Convexity, Γ-Convergence, and Gradient Flows: The Variational Toolkit

Preparatory Workshop on Gradient Flows;
13–17 July 2026, IIT Bombay.
School on Gradient flows;
31 August-11 September 2026, ICTS Bangalore.

Swarnendu Sil *

IISc Bangalore

Acknowledgement

The author warmly thanks Harsha Hutridurga (IIT Bombay) and Saikat Mazumdar (IIT Bombay) for organizing and inviting the author to speak in these two beautiful workshops and Barun Sarkar (IIT Madras), Parthanil Roy (IIT Bombay) and Adimurthi (IIT Kanpur) for stimulating discussions and feedback on the course in IIT Bombay. The author also thanks all the participants of this workshop for patiently and enthusiastically following the lectures.

The author wishes to express special thanks and gratitude to Aditi Anand (IIM Bangalore), Ameya Mane (IIT Kanpur), Jyoti Jindal (IIT Roorkee) and Kshitij Sinha (IIT Bombay) for carefully taking the notes during the lectures and diligently preparing the first typed draft for the part of the notes that were covered in IIT Bombay.

Contents

*swarnendusil@iisc.ac.in

The following topics were covered in IIT Bombay.

---

1 Overview

1.1 Direct methods in the Calculus of Variations

Calculus of variations is concerned with finding minima of a function

$$F : \mathbb{X} \to \mathbb{R}$$

where $\mathbb{X}$ is a topological space. The two main settings we are interested in are the following:

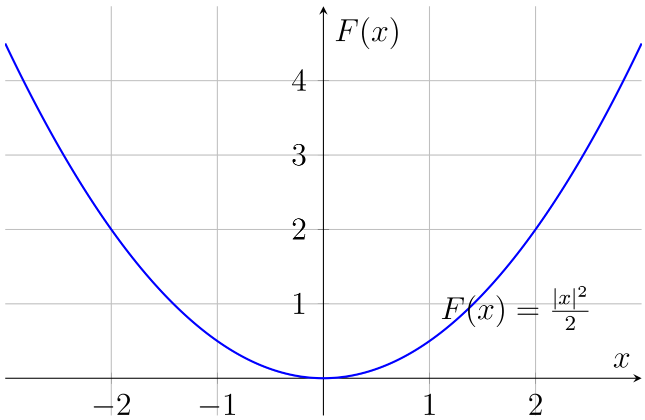
- $\mathbb{X} = (X, d)$ is a metric space, equipped with the metric topology. Typically we assume (X, d) to be a *complete separable metric space*, i.e. a *Polish space*.
- $\mathbb{X} = (X, \sigma(X; X^*))$, where X is a *separable reflexive Banach space*, equipped with the *weak topology* $\sigma(X; X^*)$.

But for the purpose of this overview, $\mathbb{X}$ is an arbitrary Hausdorff topological space. We already know how to do this when $\mathbb{X} = \mathbb{R}^d$, $d \geq 1$ integer. It is nonetheless immensely helpful to keep in mind two toy examples, both are functions $F : \mathbb{R}^d \to \mathbb{R}$ for $d \geq 1$ integer. The first one is the quadratic function

Example 1.

$$F(x) = \frac{1}{2} |x|^2 .$$

When $d = 1$, the graph of the function is sketched below.

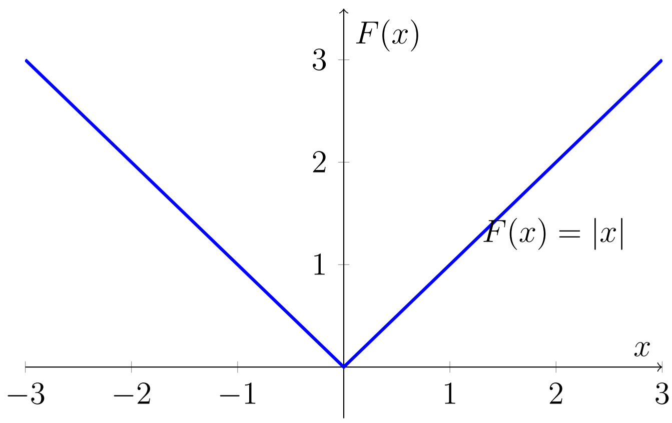


The second is the absolute value function.

Example 2.

$$G(x) = |x| \,.$$

When $d = 1$, the graph of the function is sketched below.



The first is smooth while the other is continuous, but not C^1. A glance at the graphs makes it clear that both functions have a unique minima at $x = 0$. How do we prove this? Usually, one starts by finding all *critical points* of F, i.e. points where either F is not differentiable or F is differentiable and the *derivative* ∇F *vanishes*. Why? The reason is given by the following result.

Proposition 1 (Euler-Lagrange). Let $F : \mathbb{R}^d \to \mathbb{R}$ be C^1. If $\bar{x}$ is a minimizer for F, then $\nabla F(\bar{x}) = 0$.

Proof. Let $\bar{x}$ is a minimizer. Then $F(\bar{x}) \leq F(\bar{x} + tx)$ for all $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$. We set

$$g(t) = F(\bar{x} + tx).$$

Then since F is C^1, it is easy to check that $g : \mathbb{R} \to \mathbb{R}$ is C^1 and we have $g(0) \leq g(t)$ for all $t \in (-\delta, \delta)$. Thus, g has a local minima at $t = 0$. This implies that $g'(0) = 0$. By the chain rule,

$$0 = g'(0) = \frac{\mathrm{d}}{\mathrm{d}t} F(\bar{x} + tx)\bigg|_{t=0} = \langle \nabla F(\bar{x} + tx), x \rangle \bigg|_{t=0} = \langle \nabla F(\bar{x}), x \rangle.$$

Thus, $\langle \nabla F(\bar{x}), x \rangle = 0$ for all $x \in \mathbb{R}^d$. Choosing $x = \nabla F(\bar{x})$, which implies $\|\nabla F(\bar{x})\|^2 = 0$, and therefore $\nabla F(\bar{x}) = 0$. $\square$

The proof above crucially utilized the fact that F is C^1. What can we say when F is not necessarily C^1? To see this, we notice that the above proof used only the fact that $\bar{x}$ is a *critical point* of F, not necessarily a *minima*.

Thus, we now ask: *What is the geometric condition that characterizes a (global) minima?*

Going back to the graphs of our toy examples for $d = 1$, it is clear that

- *The horizontal line passing through* $(\bar{x}, F(\bar{x}))$ touches the graph at $(\bar{x}, F(\bar{x}))$ and is *below the graph of* F *everywhere.*
- The statement above is true *only* for the point $(\bar{x}, F(\bar{x}))$, i.e. for any other point $x \neq \bar{x}$, the horizontal line passing through $(x, F(x))$ *does not lie below the graph* of F.

Thus, we may reasonably hope that this geometric observation somehow characterizes the minima. Before moving on, we notice that for the quadratic, there is *precisely one line* that passes through each point $(x, F(x))$ of the graph of F such that the graph of F *lies entirely to one side of the line.* This however, is not the case for the absolute value function. In this case, for any $x \neq 0$, there is *precisely one such line* passing through $(x, |x|)$. But at $(0, 0)$, there are *many such lines.* Such lines are important enough to deserve a definition.

Definition 1. Let $F : \mathbb{R} \to \mathbb{R}$. A real number $\alpha \in \mathbb{R}$ is called a *subdifferential of* F *at* x, denoted $\alpha \in \partial F(x)$, if

$$F(y) \geq F(x) + \alpha\,(y - x) \qquad \text{for all } y \in \mathbb{R}. \tag{1}$$

The set of all subdifferential of F at x is denoted by $\partial F(x)$. More precisely,

$$\partial F(x) := \{\alpha \in \mathbb{R} : \alpha \text{ satisfies } (1)\}.$$

Remark 1. Note that

$$L(x) = F(x) + \alpha\,(y - x)$$

is a line and the condition is the precise expression for the geometric fact that the *graph of* F *is on one side of the line.* Thus, a subdifferential of F at x is precisely the *possible slope of such a line*, if one exists. Thus, $\partial F(x)$ is the set of all possible slopes of such lines. Note that $\partial F(x)$ can very well be empty, i.e. there need not always exist such a line. As an example, consider the function $F(x) = \sin x$ and check that $\partial F(0) = \emptyset$.

Exercise 1. Let $F(x) = \sin x$ for $x \in \mathbb{R}$. Then $\partial F(0) = \emptyset$.

Exercise 2. Let $F(x) = |x|$ for $x \in \mathbb{R}$. Then

$$\partial F(x) = \begin{cases} +1 & \text{if } x > 0, \\ [-1, +1] & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

When $d > 1$, the lines in the above definition need to be replaced by hyperplanes.

Definition 2 (Subdifferential). Let $F : \mathbb{R}^d \to \mathbb{R}$. A vector $\xi \in \mathbb{R}^d$ is called a subdifferential of F at x, denoted $\xi \in \partial F(x)$, if

$$F(y) \geq F(x) + \langle \xi, y - x \rangle \qquad \text{for all } y \in \mathbb{R}^d. \tag{2}$$

The set of all subdifferential of F at x is denoted by $\partial F(x)$. More precisely,

$$\partial F(x) := \left\{ \xi \in \mathbb{R}^d : \xi \text{ satisfies (2)} \right\}.$$

We say $F : \mathbb{R}^d \to \mathbb{R}$ is *subdifferentiable* at $x \in \mathbb{R}^d$ if $\partial F(x) \neq \emptyset$.

Remark 2. Note that

$$L(x) = F(x) + \langle \xi, y - x \rangle$$

defines an d-dimensional *affine hyperplane* in $\mathbb{R}^{d+1}$ passing through the point $(x, F(x))$ and the condition is the precise expression for the geometric fact that the *graph of F is on one side of the hyperplane*. Thus, a subdifferential of F at x is precisely the *possible 'slope' of such a hyperplane*, if one exists. Thus, $\partial F(x)$ is the set of all possible slopes of such hyperplanes. Note that once again, $\partial F(x)$ can very well be empty, i.e. there need not always exist such a hyperplane.

We can now extend our notion of the *Euler-Lagrange equation* to global minima of possibly nonsmooth functions.

Proposition 2 (Subdifferential Euler-Lagrange). Let $F : \mathbb{R}^d \to \mathbb{R}$. If $\bar{x}$ is a global minimizer for F, then F is subdifferentiable at $\bar{x}$ and $0 \in \partial F(\bar{x})$.

Proof. The proof is extremely easy, thanks to our well-chosen definition of the subdifferential. Indeed, since $\bar{x}$ realizes a global minima, we have

$$F(x) \geq F(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^d.$$

But clearly $F(\bar{x}) = F(\bar{x}) + \langle 0, x - \bar{x} \rangle$ and thus we have

$$F(x) \geq F(\bar{x}) + \langle 0, x - \bar{x} \rangle \qquad \text{for all } x \in \mathbb{R}^d.$$

By definition, this means $0 \in \partial F(\bar{x})$ and as a consequence, F is subdifferentiable at $\bar{x}$. $\square$

Remark 3. Unlike the earlier Euler-Lagrange equation, this one is valid only for a global minima and not for a general critical point of F.

This condition however, is a generalization of the previous one, as proved in the following proposition.

Proposition 3. Suppose that $F : \mathbb{R}^d \to \mathbb{R}$ is both subdifferentiable and differentiable at x. Then

$$\partial F(x) = \{\nabla F(x)\}.$$

Proof. Since F is subdifferentiable at x, there exists $\xi \in \mathbb{R}^d$ such that for every $v \in \mathbb{R}^d$ and any $t \in \mathbb{R}$, we have

$$F(x + tv) \geq F(x) + t\langle \xi, v\rangle.$$

Since F is also differentiable at x, for every $v \in \mathbb{R}^d$ and any $t \in \mathbb{R}$,

$$F(x + tv) = F(x) + t\langle \nabla F(x), v\rangle + o(|t|).$$

Consequently, for $t > 0$ and dividing by t, we have

$$\langle \nabla F(x) - \xi, v\rangle + \frac{o(|t|)}{t} \geq 0.$$

Letting $t \to 0^+$ gives

$$\langle \nabla F(x) - \xi, v\rangle \geq 0.$$

On the other hand, taking $t < 0$, the same argument gives

$$\langle \nabla F(x) - \xi, v\rangle \leq 0.$$

Therefore

$$\langle \nabla F(x) - \xi, v\rangle = 0 \qquad \text{for every } v \in \mathbb{R}^d.$$

Choosing $v = \nabla F(x) - \xi$, we deduce

$$\xi = \nabla F(x).$$

Since $\xi \in \partial F(x)$ was arbitrary, we conclude that

$$\partial F(x) = \{\nabla F(x)\}.$$

□

We have seen how the *first order condition*, called (called the *Euler-Lagrange equation*) arises in connection to minimization. However, when $\mathbb{X}$ is an infinite dimensional space, say an infinite dimensional Banach space, even for C^1 functions, typically the *Euler-Lagrange equation* $DF(x) = 0$ gives us a differential (indeed, typically a PDE) equation instead of an algebraic one which is hard to solve. Instead, we want to go the other way. We want methods to directly establish the existence of a minimizer and then show, if possible, that the minimizer satisfies the Euler-Lagrange PDE. For these reason, these methods are called *direct methods in the calculus of variations.*

Let us try to prove the existence of a minima for these functions in a way which may generalize beyond finite dimensional Euclidean spaces.

Proposition 4. The minimization problem

$$m := \inf_{\mathbb{R}} G,$$

where $G(x) = |x|$, admits a minimizer.

Proof. Clearly, $G(x) \geq 0$ for all $x \in \mathbb{R}$ and hence $m \geq 0$. Let $\{x_k\}_{k\in\mathbb{N}}$ be a *minimizing sequence*, i.e.

$$G(x_k) \to m \qquad \text{as } k \to \infty.$$

Discarding finitely many terms if neceesary, we have $|x_k| = G(x_k) \leq m + 1$ for all k. Thus $\{x_k\}_{k\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}$. By Bolzano-Weierstrass theorem, up to the extraction of a subsequence that we do not relabel, we can assume $x_k \to x$ for some $x \in \mathbb{R}$. Since $t \mapsto |t|$ is continuous and in particular, lower semicontinuous, we have

$$m \leq G(x) \leq \liminf_{k\to\infty} G(x_k) = \lim_{k\to\infty} G(x_k) = m.$$

This proves $G(x) = m$ and thus the problem admits a minimizer. $\square$

Exercise 3. Prove in a similar manner that the minimization problem $m := \inf\limits_{\mathbb{R}} F$, where $F(x) = x^2/2$, admits a minimizer.

Later in Section 2, we shall generalize these methods to far more nontrivial settings.

Although we are really interested in minimizing functions $F : \mathbb{X} \to \mathbb{R}$, it would be convenient for us to allow F to take the value $+\infty$ and thus we shall consider minimization problems for functions

$$F : \mathbb{X} \to (-\infty, +\infty].$$

The reason for this is the following. We shall also consider a family of minimization problems where F and X varies. Observe that given an ambient space $\mathbb{Y}$ such that $\mathbb{X} \subset \mathbb{Y}$, we can always replace the problem of finding a minimizer for $\inf\limits_{y\in\mathbb{X}} F(y)$ by one for $\inf\limits_{y\in\mathbb{Y}} \bar{F}(y)$, where

$$\bar{F}(y) := \begin{cases} F(y) & \text{if } y \in \mathbb{X}, \\ +\infty & \text{if } y \in \mathbb{Y} \setminus \mathbb{X}. \end{cases}$$

Thus, allowing F to take the value $+\infty$ enables us to only consider family of problems where $\mathbb{X}$ is fixed and only F varies.

1.2 Γ-convergence

What kind of sequence of minimization problems do we have in mind? Again as toy examples, consider the following functions, once again real-valued functions defined on $\mathbb{R}$.

Example 3.

$$F_n(x) = \frac{1}{2}x^2 + \frac{1}{n}\sin(nx).$$

Example 4.

$$G_n(x) = |x| + \frac{1}{n}\sin(nx).$$

Example 5.

$$H_n(x) = \sqrt{x^2 + \frac{1}{n}} + \frac{1}{n}\sin(nx).$$

Exercise 4. Sketch the graphs of the three family of functions given above.

From the graphs of the families, it should be clear that if we ignore the 'wiggles', the graphs of the first and the second families look very much like the quadratic and the absolute value function we sketched at the beginning. Also, the sequence $\left\{\sqrt{x^2+\frac{1}{n}}\right\}_{n\in\mathbb{N}}$ is really a smooth approximation of $|x|$. Thus, one might expect that the minimizers of F_n might converge to the minimizer of F and the minimizers of G_n and H_n might converge to the minimizer of G. This is indeed true. Moreover, studying problems of these kind are often useful. In some cases, finding a minimizer for a given function F may be difficult, but we can find minimizers for a sequence F_n which converges to F in a suitable sense, which allows us to deduce the existence of a minimizer for F. There are also cases where the functions F_ns do not admit a minimizer, but a 'minimizing sequence for the family' converges to a minimizer of F. There are even instances, where all we care about are the limiting behavior of sequence of minimizers of F_n (this is typically the case in homogenization problems).

Thus, our main objective in these kind of problems is to find an appropriate notion of convergence

$$F_n \to F$$

such that this would imply, with some mild extra conditions if necessary, that

- *Convergence of the minimum value*:

$$\inf_{\mathbb{X}} F_n \to \inf_{\mathbb{X}} F.$$

- *Convergence of the minimizers*:

$$\arg\min_{\mathbb{X}} F_n \to \arg\min_{\mathbb{X}} F.$$

We shall se that there is a suitable notion of convergence that satisfies these requirements. This known as Γ-*convergence* or *variational convergence*. If a sequence F_n converges to F in this sense, we say F_n Γ-*converges to* F and denote this symbolically as

$$F_n \xrightarrow{\Gamma} F.$$

We shall discuss Γ-convergence later in more detail.

1.3 Gradient flows

The other topic we are interested in discussing is the notion of a *gradient flow* associated to F. This is the condition

$$0 \in \dot{x}(t) + \partial F(x(t)) \qquad \text{or equivalently} \qquad -\dot{x}(t) \in \partial F(x(t)),$$

where we are interested in finding a continuous $x : [0,T) \to \mathbb{R}^d$ for some $T \in (0,+\infty]$, such that $x(t)$ is differentiable in $(0,T)$ and $\dot{x}(t)$ denotes the derivative, i.e. the tangent vector to the curve $x(t)$ at t. By our earlier discussions, if F is both differentiable and subdifferentiable at each $x \in \mathbb{R}^d$, the gradient flow equation becomes

$$\dot{x}(t) = -\nabla F(x(t)).$$

Thus, we are looking for a curve, such that the tangent to the curve is along the (negative of) direction of the gradient of F. In other words, the particle (whose position is given by $x(t)$ at time t) *'flows' along the gradient of F* — hence the name, *gradient flow*. But why do we even want to flow along the gradient in the first place? To understand the reason, first observe that if our curve $x(t)$ *hits a global minima of F at time $t = \bar{t}$*, then the curve *stays there for eternity*. To see this, first we deduce the general fact that the *gradient flow decreases 'energy'*.

Proposition 5 (Energy dissipation). Let $F : \mathbb{R}^d \to \mathbb{R}$ be C^1 and assume F is subdifferentiable at each $x \in \mathbb{R}$. Let $\{x(t)\}_{t\geq 0}$ be a C^1 curve which is a gradient flow of F. Then

$$F(x(t)) \leq F(x(0)) \qquad \text{for all } t \geq 0.$$

Remark 4. The 'energy' at time t is $F(x(t))$.

Proof. Since F and $x(t)$ are both C^1, so is $F \circ x$ and thus, by the fundamental theorem of calculus, we have

$$\begin{aligned} F(x(t)) - F(x(0)) &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}\tau} \left[F(x(\tau))\right] \, \mathrm{d}\tau \\ &= \int_0^t \left\langle \nabla F(x(\tau)), \frac{\mathrm{d}}{\mathrm{d}\tau} x(\tau) \right\rangle \, \mathrm{d}\tau = -\int_0^t \left|\nabla F(x(\tau))\right|^2 \, \mathrm{d}\tau \leq 0. \end{aligned}$$

This proves the result. □

As a consequence, we can infer that a *gradient flow must get stuck at a global minima.*

Proposition 6. Let $F : \mathbb{R}^d \to \mathbb{R}$ be C^1 and assume F is subdifferentiable at each $x \in \mathbb{R}$. Let $\{x(t)\}_{t\geq 0}$ be a C^1 curve which is a gradient flow of F. Let $\bar{x}$ be a global minima of F. If $x(\bar{t}) = \bar{x}$ for some $\bar{t} \geq 0$, then $x(t) = \bar{x}$ for all $t \geq \bar{t}$.

