Inverse and homogenization problems for maximal monotone operators

by

Ramón Zárate Sáiz

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES
(Mathematics)

The University Of British Columbia
(Vancouver)

June 2010

Abstract

We apply self-dual variational calculus to inverse problems, optimal control problems and homogenization problems in partial differential equations.

Self-dual variational calculus allows for the variational formulation of equations which do not have to be of Euler-Lagrange type. Instead, a monotonicity condition permits the construction of a self-dual Lagrangian. This Lagrangian then permits the construction of a non-negative functional whose minimum value is zero, and its minimizer is a solution to the corresponding equation.

In the case of inverse and optimal control problems, we use the variational functional given by the self-dual Lagrangian as a penalization functional, which naturally possesses the ideal qualities for such a role. This allows for the application of standard variational techniques in a convex setting, as opposed to working with more complex constrained optimization problems. This extends work pioneered by Barbu and Kunisch.

In the case of homogenization problems, we recover existing results by dal Maso, Piat, Murat and Tartar with the use of simpler machinery. In this context self-dual variational calculus permits one to study the asymptotic properties of the potential functional using classical Γ convergence techniques which are simpler to handle than the direct techniques required to study the asymptotic properties of the equation itself. The approach also allows for the seamless handling of multivalued equations.

The study of such problems introduces naturally the study of the topological structures of the spaces of maximal monotone operators and their corresponding self-dual potentials. We use classical tools such as Γ convergence, Mosco convergence and Kuratowski-Painlevé convergence and show that these tools are well

suited for the task. Results from convex analysis regarding these topologies are extended to the more general case of maximal monotone operators in a natural way. Of particular interest is that the Γ convergence of self-dual Lagrangians is equivalent to the Mosco convergence, and this in turn implies the Kuratowski-Painlevé convergence of their corresponding maximal monotone operators; this partially extends a classical result by Attouch relating the convergence of convex functions to the convergence of their corresponding subdifferentials.

Table of Contents

Acknowledgments

This work was carried out with the support of a fellowship from CONACYT (Consejo Nacional de Ciencia y Tecnologia) and the support offered by my advisor, Nassif Ghoussoub, to whom I am most thankful for his guidance, also.

Special thanks must be given to Abbas Moameni, for the months spent working together where of great value.

I am also grateful to Philip Loewen, of my supervisory committee, with whom I had fruitful discussions at various points.

I thank the other members of my supervisory committee, Stephen Gustaphson and Richard Froese for their valuable efforts.

Finally I want to specially thank Itzia, my wife, from whom I received immense peace and support during these last years.

Chapter 1

Introduction

This work deals with the application of variational techniques to inverse and homogenization problems for certain types of partial differential equations. Even if the approach is variational, these equations do not have to be of Euler-Lagrange type (the classical variational setting), rather the variational principle used here relies on *Fenchel-Young duality*. Examples include, but are by no means limited to, skew adjoint vector fields superimposed on a gradient of some vector potential, and parabolic equations.

This variational theory and its applicability to solve partial differential equations was developed in a series of papers by Ghoussoub, culminating in the book [Gho08a], and will be the main tool. The techniques involve a dose of convex analysis on phase space, which is not a typical technique in the study of partial differential equations, but the approach unifies under the same theory the study of a big class of partial differential equations.

One of the goals of this work is to show how this unified theory can be used and applied to the study of inverse and homogenization problems associated to such families of equations.

The connection between inverse and optimal control problems and homogenization problems is that all of them call for the study of asymptotic properties of the families of equations, in particular of their topological structure. The variational setting provides interesting, well understood tools for such a purpose, and which are compatible with the spirit of generality of the theory.

Let us identify some individual topics involved in this work.

1.1 Maximal monotone operators and variational methods for partial differential equations

Monotone operators and *maximal monotone operators* were introduced as non linear generalizations of linear positive definite operators. They were studied in detail by many authors, including Brezis, Browder, Nirenberg and Rockafellar, among others. They now constitute an important tool for modern non linear analysis.

Their study is classically connected with *convex analysis*, since they can be considered as extensions of subdifferentials of convex functions: it is a well known fact that convex subdifferentials are themselves maximal monotone operators.

Of more recent origin is the fact that the connection goes beyond that: The monotonicity condition can actually be seen as a convexity condition on *phase space*, and this can be used to construct convex potentials on this larger space that "realize" a specific vector field through a variational principle, without being constrained to being the Euler-Lagrange equation of any potential.

This variational principle has two ingredients: First, and most obvious, is the minimization of certain functional. Second, the requirement that the minimum of such a functional has to be a specific value (which is naturally normalized to zero). It is in this context that the condition of self-duality is chosen in the work of Ghoussoub, as it actually grants the needed conditions for this value to be achieved, without *a priori* knowledge of existence of solutions. This very simple feature (the fact that the value of the minimized functional is actually zero) will be the main idea to exploit in the development of a general variational approach to inverse and optimal control problems.

1.2 Inverse problems

Partial differential equations are a prime tool for modelling nature. A lot of engineering, physical and biological problems are modelled with partial differential equations.

A direct problem for partial differential equations consists of, given specified parameters which represent properties in the modelled phenomena, finding a solution to the equation corresponding to those parameters, and thus, predicting the behaviour of such a system from the knowledge of its parameters.

An inverse problem consists of, given a solution for the partial differential equation, that is, a known behaviour of the modelled system, to recover the parameters defining the system. Inverse problems allow, in principle, to recover hidden or internal data from a system given experimental observations from it; this plays a fundamental role in modern science and engineering. They can be used in medicine, where, for example, Magnetic Resonance Imaging is used to recover information on the inner tissue of a patient without direct observation that would require a surgical procedure; inverse spectral methods are used in astrophysics to study the composition of stellar bodies which are completely impossible to explore by direct means. Inverse problems can also be used in the design of devices where the interaction of inner mechanics are understood, but one needs to control the behaviour of such devices by the appropriate selection of parameters that represent design decisions. Another typical example is the exploration of oil and mineral deposits under the crust of the earth, minimizing needless and costly excavation. Needless to say, inverse problems are mathematically very challenging by their very nature.

The starting point of the theory developed here is the work by Barbu and Kunisch, where they approach the problem via a least squares approach. The difficulty to such approaches is that the study of solution manifolds to families of partial differential equations is inherently complicated. To ease this difficulty the approach consists of using an appropriate functional that penalizes parameters not being in such solution set, but allowing us to work in the original ambient spaces which are typically convex, and thus allow for the use of classic variational techniques. Then as the cost of this penalization grows, one expects to find gradually better approximations to solutions to the system that minimize the penalized least squares problem.

This is where the variational principle mentioned above plays a fundamental role, as this functional is constructed in such a way that it is itself an excellent penalization to the solution set, being a positive functional whose zeroes are the solutions to the equations.

1.3 Homogenization theory

Homogenization theory for partial differential equations is the study of asymptotic behaviour of families of partial differential equations. A typical setting is when complicated behaviour in the system (say, highly oscillating parameters) can be approximated by simple behaviour (averaging of such parameters, giving *homogenized* simpler parameters), but one then needs to formally justify that such approximation is consistent. It is not enough that equations are approximated, one needs to show that the corresponding solutions are also being approximated.

Classically, for the case of variational equations (equations of Euler-Lagrange type) the potential functionals are easier to study, and from their limiting behaviour one deduces the behaviour of their minimizers, thus obtaining information about the asymptotic behaviour of the equations and solutions themselves. This obviously fails as soon as the equations involved are not variational. The approach then required is more technical and complicated.

In particular the homogenization of monotone vector fields has been widely studied by non variational techniques by many authors, including dal Maso, Murat, Tartar and others. Here we use the more general variational setting and show that traditional variational proofs, simpler in nature, work just as well for these cases. The principal idea here is that the topological structure of monotone families is tightly connected with the topological structure of their corresponding potentials. The study of such topological properties is an important part of this work.

1.4 Variational convergence

Variational convergence refers to the study of convergence properties of functionals where the notion of convergence translates well to the convergence of their minimizers. A very classical notion of convergence for spaces of compact convex sets is the *Hausdorff metric*. In order to study the topological structure of a space of convex functionals, it is of advantage to identify these functionals with their epigraphs, which are convex sets. The compactness criteria, however, is an unacceptable limitation in this case. A more suitable choice of convergence criteria of general sets is that of *Kuratowski-Painlevé* convergence. This notion, adapted to infinite dimensional spaces and their inherent topologies, gives rise to Γ (Gamma) and Mosco convergence, both already classical tools in functional analysis, known to be appropriate notions of convergence to functionals, and particularly well behaved when working with convex families. To study the convergence properties of families of monotone maps, the raw notion of Kuratowski-Painlevé convergence can be used and has been shown to be the right choice by the work of authors like Kato, Brezis, Damlamian, Attouch and others. Of particular interest is the fact that for the case of convex potentials and convex subdifferentials, Mosco convergence of the potentials corresponds to Kuratowski-Painlevé convergence of their subdifferentials (modulo a normalizing condition). In the more general context of maximal monotone operators a lot of the correlations among the topological structure is preserved, in particular, the convergence of a family of the convex functionals associated to a family of maximal monotone operators, does imply the convergence of such operators. This is the heart of the homogenization result presented here.

1.5 Chapter contents

Parts of this work are taken from [GMZ10], which is joint work with Nassif Ghoussoub and Abbas Moameni. The layout of the work is as follows:

In **Chapter 2** the basic notions of convex analysis and the theory of maximal monotone operators that will be used throughout this work are introduced. How maximal monotone operators can be represented by a certain convex potential (*Fitzpatrick's function*) and the general notion of associating convex potentials to monotone operators is introduced here, which is one of the principal ingredients for the later parts of the work. Also a standard regularization technique for convex functions and for maximal monotone operators is introduced.

In **Chapter 3** a special class of convex potentials and its properties are introduced: the class of *self-dual Lagrangians*. We show how these special potentials can be constructed using the framework of Chapter 2. Also, a few important generalizations of classical convex analysis results to general maximal monotone operators are introduced here, for later use in the work, such as a generalization of the celebrated *Brønsted-Rockafellar Lemma*. The main variational notion to be used later is introduced here, where a positive functional is constructed in such a way that its zeroes correspond to solutions to a given partial differential equation. This will play a central role in the treatment of inverse and optimal control problems. Also we describe how to deal with space dependent operators appearing in partial differential equations and show their respective self-dual Lagrangians, which allow certain types of non variational equations to be solved using the aforementioned variational principle; this will be the main setting for the homogenization result presented here.

Chapter 4 is devoted to the study of specific topologies for spaces of convex functionals and spaces of maximal monotone operators and their connections through the theory introduced earlier. The notions involved are the already mentioned Γ, Mosco and Kuratowski-Painlevé convergence. We show how Mosco and Γ convergences are the same mode of convergence when restricted to the space of self-dual Lagrangians, and show how this mode of convergence of potential functionals implies Kuratowski-Painlevé convergence of their associated maximal monotone operators; this result is a partial generalization of *Attouch's Theorem*, which states the equivalence of the notions when restricted to convex subdifferentials modulo a normalizing condition. We include here a slight reformulation of the Theorem on phase space which removes the need for any normalizing condition. The converse implication, which does not hold, at least without some restrictions, is briefly discussed. Additionally, we include in this chapter various classical and some new results regarding the continuity of the regularization notions for convex functions and monotone operators alike. The final section of this chapter contains simple compactness results in a more specific setting for special classes of functionals for later reference in chapter 6.

Chapter 5 contains the main results regarding inverse and optimal control problems. The result is stated in an abstract setting to emphasize on its generality. It provides an unified approach under very general topological conditions for existence of optimal least squares solutions, and then generalized to optimal control problems. Several versions of the results are presented to briefly outline general situations regarding the required compactness, but the results are also intended as general guides and can be adapted to other situations not considered here for the sake of simplicity. The approach consists of minimizing a positive functional constructed via the variational functional from the variational principle of chapter 3 used as a penalization. This constructed functional is defined in the whole ambient space associated to the given equation, which removes the need to work in the manifold of solutions generated by the parameters; these minimizers are then shown to be approximate solutions which then converge to the optimal true solution to the least squares problem by making the penalty parameter increase.

Then, in **Chapter 6**, sample partial differential equations applications of various types are provided to illustrate the applicability of the results from the previous chapters, as well as to illustrate its broad scope in terms of the different possible approaches for parametrizing the families of equations. This chapter also outlines how certain partial differential equations are put in the context of self-dual Lagrangians.

Finally **Chapter 7** contains a homogenization result where the variational setting is used to study the asymptotic behaviour of a specific family of equations parametrized by the period of a given, possibly multivalued, monotone, space-dependent non-linearity. These non-linearities can be assumed to be non-variational, in the sense that they do not correspond to the potential of a functional. The result recovered is a known one, but the novelty lies in the use of more direct homogenization results for convex functionals similar to those needed for the classical variational setting. The periodic setting is chosen for simplicity and to be able to easily express the limit functionals and equations, but the approach is expected to be usable in more general settings.

Chapter 2

Basics of convex analysis and the theory of monotone operators

In this chapter we present some of the basics of convex analysis and the theory for maximal monotone operators. It is provided for convenience of the reader, and as reference, in the interest of making this work as self contained as possible.

Convex analysis and maximal monotone operators are traditionally related by the fact that maximal monotone operators are known to be generalizations of convex subdifferentials. Their connection however goes deeper than that, and some aspects will be revealed in this and following chapters.

Reference on the topics for this chapter can be found in [Roc70, Att84, BP86, Phe93, ET99].

2.1 Basic notation and setting

From here on, X will denote a reflexive Banach space and X^* will denote its dual.

The space $X \times X^*$ will be referred to as *phase space*.

When working in this abstract setting, we will use x, y, z to denote variables in X, and p, q, r to denote variables in X^*.

$\|\cdot\|$ will be used to denote the norms in X and X^*.

The duality pairing between $x \in X$ and $p \in X^*$ (that is, the value of the functional p at x) will be denoted as $\langle x, p \rangle$.

The duality mapping is the (possibly multivalued) map $J : X \rightrightarrows X^*$ defined by

$$Jx := \{p \in X^* \ : \ \|p\| = \|x\| \text{ and } \langle x, p\rangle = \|x\|^2\};$$

we will assume that the Banach space X is *uniformly convex*, in such a way that the mapping J is single valued. This mapping is known to be invertible whenever X is reflexive (see [BP86, Proposition 2.16]). It is also important to point out that this mapping is not linear, in general (only in the Hilbert case).

2.2 Convex analysis

A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *convex* if for any $0 \leq \lambda \leq 1$, and any $x, y \in X$, $f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y)$. It is said to be *proper* if f is not identically $+\infty$.

The function f is said to be *lower semi continuous* if the *epigraph* of f:

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

is a closed set.

The *class of convex, proper, lower semi continuous functions* on a Banach space X will be denoted by $\Gamma(X)$.

Consider $f : X \to \mathbb{R} \cup \{+\infty\}$, a proper lower semi continuous convex function. The *subdifferential* mapping $\partial f : X \rightrightarrows X^*$ is the multivalued mapping defined by

$$\partial f(x) := \{p \in X^* \ : \ f(y) \geq f(x) + \langle y - x, p\rangle, \quad \forall y \in X\}.$$

The duality mapping J can be seen as a subdifferential: $Jx = \partial\big(\frac{\|\cdot\|^2}{2}\big)(x)$, and similarly, its inverse $J^{-1} : X^* \to X$ corresponds to $J^{-1}p = \partial\big(\frac{\|\cdot\|^2}{2}\big)(p)$.

The *Fenchel-Legendre transform* (or *conjugate*) of f, denoted by f^*, is the convex function $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(p) := \sup_{x \in X} \{\langle x, p \rangle - f(x)\}.$$

The function f^* is a proper lower semi continuous convex function on X^*. f is proper convex lower semi continuous if and only if $f^{**} = f$.

2.2.1 Fenchel-Young inequality

The following important inequality, referred to as the *Fenchel-Young inequality*, holds:

$$f(x) + f^*(p) \geq \langle x, p \rangle. \tag{2.1}$$

The case when equality is achieved in the above can be characterized the following way:

$$f(x) + f^*(p) = \langle x, p \rangle \iff p \in \partial f(x) \iff x \in \partial f^*(p). \tag{2.2}$$

The above will be referred to as the *duality case* in the Fenchel-Young inequality.

Using Fenchel-Young duality, one can parametrize the subdifferential mapping the following way:

$$\partial f(x) := \{p \in X^* \ : \ f(x) + f^*(p) - \langle x, p \rangle = 0\}. \tag{2.3}$$

We will denote the graph of the subdifferential of a convex function f by ∂f, that is:

$$\partial f := \{(x, p) \in X \times X^* \ : \ p \in \partial f(x)\}.$$

2.2.2 ε-approximation of convex subdifferentials

From (2.3), and considering inequality (2.1), it seems natural to think of a point (x, p) to be approximating ∂f if $f(x) + f^*(p) - \langle x, p \rangle \leq \varepsilon$. This notion is formalized in the following definition.

Definition 2.1 *The* ε-subdifferential *of a convex function at a point $x \in X$, denoted*

$\partial_\varepsilon f(x)$, *is defined by*

$$\partial_\varepsilon f(x) := \{p \in X^* \,:\, f(x) + f^*(p) - \langle x, p\rangle \leq \varepsilon\}.$$

It will be also denoted:

$$\partial_\varepsilon f := \{(x,p) \in X \times X^* \,:\, p \in \partial_\varepsilon f(x)\}.$$

The ε-subdifferential is a useful notion of a "perturbation" or approximation to a convex subdifferential. How near a point in $\partial_\varepsilon f$ lies to a point in ∂f is made precise by the following celebrated result, best known as the *Brøndsted-Rockafellar Lemma*:

Theorem 2.2 *If p_0 is an ε-subgradient of f at x_0, (i.e. $p_0 \in \partial_\varepsilon f(x_0)$), then for any $\lambda > 0$ there exists a pair $(x_\varepsilon^\lambda, p_\varepsilon^\lambda) \in \partial f$ such that*

- $\|x_\varepsilon^\lambda - x_0\| \leq \frac{\sqrt{\varepsilon}}{\lambda}$
- $\|p_\varepsilon^\lambda - p_0\| \leq \lambda\sqrt{\varepsilon}$

This result appears in [BR65], a proof based on *Ekeland's variational principle* can be found in [Bee93]. The proof of the above is also contained in the proof provided here for Theorem 3.9.

2.2.3 Regularization of convex functions

Here is an important way to combine convex functions.

Definition 2.3 *Given $f, g \in \Gamma(X)$, their* inf-convolution*, denoted by $f \star g$, is the mapping given by*

$$f \star g(x) := \inf_{y\in X}\{f(y) + g(x-y)\}.$$

The interaction of the inf-convolution with the Fenchel-Legendre conjugate is described by the next well known result (see for example [Att84]).

Proposition 2.4 *Let f and g be functions in $\Gamma(X)$. Then*

$$(f \star g)^* = f^* + g^*.$$

If furthermore, dom$(f)-$dom(g) contains a neighbourhood of the origin, then also

$$(f+g)^* = f^* \star g^*.$$

Inf-convolution with the square of the norm produces the *Moreau-Yosida regularization* of a convex function.

Definition 2.5 *Let f be a function in $\Gamma(X)$. For each $\lambda > 0$, the function*

$$f_\lambda(x) := \inf_{y \in X}\{f(y) + \frac{\|x-y\|^2}{2\lambda}\}$$

is the Moreau-Yosida regularization *of f with parameter λ.*

The functions f_λ approximate f in the sense that (for $f \in \Gamma(X)$):

$$f(x) = \lim_{\lambda \to 0} f_\lambda(x) = \sup_{\lambda > 0} f_\lambda(x) \qquad \forall x \in X,$$

also f_λ is locally Lipschitz (see [Att84]). The regularization also possesses additional boundedness conditions: it is easy to see that $f_\lambda(x) \leq f(0) + \frac{\|x\|^2}{2\lambda}$.

The following is a basic property (see [Att84, Proposition 2.68]):

Proposition 2.6 *Let $f \in \Gamma(X)$ and let λ_1, λ_2 be positive constants. We have*

$$\left(f_{\lambda_1}\right)_{\lambda_2} = f_{\lambda_1+\lambda_2}.$$

The above can be shown from the following fact, which will be used later in the proof of a similar proposition.

Lemma 2.7 *For λ_1, $\lambda_2 > 0$:*

$$\inf_{z \in X}\{\frac{\|x-z\|^2}{2\lambda_1} + \frac{\|y-z\|^2}{2\lambda_2}\} = \frac{1}{2(\lambda_1+\lambda_2)}\|x-y\|^2.$$

Proof: The minimizer above is reached at a point $z_* \in X$ such that $\frac{1}{\lambda_1}(z_* - x) = \frac{1}{\lambda_2}(y - z_*)$, from this we compute $z_* = \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}\left(\frac{x}{\lambda_1} + \frac{y}{\lambda_2}\right)$. Then

$$z_* - y = \frac{\lambda_1}{\lambda_1+\lambda_2}(x-y)$$

and

$$x - z_* = \frac{\lambda_2}{\lambda_1 + \lambda_2}(x - y).$$

Using the above simply compute

$$\frac{\|x - z_*\|^2}{2\lambda_1} + \frac{\|y - z_*\|^2}{2\lambda_2} = \frac{1}{2(\lambda_1 + \lambda_2)}\|x - y\|^2.$$

□

2.3 Maximal monotone operators

It is convenient to identify a multivalued operator $T : X \rightrightarrows X^*$ with its graph in $X \times X^*$, this means that T will denote the set $\{(x, p) \in X \times X^* \;:\; p \in Tx\}$.

Definition 2.8 *A multivalued operator $T : X \rightrightarrows X^*$ is said to be* monotone *if for any $(x, p) \in T$ and any $(y, q) \in T$ the following holds:*

$$\langle x - y, p - q \rangle \geq 0.$$

If furthermore, T cannot be extended by another monotone operator, then T is said to be maximal monotone.

It is a known result due to Rockafellar that a convex subdifferential is itself a maximal monotone operator; not all maximal monotone operators, however, are convex subdifferentials. We will come back to this point later on.

The following result, known as *Minty's Theorem* in the Hilbert setting and extended to reflexive Banach spaces by Rockafellar, is an important characterization of maximal monotone operators:

Theorem 2.9 *Let $T : X \rightrightarrows X^*$ be a monotone operator. The following are equivalent:*

- *T is maximal monotone.*
- *For each $\lambda > 0$, the operator $T + \frac{1}{\lambda}J$ is surjective.*

2.3.1 Regularization of maximal monotone operators

According to the previous characterization, after translation, to each $x \in X$ there exists a unique solution, z, to

$$J(z-x)+\lambda Tz \ni 0. \tag{2.4}$$

Definition 2.10 *The* resolvent *of T with parameter λ is the operator $R_\lambda^T : X \to X$ given by $R_\lambda^T x := z$ where z is given by* (2.4).

The Yosida approximation *of T with parameter λ is the mapping $T_\lambda : X \to X^*$ given by*

$$T_\lambda x := \frac{1}{\lambda} J(x - R_\lambda^T x).$$

The above operators R_λ^T and T_λ are known to be continuous (see [Att84]). Convergence properties will be stated in chapter 4.

Remark 2.11 *Observe that $T_\lambda x \in T\left(R_\lambda^T x\right)$.*

In the case of convex subdifferentials, the following relation between the regularized operator and the regularized convex potential holds.

Proposition 2.12 *Let $f \in \Gamma(X)$, then*

$$R_\lambda^{\partial f} x = \arg\min_{y \in X} \{ f(y) + \frac{\|x-y\|^2}{2\lambda} \}.$$

That is: $f_\lambda(x) = f(R_\lambda^{\partial f} x) + \frac{\|(R_\lambda^{\partial f} x) - x\|^2}{2\lambda}$.

Also $\partial\left(f_\lambda\right) = \left(\partial f\right)_\lambda$.

2.4 Convex potentials for maximal monotone operators

2.4.1 Integration of convex subdifferentials

As mentioned above, convex subdifferentials are maximal monotone operators, however, not every maximal monotone operator is a convex subdifferential. This can be made precise by using the following definition.

Definition 2.13 *A multivalued operator $T : X \rightrightarrows X^*$ is said to be* cyclically monotone *if*

$$\langle x_0 - x_1, p_0\rangle + \cdots + \langle x_{n-1} - x_n, p_{n-1}\rangle + \langle x_n - x_0, p_n\rangle \geq 0$$

for any chain $(x_i, p_i) \in T$, $i = 0, \ldots, n$. It is said to be maximal *if in addition its a maximal monotone operator.*

The following is an important result by Rockafellar, it classifies convex subdifferentials among the class of maximal monotone operators and provides an *integration formula* to recover a convex function from its subdifferential.

Theorem 2.14 *A maximal monotone mapping T is* cyclically maximal monotone *if and only if there exist a convex function f such that $T = \partial f$.*

Furthermore, by fixing $x_0 \in dom(T)$, f can be recovered by the formula

$$f(x) = f(x_0) + \sup\{\langle x - x_n, p_n\rangle + \cdots + \langle x_1 - x_0, p_0\rangle\},$$

where the sup *above is taken over all finite sets $\{(x_i, p_i)\}_{i=1}^n$ with $(x_i, p_i) \in T$, $i = 0, \ldots, n$.*

Remark 2.15 *Observe that the value $f(x_0)$ could be chosen arbitrarily. Convex functions are characterized by their subdifferentials only up to an additive constant. In other words: $\partial f = \partial g$ if and only if for some constant c, $f = g + c$.*

2.4.2 Fitzpatrick's function

Let us make a simple observation in the monotonicity condition, T is monotone if for $(x, p), (y, q) \in T$:

$$0 \leq \langle x - y, p - q\rangle = \langle x, p\rangle + \langle y, q\rangle - \langle x, q\rangle - \langle y, p\rangle,$$

this can be rewritten as

$$\langle (x, p), (y, q)\rangle_{X\times X^*} - \langle y, q\rangle \leq \langle x, p\rangle.$$

The left hand side above, fixing $(y, q) \in T$, can be seen as a linear functional on $X \times X^*$. This is effectively a convexity condition in which the following construction is based:

Definition 2.16 *Given T a maximal monotone operator, its* Fitzpatrick function*,* $F_T : X \times X^* \to \mathbb{R} \cup \{\infty\}$ *is defined by*

$$F_T(x,p) := \sup_{(y,q)\in T} \{\langle x,q\rangle + \langle y,p\rangle - \langle y,q\rangle\}.$$

This construction plays a fundamental role in the relation between the theory of maximal monotone operators and convex analysis. It appeared originally in [Fit88]. The importance of Fitzpatrick's function is that it provides a counterpart "integration formula" for maximal monotone operators.

The cyclical monotonicity condition on a vector field can be made equivalent to that of conservative. Other maximal monotone vector fields are not integrable, in this sense.

The properties of F_T are summarized in the following proposition (see [Gho08a]).

Proposition 2.17 *If T is maximal monotone, then F_T is a proper convex function on $X \times X^*$ and the following inequalities hold:*

$$\langle x,p\rangle \le F_T(x,p) \le F_T^*(p,x).$$

Furthermore:

$$\langle x,p\rangle = F_T(x,p) = F_T^*(p,x) \iff (x,p) \in T \tag{2.5}$$

Remark 2.18 *Compare* (2.5) *with* (2.2)*. The maximal monotone operator T can be parametrized as*

$$T = \{(x,p) \in X \times X^* \ : \ F_T(x,p) = \langle x,p\rangle\} \tag{2.6}$$

which is an analogue to (2.3)*.*

The above justifies the notion that the function F_T is a "convex representative" of the operator T.

Definition 2.19 *A function F is said to be a* representative *for the operator T if both:*

1. $F(x,p) \geq \langle x,p \rangle$.
2. $F(x,p) = \langle x,p \rangle$ *if and only if* $(x,p) \in T$.

The following lemma appears in [SZ04, Lemma 1.1].

Lemma 2.20 *Let T be a maximal monotone operator. If (x,p) and (y,q) are elements of $X \times X^*$ such that $(x,p) \in \partial F_T(y,q)$, then*

$$\langle x-y, p-q \rangle \leq \inf_{(z,r)\in T} \langle x-z, p-r \rangle.$$

In particular $\langle x-y, p-q \rangle \leq 0$, and $\langle x-y, p-q \rangle = 0$ if and only if $(x,p) \in T$.

