

(ii) $M = 1$;

(iii) $\overline{f} = 0$.

Indeed, a change of variable allows to prove the result for a general cube, once that for the unitary cube is established. As for (ii), this follows from the fact that weak convergence acts component-wise. Finally, assume that the result if proved under the assumption that the average of the function in a periodicity cell is zero. Then, given $f \in L^p(Q)$, consider $g := f = \overline{f}$. Then, g has zero average, and thus the function g_n, where $g_n(x) := g(nx)$, converges weakly in L^p (or weakly* in L^∞) to zero. This means that

$$\int_\Omega f_n\varphi \, dx - \overline{f} \int_\Omega \varphi \, dx \to 0,$$

for every $\varphi \in L^{p'}(Q)$ if $p < \infty$, or for every $\varphi \in L^1(Q)$ if $p = \infty$. This proves that, without loss of generality, we can reduce to the case where (i), (ii), and (iii) hold.

Step 1. Assume $\Omega = \frac{1}{k}Q$, for some $k \in \mathbb{N} \setminus \{0\}$. Let $\varphi \in C_c(\Omega)$. Let $\{z_i^n\}_{i\in\mathbb{N}}$ be an enumeration of $\frac{1}{n}\mathbb{Z}^N$. For each $n \in \mathbb{N}$, consider the grid of cubes $\{Q_i^n\}_{i\in\mathbb{N}}$, where $Q_i^n := z_i^n + \frac{1}{n}Q$. Let

$$I_n := \{i \in \mathbb{N} : Q_i^n \subset Q\}.$$

Note that a change of variables gives

$$\int_{Q_i^n} f_n(x) \, dx = \frac{1}{n^N} \int_Q f(x) \, dx, \tag{6.15}$$

for all $i \in I_n$, and all $n \in \mathbb{N}$. Since φ is a continuous function with compact support, it is uniformly continuous. Namely, for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\varphi(x) - \varphi(y)| < \varepsilon,$$

for all $x, y \in \Omega$ with $|x - y| < \delta$. Fix $\varepsilon > 0$ and let $\delta > 0$ given as above. Let $\bar{n} \in \mathbb{N}$ be such that for all $n \geq \bar{n}$ it holds $n\sqrt{N} < \delta$. Note that

$$|x - z_i^n| < \delta,$$

for all $x \in Q_i^n$, for all $i \in I_n$, and for all $n \geq \bar{n}$. Up to increasing the value of $\bar{n}$, we can also assume

$\bar{n} \geq k$. Then, for $n \geq \bar{n}$, we get

$$
\begin{aligned}
\left| \int_\Omega f_n(x)\varphi(x)\, dx \right| &= \left| \sum_{i \in I_n} \int_{Q_i^n} f_n(x)\varphi(x)\, dx \right| \\
&\leq \left| \sum_{i \in I_n} \int_{Q_i^n} f_n(x)[\varphi(x) - \varphi(z_i^n)]\, dx \right| + \left| \sum_{i \in I_n} \int_{Q_i^n} \varphi(z_i^n) f_n(x)\, dx \right| \\
&\leq \varepsilon \sum_{i \in I_n} \int_{Q_i^n} |f_n(x)|\, dx + \left| \sum_{i \in I_n} \varphi(z_i^n) \int_{Q_i^n} f_n(x)\, dx \right| \\
&\leq \varepsilon \|f\|_{L^1(Q)} + \sum_{i \in I_n} \frac{|\varphi(z_i^n)|}{n^N} \int_Q f(x)\, dx \\
&= \varepsilon \|f\|_{L^1(Q)},
\end{aligned}
$$

where the last step follows from our assumption that $\overline{f} = 0$, while in the previous to last step we used (6.15) together with the fact that $n \geq \bar{n} \geq k$. Thus, by the arbitrariness of $\varepsilon > 0$, we conclude that

$$\lim_{n \to \infty} \int_Q f_n \varphi\, dx = 0,$$

for all $\varphi \in C_c(\Omega)$.

Step 2. Assume $p \in (1, \infty]$ and let $g \in L^{p'}(\Omega)$, with $\Omega = Q$. Fix $\varepsilon > 0$. Then, thanks to Theorem 3.8, it is possible to find $\varphi \in C_c(\Omega)$ such that

$$\|g - \varphi\|_{L^{p'}(\Omega)} < \varepsilon. \tag{6.16}$$

Write

$$\int_\Omega f_n g\, dx = \int_\Omega f_n(g - \varphi)\, dx + \int_\Omega f_n \varphi\, dx. \tag{6.17}$$

By Step 1, we know that

$$\lim_{n \to \infty} \int_\Omega f_n \varphi\, dx = 0. \tag{6.18}$$

Using Hölder inequality (see Theorem 3.3) we can estimate

$$\left| \int_\Omega f_n(g - \varphi)\, dx \right| \leq \|g - \varphi\|_{L^{p'}(\Omega)} \|f_n\|_{L^p(\Omega)} \leq \varepsilon \sup_{n \in \mathbb{N}} \|f_n\|_{L^p(\Omega)}, \tag{6.19}$$

where last step follows from (6.16). We now claim that

$$\sup_{n \in \mathbb{N}} \|f_n\|_{L^p(\Omega)} < \infty. \tag{6.20}$$

Indeed, if $p = \infty$, we simply gets that

$$\|f_n\|_{L^\infty(\Omega)} = \|f\|_{L^\infty(\Omega)}$$

for all $n \in \mathbb{N}$. Assume now $p <= \infty$. Since Ω is bounded, there exists $k \in \mathbb{N}$ such that $\Omega \subset kQ$. Therefore, by using (6.15)e, we get that

$$\|f_n\|^p_{L^p(\Omega)} \leq k^N \|f\|_{L^p(Q)}.$$

Therefore, by using (6.17), (6.18), (6.19), and (6.20), we get

$$\lim_{n\to\infty} \left| \int_\Omega f_n g \, dx \right| \leq C\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we get the desired result.

Step 3. Now, assume $\Omega \subset \mathbb{R}^N$ to be a generic bounded open set, and let $\varphi \in C_c(\Omega)$. Then, it is possible to find $k \in \mathbb{N}$ such that it is possible to cover the support of φ with the closure of a finite of non-overlapping translations of the cube $\frac{1}{k}Q$, each of which lies inside Ω. Then, the result follows by Step 1 on each cube, due to the linearity of the integral.

Step 4. Assume $p = 1$. For $t > 0$ consider the truncation operator $T_t : L^1(\Omega) \to L^\infty(\Omega)$ defined as

$$T_t u(x) := \begin{cases} u(x) & \text{if } |u(x)| \leq t, \\ t & \text{if } u(x) \geq t, \\ -t & \text{if } u(x) \leq -t. \end{cases}$$

Let $g \in L^\infty(\Omega)$. Then we write

$$\int_\Omega f_n g \, dx = \int_\Omega (T_t f_n) g \, dx + \int_\Omega (f_n - T_t f_n) g \, dx = \int_\Omega (T_t f)_n g \, dx + \int_\Omega (f_n - T_t f_n) g \, dx. \quad (6.21)$$

We now estimate the two terms on the left-hand side separately. For the first term, since $T_t f \in L^\infty(Q)$ and $g \in L^1(\Omega)$, we can apply the previous step to conclude that

$$\lim_{n\to\infty} \int_\Omega (T_t f)_n g \, dx = 0. \quad (6.22)$$

In order to estimate the second term, note that

$$\left| \int_\Omega (f_n - T_t f_n) g \, dx \right| \leq \|g\|_{L^\infty(\Omega)} \|f_n - T_t f_n\|_{L^1(\Omega)}. \quad (6.23)$$

By using Lebesgue Dominated Convergence Theorem, we get that

$$\lim_{t\to\infty} \|f_n - T_t f_n\|_{L^1(\Omega)} = 0. \quad (6.24)$$

Thus, the desired result follows from (6.21), (6.22), (6.23), and (6.24). □

Remark 6.62. Note that the result holds for *any* $p \in [1, +\infty]$. Moreover, note that the same result holds for any periodicity cell Q, not necessarily a cube.

6.6.3 Integral representation formula

We are now in position to prove the desired integral representation result for the sequential relaxation of integral functionals on L^p with respect to the weak topology (see the discussion before Proposition 6.32).

Theorem 6.63. *Let $p \in (1, \infty)$, and $\Omega \subset \mathbb{R}^N$ be a bounded open set. Let $f : \mathbb{R}^M \to \mathbb{R}$ be a Borel function such that*

$$|f(z)| \leq C(|z|^p + 1),$$

for all $z \in \mathbb{R}^M$, where $C > 0$. Consider the functional $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ given by

$$F(u) := \int_\Omega f(u(x))\, dx.$$

Then, its sequential lower semi-continuous envelope $\overline{F} : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ with respect to the weak topology of L^p satisfies

$$\overline{F}(u) = \int_\Omega f^c(u(x))\, dx,$$

for all $u \in L^p(\Omega; \mathbb{R}^M)$.

Proof. Set

$$H(u) := \int_\Omega f^c(u)\, dx,$$

for $u \in L^p(\Omega; \mathbb{R}^M)$. We will prove that H satisfies (i) and (ii) of Proposition 6.11, namely the liminf and the limsup inequality.

Part 1: Liminf inequality. Since $f^c \leq f$, we get that

$$F(v) \geq H(v),$$

for all $v \in L^p(\Omega; \mathbb{R}^M)$. Thus, if $u \in L^p(\Omega; \mathbb{R}^M)$ and $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega; \mathbb{R}^M)$ is such that u_n converges to u weakly in L^p, then

$$\liminf_{n\to\infty} F(u_n) \geq \liminf_{n\to\infty} H(u_n) \geq H(u),$$

where in the last step we used the fact that H is lower semicontinuous (this can be proved by using a similar strategy as that implemented in the proof of Theorem 5.7).

Part 2: Limsup inequality. Let $u \in L^p(\Omega; \mathbb{R}^M)$. We need to build a sequence $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega; \mathbb{R}^M)$ with $u_n \rightharpoonup u$ such that

$$\limsup_{n\to\infty} F(u_n) \leq H(u).$$

In order to prove the limsup inequality, we will use the following strategy:
Step 1. We prove that it suffices to consider the case of piecewise constant function u;
Step 2. We construct a sequence of piecewise constant functions converging strongly to u;
Step 3. We construct the recovery sequence for a piecewise constant function.
We now proceed with this strategy.

Step 1: Reduction to simple functions. We claim that it is sufficient to construct a recovery sequence for piecewise constant functions.

First of all, we notice that

$$|f^c| \le C(|z|^p + 1),$$

for all $z \in \mathbb{R}^M$. Then, thanks to Proposition 6.39, we get that H is continuous with respect to the strong topology of L^p.

Let $u \in L^p(\Omega;\mathbb{R}^M)$. Let $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R}^M)$ be a sequence of piecewise constant functions converging strongly in L^p to u. Assume that

$$\overline{F}(u_n) = H(u_n), \tag{6.25}$$

for all $n \in \mathbb{N}$. Then, using the fact that $\overline{F} = \mathrm{lsc}(F)$, and thus it is lower semicontinuous, we get that

$$\overline{F}(u) \le \liminf_{n\to\infty} \overline{F}(u_n) = \liminf_{n\to\infty} H(u_n) = H(u),$$

where in the last step we used the fact that H is continuous with respect to the strong topology of L^p. Thus, we only need to prove that, given $u \in L^p(\Omega;\mathbb{R}^M)$ there exists a sequence of piecewise constant functions converging strongly in L^p to u, and that it is possible to construct a recovery sequence for every piecewise constant function.

Step 2: Construction of approximating simple functions. Let $u \in L^p(\Omega;\mathbb{R}^M)$. We now construct a sequence of piecewise constant functions $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R}^M)$ converging strongly in L^p to u.