Proof. By the same argument used in the last proposition, for any $t \geq \bar{t}$, we have

$$F(x(t)) - F(x(\bar{t})) = -\int_{\bar{t}}^t \left|\nabla F(x(\tau))\right|^2 \, \mathrm{d}\tau \leq 0.$$

Since $\bar{x}$ is a global minima, this implies $F(x(t)) = F(x(\bar{t}))$ for all $t \geq \bar{t}$. But this implies

$$0 = F(x(t)) - F(x(\bar{t})) = -\int_{\bar{t}}^t \left|\nabla F(x(\tau))\right|^2 \, \mathrm{d}\tau \qquad \text{for all } t > \bar{t}$$

and thus, we deduce

$$\nabla F(x(\tau)) \equiv 0 \qquad \text{for } \tau \in [t, \bar{t}].$$

Using the gradient flow equation, this implies $\dot{x}(\tau) = -\nabla F(x(\tau)) \equiv 0$ for $\tau \in [t, \bar{t}]$. Thus, $x(t) \equiv$ constant in $[t, \bar{t}]$ and as $x(\bar{t}) = \bar{x}$, we must have $x(t) \equiv \bar{x}$. □

From the last two propositions, we see that, at least in this simple setting, a gradient flow always decreases energy (in fact, strictly decreases energy unless the flow is 'stuck') and if the flow hits a global minima, gets stuck there. Thus, we can reasonably hope that under reasonable conditions, a gradient flow would allow us to 'flow towards the minima', starting from an arbitrary point. In fact, in principle, a gradient flow is the most efficient possible mechanism to flow towards the minima and this is the reason a gradient flow is often called a *steepest descent flow*.

Indeed, if we want to decrease the energy along a curve, we want the curve to go in the direction where energy decreases the fastest. Let $v \in \mathbb{R}^d$ be a vector. Then the change of energy, i.e. $F(x)$ along the direction v at x is given by the directional derivative

$$\langle \nabla F(x), v \rangle .$$

Since we are interested only in the direction given by the vector v, we assume $|v| = 1$. Now by Cauchy-Schwarz inequality, we have

$$|\langle \nabla F(x), v \rangle| \leq |\nabla F(x)|\, |v| = |\nabla F(x)| .$$

Thus, we deduce

$$-|\nabla F(x)| \leq \langle \nabla F(x), v \rangle \leq |\nabla F(x)| .$$

Hence, the change is energy is most negative in the direction v such that $\langle \nabla F(x), v \rangle = -|\nabla F(x)|$, i.e. along the vector

$$v = -\frac{\nabla F(x)}{|\nabla F(x)|}.$$

If time permits, we shall later study the gradient flow (more correctly, the subdifferential flow)

$$0 \in \dot{x}(t) + \partial F(x(t))$$

for a given $F : \mathbb{X} \to \mathbb{R}$, where $\mathbb{X}$ is a Hilbert space and also explore the connection of this flow with the sequence of flows

$$0 \in \dot{x}(t) + \partial F_n(x(t)),$$

when $F_n \xrightarrow{\Gamma} F$.

2 Direct method in the Calculus of Variations

2.1 Lower semicontinuity and coercivity notions

Now we discuss the direct methods in the calculus of variations in the general setting. We consider minimization problems for functions $F : \mathbb{X} \to (-\infty, +\infty]$, where $\mathbb{X}$ is a topological space. We first recall a few topological notions.

Definition 3. Let $\mathbb{X}$ be a topological space.

(i) For any $x \in \mathbb{X}$, $\mathcal{N}(x)$ denotes the set of open neighborhoods of x in $\mathbb{X}$.

(ii) $\mathbb{X}$ is called *first countable* if $\mathcal{N}(x)$ has a countable local base for each $x \in \mathbb{X}$.

(iii) $\mathbb{X}$ is called *second countable* if (the topology on) $\mathbb{X}$ has a countable local base.

(iv) A non-empty subset $\mathbb{Y} \subset \mathbb{X}$ is called *sequentially closed* in $\mathbb{X}$ if $x \in \mathbb{Y}$ whenever there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{Y}$ such that $x_n \to x$ in $\mathbb{X}$.

(v) A non-empty subset $\mathbb{Y} \subset \mathbb{X}$ is called *sequentially open* if $\mathbb{X} \setminus \mathbb{Y}$ is sequentially open in $\mathbb{X}$.

(vi) A non-empty subset $\mathbb{K} \subset \mathbb{X}$ is called *sequentially compact* if every sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{K}$ has a convergent subsequence that converges to a point in $\mathbb{K}$.

(vii) A non-empty subset $\mathbb{A} \subset \mathbb{X}$ is called *sequentially pre-compact* if its closure $\overline{\mathbb{A}}$ is sequentially compact in $\mathbb{X}$.

Reviewing the proof of Proposition 4, we see that all we have used are the following two properties:

- *Coercivity*: We have used that the closed sublevel sets $\{G(x) \le t\}$ are sequentially compact for any $t \in \mathbb{R}$, due to the Bolzano-Weierstrass theorem.
- *Lower semicontinuity*: We have used that our function G is sequentially lower semicontinuous, i.e.,

$$G(x) \le \liminf_{n\to\infty} G(x_n)$$

  for every sequence $x_n \to x$.

In view of these, we are led to the following definitions.

Definition 4 (Coercive function). A function $F : \mathbb{X} \to (-\infty, +\infty]$ is said to be *coercive* (respectively, *sequentially coercive*) if for each $t \in \mathbb{R}$, closure of the set $\{x \in \mathbb{X} : F(x) \le t\}$ is compact (respectively, sequentially compact) in $\mathbb{X}$.

Often, a slightly weaker condition also suffices.

Definition 5 (Mildly Coercive function). A function $F : \mathbb{X} \to (-\infty, +\infty]$ is said to be mildly coercive (respectively, *mildly sequentially coercive*) if there exists a compact (respectively, sequentially compact) $\mathbb{K} \subset \mathbb{X}$ such that

$$\inf_{x\in\mathbb{X}} F(x) = \inf_{x\in\mathbb{K}} F(x).$$

Definition 6 (Lower semicontinuity). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function.

(a) F is said to be *lower semicontinuous* (abbreviated as lsc) at $x \in \mathbb{X}$ if for given $t \in \mathbb{R}$ with $F(x) > t$, there exists a neighborhood $U \in \mathcal{N}(x)$ such that $F(y) > t$ for $y \in U$. F is said to be lower semicontinuous in $\mathbb{X}$ if F is lsc at each $x \in \mathbb{X}$.

(b) F is said to be *sequentially lower semicontinuous* (abbreviated as slsc) at $x \in \mathbb{X}$ if

$$F(x) \le \liminf_{n\to\infty} F(x_n)$$

for any sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{X}$ such that $x_n \to x$ in $\mathbb{X}$. F is said to be sequentially lower semicontinuous in $\mathbb{X}$ if F is slsc at each $x \in \mathbb{X}$.

Remark 5. (i) If F is lsc at $x \in \mathbb{X}$, then F is slsc.

(ii) In general, the converse of the first remark does not hold. However, if $\mathbb{X}$ is first countable, then F is slsc if and only F is lsc.

Proposition 7. If F is lsc at $x \in \mathbb{X}$, then

$$F(x) \leq \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y).$$

Proof. If the statement does not hold, then we can choose $t \in \mathbb{R}$ such that

$$F(x) > t > \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y).$$

Thus, there does not exist any $U \in \mathcal{N}(x)$ such that $F(y) > t$ for all $y \in U$. This contradicts the fact that F is lsc at x. $\square$

As a consequence, we have a formula for a lsc function.

Theorem 1. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be lsc. Then

$$F(x) = \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y). \tag{3}$$

Proof. By the last proposition, we have

$$F(x) \leq \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y).$$

But we also have

$$F(x) \geq \inf_{y \in U} F(y) \qquad \text{for any } U \in \mathcal{N}(x).$$

Taking supremum over all $U \in \mathcal{N}(x)$ yields the reverse inequality and completes the proof. $\square$

Remark 6. The right hand side of (3) defines a lsc function anyway (see Exercise 5 below). In fact, the formula gives the largest such function. More precisely, if $H : \mathbb{X} \to (-\infty, +\infty]$ is lsc and satisfies $H \leq F$ on $\mathbb{X}$, then

$$H \leq G \quad \text{on } \mathbb{X}.$$

Proof. Since H is lsc and $H \leq F$ on $\mathbb{X}$, for every $x \in \mathbb{X}$, we have

$$H(x) = \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} H(y) \leq \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y) = G(x).$$

Hence, $H \leq G$ on $\mathbb{X}$. $\square$

Exercise 5. For any function $F : \mathbb{X} \to (-\infty, +\infty]$, define

$$G(x) := \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y), \qquad x \in \mathbb{X}.$$

Then G is lsc on $\mathbb{X}$.

Definition 7 (Minorant and envelope). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function. A function Let $G : \mathbb{X} \to (-\infty, +\infty]$ is called a *minorant* of F if $G(x) \le F(x)$ for all $x \in \mathbb{X}$. G is called a *lower envelope* if it is largest minorant (or maximal as a minorant), i.e. any other minorant of F is also a minorant of G.

Remark 7. (i) In the context of the direct methods, we often drop the term 'lower' and just say an *envelope* of F, with the implicit understanding that this means the lower envelope. We shall use this convention from now on.

(ii) More often than not, one is interested in *the minorants of F satisfying a specific property*, rather than *all* minorants of F. The corresponding envelope is then defined as the largest minorant with the said property. This is very common in the direct methods. (see the lsc envelope below).

Definition 8 (Lower semicontinuous envelope). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function. The *lsc envelope of* F, denoted F_{lsc}, is the largest lower semicontinuous minorant of F. More precisely, F_{lsc} is a lower semicontinuous minorant of F with the property that any other lower semicontinuous minorant of F must also be a minorant of F_{lsc}.

Remark 8. The maximality property uniquely characterizes the lsc envelope. Indeed, we can write down a formula for F_{lsc}, proving uniqueness. Also, F and F_{lsc} has the same infimum. Thus, if we are only interested in the infimum value, one can always replace F with F_{lsc}. This is common in the context of the direct methods and usually collectively called *relaxation* (the minimization problem for F_{lsc} is called the *relaxed problem*, as opposed to the *original problem* for F).

Exercise 6. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function. Then we have

(a)
$$\inf_{x \in \mathbb{X}} F(x) = \inf_{x \in \mathbb{X}} F_{\text{lsc}}(x).$$

(b)
$$F_{\text{lsc}}(x) = \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y). \qquad \text{for all } x \in \mathbb{X}.$$

We now define the epigraph of F.

Definition 9. The *epigraph* of a function $F : \mathbb{X} \to (-\infty, +\infty]$ is defined as

$$\operatorname{epi} F := \{(x, t) \in \mathbb{X} \times \mathbb{R} \mid F(x) \le t\}.$$

Theorem 2. Let $F : \mathbb{X} \to (-\infty, +\infty]$. The following statements are equivalent:

1. F is lsc (resp, slsc) on $\mathbb{X}$.
2. For every $t \in \mathbb{R}$, the set $\{F > t\}$ is open (resp, sequentially open) in $\mathbb{X}$.
3. For every $t \in \mathbb{R}$, the set $\{F \le t\}$ is closed (resp, sequentially closed) in $\mathbb{X}$.
4. $\operatorname{epi} F$ is closed (resp, sequentially closed) in $\mathbb{X} \times \mathbb{R}$.

Proof. (1) $\Rightarrow$ (2). Fix $t \in \mathbb{R}$ and let $x \in \{F > t\}$, so that $F(x) > t$. Since F is lsc at x, by definition there exists a neighborhood $U \in \mathcal{N}(x)$ such that

$$F(y) > t \qquad \text{for all } y \in U.$$

Thus $U \subset \{F > t\}$ and consequently, $\{F > t\}$ is open in X.

(2) $\Leftrightarrow$ (3). Fix $t \in \mathbb{R}$. Since

$$\{F \leq t\} = X \setminus \{F > t\},$$

$\{F > t\}$ is open if and only if $\{F \leq t\}$ is closed.

(2) $\Rightarrow$ (4). Suppose $\{F > t\}$ is open for every $t \in \mathbb{R}$. We need to show that the complement of $\operatorname{epi} F$ in $\mathbb{X} \times \mathbb{R}$ is open. Since

$$\operatorname{epi} F := \{(x, r) \in \mathbb{X} \times \mathbb{R} : F(x) \leq r\},$$

we have

$$(\operatorname{epi} F)^C = \{(x, r) : F(x) > r\}.$$

Let $(x_0, r_0) \in (\operatorname{epi} F)^C$, i.e. $F(x_0) > r_0$. Now, choose t such that

$$r_0 < t < F(x_0).$$

Thus, $x_0 \in \{F > t\}$. Let

$$W := \{F > t\} \times (-\infty, t),$$

which is clearly open in $\mathbb{X} \times \mathbb{R}$. Let $(x, r) \in W$. Then $F(x) > t$ and $r < t$, so $F(x) > r$, i.e. $(x, r) \notin \operatorname{epi} F$. Thus, $W \subseteq (\operatorname{epi} F)^C$, so that $(\operatorname{epi} F)^C$ is open. Hence, $\operatorname{epi} F$ is closed.

(4) $\Rightarrow$ (1). Let $x \in X$ and let $t \in \mathbb{R}$ with $F(x) > t$. Then

$$(x, t) \notin \operatorname{epi} F.$$

By (4), there exist an open set $U \in \mathcal{N}(x)$ in $\mathbb{X}$ and an open interval $V \ni t$ in $\mathbb{R}$ such that

$$U \times V \subset (\mathbb{X} \times \mathbb{R}) \setminus \operatorname{epi} F,$$

so that for every $y \in U$ we have $(y, t) \in U \times V$ which gives

$$F(y) > t \qquad \text{for all } y \in U.$$

Thus, since $x \in \mathbb{X}$ and $t < F(x)$ were arbitrary, F is lower semi-continuous on $\mathbb{X}$. □

Proposition 8. (a) If F and G are lsc (resp, slsc), then $F + G$ is lsc (resp, slsc).

(b) If F_i is lsc (resp, slsc) at $x \in \mathbb{X}$ for each $i \in I$ and

$$F(x) := \sup_{i \in I} F_i(x),$$

then F is lsc (resp, slsc) at x.

(c) If F_i is lsc (resp, slsc) at $x \in \mathbb{X}$ for each $i \in I$, for I finite, and

$$F(x) := \inf_{i \in I} F_i(x),$$

then F is lsc (resp, slsc) at x.

Exercise 7. Prove Proposition 8.

Example 6. (a) $\mathbb{1}_A$ is lsc if and only if A is open.

(b) Consider the indicator function of the set A, which is denoted by

$$\iota_A(x) := \begin{cases} 0, & x \in A \\ +\infty, & x \in X \setminus A. \end{cases}$$

Then ι_A is lsc if and only if A is closed.

2.2 Tonelli's theorem and the role of weak topology

We are now ready to prove our main theorem concerning existence of minimizers.

Theorem 3 (Tonelli's Theorem). Let $\mathbb{X}$ be a topological space. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be sequentially coercive and slsc. Then F attains its minimum.

Proof. Let $m = \inf_{x \in \mathbb{X}} F(x)$. Then there exists a minimizing sequence $\{x_k\} \subset \mathbb{X}$ such that $F(x_k) \to m$. Since $F(x_k) \to m$, there exists $N \in \mathbb{N}$ such that for all $k \geq N$, $F(x_k) < m+1$. This implies

$$\{x_k\}_{k \in \mathbb{N}} \subset \{x \in \mathbb{X} : F(x) \leq m+1\} \subset \overline{\{x \in \mathbb{X} : F(x) \leq m+1\}}^{\mathbb{X}}.$$

Now, since F is sequentially coercive, the closure of the closed sublevel set

$$\overline{\{x \in \mathbb{X} : F(x) \leq m+1\}}^{\mathbb{X}}$$

is sequentially compact. Thus, we obtain a convergent subsequence $\{x_{k_l}\} \subset \{x_k\}$ and some $\bar{x} \in \mathbb{X}$ such that $x_{k_l} \to \bar{x}$ in $\mathbb{X}$. Now, since F is slsc, we have:

$$F(\bar{x}) \leq \liminf_{l \to \infty} F(x_{k_l}) = m$$

Since $m = \inf_{x \in \mathbb{X}} F(x)$, we must also have $m \leq F(\bar{x})$. From $F(\bar{x}) \leq m$ and $m \leq F(\bar{x})$, we conclude $F(\bar{x}) = m$. Thus, $\bar{x}$ is a minimizer for F. $\square$

Theorem 3 is the cornerstone of the direct methods in the calculus of variations. But in practice, it is quite difficult to verify the hypotheses. The sequential coercivity is never verified when $\mathbb{X}$ is an infinite dimensional Banach space equipped with the norm topology.

Exercise 8. Let $\mathbb{X}$ be a Banach space and let $F(x) = \|x\|_{\mathbb{X}}$. Show that F is sequentially coercive on $\mathbb{X}$ if and only if $\mathbb{X}$ is finite dimensional.

However, one can restore sequential coercivity by *passing to a weaker topology.*

The rest of the discussion in this section heavily uses a number of facts about weak and weak-$*$ topologies on normed linear spaces without proof. The proofs, along with a thorough discussion regarding these topologies can be found in [2, Chapter 3].

Exercise 9. Let $\mathbb{X}$ be a normed linear space. Then strongly closed (norm-)bounded subsets of $\mathbb{X}^*$ are weak-$*$ pre-compact, i.e. pre-compact with respect to the weak-$*$ topology $\sigma(\mathbb{X}^*;\mathbb{X})$ (cf. [2, Theorem 3.16]). Moreover, if $\mathbb{X}$ is *reflexive Banach* space, then strongly closed (norm-)bounded subsets of $\mathbb{X}$ are weakly pre-compact, i.e. pre-compact with respect to the weak topology $\sigma(\mathbb{X};\mathbb{X}^*)$ (cf. [2, Theorem 3.17]).

However, note that *pre-compactness does **not** imply sequential pre-compactness*, as neither the weak topology nor the weak-$*$ topology is first countable if $\mathbb{X}$ and $\mathbb{X}^*$ are infinite dimensional. However, if $\mathbb{X}$ is a *separable reflexive Banach* space, then the weak topology restricted to (norm-)bounded subsets of $\mathbb{X}$ are metrizable and hence first countable.

Exercise 10. Let $\mathbb{X}$ is a separable reflexive Banach space. Show that strongly closed (norm-)bounded subsets of $\mathbb{X}$ are sequentially pre-compact with respect to the weak topology $\sigma(\mathbb{X};\mathbb{X}^*)$.

Remark 9. The result is actually true even without the separability hypothesis. See [2, Theorem 3.18]

Theorem 4. Let $\mathbb{X}$ be a Banach space. If F is sequentially coercive with respect to $(\mathbb{X}, \sigma(\mathbb{X}, \mathbb{X}^*))$, then

$$F(x) \to +\infty \quad \text{as} \quad \|x\|_{\mathbb{X}} \to +\infty. \tag{$*$}$$

Proof. Suppose, to the contrary, that there exists a sequence $\{x_n\} \subset \mathbb{X}$ such that

$$\|x_n\|_{\mathbb{X}} \to +\infty \qquad \text{as } n \to \infty,$$

such that $F(x_n) \not\to +\infty$. Hence, there exist $t_0 \in \mathbb{R}$ and a subsequence (denoted by the same symbol $\{x_n\}$) such that

$$F(x_n) \le t_0 \qquad \text{for all } n,$$

which further implies

$$\{x_n\} \subset \{F \le t_0\} \subset \overline{\{F \le t_0\}}^{\sigma(\mathbb{X},\mathbb{X}^*)}.$$

Since F is sequentially coercive with respect to $\sigma(\mathbb{X}, \mathbb{X}^*)$, then for every $t \in \mathbb{R}$, the

$$\overline{\{x \in \mathbb{X} : F(x) \le t\}}^{\sigma(\mathbb{X},\mathbb{X}^*)}$$

is sequentially compact. Thus, there exists a subsequence $\{x_n\}$, and some $x \in \mathbb{X}$ such that

$$x_n \rightharpoonup x \quad \text{in } \sigma(\mathbb{X}, \mathbb{X}^*),$$

that is,

$$\langle x^*, x_n \rangle \to \langle x^*, x \rangle \text{ for all } x^* \in \mathbb{X}^*.$$

Since every weakly convergent sequence in a Banach space is bounded by the uniform boundedness principle, we have

$$\sup_{n \in \mathbb{N}} \|x_n\|_{\mathbb{X}} < +\infty.$$

This contradicts the assumption that

$$\|x_n\|_{\mathbb{X}} \to +\infty.$$

Hence, no such sequence exists, and consequently,

$$F(x) \to +\infty \qquad \text{as } \|x\|_{\mathbb{X}} \to +\infty.$$

This completes the proof. $\square$

Theorem 5. If $\mathbb{X}$ is a separable reflexive Banach space and F satisfies $(*)$, then F is sequentially coercive with respect to $\sigma(\mathbb{X},\mathbb{X}^*)$.