2.4.3 Regularization of convex functions on phase space

A direct analogue of the *Moreau-Yosida* regularization on phase space would simply be, for $F \in \Gamma(X \times X^*)$:

$$F_\lambda(x,p) := \inf_{(y,q)\in X\times X^*} \{F(y,q) + \frac{\|x-y\|^2 + \|p-q\|^2}{2\lambda}\}.$$

The following are other ways of regularizing a convex function on phase space:

Definition 2.21 *Let F be a function in $\Gamma(X \times X^*)$. Take $\lambda > 0$. Define the following:*

- $F^\lambda(x,p) := \inf_{(y,q)\in X\times X^*}\{F(y,q) + \frac{\|x-y\|^2}{2\lambda} + \frac{\lambda}{2}\|p-q\|^2\}$.
- $F_1^\lambda(x,p) := \inf_{y\in X}\{F(y,p) + \frac{\|x-y\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2\}$.
- $F_2^\lambda(x,p) := \inf_{q\in X^*}\{F(x,q) + \frac{\|x\|^2}{2\lambda} + \frac{\lambda}{2}\|p-q\|^2\}$.
- $F_{1,2}^\lambda(x,p) := \inf_{(y,q)\in X\times X^*}\{F(y,q) + \frac{\|x-y\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2 + \frac{\lambda}{2}\|p-q\|^2 + \frac{\|y\|^2}{2\lambda}\}$.

Remark 2.22 $F_{1,2}^\lambda(x,p) = \left(F_2^\lambda\right)_1^\lambda(x,p)$.

The analogue to proposition 2.6 is the following

Proposition 2.23 *For λ_1, $\lambda_2 > 0$, denote $\lambda_1 \star \lambda_2 := \frac{\lambda_1 \lambda_2}{\lambda_1+\lambda_2}$. We have*

$$\left(F_1^{\lambda_1}\right)_1^{\lambda_2} = F_1^{\lambda_1+\lambda_2}$$

and

$$\left(F_2^{\lambda_1}\right)_2^{\lambda_2} = F_2^{\lambda_1 \star \lambda_2}$$

To prove this we will use lemma 2.7 and its following analogue:

Lemma 2.24 *For λ_1, $\lambda_2 > 0$:*

$$\inf_{r\in X^*}\left\{\frac{\lambda_1}{2}\|p-r\|^2 + \frac{\lambda_2}{2}\|q-r\|^2\right\} = \frac{\lambda_1 \star \lambda_2}{2}\|p-q\|^2.$$

Proof: The minimizer above is a point $r_* \in X^*$ such that $\lambda_1(r_* - p) = \lambda_2(q - r_*)$, from this we compute $r_* = \frac{\lambda_1 p + \lambda_2 q}{\lambda_1+\lambda_2}$. Then

$$r_* - q = \frac{\lambda_2}{\lambda_1+\lambda_2}(p-q)$$

and

$$p - r_* = \frac{\lambda_1}{\lambda_1+\lambda_2}(p-q).$$

Using the above simply compute

$$\frac{\lambda_1}{2}\|p-r_*\|^2 + \frac{\lambda_2}{2}\|q-r_*\|^2 = \frac{\lambda_1 \star \lambda_2}{2}\|p-q\|^2.$$

□

Proof of proposition 2.23: Simply write

$$\begin{aligned}
\left(F_1^{\lambda_1}\right)_1^{\lambda_2}(x,p) &= \inf_{z\in X}\left\{F_1^{\lambda_1}(z,p) + \frac{\|x-z\|^2}{2\lambda_2} + \frac{\lambda_2}{2}\|p\|^2\right\} \\
&= \inf_{z\in X}\inf_{y\in X}\left\{F(y,p) + \frac{\|z-y\|^2}{2\lambda_1} + \frac{\lambda_1}{2}\|p\|^2 + \frac{\|x-z\|^2}{2\lambda_2} + \frac{\lambda_2}{2}\|p\|^2\right\} \\
&= \inf_{y\in X}\left\{F(y,p) + \inf_{z\in X}\left\{\frac{\|z-y\|^2}{2\lambda_1} + \frac{\|x-z\|^2}{2\lambda_2}\right\} + \frac{\lambda_1}{2}\|p\|^2 + \frac{\lambda_2}{2}\|p\|^2\right\},
\end{aligned}$$

which using lemma 2.7 gives

$$\begin{aligned}\left(F_1^{\lambda_1}\right)_1^{\lambda_2}(x,p) &= \inf_{y\in X}\{F(y,p)+\frac{1}{2(\lambda_1+\lambda_2)}\|x-y\|^2+\frac{\lambda_1+\lambda_2}{2}\|p\|^2\} \\ &= F_1^{\lambda_1+\lambda_2}(x,p).\end{aligned}$$

The second statement follows in an analogous fashion, using lemma 2.4.3. □

The following lemma in the spirit of proposition 2.12 regards the first regularization type and Fitzpatrick's function.

Lemma 2.25 *Let $(x,p)\in X\times X^*$ and $\lambda>0$ be such that*

$$\left(F_T\right)^{\lambda}(0,0)=F_T(x,p)+\frac{1}{2\lambda}\|x\|^2+\frac{\lambda}{2}\|p\|^2.$$

Then $(x,p)\in T$, and $p=-\frac{1}{\lambda}Jx$. That is: $x=R_\lambda^T 0$ and $p=T_\lambda 0$.

Proof: If (x,p) is as above, then we have that

$$(0,0)\in\partial F_T(x,p)+\partial\left(\frac{1}{2\lambda}\|x\|^2\right)+\partial\left(\frac{\lambda}{2}\|p\|^2\right).$$

Then there must be some (y,q) such that

$$(q,y)\in\partial F_T(x,p)$$

and

$$(-q,-y)=(\frac{1}{\lambda}Jx,\lambda J^{-1}p). \tag{2.7}$$

From lemma 2.20, we have that

$$\begin{aligned}0 \geq \langle x-y,p-q\rangle &= \langle x,-q\rangle+\langle -y,p\rangle+\langle x,p\rangle+\langle y,q\rangle \\ &= \frac{1}{\lambda}\|x\|^2+\lambda\|p\|^2+\langle x,p\rangle+\langle y,q\rangle.\end{aligned}$$

Also, since $\langle x,p\rangle\geq-\|x\|\|p\|$ and $\langle y,q\rangle\geq-\|y\|\|q\|=-\|x\|\|p\|$, we get

$$\langle x-y,p-q\rangle\geq\left(\sqrt{\lambda}\|x\|-\frac{1}{\sqrt{\lambda}}\|p\|\right)^2\geq 0.$$

Hence all inequalities are equalities above, we obtain that $(y,q) \in T$, but also

$$\frac{1}{\sqrt{\lambda}}\|x\| = \sqrt{\lambda}\|p\| = \frac{1}{\sqrt{\lambda}}\|y\| = \sqrt{\lambda}\|q\|,$$

and

$$\langle y,q\rangle = -\|\frac{1}{\sqrt{\lambda}}y\|\|\sqrt{\lambda}q\|;$$

these equalities imply that $-\frac{1}{\sqrt{\lambda}}y = J^{-1}\sqrt{\lambda}q$, that is $-\frac{1}{\lambda}Jy = q$, which implies, in view of (2.7), that $(y,q) = (x,p)$.

We can conclude $(x,p) \in T$ and $p = -\frac{1}{\lambda}Jx$. $\square$

Chapter 3

Self-dual Lagrangians and their vector fields

The notion of a *self-dual Lagrangian* is introduced here, even if the self-duality condition is not required in order to represent a given monotone operator (see definition 2.19) in many practical examples self-duality can be achieved naturally. It is also a condition that will be sufficient to guarantee existence of solutions for equations expressed by monotone operators via the variational principle introduced by Ghoussoub (without a priori knowledge of the existence of such solutions by different means).

3.1 Self-dual Lagrangians

Definition 3.1 *A function $L \in \Gamma(X \times X^*)$ is said to be a* self-dual Lagrangian *if for any $(x,p) \in X \times X^*$ it satisfies*

$$L(x,p) = L^*(p,x),$$

where L^* denotes the Fenchel-Legendre transform in $X \times X^*$.

The class of self-dual Lagrangians on $X \times X^*$ will be denoted by $\mathscr{L}(X)$.

The Fenchel-Young inequality gives that $L(x,p)+L^*(p,x) \geq \langle (x,p),(p,x)\rangle = 2\langle x,p\rangle$. From self-duality this is: $L(x,p) \geq \langle x,p\rangle$. Every self-dual Lagrangian then obeys:

$$L(x,p) \geq \langle x,p\rangle, \quad \forall (x,p) \in X \times X^*. \tag{3.1}$$

Note that self-dual Lagrangians can be translated: consider $M(x,q) := L(x,p+q) - \langle x,p\rangle$ or $N(y,p) := L(x+y,p) - \langle x,p\rangle$, which are themselves self-dual Lagrangians whenever L is.

3.2 Self-dual vector fields

The previous relation (3.1), in consideration of (2.3) and (2.6), motivates the following definition.

Definition 3.2 *Given a self-dual Lagrangian L, the associated* self-dual vector field *is the operator* $\overline{\partial}L : X \rightrightarrows X^*$ *given by*

$$\overline{\partial}Lx := \{p \in X^* \ : \ L(x,p) = \langle x,p\rangle\}.$$

The following is of fundamental importance (see [Gho08a]):

Proposition 3.3 *For any* $L \in \mathscr{L}(X)$*, the operator* $\overline{\partial}L$ *is maximal monotone.*

It can be seen now, that self-dual Lagrangians are potentials to (at least some) maximal monotone operators. The converse holds: it is possible to find a self-dual Lagrangian representing a given maximal monotone operator. This will be addressed in section 3.4.

3.3 Regularization of self-dual Lagrangians

Refering to definition 2.21, the following proposition is noteworthy. It states that self-dual Lagrangians can be regularized while preserving self-duality.

Proposition 3.4 *If* $L : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ *is a self-dual Lagrangian, then for each* $\lambda > 0$*, also* L_1^λ*,*L_2^λ *and* $L_{1,2}^\lambda$ *are self-dual Lagrangians.*

The above appears in [Gho08a, Lemma 3.2].

3.4 Construction of self-dual potentials for a maximal monotone operator

Fitzpatrick's function allows for a simple construction of a self-dual convex potential for any maximal monotone operator, by means of the *proximal average* of the Fitzpatrick's function and its conjugate.

Definition 3.5 *Consider a maximal monotone operator $T : X \rightrightarrows X^*$, fix $a > 1$ and let b be its conjugate exponent, that is: $\frac{1}{a} + \frac{1}{b} = 1$. We define the index a self-dual potential corresponding to T the following way:*

$$\begin{aligned} L_T^a(x,p) \quad = \quad & \inf\left\{\tfrac{1}{2}F_T(x_1,p_1) + \tfrac{1}{2}F_T^*(p_2,x_2) + \tfrac{1}{4a}\|x_1-x_2\|^a + \tfrac{1}{4b}\|p_1-p_2\|^b : \right. \\ & \left. (x,p) = \tfrac{1}{2}(x_1,p_1) + \tfrac{1}{2}(x_2,p_2)\right\}. \end{aligned} \tag{3.2}$$

A standard choice is $a = 2 = b$, and thus we will denote:

$$L_T := L_T^2.$$

The properties of L_T^a are summarized in the following

Proposition 3.6 *Let $T : X \rightrightarrows X^*$ be maximal monotone. Then for each $a > 1$, L_T^a is a proper lower semi continuous convex function on $X \times X^*$ and*

$$L_T^a(x,p) = \left(L_T^a\right)^*(p,x)$$

(i.e. $L_T^a \in \mathscr{L}(X)$),

$$L_T^a(x,p) \geq \langle x,p\rangle.$$

Furthermore:

$$\langle x,p\rangle = L_T^a(x,p) \iff (x,p) \in T \tag{3.3}$$

For the proof we will use the following fact (see [Gho08a, part (11.) of Proposition 2.6]): For $h(x) := \inf\{F(x_1,x_2) \ : \ x = \frac{x_1+x_2}{2}\}$, where $F \in \Gamma(X \times X)$, its Fenchel-Legendre conjugate is given by $h^*(p) = F^*(\frac{p}{2},\frac{p}{2})$.

Proof of proposition 3.3: We will show the self-duality of L_T^a first, using the previous fact. Denoting $G_1\big((x_1,p_1),(x_2,p_2)\big) := \frac{1}{2}F_T(x_1,p_1) + \frac{1}{2}F_T^*(p_2,x_2)$, and

$G_2\big((x_1,p_1),(x_2,p_2)\big) := \frac{1}{4a}\|x_1-x_2\|^a + \frac{1}{4b}\|p_1-p_2\|^b$. Hence, using the fact above:

$$\big(L_T^a\big)^*(p,x) = \Big(G_1+G_2\Big)^*\big((\frac{p}{2},\frac{x}{2}),(\frac{p}{2},\frac{x}{2})\big).$$

Direct computation yields $G_1^*\big((p_1,x_1),(p_2,x_2)\big) = \frac{1}{2}F_T^*(2p_1,2x_1)+\frac{1}{2}F_T(2x_2,2p_2)$, and

$$G_2^*\big((p_1,x_1),(p_2,x_2)\big) = \begin{cases} \frac{4^{a-1}}{a}\|x_1\|^a + \frac{4^{b-1}}{b}\|p_1\|^b & (x_1,p_1) = (-x_2,-p_2) \\ +\infty & \text{else.} \end{cases}$$

Since at least G_2 is continuous, the conjugate of the sum is the inf-convolution of the conjugates, giving that

$$\begin{aligned} \big(L_T^a\big)^*(p,x) &= \Big(G_1^*\star G_2^*\Big)\big((\frac{p}{2},\frac{x}{2}),(\frac{p}{2},\frac{x}{2})\big) \\ &= \inf_{(y_1,q_1),(y_2,q_2)}\{\frac{1}{2}F_T^*(2q_1,2y_1)+\frac{1}{2}F_T(2y_2,2q_2)+ \\ &\qquad + G_2^*\big((q_1-\frac{p}{2},y_1-\frac{x}{2}),(q_2-\frac{p}{2},y_2-\frac{x}{2})\big)\} \\ &= \inf_{(y_1,q_1)}\{\frac{1}{2}F_T^*(2q_1,2y_1)+\frac{1}{2}F_T(2x-2y_1,2p-2q_1)+ \\ &\qquad + \frac{4^{a-1}}{a}\|y_1-\frac{x}{2}\|^a + \frac{4^{b-1}}{b}\|q_1-\frac{p}{2}\|^b\}, \end{aligned}$$

which, denoting $(x_1,p_1) = (2y_1,2q_1)$ and $(x_2,p_2) = (2x-2y_1,2p-2q_1)$, finally returns

$$\big(L_T^a\big)^*(p,x) = L_T^a(x,p),$$

which establishes self-duality.

$L_T^a(x,p) \geq \langle x,p\rangle$ comes from self-duality and the Fenchel-Young inequality, as in (3.1). Given proposition 2.17, and the definition of L_T^a, it is clear that $L_T^a(x,p) = \langle x,p\rangle \iff F_T(x,p) = F_T^*(p,x) = \langle x,p\rangle \iff (x,p)\in T$. □

Remark 3.7 *Compare* (3.3) *with* (2.5) *and* (2.2). *T can be parametrized by*

$$T = \{(x,p)\in X\times X^* \;:\; L_T^a(x,p) = \langle x,p\rangle\}, \tag{3.4}$$

in an analogous fashion as in (2.3) *and* (2.6)*.*

The above listed results show that every maximal monotone operator has a self-dual Lagrangian that realizes it. Conversely, to any self-dual Lagrangian there is a maximal monotone operator being represented by it.

It must be noted that the self-dual Lagrangian representation of a maximal monotone operator is not unique (see also [BWY10] where specific examples are provided), however the maximal monotone operator corresponding to a self-dual Lagrangian is unique.

Example 3.8 *Two very illustrative and important examples of self-dual Lagrangians follow:*

1. *For $\Phi \in \Gamma(X)$, $L(x,p) := \Phi(x) + \Phi^*(p)$ is a self-dual Lagrangian and $\overline{\partial} L = \partial\Phi$.*

2. *Let now $S : X \to X^*$ be a linear continuous skew adjoint operator; $L(x,p) := \Phi(x) + \Phi^*(p - Sx)$ is a self-dual Lagrangian and $\overline{\partial} L = \partial\Phi + S$.*

3.5 ε-approximation of maximal monotone operators

As shown, inequality (3.1) can be deduced from self-duality directly from the original Fenchel-Young inequality. Going back to the definition of an ε-subdifferential, this allows for a natural way to extend that concept to maximal monotone operators:

Given T a maximal monotone operator and its corresponding potential, $L_T \in \mathscr{L}(X)$, for a parameter $\varepsilon > 0$ we can define an ε-approximant operator T_ε the following way:

$$T_\varepsilon x := \{p \in X^* \ : \ L_T(x,p) - \langle x,p \rangle \leq \varepsilon\}.$$

The importance of this notion lies in the following natural extension of the *Brøndsted-Rockafellar Lemma*:

Theorem 3.9 *Let $L : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a self-dual Lagrangian and assume that for a pair $(x_0, p_0) \in X \times X^*$, we have*

$$L(x_0, p_0) - \langle x_0, p_0 \rangle \leq \varepsilon.$$

Then, for any $\lambda > 0$, there exists a pair $(x_\varepsilon^\lambda, p_\varepsilon^\lambda) \in \bar{\partial} L$ such that

1. $\|x_\varepsilon^\lambda - x_0\| \leq \frac{\sqrt{\varepsilon}}{\lambda}$,
2. $\|p_\varepsilon^\lambda - p_0\| \leq \lambda\sqrt{\varepsilon}$,

Proof: First assume that M is a self-dual Lagrangian such that $M(0,0) \leq \varepsilon$. We claim that there exists then a pair $(y_\varepsilon^\lambda, q_\varepsilon^\lambda) \in \bar{\partial} M$ such that

1. $\|y_\varepsilon^\lambda\| \leq \frac{\sqrt{\varepsilon}}{\lambda}$,
2. $\|q_\varepsilon^\lambda\| \leq \lambda\sqrt{\varepsilon}$,

Indeed, consider J, the *duality mapping* from X to X^* and use the fact that $\bar{\partial} M$ is a maximal monotone operator and the characterization given by Theorem 2.9 to find $\tilde{x} \in X$ such that

$$-\lambda^2 J\tilde{x} \in \bar{\partial} M(\tilde{x}).$$

It follows that $M(\tilde{x}, -\lambda^2 J\tilde{x}) = \langle \tilde{x}, -\lambda^2 J\tilde{x} \rangle = -\lambda^2 \|\tilde{x}\|^2$. Now, since M is self-dual, we have

$$M(0,0) = M^*(0,0) = \sup_{(x,p) \in X \times X^*} -M(x,p) \geq -M(\tilde{x}, -\lambda^2 J\tilde{x}) = \lambda^2 \|\tilde{x}\|^2,$$

from which we obtain that $\|\tilde{x}\|^2 \leq \frac{\varepsilon}{\lambda^2}$. Since $\|\tilde{x}\| = \|J\tilde{x}\|$, it suffices to set $y_\varepsilon^\lambda := \tilde{x}$ and $q_\varepsilon^\lambda := \lambda^2 J\tilde{x}$, to obtain that $\|q_\varepsilon^\lambda\| = \lambda^2 \|y_\varepsilon^\lambda\| \leq \lambda\sqrt{\varepsilon}$. To complete the proof, we apply the above arguments to

$$M(x,p) := L(x + x_0, p + p_0) - \langle x, p_0 \rangle - \langle x_0, p \rangle - \langle x_0, p_0 \rangle,$$

which is a self-dual Lagrangian on $X \times X^*$. The hypothesis yields that

$$M(0,0) = L(x_0, p_0) - \langle x_0, p_0 \rangle \leq \varepsilon.$$

It then follows from the above that there exists a pair $(y_\varepsilon, q_\varepsilon) \in \overline{\partial} M$ such that $\|y_\varepsilon^\lambda\| \le \frac{\sqrt{\varepsilon}}{\lambda}$, $\|q_\varepsilon^\lambda\| \le \lambda\sqrt{\varepsilon}$. For $x_\varepsilon^\lambda := y_\varepsilon^\lambda + x_0$ and $p_\varepsilon^\lambda := q_\varepsilon^\lambda + p_0$, we have $L(x_\varepsilon^\lambda, p_\varepsilon^\lambda) = \langle x_\varepsilon^\lambda, p_\varepsilon^\lambda \rangle$, and therefore $(x_\varepsilon^\lambda, p_\varepsilon^\lambda) \in \overline{\partial} L$. Note also that $\|x_\varepsilon^\lambda - x_0\| \le \frac{\sqrt{\varepsilon}}{\lambda}$ and $\|p_\varepsilon^\lambda - p_0\|_* \le \lambda\sqrt{\varepsilon}$. □

Remark 3.10 *In view of item 1 in example 3.8, the above result contains the original form of the Brøndsted-Rockafellar Lemma.*

Remark 3.11 *The notion of an ε-approximation of a maximal monotone operator depends on an already given choice of self-dual representation. Given the analogous properties of Fitzpatrick's function, it should be clear that this notion could also be defined by using inequality*

$$F_T(x,p) - \langle x, p \rangle \le \varepsilon.$$

These notions are not necessarily equivalent, but the analogue Brøndsted-Rockafellar property corresponding to this notion holds, the proof (which is similar as the given above) can be found in [BS99], where Burachik & Svaiter introduce the concept this way, under the name of ε-enlargments. *For clarity, this result is stated next.*

Theorem 3.12 *Let $T : X \rightrightarrows X^*$ be a maximal monotone operator and let F_T denote its corresponding Fitzpatrick function; assume that for a pair $(x_0, p_0) \in X \times X^*$ and some $\varepsilon > 0$, we have*

$$F_T(x_0, p_0) - \langle x_0, p_0 \rangle \le \varepsilon.$$

Then, for any $\lambda > 0$, there exists a pair $(x_\varepsilon^\lambda, p_\varepsilon^\lambda) \in T$ such that

1. $\|x_\varepsilon^\lambda - x_0\| \le \frac{\sqrt{\varepsilon}}{\lambda}$,
2. $\|p_\varepsilon^\lambda - p_0\| \le \lambda\sqrt{\varepsilon}$,

3.6 A variational principle

As seen above, a maximal monotone operator T can be written as $T = \bar{\partial} L$ where L is a selfdual Lagrangian on $X \times X^*$, in such a way that solving for x in the equation

$$p \in T(x), \tag{3.5}$$

amounts to minimizing over x the non-negative functional (see (3.1))

$$I(x) := L(x,p) - \langle x, p\rangle,$$

and showing this minimum is actually zero (see proposition 3.3). This functional will play a fundamental role later in this work.

The following result originally established in [Gho07] (see also [Gho08a]), gives sufficient conditions for the infimum of self-dual Lagrangians to be attained, and (as importantly) to be zero.

Theorem 3.13 *Let L be a self-dual functional on a reflexive Banach space $X \times X^*$ such that for some $x_0 \in X$, the functional $q \to L(x_0, q)$ is bounded above on a neighborhood of the origin in X^*. Then there exists $\bar{x} \in X$ such that $I(\bar{x}) = \min_{x\in X} I(x) = 0$.*

Remark 3.14 *An important aspect of the above results is that they give conditions that yield that the* actual value of the infimum *is zero! This allows to use the variational principle for equations for which the existence of solutions is not a priori known. Self-duality plays a role in this part of the results and this was one of the original motivations to choose self-duality as an important property of the convex potentials.*

Remark 3.15 *The boundedness condition on Theorem 3.13 is usually obtained by a coersiveness condition and self-duality: For example, if for some constant C_1 we have $C_1(1 - \|x\|^2 + \|p\|^2) \leq L(x,p)$, then by conjugation we obtain that for some constant C_2 we have $L^*(x,p) \leq C_2(1 + \|x\|^2 + \|p\|^2)$.*

3.7 Coercivity conditions

Coercivity conditions are often imposed when working with variational techniques. Usually they provide the needed compactness. In this section we will show how the choice L^a_T is appropriate in the sense that it preserves coercivity conditions from those of T.

Let us begin by defining growth and boundedness conditions for monotone operators. As an example, an intuitive way to impose super linear growth conditions on T would be by requiring that for some constant $C > 0$

$$C(\|x\|^2 - 1) \leq \langle x, p \rangle, \quad \forall (x, p) \in T. \tag{3.6}$$

On the other hand, if one wishes to impose sublinear growth conditions, a direct way would be to ask, for some $\tilde{C} > 0$

$$\|p\| \leq \tilde{C}(\|x\| + 1), \quad \forall (x, p) \in T. \tag{3.7}$$

Observe that the above condition can be imposed in an alternative way: If one requires, for some $D > 0$, that for any $(x, p) \in T$, one has $D(\|p\|^2 - 1) \leq \langle x, p \rangle$, then, since $\langle x, p \rangle \leq \|x\|\|p\|$, $\|p\|$ must be bounded as in (3.7).

Hence conditions (3.6) and (3.7) can be summarized in the following condition:

$$(x, p) \in T \Rightarrow \max\{C(\|x\|^2 - 1), D(\|p\|^2 - 1)\} \leq \langle x, p \rangle.$$

As a slight generalization of the above, convenient for the setting of L^p spaces, the standard form of growth conditions for monotone operators will be given as follows: for $a > 1$, $\frac{1}{a} + \frac{1}{b} = 1$, there exists constants $C, D > 0$ such that

$$(x, p) \in T \Rightarrow \max\{C(\frac{\|x\|^a}{a} - 1), D(\frac{\|p\|^b}{b} - 1)\} \leq \langle x, p \rangle. \tag{3.8}$$

Proposition 3.16 *A given maximal monotone operator T satisfies growth conditions* (3.8) *if and only if L_T^a satisfies, for some constants $M, N > 0$,*

$$M(\|x\|^a + \|p\|^b - 1) \leq L_T^a(x, p) \leq N(\|x\|^a + \|p\|^b + 1). \tag{3.9}$$

Proof: First let us assume that L_T^a obeys (3.9). For any $(x, p) \in T$ we have that $L_T^a(x, p) = \langle x, p \rangle$, which from the left inequality in (3.9) gives that

$$M(\|x\|^a + \|p\|^b - 1) \leq \langle x, p \rangle,$$

which for some constants C and D gives (3.8).

Conversely, assume that T satisfies (3.8). Consider F_T, its Fitzpatrick's function:

$$\begin{aligned}
F_T(x,p) &= \sup_{(y,q)\in T} \{\langle x,q\rangle + \langle y,p\rangle - \langle y,q\rangle\} \\
&\leq \sup_{(y,q)\in T} \{\langle x,q\rangle + \langle y,p\rangle - \frac{C}{2a}(\|y\|^a - 1) - \frac{D}{2b}(\|q\|^b - 1)\} \\
&\leq \left(\frac{C}{2a}(\|y\|^a - 1) + \frac{D}{2b}(\|q\|^b - 1)\right)^*(p,x) \\
&= \frac{1}{a}\left(\frac{D}{2}\right)^{1-a}\|x\|^a + \frac{1}{b}\left(\frac{C}{2}\right)^{1-b}\|p\|^b + \frac{C}{2a} + \frac{D}{2b}.
\end{aligned}$$

Consider now $(0,p_0)\in T$ (some such p_0 must exist from the boundedness conditions and maximality). From the definition of L_T^a, taking $(x_2,p_2) = (0,p_0)$ and $(x_1,p_1) = (2x, 2p - p_0)$,

$$L_T^a(x,p) \leq \frac{1}{2}F_T(2x,2p-p_0) + \frac{1}{4a}\|2x\|^a + \frac{1}{4b}\|2p\|^b,$$

which using the bound obtained for F_T gives

$$L_T^a(x,p) \leq \frac{1}{a}\left(\frac{D}{2}\right)^{1-a}\|2x\|^a + \frac{1}{b}\left(\frac{C}{2}\right)^{1-b}\|2p-2p_0\|^b + \frac{C}{2a} + \frac{D}{2b} + \frac{1}{4a}\|2x\|^a + \frac{1}{4b}\|2p\|^b,$$

which gives, for some constant N :

$$L_T^a(x,p) \leq N(\|x\|^a + \|p\|^b + 1).$$

Taking the conjugate in the above inequality, since L_T^a is self-dual, yields the converse inequality. □

The above connection between "standard growth conditions" does not necessarily follow for any self-dual potential, and in particular, Fitzpatrick's functions fail to give a full connection of coercivity conditions as illustrated in the following example.

Example 3.17 *Choose $X = \mathbb{R} = X^*$ and $Tx = x$. It is easy to compute*

$$F_T(x,p) = \frac{(x+p)^2}{4}.$$

The above is not coercive in $\mathbb{R}^2$.
On the other hand

$$L_T(x,p) = \frac{x^2}{2} + \frac{p^2}{2}.$$

3.8 Space dependent maximal monotone operators

Time dependent monotone operators and self-dual Lagrangians and their "lifting" to appropriate Banach spaces of time dependent functions (path spaces) are treated in [Gho08a, Gho08b]. The case of space dependent monotone operators and space dependent Lagrangians appears in more detail in [GMZ10], from where this section is taken.