Let $(v_n)_{n\in\mathbb{N}} \in L^p(\Omega;\mathbb{R}^M) \cap C_c(\Omega;\mathbb{R}^M)$ be the sequence provided by Theorem 3.8. In particular, $v_n \to u$ strongly in L^p. Fix $n \in \mathbb{N} \setminus \{0\}$. Since v_n is a continuous function on a compact set, its image is compact. Fix $\varepsilon > 0$. Using the uniformly continuity of v_n, and of f^c on the image of v_n, there exists $\delta_n > 0$ such that

$$|v_n(x) - v_n(y)| < \varepsilon, \qquad |f^c(v_n(x)) - f^c(v_n(y))| < \varepsilon \tag{6.26}$$

for all $x, y \in \Omega$ with $|x - y| < \delta_n$. Let $(Q_i^n)_{i\in I_n}$ be a grid of cubes of diameter less than δ_n, where I_n is the set of indexes identifying the cubes with non empty intersection with Ω. Note that since Ω is bounded, I_n is finite for every $n \in \mathbb{N}$. Denote by z_i^n the center of the cube Q_i^n. Define

$$u_n(x) := v_n(z_i^n),$$

for $x \in Q_i^n$, and for all $i \in I_n$. We now claim that $u_n \to u$ as $n \to \infty$. Indeed, by triangle inequality, we get that

$$\|u_n - u\|_{L^p} \le \|u_n - v_n\|_{L^p} + \|v_n - u\|_{L^p}.$$

We then just have to estimate the first term on the right-hand side, since the last term vanishes as $n \to \infty$. We get that

$$\begin{aligned}
\int_\Omega |u_n(x) - v_n(x)|^p \, dx &= \sum_{i\in i_n} \int_{Q_i^n\cap\Omega} |u_n(x) - v_n(x)|^p \, dx \\
&= \sum_{i\in i_n} \int_{Q_i^n\cap\Omega} |v_n(z_i^n) - v_n(x)|^p \, dx \\
&\le \varepsilon^p \sum_{i\in i_n} |Q_i^n \cap \Omega| \\
&= \varepsilon^p |\Omega|,
\end{aligned}$$

where in the last inequality we used the first property in (6.26). This proves that $u_n \to u$ strongly in L^p.

Step 3: Recovery sequence for simple functions. We are thus left with construction a recovery sequence for simple functions. The functions we have to consider are constant in the open set $\Omega \cap C$, where $C \subset \mathbb{R}^N$ is a open cube. Thus, consider a constant function u in $\Omega \cap C$, say $u \equiv z$, for some $z \in \mathbb{R}^M$. Fix $\varepsilon > 0$. By using the characterization of f^c given by Proposition 6.58, it is possible to find $k \in \mathbb{N}$, $\lambda_1, \dots, \lambda_k \geq 0$ and $z_1, \dots, z_k \in \mathbb{R}^M$ with

$$\sum_{i=1}^{k} \lambda_i = 1, \qquad \sum_{i=1}^{k} \lambda_i z_i = z, \tag{6.27}$$

such that

$$f^c(z) + \varepsilon \geq \sum_{i=1}^{k} \lambda_i f(z_i). \tag{6.28}$$

Define $g : C \to \mathbb{R}^M$ as

$$g(x) := \sum_{i=1}^{k} z_i \mathbb{1}_{A_i}(x),$$

where $(A)_{i=1}^k \subset C$ is a measurable partition of C with $|A_i| = \lambda_i |C|$. Extend g to the whole $\mathbb{R}^N$ in a C-periodic way, and define $u_n : \mathbb{R}^N \to \mathbb{R}^M$ as

$$u_n(x) := g(nx).$$

The Riemann-Lebesgue lemma (see Theorem 6.61), together with (6.27), yields that u_n converges to u weakly in $L^p(\Omega; \mathbb{R}^M)$. Moreover, similar computations as those used in the proof of Theorem 6.61, allow to obtain that

$$\lim_{n\to\infty} F(v_n) = |\Omega| \sum_{i=1}^{k} \lambda_i f(z_i) \leq |\Omega| \left[f^c(z) + \varepsilon \right],$$

where in the last step we used (6.28). We thus conclude by using the arbitrariness of $\varepsilon > 0$. □

Remark 6.64. In the previous proof we employed a lot of interesting techniques. Let's highlight them. First of all, it seems that the liminf inequality was deduced easily. Of course, this is because we already had a candidate for the relaxed functional. Usually, the liminf inequality is used to *guess* possible functionals that could be the relaxation, and then use the limsup inequality to confirm that guess.

The limsup inequality was obtained by approximation. Namely, for each $u \in L^p(\Omega; \mathbb{R}^M)$, we built a sequence $\{u_n\}_{n\in\mathbb{N}} \subset L^p(\Omega; \mathbb{R}^M)$ of *simpler* objects that approximate u both configuration-wise (namely weakly in L^p) and in energy (namely $\lim_{n\to\infty} H(u_n) = H(u)$). We chose the simpler objects in such a way that the construction of the recovery sequence for each of those can be done *by hands*.

Remark 6.65. The only occasion where we used the fact that the set Ω is bounded, is in the construction of the approximating sequence in Step 2. A similar argument can be used to construct such a sequence also in the case where the assumption of Ω being bounded is dropped. We decided not to consider that case to focus on the technicalities that are more relevant.

Remark 6.66. It can be shown that the (topological) relaxation of F with respect to the weak L^1 topology satisfies the same formula. The only place where we needed $p > 1$ was in Step 1, where we used the weak lower semicontinuity of the sequential lower semicontinuous envelope.

As a direct consequence, we have also an integral representation formula for the relaxation of integral functionals, under suitable growth assumptions on the integrand f that ensures the required growth assumptions on the functional F.

Corollary 6.67. *Let $p \in (1, \infty)$, and $\Omega \subset \mathbb{R}^N$ be a bounded measurable set. Let $f : \mathbb{R}^M \to \mathbb{R}$ be a Borel function, and consider the functional $F : L^p(\Omega; \mathbb{R}^M) \to \overline{\mathbb{R}}$ given by*

$$F(u) := \int_\Omega f(u(x))\, dx.$$

Assume that f is such that

$$C|z|^p \leq |f(z)| \leq C(1 + |z|^p),$$

for all $z \in \mathbb{R}^M$, where $C > 0$. Then,

$$\mathrm{lsc}(F)(u) = \overline{F}(u) = \int_\Omega f^c(u(x))\, dx,$$

for all $u \in L^p(\Omega; \mathbb{R}^M)$, where the relaxations are with respect to the weak topology of L^p.

Proof. Note that, under the stated assumptions, the functional F is weakly coercive in L^p. Thus, applying Corollary 6.33, we get that $\mathrm{lsc}(F) = \overline{F}$. The integral representation formula follows from Theorem 6.63. □

Remark 6.68. In Theorem 6.63 we chose to work with a function f that is not allowed to assume the value $+\infty$. Sometimes this is useful if we want to incorporate a constrain on the values of u into the integrand f. In case $f : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$, we still have an integral representation result for the (sequential) relaxed functional

$$\overline{F}(u) = \begin{cases} \displaystyle\int_\Omega [\mathrm{lsc}(f)]^c(u)\, dx & \text{if } \overline{F}(u) < \infty, \\ +\infty & \text{else.} \end{cases}$$

The proof of this integral representation result is on the same lines as that of Theorem 6.63. Note that,

$$\mathrm{lsc}[(f)^c] \leq [\mathrm{lsc}(f)]^c,$$

but the two might be different. Find an example!

Finally, we state the result for the case of functionals depending on the gradient.

Theorem 6.69. *Let $p \in (1, \infty)$, and $\Omega \subset \mathbb{R}^N$ be a bounded open set. Let $f : \mathbb{R}^M \to \mathbb{R}$ be a Borel function, and consider the functional $F : W^{1,p}(\Omega) \to \overline{\mathbb{R}}$ given by*

$$F(u) := \int_\Omega f(\nabla u(x))\, dx.$$

Assume that

$$|f(z)| \le C(1 + |z|^p),$$

for all $z \in \mathbb{R}^M$*, where* $C > 0$*. Then, its sequential lower semi-continuous envelope* $\overline{F} : W^{1,p}(\Omega) \to \overline{\mathbb{R}}$ *with respect to the weak topology of* $W^{1,p}(\Omega)$ *satisfies*

$$\overline{F}(u) = \int_\Omega f^c(\nabla u(x))\, dx,$$

for all $u \in W^{1,p}(\Omega)$*. Moreover, if*

$$C|z|^p \le |f(z)| \le C(1 + |z|^p),$$

for all $z \in \mathbb{R}^M$*, where* $C > 0$*, then*

$$\mathrm{lsc}(F)(u) = \overline{F}(u) = \int_\Omega f^c(\nabla u(x))\, dx,$$

for all $u \in W^{1,p}(\Omega)$.

Remark 6.70. Note that the *upper bound* is needed in order to get the integral representation formula for $\overline{F}$, while the *lower bound* is needed to ensure coerciveness, and thus to ensure the equivalence between $\overline{F}$ and $\mathrm{lsc}(F)$.

Chapter 7

Gamma-convergence

We now consider the following situation: suppose that we have a minimization problem

$$\min\{\, F_\varepsilon(x) : x \in X_\varepsilon \,\} \tag{7.1}$$

where $F_\varepsilon : X_\varepsilon \to (-\infty, +\infty]$ is a functional parametrized by a parameter ε that is close to a parameter $\varepsilon_0 \in \overline{\mathbb{R}}$. Note that also the domain X_ε of the functional F_ε is allowed to change. We would like to understand the behaviour of minimizers $x_\varepsilon \in X_\varepsilon$ for F_ε as $\varepsilon \sim \varepsilon_0$. Since the minimization problem (7.1) might not have a solution, we can consider *almost* minimizers, namely points $x_\varepsilon \in X_\varepsilon$ such that

$$\lim_{\varepsilon \to \varepsilon_0} |F_\varepsilon(x_\varepsilon) - \inf_{X_\varepsilon} F_\varepsilon| \to 0.$$

As for the case of relaxation, we wonder if it is possible to understand the behaviour of almost minimizers $\{x_\varepsilon\}_\varepsilon$ in a *variational* way. Namely, if we can find a functional $\widetilde{F} : X \to (-\infty, +\infty]$ such that

$$\lim_{\varepsilon \to \varepsilon_0} \inf_{X_\varepsilon} F_\varepsilon = \min_X F, \tag{7.2}$$

and such that every cluster point of $\{x_\varepsilon\}$ (with respect to some topology) is a minimizer of $\widetilde{F}$.

Before continuing, note that in this case we also have to choose a proper space X where $\widetilde{F}$ is defined. This is usually not a problem, because in most of the cases it is possible to find a space X such that $X_\varepsilon \subset X$ for all ε, and thus we can equivalently consider the functionals $G_\varepsilon : X \to \overline{\mathbb{R}}$ defined as

$$G_\varepsilon(x) := \begin{cases} F_\varepsilon(x) & \text{if } x \in X_\varepsilon, \\ +\infty & \text{else.} \end{cases}$$

Thus, in what follows, we will always assume that we have a *single* space X where all of our functionals are defined.

Moreover, we will first consider *sequences* of parametrized functionals, namely we will study the behaviour of a sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ with $\varepsilon_n \to \varepsilon_0$ as $n \to \infty$. The relation with the *continuous* family of parametrized functionals F_ε will be investigated later.

7.1 Gamma-convergence in metric spaces

We recall that we are now working on a *metric* space (X, d). We would like to give a couple of examples to motivate the definitions that we are going to give. Consider the sequence of functionals as in the following figure.

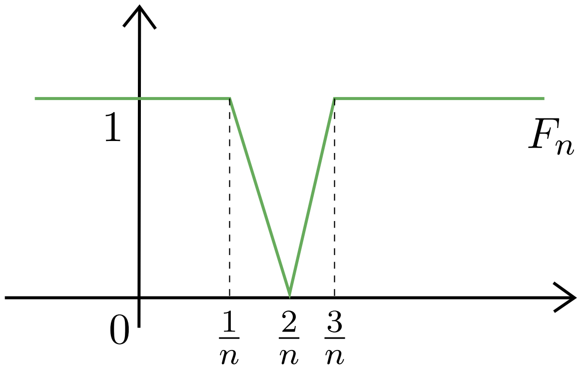


Figure 7.1: The functionals F_n all have the same minimum and the minimizers are converging to the origin.