Proof. Since F satisfies $(*)$, for every $t \in \mathbb{R}$, there exists $M = M(t) > 0$ such that

$$\{x \in \mathbb{X} : F(x) \leq t\} \subset \{x \in \mathbb{X} : \|x\|_{\mathbb{X}} \leq M(t)\}.$$

Thus, we deduce

$$\overline{\{x \in \mathbb{X} : F(x) \leq t\}}^{\sigma(\mathbb{X},\mathbb{X}^*)} \subset \overline{\{x \in \mathbb{X} : \|x\|_{\mathbb{X}} \leq M(t)\}}^{\sigma(\mathbb{X},\mathbb{X}^*)}.$$

Since $\{x \in \mathbb{X} : \|x\|_{\mathbb{X}} \leq M(t)\}$ is both strongly closed and norm-bounded, it also is sequentially pre-compact by Exercise 10, as $\mathbb{X}$ is a separable reflexive Banach space. Thus $\overline{\{x \in \mathbb{X} : \|x\|_{\mathbb{X}} \leq M(t)\}}^{\sigma(\mathbb{X},\mathbb{X}^*)}$ is sequentially compact with respect to $\sigma(\mathbb{X},\mathbb{X}^*)$. This implies $\overline{\{x \in \mathbb{X} : F(x) \leq t\}}^{\sigma(\mathbb{X},\mathbb{X}^*)}$ is sequentially compact with respect to $\sigma(\mathbb{X},\mathbb{X}^*)$, being a closed (w.r.t. $\sigma(\mathbb{X},\mathbb{X}^*)$) and thus also sequentially closed, subset of a sequentially compact set. $\square$

2.3 Weak lower semicontinuity and convexity

While *passing to a weaker topology is helpful to restore compactness*, it is in general *detrimental for preserving continuity properties.*

Example 7. Let $u_k = \sin(kx) \in L^2(0, 2\pi)$. Then $u_k \rightharpoonup 0$ in L^2, but $u_k^2 = \sin^2(kx) \rightharpoonup \frac{1}{2}$ in L^2. We infer that the L^2-norm is not weakly continuous.

However, *convex functions* stay lower semicontinuous when we pass from the strong topology to weak topology.

Definition 10. Let $\mathbb{X}$ be a vector space. A set $\mathbb{A} \subset \mathbb{X}$ is said to be convex if $tx+(1-t)y \in \mathbb{A}$ for all $t \in [0,1]$ whenever $x, y \in \mathbb{A}$.

Definition 11. Let $\mathbb{A} \subset \mathbb{X}$ be a convex subset. We say a function $F : \mathbb{A} \subset \mathbb{X} \to (-\infty, +\infty]$ is *convex* if

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) \quad \text{for all } t \in [0,1] \text{ and } x, y \in \mathbb{A}. \tag{4}$$

We define the *effective domain of* F as the set

$$\operatorname{dom} F := \{x \in \mathbb{X} : F(x) < +\infty\}. \tag{5}$$

F is called *proper* if $\operatorname{dom} F \neq \emptyset$, i.e. if $F \not\equiv +\infty$. F is called *strictly convex* if

$$F(tx + (1-t)y) < tF(x) + (1-t)F(y) \quad \text{for all } t \in (0,1) \text{ and } x, y \in \operatorname{dom} F. \tag{6}$$

Remark 10. Here we are using the convention that $\lambda \cdot (+\infty) = +\infty$ for any $\lambda > 0$, $0 \cdot (+\infty) = 0$, $(+\infty) + (+\infty) = +\infty$ and $(+\infty) \leq +\infty$. A slightly simpler (but entirely equivalent) definition of convexity, which does not require these '$+\infty$ calculus conventions', can be obtained by requiring the inequality (4) to hold only when the right hand side is finite (or equivalently, only for $x, y \in \operatorname{dom} F$).

In view of the last remark, one might wonder why we allow convex functions to take the value $+\infty$ at all. The reason is that this allows us to work only with *convex functions defined over the entire space* $\mathbb{X}$.

Exercise 11. Let $\mathbb{A} \subset \mathbb{X}$ be any non-empty subset and let $F : \mathbb{A} \subset \mathbb{X} \to (-\infty, +\infty]$ be a given function. Set

$$\bar{F}(x) := \begin{cases} F(x) & \text{if } x \in \mathbb{A}, \\ +\infty & \text{if } x \in \mathbb{X} \setminus \mathbb{A}. \end{cases}$$

Show that $\bar{F} : \mathbb{X} \to (-\infty, +\infty]$ is convex if and only if $\mathbb{A} \subset \mathbb{X}$ is convex and $F : \mathbb{A} \subset \mathbb{X} \to (-\infty, +\infty]$ is convex.

Furthermore, show that

$$\operatorname{dom} \bar{F} = \operatorname{dom} F \qquad \text{and} \qquad \inf_{x \in \mathbb{X}} \bar{F}(x) = \inf_{x \in \mathbb{A}} F(x).$$

If F is proper, the set of minimizers of F (possibly empty) is also the same as that of $\bar{F}$.

This last Exercise implies that we can forget about the subset $\mathbb{A}$ and from now on, we would often do so without comment.

The importance of convex functions stems from the following result. For a proof, see [2, Corollary 3.9, page 61]

Theorem 6. Let $\mathbb{X}$ be a Banach space and let $F : \mathbb{X} \to (-\infty, +\infty]$ be a convex function. If F is lsc with respect to norm topology, then F is lsc with respect to the weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$.

Summarizing all the discussions above, in view of Tonelli's theorem, we can prove the following result.

Theorem 7. Let $\mathbb{X}$ be a separable reflexive Banach space and let $\mathbb{A} \subset \mathbb{X}$ be a non-empty convex subset. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper, convex, lsc function such that $F(x) \to +\infty$ as $\|x\|_{\mathbb{X}} \to +\infty$. Then has a minimizer in $\mathbb{A}$. Furthermore, if F is strictly convex, then the minimizer is unique.

Proof. We leave the proof of the existence of a minimizer as a simple exercise. For the uniqueness statement, note that since F is proper, we have $m := \inf_{\mathbb{A}} F < +\infty$. Thus, any minimizer must lie in $\operatorname{dom} F$. Also, since F has at least one minimizer, $m > -\infty$ and consequently $m \in \mathbb{R}$. But if $x_1, x_2 \in \operatorname{dom} F$ are two minimizers with $x_1 \neq x_2$, then we have

$$m \leq F(\frac{1}{2}x_1 + \frac{1}{2}x_2) < \frac{1}{2}F(x_1) + \frac{1}{2}F(x_2) = \frac{1}{2}m + \frac{1}{2}m = m.$$

This contradiction proves uniqueness. $\square$

A particularly instructive example is the following.

Theorem 8 (p-Dirichlet energy). Let $\Omega \subset \mathbb{R}^d$ be open and bounded. Let $1 < p < \infty$ and let $g \in L^{p'}(\Omega)$, $h \in W^{1,p}(\Omega)$, $G \in L^{p'}\left(\Omega, \mathbb{R}^d\right)$ be given maps. Then the following minimization problem admits a unique minimizer:

$$m := \inf \left\{ \frac{1}{p} \int_{\Omega} |\nabla u|^p \, \mathrm{d}x - \int_{\Omega} gu \, \mathrm{d}x + \int_{\Omega} \langle G, \nabla u \rangle \, \mathrm{d}x : u \in g + W_0^{1,p}(\Omega) \right\}.$$

Proof. We set

$$\mathbb{A} := h + W_0^{1,p}(\Omega) \qquad \text{and} \qquad \mathbb{X} := W^{1,p}(\Omega).$$

Clearly, $\mathbb{A} \subset \mathbb{X}$ is a linear subspace and thus a convex subset. Also, $\mathbb{X}$ is a separable reflexive Banach space. We define

$$F(u) := \begin{cases} \dfrac{1}{p}\displaystyle\int_\Omega |\nabla u|^p \,\mathrm{d}x - \int_\Omega gu \,\mathrm{d}x + \int_\Omega \langle G, \nabla u\rangle \,\mathrm{d}x & \text{if } u \in \mathbb{A}, \\ +\infty & \text{if } u \in \mathbb{X} \setminus \mathbb{A}. \end{cases}$$

Since $z \mapsto |z|^p$ is strictly convex on $\mathbb{R}^d$ and the last two terms are linear, it is easy to check that F is strictly convex on $\mathbb{A}$. Also, since $g \in L^{p'}(\Omega)$ and $G \in L^{p'}\left(\Omega, \mathbb{R}^d\right)$, it is easy to check using Hölder inequality that $F(h) < +\infty$ and F is continuous and thus also lsc with respect to the strong topology of $\mathbb{X}$. Hence F is proper and also lsc and consequently, slsc with respect to the weak topology on $\mathbb{X}$. Hence it only remains to check F satisfies $(*)$. Now if $u \in A$, then $u - \varphi \in W_0^{1,p}(\Omega)$, and Poincaré inequality yields

$$\int_\Omega |u-h|^p \leq C \int_\Omega |\nabla(u-h)|^p.$$

Thus, we have

$$\begin{aligned} \int_\Omega |u|^p &\leq 2^{(p-1)} \left(\int_\Omega |u-h|^p + \int_\Omega |h|^p \right) \\ &\leq 2^{(p-1)} C \int_\Omega |\nabla(u-h)|^p + 2^{(p-1)} \int_\Omega |h|^p \leq C \left[\int_\Omega |\nabla u|^p + \int_\Omega |\nabla h|^p + \int_\Omega |h|^p \right]. \end{aligned} \tag{7}$$

Now using Young inequality with $\varepsilon > 0$ with exponents p and p', we estimate,

$$\begin{aligned} F(u) &\geq \frac{1}{p} \int_\Omega |\nabla u|^p - \left| \int_\Omega gu \right| - \left| \int_\Omega \langle G, \nabla u \right| \\ &\geq \frac{1}{p} \int_\Omega |\nabla u|^p - \varepsilon \int_\Omega |u|^p - \varepsilon \int_\Omega |\nabla u|^p - C_\varepsilon \left(\int_\Omega |g|^{p'} + \int_\Omega |G|^{p'} \right) \\ &\overset{(7)}{\geq} \left(\frac{1}{p} - \varepsilon - C\varepsilon \right) \int_\Omega |\nabla u|^p - C \left[\int_\Omega |\nabla h|^p + \int_\Omega |h|^p \right] - C_\varepsilon \left(\int_\Omega |g|^{p'} + \int_\Omega |G|^{p'} \right). \end{aligned}$$

Choosing $\varepsilon > 0$ small enough, we see that there exists a $\lambda > 0$ such that

$$F(u) \geq \lambda \int_\Omega |\nabla u|^p - C \left(\int_\Omega |\nabla h|^p + \int_\Omega |h|^p + \int_\Omega |g|^{p'} + \int_\Omega |G|^{p'} \right).$$

Using (7) once again, this implies the existence of a constant $\mu > 0$ such that

$$F(u) \geq \mu \left(\int_\Omega |\nabla u|^p + \int_\Omega |u|^p \right) - C \left(\int_\Omega |\nabla h|^p + \int_\Omega |h|^p + \int_\Omega |g|^{p'} + \int_\Omega |G|^{p'} \right).$$

This clearly implies $(*)$. □

Remark 11. One can also show that unique minimizer $u \in h + W_0^{1,p}(\Omega)$ satisfies the weak form of the Euler-Lagrange equation

$$\int_\Omega \langle |\nabla u|^{p-2}\nabla u, \nabla\phi\rangle \, \mathrm{d}x - \int_\Omega g\phi \, \mathrm{d}x + \int_\Omega \langle G, \nabla\phi\rangle \, \mathrm{d}x = 0 \qquad \text{for all } \phi \in W_0^{1,p}(\Omega)\,.$$

In other words, u is weak solution to the boundary value problem

$$\begin{cases} -\operatorname{div}\left(|\nabla u|^{p-2}\nabla u\right) = g + \operatorname{div} G & \text{in } \Omega, \\ \qquad\qquad\qquad u = h & \text{on } \partial\Omega. \end{cases}$$

A digression: Convexity is so important for the direct methods that we often work with substitutes, when a vector space is structure is unavailable. For example, if $\mathbb{X}$ is a metric space, which is not necessarily a vector space, we can not define convex functions on $\mathbb{X}$. However if $\mathbb{X}$ is a *geodesic space* such that any two points can be connected via a *unique minimizing geodesic*, then one can use the geodesics to 'lift the linear structure from $\mathbb{R}$'.

Definition 12. Let $\mathbb{X}$ be a geodesic space, i.e. a metric space such that any two points can be connected by a geodesic. Then a subset $\mathbb{A} \subset \mathbb{X}$ is said to be *geodesically convex* if for all $x, y \in \mathbb{A}$, there is a unique minimizing geodesic γ that is contained entirely within $\mathbb{A}$. A function $F : \mathbb{A} \subset \mathbb{X} \to (-\infty, +\infty]$ is called *geodesically convex* if $\mathbb{A} \subset \mathbb{X}$ is geodesically convex and $F \circ \gamma : [0, d(x,y)] \to \mathbb{A}$ is convex for any $x, y \in \mathbb{A}$, where γ is the unique unit speed geodesic connecting x and y.

We would not be using this notion in this course. However, you might encounter geodesic convexity and other such notions (e.g. displacement convexity) in the other courses in this workshops.

3 Convex Analysis

Now we study convex functions in greater detail. We start by recalling the relevant definitions.

Definition 13. Let $\mathbb{X}$ be a vector space. A set $\mathbb{A} \subset \mathbb{X}$ is said to be convex if $tx + (1-t)y \in \mathbb{A}$ for all $t \in [0,1]$ whenever $x, y \in \mathbb{A}$.

Definition 14. Let $\mathbb{A} \subset \mathbb{X}$ be a convex subset. We say a function $F : \mathbb{A} \subset \mathbb{X} \to (-\infty, +\infty]$ is *convex* if

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) \quad \text{ for all } t \in [0,1] \text{ and } x, y \in \mathbb{A}. \tag{8}$$

We define the *effective domain of* F as the set

$$\operatorname{dom} F := \{x \in \mathbb{X} : F(x) < +\infty\}\,. \tag{9}$$

F is called *proper* if $\operatorname{dom} F \neq \emptyset$, i.e. if $F \not\equiv +\infty$. F is called *strictly convex* if

$$F(tx + (1-t)y) < tF(x) + (1-t)F(y) \quad \text{ for all } t \in (0,1) \text{ and } x, y \in \operatorname{dom} F. \tag{10}$$

We have the following simple result.

Proposition 9. Let $\mathbb{X}$ be a $\mathbb{R}$-vector space and let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function. Then

(a) If F is convex, then $\{F \leq \lambda\}$ is convex for every $\lambda \in \mathbb{R}$.

(b) F is convex if and only if $\operatorname{epi} F$ is convex in $\mathbb{X} \times \mathbb{R}$.

Remark 12. The converse of statement (a) does not hold. Indeed, if F is convex and ϕ is increasing, then $\phi \circ F$ would have convex sublevel sets, i.e. $\{\phi \circ F \leq \lambda\}$ is convex for every $\lambda \in \mathbb{R}$, but $\phi \circ F$ is not convex in general.

Proof. Proof of (a): Fix $\lambda \in \mathbb{R}$. Let $x, y \in \{F \leq \lambda\}$. Then

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) \leq t\lambda + (1-t)\lambda = \lambda.$$

Thus $\{F \leq \lambda\}$ is convex.

Proof of (b): Let F be convex and let $(x, \lambda), (y, \mu) \in \operatorname{epi} F$. Then

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y) \leq t\lambda + (1-t)\mu.$$

Thus, $(tx + (1-t)y, t\lambda + (1-t)\mu) \in \operatorname{epi} F$ and hence $\operatorname{epi} F$ is convex.

Conversely, let $\operatorname{epi} F$ be convex. But $(x, F(x)), (y, F(y)) \in \operatorname{epi} F$ for every $x, y \in \mathbb{X}$. Thus, $t(x, F(x)) + (1-t)(y, F(y)) \in \operatorname{epi} F$ for every $t \in [0,1]$. Thus, we have

$$F(tx + (1-t)y) \leq tF(x) + (1-t)F(y).$$

Thus, F is convex. □

Note that in view of the fact that in a Banach space, convex subsets are strongly closed if and only if they are weakly closed, we can prove a slightly stronger version of Theorem 6 as a corollary.

Corollary 1. Let $\mathbb{X}$ be a Banach space and let $F : \mathbb{X} \to (-\infty, +\infty]$ be a convex function. Then F is lsc with respect to norm topology if and only if F is lsc with respect to the weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$.

Proof. F is lsc if and only if $\operatorname{epi} F$ is (strongly) closed, which is equivalent to $\operatorname{epi} F$ being weakly closed (as $\operatorname{epi} F$ is convex due to the convexity of F), which in turn is equivalent to F being wlsc. □

3.1 Convex duality

Let $\mathbb{X}$ be a locally convex topological vector space (over $\mathbb{R}$). Let $\mathbb{X}^*$ denote its topological strong dual, i.e. the set of all contnuous linear functionals, equipped with the strong dual topology. For readers unfamiliar with the setting of topological vector spaces, it suffices to take $\mathbb{X}$ to be a normed linear space and $\mathbb{X}^*$ to be its topological dual, equipped with the dual (operator) norm topology. For any $x^* \in \mathbb{X}^*$ and any $x \in \mathbb{X}$, we denote the duality bracket by

$$\langle x^*, x \rangle := x^*(x).$$

Now we start with the definition of an affine hyperplane. For the following, let $\mathbb{X}$ be a locally convex topological vector space (over $\mathbb{R}$).

Definition 15. H is called an *affine hyperplane* if

$$H = \{L = \alpha\} := \{x \in \mathbb{X} : L(x) = \alpha\},$$

for some linear functional $L : \mathbb{X} \to \mathbb{R}$ and some $\alpha \in \mathbb{R}$.

The following proposition is elementary.

Proposition 10. An affine hyperplane $\{L = \alpha\}$ is *closed* if and only if $L : \mathbb{X} \to \mathbb{R}$ is continuous.

We now recall the notion of separation by hyperplanes.

Definition 16. Let $\mathbb{A}, \mathbb{B} \subset \mathbb{X}$ be non-empty subsets. We say the affine hyperplane $\{L = \alpha\}$ *separates* $\mathbb{A}$ *and* $\mathbb{B}$ if

$$L(x) \leq \alpha \quad \text{for all } x \in \mathbb{A} \qquad \text{and} \qquad L(x) \geq \alpha \quad \text{for all } x \in \mathbb{B}.$$

We say $\{L = \alpha\}$ *strictly separates* $\mathbb{A}$ *and* $\mathbb{B}$ if both the inequalities above are strict.

Now we recall the *Hahn-Banach separation theorems*.

Theorem 9. Let $\mathbb{A}, \mathbb{B} \subset \mathbb{X}$ be *disjoint* non-empty *convex* subsets.

(a) If $\mathbb{A} \subset \mathbb{X}$ is *open*, then there exists a *closed affine hyperplane* that separates $\mathbb{A}$ and $\mathbb{B}$.

(b) If $\mathbb{A} \subset \mathbb{X}$ is *closed* and $\mathbb{A} \subset \mathbb{X}$ is *compact*, then there exists a *closed affine hyperplane* that strictly separates $\mathbb{A}$ and $\mathbb{B}$.

Proof. When $\mathbb{X}$ is a normed linear space, statement (a) and (b) are proved in [2, Theorem 1.6] and [2, Theorem 1.7], respectively. For the general case, see [4, Theorem 3.4]. $\square$

Theorem 10. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a function. Then the following are equivalent:

(i) F is a pointwise supremum of a family of continuous affine functions, i.e. there exists a family $\{G_i\}_{i \in I}$ such that $G_i : \mathbb{X} \to \mathbb{R}$ is continuous and affine for each $i \in I$ and

$$F(x) = \sup_{i \in I} G_i(x) \qquad \text{for each } x \in \mathbb{X}.$$

(ii) $F : \mathbb{X} \to (-\infty, +\infty]$ is convex and lsc.

Proof. $\underline{(i) \implies (ii)}$: Clearly continuous affine functions are convex and lsc. Thus, for each $i \in I$, G_i is convex and lsc. It is easy to check that both convexity and lower semicontinuity are preserved when we take pointwise supremum over an arbitrary family.