The class $M_{\Omega,p}(\mathbb{R}^N)$

We introduce a particular class of space dependent operators on a given domain with specific growth conditions. This class appears in various works (see for example [PDD90]) and is a standard choice. The key inequality below reproduces (3.8).

Definition 3.18 *For a domain* Ω *in* $\mathbb{R}^N$, $p > 1$ *and* $\frac{1}{p} + \frac{1}{q} = 1$, *we denote by* $M_{\Omega,p}(\mathbb{R}^N)$ *the class of all possibly multi-valued functions* $T : \Omega \times \mathbb{R}^N \to \mathbb{R}^N$ *with closed values, which satisfy the following conditions:*
(i) T *is measurable with respect to* $\mathscr{L}(\Omega) \times \mathscr{B}(\mathbb{R}^N)$ *and* $\mathscr{B}(\mathbb{R}^N)$ *where* $\mathscr{L}(\Omega)$ *is is the* σ*-field of all measurable subsets of* Ω *and* $\mathscr{B}(\mathbb{R}^N)$ *is the* σ*-field of all Borel subsets of* $\mathbb{R}^N$.
(ii) For a.e. $x \in \Omega$, *the map* $T(x,.) : \mathbb{R}^N \to \mathbb{R}^N$ *is maximal monotone.*
(iii) There exist non-negative constants m_1, m_2, c_1 *and* c_2 *such that for every* $\xi \in \mathbb{R}^N$ *and* $\eta \in T(\xi)$,

$$\langle \xi, \eta \rangle_{\mathbb{R}^N} \geq \max\left\{ \frac{c_1}{p}|\xi|^p - m_1, \frac{c_2}{q}|\eta|^q - m_2 \right\}, \tag{3.10}$$

holds, where $\langle .,. \rangle_{\mathbb{R}^N}$ *is the inner product in* $\mathbb{R}^N$.

In this section, we establish a correspondence between maximal monotone maps in $M_{\Omega,p}(\mathbb{R}^N)$ and a class of Ω-dependent self-dual Lagrangians.

Self-dual Lagrangians associated to $M_{\Omega,p}(\mathbb{R}^N)$

Definition 3.19 Let Ω be a domain in $\mathbb{R}^N$.

(i) A function $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is said to be an Ω-*dependent Lagrangian* on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$, if it is measurable with respect to the σ-field generated by the products of Lebesgue sets in Ω and Borel sets in $\mathbb{R}^N \times \mathbb{R}^N$.

(ii) Such a Lagrangian L is said to be *self-dual on* $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$ if for any $x \in \Omega$, the map $L_x : (a,b) \to L(x,a,b)$ is a self-dual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$, i.e., if $L^*(x,b,a) = L(x,a,b)$ for all $a,b \in \mathbb{R}^N$ where

$$L^*(x,b,a) = \sup\{\langle b,\xi\rangle_{\mathbb{R}^N} + \langle a,\eta\rangle_{\mathbb{R}^N} - L(x,\xi,\eta) : (\xi,\eta) \in \mathbb{R}^N \times \mathbb{R}^N\}.$$

The following was proved in [Gho08b] for a single maximal monotone operator.

Proposition 3.20 *If $T \in M_{\Omega,p}(\mathbb{R}^N)$ for some $p > 1$, then there exists an Ω-dependent self-dual Lagrangian $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ such that $T(x,.) = \bar{\partial} L(x,.)$ for a.e. $x \in \Omega$ and*

$$C_0(|a|^p + |b|^q - n_0(x)) \le L(x,a,b) \le C_1(|a|^p + |b|^q + n_1(x)) \qquad \text{for all } a,b \in \mathbb{R}^N. \tag{3.11}$$

where C_0 and C_1 are two positive constants and $n_0, n_1 \in L^1(\Omega)$.
Conversely, if $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is an Ω-dependent self-dual Lagrangian satisfying (3.11), then $\bar{\partial} L(x,.) \in M_{\Omega,p}(\mathbb{R}^N)$.

The proof of the above contains arguments similar to the ones used throughout this chapter, adapted to the special case of space dependent operators and Lagrangians.

Proof. Let $F : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be the Fitzpatrick function [Fit88] associated to T, i.e.,

$$F(x,a,b) := \sup\{\langle b,\xi\rangle_{\mathbb{R}^N} + \langle a-\xi,\eta\rangle_{\mathbb{R}^N}; \eta \in T(x,\xi)\}.$$

Note that measurability assumptions on T ensure that F is a normal integrand. Also, by the properties of the Fitzpatrick function [Gho08a], it follows that

$$F^*(x,b,a) \ge F(x,a,b) \ge \langle a,b\rangle_{\mathbb{R}^N} \text{ for a.e. } x \in \Omega \text{ and for all } a,b \in \mathbb{R}^N.$$

Moreover,

$$\eta \in T(x,\xi) \text{ if and only if } F^*(x,\eta,\xi) = F(x,\xi,\eta) = \langle \eta,\xi\rangle_{\mathbb{R}^N} \text{ a.e. } x \in \Omega. \tag{3.12}$$

Define $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ by

$$\begin{aligned} L(x,a,b) &= \inf\Big\{\frac{1}{2}F(x,a_1,b_1)+\frac{1}{2}F^*(x,b_2,a_2)+\frac{1}{4p}|a_1-a_2|^p+\frac{1}{4q}|b_1-b_2|^q;\\ &\qquad (a,b)=\frac{1}{2}(a_1,b_1)+\frac{1}{2}(a_2,b_2)\Big\}. \end{aligned}$$

We shall show that L is an Ω-dependent self-dual Lagrangian such that

$$F^*(x,b,a) \geq L(x,a,b) \geq F(x,a,b) \text{ for a.e. } x\in\Omega \text{ and for all } a,b\in\mathbb{R}^N. \tag{3.13}$$

Fix $a,b \in \mathbb{R}^N$. We have

$$\begin{aligned} L^*(x,b,a) &= \sup_{\xi,\eta\in\mathbb{R}^N}\{\langle \xi,b\rangle_{\mathbb{R}^N}+\langle a,\eta\rangle_{\mathbb{R}^N}-L(x,\xi,\eta)\}\\ &= \sup_{\xi,\eta\in\mathbb{R}^N}\Big\{\langle \xi,b\rangle_{\mathbb{R}^N}+\langle a,\eta\rangle_{\mathbb{R}^N}-\frac{1}{2}F(x,\xi_1,\eta_1)-\frac{1}{2}F^*(x,\xi_2,\eta_2)+\\ &\qquad -\frac{1}{4p}|\xi_1-\xi_2|^p-\frac{1}{4q}|\eta_1-\eta_2|^q\,;\,(\xi,\eta)=\frac{1}{2}(\xi_1,\eta_1)+\frac{1}{2}(\xi_2,\eta_2)\Big\}\\ &= \frac{1}{2}\sup_{\xi_1,\xi_2,\eta_1,\eta_2\in\mathbb{R}^N}\Big\{\langle \xi_1+\xi_2,b\rangle_{\mathbb{R}^N}+\langle a,\eta_1+\eta_2\rangle_{\mathbb{R}^N}-F(x,\xi_1,\eta_1)+\\ &\qquad -F^*(x,\xi_2,\eta_2)-\frac{1}{2p}|\xi_1-\xi_2|^p-\frac{1}{2q}|\eta_1-\eta_2|^q\Big\}. \end{aligned}$$

Using the fact that the Fenchel dual of the sum is the inf-convolution, we obtain

$$\begin{aligned} L^*(x,b,a) &= \frac{1}{2}\inf_{a_1,b_1\in\mathbb{R}^N}\Big\{F^*(x,b_1,a_1)+F(x,2a-a_1,2b-b_1)+\\ &\qquad +\frac{2^{q-1}}{q}|b-b_1|^q+\frac{2^{p-1}}{2p}|a-a_1|^p\Big\}. \end{aligned}$$

Setting $a_2 = 2a - a_1$ and $b_2 = 2b - b_1$ we have $a = \frac{a_1+a_2}{2}$ and $b = \frac{b_1+b_2}{2}$. It then follows that

$$\begin{aligned} L^*(x,b,a) &= \frac{1}{2} \inf_{a_1,b_1,a_2,b_2 \in \mathbb{R}^N} \Big\{ F^*(x,b_1,a_1) + F(x,a_2,b_2) + \frac{2^{q-1}}{q} |\frac{b_1-b_2}{2}|^q + \\ &\quad + \frac{2^{p-1}}{2p} |\frac{a_1-a_2}{2}|^p \,;\, (a,b) = \frac{1}{2}(a_1,b_1) + \frac{1}{2}(a_2,b_2) \Big\} \\ &= \inf \Big\{ \frac{1}{2} F^*(x,b_1,a_1) + \frac{1}{2} F(x,a_2,b_2) + \frac{1}{4q} |b_1 - b_2|^q + \\ &\quad + \frac{1}{4p} |a_1 - a_2|^p \,;\, (a,b) = \frac{1}{2}(a_1,b_1) + \frac{1}{2}(a_2,b_2) \Big\} \\ &= L(x,a,b). \end{aligned}$$

Thus, L is an Ω-dependent self-dual Lagrangian. Inequalities (3.13) simply follow from the definition and self-duality of L. We shall now prove that L satisfies the estimate (3.11). Note first that for all $\eta \in T(x,\xi)$ we have

$$\frac{1}{p}|\xi|^p + \frac{1}{q}|\eta|^p \leq m_1 + m_2 + (c_1 + c_2)\langle \xi, \eta \rangle_{\mathbb{R}^N}.$$

It follows from the definition of the Fitzpatrick function F that

$$\begin{aligned} F(x,a,b) &= \sup\{\langle b,\xi\rangle_{\mathbb{R}^N} + \langle a-\xi,\eta\rangle_{\mathbb{R}^N}; \eta \in T(x,\xi)\} \\ &\leq \sup\Big\{ \langle b,\xi\rangle_{\mathbb{R}^N} + \langle a,\eta\rangle_{\mathbb{R}^N} - \frac{1}{p(c_1+c_2)}|\xi|^p - \frac{1}{q(c_1+c_2)}|\eta|^q + \\ &\quad - \frac{m_1+m_2}{c_1+c_2}; \eta \in T(x,\xi) \Big\} \\ &\leq \sup_{\xi,\eta \in \mathbb{R}^N} \Big\{ \langle b,\xi\rangle_{\mathbb{R}^N} + \langle a,\eta\rangle_{\mathbb{R}^N} - \frac{1}{p(c_1+c_2)}|\xi|^p + \\ &\quad - \frac{1}{q(c_1+c_2)}|\eta|^q - \frac{m_1+m_2}{c_1+c_2} \Big\} \\ &= \frac{(c_1+c_2)^{p-1}}{p}|a|^p + \frac{(c_1+c_2)^{q-1}}{q}|b|^q + \frac{m_1+m_2}{c_1+c_2}. \end{aligned} \tag{3.14}$$

Let $\eta_0(x) \in T(x,0)$. By assumption $|\eta_0(x)|^q \leq m_2 + \langle 0, \eta_0(x) \rangle = m_2$ for a.e. $x \in \Omega$, from which we get $\eta_0 \in L^q(\Omega)$. It also follows from (3.12) that $F^*(x, \eta_0(x), 0) = 0$

for a.e. $x \in \Omega$. From the definition of L and (3.14), we get that

$$\begin{aligned} L(x,a,b) &\leq \frac{1}{2}F(x,2a-\eta_0(x),2b)+\frac{1}{2}F^*(x,\eta_0(x),0)+\frac{2^q}{4q}|b|^q+\frac{2^p}{4p}|a-\eta_0(x)|^p \\ &\leq C_1(|a|^p+|b|^q+n_1(x)) \ a.e. \quad x\in\Omega, \end{aligned}$$

where C_1 is a positive constant and $n_1 \in L^1(\Omega)$. The reverse inequality follows from the selfduality of L.

Conversely, let L be a Ω-dependent self-dual Lagrangian satisfying (3.11). If $\eta \in \bar{\partial}L(x,\xi)$ then

$$\langle \xi,\eta\rangle = L(x,\xi,\eta) \geq C_0(|\xi|^p+|\eta|^q-n_0(x)),$$

from which we conclude that $\bar{\partial}L(x,.) \in M_{\Omega,p}(\mathbb{R}^N)$. □

3.9 Self-dual Lagrangians on $W_0^{1,p}(\Omega)\times W^{-1,q}(\Omega)$

We now show how one can "lift" an Ω-dependent self-dual Lagrangian to a self-dual Lagrangian on the phase space $W_0^{1,p}(\Omega)\times W^{-1,q}(\Omega)$. This will allow us to give a variational formulation and resolution –via Theorem 3.13– of equations involving maximal monotone operators in divergence form. The following extends a result in [Gho08b].

Theorem 3.21 *Let $T \in M_{\Omega,p}(\mathbb{R}^N)$ for some $p>1$, then for every $w \in W^{-1,q}(\Omega)$ with $\frac{1}{p}+\frac{1}{q}=1$, there exist $\bar{u}\in W_0^{1,p}(\Omega)$ and $\bar{f}(x)\in L^q(\Omega;\mathbb{R}^N)$ such that*

$$\begin{cases} \bar{f}\in T(x,\nabla\bar{u}(x)) \quad a.e.\ x\in\Omega \\ -\mathrm{div}(\bar{f})=w. \end{cases} \tag{3.15}$$

It is obtained by minimizing the functional

$$I(u) := \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ -\mathrm{div}(f)=w}} \int_\Omega \big[L\big(x,\nabla u(x),f(x)\big)-\langle u(x),p(x)\rangle_{\mathbb{R}^N}\big]\,dx$$

on $W^{1,p}(\Omega)$, where L is an Ω-dependent self-dual Lagrangian on $\Omega\times\mathbb{R}^N\times\mathbb{R}^N$

associated to T in such a way that $\bar{\partial} L(x,\cdot) = T(x,\cdot)$ for a.e $x \in \Omega$.

The above theorem will follow from the representation of a maximal monotone map in $M_{\Omega,p}(\mathbb{R}^N)$ by an Ω-dependent self-dual Lagrangian on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$ (Proposition 3.20) combined with the following two propositions.

Proposition 3.22 *Suppose L is an Ω-dependent self-dual Lagrangian on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$ such that $L(\cdot,0,0) \in L^1(\Omega)$. Then the Lagrangian defined on $W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$ by*

$$F(u,p) := \inf\{\int_\Omega L\big(x, \nabla u(x), f(x)\big)\, dx; f \in L^q(\Omega;\mathbb{R}^N), -\mathrm{div}(f) = w\}, \quad (3.16)$$

is selfdual.

Proof: Denote $W_0^{1,p}(\Omega)$ by X and its dual $W^{-1,q}(\Omega)$ by X^*. For a fixed $(v^*,v) \in X^* \times X$, we have

$$\begin{aligned}
F^*(v^*,v) &= \sup\{\langle u, v^*\rangle + \langle p, v\rangle - F(u,p); u \in X, p \in X^*\}\\
&= \sup_{\substack{f \in L^q(\Omega;\mathbb{R}^N)\\ -\mathrm{div}(f)=w\\ (u,p)\in X\times X^*}} \left\{ \langle u, v^*\rangle + \langle p, v\rangle - \int_\Omega L\big(x, \nabla u(x), f(x)\big)\, \mathrm{d}x \right\}\\
&= \sup\{\langle u, v^*\rangle - \langle \mathrm{div}(f), v\rangle - \int_\Omega L\big(x, \nabla u(x), f(x)\big)\, \mathrm{d}x; u \in X, f \in L^q(\Omega;\mathbb{R}^N)\}\\
&= \sup\{\langle u, v^*\rangle + \langle f, \nabla v\rangle - \int_\Omega L\big(x, \nabla u(x), f(x)\big)\, \mathrm{d}x; u \in X, f \in L^q(\Omega;\mathbb{R}^N)\}.
\end{aligned}$$

Now set $E := \{g \in L^p(\Omega;\mathbb{R}^N); g = \nabla u,\, u \in X\}$ and let χ_E be the indicator function in $L^p(\Omega;\mathbb{R}^N)$, e.g.,

$$\chi_E(g) = \begin{cases} 0 & g \in E, \\ +\infty & \text{elsewhere.} \end{cases}$$

An easy computation shows that

$$\chi_E^*(f) = \begin{cases} 0 & \mathrm{div}(f) = 0, \\ +\infty & \text{elsewhere.} \end{cases}$$

Fix $f_0 \in L^q(\Omega;\mathbb{R}^N)$ with $-\mathrm{div}(f_0) = v^*$. It follows that

$$\begin{aligned} F^*(v^*,v) &= \sup\{\langle g, f_0\rangle + \langle f, \nabla v\rangle - \int_\Omega L\big(x, g(x), f(x)\big)\,\mathrm{d}x - \chi_E(g)\,; \\ &\qquad g \in L^p(\Omega;\mathbb{R}^N), f \in L^q(\Omega;\mathbb{R}^N)\} \\ &= \inf\{\int_\Omega L^*(x, f_0 - f, \nabla v)\,\mathrm{d}x + \chi_E^*(f); f \in L^q(\Omega;\mathbb{R}^N)\}. \end{aligned}$$

Here we have used the fact that $\big(\int_\Omega L\big(x,.,.\big)\,\mathrm{d}x\big)^*(g,f) = \int_\Omega L^*\big(x, f(x), g(x)\big)\,\mathrm{d}x$ that holds since $L(.,0,0) \in L^1(\Omega)$. We finally get

$$\begin{aligned} F^*(v^*,v) &= \inf\{\int_\Omega L^*(x, f_0 - f, \nabla v)\,\mathrm{d}x; f \in L^q(\Omega;\mathbb{R}^N), \mathrm{div}(f) = 0\} \\ &= \inf\{\int_\Omega L(x, \nabla v, f_0 - f)\,\mathrm{d}x; f \in L^q(\Omega;\mathbb{R}^N), \mathrm{div}(f) = 0\} \\ &= \inf\{\int_\Omega L(x, \nabla v, f)\,\mathrm{d}x; f \in L^q(\Omega;\mathbb{R}^N), -\mathrm{div}(f) = v^*\} \\ &= F(v, v^*). \end{aligned}$$

□

Here is our variational resolution for equation (3.15).

Proposition 3.23 *Suppose L is an Ω-dependent self-dual Lagrangian on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$. Assume the following coercivity condition:*

$$L(x,a,b) \geq m(x) + C(|a|^p + |b|^q) \text{ for all } a, b \in \mathbb{R}^N, \tag{3.17}$$

where $m \in L^1(\Omega)$ and C is a positive constant. Then for every $w \in W^{-1,q}(\Omega)$ the functional

$$I(u) = \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ -\mathrm{div}(f) = w}} \int_\Omega \big[L\big(x, \nabla u(x), f(x)\big) - \langle u(x), w(x)\rangle_{\mathbb{R}^N}\big]\,dx$$

attains its minimum at some $\bar{u} \in W_0^{1,p}(\Omega)$ such that $I(\bar{u}) = 0$, and there exists $\bar{f} \in L^q(\Omega;\mathbb{R}^N)$ such that

$$\begin{cases} \bar{f}(x) \in \bar{\partial} L(x, \nabla \bar{u}(x)) & a.e.\ x \in \Omega \\ -\mathrm{div}(\bar{f}) = w. \end{cases}$$

Proof. Take $f_0 \in L^q(\Omega;\mathbb{R}^N)$ with $-\mathrm{div}\big(f_0(x)\big) = w(x)$. Since L is an Ω-dependent self-dual Lagrangian, $M(x,a,b) := L(x,a,b+f_0(x)) - \langle a, f_0(x)\rangle$ is also an Ω-dependent self-dual Lagrangian on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$. It follows from the above proposition that

$$F(v,v^*) := \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ -\mathrm{div}(f)=v^*}} \int_\Omega M\big(x,\nabla v(x), f(x)\big)\,dx$$

is a self-dual Lagrangian on $W_0^{1,p}(\Omega)\times W^{-1,q}(\Omega)$. In view of the coercivity condition, Theorem 3.13 applies and there exists $\bar{u} \in W_0^{1,p}(\Omega)$ such that

$$F(\bar{u},0) = \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ -\mathrm{div}(f)=0}} \int_\Omega M\big(x,\nabla \bar{u}(x), f(x)\big)\,dx = 0.$$

Using the coercivity condition again, we get that the above infimum is attained at some $f_1 \in L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div}(f_1) = 0$. Thus,

$$\begin{aligned}
0 = F(\bar{u},0) &= \int_\Omega M\big(x,\nabla\bar{u}(x), f_1(x)\big)\,dx \\
&= \int_\Omega \big[L(x,\nabla\bar{u}(x), f_1(x)+f_0(x)) - \langle \nabla\bar{u}(x), f_0(x)\rangle_{\mathbb{R}^N}\big]\,dx \\
&= \int_\Omega \big[L(x,\nabla\bar{u}(x), f_1(x)+f_0(x)) - \langle \nabla\bar{u}(x), f_1(x)+f_0(x)\rangle_{\mathbb{R}^N}\big]\,dx.
\end{aligned}$$

Taking into consideration that $L(x,\nabla\bar{u}(x), f_1(x)+f_0(x)) - \langle \nabla\bar{u}(x), f_1(x)+f_0(x)\rangle_{\mathbb{R}^N} \geq 0$, we obtain that the latter is indeed zero, i.e.,

$$L(x,\nabla\bar{u}(x), f_1(x)+f_0(x)) - \langle \nabla\bar{u}(x), f_1(x)+f_0(x)\rangle_{\mathbb{R}^N} = 0 \qquad \text{for a.e. } x\in\Omega.$$

Setting $\bar{f} := f_1 + f_0$, we finally get that $\bar{f}(x) \in \bar{\partial} L(x,\nabla\bar{u}(x))$ for a.e. $x \in \Omega$ and that $-\mathrm{div}(\bar{f}) = w$. □

Chapter 4

Topologies on spaces of convex Lagrangians and corresponding vector fields

In this chapter we introduce and briefly study suitable topologies for the spaces of convex functions and of monotone operators. These topologies are given by the notions of Mosco convergence and Γ-convergence (*Gamma convergence*), for spaces of convex functions, and *Kuratowski-Painlevé* convergence for the space of maximal monotone operators. It will be shown that these convergence notions are related.

Γ and Mosco convergence are now popular tools in the study of asymptotic properties of families of functionals. Γ-convergence, introduced by De Giorgi, is probably best known for its role in classic homogenization theory; on the other hand Mosco convergence is known to be well behaved under the Fenchel-Legendre transform. These notions are based on the more fundamental definition of set convergence known as *Kuratowski-Painlevé*. This mode of convergence has proven ideal for the study of convergence properties of families of maximal monotone operators as shown by the work of Kato, Brezis, Attouch and Browder, among others.

4.1 Mosco and Γ-convergence

Mosco and Γ convergence are convergence notions for functions. These modes are not restricted to convex functions, but they are much better behaved in the context of proper, convex, lower semicontinuous functions, which is the natural setting we are working on.

These modes of convergence are neither stronger nor weaker than pointwise convergence, they have a more geometrical nature and are shown to be ideal notions to use in variational settings.

A good reference on the general subject is the book [Att84].

Definition 4.1 *A sequence of functions* $\{f_n\} \subset \Gamma(X)$ *is said to* Mosco converge *to a function* f*, denoted* $f_n \xrightarrow{M} f$ *if both of the following conditions are satisfied:*

1. *For each weakly convergent sequence,* $x_n \rightharpoonup x$*, we have that*

$$f(x) \leq \liminf_n f_n(x_n).$$

2. *For any* $x \in X$*, there exists a strongly convergent sequence* $x_n \to x$ *such that*

$$f(x) = \lim_n f_n(x_n).$$

The sequence $\{f_n\}$ *is said to* Γ-converge*, denoted by* $f_n \xrightarrow{\Gamma} f$*, if the above conditions hold, replacing the weak convergence in **1.** by strong convergence.*

The above will also be denoted as $f = M - \lim f_n$ and $f = \Gamma - \lim f_n$ for Mosco and Γ convergence respectively.

Conditions **1.** and **2.** are enough to guarantee uniqueness of the limits.

Definition 4.2 *A sequence satisfying condition **2.** above will be referred to as a* recovery sequence *for* $\{f_n\}$.

In terms of the epigraphs, this mode of convergence is that the set of *limit points* for the epigraphs $\{epi(f_n)\}$ is equivalent to its set of *cluster points* and this limit corresponds to $epi(f)$, in both the strong and weak topologies (the strong topology).

A fundamental property of Mosco convergence is the following:

Lemma 4.3 *$f_n \xrightarrow{M} f$ if and only if $f_n^* \xrightarrow{M} f^*$.*

The above is not the case for Γ convergence (see the defined sequence in example 4.8).

The following lemma (which appears in [AB93]) provides alternative definitions of Mosco convergence:

Lemma 4.4 *For $f_n, f \in \Gamma(X)$, the following are equivalent:*

1. *$f_n \xrightarrow{M} f$.*
2. *The following two conditions hold:*
   - *For any $x \in X$, there exists a strongly convergent sequence, $x_n \to x$, such that*
   $$f(x) = \lim_n f_n(x_n).$$
   - *For any $p \in X^*$, there exists a strongly convergent sequence, $p_n \to p$, such that*
   $$f^*(p) = \lim_n f_n^*(p_n).$$
3. *The following two conditions hold:*
   - *For any $x \in dom(\partial f)$, there exist a strongly convergent sequence, $x_n \to x$, such that*
   $$f(x) \geq \limsup_n f_n(x_n).$$
   - *For any $p \in ran(\partial f)$, there exists a strongly convergent sequence, $p_n \to p$, such that*
   $$f^*(p) \geq \limsup_n f_n^*(p_n).$$

Γ and Mosco convergence, as mentioned before, can be seen as *Kuratowski-Painlevé* convergence for epigraphs (in both the strong and weak topology in the case of Mosco convergence):

4.2 Kuratowski - Painlevé topology

For a sequence of sets $\{A_n\}$, we use the norm topology to define

Definition 4.5 *1. The* limit inferior *set, LiA_n, given by all limit points of the sequence $\{A_n\}$. In other words:*

$$x \in LiA_n \iff \lim_n \|x - a_n\| = 0 \textit{ for some sequence } \{a_n\} \textit{ with } a_n \in A_n \forall n.$$

2. The limit superior *set, LsA_n, given by all cluster points of the sequence $\{A_n\}$. In other words:*

$$x \in LsA_n \iff \liminf_n \|x - a_n\| = 0 \textit{ for some sequence } \{a_n\} \textit{ with } a_n \in A_n \forall n.$$

3. We say that the sequence converges to the set A in the sense of Kuratowski - Painlevé if

$$A = LiA_n = LsA_n.$$

This mode of convergence will be denoted by

$$A_n \xrightarrow{K-P} A.$$

This mode of convergence on the space $X \times X^*$ is ideal for a notion of convergence for maximal monotone operators, restricted to the case of convex subdifferentials it is known to match neatly with Mosco convergence, as it will be shown in the coming sections. We will also extend this relation, to some extent, to general maximal monotone operators and their corresponding convex potentials.

G convergence, introduced by Dal Maso to study certain types of homogenization problems, is based on Kuratowski-Painlevé convergence (see for example [PDD90, Definition 3.5]).

4.3 Mosco convergence vs Γ-convergence for self-dual functions

It is clear that Mosco convergence is a stronger notion of convergence than Γ-convergence, however, with the added condition of self-duality, these modes of convergence are actually equivalent. It is worth of mention that Γ-convergence in the general setting (without any restriction to self-dual functionals) combined with an equicoercivity condition is itself equivalent to Mosco convergence (that is: a family of equicoercive convex functions which is Γ-convergent will also be Mosco convergent).