We see that

$$\min_X F_n = 0, \qquad \text{and} \qquad \lim_{n\to\infty} x_n = 0, \quad \text{where} \quad F_n(x_n) = \min_X F_n.$$

Thus, we might be tempted to defined our limiting functional $\widetilde{F}$ as

$$\widetilde{F}(x) := \inf\left\{ \liminf_{n\to\infty} F_n(x_n) \; : \; x_n \to x \right\}.$$

The problem with this definition is that we cannot ensure the validity of (7.2). Indeed, consider the case where

$$F_n := (-1)^n.$$

In this case, every $x \in X$ is a minimizer of F_n for every $n \in \mathbb{N}$. Moreover, we have that $\widetilde{F} \equiv -1$. But (7.2) fails to hold. To impose the convergence of the infima to the minimum of the limiting functional, we need to make sure that

$$\liminf_{n\to\infty} \inf_{X_\varepsilon} F_n = \limsup_{n\to\infty} \inf_{X_\varepsilon} F_n.$$

Since we do not what what are the minimizers and where they converge (this is precisely why we are doing all of this!), we are forced to impose such a condition at every point. This fact will even more clear when we will give the topological characterization of Γ-limit (which is actually the more general definition of Γ-limit in topological spaces). We now hope that the following definition does not come as a surprise.

Definition 7.1. Consider a sequence of functionals $(F_n)_{n\in\mathbb{N}}$, where $F_n : X \to \overline{\mathbb{R}}$. We define the Γ-*lower limit* $\Gamma\text{-}\liminf_{n\to\infty} F_n : X \to \overline{\mathbb{R}}$ as

$$\left(\Gamma\text{-}\liminf_{n\to\infty} F_n\right)(x) := \inf\left\{ \liminf_{n\to\infty} F_n(x_n) \ : \ x_n \to x \right\},$$

and the Γ-*upper limit* $\Gamma\text{-}\limsup_{n\to\infty} F_n : X \to \overline{\mathbb{R}}$ as

$$\left(\Gamma\text{-}\limsup_{n\to\infty} F_n\right)(x) := \inf\left\{ \limsup_{n\to\infty} F_n(x_n) \ : \ x_n \to x \right\},$$

for every $x \in X$. Moreover, we say that the sequence $(F_n)_{n\in\mathbb{N}}$ Γ-converges to $F : X \to \overline{\mathbb{R}}$ if

$$F = \Gamma\text{-}\liminf_{n\to\infty} F_n = \Gamma\text{-}\limsup_{n\to\infty} F_n$$

In this case we write $F_n \overset{\Gamma-\mathrm{d}}{\longrightarrow} F$.

Remark 7.2. Note that, in general, the Γ-lower limit and the Γ-upper limit are different. Indeed, if like in the example in the discussion above, we have $F_n := (-1)^n$, we get that

$$\Gamma\text{-}\liminf_{n\to\infty} F_n \equiv -1, \qquad \Gamma\text{-}\limsup_{n\to\infty} F_n \equiv 1,$$

which are different at every point.

Remark 7.3 (The choice of the metric). We would like to stress that the definition of Γ-lower limit and the Γ-upper limit *depend* on the distance that we consider on the metric space X: different distances might (and usually do) lead to different Γ-limit inf and sup. The choice of the metric is of capital importance in the study of the limiting behaviour of almost minimizing sequences of F_n, since it *specifies* what we mean by limiting behaviour! Usually, the choice of the topology is suggested (or even forced) by compactness arguments:

Remark 7.4. The Γ-lower limit is the greatest lower semi-continuous function that is below F_n at $n = \infty$. Unfortunately, a similar interpretation of the Γ-upper limit is not available.

We first notice that the notion of relaxation is a particular instance of that of Γ-convergence.

Lemma 7.5. *Let $F : X \to \overline{\mathbb{R}}$. Set $F_n := F$, for all $n \in \mathbb{N}$. Then, $(F_n)_{n\in\mathbb{N}}$ Γ-converges to* $\mathrm{lsc}(F)$.

We now prove some basic properties of Γ-liming and Γ-limsup. First of all, a sequential characterization of the Γ-limit by a liminf and a limsup inequality in the same spirit of that for the relaxed functional (see Proposition 6.11) is available and it is usually how Γ-limit results are proved. The proof follows the same line as the proof of Proposition 6.11 and therefore we will not repeat it in here.

Lemma 7.6. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals, where $F_n : X \to \overline{\mathbb{R}}$ for each $n \in \mathbb{N}$. Then, $F_n \overset{\Gamma-\mathrm{d}}{\longrightarrow} F$ if and only if the following two conditions are satisfied:*

(i) (Liminf inequality) For every $\overline{x} \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ converging to $\overline{x}$, it holds

$$F(\overline{x}) \leq \liminf_{n\to\infty} F_n(x_n);$$

(ii) (Limsup inequality) For every $\overline{x} \in X$, there exists $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$F(\overline{x}) \geq \limsup_{n\to\infty} F_n(x_n),$$

and with $x_n \to \overline{x}$.

Remark 7.7. As for the case of the relaxed functional, the limsup inequality combined with the liminf inequality gives the existence of the so called *recovery sequence*: for every $\overline{x} \in X$ there exists $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$\lim_{n\to\infty} F_n(x_n) = F(\overline{x}),$$

and with $x_n \to \overline{x}$.

A topological characterization holds for the notion of Γ-lower and upper limit. The proof is similar to that of Proposition 6.20.

Proposition 7.8. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals, where $F_n : X \to \overline{\mathbb{R}}$ for each $n \in \mathbb{N}$. Then,

$$\Gamma\text{-}\liminf_{n\to\infty}(x) = \sup_{r>0} \liminf_{n\to\infty} \inf_{y\in B_r(x)} F_n(y) = \lim_{r\to 0} \liminf_{n\to\infty} \inf_{y\in B_r(x)} F_n(y), \tag{7.3}$$

and

$$\Gamma\text{-}\limsup_{n\to\infty}(x) = \sup_{r>0} \limsup_{n\to\infty} \inf_{y\in B_r(x)} F_n(y) = \lim_{r\to 0} \limsup_{n\to\infty} \inf_{y\in B_r(x)} F_n(y), \tag{7.4}$$

for all $x \in X$.

Remark 7.9. This topological version is the definition of Γ-lower and upper limit in general topological space, with $\sup_{U\in\mathcal{N}(x)}$ in place of $\sup_{r>0}$, where $\mathcal{N}(x)$ is the family of neighborhoods of the point x.

In a similar spirit as Proposition 7.8, we give another characterization of Γ-liminf and Γ-limsup that will be useful in the following. The idea is to *force* the minimization in (7.3) and (7.4) to be in small balls centered at the point x by adding a perturbation that is small around x, and very large outside of a ball centered at x.

Theorem 7.10. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals $F_n : X \to [0, +\infty]$. Let $\Phi : X \to [0, +\infty)$ be a continuous function such that*

(i) $\Phi(0) = 0$*;*

(ii) $\Phi(t) > 0$ *for all* $t > 0$*;*

(iii) $\liminf_{t\to\infty} \Phi(t) > 0$.

Then,

$$\Gamma\text{-}\liminf_{n\to\infty} F_n(x) = \sup_{\lambda>0} \liminf_{n\to\infty} \inf_{y\in X}[F_n(y) + \lambda\Phi(\mathrm{d}(y, x))] = \lim_{\lambda\to\infty} \liminf_{n\to\infty} \inf_{y\in X}[F_n(y) + \lambda\Phi(\mathrm{d}(y, x))],$$

and

$$\Gamma\text{-}\limsup_{n\to\infty} F_n(x) = \sup_{\lambda>0} \limsup_{n\to\infty} \inf_{y\in X}[F_n(y) + \lambda\Phi(\mathrm{d}(y, x))] = \lim_{\lambda\to\infty} \limsup_{n\to\infty} \inf_{y\in X}[F_n(y) + \lambda\Phi(\mathrm{d}(y, x))],$$

for all $x \in X$.

Remark 7.11. In particular, we get that

$$\mathrm{lsc}(F)(x) = \sup_{\lambda>0} \inf_{y\in X}[F(y) + \lambda\Phi(\mathrm{d}(y, x))],$$

for all $x \in X$.

First of all, by using the topological characterization provided in Proposition 7.8, it is possible to see that the lower and upper Γ-limit are lower semicontinuous.

Lemma 7.12. *Let* $(F_n)_{n\in\mathbb{N}}$ *be a sequence of functionals. Then,* $\Gamma\text{-}\liminf_{n\to\infty} f_n$ *and* $\Gamma\text{-}\limsup_{n\to\infty} F_n$ *are lower semicontinuous.*

We now show that, in order to compute the the Γ-liminf and the Γ-limsup, we can assume, without loss of generality, the functionals to be lower semicontinuous.

Proposition 7.13. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals, where $F_n : X \to \overline{\mathbb{R}}$ for each $n \in \mathbb{N}$. Then,

$$\Gamma\text{-}\liminf_{n\to\infty} F_n(x) = \Gamma\text{-}\liminf_{n\to\infty} \mathrm{lsc}(F_n)(x), \qquad \Gamma\text{-}\limsup_{n\to\infty} F_n(x) = \Gamma\text{-}\limsup_{n\to\infty} \mathrm{lsc}(F_n)(x).$$

for all $x \in X$.

Proof. We prove that

$$\Gamma\text{-}\liminf_{n\to\infty} F_n = \Gamma\text{-}\liminf_{n\to\infty} \mathrm{lsc}(F_n).$$

The case of the Γ-limsup follows by using similar computations. Since $\mathrm{lsc}(F_n) \le F_n$, we have that

$$\Gamma\text{-}\liminf_{n\to\infty} \mathrm{lsc}(F_n) \le \Gamma\text{-}\liminf_{n\to\infty} F_n.$$

In order to prove the opposite inequality, let $\overline{x} \in X$, and take $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \to \overline{x}$ such that

$$\lim_{n\to\infty} \mathrm{lsc}(F_n)(x_n) = \Gamma\text{-}\liminf_{n\to\infty} \mathrm{lsc}(F_n)(\overline{x}).$$

The existence of such a sequence follows by using the definition of the Γ-liminf by using the same argument employed in Lemma 6.6. For each $n \in \mathbb{N}$, we know that

$$\mathrm{lsc}(F_n)(x_n) = \inf\left\{ \liminf_{k\to\infty} F_n(y_k) \,:\, y_k \to x_n \right\}.$$

Thus, there exists $(y_i^n)_{i\in\mathbb{N}} \subset X$ with $y_i^n \to x_n$ as $i \to \infty$ such that

$$\lim_{i\to\infty} F_n(y_i^n) = \mathrm{lsc}(F_n)(x_n).$$

Therefore, for each $n \in \mathbb{N}$, we can find a point $z_n \in X$ such that

$$\mathrm{d}(z_n, x_n) < \frac{1}{n}, \qquad |F_n(z_n) - \mathrm{lsc}(F_n)(x_n)| < \frac{1}{n}.$$

Then, $z_n \to \overline{x}$. Therefore, by definition of the Γ-lower limit, we get

$$\begin{aligned}
\left(\Gamma\text{-}\liminf_{n\to\infty} F_n\right)(\overline{x}) &\le \liminf_{n\to\infty} F_n(z_n) \\
&\le \liminf_{n\to\infty} [\, \mathrm{lsc}(F_n)(x_n) + |F_n(z_n) - \mathrm{lsc}(F_n)(x_n)|\,] \\
&= \left(\Gamma\text{-}\liminf_{n\to\infty} \mathrm{lsc}(F_n =\right)(\overline{x}),
\end{aligned}$$

thus obtaining the desired inequality. □

We conclude this section by stating a result ensuring that, up to a subsequence, a sequence of functionals over a nice space admits a subsequence that Γ-converges to a limit.

Theorem 7.14. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals on a separable metric space X. Then, there exists a subsequence $(F_{n_k})_{k\in\mathbb{N}}$ that Γ converges to F, for some functional $F : X \to \overline{\mathbb{R}}$.*

7.2 Comparison with pointwise and uniform limit

In order to understand a bit better the notion of Γ-limit we now compare it with the better known notions of pointwise and uniform limit. The Γ-limit and the pointwise limit, if they exist, are, in general different, as the example in Figure 7.1 shows. Nevertheless, they always satisfy the following inequality.