$\underline{(ii) \implies (i)}$: If $F \equiv +\infty$, then F is the pointwise supremum over *all* continuous affine functions from $\mathbb{X}$ to $\mathbb{R}$. Thus, without loss of generality, we can assume $\operatorname{dom} F \neq \emptyset$. Fix $\bar{x} \in \mathbb{X}$. We want to show that for any $\bar{t} \in \mathbb{R}$ with $F(\bar{x}) > \bar{t}$, there exists a continuous affine function G such that

$$F(\bar{x}) > G(\bar{x}) > \bar{t} \qquad \text{and} \qquad G \leq F.$$

Clearly, this proves our desired result. To show this, note that since F is convex and lsc, $\operatorname{epi} F$ is closed and convex in $\mathbb{X} \times \mathbb{R}$ and also non-empty, since $\operatorname{dom} F \neq \emptyset$. Observe that

$(\bar{x}, \bar{t}) \notin \operatorname{epi} F$. Thus, $\operatorname{epi} F$ and the singleton compact set $\{(\bar{x}, \bar{t})\}$ can be strictly separated by a continuous affine hyperplane

$$H = \{(x,t) \in \mathbb{X} \times \mathbb{R} : L(x) + \alpha t = \beta\}.$$

Thus, we have

$$L(\bar{x}) + \alpha\bar{t} < \beta \qquad \text{and} \qquad L(x) + \alpha t > \beta \quad \text{for all } (x,t) \in \operatorname{epi} F.$$

If $F(\bar{x}) < +\infty$, then we take $x = \bar{x}$ and $t = F(\bar{x})$ (observe that $(\bar{x}, F(\bar{x})) \in \operatorname{epi} F$) in the inequality on the right and substract the resulting inequality from the one on the left to deduce

$$\alpha\left(\bar{t} - F(\bar{x})\right) < 0 \qquad \Longrightarrow \alpha > 0.$$

Hence, we obtain,

$$\bar{t} < \frac{\beta}{\alpha} - \frac{1}{\alpha} L(\bar{x}) < F(\bar{x}).$$

Thus, the continuous affine function

$$G(x) := \frac{\beta}{\alpha} - \frac{1}{\alpha} L(x)$$

satisfies $\bar{t} < G(\bar{x}) < F(\bar{x})$ in this case. Moreover, since $(x, F(x)) \in \operatorname{epi} F$ for all $x \in \operatorname{dom} F$, we have

$$L(x) + \alpha F(x) > \beta \qquad \Longrightarrow G(x) = \frac{\beta}{\alpha} - \frac{1}{\alpha} L(x) < F(x).$$

Now we handle the case $F(\bar{x}) = +\infty$. If $\alpha \neq 0$, the argument above still works and thus we can assume $\alpha = 0$. But this means that we have

$$L(\bar{x}) < \beta \qquad \text{and} \qquad L(x) > \beta \quad \text{for all } x \in \operatorname{dom} F.$$

Thus, we deduce

$$\beta - L(\bar{x}) > 0 \qquad \text{and} \qquad \beta - L(x) < 0 \quad \text{for all } x \in \operatorname{dom} F.$$

On the other hand, since $\operatorname{dom} F \neq \emptyset$, using the arguments in the earlier case, we know there exists a continuous affine function $G_1(x) = \gamma - m(x)$ such that

$$\gamma - m(x) < F(x) \qquad \text{for all } x \in \mathbb{X}.$$

Now we define $G(x) = \gamma - m(x) - c\,(\beta - L(x))$ for some real number c so large such that

$$G(\bar{x}) = \gamma - m(\bar{x}) - c\,(\beta - L(\bar{x})) > \bar{t}.$$

This G satisfies all the desired properties and completes the proof. □

Definition 17 (Gamma regularization). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a given function. We define *the Gamma regularization of* F, denoted ΓF, as

$$\Gamma F(x) = \sup_{G \text{ affine minorant of F}} G(x), \qquad x \in \mathbb{X}.$$

Remark 13. (a) Note that for a general function $F : \mathbb{X} \to (-\infty, +\infty]$, it is entirely possible that F does not admit even a single continuous affine minorant and thus $\Gamma F \equiv -\infty$. However, it is easy to show that as soon as F admits even a single continuous affine minorant, ΓF maps $\mathbb{X}$ to $(-\infty, +\infty]$.

(b) If F admits a continuous affine minorant, then by Theorem 10, ΓF is convex and lsc. Moreover, it is the *convex lsc (lower) envelope* of F, i.e. if G is any other convex lsc function such that $G \leq F$, then $G \leq \Gamma F$ as well. In particular, if F is convex and lsc, then $\Gamma F = F$.

(c) Assuming F admits a continuous affine minorant, it can be shown that $\Gamma F \leq F_{\text{lsc}} \leq F$ and $\Gamma F = F_{\text{lsc}}$ if F is convex.

(d) Assuming F admits a continuous affine minorant, one can show

$$\operatorname{epi} \Gamma F = \overline{\operatorname{conv} \operatorname{epi} F},$$

where $\overline{\operatorname{conv} \operatorname{epi} F}$ denotes the *closed convex hull* of $\operatorname{epi} F$, i.e. the smallest closed convex set containing $\operatorname{epi} F$.

To prove deeper into the structure of continuous affine minorants of F, we observe that a general conntinuous affine function of $\mathbb{X}$ is a map

$$x \mapsto \langle x^*, x \rangle - \alpha$$

for some $x^* \in \mathbb{X}^*$ and some $\alpha \in \mathbb{R}$. Now, clearly such a function is everywhere below F if and only if we have

$$\langle x^*, x \rangle - \alpha \leq F(x) \qquad \text{for all } x \in \mathbb{X},$$

or equivalently,

$$\alpha \geq \langle x^*, x \rangle - F(x) \qquad \text{for all } x \in \mathbb{X}.$$

This motivates the following definition.

Definition 18 (Legendre transform). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a given function. We define the *Legendre transform* or *Fenchel-Legendre conjugate* or *Fenchel dual* of F, as the function $F^* : \mathbb{X}^* \to [-\infty, +\infty]$, defined by

$$F^*(x^*) := \sup_{x \in \mathbb{X}} \{\langle x^*, x \rangle - F(x)\}, \qquad x^* \in \mathbb{X}^*. \tag{11}$$

Remark 14. (a) In general, F^* *can* take the value $-\infty$. However, one can show that if $F \not\equiv +\infty$, then $F^* > -\infty$ and thus F^* maps $\mathbb{X}^*$ to $(-\infty, +\infty]$.

(b) Since $x^* \mapsto \langle x^*, x \rangle$ defines a continuous linear functional on $\mathbb{X}^*$, by Theorem 10, F^* is always convex and lsc on $\mathbb{X}^*$, even if F is neither.

(c) It is still entirely possible for $F^* \equiv +\infty$, even when $F \not\equiv +\infty$. Indeed, it is easy to show that $F^* \not\equiv +\infty$ if and only if F admits a continuous affine minorant.

(d) It is easy to show that

$$F^*(0) = -\inf_{x \in X} F(x).$$

Example 8. Let $\mathbb{X}$ be a normed linear space over $\mathbb{R}$ and let $F(x) = \|x\|_{\mathbb{X}}$. Then $F^*(x^*) = \|x^*\|_{\mathbb{X}^*}$.

Example 9. More generally, let $\mathbb{X}$ be a normed linear space over $\mathbb{R}$ and let $\phi : \mathbb{R} \to \mathbb{R}$ be even, convex, lsc function. If $F(x) = \phi\left(\|x\|_{\mathbb{X}}\right)$, then $F^*(x^*) = \phi^*\left(\|x^*\|_{\mathbb{X}^*}\right)$.

Definition 19 (Biconjugate). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. We define the *biconjugate* of F, as the function $F^{**} : \mathbb{X} \to (-\infty, +\infty]$, defined by

$$F^{**}(x) := \sup_{x^* \in \mathbb{X}^*} \{\langle x^*, x\rangle - F^*(x^*)\}, \qquad x \in \mathbb{X}. \tag{12}$$

Remark 15. (a) If F does not admit a continuous affine minorant, then $F^* \equiv +\infty$ and in this case, one can still use the above formula to define the biconjugate, but in that case $F^{**} \equiv -\infty$.

(b) Observe that the formula for the biconjugate is not quite the conjugate of the conjugate. Indeed, the domain for the conjugate of the conjugate should be the *bidual* $\mathbb{X}^{**}$. The biconjugate is the restriction of this map to $\mathbb{X}$. When $\mathbb{X}$ is a reflexive space, then $F^{**} = (F^*)^*$.

Exercise 12. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. Show that $F^{***} = (F^{**})^*$ maps $\mathbb{X}^*$ to $(-\infty, +\infty]$ and $F^* = F^{***}$.

Theorem 11 (Young-Fenchel inequality). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. Then we have the *Young-Fenchel inequality*:

$$\langle x^*, x\rangle \leq F(x) + F^*(x^*) \qquad \text{for all } x \in \mathbb{X} \text{ and for all } x^* \in \mathbb{X}^*. \tag{13}$$

Proof. By definition of F^*, for any $x^* \in \mathbb{X}^*$, we have

$$F^*(x^*) \geq \langle x^*, x\rangle - F(x) \qquad \text{for all } x \in \mathbb{X}.$$

This immediately yields the Fenchel-Young inequality. □

Theorem 12 (Fenchel-Moreau). Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. Then $F^{**} \leq F$. In particular, if $F : \mathbb{X} \to (-\infty, +\infty]$ is *proper*, *convex* and *lsc*, then $F = F^{**}$ and thus, F has the *Fenchel-Moreau representation*, i.e.

$$F(x) = \sup_{x^* \in \mathbb{X}^*} \{\langle x^*, x\rangle - F^*(x^*)\}, \qquad \text{for all } x \in \mathbb{X}. \tag{14}$$

Proof. The Young-Fenchel inequality (13) immediately yields

$$\langle x^*, x\rangle - F^*(x^*) \leq F(x) \qquad \text{for all } x \in \mathbb{X} \text{ and for all } x^* \in \mathbb{X}^*.$$

Taking supremum over all $x^* \in \mathbb{X}^*$, we deduce $F^{**} \leq F$. Now if F is convex and lsc, then by Theorem 10, there exists a family $\{G_i\}_{i\in I}$ such that $G_i : \mathbb{X} \to \mathbb{R}$ is continuous and affine for each $i \in I$ and

$$F(x) = \sup_{i\in I} G_i(x) \qquad \text{for each } x \in \mathbb{X}.$$

Since each G_i is continuous and affine, we must have

$$G_i(x) = \langle x_i^*, x\rangle - \alpha_i, \qquad x \in \mathbb{X},$$

for some $x_i^* \in \mathbb{X}^*$ and $\alpha_i \in \mathbb{R}$. Since $G_i \leq F$, we must have

$$\alpha_i \geq F^*(x_i^*).$$

But this implies

$$G_i(x) = \langle x_i^*, x\rangle - \alpha_i \leq \langle x_i^*, x\rangle - F^*(x_i^*) \leq F^{**}(x) \qquad \text{for all } x \in \mathbb{X}.$$

Since $i \in I$ is arbitrary, we have $G_i(x) \leq F^{**}(x)$ for all $i \in I$ and for all $x \in \mathbb{X}$. Taking supremum over $i \in I$, we obtain our desired result. $\square$

Exercise 13. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. Show that $\Gamma F = F^{**}$.

3.2 Subdifferentiability and subgradients

Definition 20. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a given function and let $L : \mathbb{X} \to \mathbb{R}$ be a continuous affine minorant of F. For $\bar{x} \in \mathbb{X}$, we say L is *exact at* $\bar{x}$ if $L(\bar{x}) = F(\bar{x})$.

Definition 21. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a given function. F is said to be *subdifferentiable at* $x \in \mathbb{X}$ if F has a continuous affine minorant which is exact at x. The *slope* $x^* \in \mathbb{X}^*$ of such a continuous affine minorant is called a *subgradient of* F *at* x, denoted $x^* \in \partial F(x)$. If F is not subdifferentiable at x, then we set $\partial F(x) = \emptyset$ by convention.

Proposition 11. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function. Then $x^* \in \partial F(x)$ if and only if $F(x)$ is finite and we have

$$\langle x^*, y - x\rangle + F(x) \leq F(y) \qquad \text{for all } y \in \mathbb{X}. \tag{15}$$

Proof. Set $L(y) = \langle x^*, y-x\rangle + F(x)$. Note that when $F(x)$ is finite, L is the unique continuous affine function such that $L(x) = F(x)$. Thus, $x^* \in \partial F(x)$ implies $L \leq F$, i.e. (15) holds. Conversely, since F is proper, (15) implies $F(x)$ is finite and $L \leq F$. $\square$

The following result is an immediate corollary.

Theorem 13. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function. If $\inf_{\mathbb{X}} F \in \mathbb{R}$, then $\bar{x} \in \mathbb{X}$ is a minimizer for F if and only if $0 \in \partial F(\bar{x})$.

Theorem 14. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be a proper function such that F^* is also proper. Show that $x^* \in \partial F(x)$ if and only if

$$F(x) + F^*(x^*) = \langle x^*, x\rangle.$$

Proof. We know $x^* \in \partial F(x)$ if and only if $F(x)$ is finite and

$$\langle x^*, y - x\rangle + F(x) \leq F(y) \qquad \text{for all } y \in \mathbb{X}.$$

This is equivalent to

$$F^*(x^*) = \sup_{y\in\mathbb{X}} \langle x^*, y\rangle - F(y) \leq \langle x^*, x\rangle - F(x).$$

Thus, $x^* \in \partial F(x)$ if and only if

$$\langle x^*, x\rangle \leq F(x) + F^*(x^*).$$

Since the reverse inequality is just the Young-Fenchel inequality, we have our result. □

Remark 16. We are not going to prove it, but if F is convex, proper, lsc, then F is subdifferentiable at each $x \in \operatorname{Int}\operatorname{dom} F$, where Int denotes the (topological) interior. Also, if F is convex, finite and continuous in a neighborhood of $x \in \mathbb{X}$ and Gateaux-differentiable at x, then $\partial F(x) = \{DF(x)\}$, where $DF(x) \in \mathbb{X}^*$ is the Gateaux-derivative at x.

Theorem 15. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be proper, convex, lsc and admits a continuous affine minorant. Then for any $x \in \mathbb{X}, x^* \in \mathbb{X}^*$, the following are equivalent.

(a) $x^* \in \partial F(x)$.

(b) $x \in \partial F^*(x^*)$.

(c) Equality holds in the Young-Fenchel inequality, i.e.

$$F(x) + F^*(x^*) = \langle x^*, x\rangle.$$

Proof. Since F is proper and admits a continuous affine minorant implies F^* is proper. First observe that if any of the three holds, then $F(x)$ and $F^*(x^*)$ must be finite, as the duality pairing is always a real number and F, F^* both proper and none of them takes the value $-\infty$. Now the equivalence of (a) and (c) is just Theorem 14. Now since F is convex, proper and lsc, by Theorem 12, we have the Fenchel-Moreau representation (14)

$$F(x) = \sup_{y^*\in\mathbb{X}^*} \{\langle y^*, x\rangle - F^*(y^*)\}.$$

Thus, (c) holds if and only if

$$\langle y^*, x\rangle - F^*(y^*) \leq F(x) = \langle x^*, x\rangle - F^*(x^*) \qquad \text{for all } y^* \in \mathbb{X}^*.$$

But this is equivalent to (a) by Proposition 11 applied to F^*. □

3.3 Subdifferential as a set-valued map

We now want to study ∂F as a set-valued/multi-valued map/operator from $\mathbb{X}$ to $\mathbb{X}^*$. We denote set valued maps with "$\rightrightarrows$"

Definition 22. Let $T : \mathbb{X} \rightrightarrows \mathbb{X}^*$ be a set-valued map. The *graph of* T, denoted $\operatorname{Gr} T$, is defined as

$$\operatorname{Gr} T = \{(x, x^*) \in \mathbb{X} \times \mathbb{X}^* : x^* \in T(x)\}.$$

(a) T is called *monotone* if

$$\langle x^* - y^*, x - y \rangle \geq 0 \qquad \text{for every } (x, x^*), (y, y^*) \in \operatorname{Gr} T. \tag{16}$$

(b) T is called *maximal monotone* if

$$\begin{aligned} (z, z^*) \in \mathbb{X} \times \mathbb{X}^* \ \text{ satisfies } \ \langle z^* - x^*, z - x \rangle \geq 0 \quad \text{for every } (x, x^*) \in \operatorname{Gr} T \\ \implies (z, z^*) \in \operatorname{Gr} T. \end{aligned} \tag{17}$$

(c) T is called *cyclically monotone* if for every integer $N \in \mathbb{N}$, we have

$$\begin{aligned} &\sum_{i=1}^{N} \langle x_i^*, x_{i+1} - x_i \rangle \leq 0 \\ &\qquad \text{whenever } (x_i, x_i^*) \in \operatorname{Gr} T \quad \text{for all } 1 \leq i \leq N, \ \text{and } x_{N+1} = x_1. \end{aligned} \tag{18}$$

(d) T is called *maximally cyclically monotone* if

$$\begin{aligned} (z, z^*) \in &\mathbb{X} \times \mathbb{X}^* \ \text{ satisfies} \\ &\sum_{i=1}^{N} \langle x_i^*, x_{i+1} - x_i \rangle + \langle z^*, x_1 - z \rangle \leq 0, \qquad \text{for every } N \in \mathbb{N}, \\ &\quad \text{whenever } (x_i, x_i^*) \in \operatorname{Gr} T \quad \text{for all } 1 \leq i \leq N, \ \text{with } x_{N+1} = z, \\ &\qquad \implies (z, z^*) \in \operatorname{Gr} T. \end{aligned} \tag{19}$$

Remark 17. (a) The definition of maximal monotonicity implies that $\operatorname{Gr} T$ is *maximal as a subset of* $\mathbb{X} \times \mathbb{X}^*$ *enjoying the monotonicity property.* More precisely, there does not exist a subset $A \subset \mathbb{X} \times \mathbb{X}^*$ such that $\operatorname{Gr} T \subset A$ with $\operatorname{Gr} T \neq A$, such that

$$\langle x^* - y^*, x - y \rangle \geq 0 \qquad \text{for every } (x, x^*), (y, y^*) \in A.$$

(b) Similarly, he definition of maximal cyclical monotonicity implies that $\operatorname{Gr} T$ is *maximal as a subset of* $\mathbb{X} \times \mathbb{X}^*$ *enjoying the cyclical monotonicity property.* More precisely, there does not exist a subset $A \subset \mathbb{X} \times \mathbb{X}^*$ such that $\operatorname{Gr} T \subset A$ with $\operatorname{Gr} T \neq A$, such that

$$\begin{aligned} &\sum_{i=1}^{N} \langle x_i^*, x_{i+1} - x_i \rangle \leq 0 \\ &\qquad \text{whenever } (x_i, x_i^*) \in A \quad \text{for all } 1 \leq i \leq N, \ \text{and } x_{N+1} = x_1. \end{aligned}$$

(c) Note that cyclic montonicity implies monotonicity. Indeed, for $N = 2$, cyclic monotonicity inequality becomes $\langle x^*, y - x \rangle + \langle y^*, x - y \rangle \leq 0$ for every $(x, x^*), (y, y^*) \in \operatorname{Gr} T$, which is just montonicity.

(d) The motivation for the definitions are of course, C^2 convex functions $f : \mathbb{R}^d \to \mathbb{R}$. Indeed, if $f : \mathbb{R}^d \to \mathbb{R}$ is C^2, then $\nabla^2 f \geq 0$ as matrices, i.e. $\nabla^2 f(x)$ is a positive definite matrix for every $x \in \mathbb{R}$. Thus, by the fundamental theorem of calculus, we have

$$\begin{aligned}\langle \nabla f(x) - \nabla f(y), x - y\rangle &= \left\langle \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\left[\nabla f(y + t(x-y))\right]\ \mathrm{d}t, x - y\right\rangle \\ &= \int_0^1 \left\langle \nabla^2 f(y + t(x-y))(x-y), x-y\right\rangle\ \mathrm{d}t \geq 0.\end{aligned}$$

This implies ∇f is monotone. To get the cyclic monotonicity, define the function $g : [0,1] \to \mathbb{R}$ by $g(t) = f(ty + (1-t)x)$. Then $g''(t) = \langle \nabla^2 f(y + t(x-y))(x-y), x-y\rangle \geq 0$. Thus, g' is nondecreasing and hence $g'(t) \geq g'(0)$ for all $t \in [0,1]$. Thus, by the fundamental theorem of calculus, we have

$$f(y) - f(x) = g(1) - g(0) = \int_0^1 g'(t)\ \mathrm{d}t \geq g'(0) = \langle \nabla f(x), y - x\rangle.$$

Thus, we have

$$f(x_{i+1}) - f(x_i) \geq \langle \nabla f(x_i), x_{i+1} - x_i\rangle \qquad \text{for all } 1 \leq i \leq N.$$

Thus, summing up from $1 \leq i \leq N$ and recalling $x_{N+1} = x_1$, we have

$$0 = f(x_{N+1}) - f(x_1) = \sum_{i=1}^{N} f(x_{i+1}) - f(x_i) \geq \sum_{i=1}^{N} \langle \nabla f(x_i), x_{i+1} - x_i\rangle.$$

It is also easy to see that ∇f is also maximal cyclically monotone. Indeed, even using the definition for $N = 1$, we see that if $(p, z) \in \mathbb{R}^d \times \mathbb{R}^d$ satisfies

$$\langle \nabla f(x), z - x\rangle + \langle p, x - z\rangle \leq 0, \qquad \text{for all } x \in \mathbb{R}^d,$$

then we have

$$\langle \nabla p - f(x), x - z\rangle \leq 0, \qquad \text{for all } x \in \mathbb{R}^d.$$

Choosing $x = z + tv$ for some $v \in \mathbb{R}^d$ and $t > 0$, we get $\langle \nabla p - f(z + tv), v\rangle \leq 0$. Letting $t \to 0+$, we have $\langle \nabla p - f(z), v\rangle \leq 0$. Similarly, choosing $t < 0$ and letting $t \to 0-$, we deduce $\langle \nabla p - f(z), v\rangle \geq 0$. Thus, $\langle \nabla p - f(z), v\rangle = 0$. But since $v \in \mathbb{R}^d$ is arbitrary, this implies $p = f(z)$.