Theorem 4.6 *Let $\{L_n\}$ be a family of self-dual Lagrangians on $X \times X^*$, and let L be a function on $X \times X^*$. The following statements are then equivalent:*

1. *$\{L_n\}$ is Mosco convergent to L.*
2. *L is self-dual and $\{L_n\}$ Γ-converges to L.*
3. *L is self-dual and for any $(x,p) \in X \times X^*$ there exists a sequence (x_n, p_n) converging strongly to (x,p) in $X \times X^*$ such that*

$$\limsup_n L_n(x_n, p_n) \leq L(x,p).$$

Proof: For $(1) \to (2)$ we just need to prove that L is self-dual since Mosco convergence clearly implies Γ-convergence. Since L is the Mosco limit of L_n, it follows from Lemma 4.3 that L^* is a Mosco limit of L_n^*. Denoting

$$L_n^T(p,x) := L_n(x,p) \text{ and } L^T(p,x) := L(x,p),$$

it follows that L^T is a Mosco-limit of L_n^T on $X^* \times X$. On the other hand, by self-duality of L_n we have that $L_n^T = L_n^*$ from which we obtain that $L^T = M - \lim_n L_n^T = M - \lim_n L_n^* = L^*$, and therefore $L^T = L^*$, and L is therefore self-dual.
(2)$\to$(3) follows from the definition of Γ-convergence.
For (3)$\to$(1) we let $(p,x) \in X^* \times X$ and consider a sequence $\{(p_n,x_n)\} \subset X^* \times X$ such that $(p_n,x_n) \rightharpoonup (p,x)$ weakly in $X^* \times X$. By the definition of Fenchel-Legendre

duality we have

$$\liminf_n L_n^*(p_n,x_n) = \liminf_n \sup_{(y,q)\in X\times X^*} \{\langle x_n,q\rangle + \langle y,p_n\rangle - L_n(y,q)\}. \tag{4.1}$$

Consider now an arbitrary pair $(\tilde{x},\tilde{p})$ and let $\{(\tilde{x}_n,\tilde{p}_n)\}$ be the recovery sequence given by item (3.). It follows from (4.1) that

$$\liminf_n L_n^*(p_n,x_n) \geq \liminf_n \Big(\langle x_n,\tilde{p}_n\rangle + \langle \tilde{x}_n,p_n\rangle - L_n(\tilde{x}_n,\tilde{p}_n)\Big) \geq \langle x,\tilde{p}\rangle + \langle \tilde{x},p\rangle - L(\tilde{x},\tilde{p}).$$

Since $(\tilde{x},\tilde{p})$ is arbitrary, taking the supremum over all $(\tilde{x},\tilde{p})$ yields

$$\liminf_n L_n^*(p_n,x_n) \geq L^*(p,x).$$

Since both L_n and L are self-dual, this implies that

$$\liminf_n L_n(x_n,p_n) \geq L(x,p),$$

and therefore that L is a Mosco-limit of L_n. □

Remark 4.7 *Note that while the Mosco convergence of self-dual Lagrangians automatically implies that the limiting Lagrangian L is itself self-dual, this fails for Γ-convergence as shown in the following example.*

The following example is an adaptation of an counter example found in [DMS08].

Example 4.8 *Let H be an infinite dimensional Hilbert space. Consider a set $\{e_n\}$ with $\|e_n\| = 1$ and $e_n \rightharpoonup 0$ (For example, the orthonormal basis of the space). Define*

$$L_n(x,p) := \frac{1}{2}\|x-e_n\|^2 + \frac{1}{2}\|p\|^2 + \langle p,e_n\rangle.$$

Notice that L_n is self-dual. It can be checked directly that for any strongly convergent sequence $(x_n,p_n) \to (x,p)$ in $H\times H$ we have $\lim_n L_n(x_n,p_n) = L(x,p)$, where

$$L(x,p) := \frac{1}{2}\|x\|^2 + \frac{1}{2}\|p\|^2 + \frac{1}{2}.$$

This means that L is a Γ-limit of L_n. On the other hand, it is easily seen that

$L^*(p,x) = L(x,p) - 1$, so L is not self-dual and therefore we do not have Mosco convergence.

4.4 Kuratowski-Painlevé convergence for maximal monotone operators

In a spirit similar to Theorem 4.6, where the definition for Mosco convergence can be simplified in the self-dual case, the following lemma is a standard result and provides a shorter definition of Kuratowski-Painlevé convergence for maximal monotone operators:

Lemma 4.9 *Let $\{T_n\}$ be a sequence of maximal monotone operators and T a maximal monotone operator. Then the following are equivalent:*

1. $T \subset LiT_n$.
2. $T_n \xrightarrow{K-P} T$.

Proof: Assume that $T \subset LiT_n$, we need to show that $Ls(T_n) \subset T$. Consider $(y,q) \in Ls(T_n)$, thus, there exist some sequence $(y_n, q_n) \in T_n$ such that for some subsequence $(y_{n(k)}, q_{n(k)}) \to (y,q)$. Now take an arbitrary $(x,p) \in T$. From our assumption that $T \subset LiT_n$, there exist a sequence $(x_n, p_n) \in T_n$ such that $(x_n, p_n) \to (x,p)$. For each k we have

$$\langle x_{n(k)} - y_{n(k)}, p_{n(k)} - q_{n(k)} \rangle \geq 0,$$

and as $k \to \infty$ we get

$$\langle x - y, p - q \rangle \geq 0.$$

The above is for any $(x,p) \in T$, so by maximality, we must have that $(y,q) \in T$, and thus $Ls(T_n) \subset T$. We conclude $T_n \xrightarrow{K-P} T$.

The converse implication follows simply by definition. □

4.5 Continuity of the regularization of self-dual Lagrangians

The following well known result (see [Att84]) relates convergence of a sequence of convex functions with convergence of their Moreau-Yosida regularizations:

Proposition 4.10 *Let $\{f_n\} \subset \Gamma(X)$ and $f \in \Gamma(X)$. The following are equivalent*

1. $f_n \xrightarrow{M} f$.
2. *For some* $\lambda_0 > 0$, $(f_n)_{\lambda_0} \xrightarrow{M} (f)_{\lambda_0}$.
3. *For any* $\lambda > 0$, $(f_n)_\lambda \xrightarrow{M} (f)_\lambda$.

In a similar spirit, it can be shown that some of the regularizations given in definition 2.21 also preserve Mosco convergence.

Proposition 4.11 *Let $\{L_n\}$ and L be a sequence and an element of $\mathscr{L}(X)$. The following are equivalent*

1. $L_n \xrightarrow{M} L$.
2. *For some* $\lambda_0 > 0$, $(L_n)_1^{\lambda_0} \xrightarrow{M} (L)_1^{\lambda_0}$.
3. *For any* $\lambda > 0$, $(L_n)_1^{\lambda} \xrightarrow{M} (L)_1^{\lambda}$.
4. *For some* $\lambda_0 > 0$, $(L_n)_2^{\lambda_0} \xrightarrow{M} (L)_2^{\lambda_0}$.
5. *For any* $\lambda > 0$, $(L_n)_2^{\lambda} \xrightarrow{M} (L)_2^{\lambda}$.

The proof of the above will be split into various smaller results. Recalling definition 2.21, let us introduce the following:

Definition 4.12 *For any $\lambda > 0$ and any $L \in \mathscr{L}(X)$, let*

$$R_L^\lambda(x,p) := \operatorname*{arg\,min}_{y\in X} \left\{ L(y,p) + \frac{\|x-y\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2 \right\}$$

and

$$T_L^\lambda(x,p) := \operatorname*{arg\,min}_{q\in X^*} \left\{ L(x,q) + \frac{\|x\|^2}{2\lambda} + \frac{\lambda}{2}\|p-q\|^2 \right\}.$$

Also, the following important boundedness property of Mosco convergence will be useful (a proof can be found in [DMS08]).

Lemma 4.13 *If $\{L_n\}$ is a Mosco converging sequence of proper convex functionals on $X \times X^*$, then there exist positive constants a and b such that*

$$L_n(x,p) + a(\|x\| + \|p\|) + b \geq 0$$

for all $(x,p) \in X \times X^*$.

We will begin by showing the following.

Proposition 4.14 *If $L \in \mathscr{L}(X)$ is the Mosco-limit for some sequence $\{L_n\}$ of self-dual Lagrangians:*

1. *For every $\lambda > 0$*

$$L_1^\lambda = M - \lim \left(L_n\right)_1^\lambda \text{ and } L_2^\lambda = M - \lim \left(L_n\right)_2^\lambda.$$

2a. *For any converging sequence on X, $x_n \to x$, and any $p \in X^*$, there exist a converging sequence on X^*, $p_n \to p$ such that:*

$$\left(L_n\right)_1^\lambda (x_n, p_n) \to L_1^\lambda (x,p).$$

2b. *For any converging sequence on X^*, $p_n \to p$, and any $x \in X$, there exist a converging sequence on X, $x_n \to x$ such that:*

$$\left(L_n\right)_2^\lambda (x_n, p_n) \to L_2^\lambda (x,p).$$

3. *For any converging sequence $(x_n, p_n) \to (x,p)$:*

$$R_{L_n}^\lambda (x_n, p_n) \to R_L^\lambda (x,p) \text{ and } T_{L_n}^\lambda (x_n, p_n) \to T_L^\lambda (x,p).$$

Proof: Consider any $(x,p) \in X \times X^*$ and any converging sequence on X, $x_n \to x$. Denote $y := R_L^\lambda (x,p)$. Let $(y_n, p_n) \to (y,p)$ be a corresponding recovery sequence for $\{L_n\}$, that is:

$$L_n(y_n, p_n) \to L(y, p).$$

We have

$$\left(L_n\right)_1^{\lambda}(x_n, p_n) \leq L_n(y_n, p_n) + \frac{\|x_n - y_n\|^2}{2\lambda} + \frac{\lambda}{2}\|p_n\|^2,$$

and thus, since all the sequences $\{x_n\}$, $\{y_n\}$ and $\{p_n\}$ are strongly convergent:

$$\limsup_n \left(L_n\right)_1^{\lambda}(x_n, p_n) \leq L(y, p) + \frac{\|x - y\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2.$$

Since $y = R_L^{\lambda}(x, p)$, the above is:

$$\limsup_n \left(L_n\right)_1^{\lambda}(x_n, p_n) \leq L_1^{\lambda}(x, p).$$

In view of the third characterization for Mosco convergence given in Theorem 4.6 and the self-duality of $\left(L_n\right)_1^{\lambda}$ and L_1^{λ} (proposition 3.4), this inequality establishes the first equation in item **1**:

$$L_1^{\lambda} = M(X \times X^*) - \lim \left(L_n\right)_1^{\lambda}.$$

This in turn grants, by the $\liminf$ property of Mosco convergence, that for the sequence shown above:

$$\left(L_n\right)_1^{\lambda}(x_n, p_n) \to L_1^{\lambda}(x, p).$$

This is conclusion **2a**.

To complete this proof, we turn to show that for any converging sequence $(x_n, p_n) \to (x, p)$ such that $\left(L_n\right)_1^{\lambda}(x_n, p_n) \to L_1^{\lambda}(x, p)$, we have

$$R_{L_n}^{\lambda}(x_n, p_n) \to R_L^{\lambda}(x, p).$$

Considering such (x_n, p_n), define now $y_n := R^\lambda_{L_n}(x_n, p_n)$. We have

$$\big(L_n\big)^\lambda_1(x_n, p_n) = L_n(y_n, p_n) + \frac{\|x_n - y_n\|^2}{2\lambda} + \frac{\lambda}{2}\|p_n\|^2. \tag{4.2}$$

Using the uniform boundedness from below for Mosco convergence (see lemma 4.13), since we have that $\{L_n\}$ is a Mosco converging sequence, we have, for some positive constants a and b:

$$\big(L_n\big)^\lambda_1(x_n, p_n) \geq \frac{\|x_n - y_n\|^2}{2\lambda} + \frac{\lambda}{2}\|p_n\|^2 - a(\|y_n\| + \|p_n\|) - b.$$

Since the left hand side of the above inequality is a converging sequence, and so is $\{(x_n, p_n)\}$, then the sequence $\{y_n\}$ must be bounded. Let us then consider a weakly convergent subsequence which we will not relabel for simplicity: $y_n \rightharpoonup \tilde{y}$. Take the lower limit on (4.2) to get

$$L^\lambda_1(x, p) \geq \liminf_n L_n(y_n, p_n) + \liminf_n \frac{\|x_n - y_n\|^2}{2\lambda} + \liminf_n \frac{\lambda}{2}\|p_n\|^2,$$

which from Mosco convergence of $\{L_n\}$ and lower semicontinuity of the norm yields

$$L^\lambda_1(x, p) \geq L(\tilde{y}, p) + \frac{\|x - \tilde{y}\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2;$$

The above must then, be an equality and this gives: $\tilde{y} = R^\lambda_L(x, p)$.

Furthermore, we can assume, without loss of generality, that the subsequence was taken in such a way that $L_n(y_n, p_n) \to \liminf_n L_n(y_n, p_n)$, which gives that also

$$\frac{\|x - y_n\|^2}{2\lambda} \to \frac{\|x - \tilde{y}\|^2}{2\lambda},$$

and hence $\|y_n\| \to \|\tilde{y}\|$, so that subsequence actually converges strongly to $R^\lambda_L(x, p)$.

This argument can be done for any subsequence of the original sequence $\{y_n\}$, and hence we have shown the desired result, namely,

$$R^\lambda_{L_n}(x_n, p_n) \to R^\lambda_L(x, p).$$

This is the first part of conclusion **3**. The proofs for the related results for $\left(L_n\right)_2^\lambda$, L_2^λ and T_L^λ are directly analogous to the above. □

Proposition 4.15 *If for some* $\lambda_0 > 0$, *we have one of* $L_1^{\lambda_0} = M - \lim \left(L_n\right)_1^{\lambda_0}$ *or* $L_2^{\lambda_0} = M - \lim \left(L_n\right)_2^{\lambda_0}$, *then*

$$L = M - \lim L_n.$$

Proof: Consider $(x,p) \in X \times X^*$. Assume that $(x,p) \in dom(\partial L)$. Take $q \in \partial_1 L(x,p)$ ($y \in \partial_2 L(x,p)$). Define, for some $\varepsilon > 0$, $\lambda = \lambda_0 + \varepsilon$ ($\lambda = \lambda_0 \star \varepsilon$). Define

$$y := x - \lambda J^{-1} q \ \big(q := p - \lambda J y\big).$$

It can be checked that $x = R_L^\lambda(y,p)$ ($p = T_L^\lambda(x,q)$). From proposition 2.23, we can apply proposition 4.13 to get that there exists a sequence $(y_n, p_n) \to (y,p)$ such that $\left(L_n\right)_1^\lambda(y_n,p_n) \to \left(L\right)_1^\lambda(y,p)$ (a sequence $(x_n,q_n) \to (x,q)$ such that $\left(L_n\right)_2^\lambda(x_n,q_n) \to \left(L\right)_2^\lambda(x,q)$) and $R_{L_n}^\lambda(y_n,p_n) \to R_L^\lambda(y,p)$ ($T_{L_n}^\lambda(x_n,q_n) \to T_L^\lambda(x,q)$). Define

$$x_n := R_{L_n}^\lambda(y_n,p_n) \ \big(p_n := T_{L_n}^\lambda(x_n,q_n)\big).$$

Simply write now

$$\left(L_n\right)_1^\lambda(y_n,p_n) = L_n(x_n,p_n) + \frac{\|x_n - y_n\|^2}{2\lambda} + \frac{\lambda}{2}\|p_n\|^2,$$

$$\left(\left(L_n\right)_2^\lambda(x_n,q_n) = L_n(y_n,q_n) + \frac{\|x_n\|^2}{2\lambda} + \frac{\lambda}{2}\|p_n - q_n\|^2\right).$$

Then

$$\lim_n L_n(x_n,p_n) = \left(L\right)_1^\lambda(y,p) - \frac{\|x-y\|^2}{2\lambda} - \frac{\lambda}{2}\|p\|^2,$$

$$\left(\lim_n L_n(y_n,q_n) = \left(L\right)_2^\lambda(x,q) - \frac{\|x\|^2}{2\lambda} - \frac{\lambda}{2}\|p-q\|^2\right),$$

this is

$$L_n(x_n,p_n) \to L(x,p)$$

and, since L_n and L are self-dual, this is enough to guarantee

$$L = M - \lim L_n.$$

□

Corollary 4.16 *Given $L_n, L \in \mathscr{L}(X)$ and $\lambda > 0$,*

$$L = M - \lim L_n \iff \left(L\right)_1^\lambda = M - \lim \left(L_n\right)_1^\lambda \iff \left(L\right)_2^\lambda = M - \lim \left(L_n\right)_2^\lambda.$$

4.6 Continuity of regularizations of maximal monotone operators

Results analogous to Proposition 4.11 are already known for maximal monotone operators. The following can be found in [Att84].

Proposition 4.17 *Let $\{T_n\}$ be a sequence of maximal monotone operators and let T be a maximal monotone operator. The following are equivalent:*

1. $T_n \xrightarrow{K-P} T$.
2. *For some* $\lambda_0 > 0$, $(T_n)_{\lambda_0} \xrightarrow{K-P} T_{\lambda_0}$.
3. *For each* $\lambda > 0$, $(T_n)_\lambda \xrightarrow{K-P} T_\lambda$.
4. *For some* $\lambda_0 > 0$, $R_{\lambda_0}^{T_n} \xrightarrow{K-P} R_{\lambda_0}^T$.
5. *For each* $\lambda > 0$, $R_\lambda^{T_n} \xrightarrow{K-P} R_\lambda^T$.

4.7 Attouch's theorem: on the bicontinuity of the application $f \to \partial f$

For each $f \in \Gamma(X)$, define

$$\Delta(f) := \{(x, f(x), p) \mid (x, p) \in \partial f\}.$$

Recall *Attouch's Theorem*, it appears in [Att84] and [AB93]:

Theorem 4.18 *Let X be a reflexive Banach space and $f, f_1, f_2, \ldots$ proper convex functions on X. The following are equivalent:*

1. $f_n \xrightarrow{M} f$.

2. $\partial f_n \xrightarrow{K-P} \partial f$ *and* $\Delta(f) \cap Li\Delta(f_n) \neq \emptyset$.

Essentially, this result establishes the equivalence of Mosco convergence of proper convex functions on X (equivalently, their Fenchel-Legendre conjugates on X^*) and Kuratowski - Painlevé convergence on $X \times X^*$ of their subdifferentials. The condition on $\Delta(f)$ is just a normalizing condition, since a convex functional is uniquely defined by its subdifferential mapping only up to an additive constant.

Remark 4.19 *As a trivial example of the role of the condition on* $\Delta(f)$*, consider* f *a fixed convex function in* $\Gamma(X)$ *and the sequence given by*

$$f_n(x) := f(x) + n.$$

Clearly f_n *does not Mosco converge to* f*, but* ∂f_n *remains constant (and thus converges in the sense of Kuratowski-Painlevé). Observe we have*

$$f_n^*(p) = f^*(p) - n,$$

and hence

$$f_n + f_n^* = f + f^*.$$

Observe that the problem lies in the additive constant that preserves the subdifferential.

This very simple idea will be used later, when we present a "self-dual" formulation of Attouch's Theorem on phase space.

Proofs of Attouch's Theorem can be found on [Att84] and [AB93]. However, we present here the proof for convenience of the reader and since its arguments are of interest to us.

Remark 4.20 *Observe that the proof relies on two key ingredients: The Brønsted - Rockafellar Lemma, and Rockafellar's integration formula.*

Proof of Theorem 4.18: Assume first that $f_n \xrightarrow{M} f$. Fix $(x,p) \in \partial f$, there exists, by the presented characterizations of Mosco convergence on lemma 4.4, a sequence $x_n \to x$ and $p_n \to p$, each converging strongly in X and X^* respectively, such that $f_n(x_n) \to f(x)$ and $f_n^*(p_n) \to f^*(p)$.

We have then $f(x) + f^*(p) = \langle x, p \rangle = \lim_n \langle x_n, p_n \rangle$, and hence, if we define $\varepsilon_n := f_n(x_n) + f_n^*(p_n) - \langle x_n, p_n \rangle$, we obtain that $\lim_n \varepsilon_n = 0$ and that p_n is an ε_n-subdifferential of f_n at x_n.

Hence, by Theorem 2.2, we have the existence of a pair $(\tilde{x}_n, \tilde{p}_n) \in \partial f_n$ such that $\|x_n - \tilde{x}_n\| < \sqrt{\varepsilon_n}$ and $\|p_n - \tilde{p}_n\| < \sqrt{\varepsilon_n}$. Clearly, since $\varepsilon_n \to 0$: $\tilde{x}_n \to x$ and $\tilde{p}_n \to p$.

This shows that $\partial f \subset Li(\partial f_n)$.

Now, since we have $f_n \xrightarrow{M} f$, this gives that

$$f(x) \leq \liminf_n f_n(\tilde{x}_n) \tag{4.3}$$

and

$$f^*(p) \leq \liminf_n f_n^*(\tilde{p}_n),$$

and since

$$\lim_n \Big(f_n(\tilde{x}_n) + f_n^*(\tilde{p}_n) \Big) = \lim_n \Big(\langle \tilde{x}_n, \tilde{p}_n \rangle \Big) = \langle x, p \rangle = f(x) + f^*(p),$$

we can write now

$$\limsup_n f_n(\tilde{x}_n) \leq \limsup_n \Big(f_n(\tilde{x}_n) + f_n^*(\tilde{p}_n) \Big) + \limsup_n \Big(- f_n^*(\tilde{p}_n) \Big) =$$

$$= f(x) + f^*(p) - \liminf_n f_n^*(\tilde{p}_n) \leq f(x).$$

In conjunction with (4.3), this shows $\lim f_n(\tilde{x}_n) = f(x)$.

We have shown $(\tilde{x}_n, f_n(\tilde{x}_n), \tilde{p}_n) \to (x, f(x), p)$, where $(\tilde{x}_n, \tilde{p}_n) \in \partial f_n$. This shows that $\Delta(f) \cap Li\Delta(f_n) \neq \emptyset$.

To finalize this part of the proof, we just need to show that $Ls(\partial f_n) \subset \partial f$. But this follows from the already established fact $\partial f \subset Li(\partial f_n)$, the maximal monotonicity of convex subdifferentials, and Lemma 4.9.

Now we establish the converse implication. Assume that $\partial f_n \xrightarrow{K-P} \partial f$ and that

$$\Delta(f) \cap Li\Delta(f_n) \neq \emptyset.$$

Fix some $(y,q) \in \partial f$. From the assumption, we have a sequence $(y_n, q_n) \in \partial f_n$ converging to (y,q).

Pick now (x,p) such that $(x, f(x), p) \in \Delta(f) \cap Li\Delta(f_n)$. Let this condition be realized by the sequence $(x_n, p_n) \in \partial f_n$.

Choose an arbitrary chain from y to x, that is, some finite set $\{x^i\}_{i=1}^k$ with $x^1 = y$ and $x^k = x$. Pick corresponding $p^i \in \partial f(x^i)$, with $p^1 = q$ and $p^k = p$.

Since $\partial f_n \xrightarrow{K-P} \partial f$, we can consider, also, sequences $\{(x_n^i, p_n^i)\}_n$ with $(x_n^i, p_n^i) \in \partial f_n$ and $(x_n^i, p_n^i) \to (x^i, p^i)$, where we require $(x_n^k, p_n^k) = (x_n, p_n)$, with (x_n, p_n) as given above.

Write now

$$f_n(x_n) - f_n(y_n) = \sum_{i=1}^{k-1} \left(f_n(x_n^{i+1}) - f_n(x_n^i) \right).$$

From the subgradient inequality,

$$f_n(x_n) - f_n(y_n) \geq \sum_{i=1}^{k-1} \langle x_n^{i+1} - x_n^i, p_n^i \rangle.$$

Now, since we chose x_n in such a way that $\lim_n f_n(x_n) = f(x)$, we apply $\limsup_n$ to the above to get

$$f(x) \geq \limsup_n f_n(y_n) + \sum_{i=1}^{k-1} \langle x^{i+1} - x^i, p^i \rangle,$$

and, taking the sup over all chains, Rockafellar's formula (see Theorem 2.14) yields

$$f(x) \geq \limsup_n f_n(y_n) + f(x) - f(y),$$

which finally gives us the first condition in item 3. of Lemma 4.4:

$$f(y) \geq \limsup_n f_n(y_n).$$

In an analogous way, using $\Delta(f^*) = \{(p, f^*(p), x) \,:\, (x, p) \in \partial f\}$, we can show the second condition in item 3. of Lemma 4.4, namely,

$$f^*(p) \geq \limsup_n f_n^*(p_n).$$

This completes the proof. $\square$

4.7.1 Attouch's theorem on phase space

We propose the following reformulation of Attouch's Theorem as a first step towards a generalization of the result for maximal monotone operators and self-dual Lagrangians. Note that the need of a normalizing condition is gone and the result in this form directly relates Mosco convergence (in phase space) to Kuratowski-Painlevé convergence:

Theorem 4.21 *Let X be a reflexive Banach space and $f, f_1, f_2, \ldots$ functions in $\Gamma(X)$. The following are equivalent:*

1. $f_n + f_n^* \xrightarrow{M} f + f^*$.
2. $\partial f_n \xrightarrow{K-P} \partial f$.

Proof: The proof that the first condition implies the second condition is identical to the proof used for Theorem 4.18.

Assume, then, that $\partial f_n \xrightarrow{K-P} \partial f$. Fix any $(x_0, p_0) \in \partial f$ and any sequence $(x_n, p_n) \to (x_0, p_0)$ with $(x_n, p_n) \in \partial f_n$.

Now define $g_n(x) := f_n(x) - f_n(x_n)$ and $g(x) := f(x) - f(x_0)$.

Observe that $\partial g_n = \partial f_n$ and $\partial g = \partial f$, which, given the current assumption, implies

$$\partial g_n \xrightarrow{K-P} \partial g.$$

By construction, we also have $g_n(x_n) = g(x_0) = 0$, which yields $(x_n, g_n(x_n), p_n) \to (x_0, g(x_0), p_0)$, and thus

$$\Delta(g) \cap Li\Delta(g_n) \neq \emptyset$$

so we can use Theorem 4.18 to obtain right away that, on X, $g_n \xrightarrow{M} g$ which in turn yields that also, on X^*, $g_n^* \xrightarrow{M} g^*$.

Clearly this implies, in $X \times X^*$,

$$g_n + g_n^* \xrightarrow{M} g + g^*,$$

and since $g_n^*(p) = f_n^*(p) + f_n(x_n)$ and $g^*(p) = f^*(p) + f(x_0)$, we have $g_n + g_n^* = f_n + f_n^*$ and $g + g^* = f + f^*$, that is:

$$f_n + f_n^* \xrightarrow{M} f + f^*$$

which completes the proof. □

4.8 $L \mapsto \overline{\partial} L$ and $F_T \mapsto T$ are continuous

Our first result shows the continuity of the application $L \mapsto \overline{\partial} L$, as well as its analogue for Fitzpatrick's functions.

Theorem 4.22 *Let X be a reflexive Banach space and suppose $\{L_n\}$ is a family of selfdual Lagrangians on $X \times X^*$. If $L : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ is a selfdual Lagrangian that is a Mosco limit of $\{L_n\}$, then the graph of $\bar{\partial} L_n$ converge to the graphs of $\bar{\partial} L$ in the sense of Kuratowski-Painlevé.*

The following is the corresponding analogue for Fitzpatrick's functions. Its proof is directly analogous to the one for the previous result, using the Brønsted-Rockafellar property of the Fitzpatrick function (see Remark 3.11), and will not be specifically written.

Theorem 4.23 *Let X be a reflexive Banach space and suppose that T and the elements of a family $\{T_n\}$ are maximal monotone operators on X. If F_T, the Fitzpatrick function for T, is the Mosco-limit of $\{F_{T_n}\}$, where F_{T_n} is the Fitzpatrick function for $\{T_n\}$, then $\{T_n\}$ converges to T in the sense of Kuratowski-Painlevé.*

Proof of Theorem 4.22. Fix $(x, p) \in \overline{\partial} L$. Then in view of the Mosco convergence, there exists a sequence (x_n, p_n) converging strongly to (x, p) in $X \times X^*$ such that $L_n(x_n, p_n) \to L(x, p)$. We then have $L(x, p) = \langle x, p \rangle = \lim_n \langle x_n, p_n \rangle$, and therefore if we define $\varepsilon_n := L_n(x_n, p_n) - \langle x_n, p_n \rangle$, we obtain that $\lim_n \varepsilon_n = 0$. Hence, by Lemma 3.9, we have the existence of a pair $(\tilde{x}_n, \tilde{x}_n^*) \in \overline{\partial} L_n$ such that $\|x_n - \tilde{x}_n\| < \sqrt{\varepsilon_n}$

and $\|p_n - \tilde{x}_n^*\|_* < \sqrt{\varepsilon_n}$. Clearly $\tilde{x}_n \to x$ and $\tilde{x}_n^* \to p$ as $\varepsilon_n \to 0$. This shows that $\overline{\partial} L \subset Li(\overline{\partial} L_n)$, and the proof is complete, in view of Lemma 4.9. □

4.9 On the continuity of $\overline{\partial} L \mapsto L$

Before attempting a converse result, namely, stating the continuity of the application $\overline{\partial} L \mapsto L$, it is important to be aware that, as mentioned before, self-dual Lagrangians are not unique, and hence we cannot expect any perfect analogy to Theorem 4.21.