Lemma 7.15. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals. Then,*

$$\Gamma\text{-}\liminf_{n\to\infty} F_n \leq \liminf_{n\to\infty} F_n, \qquad \Gamma\text{-}\limsup_{n\to\infty} F_n \leq \limsup_{n\to\infty} F_n.$$

In particular, if $(F_n)_{n\in\mathbb{N}}$ converges pointwise to F, then

$$\Gamma\text{-}\liminf_{n\to\infty} F_n \leq \Gamma\text{-}\limsup_{n\to\infty} F_n \leq F.$$

Proof. It follows directly from the definition of Γ-lower limit and Γ-upper limit by taking the constant sequence $x_n = x$ for all $n \in \mathbb{N}$. □

Remark 7.16. The Γ-limit and the pointwise limit of a sequence of functionals $(F_n)_{n\in\mathbb{N}}$ might both exist and be different at every point! Indeed, given an enumeration of the rationals $(q_n)_{n\in\mathbb{N}}$ of $\mathbb{R}$, consider the functionals

$$F_n(x) := \begin{cases} 0 & \text{if } x = q_n, \\ 1 & \text{else.} \end{cases}$$

Then, by the density of $\mathbb{Q}$ in $\mathbb{R}$, we get that $(F_n)_{n\in\mathbb{N}}$ Γ-converges to $F \equiv 0$, but converges pointwise to $G \equiv 1$.

Remark 7.17. There are situations in which the pointwise limit does not exists, but the Γ-limit does. Indeed, consider the case where $F_n(x) := \sin(nx)$. Then, $(F_n)_{n\in\mathbb{N}}$ does *not* converge pointwise to anything, but Γ-converges to $F \equiv -1$. This example shows that the notion of Γ-convergence can be used to treat situations in which high oscillations take places, and that cannot be handled by the notion of pointwise limit.

We now investigate three cases in which it is possible to state a better relation between the Γ-limit and the pointwise limit:the case of monotone sequences, the case of a sequence converging uniformly, and the case of convex functionals. We start by investigating the case of monotone sequences. First of all, note that a monotone sequence admits a pointwise limit. We will see that there is an essential difference between increasing and decreasing sequences. We start by investigating the former case.

Proposition 7.18. Let $(F_n)_{n\in\mathbb{N}}$ be a decreasing sequence of functionals. Then,

$$\Gamma\text{-}\lim_{n\to\infty} F_n = \mathrm{lsc}(F),$$

where F is the pointwise limit of $(F_n)_{n\in\mathbb{N}}$.

Proof. For each $x \in X$ we have that

$$\begin{aligned}
\operatorname{lsc}(F)(x) &= \inf\left\{ \liminf_{n\to\infty} F(x_n) \;:\; x_n \to x \right\} \\
&\le \inf\left\{ \liminf_{n\to\infty} F_n(x_n) \;:\; x_n \to x \right\} \\
&= \Gamma\text{-}\liminf_{n\to\infty} F_n(x) \\
&\le \Gamma\text{-}\limsup_{n\to\infty} F_n(x) \\
&\le F(x),
\end{aligned}$$

where in the second step we used the fact that $F \le F_n$ for all $n \in \mathbb{N}$, while in last step we used Lemma 7.15. Therefore, the Γ-lower limit and the Γ-upper limit are lower semicontinuous functionals (see Lemma 7.12) in between $\operatorname{lsc}(F)$ and F. Since $\operatorname{lsc}(F)$ is the greatest lower semicontinuous functional majorized by F, we get that all of the inequalities above are actually equalities. Therefore, the Γ-limit exists, and is equal to $\operatorname{lsc}(F)(x)$. □

Remark 7.19. The above result states that, in order to compute the Γ-limit of a decreasing sequence of functionals, you first take the pointwise limit, and then you relax it.

We would like to stress that if we first take the relaxation, and then the pointwise limit we get, in general, a different functional. Indeed,

$$\inf_{n\in\mathbb{N}} \operatorname{lsc}(F_n) \neq \operatorname{lsc}\left(\inf_{n\in\mathbb{N}} F_n\right) = \operatorname{lsc}(F).$$

From the technical point of view, the difference between the two objects is due to the fact that we cannot interchange a supremum and an infimum. Consider, for instance, the case $f_n : (-1,1) \to \mathbb{R}$ defined as

$$f_n(x) := \begin{cases} 1 & \text{if } x \le 0, \\ 1 - nx & \text{if } 0 \le x \le \frac{1}{n}, \\ 0 & \text{else.} \end{cases}$$

Then, each f_n is continuous, and $f_n \to f$ pointwise, where

$$f(x) := \begin{cases} 1 & \text{if } x \le 0, \\ 0 & \text{else.} \end{cases}$$

On the other hand, f is not lower semicontinuous. In particular, $\operatorname{lsc}(f)(0) < f(0)$.

We now study what happens for increasing sequences of functionals.

Proposition 7.20. Let $(F_n)_{n\in\mathbb{N}}$ be an increasing sequence of functionals. Then,

$$\left(\Gamma\text{-}\lim_{n\to\infty} F_n\right)(x) = \lim_{n\to\infty} \operatorname{lsc}(F_n)(x) = \sup_{n\in\mathbb{N}} \operatorname{lsc}(F_n)(x),$$

for all $x \in X$.

Proof. First of all, thanks to Proposition 7.13, we have that

$$\Gamma\text{-}\lim_{n\to\infty} F_n = \Gamma\text{-}\lim_{n\to\infty} \operatorname{lsc}(F_n).$$

Let
$$G := \sup_{n\in\mathbb{N}} \operatorname{lsc}(F_n).$$
By Lemma 7.15, we have that
$$\Gamma\text{-}\liminf_{n\to\infty} \operatorname{lsc}(F_n) \leq G.$$
We now prove the opposite inequality. Fix $x \in X$ and let $t < G(x)$. Note that G is lower semicontinuous, since it is the supremum of lower semicontinuous functionals (see Proposition 6.14). Therefore, by (ii) of Proposition 6.21 we have that there exists $r > 0$ such that $G(y) > t$ for all $y \in B(x, r)$. In particular, this means that for all $y \in B(x, r)$ there exists $\bar{n}(y) \in \mathbb{N}$ such that

$$\operatorname{lsc}(F_n)(y) > t,$$

for all $n \geq \bar{n}(y)$. By using the topological characterization given in Proposition 7.8, this yields that

$$\Gamma\text{-}\liminf_{n\to\infty} \operatorname{lsc}(F_n)(x) = \sup_{r>0} \liminf_{n\to\infty} \inf_{y\in B_r(x)} \operatorname{lsc}(F_n)(y) \geq t.$$

Since $t < G(x)$ is arbitrary, we conclude that $\Gamma\text{-}\liminf_{n\to\infty} F_n \geq G$. □

Remark 7.21. In particular, if in the above proposition all of the F_n's are lower semicontinuous, then the Γ-limit equals the pointwise limit.

Remark 7.22. If the lower semicontinuity of the F_n's does not hold, we cannot claim that

$$\Gamma\text{-}\lim_{n\to\infty} F_n = \operatorname{lsc}\left(\sup_{n\in\mathbb{N}} F_n\right),$$

since, in general,
$$\sup_{n\in\mathbb{N}} \operatorname{lsc}(F_n) \neq \operatorname{lsc}\left(\sup_{n\in\mathbb{N}} F_n\right).$$
Consider, for instance, the case $f_n : (0, 1) \to \mathbb{R}$ defined as

$$f_n(x) := \begin{cases} 0 & \text{if } x = q_k,\ k \geq n, \\ 1 & \text{else,} \end{cases}$$

where $(q_n)_{n\in\mathbb{N}}$ is an enumeration of $\mathbb{Q} \cap (0, 1)$. Then, it is possible to see that $(F_n)_{n\in\mathbb{N}}$ converges pointwise to $F \equiv 1$. On the other hand, by using Proposition 7.20, we get that $(F_n)_{n\in\mathbb{N}}$ Γ-converges to $G \equiv 0$.

Remark 7.23. The *metric* that is used to compute the Γ-limit of monotone sequences of functional is seen when we take the relaxation.

In case of a very strong pointwise converge, namely the uniform convergence, the Γ-limit and the uniform limit agree. The proof follows easily from the definition of uniform convergence and that of Γ-limit.

Lemma 7.24. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals converging to some $F : X \to \overline{\mathbb{R}}$ uniformly. Then,*
$$\left(\Gamma\text{-}\lim_{n\to\infty} F_n\right)(x) = \operatorname{lsc}(F)(x),$$
for all $x \in X$.

Finally, there is a particular class of functionals, namely *convex* functionals, for which the pointwise limit, even for sequences that are not monotonically nor uniformly converging, coincides with the Γ-limit. We need though a bit of more structure for the space X in order to talk about convexity, since we need convex combinations of elements of X.

Proposition 7.25. Let $(X, \|\cdot\|)$ be a *normed* vector space. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of convex functionals and assume that they are equibounded in a neighbourhood of every point $x \in X$, namely for each $x \in X$ there exists $R > 0$ and $C > 0$ such that, for all $y \in B(x, R)$, it holds

$$|F_n(y)| \leq C.$$

Then, $(F_n)_{n\in\mathbb{N}}$ Γ-converges to $F : X \to \overline{\mathbb{R}}$ if and only if $(F_n)_{n\in\mathbb{N}}$ converges pointwise to F.

Proof. Step 1. First of all note that the proof of Proposition 6.48 works also in a general normed space X, not necessarily $\mathbb{R}^N$. In particular, thanks to the uniformity of the constants C and R in the condition of equiboundness, this implies that for every $x \in X$ there exist $R > 0$ and $K > 0$ such that

$$|F_n(x) - F_n(y)| \leq K\|x - y\|, \tag{7.5}$$

for all $y \in B(x, R)$, and all $n \in \mathbb{N}$.

Step 1. Let $(x_n)_{n\in\mathbb{N}} \subset X$ be such that $x_n \to x$. We claim that

$$\liminf_{n\to\infty} F_n(x) \leq \Gamma\text{-}\liminf_{n\to\infty} F_n(x), \tag{7.6}$$

and that

$$\limsup_{n\to\infty} F_n(x) \leq \Gamma\text{-}\limsup_{n\to\infty} F_n(x). \tag{7.7}$$

We prove (7.6). Note that (7.7) follows by using similar computations. We can assume, without loss of generality, that $\|x_n - x\| < R$ for all $n \in \mathbb{N}$. Thus,

$$\begin{aligned}
\liminf_{n\to\infty} F_n(x) &= \liminf_{n\to\infty} [\, F_n(x_n) + F_n(x) - F_n(x_n) \,] \\
&\leq \liminf_{n\to\infty} [\, F_n(x_n) + |F_n(x) - F_n(x_n)| \,] \\
&= \liminf_{n\to\infty} F_n(x_n),
\end{aligned}$$

where in the last step we used (7.5) together with the fact that $\|x_n - x\| \to 0$ as $n \to \infty$. We then obtain (7.6) by using the definition of $\Gamma\text{-}\liminf_{n\to\infty} F_n(x)$.

Step 3. By using Lemma 7.15, together with (7.6) and (7.7), we get that

$$\Gamma\text{-}\liminf_{n\to\infty} F_n \leq \liminf_{n\to\infty} F_n \leq \Gamma\text{-}\liminf_{n\to\infty} F_n \leq \Gamma\text{-}\limsup_{n\to\infty} F_n \leq \limsup_{n\to\infty} F_n \leq \Gamma\text{-}\limsup_{n\to\infty} F_n.$$

This shows that, in this case, pointwise convergence is equivalent to Γ-convergence. □

7.3 Γ-limit of sums of functionals

We now investigate when the operation of Γ-limit and the sum of functionals commute. The proofs of these results follow directly from Definition 7.1, properties of liminf and limsup, and Lemma 7.15.