Theorem 16. Let $\mathbb{X}$ be a reflexive Banach space and let $F : \mathbb{X} \to (-\infty, +\infty]$ be convex, lsc and proper. Then $\operatorname{Gr} \partial F \subset \mathbb{X} \times \mathbb{X}^*$ is strong-weak $*$-closed. More precisely, if $\{(x_n, \xi_n)\}_{n \in \mathbb{N}} \subset \operatorname{Gr} \partial F$ be a sequence such that $x_n \to x$ strongly in $\mathbb{X}$ and $\xi_n \overset{*}{\rightharpoonup} \xi$ weakly $*$ in $\mathbb{X}^*$, then $(x, \xi) \in \operatorname{Gr} \partial F$, i.e. $\xi \in \partial F(x)$.

Proof. Since $(x_n, \xi_n) \in \operatorname{Gr} \partial F$ means $\xi_n \in \partial F(x_n)$ for all $n \in \mathbb{N}$, by Proposition 11, we get

$$F(x_n) \leq F(y) - \langle \xi_n, y - x_n\rangle = F(y) - \langle \xi_n, y - x\rangle - \langle \xi_n, x - x_n\rangle \qquad \text{for all } y \in \mathbb{X}. \tag{20}$$

Now since $\xi_n \overset{*}{\rightharpoonup} \xi$, we have

$$\lim_{n \to \infty} \langle \xi_n, y - x\rangle = \langle \xi, y - x\rangle.$$

Also, since $\{\xi_n\}_{n\in\mathbb{N}}$ is weak $*$ convergent, by the uniform boundedness principle we have $\|\xi_n\|_{\mathbb{X}^*} \leq C$ for some $C > 0$ for all $n \in \mathbb{N}$. Thus, we deduce

$$|\langle \xi_n, x - x_n| \leq \|\xi_n\|_{\mathbb{X}^*} \|x - x_n\|_{\mathbb{X}} \leq C \|x - x_n\|_{\mathbb{X}} \to 0,$$

as $x_n \to x$ strongly in $\mathbb{X}$. Thus, letting $n \to \infty$ in (20), we arrive at

$$\limsup F(x_n) \leq F(y) - \langle \xi, y - x \rangle \qquad \text{for all } y \in \mathbb{X}.$$

Now since F is lsc, we have

$$F(x) \leq \liminf F(x_n) \leq \limsup F(x_n) \leq F(y) - \langle \xi, y - x \rangle \qquad \text{for all } y \in \mathbb{X}.$$

Since F is proper, this implies $F(x)$ is finite and we have

$$\langle \xi, y - x \rangle + F(x) \leq F(y) \qquad \text{for all } y \in \mathbb{X}.$$

By Proposition 11, this implies $\xi \in \partial F(x)$. ☐

Theorem 17. Let $\mathbb{X} = \mathbb{H}$ be a Hilbert space and let $F : \mathbb{H} \to (-\infty, +\infty]$ be convex, lsc and proper. Then $\partial F : \mathbb{H} \rightrightarrows \mathbb{H}$ is maximal cyclically monotone. Here we have identified $\mathbb{H}^*$ with $\mathbb{H}$ via the Riesz isomorphism.

Proof. First let us prove that ∂F is cyclically monotone. Let $N \in \mathbb{N}$ be an integer and let $(x_i, x_i^*) \in \operatorname{Gr} \partial F$ for all $1 \leq i \leq N$. By Proposition 11, this implies

$$f(y) \geq f(x_i) + \langle x_i^*, y - x_i \rangle \qquad \text{for all } y \in \mathbb{H}, \text{ for all } 1 \leq i \leq N.$$

Plugging $y = x_{i+1}$, where $x_{N+1} = x_1$ and summing up over $1 \leq i \leq N$, we deduce

$$\sum_{i=1}^{N} \langle x_i^*, x_{i+1} - x_i \rangle \leq \sum_{i=1}^{N} [f(x_{i+1}) - f(x_i)] = f(x_{N+1}) - f(x_1) = 0.$$

Thus, ∂F is cyclically monotone.

Now since cyclical monotonicity implies monotonicity, to prove ∂F is maximal cyclically monotone, it is enough to show that it is maximal monotone. Now let $(z, z^*) \in \operatorname{Gr} \partial F$ be such that

$$\langle z^* - x^*, z - x \rangle \leq 0 \qquad \text{for all } (x, x^*) \in \operatorname{Gr} \partial F. \tag{21}$$

We want to show this implies $z^* \in \partial F(z)$. We define $G(x) = F(x) - \langle z^*, x \rangle$ for all $x \in \mathbb{H}$. Then it is easy to see that G is convex, proper, lsc and we have $\partial G(x) = \partial F(x) - z^*$, i.e. $\xi \in \partial G(x)$ if and only if $\xi + z^* \in \partial F(x)$. Thus, the inequality (21) is equivalent to

$$\langle \xi, z - x \rangle \leq 0 \qquad \text{for all } \xi \in \partial G(x). \tag{22}$$

Now we define

$$\Phi(x) = G(x) + \frac{1}{2} \|x - z\|_{\mathbb{H}}^2 \qquad \text{for all } x \in \mathbb{H}.$$

Clearly Φ is convex, lsc, proper, since G is. Also, the term $\frac{1}{2} \|x - z\|_{\mathbb{H}}^2$ is strictly convex and hence so is Φ. We claim Φ is also coercive. Indeed, since G is convex, proper, lsc, G admits a continuous affine minorant. Let $\langle p, x\rangle_{\mathbb{H}} + \alpha$ be a continuous affine minorant of G, for some $p \in \mathbb{H}$ and some $\alpha \in \mathbb{R}$. Then we have

$$\begin{aligned}
\Phi(x) &= G(x) + \frac{1}{2}\|x-z\|_{\mathbb{H}}^2 \\
&\geq \langle p, x\rangle_{\mathbb{H}} + \alpha + \frac{1}{2}\|x-z\|_{\mathbb{H}}^2 \\
&= \frac{1}{2}\left[\|x-z\|_{\mathbb{H}}^2 + 2\langle p, x-z\rangle_{\mathbb{H}} + \|p\|_{\mathbb{H}}^2\right] + \langle p, z\rangle_{\mathbb{H}} - \frac{1}{2}\|p\|_{\mathbb{H}}^2 \\
&= \frac{1}{2}\|x-z+p\|_{\mathbb{H}}^2 + \langle p, z\rangle_{\mathbb{H}} - \frac{1}{2}\|p\|_{\mathbb{H}}^2 .
\end{aligned}$$

But, we also have

$$\|x\|_{\mathbb{H}}^2 \leq 2\left(\|x-z+p\|_{\mathbb{H}}^2 + \|z-p\|_{\mathbb{H}}^2\right).$$

Combining, we deduce

$$\begin{aligned}
\Phi(x) &\geq \frac{1}{2}\|x-z+p\|_{\mathbb{H}}^2 + \langle p, z\rangle_{\mathbb{H}} - \frac{1}{2}\|p\|_{\mathbb{H}}^2 \\
&\geq \frac{1}{4}\|x\|_{\mathbb{H}}^2 - \frac{1}{2}\|z-p\|_{\mathbb{H}}^2 + \langle p, z\rangle_{\mathbb{H}} - \frac{1}{2}\|p\|_{\mathbb{H}}^2 .
\end{aligned}$$

Thus, $\Phi(x) \to +\infty$ as $\|x\|_{\mathbb{H}} \to +\infty$. Thus, Φ has a unique minimizer $x_0 \in \mathbb{H}$. But then $0 \in \partial\Phi(x_0)$. Since $\partial\Phi(x) = \partial G(x) + x - z$, we infer that there exists $\xi_0 \in \partial G(x_0)$ such that $\xi_0 + x_0 - z = 0$. But by (22), we have

$$\langle \xi_0, z - x_0\rangle_{\mathbb{H}} \leq 0.$$

Thus, we arrive at $\|z-x_0\|_{\mathbb{H}}^2 \leq 0$ and consequently $z = x_0$ and $\xi_0 = 0$. Thus, we have established that $0 \in \partial G(z)$ and equivalently, $z^* \in \partial F(z)$. This completes the proof. □

Remark 18. The converse is also true, even on Banach spaces and is known as Rockafellar theorem.

The following topics were covered in ICTS.

---

4 Gamma convergence

4.1 What do we wish from a variational convergence?

We consider sequence of functions $F_n : \mathbb{X} \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} = [-\infty, +\infty]$. We are still interested in functions taking values in $(-\infty, +\infty]$, but there are certain technical advantages in working with $\overline{\mathbb{R}}$. Of course, as before, $\mathbb{X}$ is a topological space.

Now we want a notion of convergence for the sequence F_n such that we have the following (under reasonable hypotheses):

Wishlist

$F_n \to F$ implies

(a) minimizer of F exists

(b) the infimas converge to infimum, i.e.

$$\inf_{\mathbb{X}} F_n \to \inf_{\mathbb{X}} F.$$

(c) sequence of minimizers for F_n converge to a minimizer for F, i.e.

$$\arg\min_{\mathbb{X}} F_n \to \arg\min_{\mathbb{X}} F.$$

or we might even wish for the stronger property that any *minimizing sequence for the family* $\{F_n\}_{n\in\mathbb{N}}$ converges to a minimizer of F.

What is a *minimizing sequence for the family* $\{F_n\}_{n\in\mathbb{N}}$? We have already seen minimizing sequences for a single function F. Now we generalize to a countable family $\{F_n\}_{n\in\mathbb{N}}$.

Definition 23. A sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{X}$ is called a *minimizing sequence for the family* $\{F_n\}_{n\in\mathbb{N}}$ if

$$\lim_{n\to\infty} F_n(x_n) = \lim_{n\to\infty} \inf_{\mathbb{X}} F_n.$$

Remark 19. Note that if x_n is a minimizer of F_n for each $n \in \mathbb{N}$, then $\{x_n\}_{n\in\mathbb{N}}$ is clearly a minimizing sequence for $\{F_n\}_{n\in\mathbb{N}}$.

We also define *almost minimizers* of a function F.

Definition 24. Let $F : \mathbb{X} \to \overline{\mathbb{R}}$ and let $\varepsilon > 0$. Then $x \in \mathbb{X}$ is called ε-almost minimizer for F in $\mathbb{X}$ if

$$F(x) \leq \max\left\{\left[\inf_{\mathbb{X}} F + \varepsilon\right], -\frac{1}{\varepsilon}\right\}.$$

The set of all ε-(almost) minimizers is denoted by $\mathcal{M}_\varepsilon(F)$. The set of minimizers of F is denoted by $\mathcal{M}(F)$.

Remark 20. (i) $\mathcal{M}(F)$ can be empty, but $\mathcal{M}_\varepsilon(F)$ is never empty. Also, clearly

$$\mathcal{M}(F) = \bigcap_{\varepsilon>0} \mathcal{M}_\varepsilon(F).$$

(ii) The term $-1/\varepsilon$ is there to cater to the case where $\inf_{\mathbb{X}} F = -\infty$. If F is bounded below, then x is an ε-minimizer (for small $\varepsilon > 0$) if and only if

$$F(x) \leq \inf_{\mathbb{X}} F + \varepsilon.$$

Definition 25. Let $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset (0,1)$ be a sequence of positive reals such that $\varepsilon_n \to 0$. A sequence $\{x_n\} \subset \mathbb{X}$ is called an ε_n*-almost minimizing* (sequence) *for* $\{F_n\}_{n\in\mathbb{N}}$ if

$$F_n(x_n) \leq \max\left\{\left[\inf_{\mathbb{X}} F_n + \varepsilon_n\right], -\frac{1}{\varepsilon_n}\right\}.$$

Note: From here onwards, we shall always assume any minimizing sequence for $\{F_n\}_{n\in\mathbb{N}}$ is ε_n-almost minimizing for some sequence $\{\varepsilon_n\}_{n\in\mathbb{N}} \subset (0,1)$ with $\varepsilon_n \to 0$.

Now suppose $\{\bar{x}_n\}$ be ε_n-almost minimizing for $\{F_n\}$ and suppose $\bar{x}_n \to \bar{x}$ in $\mathbb{X}$. Since we would very much like to conclude $\bar{x}$ is a minimizer for F, where F is the *limit* under our desired convergence, from our experience with the direct methods, it seems reasonable, indeed one might suspect probably even necessary, to assume some type of sequential *lower semicontinuity* for the family $\{F_n\}$. Thus, it seems logical to require the following:

- $(\mathbf{1})$ **: lim inf inequality:** For each $x \in \mathbb{X}$ and *every* sequence $x_n \to x$ in $\mathbb{X}$, we have

  $$F(x) \leq \liminf_n F_n(x_n).$$

  But obviously we also can not allow F to stay well below the infimas of F_n. Thus, it seems reasonable to also require an *upper bound.*

- $(\mathbf{2})$ **: lim sup inequality:** For each $x \in \mathbb{X}$, *there exists* a sequence $x_n \to x$ in $\mathbb{X}$ such that

  $$F(x) \geq \limsup_n F_n(x_n).$$

  In view of $(\mathbf{1})$, it is easy to see that $(\mathbf{2})$ is equivalent to

  $(\mathbf{2'})$ **: recovery/realizing sequence:** For each $x \in \mathbb{X}$, *there exists* a sequence $x_n \to x$ in $\mathbb{X}$ such that

  $$F(x) = \lim_n F_n(x_n).$$

Are these two conditions (either $(\mathbf{1})$ and $(\mathbf{2})$, or, equivalently, $(\mathbf{1})$and $(\mathbf{2'})$) enough? To get a feel for what we are aiming for, let us discuss a plausible argument to 'prove' the statements in our wishlist.

Arguments: Let $\{\bar{x}_n\}$ be ε_n-almost minimizing for $\{F_n\}$ and assume $\bar{x}_n \to \bar{x}$ in $\mathbb{X}$. Now, for every $x \in \mathbb{X}$, by $(\mathbf{2})$, there exists a sequence $y_n \to x$ such that

$$\limsup F_n(y_n) \leq F(x).$$

But clearly,

$$\inf_{\mathbb{X}} F_n \leq F_n(y_n).$$

Thus, we have

$$\limsup \left(\inf_{\mathbb{X}} F_n \right) \leq \limsup F_n(y_n) \leq F(x).$$

As this must hold for every $x \in \mathbb{X}$, taking inf over $x \in \mathbb{X}$, we deduce

$$\limsup \left(\inf_{\mathbb{X}} F_n \right) \leq \inf_{\mathbb{X}} F.$$

Now since $\{\bar{x}_n\}$ be ε_n-almost minimizing and $\bar{x}_n \to \bar{x}$, by (**1**), we deduce

$$\begin{aligned}
\inf_{\mathbb{X}} F &\leq F(\bar{x}) \\
&\leq \liminf F_n(\bar{x}_n) && \text{[By } (\mathbf{1})] \\
&\leq \liminf \left[\inf_{\mathbb{X}} F_n + \varepsilon_n\right] && [\text{As } \{\bar{x}_n\} \text{ is } \varepsilon_n\text{-minimizing}] \\
&\liminf \inf_{\mathbb{X}} F_n && [\text{As } \varepsilon_n \to 0] \\
&\leq \limsup \inf_{\mathbb{X}} F_n \\
&\leq \inf_{\mathbb{X}} F.
\end{aligned}$$

Thus, all of the above inequalities must be equalities and we infer that $\inf_{\mathbb{X}} F_n \to \inf_{\mathbb{X}} F$ and $\bar{x}$ is a minimizer for F.

Obviously, there is no reason why a ε_n-almost minimizing sequence must converge, or even admit a convergent subsequence. But as usual, this type of compactness properties can be ensured by some fairly reasonable hypotheses.

4.2 Γ-convergence: Definitions

Thus, with the above arguments in mind, we now define the notion of *gamma convergence*, often also known as *variational convergence* or *epi-convergence* (the name stems from the *epigraph* of a function we defined earlier) and typically written as Γ-convergence.

Definition 26 (Γ-convergence). Let $F_n : \mathbb{X} \to \overline{\mathbb{R}}$ be a sequence of functions. We define the *lower* Γ*-limit* or Γ-lim inf and the *upper* Γ*-limit* or Γ-lim sup, of $\{F_n\}$, as

$$\begin{aligned}
\underline{F}(x) := \left(\Gamma\text{-}\liminf F_n\right)(x) &:= \sup_{U \in \mathcal{N}(x)} \liminf_n \inf_{y \in U} F_n(y), \\
\overline{F}(x) := \left(\Gamma\text{-}\limsup F_n\right)(x) &:= \sup_{U \in \mathcal{N}(x)} \limsup_n \inf_{y \in U} F_n(y).
\end{aligned}$$

We say F_n Γ*-converges to* F, denoted $F_n \xrightarrow{\Gamma} F$, or F *is the* Γ*-limit of* F_n, if

$$\Gamma\text{-}\liminf F_n = \Gamma\text{-}\liminf F_n := \Gamma\text{-}\lim F_n$$

and $F = \Gamma\text{-}\lim F_n$.

Remark 21. (i) Recall the formula for the lsc envelope

$$F_{\mathrm{lsc}} = \sup_{U \in \mathcal{N}(x)} \inf_{y \in U} F(y)$$

and compare with the formulas above. This trick is quite useful for remembering the formulas. This also hints at the conclusion of Proposition 12.

(ii) The advantage of working with $\overline{\mathbb{R}}$ is already apparent in the fact that $\Gamma\text{-}\liminf F_n$ and $\Gamma\text{-}\limsup F_n$ always exist as extended real-valued functions.

Proposition 12. $\Gamma\text{-}\liminf F_n$ and $\Gamma\text{-}\limsup F_n$ are always lsc (with respect to the topology on $\mathbb{X}$ that is used to define the upper and lower Γ-limits). In particular, $\Gamma\text{-}\lim F_n$ is always lsc even if none of the F_ns are.

Exercise 14. Prove Proposition 12.

When the topology is first countable, we can obviously characterize these in terms of sequences.

Proposition 13. Let $\mathbb{X}$ be first countable. Then

$\Gamma\text{-}\liminf F_n$ is *characterized* by the following two properties:

(a) For every $x \in \mathbb{X}$ and every sequence $x_n \to x$ in $\mathbb{X}$, we have

$$\Gamma\text{-}\liminf F_n(x) \leq \liminf_n F_n(x_n).$$

(b) For every $x \in \mathbb{X}$, there exists a sequence $x_n \to x$ in $\mathbb{X}$ such that

$$\Gamma\text{-}\liminf F_n(x) = \liminf_n F_n(x_n).$$

$\Gamma\text{-}\limsup F_n$ is *characterized* by the following two properties:

(c) For every $x \in \mathbb{X}$ and every sequence $x_n \to x$ in $\mathbb{X}$, we have

$$\Gamma\text{-}\limsup F_n(x) \leq \limsup_n F_n(x_n).$$

(d) For every $x \in \mathbb{X}$, there exists a sequence $x_n \to x$ in $\mathbb{X}$ such that

$$\Gamma\text{-}\limsup F_n(x) = \limsup_n F_n(x_n).$$

In particular, $F_n \xrightarrow{\Gamma} F$ if and only if we have the following two conditions:

(1) $\liminf$ *inequality*: For every $x \in \mathbb{X}$ and every sequence $x_n \to x$ in $\mathbb{X}$, we have

$$F(x) \leq \liminf_n F_n(x_n).$$

(2) *recovery sequence*: For every $x \in \mathbb{X}$, there exists a sequence $x_n \to x$ in $\mathbb{X}$ such that

$$F(x) = \lim_n F_n(x_n).$$

Proof. Since $\mathbb{X}$ is first countable, the topology can be characterized by sequences. Hence it is easy to see that the definitions in this case implies

$$\begin{aligned}
\Gamma\text{-}\liminf F_n(x) &= \inf_{x_n \to x} \liminf_n F_n(x_n), \\
\Gamma\text{-}\limsup F_n(x) &= \inf_{x_n \to x} \limsup_n F_n(x_n), \\
\Gamma\text{-}\lim F_n(x) &= \inf_{x_n \to x} \lim_n F_n(x_n).
\end{aligned}$$

The statements easily follows from this. The details are left to the reader. $\square$

Before we see some examples of Γ-convergent sequences, we first define some other natural notion of convergences.