The first obvious ingredient we are missing is an integration formula for a given maximal monotone operator T. A direct substitute for it is the selection of L_T^a, which is uniquely determined. Indeed the following result can be shown:

Theorem 4.24 *Let $\{T_n\}$ be a sequence of maximal monotone operators, and let T be a maximal monotone operator. If the sequence $\{F_{T_n}\}$ Mosco converges to F_T, then, for each fixed $a > 1$, we have that $L_{T_n}^a \xrightarrow{M} L_T^a$.*

The proof follows immediately from the general fact below:

Proposition 4.25 *Let $\{F_n\}$ be a sequence of functions in $\Gamma(X \times X^*)$ Mosco convergent to F. For any constants A, B and $a, b > 1$, we have that the sequence of functions $\{G_n\}$, defined by*

$$G_n(x,p) = \inf_{(y,q)} \{F_n(y,q) + F_n^*(2p-q, 2x-y) + A\|x-y\|^a + B\|p-q\|^b\},$$

Mosco converges to the function G, given by

$$G(x,p) = \inf_{(y,q)} \{F(y,q) + F^*(2p-q, 2x-y) + A\|x-y\|^a + B\|p-q\|^b\}.$$

Proof: Fix (x,p). There exists some (y,q) such that

$$G(x,p) = F(y,q) + F^*(2p-q, 2x-y) + A\|x-y\|^a + B\|p-q\|^b.$$

Observe that since $\{F_n\}$ is Mosco convergent, so is $\{F_n^*\}$. Define $z := 2x - y$ and $r := 2p - q$. There exist sequences (y_n, q_n) and (z_n, r_n), strongly convergent to (y,q)

and (z,r) respectively, and such that

$$\lim_n F_n(y_n,q_n) = F(y,q)$$

and

$$\lim_n F_n^*(r_n,z_n) = F^*(r,z).$$

Define $x_n := \frac{y_n+z_n}{2}$ and $p_n := \frac{q_n+r_n}{2}$. It can be checked that (x_n,p_n) is convergent to (x,p). Also

$$\begin{aligned} G_n(x_n,p_n) &\leq F_n(y_n,q_n)+F_n^*(2p_n-q_n,2x_n-y_n)+A\|x_n-y_n\|^p+B\|p_n-q_n\|^q \\ &= F_n(y_n,q_n)+F_n^*(r_n,z_n)+A\|x_n-y_n\|^p+B\|p_n-q_n\|^q. \end{aligned}$$

Taking the upper limit yields

$$\limsup_n G_n(x_n,p_n) \leq F(y,q)+F^*(r,z)+A\|x-y\|^p+B\|p-q\|^q = G(x,p).$$

The above can be repeated for G^*, yielding the result. □

However, whether or not the convergence of $\{T_n\}$ to T implies the convergence of their respective Fitzpatrick's functions, remains without proof, to this point.

4.10 A compact class in L^p

In the following u will denote a function in $L^p(\Omega)$, $1<p<\infty$, where Ω is a domain on $\mathbb{R}^n$. For q, the conjugate exponent of p (i.e. such that $\frac{1}{p}+\frac{1}{q}=1$), v will denote a function in $L^q(\Omega)$.

We will consider a class of functionals of the form

$$\mathscr{F} = \{F(u) = \int_\Omega f\big(u(x)\big)\,\mathrm{d}x \,:\, f\in\tilde{\mathscr{F}}\},$$

where the class $\tilde{\mathscr{F}}\subset\Gamma(\mathbb{R})$ is specified by constants β_i,γ_i, $i=1,2$:

$$\tilde{\mathscr{F}} = \{f:\mathbb{R}\to\mathbb{R} \,:\, \beta_1|x|^p-\gamma_1 < f(x) < \beta_2|x|^p+\gamma_2 \text{ and } f\in\Gamma(\mathbb{R})\}.$$

Proposition 4.26 *The class $\mathscr{F}$ is Γ compact in $L^p(\Omega)$.*

Proof: Consider a sequence $\{F_n\}$ in $\mathscr{F}$; by taking $F_n(u) = \int_\Omega f_n(u(x))\,\mathrm{d}x$, this sequence corresponds in a natural way to a sequence $\{f_n\}$ in $\tilde{\mathscr{F}}$. Since the class $\tilde{\mathscr{F}}$ consists entirely of convex functions, we can show they are also equicontinuous on compact subsets of $\mathbb{R}$: consider $v(x) \in \partial f(x)$, we have for any $h \in \mathbb{R}$

$$\langle v(x), h\rangle \leq f(x+h) - f(x),$$

taking the sup over $\{|h| = 1\}$ and considering the definition of the class, we get that there must be a constant C such that

$$|v(x)| \leq C(|x|^p + 1).$$

This gives that the subdifferentials have uniformly bounded norms on compact subsets of $\mathbb{R}$, which implies equicontinuity on compact subsets of $\mathbb{R}$. By the *Arzela-Ascoli lemma*, for any compact $K \subset \mathbb{R}$, there exists a subsequence that converges uniformly on K to some $f_K \in \tilde{\mathscr{F}}$. By covering $\mathbb{R}$ with an increasing compact cover, we can now construct an appropriate subsequence (denoted by the same symbol) $\{f_n\}$ such that

$$f_n(x) \to f_\infty(x)$$

for each $x \in \mathbb{R}$ and uniformly on compact subsets of $\mathbb{R}$. Clearly $f_\infty(x) \in \tilde{\mathscr{F}}$.

Consider now any subsequence $\{u_n\}$ in $L^p(\Omega)$ converging in norm to u. Focus on an a.e. converging subsequence, again denoted by $\{u_n\}$. By simply applying *Fatou's lemma* we get

$$\int_\Omega f_\infty\Big(u(x)\Big)\,\mathrm{d}x \leq \liminf \int_\Omega f_n\Big(u_n(x)\Big)\,\mathrm{d}x$$

which, taking $F_\infty(u) = \int_\Omega f_\infty\Big(u(x)\Big)\,\mathrm{d}x$, gives

$$F_\infty(u) \leq \liminf F_n(u_n).$$

Consider now any $u \in L^p(\Omega)$. Since $f_n \to f$ pointwise, and again, considering the bounds defining the class, by the *dominated convergence theorem* we get that

$$F_n(u) = \int_\Omega f_n\Big(u(x)\Big)\,\mathrm{d}x \to \int_\Omega f_\infty\Big(u(x)\Big)\,\mathrm{d}x = F_\infty(u).$$

We have shown that $F_n \xrightarrow{\Gamma} F_\infty$. $\square$

We can also define the class

$$\mathscr{F}^* := \{F^* \,:\, F \in \mathscr{F}\}.$$

It is the case that

$$\mathscr{F}^* = \{G(v) = \int_\Omega f^*\big(v(x)\big)\,\mathrm{d}x \,:\, f \in \tilde{\mathscr{F}}\},$$

A direct computation can show that if $f \in \tilde{\mathscr{F}}$, then

$$\Big(\beta_2|x|^p - \gamma_2\Big)^* \geq f^*(x) \geq \Big(\beta_1|x|^p - \gamma_1\Big)^*.$$

Corollary 4.27 *The class $\mathscr{F}^*$ is Γ compact.*

In order to show that the class $\mathscr{F}$ has the liminf property also for weakly convergent sequences, Barbu and Kunisch used the following lemma (see [BK95, BK96]):

Lemma 4.28 *Consider a sequence of functions $\{f_n\} \subset \tilde{\mathscr{F}}$ converging to $f_0 \in \tilde{\mathscr{F}}$ uniformly on compact sets of $\mathbb{R}$. Let $v \in L^p(\Omega)$. Then*

$$(1+\lambda\partial f_n)^{-1}v \to (1+\lambda\partial f_0)^{-1}v$$

strongly on $L^p(\Omega)$. Consequently there is a subsequence $\{f_{n_k}\}$ such that

$$(1+\lambda\partial f_{n_k})^{-1}v(x) \to (1+\lambda\partial f_0)^{-1}v(x)$$

for a.e. $x \in \Omega$.

We can work without this result by using the properties of Mosco convergence instead. (Of course, our proofs used similar machinery as the lemma above).

Corollary 4.29 *The class $\mathscr{F}$ is Mosco compact.*

Proof: From Proposition 4.26 and corollary 4.27, for any sequence $\{F_n\} \subset \mathscr{F}$ there exists a subsequence, denoted by the same symbol, such that for any $u \in L^p(\Omega)$

$$\lim_n F_n(u) = F(u)$$

and for any $v \in L^q(\Omega)$

$$\lim_n F_n^*(v) = F^*(v).$$

Referring to item 2 in lemma 4.4, we have that

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

□

The following simple lemma will be useful in the sequel:

Lemma 4.30 *If $\{F_n\} \subset \mathscr{F}$ is such that $F_n \xrightarrow{M} F$, then:*

1. *For any strongly convergent sequence in X^*, $p_n \to p$,*

$$F_n^*(p_n) \to F^*(p).$$

2. *For any strongly convergent sequence in X, $x_n \to x$,*

$$F_n(x_n) \to F(x).$$

Proof: Consider a sequence $p_n \to p$. Define x_n to be an element such that

$$\langle x_n, p_n \rangle - F_n(x_n) = F_n^*(p_n).$$

Such x_n exists thanks to the boundedness conditions on $\mathscr{F}$, and for the same conditions we must have that the sequence $\{x_n\}$ remains bounded in X. Hence we can extract a weakly convergent subsequence, again denoted by the same symbol, $x_n \rightharpoonup x$. We have for this subsequence

$$\limsup_n F^*(p_n) \leq \langle x, p \rangle - \liminf_n F_n(x_n) \leq \langle x, p \rangle - F(x) \leq F^*(p).$$

Since we also have that $\liminf_n F_n^*(p_n) \geq F^*(p)$, we obtain the desired conclusion, that $F_n^*(p_n) \to F^*(p)$.

Observing the analogue bound conditions for $\mathscr{F}^*$ and that $F_n \xrightarrow{M} F$ if and only if $F_n^* \xrightarrow{M} F^*$, the same argument as above yields $F_n(x_n) \to F(x)$ for any sequence $x_n \to x$. $\square$

Chapter 5

A self-dual approach to inverse problems and optimal control

In this chapter we present a unified approach to inverse problems which involve the identification of non-linearities in a given family of monotone PDE. The approach is based on work by Barbu and Kunisch ([BK96] and [BK95]), distilled into a very general form which is perfectly suited to the existing variational theory developed by Ghoussoub in [Gho08a]. The approach relies on a penalization method, where the penalty term is given by a variational potential naturally corresponding to the equation itself (see section 3.6). This approach is easily generalized to more general optimal control problems.

More specific applications are provided in the next chapter.

5.1 An inverse problem via a variational, penalization based approach

To better illustrate the general type of problem and the spirit of the general results, let us outline briefly a simple inverse problem and the approach taken in [BK96]:

For a convex function Φ, which we will assume is Frechet differentiable, consider the problem

$$\begin{aligned} -\Delta u + D\Phi(u) &= f \quad x \in \Omega. \\ u &= 0 \quad x \in \partial\Omega. \end{aligned} \tag{5.1}$$

This problem can be shown to have a solution using classic variational methods. Simply consider the functional

$$\tilde{J}(u) := \Phi(u) + \frac{1}{2}\int_\Omega |\nabla u|^2 - \int_\Omega fu$$

which can be shown to have a minimizer u_Φ in the space $H_0^1(\Omega)$ and that minimizer corresponds precisely to a solution of the differential equation (5.1). This is the classic direct problem.

Observe here that as an alternative to $\tilde{J}$ we could have considered, in view of Fenchel-Young duality (2.2), the functional

$$\begin{aligned} J(u) &:= \Phi(u) + \Phi^*(\Delta u + f) - \langle u, \Delta u + f \rangle = \\ &= \Phi(u) + \Phi^*(\Delta u + f) + \int_\Omega |\nabla u|^2 - \int_\Omega fu. \end{aligned}$$

At first contact, this might seem like a needlessly complicated alternative, but J has the following desirable properties, which will be fundamental in the following:

1. $J(u) \geq 0$.
2. $J(u) = 0$ if and only if u solves (5.1).

Suppose we are given a class $\mathscr{K}$ of convex functions, and for each $\Phi \in \mathscr{K}$ we have a PDE, given by (5.1).

The inverse problem consists of the following: Given some function u_0, can we recover a function Φ such that u_0 corresponds to the solution of the corresponding differential equation (5.1)?

There is an important restriction we must observe first: that it may not be possible to find such a Φ, depending on the chosen u_0 (maybe u_0 is a noisy measurement of the actual solution, for example). But we can still hope to find the *closest solution* by some least square criterion, that is, a function Φ such that its corresponding solution u_Φ minimizes the functional

$$g(u) := \int_\Omega |u_0 - u|^2 \,\mathrm{d}x. \tag{5.2}$$

This leads to the following minimization problem: find Φ_* such that if u_* is its corresponding solution to (5.1), then

$$\|u_0 - u_*\|^2 = \inf\{\|u_0 - u\|^2 : u \text{ solves (5.1) for some } \Phi \in \mathscr{K}\}. \tag{5.3}$$

The constraint above makes the minimization problem cumbersome to handle, which is where the choice of variational functional comes into play:

Observe first, that the previously defined functional J can be seen as a functional on u and Φ variables:

$$J(u,\Phi) = \Phi(u) + \Phi^*(\Delta u + f) + \int_\Omega |\nabla u|^2 - \int_\Omega fu.$$

This functional is positive and is directly normalized in such a way that its infimum is actually zero and is achieved only at a solution. We will use this functional to penalize the couple (u,Φ) *not being a solution* to equation (5.1).

Define now

$$I_\varepsilon(u,\Phi) := g(u) + \frac{1}{\varepsilon} J(u,\Phi).$$

The functional I_ε can be minimized for each fixed ε to obtain $(u_\varepsilon, \Phi_\varepsilon)$, note that $J(u_\varepsilon, \Phi_\varepsilon)$ is not necessarily zero, which means the pair $(u_\varepsilon, \Phi_\varepsilon)$ may not be a solution to (5.1), however, as ε goes to zero, it is enforced that $J(u_\varepsilon, \Phi_\varepsilon)$ goes to zero, which, if a limit exists, would yield a solution.

This approach will be formalized in the coming results.

5.2 Main result

The general problem and approach will be outlined before stating the result.

Suppose we are given $\mathscr{L}$, a family of self-dual Lagrangians. This family corresponds to a family of monotone equations given by: for some fixed $p \in X^*$, for each $L \in \mathscr{L}$ we have

$$p \in \overline{\partial} L(x). \tag{5.4}$$

Given a fixed $x_0 \in X$, we are interested on identifying a particular L_* in $\mathscr{L}$ such that if x_* satisfies the corresponding equation (5.4), then this solution is the best possible approximation to x_0, in the sense that it satisfies

$$\|x_0 - x_*\|^2 = \inf\{\|x_0 - x\|^2 : p \in \overline{\partial} L(x) \text{ for some } L \in \mathscr{L}\}. \tag{5.5}$$

To overcome the complicated constraint space given above, we will make use of the following functional, given naturally by the self-dual potential (see section 3.6):

$$J(x, L) := L(x, p) - \langle x, p \rangle. \tag{5.6}$$

This functional, in view of (3.1) and definition 3.2, satisfies

1. $J(x, L) \geq 0$.
2. $J(x, L) = 0$ if and only if the pair (x, L) solves (5.4).

The idea is to then obtain approximate solutions by minimizing the functional

$$I_\varepsilon(x, L) := \|x_0 - x\|^2 + \frac{1}{\varepsilon} J(x, L). \tag{5.7}$$

This functional's important qualities are summarized below.

Proposition 5.1 *The functional I_ε is everywhere non negative, and (x_*, L_*) are such that $I_\varepsilon(x_*, L_*) = 0$ if and only if both*

1. $x_* = x_0$

2. $p \in \overline{\partial} L_*(x_0)$.

Proof: I_ε is the sum of $\|x_0 - x\|^2$ and $\frac{1}{\varepsilon}\big(L(x, p) - \langle x, p \rangle\big)$, both positive functionals, and hence is clearly positive. Also, it is zero if and only if *both* of the previous are zero, which yields the conclusion. □

The above means that if the given x_0 is achieved in $\mathscr{L}$, meaning that x_0 solves (5.4) for some $L \in \mathscr{L}$, then the minimum value of I_ε, for any $\varepsilon > 0$, is actually zero and vice versa!

If on the contrary, x_0 is not achievable in the class $\mathscr{L}$ then the hope is to obtain a best possible approximate solution as a limit of the minimizers $(x_\varepsilon, L_\varepsilon)$ by making $\varepsilon \to 0$, which should enforce $J(x_\varepsilon, L_\varepsilon) \to 0$, then if a limit for $(x_\varepsilon, L_\varepsilon)$ exists, it would be to a minimizing solution.

Indeed, the following holds:

Theorem 5.2 *Let $p \in X^*$ be given. Assume that the class $\mathscr{L}$ is* Mosco compact*. Assume also that for some $L \in \mathscr{L}$, problem* (5.4) *has a solution. Fix $x_0 \in X$.*
Then there exists $L_* \in \mathscr{L}$ such that a corresponding solution to (5.4), x_*, is the best possible approximation to x_0 in the sense that it satisfies (5.5).
Furthermore, the pair (x_*, L_*) can be obtained as a limit as $\varepsilon \to 0$ of minimizers of the functional I_ε.

Proof: The functional I_ε is non-negative, thus, we can consider a minimizing sequence $\{(x_n, L_n)\}$ for I_ε. Since J is positive, we have that the term $\|x_0 - x\|^2$ must remain bounded, since the sequence is minimizing. Hence we can extract from $\{x_n\}$ a weakly convergent subsequence, denoted by the same symbol and converging to some $x_\varepsilon \in X$. Also because of the compactness of the class $\mathscr{L}$, there exists some subsequence of $\{L_n\}$ denoted by the same symbol, Mosco converging to some $L_\varepsilon \in \mathscr{L}$.

Then, from the Mosco convergence, $L_\varepsilon(x_\varepsilon, p) \leq \liminf L_n(x_n, p)$. This gives, since the norm is weakly lower semicontinuous and $\langle \cdot, p \rangle$ is weakly continuous, that

$$I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq \liminf I_\varepsilon(x_n, L_n) = \inf_{\mathscr{L} \times X} I_\varepsilon(x, L),$$

hence

$$I_\varepsilon(x_\varepsilon, L_\varepsilon) = \inf_{\mathscr{L} \times X} I_\varepsilon(x, L).$$

Now make $\varepsilon \to 0$. Observe that since for some $L \in \mathscr{L}$ there is some x that satisfies $p \in \overline{\partial} L(x)$, then for any $\varepsilon > 0$:

$$I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq I_\varepsilon(x, L) = \|x_0 - x\|^2,$$

so as $\varepsilon \to 0$, we get, similar as before, that $\|x_\varepsilon\|$ remains bounded. Furthermore, since the bound in the previous expression does not depend on ε we must have that as $\varepsilon \to 0$

$$J_\varepsilon(x_\varepsilon, L_\varepsilon) = L_\varepsilon(x_\varepsilon, p) - \langle x_\varepsilon, p\rangle \to 0.$$

Again, by Mosco compactness of the class $\mathscr{L}$ we get that there exist some $x_* \in X$ and $L_* \in \mathscr{L}$ such that at least for some sequence $\{\varepsilon_n\}$ converging to zero:

$$x_{\varepsilon_n} \rightharpoonup x_*$$

weakly in X and

$$L_*(x_*, p) \leq \liminf L_{\varepsilon_n}(x_{\varepsilon_n}, p).$$

Now, since $\langle\cdot, p\rangle$ is weakly continuous on X:

$$L_*(x_*, p) = \langle x_*, p\rangle$$

which yields both

$$p \in \overline{\partial} L_*(x_*)$$

and for any $(x, L) \in X \times \mathscr{L}$ that solves (5.4):

$$\|x_0 - x_*\|^2 = I_\varepsilon(x_*, L_*) \leq I_\varepsilon(x, L) = \|x_0 - x\|^2,$$

which corresponds to (5.5). The proof is complete. □

Remark 5.3 *By proposition 5.1, if x_0 is achievable in the class, meaning that there is some $L_0 \in \mathscr{L}$ such that the pair (x_0, L_0) solves* (5.4)*, then $\inf_{X\times\mathscr{L}} I_\varepsilon = 0 = I_\varepsilon(x_0, L_0)$. This means that any minimizer in the approximate problem will be the optimal solution, and there is no need to take the limit $\varepsilon \to 0$.*

5.3 Main results for optimal control problems

The approach taken in the previous section can be generalized. We present now analogous results in the context of Optimal Control.

Consider first a fixed, bounded below, lower semicontinuous cost

$$g : X \to \mathbb{R}.$$

The optimal control problem associated to this cost and the family of equations parametrized by $\mathscr{L}$ given by (5.4) is to find an $L_* \in \mathscr{L}$ providing the *cheapest solution* for the cost g. That is, to identify L_* such that the pair (x_*, L_*) satisfies (5.4) and also satisfies

$$g(x_*) = \inf\{g(x) \,:\, p \in \overline{\partial} L(x) \text{ for some } L \in \mathscr{L}\}. \tag{5.8}$$

As before, the complicated constraint set above will be dealt with by instead minimizing the following functional:

$$I_\varepsilon(x, L) := g(x) + \frac{1}{\varepsilon} J(x, L), \tag{5.9}$$

where J is defined again by (5.6).

Theorem 5.4 *Let $p \in X^*$ be given. Assume g is a bounded below, weakly lower semicontinuous and coercive cost in X. Assume that the class $\mathscr{L}$ is* Mosco compact. *Assume also that for some $L \in \mathscr{L}$, problem* (5.4) *has a solution, denoted x_L. Then there exists $L_* \in \mathscr{L}$ such that the corresponding solution to (5.4), x_*, satisfies* (5.8)*. Furthermore, the pair (x_*, L_*) can be obtained as a limit as $\varepsilon \to 0$ of minimizers of* (5.9).

Proof: This proof is very similar to the one for Theorem 5.2. The function $g(x)$ takes the place of $\|x - x_0\|^2$ and has the analogue properties needed in the proof. □

The coercivity condition can be transferred from the cost functional to the class of potentials: The coercivity condition on g can be removed if the class $\mathscr{L}$ is known to be *equicoercive* in the space variable.

Definition 5.5 *A class of functionals $\mathscr{F}$ is said to be equicoercive if for some fixed constants $M, C > 0$ and $a > 1$ and every $F \in \mathscr{F}$, $F : X \to \mathbb{R}$, the following holds:*

$$\|x\| > M \Rightarrow C(\|x\|^a - 1) \leq F(x).$$

The following holds:

Theorem 5.6 *Let $p \in X^*$ be given. Assume g is a bounded below, weakly lower semicontinuous cost in X. Assume that the class $\mathscr{L}$ is* Mosco compact *and equico-*

ercive in the first variable. Assume also that for some $L \in \mathscr{L}$, problem (5.4) *has a solution, denoted x_L.*
Then there exists $L_* \in \mathscr{L}$ such that the corresponding solution to (5.4), x_*, satisfies (5.8)*. Furthermore, the pair (x_*, L_*) can be obtained as a limit as $\varepsilon \to 0$ of minimizers of* (5.9)*.*

Proof: As before, since the functional I_ε is bounded below, we can consider $\{(x_n, L_n)\}$ a minimizing sequence for I_ε. Since $\{g(x_n)\}$ is bounded below, $J(x_n, L_n) = L_n(x_n, p) - \langle x_n, p\rangle$ remains bounded. This implies that $\{x_n\}$ is bounded in X, since $\mathscr{L}$ is equicoercive. There exist, then, a weakly convergent subsequence denoted again by $\{x_n\}$ converging to some $x_\varepsilon \in X$. The rest of the proof remains intact: Since the class $\mathscr{L}$ is assumed Mosco compact, there exists some L_ε such that for some subsequence of $\{L_n\}$, denoted by the same symbol, one has $L_\varepsilon(x_\varepsilon, p) \leq \liminf L_n(x_n, p)$. This, along with the weak lower semicontinuity of g, gives $I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq \liminf I_\varepsilon(x_n, L_n)$ and so

$$I_\varepsilon(x_\varepsilon, L_\varepsilon) = \inf_{X \times \mathscr{L}} I_\varepsilon(x, L).$$

Now let $\varepsilon \to 0$. Since for some $L \in \mathscr{L}$ there is some x satisfying $L(x, p) - \langle x, p\rangle = 0$, we have

$$I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq I_\varepsilon(x, L) = g(x)$$

which gives a uniform bound, implying that $\{x_\varepsilon\}$ is bounded in X as $\varepsilon \to 0$. It also implies that $L_\varepsilon(x_\varepsilon, p) - \langle x_\varepsilon, p\rangle \to 0$.

As before, we can thus assume the existence of $(x_*, L_*) \in X \times \mathscr{L}$ such that, at least for some sequence $\{\varepsilon_n\}$ converging to zero, $x_{\varepsilon_n} \rightharpoonup x_*$ weakly in X and $L_0(x_*, p) \leq \liminf L_\varepsilon(x_\varepsilon, p)$. Thus, we have that $p \in \overline{\partial} L_*(x_*)$, and for any pair $(x, L) \in X \times \mathscr{L}$ satisfying (5.4):

$$g(x_*) \leq \liminf I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq I_\varepsilon(x, L) = g(x),$$

which completes the proof. □

5.4 Further generalizations and comments on coercivity and compactness conditions

Also, instead of fixing $p \in X^*$, we can also associate a cost functional to it:

$$h : X^* \to \mathbb{R}.$$

The problem consists now on finding a triplet (x_*, p_*, L_*) solving (5.4) such that

$$g(x_*) + h(p_*) = \inf\{g(x) + h(p) \; : \; p \in \overline{\partial} L(x) \text{ for some } L \in \mathscr{L}\}. \tag{5.10}$$

However, there is an issue that must be observed; it may be apparent that asking g and h to be coercive along with the compactness of $\mathscr{L}$ is enough to warranty existence of such an optimal solution, but if the proof is attempted one gets sequences $\{x_n\}$ and $\{p_n\}$ which are only weakly convergent. The problem now is that the limit behaviour of $\langle x_n, p_n \rangle$ cannot be controlled.

What is happening, in terms of the Mosco topology, is a loss of compactness in the following way: Allowing p to vary accounts for actually reparametrizing the self-dual potential. Consider the following class

$$\mathscr{M} := \{M \; : \; M(x,q) = L(x, q+p) - \langle x, p \rangle \text{ for } p \in X^* \text{ and some } L \in \mathscr{L}\},$$

this class contains only self-dual Lagrangians, and solving

$$p \in \overline{\partial} L(x) \text{ for some } L \in \mathscr{L},$$

is equivalent to solving

$$0 \in \overline{\partial} M(x) \text{ for some } M \in \mathscr{M}.$$

This way the cost on X^*, $h(p)$ can be seen as a cost on $\mathscr{M}$, $H(M)$. However, the class $\mathscr{M}$ *is not Mosco compact*! (The reason being precisely that $\langle \cdot, \cdot \rangle$ cannot be controlled for jointly weakly convergent pairs).

To get around this, two approaches will be outlined here: The first is to enforce compactness of the class $\mathscr{M}$ by restricting the variable p to some compact class.

The second is to ask for stronger compactness conditions on X, via a compactly embedded subspace and keeping a coercivity condition on h.

The first alternative is contained in the following:

Theorem 5.7 *Assume g is a bounded below, weakly lower semicontinuous and coercive cost on X. Let $P \subset X^*$ be a compact set. Let h be a bounded below, lower semicontinuous cost on P. Assume that the class $\mathscr{L}$ is* Mosco compact. *Assume also that for some $L \in \mathscr{L}$ and some $p \in P$, problem* (5.4) *has a solution $x \in X$. Then there exist $(x_*, p_*, L_*) \in X \times P \times \mathscr{L}$ such that they solve (5.4), and they satisfy* (5.10). *Furthermore, the triple (x_*, p_*, L_*) can be obtained as a limit as $\varepsilon \to 0$ of minimizers of*

$$I_\varepsilon(x,L) := g(x) + h(p) + \frac{1}{\varepsilon}\big(L(x,p) - \langle x,p\rangle\big). \tag{5.11}$$

Proof: The functional I_ε is bounded below, since g and h were assumed bounded below and $L(x,p) - \langle x,p\rangle$ is a positive functional. Thus we can consider $\{(x_n, p_n, L_n)\}$ a minimizing sequence for I_ε. Again, since $L(x,p) - \langle x,p\rangle \geq 0$ and h is bounded below, we must have that $\{g(x_n)\}$ is bounded, and by the coercivity of g, this implies that $\{x_n\}$ is bounded in X. There exist, then, a weakly convergent subsequence, denoted again by $\{x_n\}$, converging to some $x_\varepsilon \in X$. Since P is compact, we can extract some further subsequence, denoted by the same symbol, such that $p_n \to p_\varepsilon$ strongly in X^* for some $p_\varepsilon \in X^*$. Since the class $\mathscr{L}$ is assumed Mosco compact, there exists some L_ε such that for some subsequence of $\{L_n\}$, denoted by the same symbol, one has $L_\varepsilon(x_\varepsilon, p_\varepsilon) \leq \liminf L_n(x_n, p_n)$. This, along with the weak lower semicontinuity of g and h, gives $I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq \liminf I_\varepsilon(x_n, L_n)$ (here we are using that $\langle x_\varepsilon, p_\varepsilon\rangle = \lim\langle x_n, p_n\rangle$, which is where the strong convergence of $\{p_n\}$ is required) and so

$$I_\varepsilon(x_\varepsilon, p_\varepsilon, L_\varepsilon) = \inf_{X \times X^* \times \mathscr{L}} I_\varepsilon(x,p,L).$$

Now let $\varepsilon \to 0$. Since for some $L \in \mathscr{L}$ there is some x and some $p \in P$ satisfying $L(x,p) - \langle x,p\rangle = 0$, we have

$$I_\varepsilon(x_\varepsilon, p_\varepsilon, L_\varepsilon) \leq I_\varepsilon(x,p,L) = g(x) + h(p)$$

which gives a uniform bound, implying that $\{x_\varepsilon\}$ is bounded in X as $\varepsilon \to 0$. It also implies that $L_\varepsilon(x_\varepsilon, p_\varepsilon) - \langle x_\varepsilon, p_\varepsilon\rangle \to 0$.