Lemma 7.26. *Let $(F_n)_{n\in\mathbb{N}}, (G_n)_{n\in\mathbb{N}}$ be two sequences of functionals. Then,*

$$\Gamma\text{-}\liminf_{n\to\infty}(F_n + G_n) \geq \Gamma\text{-}\liminf_{n\to\infty} F_n + \Gamma\text{-}\liminf_{n\to\infty} G_n,$$

and

$$\Gamma\text{-}\limsup_{n\to\infty}(F_n + G_n) \geq \Gamma\text{-}\limsup_{n\to\infty} F_n + \Gamma\text{-}\liminf_{n\to\infty} G_n,$$

provided that everything is well defined.

Remark 7.27. Note that the above inequalities can be strict, even if $(F_n)_{n\in\mathbb{N}}$ and $(G_n)_{n\in\mathbb{N}}$ are Γ-converging. Consider for instance $F_n(x) := \sin(nx)$, and $G_n(x) := -\sin(nx)$. Then

$$0 \equiv \Gamma\text{-}\lim(F_n + G_n) > -1 - 1 = \Gamma\text{-}\lim F_n + \Gamma\text{-}\lim G_n.$$

Remark 7.28. Note that by combining the first lower bound given by Lemma 7.26 together with the fact that

$$\Gamma\text{-}\limsup_{n\to\infty}(F_n + G_n) \leq \Gamma\text{-}\limsup_{n\to\infty} F_n + \Gamma\text{-}\limsup_{n\to\infty} G_n,$$

we get

$$\begin{aligned}\Gamma\text{-}\liminf_{n\to\infty} F_n + \Gamma\text{-}\liminf_{n\to\infty} G_n &\leq \Gamma\text{-}\liminf_{n\to\infty}(F_n + G_n)\\ &\leq \Gamma\text{-}\limsup_{n\to\infty}(F_n + G_n) \leq \Gamma\text{-}\limsup_{n\to\infty} F_n + \Gamma\text{-}\limsup_{n\to\infty} G_n.\end{aligned}$$

The reason why upper bounds are not usually useful in the study of Γ-limits is because usually the Γ-limit is guessed by first bounding from below the sequence of functionals and then confirming that guess by proving the limsup inequality.

In order to have equality we need stronger assumptions. We present two situations in which we can ensure the commutative property of Γ-limits and sums.

Proposition 7.29. Let $(F_n)_{n\in\mathbb{N}}, (G_n)_{n\in\mathbb{N}}$ be two sequences of functionals. Assume that the G_n's are everywhere finite on X and converges uniformly to $G : X \to \mathbb{R}$. Then,

$$\Gamma\text{-}\liminf_{n\to\infty}(F_n + G_n) = \Gamma\text{-}\liminf_{n\to\infty} F_n + G,$$

and

$$\Gamma\text{-}\limsup_{n\to\infty}(F_n + G_n) = \Gamma\text{-}\limsup_{n\to\infty} F_n + G.$$

Remark 7.30. Note that in the above result, it is crucial that the functionals are finite everywhere. This, in particular, prevents to use the above result to treat cases where a constraint is inserted into the functional as a $+\infty$ penalization. A particular case of Proposition 7.29 is when $G_n \equiv G$ for some $G : X \to \mathbb{R}$.

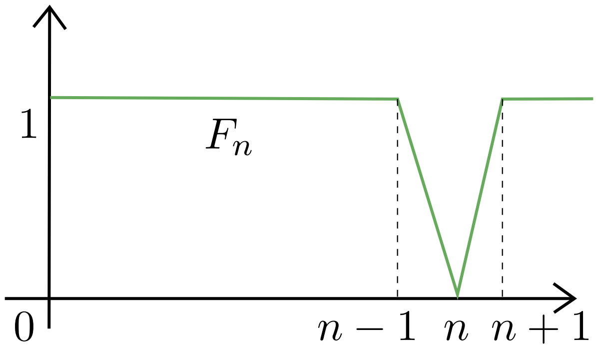


Figure 7.2: An example of a sequence of functionals for which minimizers are *escaping at infinity*.

Proposition 7.31. Let $(F_n)_{n\in\mathbb{N}}, (G_n)_{n\in\mathbb{N}}$ be two sequences of functionals. Assume that there exist $F, G : X \to \overline{\mathbb{R}}$ such that

(i) $F_n \xrightarrow{\Gamma-d} F$, and $F_n \to F$ pointwise;

(ii) $G_n \xrightarrow{\Gamma-d} G$, and $G_n \to G$ pointwise.

Then

$$F_n + G_n \xrightarrow{\Gamma-d} F + G, \qquad F_n + G_n \to F + G \text{ pointwise},$$

if everything is well defined.

7.4 Convergence of minima and (local) minimizers

The convergence of minima and minimizers of Γ-converging sequences is more delicate than in the case of the relaxation. The reason is that, since we have a *sequence* of functionals, the minimizers can *escape at infinity*, and we might have a gap in the value of minima. Consider, for instance, the functionals in Figure 7.2. Then, it holds that

$$F_n \xrightarrow{\Gamma-d} F \equiv 1,$$

but

$$\min_X F_n \equiv 0, \qquad \text{while} \qquad \min_X F = 1.$$

Moreover, we see that there is no cluster point for the sequence of minimizers for F_n, and none of the minimizer for F is the limit of any sequence of minimizers for F_n. The goal of this section is to find sufficient conditions that prevent from this situation to happen.

First of all, from Definition 7.1, we easily obtain that the sequences of infima of F_n and the infima of the Γ-lower limit and Γ-upper limit are always related as follows.

Lemma 7.32. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals. Then,*

$$\inf_X \left(\Gamma\text{-}\liminf_{n\to\infty} F_n \right) \geq \liminf_{n\to\infty} \inf_X F_n, \qquad \inf_X \left(\Gamma\text{-}\limsup_{n\to\infty} F_n \right) \geq \limsup_{n\to\infty} \inf_X F_n.$$

We start with investigating the convergence of minima. A way to avoid the issue with minimizers escaping at infinity, is to require explicitly all of them to stay in a compact set.

Theorem 7.33. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals. Assume that there exists a compact set $K \subset X$ such that*

$$\inf_X F_n = \inf_K F_n, \tag{7.8}$$

for all $n \in \mathbb{N}$. Then, the minimization problem

$$\min_X \left(\Gamma\text{-}\liminf_{n\to\infty} F_n \right)$$

has a solution, and it holds that

$$\min_X \left(\Gamma\text{-}\liminf_{n\to\infty} F_n \right) = \liminf_{n\to\infty} \inf_X F_n. \tag{7.9}$$

Moreover, if $(F_n)_{n\in\mathbb{N}}$ Γ-converges to some functional F, then the minimization problem

$$\min_X \left(\Gamma\text{-}\lim_{n\to\infty} F_n \right)$$

has a solution, and it holds that

$$\min_X \left(\Gamma\text{-}\lim_{n\to\infty} F_n \right) = \lim_{n\to\infty} \inf_X F_n. \tag{7.10}$$

Proof. By using Proposition 7.13 and Theorem 6.10 we can assume, without loss of generality, that each F_n is lower semicontinuous. Let $(F_{n_k})_{k\in\mathbb{N}}$ be a subsequence such that

$$\liminf_{n\to\infty} \inf_X F_n = \lim_{k\to\infty} \inf_X F_{n_k}. \tag{7.11}$$

Step 1. For each $n \in \mathbb{N}$, since F_n lower semicontinuity of and coercive (thanks to (7.8)) by Theorem 5.2 we get that the minimization problem

$$\min_X F_n$$

admits a solution. In particular, for each $k \in \mathbb{N}$ there exists $x_{n_k} \in K$ such that

$$F_{n_k}(x_{n_k}) = \min_X F_{n_k}.$$

By using the compactness of K, it is possible to extract a subsequence $(x_{n_{k_i}})_{i\in\mathbb{N}}$ such that $x_{n_{k_i}} \to \overline{x}$ as $i \to \infty$, for some $\overline{x} \in K$. By using Lemma 7.32, we get

$$\begin{aligned} \liminf_{n\to\infty} \inf_X F_n &\leq \inf_X \left(\Gamma\text{-}\liminf_{n\to\infty} F_n \right) \leq \left(\Gamma\text{-}\liminf_{n\to\infty} F_n \right)(\overline{x}) \\ &\leq \liminf_{i\to\infty} F_{n_{k_i}}(x_{n_{k_i}}) = \liminf_{i\to\infty} \inf_X F_{n_{k_i}} \\ &= \liminf_{n\to\infty} \inf_X F_n, \end{aligned}$$

where last equality follows from (7.11), while in the third inequality we used the definition of Γ-liminf together with the fact that

$$\liminf_{n\to\infty} F_n(x_n) \le \liminf_{i\to\infty} F_{n_{k_i}}(x_{n_{k_i}}),$$

where $x_n := x_{n_{k_i}}$ if $n = n_{k_i}$, and $x_n := x$ else. This yields that $\overline{x}$ is a minimum point for $\Gamma\text{-}\liminf_{n\to\infty} F_n$ and the validity of (7.9).

Step 2. Assume that $(F_n)_{n\in\mathbb{N}}$ Γ-converges to some functional F. Then, from Step 1 we get that the minimization problem

$$\min_X \left(\Gamma\text{-}\lim_{n\to\infty} F_n \right)$$

admits a solution, and that

$$\min_X \left(\Gamma\text{-}\lim_{n\to\infty} F_n \right) = \liminf_{n\to\infty} \inf_X F_n.$$

To prove that the liminf on the right-hand side is a limit, we observe that

$$\liminf_{n\to\infty} \inf_X F_n = \min_X \left(\Gamma\text{-}\lim_{n\to\infty} F_n \right) \ge \limsup_{n\to\infty} \inf_X F_n,$$

where the last inequality follows from Lemma 7.32. This proves (7.10). □

Remark 7.34. In general, it is *not* true that

$$\min_X \left(\Gamma\text{-}\limsup_{n\to\infty} F_n \right) = \limsup_{n\to\infty} \inf_X F_n.$$

Indeed, for each $n \in \mathbb{N}$, consider the functionals $F_n : [-1, 1] \to \mathbb{R}$ defined as

$$F_n(x) := \begin{cases} (-1)^n & \text{if } x \in [-1, 0), \\ \\ (-1)^{n+1} & \text{if } x \in [0, 1]. \end{cases}$$

Then, we get that

$$\Gamma\text{-}\liminf_{n\to\infty} F_n = -1, \qquad \Gamma\text{-}\limsup_{n\to\infty} F_n = 1.$$

Therefore, for all $n \in \mathbb{N}$, we get that

$$\inf_{[-1,1]} F_n = -1, \qquad \inf_{[-1,1]} \Gamma\text{-}\limsup_{n\to\infty} F_n = 1.$$

Therefore, there cannot be a similar formula connecting the minimum of the Γ-upper limit and the sequence of infima of F_n.

We now consider the more delicate case of convergence of minimizers. As for the convergence of infima, we need to make sure that minimizers (if they exist) do not escape at infinity. This will ensure that cluster points of minimizing sequences will be minimizers of the limiting functional.

We wonder if every minimizer of the limiting functional can be approximated by a sequence of minimizers for F_n. The answer is simply no, as it can easily seen by considering the functions $F_n : \mathbb{R} \to \mathbb{R}$ given by

$$F_n(x) := \frac{x^2}{n}.$$

Then, for each $n \in \mathbb{N}$, the only minimizer of F_n is $x = 0$. On the other hand, $(F_n)_{n\in\mathbb{N}}$ Γ-converges to $F \equiv 0$ (since it converges to F uniformly on every compact set). But if we take $x \neq 0$, we cannot approximate it by any sequence $(x_n)_{n\in\mathbb{N}}$, where each x_n is a minimizer of F_n.

In order to get the desired approximation property of *every* minimizer of the limiting functional, we need to introduce the notion of quasi-minimizers.

Definition 7.35. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals. We say that $(x_n)_{n\in\mathbb{N}} \subset X$ is a sequence of *quasi-minimizers* for $(F_n)_{n\in\mathbb{N}}$ if there exists $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0, 1]$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that

$$F_n(x_n) \leq \left[\inf_X F_n + \varepsilon_n\right] \vee -\frac{1}{\varepsilon_n}$$

for all $n \in \mathbb{N}$.