Definition 27. We say F_n converges *pointwise* to F, denoted $F_n \overset{\text{pw}}{\to} F$, if

$$F_n(x) \to F(x) \qquad \text{for all } x \in \mathbb{X}.$$

F_n is said to *continuously converge* to F, denoted $F_n \overset{\text{cc}}{\to} F$, if for every $x \in \mathbb{X}$ and every neighborhood V of $F(x)$ in $\overline{\mathbb{R}}$, there exists $k \in \mathbb{N}$ and $U \in \mathcal{N}(x)$ such that

$$F_n(y) \in V \quad \text{ for every } n \geq k \text{ and every } y \in U.$$

Remark 22. (i) *continuous convergence* is *stronger* than Γ-convergence. Indeed, it is not difficult to show from the definitions that

$$F_n \overset{\text{cc}}{\to} F \qquad \text{if and only if} \qquad F_n \overset{\Gamma}{\to} F \quad \text{and} \quad (-F_n) \overset{\Gamma}{\to} (-F).$$

(ii) If $\mathbb{X}$ is first countable, then

$$F_n \overset{\text{cc}}{\to} F \qquad \text{if and only if} \qquad F_n(x_n) \to F(x) \quad \text{for every sequence } x_n \to x \text{ in } \mathbb{X}.$$

(iii) If $\mathbb{X}$ is a Banach space, equipped with weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$, then

$$F_n \overset{\text{cc}}{\to} F \qquad \text{if and only if} \qquad F_n(x_n) \to F(x) \quad \text{for every sequence } x_n \rightharpoonup x \text{ in } \mathbb{X}.$$

(iv) It is easy to see that *continuous convergence* is *stronger* than *pointwise convergence.* Indeed, when $F_n = F$ for all n, $F_n \overset{\text{pw}}{\to} F$, but $F_n \overset{\text{cc}}{\to} F$ amounts to *sequential continuity* of F.

Remark 23. We will not bother about even stronger notions of convergence, such that *uniform convergence.* F_n is said to converge uniformly to F if $\lim\limits_{n} \sup\limits_{x\in\mathbb{X}} d\left(F_n(x), F(x)\right) = 0$, where d is any metric on $\overline{\mathbb{R}}$ that induces the topology on $\overline{\mathbb{R}}$. Convergences like these are too strong to be useful in practice.

Notation: As Γ-convergence can depend on the topology, when $\mathbb{X}$ is a Banach space, we would write $F_n \overset{\Gamma}{\to} F$ when we mean Γ-convergence with respect to the strong (norm) topology and write $F_n \overset{\Gamma}{\rightharpoonup} F$ when we mean Γ-convergence with respect to the weak topology. Sometimes we also write $\Gamma_{\text{s}}\text{-}\lim F_n$ and $\Gamma_{\text{w}}\text{-}\lim F_n$ to denote the Γ-limits with respect to the strong and the weak topology, respectively.

Definition 28. Let $\mathbb{X}$ be a Banach space. We say F_n *Mosco-converges to* F, denoted $F_n \overset{\text{M}}{\to} F$ if

$$F_n \overset{\Gamma}{\to} F \qquad \text{and} \qquad F_n \overset{\Gamma}{\rightharpoonup} F.$$

Remark 24. It is easy to see that to show $F_n \overset{\text{M}}{\to} F$, it is enough to show that lim inf inequality for *every weakly convergent sequence* $x_n \rightharpoonup x$ and to find a *strongly convergent recovery sequence* $x_n \to x$ for every x.

4.3 Examples and comparison with other convergences

Example 10. Take $\mathbb{X} = \mathbb{R}$. Let

$$F_n(x) = nxe^{-2n^2x^2}.$$

Then clearly, $F_n \overset{\text{pw}}{\to} 0$, i.e. the identically zero function, as the exponential beats the linear term. In order to compute the Γ-limit, a good starting point is to *start by looking for a lower bound. If our lower bound is sharp enough, that would usually be the* Γ*-limit.* Clearly, if $x \neq 0$, $F_n(x) \to 0$. Thus, if $x_n \to x \neq 0$, then $F_n(x_n) \to 0$. To find a lower bound when $x = 0$, let us put $t = nx$ and set $f(t) = te^{-2t^2}$. It is easy to establish that $f(t)$ has a global minima at $t = -\frac{1}{2}$ and $f(-\frac{1}{2}) = -\frac{1}{2}e^{-\frac{1}{2}}$. Thus, we must have

$$F_n(x_n) \geq -\frac{1}{2}e^{-\frac{1}{2}} \qquad \text{for all } n \in \mathbb{N} \text{ and any sequence } \{x_n\}.$$

For constructing the recovery sequence, note that f takes this value when $t = -\frac{1}{2}$, thus taking $nx_n = -\frac{1}{2}$, i.e. choosing $x_n = -\frac{1}{2n}$ gives the recovery sequence, as

$$-\frac{1}{2n} \to 0 \qquad \text{and} \qquad F_n(-\frac{1}{2n}) = -\frac{1}{2}e^{-\frac{1}{2}} \quad \text{for all } n \in \mathbb{N}.$$

Thus, we have $F_n \overset{\Gamma}{\to} F$, where

$$F(x) = \begin{cases} 0 & \text{if } x \neq 0, \\ -\frac{1}{2}e^{-\frac{1}{2}} & \text{if } x = 0. \end{cases}$$

Thus the Γ-limit is different from the pointwise limit.

Example 11. Take $\mathbb{X} = \mathbb{R}$. Let

$$F(x) = \begin{cases} nxe^{-2n^2x^2} & \text{if } n \text{ is odd}, \\ 2nxe^{-2n^2x^2} & \text{if } n \text{ is even}. \end{cases}$$

Then it is easy to check that $F_n \overset{\text{pw}}{\to} 0$, but Γ-limit does not exist, as we have

$$\Gamma\text{-}\liminf F_n = \begin{cases} -e^{-\frac{1}{2}} & \text{if } x = 0, \\ 0 & \text{if } x \neq 0, \end{cases} \qquad \text{and} \qquad \Gamma\text{-}\liminf F_n = \begin{cases} -\frac{1}{2}e^{-\frac{1}{2}} & \text{if } x = 0, \\ 0 & \text{if } x \neq 0. \end{cases}$$

Example 12. Take $\mathbb{X} = \mathbb{R}$. Let $F_n(x) = nxe^{nx}$. Then $F_n \overset{\Gamma}{\to} F$, but $F_n \overset{\text{pw}}{\to} \widetilde{F}$, where

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ -\frac{1}{e} & \text{if } x = 0, \\ +\infty & \text{if } x > 0, \end{cases} \qquad \text{and} \qquad \widetilde{F}(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ +\infty & \text{if } x > 0. \end{cases}$$

Example 13. Take $\mathbb{X} = \mathbb{R}$. Let $F_n(x) = \tan^{-1}(nx)$. Then $F_n \overset{\Gamma}{\to} F$, but $F_n \overset{\text{pw}}{\to} \widetilde{F}$, where

$$F(x) = \begin{cases} -\frac{\pi}{2} & \text{if } x \leq 0, \\ +\frac{\pi}{2} & \text{if } x > 0, \end{cases} \qquad \text{and} \qquad \widetilde{F}(x) = \begin{cases} -\frac{\pi}{2} & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ +\frac{\pi}{2} & \text{if } x > 0. \end{cases}$$

Now we give two simple examples in a general topological spaces.

Example 14. Let $F_n = a_n$ for all $n \in \mathbb{N}$, where $\{a_n\} \subset \mathbb{R}$. Then

$$\Gamma\text{-}\liminf F_n = \liminf_n a_n \qquad \text{and} \qquad \Gamma\text{-}\limsup F_n = \limsup_n a_n.$$

Thus, F_n Γ-converges if and only if $\{a_n\}$ is convergent and in that case, the Γ-limits of F_n is the constant function $F \equiv \lim_n a_n$.

Example 15. Let $F_n = F$ for all $n \in \mathbb{N}$. Then $F_n \overset{\text{pw}}{\to} F$ and $F_n \overset{\Gamma}{\to} F_{\text{lsc}}$. Furthermore, if F is continuous, then $F_n \overset{\text{cc}}{\to} F$.

This last example should be obvious from the fact that Γ-limits, indeed even upper and lower Γ-limits are automatically lsc.

Γ-convergence is tailored to minimization. The price we pay for this is the fact that unlike usual limits, Γ-limits are not linear. Γ-limits are more prone to 'going down'(*towards the minima* if you will) than 'going up'. To see this, we again return to $\mathbb{R}$.

Example 16. Let $\mathbb{X} = \mathbb{R}$ and let $F_n(x) = \sin(nx)$ and $G_n = -\sin(nx)$. Now note that $\sin(nx) \geq -1$ for all n and all x. Now fix $x \in \mathbb{R}$ and we want to construct a recovery sequence. Since $\sin(-\frac{\pi}{2} + 2\pi k) = -1$ for any $k \in \mathbb{Z}$, we want to find integers k_n such that

$$x_n = x + \frac{1}{n}\left[-nx - \frac{\pi}{2} + 2\pi k_n\right] \to x.$$

Obviously, it is sufficient to find integers k_n such that $-nx - \frac{\pi}{2} + 2\pi k_n$ stays bounded. But this is easy to do and left as an exercise to the reader to find a formula for k_n involving the floor function (greatest integer less than or equal to). Thus, $F_n \overset{\Gamma}{\to} -1$. The same argument also proves $G_n \overset{\Gamma}{\to} -1$. Clearly, neither F_n or G_n converge pointwise. However, $F_n + G_n = 0$ for all n and thus, $F_n + G_n \overset{\text{cc}}{\to} 0$. But this implies

$$F_n + G_n \overset{\Gamma}{\to} 0 \neq -2 = \Gamma\text{-}\lim F_n + \Gamma\text{-}\lim G_n.$$

In fact, sums of two Γ-convergent sequence need to Γ-converge.

Example 17. Let $\mathbb{X} = \mathbb{R}$ and let $F_n(x) = \sin(nx)$ and $G_n = (-1)^n \sin(nx)$. Then as above, we have $F_n, G_n \overset{\Gamma}{\to} -1$. But using the same argument, it is easy to check that

$$\Gamma\text{-}\liminf\,(F_n + G_n) = -2 \qquad \text{and} \qquad \Gamma\text{-}\limsup\,(F_n + G_n) = 0.$$

Though Γ-convergence 'misbehaves' in many aspects, the one good thing is that usually they already misbehave in the simplest possible situation, here, already on $\mathbb{R}$, so at least the examples are simple. In general, one can only prove if $F_n \overset{\Gamma}{\to} F$, $G_n \overset{\Gamma}{\to} G$ *and* $F_n + G_n \overset{\Gamma}{\to} H$, then $F + G \leq H$. Note that this is again in accordance with the principle that Γ-limits are more prone to 'going down'. We will not prove this result (interested readers can see Dal Maso [3]), but prove a simpler one.

Proposition 14. Let $F_n \overset{\Gamma}{\to} F$ and $G_n \overset{\text{cc}}{\to} G$. Then $F_n + G_n \overset{\Gamma}{\to} F + G$. In particular, if G is continuous, then $F_n + G \overset{\Gamma}{\to} F + G$.

Proof. Note that $F_n \xrightarrow{\Gamma} F$ and $G_n \xrightarrow{\mathrm{cc}} G$ implies that for any sequence $x_n \to x$, we have

$$F(x) \leq \liminf_n F_n(x_n) \qquad \text{and} \qquad G(x) = \lim_n G_n(x_n).$$

Adding the two yields the liminf inequality for $F + G$. For the recovery sequence, we can just use the recovery sequence for F. The last conclusion follows from the fact that if G is continuous, then taking $G_n = G$ for all n, we have $G_n \xrightarrow{\mathrm{cc}} G$. $\square$

As a consequence, we have the following important corollary.

Corollary 2. Let $\mathbb{X}$ be a Banach space. Let $F_n \xrightarrow{\Gamma} F$. Let $\xi \in \mathbb{X}^*$ and set

$$G_n^\xi(x) := F_n(x) - \langle \xi, x \rangle \qquad \text{for } x \in \mathbb{X},$$

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

$$G_n^\xi \xrightarrow{\Gamma} G = F(\cdot) - \langle \xi, \cdot \rangle .$$

Proof. Follows immediately from Proposition 14 in view of the fact that $x \mapsto \langle \xi, x \rangle$ is continuous for $\xi \in \mathbb{X}^*$. $\square$

Remark 25. The same argument also shows that when $\mathbb{X}$ is Banach and $F_n \overset{\Gamma}{\rightharpoonup} F$, then for $\xi \in \mathbb{X}^*$, we have

$$F_n(\cdot) - \langle \xi, \cdot \rangle \overset{\Gamma}{\rightharpoonup} F(\cdot) - \langle \xi, \cdot \rangle .$$

The following example is also a consequence of Proposition 14.

Example 18. Let $\mathbb{X} = \mathbb{R}$. Then

$$F_n(x) = \frac{1}{2}x^2 - \cos(nx) \quad \xrightarrow{\Gamma} \quad F = \frac{1}{2}x^2 - 1.$$

For any $\lambda \in \mathbb{R}$, we have

$$F_n(x) = \frac{\lambda}{2}x^2 - \lambda\cos(nx) \quad \xrightarrow{\Gamma} \quad F = \frac{\lambda}{2}x^2 - |\lambda|.$$

This last statement shows the failure of linearity under multiplication by a scalar as well.

The conclusions can go horribly wrong if G is not continuous.

Example 19. Let $\mathbb{X} = \mathbb{R}$ and let $F_n = \min\{|nx - 1|, 1\}$ and let

$$G(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \neq 0. \end{cases}$$

Then $F_n \xrightarrow{\Gamma} G$, but $F_n + G \xrightarrow{\Gamma} G + 1$, which differs from $G + G$ at $x = 0$, the point of discontinuity of G.

Example 20. Let $\mathbb{X} = \mathbb{R}$ and let $F_n = \tan^{-1}(nx + (-1)^n)$. Let

$$G(x) = \begin{cases} \pi & \text{if } x < 0, \\ 0 & \text{if } x \geq 0, \end{cases} \qquad \text{and} \qquad F(x) = \begin{cases} -\dfrac{\pi}{2} & \text{if } x \leq 0, \\ \dfrac{\pi}{2} & \text{if } x > 0. \end{cases}$$

Then $F_n \xrightarrow{\Gamma} F$, but

$$\Gamma\text{-}\liminf (F_n + G)(x) = \begin{cases} -\dfrac{\pi}{4} & \text{if } x = 0, \\ \dfrac{\pi}{2} & \text{if } x \neq 0, \end{cases}$$

and

$$\Gamma\text{-}\limsup (F_n + G)(x) = \begin{cases} \dfrac{\pi}{4} & \text{if } x = 0, \\ \dfrac{\pi}{2} & \text{if } x \neq 0. \end{cases}$$

Finally, we wish to show by an example that Γ-convergence depends crucially on the chosen topology. For this, we need a standard result about weak convergence in L^p.

Lemma 1 (Riemann-Lebesgue lemma). Let $1 \leq p \leq \infty$ and let $\Omega \subset \mathbb{R}^d$ be open and bounded. Let $f \in L^p(Q)$ be extended Q-periodically to all of $\mathbb{R}^d$. Define $f_n(x) = f(nx)$. Then

$$\begin{aligned} f_n &\rightharpoonup \bar{f} && \text{weakly in } L^p(\Omega) \text{ if } 1 \leq p < \infty, \\ f_n &\overset{*}{\rightharpoonup} \bar{f} && \text{weakly } * \text{ in } L^\infty(\Omega), \end{aligned}$$

where

$$\bar{f} \equiv \fint_Q f(y)\ \mathrm{d}y.$$

Example 21. Let $\Omega \subset \mathbb{R}^d$ be open, bounded and Lipschitz. Let $\mathbb{X} = L^2\left(\Omega; \mathbb{R}^d\right)$ and let $A \in L^\infty\left(\mathbb{R}^d; \mathbb{R}^{d\times d}_{\text{symm}}\right)$ be 1-periodic and uniformly elliptic, i.e. there exist constants $0 < \lambda \leq \Lambda < \infty$ such that

$$\lambda|\xi|^2 \leq \langle A(y)\xi, \xi\rangle \leq \Lambda|\xi|^2 \qquad \text{for all } \xi \in \mathbb{R}^d \text{ and for a.e. } x \in \mathbb{R}^d.$$

Let $A_n(x) := A(nx)$ and

$$F_n(u) := \frac{1}{2}\int_\Omega \langle A(nx)u(x), u(x)\rangle\ \mathrm{d}x \qquad \text{for } u \in \mathbb{X}.$$

We set

$$A_{\text{arith}} := \fint_{[0,1]^d} A(y)\ \mathrm{d}y \quad \text{and} \quad A_{\text{harm}} := \left(\fint_{[0,1]^d} A^{-1}(y)\ \mathrm{d}y\right)^{-1}.$$

We also denote

$$F_{\text{arith}}(u) := \frac{1}{2}\int_\Omega \langle A_{\text{arith}}u(x), u(x)\rangle\ \mathrm{d}x \quad \text{and} \quad F_{\text{harm}}(u) := \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u(x)\rangle\ \mathrm{d}x.$$

Then $F_n \xrightharpoonup{\Gamma} F_{\text{harm}}$ and $F_n \xrightarrow{\text{cc}} F_{\text{arith}}$ with respect to the strong topology of L^2 and as a consequence, $F_n \xrightarrow{\Gamma} F_{\text{arith}}$.

We first show $F_n \xrightharpoonup{\Gamma} F_{\text{harm}}$. Assume $u_n \rightharpoonup u$ in L^2. Then we have

$$\begin{aligned}
F_n(u_n) &= \frac{1}{2}\int_\Omega \langle A_n(x)u_n(x), u_n(x)\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \left\langle A_n(x)u_n(x), u_n(x) - [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&\qquad\qquad + \frac{1}{2}\int_\Omega \left\langle A_n(x)u_n(x), [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \left\langle A_n(x)\left(u_n(x) - [A_n(x)]^{-1} A_{\text{harm}}u(x)\right), u_n(x) - [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&\qquad + \frac{1}{2}\int_\Omega \left\langle A_n(x)\,[A_n(x)]^{-1} A_{\text{harm}}u(x), u_n(x) - [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&\qquad\qquad\qquad + \frac{1}{2}\int_\Omega \left\langle A_n(x)u_n(x), [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&:= I_1^n + I_2^n + I_3^n.
\end{aligned}$$

Now we have $I_1^n \geq 0$ for all n, by uniform ellipticity of A. For I_2^n, we compute

$$\begin{aligned}
I_2^n &= \frac{1}{2}\int_\Omega \left\langle A_{\text{harm}}u(x), u_n(x) - [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u_n(x)\rangle \,\mathrm{d}x - \frac{1}{2}\int_\Omega \left\langle A_{\text{harm}}u(x), [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&\to \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u(x)\rangle \,\mathrm{d}x - \frac{1}{2}\int_\Omega \left\langle A_{\text{harm}}u(x), [A_{\text{harm}}]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u(x)\rangle \,\mathrm{d}x - \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u(x)\rangle \,\mathrm{d}x = 0,
\end{aligned}$$

since $u_n \rightharpoonup u$ weakly in L^2, $A_n^{-1} \xrightharpoonup{*} A_{\text{harm}}^{-1}$ weakly $*$ in L^∞ by Lemma 1 and the inner product is symmetric. Now using the symmetry of A and that of the inner product again, we have

$$\begin{aligned}
I_3^n &= \frac{1}{2}\int_\Omega \left\langle A_n(x)u_n(x), [A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \left\langle u_n(x), A_n(x)\,[A_n(x)]^{-1} A_{\text{harm}}u(x)\right\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \langle u_n(x), A_{\text{harm}}u(x)\rangle \,\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u_n(x)\rangle \,\mathrm{d}x \to \frac{1}{2}\int_\Omega \langle A_{\text{harm}}u(x), u(x)\rangle \,\mathrm{d}x.
\end{aligned}$$

Thus, we have the lim inf inequality

$$F_{\text{harm}}(u) \leq \liminf_n F_n(u_n).$$

Clearly, this im inf would be a limit if we have $I_1^n = 0$ for all n. Thus, for the recovery sequence, given $u \in \mathbb{X}$, it is natural to choose $u_n(x) = [A_n(x)]^{-1} A_{\text{harm}}u(x)$. Then $I_1^n = 0$ for all n and we have

$$F_{\text{harm}}(u) = \lim_n F_n(u_n).$$

On the other hand, it is easy to see that $u_n \rightharpoonup u$. Indeed, for any $\phi \in \mathbb{X}$, we have

$$\int_\Omega \left\langle [A_n(x)]^{-1} A_{\text{harm}} u(x), \phi(x) \right\rangle \,\mathrm{d}x \to \int_\Omega \left\langle A_{\text{harm}}^{-1} A_{\text{harm}} u(x), \phi(x) \right\rangle \,\mathrm{d}x = \int_\Omega \langle u(x), \phi(x) \rangle \,\mathrm{d}x,$$

where we have used the fact that components of $A_{\text{harm}} u$ and ϕ are in L^2 and thus by Hölder inequality, their products are in L^1, and can thus be used as test functions for the weak $*$ convergence $A_n^{-1} \overset{*}{\rightharpoonup} A_{\text{harm}}^{-1}$ in L^∞. This establishes $F_n \overset{\Gamma}{\rightharpoonup} F_{\text{harm}}$.