As before, we can thus assume the existence of $(x_*, p_*, L_*) \in X \times X^* \times \mathscr{L}$ such that, at least for some sequence $\{\varepsilon_n\}$ converging to zero, $x_{\varepsilon_n} \rightharpoonup x_*$ weakly in X, $p_n \to p_*$ strongly in X^* and $L_*(x_*, p_*) \leq \liminf L_\varepsilon(x_\varepsilon, p_\varepsilon)$. Thus, we have that $p_* \in \overline{\partial} L_*(x_*)$, and for any triple $(x, p, L) \in X \times \mathscr{L}$ satisfying (5.4):

$$g(x_*) + h(p_*) \leq \liminf I_\varepsilon(x_\varepsilon, L_\varepsilon) \leq I_\varepsilon(x, L) = g(x) + h(p),$$

which completes the proof. □

The second alternative, to impose growth conditions on a compactly embedded subspace, follows. As in the case of Theorems 5.4 and 5.6, the coercivity conditions can be imposed on the cost g or on the class $\mathscr{L}$.

Theorem 5.8 *Assume Y is a compactly embedded closed subspace of X. Assume that either: g is coercive as a functional on Y and bounded below and lower semicontinuous in X, or $\mathscr{L}$ is equicoercive in Y and g bounded below and lower semicontinuous in X. Let h be a bounded below, weakly lower semicontinuous cost on X^*. Assume that the class $\mathscr{L}$ is* Mosco compact*. Assume also that for some $L \in \mathscr{L}$ and some $p \in X^*$, problem* (5.4) *has a solution $x \in X$.*
Then there exist $(x_*, p_*, L_*) \in X \times X^* \times \mathscr{L}$ such that they solve (5.4), and they satisfy (5.10)*. Furthermore, the triple (x_*, p_*, L_*) can be obtained as a limit as $\varepsilon \to 0$ of minimizers of* (5.11)*.*

Proof: I_ε is bounded below. Consider then a minimizing subsequence $\{(x_n, p_n, L_n)\}$. By the coercivity condition, $\{x_n\}$ is bounded on Y, which gives the existence of some $x_\varepsilon \in X$ such that $x_n \to x_\varepsilon$ strongly on X. By the coercivity of h, we can recover some $p_\varepsilon \in X^*$ such that (for some subsequence) $p_n \rightharpoonup p_\varepsilon$ weakly in X^*. By the Mosco compactness, there exists some $L_\varepsilon \in \mathscr{L}$ such that $L_\varepsilon(x_\varepsilon, p_\varepsilon) \leq \liminf L_n(x_n, p_n)$ and since g and h are lower semicontinuous, we get

$$I_\varepsilon(x_\varepsilon, p_\varepsilon, L_\varepsilon) \leq \liminf I_\varepsilon(x_n, p_n, L_n)$$

and then

$$I_\varepsilon(x_\varepsilon, p_\varepsilon, L_\varepsilon) = \inf_{\mathscr{L} \times X} I_\varepsilon(x, p, L).$$

Now let $\varepsilon \to 0$. For some (x, p, L) which solves (5.4) this gives

$$I_\varepsilon(L_\varepsilon, x_\varepsilon) \leq I_\varepsilon(x, p, L) = g(x) + h(p)$$

which implies that, as $\varepsilon \to 0$,

$$L_\varepsilon(x_\varepsilon, p_\varepsilon) - \langle x_\varepsilon, p_\varepsilon \rangle \to 0$$

and $g(x_\varepsilon)$ bounded, which in turn implies that $\{x_\varepsilon\}$ is a bounded set in Y, which gives the existence of some x_* such that, at least for some sequence $\{\varepsilon_n\}$ converging to zero, $x_{\varepsilon_n} \to x_*$ strongly on X. As before, we can recover some $p_* \in X^*$ such that for some subsequence $p_{\varepsilon_n} \rightharpoonup p_*$ weakly in X^*.By using now the Mosco compactness, this gives the existence of some L_* such that

$$L_*(x_*, p_*) \leq \liminf L_\varepsilon(x_\varepsilon, p_\varepsilon)$$

which implies

$$L_*(x_*, p_*) - \langle x_*, p_* \rangle = 0$$

and by the lower semicontinuity of g and h, we get, for any triple (x, p, L) solving (5.4),

$$g(x_*) + h(p_*) \leq \liminf I_\varepsilon(x_\varepsilon, p_\varepsilon, L_\varepsilon) \leq g(x) + h(p),$$

which completes the proof. □

Chapter 6

Inverse and optimal control problems

In this chapter we provide sample applications of the results of the previous chapter. These applications are by no means exhaustive, and are meant to illustrate and explore briefly the generality and usability of the abstract setting developed so far.

6.1 Identification of a convex potential and a transport term

Here we consider a slight modification of one of the problems treated in [BK96]. Consider the following (non variational) equation:

$$\begin{cases} -\Delta u + b(u) & \ni & f + \vec{a} \cdot \nabla u & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega, \end{cases} \tag{6.1}$$

where $\vec{a} : \Omega \to \mathbb{R}^n$ is a vector field in Ω with positive divergence and $b : \mathbb{R} \rightrightarrows \mathbb{R}$ is a monotone (possibly multivalued) function. The equation will be parametrized by $\vec{a}$ and b, and the inverse problem will consist of recovering $\vec{a}$ and b from a given solution.

We will now put the problem in the context of Theorem 5.2.

b will vary in a class $\mathscr{B}$ such that for every $b \in \mathscr{B}$, for some positive constants α_1 and α_2: $\alpha_1(|x|-1) \leq b(x) \leq \alpha_2(|x|+1)$, so we will choose to parametrize it by a class of convex function with quadratic bounds by considering

$$\varphi(x) := \int_0^x b(t)\,\mathrm{d}t,$$

the function $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function such that $b = \partial\varphi$. This allows to define the function $\Phi : L^2(\Omega) \to \mathbb{R}$ by

$$\Phi(u) := \int_\Omega \varphi\big(u(x)\big)\,\mathrm{d}x.$$

It is known that Φ as defined above is a convex functional, and its subdifferential map is given by

$$\partial\Phi(u) = \{g(x) \in L^2(\Omega)\ :\ g(x) \in \partial\varphi\big(u(x)\big) = b\big(u(x)\big) \text{ for a.e. } x \in \Omega\}$$

(see for example [Gho08a, Proposition 2.7]). Taking the above into consideration, equation (6.1) can be equivalently written as

$$\begin{cases} -\Delta u + \partial\Phi(u) & \ni & f + \vec{a}\cdot\nabla u & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega \end{cases}$$

Hence $\mathscr{B}$ will be chosen as $\mathscr{B} := \{\partial\varphi(x)\ :\ \varphi \in \tilde{\mathscr{F}}\}$, where

$$\tilde{\mathscr{F}} = \{\varphi : \mathbb{R} \to \mathbb{R} \mid \beta_1|x|^2 - \gamma_1 < \varphi(x) < \beta_2|x|^2 + \gamma_2 \text{ and } \varphi \text{ is convex}\}$$

and so we can substitute for $\mathscr{B}$, the following class of convex functionals:

$$\mathscr{F} = \{\Phi(u) = \int_\Omega \varphi\big(u(x)\big)\,\mathrm{d}x\ :\ \varphi \in \tilde{\mathscr{F}}\}.$$

The field $\vec{a}$ will be chosen among an unspecified class $\mathscr{A}$ of *positive divergence vector fields* in Ω which is assumed to be compact with respect to the C^∞ topology (for example each component of the vector field smooth on $\bar{\Omega}$ and with all derivatives uniformly bounded in the C norm).

The choice for our main Banach space will be $L^2(\Omega)$, which means all conjugates will be under the L^2 pairing and norms will denote the L^2 norm. Also, Mosco convergence will be with respect to the L^2 topology.

The self-dual Lagrangian associated to equation (6.1) is given by

$$L_{\Phi,\vec{a}}(u,f) := \Psi_{\Phi,\vec{a}}\big(u\big) + \Psi^*_{\Phi,\vec{a}}\big(f + \vec{a}\cdot\nabla u + \frac{1}{2}div(\vec{a})u\big),$$

where

$$\Psi_{\Phi,\vec{a}}\big(u\big) := \begin{cases} +\infty, & u \notin H^1_0(\Omega), \\ \frac{1}{2}\|\nabla u\|^2 + \Phi(u) + \frac{1}{4}\int_\Omega div(\vec{a})u^2(x)\,\mathrm{d}x, & u \in H^1_0(\Omega). \end{cases}$$

(In order to be convinced of the self-duality, observe that $\vec{a}\cdot\nabla u + \frac{1}{2}div(\vec{a})u$ is the skew part of the linear operator given by $\vec{a}\cdot\nabla u$, and compare $L_{\Phi,\vec{a}}$ with item 2 in example 3.8).

Remark 6.1 *We could have chosen instead of $L_{\Phi,\vec{a}}$ as defined above, the functional $\Phi(u)+\Phi^*(\Delta u+\vec{a}\cdot\nabla u+f)$, which would more closely relate to the original choice made in [BK95, BK96]. This functional is not self-dual, but it is an appropriate representative for equation 6.1 (remember definition 2.19) and it has the needed coercivity conditions. The self-duality condition is not essential in the results from chapter 5.*

From the above, equation (6.1) is equivalent to

$$f \in \overline{\partial} L_{\Phi,\vec{a}}(u).$$

With this, we can parametrize the family of equations given by (6.1) by the class of self-dual Lagrangians $\mathscr{L}$, defined by

$$\mathscr{L} := \{L_{\Phi,\vec{a}} \,:\, \Phi \in \mathscr{F} \text{ and } \vec{a} \in \mathscr{A}\}.$$

The corresponding inverse problem in the above setting is the following:

Given $u_0 \in L^2(\Omega)$, find a pair $(\Phi_*, \vec{a}_*) \in \mathscr{F}\times\mathscr{A}$ such that for some u_* the triple $(u_*, \Phi_*, \vec{a}_*)$ solves (6.1) and it satisfies

$$\|u_0 - u_*\|^2 = \inf\{\|u_0 - u\|^2 \,:\, u \text{ solves (6.1) for some } (\Phi,\vec{a}) \in \mathscr{F}\times\mathscr{A}\}.$$

The above amounts to finding the optimal $L_* := L_{\Phi_*,\vec{a}_*}$ in the class $\mathscr{L}$.

In order to apply Theorem 5.2 to the corresponding inverse problem, we will need to show the following

Proposition 6.2 *The class $\mathscr{L}$ is Mosco compact.*

The above follows from the Mosco compactness of the class $\mathscr{F}$, the compactness of $\mathscr{A}$ and the following two lemmas.

Lemma 6.3 *Let $\{\Phi_n\} \subset \mathscr{F}$ and $\{\vec{a}_n\} \subset \mathscr{A}$ be sequences such that*

1. $\Phi_n \xrightarrow{M} \Phi_*$.
2. $\vec{a}_n \to \vec{a}_*$ *in the C^∞ topology.*

Then $\Psi_{\Phi_n,\vec{a}_n} \xrightarrow{M} \Psi_{\Phi_*,\vec{a}_*}$.

Proof: Let us denote $\Psi_n := \Psi_{\Phi_n,\vec{a}_n}$ and $\Psi_* := \Psi_{\Phi_*,\vec{a}_*}$. Fix any $u \in L^2(\Omega)$. Since $\Phi_n \in \mathscr{F}$, lemma 4.30 yields that $\Phi_n(u) \to \Phi_*(u)$. Looking at the definition of Ψ_n and Ψ_* this trivially gives that $\Psi_n(u) \to \Psi_*(u)$.

Let us then move on to prove the liminf property for $\{\Psi_n\}$. Consider a weakly convergent sequence $\{u_n\}$. If it is impossible to extract a subsequence which is bounded in $H_0^1(\Omega)$, then we trivially obtain $\liminf_n \Psi_n(u_n) = +\infty \geq \Psi_*(u_*)$.

Let us then assume that the sequence $\{u_n\}$ is bounded in $H_0^1(\Omega)$, hence, up to a subsequence, denoted by the same symbol, $u_n \to u_*$ strongly in $L^2(\Omega)$ and $u_n \rightharpoonup u_*$ weakly in $H_0^1(\Omega)$, hence, by weak lower semicontinuity of the norm, the strong L^2 convergence and the Mosco convergence of $\{\Phi_n\}$:

$$\liminf_n \Psi_n(u_n) \leq \frac{1}{2}\|\nabla u_*\|^2 + \Phi_*(u_*) + \frac{1}{4}\int_\Omega div(\vec{a})u_*^2(x)\,\mathrm{d}x = \Psi_*(u_*).$$

We have established that $\Psi_{\Phi_n,\vec{a}_n} \xrightarrow{M} \Psi_{\Phi_*,\vec{a}_*}$. □

Lemma 6.4 *Let $\{\Phi_n\} \subset \mathscr{F}$ and $\{\vec{a}_n\} \subset \mathscr{A}$ be sequences such that*

1. $\Phi_n \xrightarrow{M} \Phi_*$.
2. $\vec{a}_n \to \vec{a}_*$ *in the C^∞ topology.*

Then $L_{\Phi_n,\vec{a}_n} \xrightarrow{M} L_{\Phi_*,\vec{a}_*}$.

Proof: We have established in the previous lemma that under the above hypotheses $\Psi_{\Phi_n,\vec{a}_n} \xrightarrow{M} \Psi_{\Phi_*,\vec{a}_*}$. An analogous argument as that of lemma 4.30, given the growth conditions, will show that for any sequence $g_n \to g$ in $L^2(\Omega)$, $\Psi^*_{\Phi_n,\vec{a}_n}(g_n) \to \Psi^*_{\Phi_*,\vec{a}_*}(g)$. Fix any u in $H^1_0(\Omega)$ and f in $L^2(\Omega)$, denote $g_n := f + \vec{a}_n \cdot \nabla u + \frac{1}{2} div(\vec{a}_n)u$ and $g := f + \vec{a}_* \cdot \nabla u + \frac{1}{2} div(\vec{a}_*)u$; clearly $g_n \to g$ in $L^2(\Omega)$. Hence

$$L_{\Phi_n,\vec{a}_n}(u,f) = \Psi_{\Phi_n,\vec{a}_n}(u) + \Psi^*_{\Phi_n,\vec{a}_n}(g_n) \to \Psi_{\Phi_*,\vec{a}_*}(u) + \Psi^*_{\Phi_*,\vec{a}_*}(g).$$

The above is $L_{\Phi_n,\vec{a}_n}(u,f) \to L_{\Phi_*,\vec{a}_*}(u,f)$, which, combined with self-duality, yields Mosco convergence. □

Since given the conditions on $\mathscr{F}$, equation (6.1) always has a solution, we have all the needed conditions to apply Theorem 5.2, from which we have the following

Corollary 6.5 *There exists $\Phi_* \in \mathscr{F}$ and $a_* \in \mathscr{A}$ such that for $b_* = \partial\Phi$, and some u_* we have both*

$$\begin{cases} -\Delta u_* + b_*(u_*) & \ni & f + \vec{a} \cdot \nabla u & \text{in } \Omega \\ u_* & = & 0 & \text{on } \partial\Omega, \end{cases}$$

and

$$\|u_0 - u_*\|^2 = \inf\{\|u_0 - u\|^2 \ : \ u \text{ solves } (6.1) \text{ for some } (\Phi,\vec{a}) \in \mathscr{F} \times \mathscr{A}\}.$$

Furthermore, the triple $(\vec{a}_*, \Phi_*, u_*)$ can be recovered as a limit as $\varepsilon \to 0$ of the minimizers $(\vec{a}_\varepsilon, \Phi_\varepsilon, u_\varepsilon)$ for the functionals

$$I_\varepsilon(\vec{a}, \Phi, u) := \|u_0 - u\|^2 + \frac{1}{\varepsilon}\Big(L_{\Phi,\vec{a}}(u,f) - \int_\Omega uf \, \mathrm{d}x\Big).$$

6.2 Identification of a non conservative diffusion coefficient

Consider now the following equation

$$\begin{cases} -div\big(T(\nabla u)\big) & \ni & f & \text{in } \Omega \\ u & = & 0 & \text{on } \partial\Omega \end{cases} \tag{6.2}$$

where the (possibly multivalued) operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a (maximal) monotone vector field, not necessarily conservative (hence not the gradient of a corresponding potential function). This problem was considered in [BK96] in the case where T is assumed to be the subdifferential of a convex potential. The monotone non linearity T will be assumed to satisfy growth conditions (3.8), that is,

$$(x,p) \in T \Rightarrow \max\{C(\frac{\|x\|}{2}^2 - 1), D(\frac{\|p\|}{2}^2 - 1)\} \leq \langle x, p\rangle.$$

The corresponding condition for its self-dual Lagrangian L_T, constructed as in definition 3.2, is (3.9):

$$M(\|x\|^2 + \|p\|^2 - 1) \leq L_T(x,p) \leq N(\|x\|^2 + \|p\|^2 + 1).$$

As such we will parametrize equation (6.2) with the class

$$\mathscr{T} := \{T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \ : \ T \text{ is maximal monotone and satisfies (3.8)}\}.$$

We have the corresponding class of self-dual Lagrangians:

$$\hat{\mathscr{L}} := \{L_T \ : \ T \in \mathscr{T}\},$$

and each $L \in \hat{\mathscr{L}}$ satisfies growth conditions (3.9).

In order to write the self-dual Lagrangian corresponding to equation (6.2), we look at proposition 3.22, and see that

$$F_T(u,f) := \inf\{\int_\Omega L_T\big(\nabla u(x), w(x)\big)\,dx; w \in L^2(\Omega;\mathbb{R}^N), -\mathrm{div}(w) = f\}$$

is self-dual, and that any solution u of $f \in \overline{\partial} F_T(u)$ will be a solution to (6.2).

Then, in order to use our results we will consider the class

$$\mathscr{L} := \{F(u,f) = \inf\{\int_\Omega L\big(\nabla u(x), w(x)\big)\,\mathrm{d}x; w \in L^2(\Omega;\mathbb{R}^N), -\mathrm{div}(w) = f\} : L \in \hat{\mathscr{L}}\},$$

and show

Proposition 6.6 *The class $\mathscr{L}$ is Mosco compact.*

Proof: Combining propositions 4.29 and 4.30, given the growth conditions on $\hat{\mathscr{L}}$, it follows that for any sequence $\{L_n\}$ we can extract a subsequence such that for some L and for any $w \in L^2(\Omega;\mathbb{R}^N)$, with $-\mathrm{div}(w) = f$:

$$\limsup_n F_n(u,f) \le \limsup_n \int_\Omega L_n\big(\nabla u(x), w(x)\big)\,\mathrm{d}x = \int_\Omega L\big(\nabla u(x), w(x)\big)\,\mathrm{d}x.$$

Since the above holds for any such w, we get

$$\limsup_n F_n(u,f) \le F(u,f),$$

and we are done. □

Then we can simply use Theorem 5.2 to get the existence of an optimal solution to the inverse problem, that is

Corollary 6.7 *For any given $u_0 \in L^2(\Omega)$, there exists $T_* \in \mathscr{T}$ such that for some u_* we have both*

$$\begin{cases} -div\big(T_*(\nabla u_*)\big) & \ni & f & in\ \Omega \\ u_* & = & 0 & on\ \partial\Omega \end{cases}$$

and

$$\|u_0 - u_*\|^2 = \inf\{\|u_0 - u\|^2 \ :\ u \text{ solves } (6.2) \text{ for some } T \in \mathscr{T}\}.$$

Furthermore, the couple (T_*, u_*) can be recovered as a limit as $\varepsilon \to 0$ of the minimizers $(T_\varepsilon, u_\varepsilon)$ of the functional

$$I_\varepsilon(T,u) := \|u_0 - u\|^2 + \frac{1}{\varepsilon}\Big(F_T(u,f) - \int_\Omega uf\,\mathrm{d}x\Big).$$

6.3 Obtaining the cheapest temperature source control for a desired temperature profile

Parabolic equations can be handled via self-dual Lagrangians (see [Gho08a] and [Gho08b] for details on how different time-boundary conditions are handled). Now we present the following control problem, involving the heat equation.

Consider the heat equation

$$\begin{array}{rcll} \dot{u}(x,t)-\Delta u(x,t) &=& f(x) & \text{in } \Omega\times[0,1] \\ u(x,t) &=& 0 & \text{on } \partial\Omega\times[0,1] \\ u(x,0) &=& g(x) & \text{in } \Omega \end{array} \tag{6.3}$$

Suppose we are given a specified $u_0(x,t)$, with $u_0(x,0)=g(x)$, which represents a desired temperature profile to be achieved over Ω along the time interval $[0,1]$. We can control the temperature by specifying the heat source f over the domain Ω.

Assume the cost of maintaining such temperature is given by

$$C(f)=\int_0^1 \|f\|^2\,\mathrm{d}t.$$

Assuming we want to achieve an equilibrium between the cost of f and achieve the closest possible behaviour to the profile u_0, we want to minimize

$$\int_0^1 \|u(x,t)-u_0(x,t)\|^2+\|f\|^2\,\mathrm{d}t$$

among all possible solutions u of (6.3) for some f.

The self-dual Lagrangian (translated by f) associated to (6.3) is

$$\begin{aligned} L_f(u,0)=&\ \frac{1}{2}\int_0^1\int_\Omega\Big(|\nabla u(t,x)|^2+\Big|\nabla(-\Delta)^{-1}\big(f(x)-\dot{u}(t,x)\big)\Big|^2-2f(x)u(t,x)\Big)\,\mathrm{d}x\,\mathrm{d}t \\ &+\int_\Omega |g(x)|^2\,\mathrm{d}x-2\int_\Omega u(0,x)g(x)\,\mathrm{d}x+\frac{1}{2}\int_\Omega(|u(0,x)|^2+|u(1,x)|^2)\,\mathrm{d}x. \end{aligned}$$

Equation (6.3), then, is solved by the minimizer of

$$J(u,f)=L_f(u,0).$$

We must observe that $J(u,f)$ *is not a self-dual function* on (u,f), see the discussion in section 5.4.

We will then assume that f is further restricted to a compact class $P \subset L^2(\Omega)$. Fix $f_i \in L^2(\Omega)$, $i = 1,\dots,m$, and choose

$$P := \{f = \sum_{i=1}^{m} \alpha_i f_i \ :\ \alpha_i \in [0,1]\}.$$

This class represents the ability to control the heat source by choosing a finite set of parameters. It is obviously compact in $L^2(\Omega)$.

Under this compactness assumption, Theorem 5.7 immediately yields

Corollary 6.8 *There exists $f_* \in P$ and u_* such that the pair solves* (6.3) *and*

$$\begin{aligned}\int_0^1 \|u_*(x,t) - u_0(x,t)\|^2 + \|f_*\|^2 \,\mathrm{d}t \quad &= \quad \inf\{\int_0^1 \|u(x,t) - u_0(x,t)\|^2 + \|f\|^2 \,\mathrm{d}t \ : \\ &\qquad (u,f) \text{ solve } (6.3) \text{ for some } f \in P\}.\end{aligned}$$

Furthermore, the couple (f_*, u_*) can be recovered as a limit as $\varepsilon \to 0$ of the minimizers $(f_\varepsilon, u_\varepsilon)$ of the functional

$$I_\varepsilon(f,u) := \|u_0 - u\|^2 + \frac{1}{\varepsilon} J(u,f).$$

6.4 Identification of a convex potential and a transport term in a dynamical setting

Here we consider the dynamical version of (6.1), that is

$$\begin{cases} \dot{u} - \Delta u + b(u) & \ni & f + \vec{a}\cdot\nabla u & \text{in } \Omega \times [0,T] \\ u(x,t) & = & 0 & \text{on } \partial\Omega \times [0,T], \\ u(x,0) & = & g(x) & \text{in } \Omega. \end{cases} \tag{6.4}$$

The procedure of moving from stationary to dynamic equations is detailed in [Gho08b, chapter 6], which we follow here.

Consider the stationary counterpart of (6.4), namely (6.1). Starting with the self-dual Lagrangian for (6.1), $L_{\Phi,\vec{a}}$, we extend this for u in the space $A^2_{H^1_0(\Omega)}[0,T]$

and $p = (p_1, p_0)$ in the dual of $A^2_{H^1_0(\Omega)}[0,T]$ (identified as $H^1_0(\Omega) \times L^2_{L^2(\Omega)}$) as follows

$$\begin{aligned} F_{\Phi,\vec{a}}(u,p) &= \int_0^1 L_{\Phi,\vec{a}}\big(u(t)+p_0(t), -\dot{u}(t)\big)\,\mathrm{d}t + \\ &+ \frac{1}{4}\|u(1)-u(0)+p_1\|^2 + \langle u(1)-u(0)+p_1, g\rangle + \|\frac{u(0)+u(1)}{2}+g\|^2, \end{aligned}$$

where the term on the second line is the self-dual Lagrangian corresponding to the boundary conditions being imposed.