Remark 7.36. The apparently strange definition above is to take also in consideration the case where $\inf_X F_n = -\infty$. In case $\inf_X F_n > -\infty$, the condition for a quasi-minimizer writes as

$$F_n(x_n) \leq \inf_X F_n + \varepsilon_n,$$

for all $n \in \mathbb{N}$.

We are now in position to prove the result concerning the limiting behaviour of quasi-minimizers. We will focus on the case where the sequence has a Γ-limit. More delicate results in the case where the Γ-limit does not exists hold for the Γ-lower and upper limit.

Theorem 7.37. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals. Assume that there exists a compact set $K \subset X$ such that*

$$\inf_X F_n = \inf_K F_n,$$

for all $n \in \mathbb{N}$, and that $(F_n)_{n\in\mathbb{N}}$ Γ converges to F. If $(x_n)_{n\in\mathbb{N}}$ is a sequence of quasi-minimizers for $(F)_{n\in\mathbb{N}}$, then every of its cluster points is a minimum for F. Moreover, every minimizer of F is the limit of a sequence of quasi-minimizers for $(F_n)_{n\in\mathbb{N}}$.

Proof. By using Proposition 7.13 and Theorem 6.10, we can assume, without loss of generality, that each F_n is lower semicontinuous.

Step 1. The proof that every cluster point of a quasi-minimizing sequence $(x_n)_{n\in\mathbb{N}}$ for $(F_n)_{n\in\mathbb{N}}$ is a minimizer for F uses a similar argument to that employed in the proof of Theorem 7.33.

Step 2. Let $\overline{x} \in X$ be a minimizer of F. Assume by contradiction that there is no sequence of quasi-minimizers for $(F_n)_{n\in\mathbb{N}}$ having $\overline{x}$ as a cluster point. Then it is possible to find $\delta > 0$, and $n_1 \in \mathbb{N}$ such that

$$F(\overline{x}) > \inf_X F_n + \delta,$$

for all $n \geq n_1$. By using the topological characterization of the Γ-limit (see Proposition 7.8), we get that there exists $n_2 \in \mathbb{N}$ (without loss of generality, we can assume $n_2 \geq n_1$), and r_0 such that

$$\inf_{B(\overline{x},r)} F_n \geq \inf_X F_n + \delta, \tag{7.12}$$

for all $n \geq n_2$. Now, let $(x_n)_{n\in\mathbb{N}} \subset X$ be a recovery sequence for $F(\overline{x})$, namely

$$\lim_{n\to\infty} F_n(x_n) = F(\overline{x}).$$

Let $\overline{n} \in \mathbb{N}$ be such that $x_n \in B(x, r)$ for all $n \geq \overline{n}$. By using (7.12) we get

$$\lim_{n\to\infty} \inf_X F_n + \delta \leq \lim_{n\to\infty} F_n(x_n) = F(\overline{x}) = \min_X F = \lim_{n\to\infty} \inf_X F_n,$$

where in the last equality we used the result of Theorem 7.33. This gives the desired contradiction. □

Remark 7.38. In the previous result the metric of the space X plays the crucial role of relating the notion of compactness to that of Γ-limits.

The condition that there exists a compact set $K \subset X$ such that

$$\inf_X F_n = \inf_K F_n$$

for all $n \in \mathbb{N}$, is usually hard to check. Usually ones uses a stronger condition, but with the advantage of being easier to check.

Corollary 7.39. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of equi-coercive functionals. Then, all of the results of Theorem 7.33 and Theorem 7.37 hold.*

The above result, together with Theorem 7.10, allows us to give a characterization of the Γ-limit in terms of convergence of minimum problems. We first need to recall the notion of equi-coerciveness for a sequence of functionals.

Definition 7.40. We say that a sequence of functionals $(F_n)_{n\in\mathbb{N}}$ is *equi-coercive* if there exists a lower semi-continuous coercive functional $\Psi : X \to \overline{\mathbb{R}}$ such that $F_n \geq \Psi$ for all $n \in \mathbb{N}$.

Theorem 7.41. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of equi-coercive functionals $F_n : X \to [0, +\infty]$. Let $\Phi : X \to [0, +\infty]$ be a continuous function satisfying the assumption in Theorem 7.10. Let $(\lambda_j)_{j\in\mathbb{N}}$ be an increasing sequence of real numbers going to $+\infty$. Assume that the sequence of functionals*

$$y \mapsto F_n(y) + \lambda_1\Phi(\mathrm{d}(y, x))$$

is equicoercive for all $x \in X$. Then, the followings are equivalent:

(i) $(F_n)_{n\in\mathbb{N}}$ *Γ-converges to a lower semicontinuous functional $F : X \to [0, +\infty]$;*

(ii) It holds that

$$\lim_{n\to\infty} \inf_{y\in X} \left[F_n(y) + \lambda_j\Phi(\mathrm{d}(y, x))\right] = \inf_{y\in X} \left[F(y) + \lambda_j\Phi(\mathrm{d}(y, x))\right],$$

for every $j \in \mathbb{N}$ and every $x \in X$.

Proof. Step 1. We prove that (i) implies (ii). Then, for every $j \in \mathbb{N}$ and $x \in X$, the functional

$$y \mapsto \lambda_j\Phi(\mathrm{d}(y, x))$$

is continuous and everywhere finite. Therefore, from Proposition 7.29 we get that

$$F_n + \lambda_j\Phi(\mathrm{d}(\cdot, x)) \xrightarrow{\Gamma} F + \lambda_j\Phi(\mathrm{d}(\cdot, x)).$$

By assumption, we get that, for each $j \in \mathbb{N}$, the sequence of functional

$$y \mapsto F_n(y) + \lambda_j \Phi(\mathrm{d}(y, x))$$

is equicoercive. Thus, from Corollary 7.39 we get the desired convergence of minima.

Step 2. We prove that (i) implies (ii). Fix $x \in X$. Thanks to Theorem 7.10, we get that

$$\begin{aligned}
\Gamma\text{-}\liminf_{n\to\infty} F_n(x) = \Gamma\text{-}\limsup_{n\to\infty} F_n(x) &= \lim_{j\to\infty}\lim_{n\to\infty}\inf_{y\in X}(F(y) + \lambda_j\Phi(\mathrm{d}(y, x))) \\
&= \lim_{j\to\infty}\inf_{y\in X}(F(y) + \lambda_j\Phi(\mathrm{d}(y, x))) \\
&= F(x),
\end{aligned}$$

where in the last step we used Remark 7.11. □

Finally, we consider the case of the approximation of *local* minimizers. This is very important because in a lot of situations, the (physical) *system* that you (observe) are interested in is *locked* in a local minimum of your functional.

Theorem 7.42. *Let $(F_n)_{n\in\mathbb{N}}$ be an equi-coercive sequence of lower semicontinuous functional* Γ-*converging to F. Let $\overline{x} \in X$ be an* isolated *local minimizer of F. Then, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$, where each x_n is a local minimizer of F_n, such that $x_n \to \overline{x}$ as $n \to \infty$.*

Proof. Note that the fact that $\overline{x}$ is an isolated local minimizer of F implies that $F(\overline{x}) < +\infty$. Let $r > 0$ such that

$$F(\overline{x}) = \min\{F(y) \,:\, \mathrm{d}(y, \overline{x}) \le r\}.$$

For each $n \in \mathbb{N}$, consider the constrained minimization problem

$$\min\{F_n(y) \,:\, \mathrm{d}(y, \overline{x}) \le r\}. \tag{7.13}$$

We claim that there exists $\overline{n} \in \mathbb{N}$ depending on $r > 0$, such that for all $n \ge \overline{n}$, there exists a solution x_n to (7.13) satisfying

$$\mathrm{d}(x_n, \overline{x}) < r. \tag{7.14}$$

Namely, x_n is a local minimizer of F_n, since

$$F_n(x_n) = \min\{F_n(y) \,:\, \mathrm{d}(y, x_n) \le r - \mathrm{d}(x_n, \overline{x}).$$

This allows us to conclude. Indeed, assume that the claim is true, and consider a subsequence $(r_n)_{n\in\mathbb{N}} \subset (0, 1)$ be such that $r_n \to 0$ as $n \to \infty$. Then, for each $n \in \mathbb{N}$, thanks to (7.14), there exists a local minimizer x_n of F_n with $\mathrm{d}(x_n, \overline{x}) < r_n$. Thus, $(x_n)_{n\in\mathbb{N}}$ is the required sequence of local minimizers converging to $\overline{x}$.

Let us prove the claim. Assume by contradiction that this is not the case. Then, for each $n \in \mathbb{N}$, we would have that

$$\min\{F_n(y) \,:\, \mathrm{d}(y, \overline{x}) \le r\} = \min\{F_n(y) \,:\, \mathrm{d}(y, \overline{x}) = r\} < \min\{F_n(y) \,:\, \mathrm{d}(y, \overline{x}) \le r/2\}. \tag{7.15}$$

By using the sequential characterization of the Γ-limit, there exists a sequence $(y_n)_{n\in\mathbb{N}}$ with $y_n \to \overline{x}$ such that $F_n(y_n) \to F(\overline{x})$. In particular, for n large $y_n \in B_r(\overline{x})$, and

$$F_n(y_n) \le F(\overline{x}) + 1 < +\infty.$$

Thus, there exists $\overline{n} \in \mathbb{N}$ such that

$$\sup_{n \geq \overline{n}} F_n(x_n) < +\infty.$$

Since the sequence $(F_n)_{n\in\mathbb{N}}$ is equi-coercive, it is possible to find a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $x_{n_k} \to z \in X$. We claim that $z = \overline{x}$, and thus the contradiction, since $\mathrm{d}(x_{n_k}, \overline{x}) = r$ for all $k \in \mathbb{N}$. Assume that $z \neq \overline{x}$. Then

$$F(z) \leq \liminf_{k\to\infty} F_{n_k}(x_{n_k}) \leq \limsup_{k\to\infty} \inf_{B_{r/2}}(\overline{x}) F_{n_k} \leq F(\overline{x}),$$

where in the last step we used (7.15). This contradicts the isolated local minimality of $\overline{x}$. Thus, we get that $z = \overline{x}$ and we conclude the proof. □

Remark 7.43. Note that, in general, it is not true that an isolated local minimizer can be approximated by *isolated* local minimizers. Indeed, consider the case where

$$F_n(x) := \begin{cases} \frac{1}{n} & \text{if } x \in [-1/n, 1/n], \\ & \\ |x| & \text{else.} \end{cases}$$

Then, the sequence $(F_n)_{n\in\mathbb{N}}$ Γ converges to $F : \mathbb{R} \to \mathbb{R}$, where $F(x) := |x|$, but the isolated local minimizer (actually, the only global minimizer) $x = 0$ of F cannot be approximated by *isolated* local minimizers of F_n.

Moreover, if $\overline{x} \in X$ is a local minimizer, not isolated, in general, it is not the limit of local minimizers of F_n. This is the same situation as for global minimizers. We would need to consider the notion of quasi-local minimizer.

7.5 Γ-convergence and weak topologies

We now consider the notion of Γ-convergence for weak topologies. As mentioned previously, in a general topological space (X, τ), it is possible to define the notion of Γ-inferior limit and Γ-superior limit by using the same idea behind the indentitys of Proposition 7.8.

Definition 7.44. Given a sequence of functionals $(F_n)_{n\in\mathbb{N}}$, we define

$$F^-(x) := \sup_{U\in\mathcal{N}(x)} \liminf_{n\to\infty} \inf_U F_n,$$

and

$$F^+(x) := \sup_{U\in\mathcal{N}(x)} \limsup_{n\to\infty} \inf_U F_n,$$

where $\mathcal{N}(x)$ is the family of neighborhoods of the point x in the topology τ. Moreover we say that the sequence $(F_n)_{n\in\mathbb{N}}$ Γ-converges to a functional F with respect to the topology τ, if $F = \mathbb{F}^- = \mathbb{F}^+$.

These definitions are not very useful when one has to compute the Γ-limit. That is why it would be nice to be able to use the sequential characterization of the Γ-limit provided by Lemma 7.6. In general, the objects obtained by the topological definition and by the sequential definition do no coincide. A condition ensuring that they are the same is to ask the topological space X to satisfy the first axiom of contability. Unfortunately, weak topologies does not satisfy this assumption.