Now for any $u_n \to u$ in $\mathbb{X}$, we have

$$\begin{aligned}
F_n(u_n) &= \frac{1}{2} \int_\Omega \langle A_n(x) u_n(x), u_n(x) \rangle \,\mathrm{d}x \\
&= \frac{1}{2} \int_\Omega \langle A_n(x) u_n(x), u_n(x) - u(x) \rangle \,\mathrm{d}x + \frac{1}{2} \int_\Omega \langle A_n(x) u_n(x), u(x) \rangle \,\mathrm{d}x \\
&= \frac{1}{2} \int_\Omega \langle A_n(x) u_n(x), u_n(x) - u(x) \rangle \,\mathrm{d}x + \frac{1}{2} \int_\Omega \langle A_n(x) \left[u_n(x) - u(x) \right], u(x) \rangle \,\mathrm{d}x \\
&\qquad + \frac{1}{2} \int_\Omega \langle A_n(x) u(x), u(x) \rangle \,\mathrm{d}x \\
&:= J_1^n + J_2^n + J_3^n.
\end{aligned}$$

Now we have

$$\begin{aligned}
|J_1^n| &\le \frac{1}{2} \|A_n\|_{L^\infty} \|u\|_{L^2} \|u_n - u\|_{L^2} \to 0, \\
|J_2^n| &\le \frac{1}{2} \|A_n\|_{L^\infty} \|u\|_{L^2} \|u_n - u\|_{L^2} \to 0.
\end{aligned}$$

As components of u are in L^2, by Hölder inequality, their products are in L^1 and thus be used as test functions for the weak $*$ convergence $A_n \overset{*}{\rightharpoonup} A_{\text{arith}}$ in L^∞ Thus,

$$J_3^n = \frac{1}{2} \int_\Omega \langle A_n(x) u(x), u(x) \rangle \,\mathrm{d}x \to \frac{1}{2} \int_\Omega \langle A_{\text{arith}} u(x), u(x) \rangle \,\mathrm{d}x.$$

This implies $F_{\text{arith}}(u) = \lim_n F_n(u_n)$ for every sequence $u_n \to u$. This establishes the continuous convergence and the result.

‘

4.4 Compactness of Γ-convergence

In this section, we want to investigate the compactness properties of Γ-convergence. In other words, we want to investigate when a sequence admits a Γ-convergent subsequences.

Theorem 18 (Abstract compactness theorem). Let $\mathbb{X}$ be second countable. Then any sequence of functions $F_n : \mathbb{X} \to \overline{\mathbb{R}}$ admits a Γ-convergent subsequence.

Proof. The proof is just a diagonal argument. Let $\mathcal{B} := \{U_j\}_{j\in\mathbb{N}}$ be a countable base for the topology on $\mathbb{X}$. Since $\overline{\mathbb{R}}$ is compact, for every $j \in \mathbb{N}$, there exists a sunsequence $\{F_{n_{k,j}}\}$ such that

$$\lim_{k\to\infty} \inf_{y\in U_j} F_{n_{k,j}}(y) \qquad \text{exists in } \overline{\mathbb{R}}.$$

Thus, by a diagonal argument, we can extract a subsequence $\{F_{n_i}\}$ such that

$$\lim_{i\to\infty} \inf_{y\in U} F_{n_i}(y) \qquad \text{exists in } \overline{\mathbb{R}} \text{ for every } U \in \mathcal{B}.$$

Define

$$F(x) = \sup_{U\in\mathcal{B}(x)} \lim_{i\to\infty} \inf_{y\in U} F_{n_i}(y),$$

where $\mathcal{B}(x)$ denotes the collection of all elements in $\mathcal{B}$ that contains x. Then it is easy to check that we have $F_{n_i} \xrightarrow{\Gamma} F$. This completes the proof. □

This abstract result is not that useful in practice, as we are really only interested in the situation where the Γ-limit F is proper and $F : \mathbb{X} \to (-\infty, +\infty]$. Observe that we have used the compactness of $\overline{\mathbb{R}}$ in a crucial way. As a result, even if each F_n takes values in $(-\infty, +\infty)$ and is proper, the subsequential Γ-limit F can still take the value $-\infty$ and can even be identically equal to $-\infty$. This extremely pathological situation can be avoided if we assume a uniform lower bound for $\{F_n\}$, but F can still be identically $+\infty$. To this end, we start with the following theorem.

Theorem 19 (Compactness theorem in separable metric spaces). Let $(\mathbb{X}, d)$ be a *separable metric space*. Let $F_n : \mathbb{X} \to (-\infty, +\infty]$ be a sequence such that

(a) $\inf_n \inf_{\mathbb{X}} F_n > -\infty$ and

(b) there exists $x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$.

Then there exists a subsequence $\{F_{n_k}\}$ such that $F_{n_k} \xrightarrow{\Gamma_d} F$, for some $F : \mathbb{X} \to (-\infty, +\infty]$, which is proper.

Proof. The result is an easy consequence of our abstract theorem above. Since every separable metric space is second countable, we can apply Theorem 18 to conclude the existence of a subsequence and a Γ-limit F, where $F : \mathbb{X} \to \overline{\mathbb{R}}$. Thus, it only remains to show that such an F can never take the value $-\infty$ and has a finite value at least at one point. Observe that by (b), we have $\inf_n \inf_{\mathbb{X}} F_n < +\infty$. Combining this with (a), we get $\inf_n \inf_{\mathbb{X}} F_n := c \in \mathbb{R}$. We claim $F \geq c$. Indeed, for any $x \in \mathbb{X}$, there exists a recovery sequence $\{y_{n_k}\}$. Thus, we have

$$F(x) = \lim_k F_{n_k}(y_{n_k}) \geq \inf_n \inf_{\mathbb{X}} F_n = c.$$

To prove F is proper, we take the constant sequence $\{x_n = x_0\}$ and use the lim inf inequality to deduce

$$F(x_0) \leq \liminf_k F_{n_k}(x_0) \leq \sup_n F_n(x_0) < +\infty.$$

This completes the proof. □

Now we consider the case of Banach spaces.

Theorem 20. Let $\mathbb{X}$ be a separable reflexive Banach space. Let $F_n : \mathbb{X} \to (-\infty, +\infty]$ be a sequence such that there exists $\psi : \mathbb{X} \to (-\infty, +\infty]$ with the property that

$$\psi(x) \to +\infty \quad \text{as } \|x\|_{\mathbb{X}} \to +\infty \qquad \text{and} \qquad F_n \geq \psi \qquad \text{for all } n.$$

Assume

(a) $\inf_n \inf_{\mathbb{X}} F_n > -\infty$, (in particular, ψ is bounded below) and

(b) there exists $x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$.

Then there exists a subsequence $\{F_{n_k}\}$ such that $F_{n_k} \xrightharpoonup{\Gamma} F$, for some $F : \mathbb{X} \to (-\infty, +\infty]$, which is proper.

Proof. We divide the proof into three steps.

Step 1: Since $\mathbb{X}$ is separable reflexive space, there exists a metric d on $\mathbb{X}$ such that $(\mathbb{X}, d)$ is a separable metric space and the metric topology restricted to norm bounded subsets coincides with the weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$ restricted to these sets. Now we apply Theorem 18 to deduce that $\{F_n\}$ admits a Γ_d-convergent subsequence.

Step 2: Now we wish to show that $\{F_n\}$ also admits a Γ_{w}-convergent subsequence. If $\{F_{n_k}\}$ is the Γ_d-convergent subsequence, then working with $\{F_{n_k}\}$ instead of $\{F_n\}$, we can assume, without loss of generality, that $F_n \xrightarrow{\Gamma_d} F$, for some $F : \mathbb{X} \to \overline{\mathbb{R}}$. Now if $x_n \rightharpoonup x$, then by uniform boundedness principle, we have $\sup \lim_n \|x_n\|_{\mathbb{X}} < +\infty$. This implies $\{x_n\}$ lives in a norm-bounded subset and thus $x_n \rightharpoonup x$ implies $d(x_n, x) \to 0$. Hence the lim inf inequality for the weak topology follows from the one for d-topology. For the recovery sequence, observe that we know for any $x \in \mathbb{X}$, there exists a sequence $\{y_n\}$ such that $d(y_n, x) \to 0$ and $F(x) = \lim_n F_n(y_n)$. If $F(x) = +\infty$, we can just take the constant sequence $\{x_n = x\}$ as our recovery sequence. Hence, we suppose $F(x) < +\infty$. Thus, there exists $c \in \mathbb{R}$ such that up to discarding finitely many terms if necessary, we have

$$F_n(y_n) \leq c \qquad \text{for all } n \in \mathbb{N}.$$

This implies

$$\psi(y_n) \leq F_n(y_n) \leq c \qquad \text{for all } n \in \mathbb{N}.$$

As $\psi(x) \to +\infty$ as $\|x\|_{\mathbb{X}} \to +\infty$, this implies $\|y_n\|_{\mathbb{X}}$ is uniformly bounded. Since $\mathbb{X}$ is separable and reflexive, this implies $\{y_n\}$ admits a weakly convergent subsequence, say $\{y_{n_k}\}$. But since $d(y_n, x) \to 0$ and d induces the weak topology on norm-bounded sets, this implies that the weak limit can only be x. Hence there exists a subsequence $\{y_{n_k}\}$ such that $y_{n_k} \rightharpoonup x$ and we have

$$F(x) = \lim_n F_n(y_n) = \lim_k F_{n_k}(y_{n_k}).$$

This implies $F_{n_k} \xrightharpoonup{\Gamma} F$.

Step 3: Finally, we wish to show that F can not take the value $-\infty$ and must be proper. If ψ is bounded below, we have

$$\inf_n \inf_{\mathbb{X}} F_n \geq \inf_{\mathbb{X}} \psi > -\infty.$$

Now both conclusions follow exactly as in the proof of Theorem 19. This completes the proof. $\square$

For later reference, we record a few definitions.

Definition 29. Let $F_n : \mathbb{X} \to \overline{\mathbb{R}}$. We say $\{F_n\}$ is *(sequentially) equi-coercive* on $\mathbb{X}$ if for every $t \in \mathbb{R}$, there exists a closed and (sequentially) compact subset $K_t \subset \mathbb{X}$ such that

$$\{F_n \leq t\} \subset K_t \qquad \text{for all } n \in \mathbb{N}.$$

$\{F_n\}$ is called *(sequentially) equi-mildly coercive* if there exists a closed and (sequentially) compact subset $K \subset \mathbb{X}$ such that

$$\inf_{\mathbb{X}} F_n = \inf_K F_n \qquad \text{for all } n \in \mathbb{N}.$$

Remark 26. Recall and compare with Definition 4 and Definition 5.

Using the same arguments as in Theorem 4 and Theorem 5, one can prove the following result.

Proposition 15. Let $\mathbb{X}$ be a separable reflexive Banach space. Let $F_n : \mathbb{X} \to (-\infty, +\infty]$. Then $\{F_n\}$ is sequentially equi-coercive with respect to the weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$ if and only there exists a $\psi : \mathbb{X} \to (-\infty, +\infty]$ with the property that

$$\psi(x) \to +\infty \quad \text{as } \|x\|_{\mathbb{X}} \to +\infty \qquad \text{and} \qquad F_n \geq \psi \qquad \text{for all } n.$$

We will also need a slightly stronger property.

Definition 30 (Superlinear function). A function $\phi : [0, \infty) \to (-\infty, +\infty]$ is called *superlinear* if

$$\lim_{t \to +\infty} \frac{\phi(t)}{t} = +\infty.$$

Definition 31. Let $\mathbb{X}$ be a separable reflexive Banach space. A sequence of functions $F_n : \mathbb{X} \to (-\infty, +\infty]$ is called *equi-superlinear* if there exists a *superlinear function* $\phi : [0, \infty) \to (-\infty, +\infty]$ such that

$$F_n(x) \geq \phi\left(\|x\|_{\mathbb{X}}\right) \qquad \text{for all } x \in \mathbb{X} \text{ and for all } n \in \mathbb{N}.$$

Remark 27. By setting $\psi(x) = \phi\left(\|x\|_{\mathbb{X}}\right)$, it is easy to see that equi-superlinearity implies sequential equi-coercivity with respect to the weak topology. However, a uniform lower bound by a function $\psi(x) = \|x\|_{\mathbb{X}}^{\alpha}$ with $\alpha \in (0, 1]$, would still imply sequential equi-coercivity with respect to the weak topology, but not equi-superlinearity.

4.5 Convergence of minima and minimizers

We now discuss the convergence of minima and minimizers under Γ-convergence. We have already discussed the arguments at the beginning as a motivation for the definition of Γ-convergence. So we would state the precise results, but only prove one of them.

Theorem 21 (Metric space selection principle version). Let $(\mathbb{X}, d)$ be a metric space. Let $F_n, F : \mathbb{X} \to (-\infty, +\infty]$ be such that $F_n \xrightarrow{\Gamma} F$. Assume there exists $x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$ and $\{F_n\}$ is equi-mildly coercive. Let $K \subset \mathbb{X}$ be the compact set given by equi-mild coercivity. Then

(i) F is proper and attains its minimum in K,

(ii) $\min_{\mathbb{X}} F = \lim_n \inf_{\mathbb{X}} F_n$,

(iii) there exists a $\{\frac{1}{n}\}$-almost minimizing sequence $\{x_n\} \subset K$ for $\{F_n\}$ such that $\{x_n\}$ admits a convergent subsequence and any subsequential limit of $\{x_n\}$ converges to a minimizer of F.

Further, every minimizer of F can be recovered, i.e. for each $\bar{x} \in \mathcal{M}(F)$, there exists a sequence $\{y_n\} \subset \mathbb{X}$ such that $y_n \to \bar{x}$ and $\lim_n F_n(y_n) = F(\bar{x}) = \inf_{\mathbb{X}} F$.

Proof. First note that by the existence of recovery sequence, the last conclusion is valid for any minimizer of F. Furthermore, for every $x \in \mathbb{X}$, there exists a sequence $y_n \to x$ such that

$$\limsup F_n(y_n) \leq F(x).$$

But clearly,

$$\inf_{\mathbb{X}} F_n \leq F_n(y_n).$$

Thus, we have

$$\limsup \left(\inf_{\mathbb{X}} F_n \right) \leq \limsup F_n(y_n) \leq F(x).$$

As this must hold for every $x \in \mathbb{X}$, taking inf over $x \in \mathbb{X}$, we deduce

$$\limsup \left(\inf_{\mathbb{X}} F_n \right) \leq \inf_{\mathbb{X}} F. \tag{23}$$

Now since

$$\inf_{\mathbb{X}} F_n = \inf_K F_n \qquad \text{for all } n \in \mathbb{N},$$

for each n, we can find $x_n \in K$ such that

$$F_n(x_n) \leq \inf_K F_n + \frac{1}{n}.$$

Since K is compact and thus sequentially compact (as $(\mathbb{X}, d)$ is a metric space), $\{x_n\}$ admits a convergent subsequence. Suppose $\{x_{n_k}\}$ be *any* convergent subsequence of $\{x_n\}$ and suppose $x_{n_k} \to \bar{x}$. Then we have

$$\begin{aligned}
\inf_{\mathbb{X}} F \leq F(\bar{x}) \leq \liminf_k F_{n_k}(x_{n_k}) &\leq \liminf_k \left[\inf_K F_{n_k} + \frac{1}{n_k} \right] \\
&= \liminf_k \inf_K F_{n_k} \\
&\leq \limsup_k \inf_K F_{n_k} \\
&\leq \limsup_n \inf_K F_n = \limsup_n \inf_{\mathbb{X}} F_n \overset{(23)}{\leq} \inf_{\mathbb{X}} F.
\end{aligned}$$

Thus, all of the above inequalities must be equalities. Using the constant sequence $\{x_n = x_0\}$, we have

$$F(x_0) \leq \liminf_n F_n(x_0) \leq \sup_n F_n(x_0) < +\infty.$$

This completes the proof. $\square$

Remark 28. In general, one can replace the assumption $\sup_n F_n(x_0) < +\infty$ by slightly weaker one, e.g. one can only assume there is a subsequence $\{F_{n_k}\}$ such that $\sup_k F_{n_k}(x_0) < +\infty$. In this case, we just replace $\{F_n\}$ by $\{F_{n_k}\}$. But we can not weaken this assumption too much. As an example, if we only assume there exists a sequence $\{x_n\}$ such that $\sup_n F_n(x_n) < +\infty$, this is not enough to ensure F is proper. Indeed, one can take $\mathbb{R}$ and take F_n as the function which is $+\infty$ everywhere except the point $x = n$, with $F_n(n) = 0$. Then the Γ-limit F is identically $+\infty$. On the other hand, if one assume the existence of a *convergent sequence*, i.e. $\{x_n\}$ such that $\sup_n F_n(x_n) < +\infty$ and $x_n \to x_0$, then this is enough. Indeed, by the liminf inequality, we deduce

$$F(x_0) \leq \liminf_n F_n(x_n) \leq \sup_n F_n(x_n) < +\infty.$$

Remark 29. Note that the theorem *does not claim* that *every minimizer of F can be recovered by a sequence in K.* Indeed, take $\mathbb{X} = \mathbb{R}$ and let $F_n = F$ for all n, where $F(x) = x^2(x-1)^2$. Then $\mathcal{M}(F) = \{0, 1\}$. If we take $K = \{0\}$, then 1 can not be recovered by sequences in K.

Remark 30. The theorem also *does not claim* that every ε_n-almost minimizing sequence for F_n will admit a subsequence. The theorem only asserts the existence of one such sequence. Take $\mathbb{X} = \mathbb{R}$ and let

$$F_n(x) = \begin{cases} 0 & \text{if } x = 0, \\ \dfrac{1}{n} & \text{if } x = n, \\ 1 & \text{otherwise.} \end{cases}$$

Take $K = \{0\}$. Then $\inf_{\mathbb{X}} F_n = \inf_K F_n = 0$. But the sequence $\{x_n = n\}$ is $\{\frac{1}{n}\}$-almost minimizing, but admits no convergent subsequence.

The above result describes the typical situation where Γ-convergence can be used as a *selection principle* to select a minimizer for F. By now, it should be clear how to prove the following results. We leave this as an exercise.

Theorem 22 (Metric space equi-coercive version). Let $(\mathbb{X}, d)$ be a metric space. Let $F_n, F : \mathbb{X} \to (-\infty, +\infty]$ be such that $F_n \xrightarrow{\Gamma} F$. Assume there exists $x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$ and $\{F_n\}$ is equi-coercive. Then

(i) F is proper and attains its minimum,

(ii) $\min_{\mathbb{X}} F = \lim_n \inf_{\mathbb{X}} F_n$,

(iii) any $\{\varepsilon_n\}$-almost minimizing sequence for $\{F_n\}$ admits a convergent subsequence and any such subsequential limit is a minimizer of F.

Further, every minimizer of F can be recovered.

Theorem 23 (Banach space equi-coercive version). Let $\mathbb{X}$ be a separable reflexive Banach space. Let $F_n, F : \mathbb{X} \to (-\infty, +\infty]$ be such that $F_n \xrightarrow{\Gamma} F$. Assume there exists

$x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$ and there exists a $\psi : \mathbb{X} \to (-\infty, +\infty]$ with the property that

$$\psi(x) \to +\infty \quad \text{as } \|x\|_{\mathbb{X}} \to +\infty \qquad \text{and} \qquad F_n \geq \psi \qquad \text{for all } n.$$

Then

(i) F is proper and attains its minimum,

(ii) $\min_{\mathbb{X}} F = \lim_n \inf_{\mathbb{X}} F_n$,

(iii) any $\{\varepsilon_n\}$-almost minimizing sequence for $\{F_n\}$ admits a convergent subsequence and any such subsequential limit is a minimizer of F.

Further, every minimizer of F can be recovered.