Choosing $p_f := \big((-\Delta^{-1})f, 0\big)$, equation (6.4) is equivalent to equation

$$p_f \in \overline{\partial} F_{\Phi,\vec{a}}(u).$$

As in section 6.1, we will consider $\vec{a} \in \mathscr{A}$ and $b = \partial\Phi$, with $\Phi \in \mathscr{F}$. The classes $\mathscr{F}$ and $\mathscr{A}$ are taken as in section 6.1. In this case the class of self-dual Lagrangians parametrizing the equations is then given by

$$\begin{aligned} \mathscr{L} &:= \{F_{\Phi,\vec{a}}(u,p) = \int_0^1 L_{\Phi,\vec{a}}\big(u(t)+p_0(t), -\dot{u}(t)\big)\,\mathrm{d}t + \\ &\quad + S(u(1)-u(0)+p_1, \frac{u(0)+u(1)}{2}) \;:\; \Phi \in \mathscr{F} \text{ and } \vec{a} \in \mathscr{A}\}. \end{aligned}$$

Proposition 6.9 *$\mathscr{L}$ is Mosco compact.*

Proof: In the proof of proposition 6.2, the following was established: If $\Phi_n \xrightarrow{M} \Phi_*$ and $\vec{a}_n \xrightarrow{C^\infty} \vec{a}_*$, then $L_{\Phi_n,\vec{a}_n}(\tilde{u},\tilde{p}) \to L_{\Phi_*,\vec{a}_*}(\tilde{u},\tilde{p})$. Hence, for such $\{\Phi_n\}$ and $\{\vec{a}_n\}$, for a.e. $t \in [0,1]$:

$$L_{\Phi_n,\vec{a}_n}\big(u(t),p(t)\big) \to L_{\Phi_*,\vec{a}_*}\big(u(t),p(t)\big).$$

Given the boundedness conditions, the dominated convergence theorem now yields

$$\int_0^1 L_{\Phi_n,\vec{a}_n}\big(u(t),p(t)\big)\,\mathrm{d}t \to \int_0^1 L_{\Phi_*,\vec{a}_*}\big(u(t),p(t)\big)\,\mathrm{d}t.$$

Since $\mathscr{F}$ and $\mathscr{A}$ are compact in their respective topologies, and given the self-duality of the above functionals, this completes the proof. □

We can thus apply Theorem 5.2, which yields

Corollary 6.10 *Given any $u_0 \in L^2_{L^2(\Omega)}$, there exists $\Phi_* \in \mathscr{F}$ and $a_* \in \mathscr{A}$ such that for $b_* = \partial\Phi$, and some u_* we have both*

$$\begin{cases} \dot{u}_* - \Delta u_* + b(u_*) & \ni & f + \vec{a}\cdot\nabla u_* & \text{in } \Omega\times[0,T] \\ u_*(x,t) & = & 0 & \text{on } \partial\Omega\times[0,T], \\ u_*(x,0) & = & g(x) & \text{in } \Omega. \end{cases}$$

and

$$\int_0^1 \|u_0(t)-u_*(t)\|^2\,\mathrm{d}t \quad = \quad \inf\{\int_0^1 \|u_0(t)-u(t)\|^2\,\mathrm{d}t \ : \ u \text{ solves } (6.4) \text{ for some } (\Phi,\vec{a})\in\mathscr{F}\times\mathscr{A}\}.$$

Furthermore, the triple $(\vec{a}_*,\Phi_*,u_*)$ can be recovered as a limit as $\varepsilon\to 0$ of the minimizers $(\vec{a}_\varepsilon,\Phi_\varepsilon,u_\varepsilon)$ of the functional

$$I_\varepsilon(\vec{a},\Phi,u) := \int_0^1 \|u_0-u\|^2\,\mathrm{d}t + \frac{1}{\varepsilon}\Big(F_{\Phi,\vec{a}}(u,p_f) - \int_0^1\int_\Omega uf\,\mathrm{d}x\,\mathrm{d}t\Big).$$

6.5 Identification of a non conservative diffusion coefficient in a dynamical setting

Here we consider the dynamical version of (6.2):

$$\begin{cases} \dot{u} - div\Big(T\big(\nabla_x u(x,t)\big)\Big) & \ni & f & \text{in } \Omega\times[0,1], \\ u(x,t) & = & 0 & \text{on } \partial\Omega\times[0,1], \\ u(x,0) & = & 0 & \text{in } \Omega. \end{cases} \tag{6.5}$$

What follows is similar to what was done in the previous section, in particular the choice of spaces will be the same. Given the corresponding self-dual Lagrangian for equation (6.2), F_T, we lift this to

$$\begin{aligned} G_T(u,p) &= \int_0^1 F_T\big(u(t)+p_0(t), -\dot{u}(t)\big)\,\mathrm{d}t + \\ &+ \frac{1}{4}\|u(1)-u(0)+p_1\|^2 + \langle u(1)-u(0)+p_1, g\rangle + \|\frac{u(0)+u(1)}{2}+g\|^2, \end{aligned}$$

which is a self-dual Lagrangian, corresponding to equation (6.5). As in the previous section, choosing $p_f := \big((-\Delta^{-1})f, 0\big)$, equation (6.5) is equivalent to equation

$$p_f \in \overline{\partial} G_T(u).$$

Then we take $\mathscr{L}$ as in Section 6.2 and consider show that the class

$$\mathscr{G} := \{G_T \ : \ F_T \in \mathscr{L}\}.$$

Proposition 6.11 *$\mathscr{G}$ is Mosco compact.*

Proof: In the proof of proposition 6.6 it was established that for any sequence $\{F_n\}$ in $\mathscr{L}$ Mosco converging to F_*, we have: $\limsup_n F_n(u,f) \leq F_*(u,f)$. This is equivalent to, given the Mosco convergence,

$$\lim_n F_n(u,f) = F_*(u,f).$$

Hence, for such $\{F_n\}$, for a.e. $t \in [0,1]$:

$$F_n\big(u(t), p(t)\big) \to F_*\big(u(t), p(t)\big).$$

Given the boundedness conditions, the dominated convergence theorem now yields

$$\int_0^1 F_n\big(u(t), p(t)\big)\,\mathrm{d}t \to \int_0^1 F_*\big(u(t), p(t)\big)\,\mathrm{d}t.$$

Since $\mathscr{L}$ is also Mosco compact, and given the self-duality of the above functionals, this completes the proof. □

We can, again, apply Theorem 5.2, which yields

Corollary 6.12 *Given any $u_0 \in L^2_{L^2(\Omega)}$, there exists $T_* \in \mathscr{T}$ such that for some u_* we have both*

$$\begin{cases} \dot{u}_* - div\Big(T\big(\nabla_x u_*(x,t)\big)\Big) & \ni & f \quad in\ \Omega \times [0,1] \\ u_*(x,t) & = & 0 \quad on\ \partial\Omega \times [0,1] \\ u_*(x,0) & = & 0 \quad in\ \Omega \end{cases}$$

and

$$\int_0^1 \|u_0(t) - u_*(t)\|^2 \,\mathrm{d}t = \inf\{\int_0^1 \|u_0(t) - u(t)\|^2 \,\mathrm{d}t \,:\, u \text{ solves } (6.5) \text{ for some } T \in \mathscr{T}\}.$$

Furthermore, the couple (T_*, u_*) can be recovered as a limit as $\varepsilon \to 0$ of the minimizers $(T_\varepsilon, u_\varepsilon)$ of the functional

$$I_\varepsilon(T,u) := \int_0^1 \|u_0 - u\|^2 \,\mathrm{d}t + \frac{1}{\varepsilon}\Big(G_T(u, p_f) - \int_0^1 \int_\Omega uf \,\mathrm{d}x\,\mathrm{d}t\Big).$$

Chapter 7

Homogenization of periodic maximal monotone operators via self-dual calculus

In this chapter we consider the homogenization of the problem

$$\begin{cases} \tau_n(x) & \in & T(\frac{x}{\varepsilon_n}, \nabla u_n(x)) & x \in \Omega, \\ -\mathrm{div}(\tau_n(x)) & = & u_n^*(x) & x \in \Omega, \\ u_n(x) & = & 0 & x \in \partial\Omega, \end{cases} \tag{7.1}$$

where Ω is a bounded domain of $\mathbb{R}^N$ and $T : \Omega \times \mathbb{R}^N \to \mathbb{R}^N$ is a measurable map on $\Omega \times \mathbb{R}^N$ such that $T(x, \cdot)$ is maximal monotone on $\mathbb{R}^N$ for almost all $x \in \Omega$, and such that $T(., \xi)$ is Q-periodic for an open non-degenerate parallelogram Q in $\mathbb{R}^n$. This problem has been investigated in recent years by many authors. We refer the interested reader to [Att84, BC94, BCD92, PDD90, DMV07, FMT09, FM87, PD90] for related results.

The particular case where the maximal monotone operator is a subdifferential of the form

$$T(x, \xi) = \partial_\xi \psi(x, \xi), \tag{7.2}$$

with $\psi : \Omega \times \mathbb{R}^N \to \mathbb{R}$ being a convex function in the second variable is particularly appealing and completely understood. Indeed, under appropriate boundedness and coercivity conditions on ψ, say

$$C_0(|\xi|^p - 1) \le \psi(x,\xi) \le C_1(|\xi|^p + 1) \qquad \text{for all } (x,\xi) \in \Omega \times \mathbb{R}^N,$$

where $1 < p < \infty$ and C_0, C_1 are positive constants, one can use a variational approach to identify for a given $u^* \in W^{-1,p}(\Omega)$, the solution (u,τ) of (7.1) as the respective minima of the problems

$$\inf\left\{\int_\Omega \psi(x,\nabla u(x))\,dx - \int_\Omega u^*(x)u(x)\,dx;\ u \in W_0^{1,p}(\Omega)\right\}, \tag{7.3}$$

and

$$\inf\left\{\int_\Omega \psi^*(x,\tau(x))\,dx;\ \mathrm{div}(\tau) = u^*\right\}, \tag{7.4}$$

where ψ^* is the Fenchel-Legendre dual (in the second variable) of ψ. In this case, the classical concept of Γ-convergence –introduced by DeGiorgi– can be used to show that if $u_n^* \to u^*$ strongly in $W^{-1,q}(\Omega)$ with $q = \frac{p}{p-1}$, then up to a subsequence $u_n \to u$ weakly in $W_0^{1,p}(\Omega)$ and $\tau_n \to \tau$ weakly in $L^q(\Omega;\mathbb{R}^N)$, where u is a solution and τ is a momentum of the homogenized problem

$$\begin{cases} \tau(x) \in T_{hom}(\nabla u(x)) & a.e.\ x \in \Omega, \\ -\mathrm{div}(\tau(x)) = u^*(x) & a.e.\ x \in \Omega. \end{cases} \tag{7.5}$$

Here T_{hom} can be defined variationally as follows: for $\xi \in \mathbb{R}^N$, $T_{hom}(\xi) = \partial\psi_{hom}(\xi)$, where

$$\psi_{hom}(\xi) := \min_{\varphi \in W_\#^{1,p}(Q)} \frac{1}{|Q|}\int_Q \psi\big(x, \xi + \nabla\varphi(x)\big)\,\mathrm{d}x, \tag{7.6}$$

and

$$W_\#^{1,p}(Q) = \{u \in W^{1,p}(Q); \int_Q u(x)\,dx = 0 \text{ and } u \text{ is } Q-\text{periodic}\}. \tag{7.7}$$

A similar result can be obtained for general maximal monotone maps $T : \Omega \times \mathbb{R}^N \to \mathbb{R}^N$ with appropriate boundedness conditions (see below), by using the more cumbersome graph convergence (or G-convergence, a notion based on classic Kuratowski-Painlevé convergence) methods. In this case, T_{hom} is defined by the following non-variational formula

$$T_{hom}(\xi) = \Big\{ \int_Q g(y)\,dy \in \mathbb{R}^N\,; g \in L^q_\#(Q;\mathbb{R}^N),\, g(y) \in T(y, \xi + \nabla\psi(y)) \text{ a.e. in } Q \text{ for some } \psi \in W^{1,p}_\#(Q)\Big\}, \tag{7.8}$$

where

$$L^q_\#(Q;\mathbb{R}^N) := \left\{ g \in L^q(Q;\mathbb{R}^N);\, \int_Q \langle g(y), \nabla\varphi(y)\rangle_{\mathbb{R}^N}\,dy = 0 \text{ for every } \varphi \in W^{1,p}_\#(Q) \right\}. \tag{7.9}$$

The goal here is to describe how the variational approach to maximal monotone operators is particularly well suited to deal with the homogenization of such equations, first by showing that –just as in the case of a convex potential (7.2)– the limiting process can be handled through Γ-convergence of associated self dual Lagrangians, and secondly by giving a variational characterization for the limiting vector field (7.8) in the same spirit as in (7.6).

The following is the main result for this chapter:

Theorem 7.1 *Let Ω be a domain in $\mathbf{R^N}$, $q, p > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, and assume $u_n^* \to u^*$ strongly in $W^{-1,q}(\Omega)$. Let u_n (resp., τ_n) be (weak) solutions in $W_0^{1,p}(\Omega)$ (resp., momenta in $L^q(\Omega;\mathbb{R}^N)$) for the Dirichlet boundary value problems (7.1), where $T : \Omega \times \mathbb{R}^N \to \mathbb{R}^N$ belongs to $M_{\Omega,p}(\mathbb{R}^N)$.*

If $T(.,\xi)$ is Q-periodic for an open non-degenerate parallelogram Q in $\mathbb{R}^n$ then, up to a subsequence

$$u_n \to u \qquad \text{weakly in } W_0^{1,p}(\Omega),$$

$$\tau_n \to \tau \qquad \text{weakly in } L^q(\Omega;\mathbb{R}^N),$$

where u is a solution and τ is a momentum of the homogenized problem

$$\begin{cases} \tau(x) \in T_{hom}(\nabla u(x)) & a.e. \quad x \in \Omega, \\ -\mathrm{div}(\tau(x)) = u^*(x) & a.e. \quad x \in \Omega, \\ u \in W_0^{1,p}(\Omega). \end{cases} \tag{7.10}$$

Here $T_{hom} = \bar{\partial} L_{hom}$, with L_{hom} being a self dual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$ defined *by*

$$L_{hom}(a,b) := \min_{\substack{\varphi \in W^{1,p}_{\#}(Q)\\ g \in L^q_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|}\int_Q L\big(x, a + D\varphi(x), b + g(x)\big)\,\mathrm{d}x, \tag{7.11}$$

where for each $x \in \Omega$, $L(x,\cdot,\cdot)$ is a self dual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$ such that

$$T(x,\cdot) = \bar{\partial} L(x,\cdot). \tag{7.12}$$

We include a homogenization result via Γ-convergence for general Q-periodic Lagrangians which are not necessarily self dual. This is then applied to obtain the result claimed in Theorem 7.1 above in the case of self dual Lagrangians.

7.1 Auxiliary results

We open with results used in some of the following proofs.

Lemma 7.2 *Assume $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a convex function such that $C_0(|a|^p + |b|^q - 1) \le L(a,b) \le C_1(|a|^p + |b|^q + 1)$ for all $a,b \in \mathbb{R}^N$ where $p,q > 1$ are two constants. Suppose Ω is a bounded open domain in $\mathbb{R}^N$ and $\tau_1 \in L^p(\Omega;\mathbb{R}^N)$ and $\tau_2 \in L^q(\Omega;\mathbb{R}^N)$ are two piecewise constant functions such that*

$$\tau_1(x) = a_i, \quad x \in \Omega_i, \text{ and } \tau_2(x) = b_i, \quad x \in \Omega_i,$$

where $\{\Omega_i\}_{i\in I}$ is a finite polyhedral partitions of Ω, and $\{a_i\}_{i\in I}, \{b_i\}_{i\in I}$ are two sequences $\in \mathbb{R}^N$. Then

$$\min_{\substack{f \in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f = 0}} \int_\Omega L\big(\tau_1, \tau_2(x) + f(x)\big)\,\mathrm{d}x \ge \sum_{i\in I} |\Omega_i| \inf_{\eta_i \in \mathbb{R}^n} L(a_i, b_i + \eta_i).$$

Proof We first prove a stronger result (actually an equality) when the set index I is a singleton. For any constant $\eta \in \mathbb{R}^N$ we have

$$\min_{\substack{f \in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f = 0}} \int_\Omega L\big(a, b + f(x)\big)\,\mathrm{d}x \le \int_\Omega L(a, b + \eta)\,\mathrm{d}x = |\Omega| L(a, b + \eta),$$

from which we obtain

$$\min_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \mathrm{div} f=0}} \int_\Omega L\big(a,b+f(x)\big)\,\mathrm{d}x \quad \leq \quad \inf_{\eta\in\mathbb{R}^N} |\Omega| L(a,b+\eta),$$

Let now $\tilde{f}$ be the element in $L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div}\tilde{f}=0$ such that

$$\int_\Omega L\big(a,b+\tilde{f}(x)\big)\,\mathrm{d}x = \min_{\substack{f\in L^q(\Omega;\mathbb{R}^N))\\ \mathrm{div} f=0}} \int_\Omega L\big(a,b+f(x)\big)\,\mathrm{d}x.$$

Using Jensen's inequality, we obtain

$$\begin{aligned}
\inf_{\eta\in\mathbb{R}^N} |\Omega| L(a,b+\eta) &\leq |\Omega| L\big(a,b+\frac{1}{|\Omega|}\int_\Omega \tilde{f}(x)\,\mathrm{d}x\big) \\
&= |\Omega| L\big(\frac{1}{|\Omega|}\int_\Omega a\,\mathrm{d}x, \frac{1}{|\Omega|}\int_\Omega b+\tilde{f}(x)\,\mathrm{d}x\big) \\
&\leq \int_\Omega L\big(a,b+\tilde{f}(x)\big)\,\mathrm{d}x \\
&= \min_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \mathrm{div} f=0}} \int_\Omega L\big(a,b+f(x)\big)\,\mathrm{d}x.
\end{aligned}$$

This completes the proof when I is a singleton. Now we prove it for the general case. Note first that, using the above argument on each Ω_i, we have

$$\inf_{\substack{g\in L^q(\Omega_i;\mathbb{R}^N)\\ \mathrm{div} g=0}} \int_{\Omega_i} L\big(a_i,b_i+g(x)\big)\,\mathrm{d}x = \inf_{\eta_i\in\mathbb{R}^N} |\Omega_i| L(a_i,b_i+\eta_i). \tag{7.13}$$

One also can easily deduce that

$$\inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \mathrm{div} f=0}} \int_\Omega L\big(\tau_1(x),\tau_2(x)+f(x)\big)\,\mathrm{d}x \geq \sum_i \inf_{\substack{f_i\in L^q(\Omega_i;\mathbb{R}^N)\\ \mathrm{div} f_i=0}} \int_{\Omega_i} L\big(a_i,b_i+f_i(x)\big)\,\mathrm{d}x \tag{7.14}$$

In fact if $\inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L\big(\tau_1(x),\tau_2(x)+f(x)\big)\,\mathrm{d}x = \int_\Omega L\big(\tau_1(x),\tau_2(x)+\bar f(x)\big)\,\mathrm{d}x$ for some $\bar f \in L^q(\Omega;\mathbb{R}^N)$ with $\operatorname{div}(\bar f)=0$, then

$$\begin{aligned}\int_\Omega L\big(\tau_1(x),\tau_2(x)+\bar f(x)\big)\,\mathrm{d}x &= \sum_{i\in I}\int_{\Omega_i} L\big(a_i,b_i+\bar f(x)\big)\,\mathrm{d}x\\ &\geq \sum_{i\in I}\inf_{\substack{f_i\in L^q(\Omega_i;\mathbb{R}^N)\\ \operatorname{div} f_i=0}}\int_{\Omega_i} L\big(a_i,b_i+f_i(x)\big)\,\mathrm{d}x.\end{aligned}$$

The proof therefore follows from combining (7.13) and (7.14). □

The following three Lemmas are standard and we refer to [Suq82] for the proof.

Lemma 7.3 *Let $r\geq 1$ and $f\in L^r(Q)$. Then f can be extended by periodicity to a function (still denoted by f) belonging to $L^r_{loc}(\mathbb{R}^N)$. Moreover, if (ε_k) is a sequence of positive real numbers converging to 0 and $g_k(x)=g(\frac{x}{\varepsilon_k})$, then*

$$\text{if } 1\leq r<\infty, \text{ then } f_k\to M(f)=\tfrac{1}{|Q|}\int_Q f(x)\,\mathrm{d}x \text{ weakly in } L^r_{loc}(\mathbb{R}^N),$$

and

$$\text{if } r=\infty, \text{ then } f_k\to M(f) \text{ weak}^* \text{ in } L^\infty(\mathbb{R}^N).$$

Lemma 7.4 *Let $r>1$ and $u\in W^{1,r}_\#(Q)$, then u can be extended by periodicity to an element of $W^{1,r}_{loc}(\mathbb{R}^N)$.*

Lemma 7.5 *Let $r>1$ and $r'=\frac{r}{r-1}$. Let $g\in L^{r'}(Q;\mathbb{R}^N)$ such that $\int_Q\langle g(x),\nabla v(x)\rangle\,dx = 0$ for every $v\in W^{1,r}_\#(Q)$. Then g can be extended by periodicity to an element of $L^{r'}_{loc}(\mathbb{R}^N;\mathbb{R}^N)$, still denoted by g, such that $\operatorname{div}(g)=0$ in $D'(\mathbb{R}^N)$.*

7.2 Variational formula for the homogenized maximal monotone vector field

Given a maximal monotone family T in $M_{\Omega,p}(\mathbb{R}^N)$ that is Q-periodic for an open non-degenerate parallelogram Q in $\mathbb{R}^n$, its homogenization T_{hom} is normally given by the non-variational formula (7.8). In this section, we shall give a variational formulation for the vector field T_{hom} in terms of a suitably homogenized self dual

Lagrangian L_{hom} derived from the Ω-dependent self dual Lagrangian associated to T.

Theorem 7.6 *Assume $T \in M_{\Omega,p}(\mathbb{R}^N)$ is Q-periodic and let L be an Ω-dependent self dual Lagrangian such that $T(x,.) = \bar{\partial} L(x,.)$ given by Proposition 3.20. If the operator T_{hom} is given by (7.8), then $T_{hom} = \bar{\partial} L_{hom}$ where L_{hom} is the self dual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$ given by*

$$L_{hom}(\xi,\eta) = \min_{\substack{\varphi \in W^{1,p}_{\#}(Q)\\ g \in L^q_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|} \int_Q L\big(x, \xi + \nabla\varphi(x), \eta + g(x)\big)\,\mathrm{d}x. \tag{7.15}$$

The proof relies on the following propositions. First, we show that the homogenized Lagrangian L_{hom} inherits many of the properties of the original Ω-dependent Lagrangian L such as convexity, boundedness and coercivity.

Proposition 7.7 *Assume L is an Ω-dependent Lagrangian on $\Omega \times \mathbb{R}^N \times \mathbb{R}^N$ satisfying (3.11) for some $p,q > 1$. Then L_{hom} is convex and lower semi continuous, and for every $a^*, b^* \in \mathbb{R}^n$,*

$$L^*_{hom}(a^*,b^*) = \inf_{\substack{\varphi \in W^{1,q'}_{\#}(Q)\\ g \in L^{p'}_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|} \int_Q L^*\big(x, a^* + g(x), b^* + \nabla\varphi(x)\big)\,\mathrm{d}x, \tag{7.16}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} + \frac{1}{q'} = 1$. Furthermore,

$$C_0(|a|^p + |b|^q - 1) \le L_{hom}(a,b) \le C_1(1 + |a|^p + |b|^q) \qquad \textit{for all } a,b \in \mathbb{R}^n. \tag{7.17}$$

The following gives the relation between the subdifferentials of L_{hom} and of L.

Proposition 7.8 *For each $a,b \in \mathbb{R}^n$, the subdifferential map $\partial L_{hom}(a,b)$ is given by*

$$\partial L_{hom}(a,b) = \frac{1}{|Q|} \int_Q \partial L\big(y, a + \nabla\tilde{\varphi}(y), b + \tilde{g}(y)\big)\,\mathrm{d}y,$$

where $\tilde{\varphi} \in W^{1,p}_{\#}(Q)$ and $\tilde{g} \in L^q_{\#}(Q;\mathbb{R}^N)$ are such that

$$L_{hom}(a,b) = \frac{1}{|Q|} \int_Q L\big(y, a + \nabla\tilde{\varphi}(y), b + \tilde{g}(y)\big)\,\mathrm{d}y.$$

We need a few preliminary facts. For each $1 < r < \infty$, set

$$E_r := \{f = \nabla u \in L^r(Q;\mathbb{R}^N);\text{ for some } u \in W^{1,r}_{\#}(Q)\}$$

and

$$E_r + \mathbb{R}^n := \{f + \eta \ :\ f \in E_r,\, \eta \in \mathbb{R}^n\}.$$

The Poincaré-Wirtenger inequality which states that for D bounded open and convex, there exists $K := K(r,D) > 0$ such that

$$\|u - \tfrac{1}{|D|}\int_D u\|_{L^r(D)} \leq K\|\nabla u\|_{W^{1,r}(D)} \text{ for every } u \in W^{1,r}(D),$$

implies that $E_r + \mathbb{R}^n$ is a convex weakly closed subset of $L^r(Q;\mathbb{R}^N)$. The indicator function of $E_r + \mathbb{R}^n$,

$$\chi_{E_r+\mathbb{R}^n}(f) = \begin{cases} 0 & f \in E_r + \mathbb{R}^n, \\ +\infty & f \in L^r(Q;\mathbb{R}^N) \setminus (E_r + \mathbb{R}^n), \end{cases}$$

is therefore convex and lower semi-continuous in $L^r(Q;\mathbb{R}^N)$. Assuming that r' is the conjugate of r, i.e., $\frac{1}{r} + \frac{1}{r'} = 1$, define

$$E_{r'}^{\perp} := \Big\{g \in L^{r'}(Q;\mathbb{R}^N);\ \int_\Omega \langle f(x), g(x)\rangle_{\mathbb{R}^N}\, dx = 0 \quad \text{for all } f \in E_r + \mathbb{R}^n\Big\}.$$

The Fenchel-Legendre dual $\chi^*_{E_r+\mathbb{R}^n}$ of $\chi_{E_r+\mathbb{R}^n}$ is then given by,

$$\begin{aligned} \chi^*_{E_r+\mathbb{R}^n}(g) &= \sup_{f \in L^r(Q;\mathbb{R}^N)} \Big\{ \int_Q \langle f(x), g(x)\rangle_{\mathbb{R}^N}\, dx - \chi_{E+\mathbb{R}^n}(f) \Big\} \\ &= \sup_{f \in E_r+\mathbb{R}^n} \int_Q \langle f(x), g(x)\rangle_{\mathbb{R}^N}\, dx = \chi_{E_{r'}^{\perp}}(g), \end{aligned}$$

for all $g \in L^{r'}(Q;\mathbb{R}^N)$. Also due to the convexity and lower semi-continuity of $\chi_{E_r+\mathbb{R}^n}$ one has $\chi^*_{E_{r'}^{\perp}} = \chi_{E_r+\mathbb{R}^n}$. Similarly one can deduce that,

$$\chi^*_{E_{r'}^{\perp}+\mathbb{R}^n}(f) = \chi_{E_r}(f)$$

for all $f \in L^r(Q;\mathbb{R}^N)$. Note also that E_r is the isometric image of $W^{1,r}_{\#}(Q)$ by ∇ and $E_r^{\perp} = L^r_{\#}(Q;\mathbb{R}^N)$.

Proof of Proposition 7.7. We first prove (7.16). Fix $(a^*,b^*) \in \mathbb{R}^n \times \mathbb{R}^n$ and write

$$L^*_{hom}(a^*,b^*) = \sup_{(a,b)\in\mathbb{R}^n\times\mathbb{R}^n} \{\langle a,a^*\rangle_{\mathbb{R}^N} + \langle b,b^*\rangle_{\mathbb{R}^N} - L_{hom}(a,b)\} =$$

$$= \sup_{\substack{(a,b)\in\mathbb{R}^n\times\mathbb{R}^n \\ (\varphi,g)\in W^{1,p}_{\#}(Q)\times L^q_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|}\int_Q \Big[\langle a,a^*\rangle_{\mathbb{R}^N} + \langle b,b^*\rangle_{\mathbb{R}^N} - L\big(x,a+\nabla\varphi(x),b+g(x)\big)\Big]\,\mathrm{d}x =$$

$$= \sup_{\substack{(a,b)\in\mathbb{R}^n\times\mathbb{R}^n \\ (f,g)\in E_p\times E_q^{\perp}}} \frac{1}{|Q|}\int_Q \Big[\langle a,a^*\rangle_{\mathbb{R}^N} + \langle b,b^*\rangle_{\mathbb{R}^N} - L\big(x,a+f(x),b+g(x)\big)\Big]\,\mathrm{d}x.$$

Setting $A(x) = a + f(x)$, $B(x) = b + g(x)$ and substituting above we have

$$\begin{aligned} L^*_{hom}(a^*,b^*) &= \sup_{\substack{A\in E_p+\mathbb{R}^n \\ B\in E_q^{\perp}+\mathbb{R}^n}} \frac{1}{|Q|}\int_Q \Big[\langle A,a^*\rangle_{\mathbb{R}^N} + \langle B,b^*\rangle_{\mathbb{R}^N} - L\big(x,A(x),B(x)\big)\Big]\,\mathrm{d}x \\ &= \sup_{\substack{A\in L^p(\Omega;\mathbb{R}^n) \\ B\in L^q(\Omega;\mathbb{R}^n)}} \Big\{\frac{1}{|Q|}\int_Q \Big[\langle A,a^*\rangle_{\mathbb{R}^N} + \langle B,b^*\rangle_{\mathbb{R}^N} - L\big(x,A(x),B(x)\big)\Big]\,\mathrm{d}x \\ &\quad -\chi_{E_p+\mathbb{R}^n}(A) - \chi_{E_q^{\perp}+\mathbb{R}^n}(B)\Big\}. \end{aligned}$$

Now using the fact that the Fenchel dual of a sum is their *inf-convolution*, we obtain

$$\begin{aligned} L^*_{hom}(a^*,b^*) &= \inf_{\substack{f\in L^{q'}(Q;\mathbb{R}^n) \\ g\in L^{p'}(Q;\mathbb{R}^n)}} \Big\{\frac{1}{|Q|}\int_Q L^*\big(x,a^*-g(x),b^*-f(x)\big)\,\mathrm{d}x + \chi_{E^{\perp}_{p'}}(g) + \chi_{E_{q'}}(f)\Big\} \\ &= \inf_{\substack{f\in E_{q'} \\ g\in E^{\perp}_{p'}}} \frac{1}{|Q|}\int_Q L^*\big(x,a^*-g(x),b^*-f(x)\big)\,\mathrm{d}x \\ &= \inf_{\substack{\varphi\in W^{1,q'}_{\#}(Q) \\ g\in L^{p'}_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|}\int_Q L^*\big(x,a^*+g(x),b^*+\nabla\varphi(x)\big)\,\mathrm{d}x. \end{aligned}$$

This proves (7.16), which then implies that $L^{**}_{hom} = L_{hom}$ and therefore L_{hom} is convex and lower semi-continuous.