Nevertheless, there are conditions ensuring that they are the same (cfr. Section 6.3). We present here the statement for a Banach space X whose dual space X' is separable. Note that this leaves out the case of $X = L^1$.

Theorem 7.45. *Let X be a Banach space such that X' is separable. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals $F_n : X \to \overline{\mathbb{R}}$ such that there exists $\Psi : X \to \overline{\mathbb{R}}$ with*

$$\lim_{\|x\|\to\infty} \Psi(x) = +\infty$$

such that $F_n \geq \Psi$ for all $n \in \mathbb{N}$. Let $\mathrm{d} : X \times X \to [0,\infty)$ be any distance that induces the weak topology on bounded sets. Namely

$$x_n \rightharpoonup x \qquad \text{if and only if} \qquad (x_n)_{n\in\mathbb{N}} \quad \text{is bounded and } \lim_{n\to\infty} \mathrm{d}(x_n, x) = 0.$$

Then, by considering the weak topology on X, it holds that

$$\Gamma\text{-}\liminf_{n\to\infty} F_n(x) = F^-(x), \qquad \Gamma\text{-}\limsup_{n\to\infty} F_n(x) = F^+(x),$$

where the left-hand sides are computed with respect to the distance d*. In particular, $(F_n)_{n\in\mathbb{N}}$ Γ-converges to F with respect to the weak topology if and only if the followings hold:*

(i) (Liminf inequality) For every $\overline{x} \in X$, and $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \rightharpoonup \overline{x}$, it holds

$$F(\overline{x}) \leq \liminf_{n\to\infty} F_n(x_n);$$

(ii) (Limsup inequality) For every $\overline{x} \in X$, there exists $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$F(\overline{x}) \geq \limsup_{n\to\infty} F_n(x_n),$$

and with $x_n \rightharpoonup \overline{x}$.

Under the previous assumptions also the Urysohn properties and compactness hold. Namely, the followings hold.

Proposition 7.46. Let X be a Banach space such that X' is separable, and let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals $F_n X \to \overline{\mathbb{R}}$. Then $(F_n)_{n\in\mathbb{N}}$ Γ converges to F if and only if every subsequence $(F_{n_k})_{k\in\mathbb{N}}$ has a further subsequence Γ converging to F.

Theorem 7.47. *Let X be a Banach space such that X' is separable, and let $(F_n)_{n\in\mathbb{N}}$ be a sequence of functionals $F_n : X \to \overline{\mathbb{R}}$. Then, there exists a subsequence $(F_{n_k})_{k\in\mathbb{N}}$ such that F_{n_k} Γ converges to F, for some functional $F : X \to \overline{\mathbb{R}}$.*

7.6 Integral representation of Γ-limits of integral functionals on L^p

In this section we consider a sequence of integral functionals $F_n : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$, for $p \in (1,+\infty)$, and study what kind of Γ-limits we can expect. Namely, write

$$F_n(u) := \int_\Omega f_n(u(x))\, dx,$$

for some Borel function $f_n : \mathbb{R}^M \to \mathbb{R}$. Is is true that if $(F_n)_{n\in\mathbb{N}}$ Γ-converges with respect to the weak topology of L^p, to some functional $F : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R}\cup\{+\infty\}$, then also F is an integral functional? That is, there exists some Borel function $f : \mathbb{R}^M \to \mathbb{R}$ such that

$$F(u) = \int_\Omega f(u(x))\,dx,$$

for all $u \in L^p(\Omega;\mathbb{R}^M)$? And if so, is it possible to characterize the limiting energy density f by using the sequence $(f_n)_{n\in\mathbb{N}}$? These are respectively known as *integral representation* of the Γ-limit, and *representation formula* for the limiting density.

First, we need to introduce a technical tool that is interesting in itself, namely the L-Yosida approximation of a function: the greatest Lipschitz function with Lipschitz constant L that is majorized the given function.

Definition 7.48. Let (X,d) be a metric space, $L > 0$, and let $f : X \to \mathbb{R}\cup\{\infty\}$. Define the *$L$-Yosida transform of f $f_L : X \to \mathbb{R}$* as

$$f_L(x) := \inf_{y\in X}[f(y) + L\,\mathrm{d}(x,y)]\,.$$

The Yosida approximation enjoys some properties whose easy proof is left to the reader.

Proposition 7.49. Let (X,d) be a metric space. Let $f : X \to \mathbb{R}\cup\{\infty\}$ be lower semicontinuous and bounded from below, namely such that there exists $C > -\infty$ such that $f(x) \geq C$ for each $x \in X$. Then, it holds that:

(i) f_L is Lipschitz with Lipschitz constant L;

(ii) f_L is increasing and $f_L \leq f$ for each $L > 0$;

(iii) f_L converges to f as $L \to \infty$.

Moreover, if f is convex, then f_L is convex for all $L > 0$.

Remark 7.50. The previous result is false if the assumption of lower semicontinuity of f is dropped, as it can be easily seen by considering the function

$$f(x) := \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

In that case $f_L(0) = 0$ for all $L > 0$, and thus it cannot converge to $f(0) = 1$.

We are now in position to prove the main result of this section.

Theorem 7.51. *Let $p \in (1,\infty)$, and let $f_n : \mathbb{R}^M \to \mathbb{R}$ be Borel functions such that*

$$\sup_{n\in\mathbb{N}} f_n(0) < +\infty,$$

and such that there exists $c > 0$ for which

$$f_n(\xi) \geq c(|\xi|^p - 1) \tag{7.16}$$

for all $\xi \in \mathbb{R}^M$. Consider the functionals $F_n : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ defined as

$$F_n(u) := \int_\Omega f_n(u(x))\, dx,$$

where $\Omega \subset \mathbb{R}^N$ is a measurable set. Then, the sequence $(F_n)_{n\in\mathbb{N}}$ Γ-converges to some functional $F : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ with respect to the weak topology of L^p if and only if the sequence $(f_n^c)_{n\in\mathbb{N}}$ Γ-converges to some $f : \mathbb{R}^M \to \mathbb{R}$ and in that case we have

$$F(u) = \int_\Omega f(u(x))\, dx,$$

for all $u \in L^p(\Omega;\mathbb{R}^M)$.

Proof. First of all, note that Proposition 7.13 tells us that the Γ-limit of the sequence $(F_n)_{n\in\mathbb{N}}$ with respect to the weak topology of L^p is the same as that of $(\mathrm{lsc}(F)_n)_{n\in\mathbb{N}}$. Moreover, thanks to Theorem 6.63 we have that

$$\mathrm{lsc}(F_n)(u) = \int_\Omega f_n^c(u(x))\, dx.$$

Note that we can apply that result thanks to the uniform bound from below (7.16) of the f_n's.

Step 1. Assume that the sequence $(f_n^c)_{n\in\mathbb{N}}$ Γ-converges to some $f : \mathbb{R}^M \to \mathbb{R}$. We need to prove that $(F_n)_{n\in\mathbb{N}}$ Γ-converges with respect to the weak topology of L^p to the functional $F : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ defined as

$$F(u) := \int_\Omega f(u(x))\, dx.$$

Thanks to the uniform growth (7.16) of the f_n's, we will use the sequential characterization provided by Theorem 7.45.

Step 1.1 We start by proving the liminf inequality. Let $u \in L^p(\Omega;\mathbb{R}^M)$ and let $(u_n)_{n\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R}^M)$ such that $u_n \rightharpoonup u$ weakly in L^p. We claim that

$$F(u) \le \liminf_{n\to\infty} \mathrm{lsc}(F_n)(u_n).$$

Since each f_n^c is convex, it is easy to see that also the function f is convex. This implies that the functional F is lower semicontinuous with respect to the weak topology of L^p (see Theorem **??**). Therefore,

$$F(u) \le \liminf_{n\to\infty} F(u_n). \tag{7.17}$$

The idea is to *substitute*, on the right-hand side, $\mathrm{lsc}(F_n)$ in place of F. This is not possible in general, since there is no *uniform* convergence of f_n^c to f that allows us to control the error

$$|F(u_n) - \mathrm{lsc}(F_n)(u_n)|$$

uniformly in $n \in \mathbb{N}$. The idea is the following: let

$$F^k(u) := \int_\Omega T_k(f)(u)\, dx, \qquad F_n^k(u) := \int_\Omega T_k(f_n^c)(u)\, dx,$$

where $T_k(f_n^c)$ and $T_k(f)$ denotes the k-Yosida transform of f_n^c and of f respectively. Once we establish this inequality, we obtain (7.17) by pass to the limit for $k \to \infty$.

First of all, we note that:

(i) For every $k > 0$, there exists $R > 0$ such that $T_k(f_n^c)$ is linear outside of $B(0, R)$, with slope k;

(ii) $T_k(f_n^c)$ and $T_k(f)$ is convex are convex;

(iii) $T_k(f_n^c)$ converges pointwise to $T_k(f)$ as $n \to \infty$.

Indeed, (i) is a consequence of the uniform bound from below on the f_n's (and, in turn, on the f_n^c's); (ii) follows from Proposition 7.49, since both f_n^c and f are convex. Finally, (iii) follow by using the definition of the k-Yosida transform and by using the convergence of infima given by Theorem 7.41, with Φ the identity map, together with the uniform growth condition on the f_n's (and, in turn, on the f_n^c's) which ensures the equicoerciveness of the maps

$$y \mapsto f_n^c(y) + k\mathrm{d}(y, x),$$

for all $x \in \mathbb{R}^M$. Thus, from (i), (ii), and (iii) we obtain that $T_k(f_n^c)$ converges *uniformly* to $T_k f$. In particular, this means that, for each $\varepsilon > 0$, there exists $\overline{k} > 0$ such that, for all $k \geq \overline{k}$ it holds that

$$|T_k(f_n^c)(\xi) - T_k(f^c)(\xi)| \leq \varepsilon, \tag{7.18}$$

for all $\xi \in \mathbb{R}^M$. Thus, for all $k > \overline{k}$, it holds that

$$\begin{aligned}
\liminf_{n\to\infty} F_n^k(u_n) &= \liminf_{n\to\infty} \int_\Omega T_k(f_n^c)(u_n)\, dx \\
&= \liminf_{n\to\infty} \left[\int_\Omega T_k(f^c)(u_n)\, dx + \int_\Omega \left[T_k(f_n^c)(u_n) - T_k(f^c)(u_n) \right] \, dx \right] \\
&\geq \liminf_{n\to\infty} \int_\Omega T_k(f^c)(u_n)\, dx - \varepsilon \\
&\geq \int_\Omega T_k(f^c)(u)\, dx - \varepsilon,
\end{aligned}$$

where in the previous to last step we used (7.18), while last inequality is due to the lower semicontinuity of the functional

$$u \mapsto \int_\Omega T_k(f^c)(u)\, dx,$$

since $T_k(f^c)$ is convex. Thus, by taking the limit as $k \to \infty$ on both sides of the above inequality, thanks to Proposition 7.49, and the Monotone Convergence Theorem, we get

$$F(u) - \varepsilon \leq \liminf_{n\to\infty} F(u_n)$$

Thus, from the arbitrariness of $\varepsilon > 0$, we obtain (7.17).

Step 1.2 We now prove the limsup inequality. We first use a similar argument as that implemented in Step 2 of the proof of Theorem 6.63 to show that it suffices to construct a recovery sequence for a piecewise constant function. Let $u \in L^p(\Omega; \mathbb{R}^M)$, and let $(u_k)_{k\in\mathbb{N}}$ be a sequence of piecewise constant functions that converges strongly to u in L^p. Assume that

$$[\Gamma\text{-}\limsup_{n\to\infty} F_n](u_k) = F(u_k),$$

for all $k \in \mathbb{N}$. Then,

$$\Gamma\text{-}\limsup_{n\to\infty} F_n(u) \le \liminf_{k\to\infty} [\Gamma\text{-}\limsup_{n\to\infty} F_n](u_k) = \liminf_{k\to\infty} F(u_k) = F(u),$$

where the last equality follows from the continuity of F with respect to the strong convergence of L^p. This proves that, if we know how to obtain the limsup inequality for piecewise constant functions, we know how to conclude its validity for all functions in $L^p(\Omega; \mathbb{R}^M)$.