Theorem 24 (Banach space equi-superlinear version). Let $\mathbb{X}$ be a separable reflexive Banach space. Let $F_n, F : \mathbb{X} \to (-\infty, +\infty]$ be such that $F_n \xrightarrow{\Gamma} F$. Assume there exists $x_0 \in \mathbb{X}$ such that $\sup_n F_n(x_0) < +\infty$ and $\{F_n\}$ is equi-superlinear. For any $\xi \in \mathbb{X}^*$, define

$$G_n^\xi(x) = F_n(x) - \langle \xi, x \rangle \qquad \text{and} \qquad G^\xi(x) = F(x) - \langle \xi, x \rangle \qquad \text{for all } x \in \mathbb{X}.$$

Then

(i) $G^\xi : \mathbb{X} \to (-\infty, +\infty]$ is proper and attains its minimum,

(ii) $\min_{\mathbb{X}} G^\xi = \lim_n \inf_{\mathbb{X}} G_n^\xi$,

(iii) any $\{\varepsilon_n\}$-almost minimizing sequence for $\{G_n^\xi\}$ admits a convergent subsequence and any such subsequential limit is a minimizer of G^ξ.

Further, every minimizer of G^ξ can be recovered.

Proof. Clearly, For any $\xi \in \mathbb{X}^*$, $G_n^\xi \xrightarrow{\Gamma} G^\xi$. Since $F : \mathbb{X} \to (-\infty, +\infty]$, it is easy to see that G^ξ also takes values in $(-\infty, +\infty]$. We also have

$$\sup_n G_n^\xi(x_0) = \sup_n \left[F_n(x_0) - \langle \xi, x_0 \rangle\right] \leq \sup_n F_n(x_0) + \|\xi\|_{\mathbb{X}^*} \|x_0\|_{\mathbb{X}} < +\infty.$$

Furthermore, by equi-superlinearity, we deduce

$$\begin{aligned} G_n^\xi(x) &= F_n(x) - \langle \xi, x \rangle \\ &\geq F_n(x) - |\langle \xi, x \rangle| \\ &\geq F_n(x) - \|\xi\|_{\mathbb{X}^*} \|x\|_{\mathbb{X}} \geq \phi\left(\|x\|_{\mathbb{X}}\right) - \|\xi\|_{\mathbb{X}^*} \|x\|_{\mathbb{X}} = \|x\|_{\mathbb{X}} \left[\frac{\phi\left(\|x\|_{\mathbb{X}}\right)}{\|x\|_{\mathbb{X}}} - \|\xi\|_{\mathbb{X}^*}\right]. \end{aligned}$$

Thus, we can apply the last theorem to G_n^ξ to deduce the result. $\square$

4.6 Mosco and Γ-convergence: effect of compact embeddings

Now we discuss an useful result, which concerns the effect of compact embeddings on Γ-convergence.

Theorem 25. Let $\mathbb{X}, \mathbb{Y}$ be reflexive Banach spaces such that the embedding of $\mathbb{Y} \hookrightarrow \mathbb{X}$ is *compact.* Let $F_n : \mathbb{Y} \to (-\infty, +\infty]$ be extended to all of $\mathbb{X}$ by defining $F_n \equiv +\infty$ in $\mathbb{X} \setminus \mathbb{Y}$. Assume there exists a $\psi : \mathbb{Y} \to (-\infty, +\infty]$ such that

$$\psi(y) \to +\infty \quad \text{as} \quad \|y\|_{\mathbb{Y}} \to +\infty \qquad \text{and} \qquad F_n \geq \psi \quad \text{for all } n \in \mathbb{N}.$$

Then $F_n \overset{\Gamma}{\rightharpoonup} F$ in $\mathbb{Y}$ if and only if $F_n \overset{\mathrm{M}}{\to} F$ in $\mathbb{X}$.

Proof. We first show the 'if' part. This does not need the compactness of the embedding. Assume $F_n \overset{\mathrm{M}}{\to} F$ in $\mathbb{X}$. Choose a sequence $x_n \rightharpoonup x$ in $\mathbb{Y}$. As long as $\mathbb{Y} \hookrightarrow \mathbb{X}$ is *continuous*, this implies $x_n \rightharpoonup x$ in $\mathbb{X}$ as well, by comparing the weak topology of $\mathbb{Y}$ itself and the restriction of the weak topology of $\mathbb{X}$ restricted to $\mathbb{Y}$. Thus, the lim inf inequality follows. Now let $y \in \mathbb{Y}$. If $F(y) = +\infty$, then we take the constant sequence $\{x_n = y\}$ as the recovery sequence. If $F(y) < +\infty$, then as $F_n \overset{\mathrm{M}}{\to} F$ in $\mathbb{X}$, there exists a sequence $\{x_n\} \subset \mathbb{X}$ such that $x_n \to x$ and $F(y) = \lim_n F_n(x_n)$. Since $F(y) < +\infty$ and $F(y) = \lim_n F_n(x_n)$, we infer that $\{F_n(x_n)\}$ is uniformly bounded. By our assumptions, this implies $\{x_n\} \subset \mathbb{Y}$ (as F_ns are identically $+\infty$ outside $\mathbb{Y}$) and $\{\psi(x_n)\}$ is uniformly bounded and consequently, $\{\|x_n\|_{\mathbb{Y}}\}$ is uniformly bounded. As $\mathbb{Y}$ is reflexive, this implies there exists a subsequence $\{x_{n_k}\}$ such that x_{n_k} is convergent. But the weak limit can only be y. Hence $x_{n_k} \rightharpoonup y$ in $\mathbb{Y}$ and this is a recovery sequence for y.

Now we prove the 'only if' part. Assume $F_n \overset{\Gamma}{\rightharpoonup} F$ in $\mathbb{Y}$ and take any sequence $x_n \rightharpoonup x$ in $\mathbb{X}$. If $\liminf_n \|x_n\|_{\mathbb{Y}} = +\infty$ (with the convention that we write $\|z\|_{\mathbb{Y}} = +\infty$ if $z \in \mathbb{X} \setminus \mathbb{Y}$), then $F_n(x_n) \geq \psi(x_n) \to +\infty$ and the lim inf inequality holds trivially, as the right hand side is $+\infty$. On the other hand, if there exists a subsequence $\{x_{n_k}\}$ such that $\|x_{n_k}\|_{\mathbb{Y}}$ is uniformly bounded. This implies, by the same argument as above, that there exists a further subsequence $x_{n_{k_i}} \rightharpoonup x$ in $\mathbb{Y}$. Then since $F_n \overset{\Gamma}{\rightharpoonup} F$ in $\mathbb{Y}$, we have

$$F(x) \leq \liminf_i F_{n_{k_i}}(x_{n_{k_i}}).$$

But applying this to the subsequence that realizes the lim inf in $\liminf_n F_n(x_n)$, we have our desired lim inf inequality. For the recovery sequence, take $x \in \mathbb{X}$. If $x \in \mathbb{X} \setminus \mathbb{Y}$, the we can choose the constant sequence $\{x_n = x\}$. If $x \in \mathbb{Y}$, then as $F_n \overset{\Gamma}{\rightharpoonup} F$ in $\mathbb{Y}$, we know there exists a sequence $\{x_n\} \subset \mathbb{Y}$ such that $x_n \rightharpoonup x$ in $\mathbb{Y}$ and $F(x) = \lim_n F_n(x_n)$. But by the compactness of the embedding,

$$x_n \rightharpoonup x \quad \text{in} \quad \mathbb{Y} \implies x_n \to x \quad \text{in} \quad \mathbb{X}.$$

This gives our desired recovery sequnce and completes the proof. □

4.7 Γ-convergence and convex duality

4.7.1 Convexity and Γ-convergence

So far, our discussion of Γ-convergence did not involve convexity at all. This is useful for applications, as one often employs Γ-convergence to study minimization of nonconvex variational problems, where one can not apply the direct methods directly to F. But in this section, we shall be discussing Γ-convergence in the specific context of *dissipation potentials* (for gradient systems).

We begin with a result that is easy to prove.

Proposition 16. Let $\mathbb{X}$ be a locally convex Hausdorff topological vector space. Let $F_n : \mathbb{X} \to (-\infty, +\infty]$ is convex for all $n \in \mathbb{N}$. Then $\Gamma\text{-}\limsup F_n$ is convex. In particular, if $F_n \xrightarrow{\Gamma} F$, then F is convex.

Remark 31. In general, $\Gamma\text{-}\liminf F_n$ need not be convex. As an example, take $\mathbb{X} = \mathbb{R}$ and let

$$F_n(x) := \begin{cases} x & \text{if } n \text{ is even,} \\ -x & \text{if } n \text{ is odd.} \end{cases}$$

Then each F_n is convex (indeed affine) and continuous, but $\Gamma\text{-}\liminf F_n(x) = -|x|$, which is nonconvex.

Exercise 15. Prove Proposition 16. If you are unfamiliar with topological vector spaces, at least prove this result when $(\mathbb{X}, d)$ is metric and when $\mathbb{X}$ is a Banach space equipped with the weak topology $\sigma(\mathbb{X}; \mathbb{X}^*)$.

4.7.2 Dissipation potentials

Recall that Γ-limits are always lsc with respect to the topology.

Definition 32 (Dissipation potentials). Let $\mathbb{X}$ be a Banach space. A function $F : \mathbb{X} \to [0, +\infty]$ is called a *dissipation potential* on $\mathbb{X}$ if F is convex, lsc with respect to the strong topology on $\mathbb{X}$ and $F(0) = 0$.

Theorem 26. Let $\mathbb{X}$ be a Banach space. Let $\{F_n\}$ be a sequence of dissipation potentials on $\mathbb{X}$. Then $\{F_n^*\}$ defines a sequence of dissipation potential on $\mathbb{X}^*$. Furthermore, if $F_n \xrightharpoonup{\Gamma} F$, then F is a dissipation potential on $\mathbb{X}$ and F^* is a dissipation potential on $\mathbb{X}^*$.

Proof. Fix n. Now F_n^* is automatically convex and strongly lsc on $\mathbb{X}^*$, by properties of Legendre-transforms. Thus, we only need to show F_n^* is nonnegative and $F_n^*(0) = 0$. Now, since $F_n(0) = 0$, we have

$$F_n^*(x^*) = \sup_{x \in \mathbb{X}} \{\langle x^*, x\rangle - F_n(x)\} \geq \langle x^*, 0\rangle - F_n(0) = 0.$$

Furthermore, since F_n is nonnegative and $F_n(0) = 0$, we have

$$F_n^*(0) = \sup_{x \in \mathbb{X}} \{\langle 0, x\rangle - F_n(x)\} = -\inf_{x \in \mathbb{X}} F_n(x) = 0.$$

This proves F_n is a dissipation potential. But n is arbitrary.

Now since $F_n \xrightharpoonup{\Gamma} F$, F is automatically convex and wlsc and consequently lsc. SInce F_ns are all nonnegative, by the existence of recovery sequences, it is easy to show that F is nonnegative as well. Moreover, by choosing the constant sequence $\{x_n = 0\}$ in the lim inf inequality, we have $F(0) \leq \liminf F_n(0) = 0$. Since F is nonenagtiuve, this shows $F(0) = 0$. Once again, F^* is convex and strongly lsc on $\mathbb{X}^*$, by properties of Legendre transform. Then F^* is nonnegative and $F^*(0) = 0$ follows by the same arguments that we used above for F_n^*. This completes the proof. $\square$

4.7.3 Continuity of Legendre transform with respect to Γ-convergence

Now we want to discuss the relation between Γ-convergence of $\{F_n\}$ and Γ-convergence of $\{F_n^*\}$. But before stating the result, first let us discuss what kind of a result would be useful? First of all, our hypothesis should be Γ-convergence of $\{F_n\}$ *with respect to the weak topology*, as this is really the situation in practice. What is the conclusion we are looking for? If we obtain

$$F_n \xrightharpoonup{\Gamma} F \implies F_n^* \xrightharpoonup{\Gamma} F^*,$$

for a separable refelxive Banach space (or replacing weak topology by the weak $*$ topology in $\mathbb{X}^*$), would this be enough for applications. Unfortunately no. The reason is that duality bracket between $\mathbb{X}$ and $\mathbb{X}^*$ is *strong-weak* and *weak-strong* continuous, but not *weak-weak* continuous.

Proposition 17. Let X be a Banach space. Then $\langle \xi_n, x_n\rangle \to \langle \xi, x\rangle$ if either of the following holds.

(i) $\xi_n \xrightharpoonup{*} \xi$ in $\mathbb{X}^*$ and $x_n \to x$ in $\mathbb{X}$.

(ii) $x_n \rightharpoonup x$ in $\mathbb{X}$ and $\xi_n \to \xi$ in $\mathbb{X}^*$.

Proof. We just prove (ii), as the other one is similar. We have

$$\langle \xi_n, x_n\rangle = \langle \xi_n - \xi, x_n\rangle + \langle \xi, x_n\rangle \leq |\langle \xi_n - \xi, x_n\rangle| + \langle \xi, x_n\rangle \leq \|\xi_n - \xi\|_{\mathbb{X}^*}\|x_n\|_{\mathbb{X}} + \langle \xi, x_n\rangle.$$

As $x_n \rightharpoonup x$, by Uniform Boundedness principle, $\{\|x_n\|_{\mathbb{X}}\}$ is uniformly bounded. Thus, the first term converges to zero. The second term obviously converges to $\langle \xi, x\rangle$. □

Remark 32. It is easy to see that the duality bracket is not weak-weak continuous. Indeed, even if $\mathbb{X} = H$, a separable Hilbert space, where we identify $H \simeq H^*$ via the Riesz isomorphism. Let $\{e_n\} \subset H$ be an orthonormal basis. Then taking $x_n = e_n = \xi_n$, we see that $e_n \rightharpoonup 0$. Indeed, for any $x \in H$, by the Parseval identity, we have

$$\|x\|_H^2 = \sum_n |\langle x, e_n\rangle|^2.$$

The summability of the series implies we must have $\langle x, e_n\rangle \to 0$. Since this holds for any $x \in H$, this means $x_n \rightharpoonup 0$. But clearly $\langle e_n, e_n\rangle = \|e_n\|_H^2 = 1 \neq 0$.

In view of this discussion, the following result should seem well-motivated.

Theorem 27. Let $\mathbb{X}$ be a separable reflexive Banach space. Let $\{F_n\}$ be a sequence of dissipation potentials on $\mathbb{X}$. Then

$$F_n \xrightharpoonup{\Gamma} F \qquad \text{if and only if} \qquad F_n^* \xrightarrow{\Gamma} F^*.$$

This is a deep result and we are not going discuss a complete proof of this result. Interested readers can see [1, Chapter 3]. However, we are going to discuss a bit the general strategy.

The proof goes via proving the following four implications.

(a) weak - lim inf for F_n $\implies$ strong - lim sup for F_n^*.

$$\text{(b) strong - }\limsup \text{ for } F_n \quad \Longrightarrow \quad \text{strong - }\liminf \text{ for } F_n^*.$$

$$\text{(c) strong - }\liminf \text{ for } F_n^* \quad \Longrightarrow \quad \text{weak - }\limsup \text{ for } F_n.$$

$$\text{(d) strong - }\limsup \text{ for } F_n^* \quad \Longrightarrow \quad \text{weak - }\liminf \text{ for } F_n.$$

Here weak and strong refers to the weak and strong topology, respectively. Also, lim inf denotes the lim inf inequality and lim sup stands for the existence of a recovery sequence. Note that as the topologies are different, we can not just prove two of them and them dualize to get the other two. As a rule , lim sup implies lim inf type implications are much easier to prove, as we not have to prove existence of a sequence. Thus, we first prove (b). Proof of (d) is similar and simpler.

Proof of (b): Consider $\xi_n \to \xi$ in $\mathbb{X}^*$. Now, given any $\varepsilon > 0$, there exists $x_0 \in \mathbb{X}$ such that

$$F^*(\xi) := \sup_{x\in\mathbb{X}} [\langle \xi, x\rangle - F(x)] \le \varepsilon + \langle \xi, x_0\rangle - F(x_0).$$

Now by (w-lim sup for F_n), there exists a recovery sequence $x_n^0 \rightharpoonup x_0$ in $\mathbb{X}$ such that $F(x_0) = \lim_n F_n(x_n^0)$. Thus, we have

$$\begin{aligned}
\liminf_n F_n^*(\xi_n) &\ge \liminf_n \left[\langle \xi_n, x_n^0\rangle - F_n(x_n^0)\right] \\
&= \lim_n \langle \xi_n, x_n^0\rangle - \lim_n F_n(x_n^0) = \langle \xi, x_0\rangle - F(x_0) \ge F^*(\xi) - \varepsilon.
\end{aligned}$$

As $\varepsilon > 0$ is arbitrary, we have our result.

Proof of (a) assuming equi-superlinearity: Proof of (a) and (c) are much more difficult. To explain the main idea, we prove (a) under the additional assumption that $\{F_n\}$ is equi-superlinear. Under this assumption, we can use constant sequences as our recovery sequences. Indeed, let $\xi \in \mathbb{X}^*$ and consider the constant sequence $\{\xi_n = \xi\}$. Since each F_n is superlinear, lsc and convex, the map $x \mapsto F_n(x) - \langle \xi, x\rangle$ has at least one minimizer. Consequently, there exists $\{\bar{x}_n\}$ such that

$$F_n^*(\xi) = \langle \xi, \bar{x}_n\rangle - F_n(\bar{x}_n) \qquad \text{for all } n \in \mathbb{N}.$$

By equi-superlinearity, $\{\|\bar{x}_n\|_{\mathbb{X}}\}$ is uniformly bounded. Now let $\{F_{n_k}^*\}$ be a subsequence that realizes the lim sup in $\limsup F_n^*(\xi)$, i.e.

$$F_{n_k}^*(\xi) \to \limsup F_n^*(\xi).$$

Since $\{\|\bar{x}_{n_k}\|_{\mathbb{X}}\}$ is uniformly bounded, we can extract a further subsequence $\bar{x}_{n_{k_i}}$ such that $\bar{x}_{n_{k_i}} \rightharpoonup \tilde{x}$. Now we have

$$\begin{aligned}
-F^*(\xi) &= \inf_{x\in\mathbb{X}} [F(x) - \langle \xi, x\rangle] \\
&\le F(\tilde{x}) - \langle \xi, \tilde{x}\rangle \\
&\le \liminf \left[F_{n_{k_i}}(\bar{x}_{n_{k_i}}) - \langle \xi, \bar{x}_{n_{k_i}}\rangle\right] \\
&= \liminf \left[-F_{n_{k_i}}^*(\xi)\right] = -\limsup F_{n_{k_i}}^*(\xi) = -\limsup F_n^*(\xi).
\end{aligned}$$

This proves that the constant sequence $\{\xi_n = \xi\}$ is a recovery sequence and completes the proof of (a).

The difficulty of proving the general result lies in the fact that the map $x \mapsto F_n(x) - \langle \xi, x \rangle$ need not admit a minimizer. For (c), we can not ensure that the map $x^* \mapsto F_n^*(x^*) - \langle x^*, x \rangle$ admits a minimizer. What is even worse is that for (c), assuming something like equi-superlinearity is not reasonable, as assumptions that imply equi-superlinearity for $\{F_n\}$ as well as for $\{F_n^*\}$ are way too strong to be useful.

The basic strategy depends on using what are known as *Moreau-Yosida approximations.*

Theorem 28 (Properties of Moreau-Yosida approximations). Let $\mathbb{X}$ be a normed linear space. Let $F : \mathbb{X} \to (-\infty, +\infty]$ be convex, lsc and proper. Then for every $\lambda > 0$, the function defined by

$$F_\lambda(x) := \inf_{z \in \mathbb{X}} \left[F(z) + \frac{1}{2\lambda} \|x - z\|_{\mathbb{X}}^2 \right], \tag{24}$$

is convex and continuous. Its convex conjugate is given by

$$(F_\lambda)^* (\xi) = F^*(\xi) + \frac{\lambda}{2} \|\xi\|_{\mathbb{X}^*}^2 .$$

Moreover, the sequence $\{F_\lambda\}_{\lambda > 0}$ increases to F as λ decreases to 0.

Remark 33. The functions defined by (24) are called *Moreau-Yosida approximations* of F.

In view of these properties, the strategy to prove Theorem 27 in the general case is to replace $\{F_n\}$ by their Moreau-Yosida approximations $\{(F_n)_\lambda$ for a fixed $\lambda > 0$. Then the sequences are equi-superlinear. In the end, we pass to the limit as $\lambda \to 0$.

References

[1] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[2] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011.

[3] Gianni Dal Maso. *An introduction to* Γ*-convergence*, volume 8 of *Progress in Nonlinear Differential Equations and their Applications.* Birkhäuser Boston, Inc., Boston, MA, 1993.

[4] Walter Rudin. *Functional analysis.* McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973.