Thus, consider a piecewise constant function u. In particular, thanks to Step 2 of the proof of Theorem 6.63, we can write

$$u(x) = \sum_{j=1}^{k} \xi^j \mathbb{1}_{A^j}(x),$$

where each $A^j \subset \Omega$ is an open set, and they are pairwise disjoint. Since $(f_n^c)_{n\in\mathbb{N}}$ is Γ-converging to f, thanks to Lemma 7.6, for each $j = 1, \dots, k$, there exists a sequence $(\xi_n^j)_{n\in\mathbb{N}}$ with $\xi_n^j \to \xi^j$ as $n \to \infty$, such that

$$\lim_{n\to\infty} f_n^c(\xi_n^j) = f(\xi^j).$$

Define, for each $n \in \mathbb{N}$, the function $v_n \in L^p(\Omega; \mathbb{R}^M)$ as

$$v_n(x) := \sum_{j=1}^{k} \xi_j \mathbb{1}_{A^j}(x).$$

Then, $v_n \to u$ strongly in $L^p(A; \mathbb{R}^M)$, and

$$\lim_{n\to\infty} \operatorname{lsc}(F_n)(v_n) = F(u).$$

This proves the Γ-limsup inequality.

Step 2. Assume that F_n Γ-converges to some functional $F : L^p(\Omega; \mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ with respect to the weak topology of L^p. We need to show that there exists a Borel function $f : \mathbb{R}^M \to \mathbb{R}$ such that

$$F(u) = \int_\Omega f(u(x))\, dx,$$

for all $u \in L^p(\Omega; \mathbb{R}^M)$.

To do that, we consider the sequence $(f_n^c)_{n\in\mathbb{N}}$. Thanks to Theorem 7.14 it is possible to extract a subsequence $(f_{n_k}^c)_{k\in\mathbb{N}}$ which Γ-converges to some Borel function $f : \mathbb{R}^M \to \mathbb{R}$, possibly depending on the chosen subsequence. Thanks to Step 1, we get that

$$F(u) = \int_\Omega f(u(x))\, dx, \tag{7.19}$$

for all $u \in L^p(\Omega; \mathbb{R}^M)$. If we now take another subsequence $(f_{n_i}^c)_{i\in\mathbb{N}}$ that is Γ-converging to some Borel function $g : \mathbb{R}^M \to \mathbb{R}$, we get that

$$F(u) = \int_\Omega g(u(x))\, dx, \tag{7.20}$$

for all $u \in L^p(\Omega;\mathbb{R}^M)$. Thus, from (7.19) and (7.20) we get that $f = g$. This implies that the whole sequence $(f_n^c)_{n\in\mathbb{N}}$ Γ-converges to some Borel function $f : \mathbb{R}^M \to \mathbb{R}$, and that

$$F(u) = \int_\Omega f(u(x))\,dx$$

for all $u \in L^p(\Omega;\mathbb{R}^M)$. This concludes the proof. □

A special case of the above result is when the sequence of integrands $(F_n)_{n\in\mathbb{N}}$ is locally equi-bounded. In that case, the Γ converges reduces to pointwise converges.

Corollary 7.52. *Let $p \in (1,\infty)$, and let $f_n : \mathbb{R}^M \to \mathbb{R}$ be Borel functions such that*

$$\sup_{n\in\mathbb{N}} f_n(0) < +\infty,$$

and such that there exists $c > 0$ for which

$$f_n(\xi) \geq c(|\xi|^p - 1),$$

for all $\xi \in \mathbb{R}^M$. Moreover, assume that for each compact set $K \subset \Omega$ there exists $M_K > 0$ such that

$$f_n(\xi) \leq M$$

for all $\xi \in K$. Consider the functionals $F_n : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ defined as

$$F_n(u) := \int_\Omega f_n(u(x))\,dx,$$

where $\Omega \subset \mathbb{R}^N$ is a measurable set. Then F_n Γ-converges to some functional $F : L^p(\Omega;\mathbb{R}^M) \to \mathbb{R} \cup \{+\infty\}$ with respect to the weak topology of L^p if and only if the sequence $(f_n^c)_{n\in\mathbb{N}}$ converges pointwise to some $f : \mathbb{R}^M \to \mathbb{R}$ and in that case we have

$$F(u) = \int_\Omega f(u(x))\,dx,$$

for all $u \in L^p(\Omega;\mathbb{R}^M)$.

Remark 7.53. The above theorem can be extended to the case where

$$F_n(u) := \int_\Omega f_n(x, u(x))\,dx,$$

for some Charathéodory function $f_n : \mathbb{R}^M \to \mathbb{R}$. In that case, the convergence is of $f_n^c(x,\cdot)$ to $f(x,\cdot)$, for a.e. $x \in \Omega$.

7.7 Integral representation of Γ-limits of integral functionals on $W^{1,p}(\Omega)$

In this section we want to perform similar investigations as those carried out in the previous section, but for integral functionals defined on $W^{1,p}(\Omega)$. For such functionals, the proofs are extremely more involved, and therefore we will not report them in here. The interested reader in the proof of the two results can consult [6] and [3] respectively.

The Γ-limit will be with respect to the strong convergence of L^p for $p \in (1, +\infty)$. The reason being that usually the bounds on the functionals ensure only compactness in that topology. Moreover, for functionals depending on the gradient, usually also a uniform upper bound on the integrands is necessary. For this reason, we introduce the following class. First of all, we deal with the issue of integral representation of the Γ limit.

Theorem 7.54. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set. Let $1 < p < +\infty$ and consider Borel functions $f_n : \Omega \times \mathbb{R}^M \times \mathbb{R}^{N\times M} \to [0, \infty)$ such that there exists $c > 0$ for which*

$$\frac{1}{c}|\xi|^p \leq f_n(x, z, \xi) \leq c|\xi|^p$$

for almost all $x \in \Omega$, and for all $z \in \mathbb{R}^M$ and $\xi \in \mathbb{R}^{N\times M}$. Consider the functionals $F_n : W^{1,p}(\Omega; \mathbb{R}^M) \to [0, +\infty]$ defined as

$$F_n(u) := \int_\Omega f_n(x, u(x), Du(x))\, dx.$$

Then, if F_n Γ-converges to some $F : L^p(\Omega; \mathbb{R}^M) \to [0, +\infty]$ with respect to the L^p topology, then

$$F(u) = \begin{cases} \displaystyle\int_\Omega f(x, u(x), Du(x))\, dx & \text{if } u \in W^{1,p}(\Omega; \mathbb{R}^M), \\ +\infty & \text{else in } L^p(\Omega; \mathbb{R}^M), \end{cases}$$

where $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{N\times M} \to [0, \infty)$ is a Borel function.

For integral functionals the pointwise convergence of hte integrands yields the Γ convergence in L^p of the functionals, but the opposite is not true in general. We will see a counterexample to this in the next section.

Theorem 7.55. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set. Let $1 < p < +\infty$ and consider Borel functions $f_n : \Omega \times \mathbb{R}^M \times \mathbb{R}^{N\times M} \to \mathbb{R}$ such that there exists $c > 0$ for which*

$$\frac{1}{c}\left(|\xi|^p - 1\right) \leq f_n(x, z, \xi) \leq c\left(|\xi|^p + 1\right)$$

for almost all $x \in \Omega$, and for all $z \in \mathbb{R}^M$ and $\xi \in \mathbb{R}^{N\times M}$. Consider the functionals $F_n : W^{1,p}(\Omega; \mathbb{R}^M) \to [0, +\infty]$ defined as

$$F_n(u) := \int_\Omega f_n(x, u(x), Du(x))\, dx.$$

Assume that there exists a Borel function $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{N\times M} \to [0,\infty)$ *such that for all* $z \in \mathbb{R}^M$ *and* $\xi \in \mathbb{R}^{N\times M}$ *it holds that* $f_n(\cdot, z, \xi)$ *converge pointwise almost everywhere in* Ω *to* $f(\cdot, z, \xi)$. *Then, the functionals* $F_n : W^{1,p}(\Omega;\mathbb{R}^M) \to [0,+\infty]$ *defined as*

$$F_n(u) := \int_\Omega f_n(x, u(x), Du(x))\, dx$$

Γ *converge with respect to the* L^p *topology to the functional* $F : L^p(\Omega;\mathbb{R}^M) \to [0,+\infty]$ *defined as*

$$F(u) := \int_\Omega f(x, u(x), Du(x)),$$

if $u \in W^{1,p}(\Omega;\mathbb{R}^M)$*, and* $+\infty$ *else.*

Remark 7.56. Note that the opposite is *not* true in general!

In the one dimensional scalar case, it is possible to completely characterize the Γ convergence of the functionals in terms of converges of the integrands, in the same spirit as Theorem 7.51. In order to state the result, we need a notion from convex analysis.

Definition 7.57. Let $g : \mathbb{R} \to \mathbb{R}$ be a function. We define the *conjugate* $g^* : \mathbb{R} \to \mathbb{R}$ of g by

$$g^*(z) := \sup\{zt - g(t) \ : \ t \in \mathbb{R}\}.$$

Theorem 7.58. *Let* $(a,b) \subset \mathbb{R}$*, and* $1 < p < +\infty$. *Consider a sequence of Borel functions* $f_n : (a,b) \times \mathbb{R} \to \mathbb{R}$ *satisfying*

$$\frac{1}{c}\left(|\xi|^p - 1\right) \le f_n(x,\xi) \le c\left(|\xi|^p + 1\right),$$

and let $f(a,b) \times \mathbb{R} \to \mathbb{R}$ *be a Borel function. For each* $n \in \mathbb{N}$ *each open subinterval* $I \subset (a,b)$*, defined the functionals*

$$F_n(u; I) := \int_I f_n(x, u'(x))\, dx,$$

and

$$F(u; I) := \int_I f(x, u'(x))\, dx.$$

Then, the followings are equivalent:

(i) For each open subinterval $I \subset (a,b)$ *the functionals* $F_n(\cdot; I)$ Γ*-converge to the function* $F(\cdot; I)$ *with respect to the strong topology of* L^p*;*

(ii) For all $z \in \mathbb{R}$*, the function* $f^*(\cdot, z)$ *is the limit in the weak* topology of* L^∞ *of the sequence* $(f_n^*(\cdot, z))_{n\in\mathbb{N}}$.

Bibliography

[1] Alt, H. W., Linear functional analysis: An application-oriented introduction, Springer, 2016.

[2] Bandle, C. Notices Amer. Math. Soc. 64 (2017).

[3] Braides, A., Γ-convergence for beginners, Oxford Lecture Series in Mathematics and its Applications 22, Oxford University Press, Oxford, 2002.

[4] Dacorogna, B., Direct Methods in the Calculus of Variations, Applied Mathematical Sciences 78, Springer-Verlag, Berlin, 1989.

[5] Dacorogna, B., Introduction to the Calculus of Variations, 3rd Edition, Imperial College Press, Cambridge, 2015.

[6] Dal Maso, G., An introduction to Γ-convergence, Birkh auser, 1993.

[7] Evans, L., Partial Differential Equations, Graduate Studies in Mathematics 19, American Mathematical Society, Providence, RI, 1998.

[8] Fonseca, I., Leoni, G., Modern methods in the calculus of variations: L^p spaces, Springer Monographs in Mathematics, 2007,

[9] Giaquinta, M., Hildebrandt, S., Calculus of variations I. The Lagrangian formalism, Springer-Verlag, Berlin, 1996.

[10] Jost, J., Li-Jost, X., Calculus of Variations, Cambridge Studies in Advanced Mathematics, 64, Cambridge University Press, Cambridge, 1998.

[11] Kielhöfer, H., Calculus of variations. An introduction to the one-dimensional theory with examples and exercises, Translated from the 2010 German original, Texts in Applied Mathematics 67, Springer, Cham, 2018.

[12] Rindler, F., Calculus of Variations, Universitext, Springer, 2018.

[13] Schmidt, B., Variationsrechnung, Lecture notes, Universität Augsburg, 2013.

[14] Werner, D., Funktionalanalysis, 3rd, revised and extended edition, Springer-Verlag, Berlin, 2000.