We now prove estimate (7.17). In fact, the upper bound simply follows from

$$L_{hom}(a,b) \le \frac{1}{|Q|}\int_Q L(x,a,b)\,\mathrm{d}x \le C_1(|a|^p+|b|^q+1).$$

For the lower bound, note first that since $C_0(|a|^p+|b|^q-1) \le L(x,a,b)$ for all $a,b \in \mathbb{R}^N$, it follows that

$$L^*(x,a,b) \le \frac{C_0(p-1)}{(C_0p)^{p'}}|a|^{p'} + \frac{C_0(q-1)}{(C_0q)^{q'}}|b|^{q'} + C_0 \quad \text{for all} \quad a,b \in \mathbb{R}^N.$$

One then gets from (7.16) that

$$L^*_{hom}(a,b) \le \frac{1}{|Q|}\int_Q L^*(x,a,b)\,\mathrm{d}x \le \frac{C_0(p-1)}{(C_0p)^{p'}}|a|^{p'} + \frac{C_0(q-1)}{(C_0q)^{q'}}|b|^{q'} + C_0$$

for all $a,b \in \mathbb{R}^N$, from which we get that

$$L_{hom}(a,b) = L^{**}_{hom}(a,b) \ge C_0(|a|^p+|b|^q-1)$$

for all $a,b \in \mathbb{R}^N$. □

Proof of Proposition 7.8. Setting $A(a,b) := \frac{1}{|Q|}\int_Q \partial L\big(y,a+\nabla\tilde{\varphi}(y),b+\tilde{g}(y)\big)\,\mathrm{d}y$, we shall first show that $A \subset \partial L_{hom}$. For that consider $(a_1,b_1) \in \mathbb{R}^N \times \mathbb{R}^N, \varphi \in W^{1,p}_{\#}(Q)$ and $g \in L^q_{\#}(Q;\mathbb{R}^N)$. From the convexity of L:

$$\begin{aligned} L\big(y,a_1+\nabla\varphi(y),b_1+g(y)\big) &\ge L\big(y,a+\nabla\tilde{\varphi}(y),b+\tilde{g}(y)\big) \\ + \quad &\langle \partial_1 L\big(y,a+\nabla\tilde{\varphi}(y),b+\tilde{g}(y)\big), a_1+\nabla\varphi(y)-a-\nabla\tilde{\varphi}(y)\rangle_{\mathbb{R}^N} \quad + \\ + \quad &\langle \partial_2 L\big(y,a+\nabla\tilde{\varphi}(y),b+\tilde{g}(y)\big), b_1+g(y)-b-\tilde{g}(y)\rangle_{\mathbb{R}^N}. \end{aligned}$$

Averaging the above on Q implies that

$$\frac{1}{|Q|}\int_Q L\big(y,a_1+\nabla\varphi(y),b_1+g(y)\big)\,\mathrm{d}y \ge L_{hom}(a,b) + \langle A(a,b),(a_1-a,b_1-b)\rangle_{\mathbb{R}^N\times\mathbb{R}^N},$$

from which we get

$$L_{hom}(a_1,b_1) \ge L_{hom}(a,b) + \langle A(a,b),(a_1-a,b_1-b)\rangle_{\mathbb{R}^N\times\mathbb{R}^N}.$$

This implies that $A \subset \partial L_{hom}$. To prove the reverse inclusion, let (d,c) be in $\partial L_{hom}(a,b)$. Since L_{hom} is convex and lower semi-continuous, we have

$$L_{hom}(a,b) + L^*_{hom}(d,c) = \langle a,d\rangle_{\mathbb{R}^N} + \langle b,c\rangle_{\mathbb{R}^N}.$$

It follows from Proposition 7.7 that there exist $\varphi \in W^{1,q'}_{\#}(Q)$ and $g \in L^{p'}_{\#}(Q;\mathbb{R}^N)$ such that

$$L^*_{hom}(a^*,b^*) = \frac{1}{|Q|}\int_Q L^*\big(x, a^* + g(x), b^* + \nabla\varphi(x)\big)\,\mathrm{d}x,$$

and therefore

$$\frac{1}{|Q|}\int_Q \Big(L\big(y, a+\nabla\tilde{\varphi}(y), b+\tilde{g}(y)\big) + L^*\big(x, d+g(x), c+\nabla\varphi(x)\big)\Big)\,\mathrm{d}x = \langle a,d\rangle_{\mathbb{R}^N} + \langle b,c\rangle_{\mathbb{R}^N}.$$

On the other hand,

$$\langle a,d\rangle_{\mathbb{R}^N} + \langle b,c\rangle_{\mathbb{R}^N} = \frac{1}{|Q|}\int_Q \langle a+\nabla\tilde{\varphi}(y), d+g(y)\rangle_{\mathbb{R}^N} + \langle b+\tilde{g}(y), c+\nabla\varphi(y)\rangle_{\mathbb{R}^N}\,dy,$$

which together with the previous equality yield

$$\begin{aligned}\int_Q \big[L\big(y, a+\nabla\tilde{\varphi}(y), b+\tilde{g}(y)\big) + L^*\big(y, d+g(y), c+\nabla\varphi(y)\big)\big]\,dy \quad &= \\ \int_Q \big[\langle a+\nabla\tilde{\varphi}(y), d+g(y)\rangle_{\mathbb{R}^N} - \langle b+\tilde{g}(y), c+\nabla\varphi(y)\rangle_{\mathbb{R}^N}\big]\,dy.&\end{aligned}$$

Taking into account that, pointwise, on the previous equation, the integrand above is greater then the integrand below almost everywhere, we obtain

$$\begin{aligned}L\big(y, a+\nabla\tilde{\varphi}(y), b+\tilde{g}(y)\big) + L^*\big(y, d+g(y), c+\nabla\varphi(y)\big) \quad &= \\ \langle a+\nabla\tilde{\varphi}(y), d+g(y)\rangle_{\mathbb{R}^N} - \langle b+\tilde{g}(y), c+\nabla\varphi(y)\rangle_{\mathbb{R}^N}&\end{aligned}$$

for almost all $y \in Q$. This implies that

$$(d+g(y), c+\nabla\varphi(y)) \in \partial L\big(y, a+\nabla\tilde{\varphi}(y), b+\tilde{g}(y)\big) \qquad \text{a.e. } y \in Q.$$

Integrating the above over Q implies that

$$(d,c) \in \frac{1}{|Q|}\int_Q \partial L\big(y, a+\nabla\tilde{\varphi}(y), b+\tilde{g}(y)\big),$$

which completes the proof. □

Proof of Theorem 7.6 Let $\eta \in \bar{\partial} L_{hom}(\xi)$ in such a way that $L_{hom}(\xi,\eta) = \langle \xi,\eta\rangle_{\mathbb{R}^N}$. From the definition of L_{hom}, we have

$$L_{hom}(\xi,\eta) = \min_{\substack{\varphi\in W^{1,p}_{\#}(Q)\\ g\in L^q_{\#}(Q;\mathbb{R}^N)}} \frac{1}{|Q|}\int_Q L\big(x, \xi+\nabla\varphi(x), \eta+g(x)\big)\,\mathrm{d}x.$$

From the coercivity assumptions on L, it follows that there exist $\varphi \in W^{1,p}_{\#}(Q)$ and $g \in L^q_{\#}(Q;\mathbb{R}^N)$ such that

$$L_{hom}(p,q) = \frac{1}{|Q|}\int_Q L\big(x, p+D\varphi(x), q+g(x)\big)\,\mathrm{d}x.$$

Hence

$$\begin{aligned} 0 &= L_{hom}(\xi,\eta) - \langle \xi,\eta\rangle_{\mathbb{R}^N} \\ &= \frac{1}{|Q|}\int_Q L\big(x,\xi+\nabla\varphi(x),\eta+g(x)\big)\,\mathrm{d}x - \langle \xi,\eta\rangle_{\mathbb{R}^N} \\ &= \frac{1}{|Q|}\int_Q \big[L\big(x,\xi+\nabla\varphi(x),\eta+g(x)\big) - \langle \xi+\nabla\varphi(x),\eta+g(x)\rangle_{\mathbb{R}^N}\big]\,\mathrm{d}x, \end{aligned}$$

and since the integrand in non-negative we obtain

$$L\big(x,\xi+\nabla\varphi(x),\eta+g(x)\big) - \langle \xi+\nabla\varphi(x),\eta+g(x)\rangle_{\mathbb{R}^N} = 0 \ \text{ for a.e. } x\in Q,$$

from which we have

$$\eta+g(x) \in \bar{\partial} L(x,\xi+\nabla\varphi(x)) = T(x,\xi+\nabla\varphi(x))$$

and finally $\eta = \int_Q (\eta+g(x))\,dx$. This implies that $\bar{\partial} L_{hom} \subset T_{hom}$ and the equality follows since $\bar{\partial} L_{hom}$ is itself a maximal monotone operator. □

7.3 A variational approach to homogenization

We start by studying the homogenization of a class of Lagrangians that is more general than the one introduced in Proposition 3.22. We shall then apply this result to deduce Theorem 7.1 announced in the introduction.

7.3.1 The homogenization of general Lagrangians on $W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$

The following homogenization result does not require the Ω-dependent Lagrangian L to be self dual nor that the exponents p and q to be conjugate.

Theorem 7.9 *Let Ω be a regular bounded domain and Q an open non-degenerate parallelogram in $\mathbb{R}^n$. Let $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be an Ω-dependent Lagrangian such that:*
(1) For each $a,b \in \mathbb{R}^N$ the function $x \to L(x,a,b)$ is Q-periodic.
(2) There exist constants $C_0, C_1 \geq 0$ and exponents $p,q > 1$ such that for every $x \in \Omega$,

$$C_0(|a|^p + |b|^q - 1) \leq L(x,a,b) \leq C_1(|a|^p + |b|^q + 1).$$

Let $\{G_\varepsilon; \varepsilon > 0\}$ be the family of functionals on $W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ defined by

$$G_\varepsilon(u,\tau) := \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ \mathrm{div} f = 0}} \int_\Omega L\big(\frac{x}{\varepsilon}, \nabla u(x), \tau(x) + f(x)\big)\,\mathrm{d}x,$$

and set

$$L_{hom}(a,b) := \min_{\substack{\varphi \in W^{1,p}_\#(Q) \\ g \in L^q_\#(Q;\mathbb{R}^N)}} \frac{1}{|Q|} \int_Q L\big(x, a + \nabla\varphi(x), b + g(x)\big)\,\mathrm{d}x. \tag{7.18}$$

Equip $L^q(\Omega;\mathbb{R}^N)$ with the following topology denoted by $\mathscr{T}$,

$$\tau_n \to \tau \textit{ for } \mathscr{T} \quad \text{if and only if}$$

$$\tau_n \to \tau \textit{ weakly in } L^q(\Omega;\mathbb{R}^N) \textit{ and}$$

$$\mathrm{div}(\tau_n) \to \mathrm{div}(\tau) \textit{ strongly in } W^{-1,q}(\Omega),$$

There exists then a Lagrangian G_{hom} on $W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ that is a Γ-limit of $\{G_\varepsilon; \varepsilon > 0\}$ as $\varepsilon \to 0$. Moreover, G_{hom} is given by the formula

$$G_{hom}(u,\tau) := \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ \mathrm{div} f = 0}} \int_\Omega L_{hom}\big(\nabla u(x), \tau(x) + f(x)\big)\, \mathrm{d}x, \tag{7.19}$$

Remark 7.10 Note that when the Lagrangian L is independent of the third variable, i.e.,

$$L(x,a,b) = \varphi(x,a) \quad \text{for all } (x,a,b) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^N,$$

for some function $\varphi : \Omega \times \mathbb{R}^N \to \mathbb{R}$, this homogenization problem is completely understood. Also, when the Lagrangian L is independent of the second variable then this problem can be dealt using the bi-continuity of the Fenchel dual (see for instance [Att84, DMS08]). The proof for the general Lagrangians consists of two parallel parts corresponding to each of these variables and should be done simultaneously for both. The part regarding the first variable is rather standard and the same argument can be found for instance in [Att84].

The proof of Theorem 7.9 will follow from the following two lemmas.

Lemma 7.11 *For any $(u,\tau) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$, there exists a sequence $(u_\varepsilon, \tau_\varepsilon) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ such that $u_\varepsilon \to u$ strongly in $L^p(\Omega)$, $\tau_\varepsilon \to \tau$ strongly in $L^q(\Omega;\mathbb{R}^N)$ and*

$$\limsup_{\varepsilon \to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \le G_{hom}(u,\tau).$$

Lemma 7.12 *Let $f \in L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div}(f) = 0$. For any $(u,\tau) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ and any sequence $(u_\varepsilon, \tau_\varepsilon)$ such that $u_\varepsilon \to u$ strongly in $L^p(\Omega)$ and $\tau_\varepsilon \to \tau$ with the $\mathscr{T}$-topology in $L^q(\Omega;\mathbb{R}^N)$, we have*

$$\liminf_{\varepsilon \to 0} \int_\Omega L\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x), \tau_\varepsilon(x) + f(x)\Big)\, \mathrm{d}x \ge \int_\Omega L_{hom}\big(\nabla u(x), \tau(x) + f(x)\big)\, \mathrm{d}x.$$

We first show how Theorem 7.9 follows from the two lemmas above. The limsup property in the definition of Γ-convergence readily follows from Lemma 7.11. For the liminf property we must show that for any $(u,\tau) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ and any sequence $\{(u_\varepsilon,\tau_\varepsilon)\} \subset W^{1,p}(\Omega)\times L^q(\Omega;\mathbb{R}^N)$ such that

$$u_\varepsilon \to u \text{ strongly in } L^p(\Omega) \quad \text{and} \quad \tau_\varepsilon \to \tau \quad \text{in the } \mathscr{T}-\text{topology},$$

we have that

$$\liminf_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) \geq G_{hom}(u,\tau).$$

By Lemma 7.12 we have

$$\liminf_{\varepsilon\to 0}\int_\Omega L(\frac{x}{\varepsilon},\nabla u_\varepsilon,\tau_\varepsilon+f)\,\mathrm{d}x \geq \int_\Omega L_{hom}\big(\nabla u,\tau+f\big)\,\mathrm{d}x,$$

for every $f\in L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div}(f)=0$. Since

$$\inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \mathrm{div} f=0}} \liminf_{\varepsilon\to 0}\int_\Omega L(\frac{x}{\varepsilon},\nabla u_\varepsilon,\tau_\varepsilon+f)\,\mathrm{d}x = \liminf_{\varepsilon\to 0} \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \mathrm{div} f=0}}\int_\Omega L(\frac{x}{\varepsilon},\nabla u_\varepsilon,\tau_\varepsilon+f)\,\mathrm{d}x,$$

we obtain that $\liminf_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon)\geq G_{hom}(u,\tau)$, as desired. □

Proof of Lemma 7.11. Note that without loss of generality we may assume $L\geq 0$. Assume first that u is an affine function and τ is constant on Ω, that is

$$u(x) = \langle a,x\rangle + \alpha \text{ and } \tau(x) = b,$$

for some a and b in $\mathbb{R}^n$ and $\alpha\in\mathbb{R}$. Fix $\eta\in\mathbb{R}^n$ and let $\tilde{\varphi}$ and $\tilde{g}$ to be the minimizers on the formula for L_{hom} given by (7.18):

$$L_{hom}(a,b+\eta) = \frac{1}{|Q|}\int_Q L\big(x,a+\nabla\tilde{\varphi}(x),b+\eta+\tilde{g}(x)\big). \tag{7.20}$$

Define

$$u_\varepsilon(x) := u(x)+\varepsilon\tilde{\varphi}(\tfrac{x}{\varepsilon}) \quad \text{and} \quad \tau_\varepsilon(x) := \tau.$$

Note that by Lemma 7.5 in section 7.1, $\tilde{g}$ can be extended by periodicity to an element of $L^q_{loc}(\mathbb{R}^N;\mathbb{R}^N)$, still denoted by $\tilde{g}$ such that $\operatorname{div}(\tilde{g}) = 0$. It follows that

$$\begin{aligned}
\limsup_{\varepsilon} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) &= \limsup_{\varepsilon} \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L\big(\frac{x}{\varepsilon}, a+\nabla\tilde{\varphi}(\frac{x}{\varepsilon}), b+f(x)\big)\,\mathrm{d}x \\
&\leq \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \limsup_{\varepsilon} \int_\Omega L\big(\frac{x}{\varepsilon}, a+\nabla\tilde{\varphi}(\frac{x}{\varepsilon}), b+f(x)\big)\,\mathrm{d}x \\
&\leq \limsup_{\varepsilon} \int_\Omega L\big(\frac{x}{\varepsilon}, a+\nabla\tilde{\varphi}(\frac{x}{\varepsilon}), b+\eta+\tilde{g}(\frac{x}{\varepsilon})\big)\,\mathrm{d}x.
\end{aligned}$$

By Lemma 7.3 of section 7.1 we have as $\varepsilon \to 0$,

$$\int_\Omega L\big(\frac{x}{\varepsilon}, a+\nabla\tilde{\varphi}(\frac{x}{\varepsilon}), b+\eta+\tilde{g}(\frac{x}{\varepsilon})\big)\,\mathrm{d}x \to \frac{|\Omega|}{|Q|}\int_Q L\big(y, a+\nabla\tilde{\varphi}(y), b+\eta+\tilde{g}(y)\big)\,\mathrm{d}y.$$

It then follows from (7.20) that

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \leq |\Omega| L_{hom}(a, b+\eta),$$

and since η is arbitrary, we have that

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \leq \inf_{\eta\in\mathbb{R}^n} |\Omega| L_{hom}(a, b+\eta).$$

By Lemma 7.2 we have

$$\inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L_{hom}\big(a, b+f(x)\big)\,\mathrm{d}x \geq \inf_{\eta\in\mathbb{R}^n} |\Omega| L_{hom}(a, b+\eta),$$

and thus we conclude, as desired

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \leq \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L_{hom}\big(a, b+f(x)\big)\,\mathrm{d}x = G_{hom}(u,\tau).$$

Assume now that u is a piecewise affine function and τ is a piecewise constant function on Ω, that is for $\{\hat{\Omega}_j\}_{j\in I_1}$ and $\{\tilde{\Omega}_k\}_{k\in I_2}$, both finite polyhedral partitions of Ω, we have

$$u(x) = \langle a_j, x\rangle + \alpha_j \text{ for } x \in \hat{\Omega}_j \quad \text{and} \quad \tau(x) = b_k \text{ for } x \in \tilde{\Omega}_k,$$

for fixed $a_j \in \mathbb{R}^n$ and $b_k \in \mathbb{R}^n$ and constants α_j. By considering non-empty intersections $\hat{\Omega}_j \cap \tilde{\Omega}_k$ and re-indexing them, we can consider $\{\Omega_i\}_{i\in I}$ a polyhedral partition of Ω such that

$$u(x) = \langle a_i, x\rangle + \alpha_i \text{ for } x \in \Omega_i \quad \text{and} \quad \tau(x) = b_i \text{ for } x \in \Omega_i.$$

Analogous to what was done previously, fix $\{\eta_i\} \subset \mathbb{R}^N$ and let $\tilde{\varphi}_i$ and $\tilde{g}_i$ be such that

$$L_{hom}(a_i, b_i + \eta_i) = \frac{1}{|Q|}\int_Q L\big(x, a_i + \nabla\tilde{\varphi}_i(x), b_i + \eta_i + \tilde{g}_i(x)\big)\,\mathrm{d}x,$$

and set $u^i_\varepsilon(x) := u(x) + \varepsilon\tilde{\varphi}_i(\frac{x}{\varepsilon})$.

Unfortunately, we cannot consider u_ε as the piecewise construction of the above functions, as the φ_i won't necessarily match along the interface between the Ω_i and thus will not in general be a function in $W^{1,p}(\Omega)$. This can be remedied by the following standard construction (see for instance [Att84]): Let Σ be the interface set between the Ω_i, and define for $\delta > 0$, $\Sigma_\delta := \{x \in \Omega \,:\, d(x,\Sigma) \le \delta\}$. Consider a smooth function Ψ_δ so that

$$\Psi_\delta(x) = \begin{cases} 1 & x \in \Sigma_\delta \\ 0 & x \in \Omega \setminus \Sigma_{2\delta}, \end{cases}$$

and define

$$u^\delta_\varepsilon(x) := \big(1 - \Psi_\delta(x)\big)u^i_\varepsilon(x) + \Psi_\delta(x)u(x) \ \text{ for } x \in \Omega_i \quad \text{and} \quad \tau_\varepsilon := \tau.$$

It can be checked that the function u^δ_ε lies in $W^{1,p}(\Omega)$. Note that by Lemma 7.5, each $\tilde{g}_i$ can be extended by periodicity to an element of $L^q_{loc}(\mathbb{R}^N;\mathbb{R}^N)$, still denoted by $\tilde{g}_i$ such that $\operatorname{div}(\tilde{g}_i) = 0$. Thus $\operatorname{div}(\eta_i + \tilde{g}_i(\frac{x}{\varepsilon})) = 0$ on $\mathbb{R}^N$ and in particular on $\Omega_i \setminus \Sigma_\delta$. Define $f_{\varepsilon,\delta}(x) = \eta_i + \tilde{g}_i(\frac{x}{\varepsilon})$ on $\Omega_i \setminus \Sigma_\delta$. One can also extend (using Theorem 2.5 and Corollary 2.8 in [VG86]) $f_{\varepsilon,\delta}$ to an element in $L^q(\Omega;\mathbb{R}^N)$, still denoted by $f_{\varepsilon,\delta}$ such that $\|f_{\varepsilon,\delta}\|_{L^q(\Omega;\mathbb{R}^N)}$ is bounded and $\operatorname{div}(f_{\varepsilon,\delta}) = 0$. Take now any $0 < t < 1$, then

$$
\begin{aligned}
G_\varepsilon(tu_\varepsilon^\delta,\tau_\varepsilon) &= \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L\Big(\nabla tu_\varepsilon^\delta,\tau_\varepsilon+f\Big)\,\mathrm{d}x \\
&\le \int_\Omega L\Big(\nabla tu_\varepsilon^\delta,\tau_\varepsilon+f_{\varepsilon,\delta}\Big)\,\mathrm{d}x =
\end{aligned}
$$

$$
\sum_i \int_{\Omega_i\setminus\Sigma_\delta} L\Big(\frac{x}{\varepsilon}, t\big(1-\Psi_\delta\big)\nabla u_\varepsilon^i + t\Psi_\delta\nabla u + \frac{t(1-t)}{(1-t)}(u-u_\varepsilon^i)\nabla\Psi_\delta, b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x +
$$

$$
+\int_{\Sigma_\delta} L\Big(\nabla tu_\varepsilon^\delta,\tau_\varepsilon+f_{\varepsilon,\delta}\Big)\,\mathrm{d}x
$$

Since L is convex in the middle variable and since $t(1-\Psi_\delta)+t\Psi_\delta+(1-t)=1$, we obtain

$$
\begin{aligned}
G_\varepsilon(tu_\varepsilon^\delta,\tau_\varepsilon) &\le \sum_i \int_{\Omega_i\setminus\Sigma_\delta} t(1-\Psi_\delta) L\Big(\frac{x}{\varepsilon}, a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x \\
&+ \sum_i \int_{\Omega_i\setminus\Sigma_\delta} (1-t) L\Big(\frac{x}{\varepsilon}, \frac{t}{(1-t)}(u-u_\varepsilon^i)\nabla\Psi_\delta, b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon}\Big)\,\mathrm{d}x \\
&+ \int_{\Sigma_{2\delta}\setminus\Sigma_\delta} t\Psi_\delta L\Big(\frac{x}{\varepsilon}, \nabla u, b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon}\Big)\,\mathrm{d}x \\
&+ \int_{\Sigma_\delta} L\Big(\nabla tu_\varepsilon^\delta,\tau_\varepsilon+f_{\varepsilon,\delta}\Big)\,\mathrm{d}x.
\end{aligned}
$$

For the first term on the right hand side of this inequality we have

$$
\int_{\Omega_i\setminus\Sigma_\delta} t(1-\Psi_\delta) L\Big(\frac{x}{\varepsilon}, a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x \le
$$

$$
\int_{\Omega_i\setminus\Sigma_\delta} L\Big(\frac{x}{\varepsilon}, a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x.
$$

Using the boundedness of L we get the following estimate for the second term,

$$
\int_{\Omega_i\setminus\Sigma_\delta} (1-t) L\Big(\frac{x}{\varepsilon}, \frac{t}{(1-t)}(u-u_\varepsilon^i)\nabla\Psi_\delta, b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x \le
$$

$$
C_1(1-t)\int_{\Omega_i\setminus\Sigma_\delta} \Big(|\frac{t}{(1-t)}(u-u_\varepsilon^i)\nabla\Psi_\delta|^p + |b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})|^q + 1\Big)\,\mathrm{d}x,
$$

and similarly

$$\int_{\Sigma_{2\delta}\setminus\Sigma_\delta} t\Psi_\delta L\Big(\frac{x}{\varepsilon},\nabla u, b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x \le C_1\int_{\Sigma_{2\delta}\setminus\Sigma_\delta}\Big(1+|\nabla u|^p+|b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})|^q\Big)\,\mathrm{d}x,$$

as well as

$$\int_{\Sigma_\delta} L\Big(\nabla t u_\varepsilon^\delta, \tau_\varepsilon+f_{\varepsilon,\delta}\Big)\,\mathrm{d}x \le C_1\int_{\Sigma_\delta}\Big(1+|\nabla t u_\varepsilon^\delta|^p+|\tau_\varepsilon+f_{\varepsilon,\delta}|^q\Big)\,\mathrm{d}x.$$

It then follows that

$$\begin{aligned}
G_\varepsilon(tu_\varepsilon^\delta,\tau_\varepsilon) \quad \le \quad & \sum_i \int_{\Omega_i\setminus\Sigma_\delta} L\Big(\frac{x}{\varepsilon}, a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\Big)\,\mathrm{d}x \\
+ \quad & C_1(1-t)\sum_i \int_{\Omega_i\setminus\Sigma_\delta}\Big(|\frac{t(u-u_\varepsilon^i)}{(1-t)}\nabla\Psi_\delta|^p+|b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})|^q+1\Big)\,\mathrm{d}x \\
+ \quad & C_1\int_{\Sigma_{2\delta}\setminus\Sigma_\delta}\Big(1+|\nabla u|^p+|b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})|^q\Big)\,\mathrm{d}x \\
+ \quad & C_1\int_{\Sigma_\delta}\Big(1+|\nabla t u_\varepsilon^\delta|^p+|\tau_\varepsilon+f_{\varepsilon,\delta}|^q\Big)\,\mathrm{d}x.
\end{aligned}$$

By taking $\limsup_{\varepsilon\to 0}$ on both sides and considering $u_\varepsilon^i \to u$ on $L^p(\Omega_i)$, and then letting $t\to 1$ and $\delta\to 0$ we finally get,

$$\limsup_{\substack{t\to 1\\ \delta\to 0}}\limsup_{\varepsilon\to 0} G_\varepsilon(tu_\varepsilon^\delta,\tau_\varepsilon) \le \sum_i \frac{|\Omega_i|}{|Q|}\int_Q L\big(x, a_i+\nabla\tilde{\varphi}_i(x), b_i+\eta_i+\tilde{g}_i(x)\big)\,\mathrm{d}x. \tag{7.21}$$

Also note that,

$$\sum_i \frac{|\Omega_i|}{|Q|}\int_Q L\big(x, a_i+\nabla\tilde{\varphi}_i(x), b_i+\eta_i+\tilde{g}_i(x)\big)\,\mathrm{d}x = \sum_i |\Omega_i| L_{hom}(a_i, b_i+\eta_i).$$

A diagonalization argument yields from limit (7.21) the existence of some $t(\varepsilon)$ and $\delta(\varepsilon)$ such that $t(\varepsilon)\to 1$ and $\delta(\varepsilon)\to 0$ as $\varepsilon\to 0$. Defining $u_\varepsilon := t(\varepsilon)u_\varepsilon^{\delta(\varepsilon)}$, we obtain

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) \le \sum_i |\Omega_i| L_{hom}(a_i, b_i+\eta_i),$$

and since the $\{\eta_i\}$ is arbitrary one has

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) \le \sum_i |\Omega_i| \inf_{\eta_i\in\mathbb{R}^n} L_{hom}(a_i, b_i+\eta_i).$$

Now we use Lemma 7.2 to obtain

$$\sum_i |\Omega_i| \inf_{\eta_i\in\mathbb{R}^n} L_{hom}(a_i, b_i+\eta_i) \le \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L_{hom}\big(\nabla u(x), \tau(x)+f(x)\big)\,\mathrm{d}x,$$

from which we get $\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) \le G_{hom}(u,\tau)$.

Finally, consider any $(u,\tau)\in W^{1,p}(\Omega)\times L^q(\Omega;\mathbb{R}^N)$. There exists then a sequence $\{u_n\}$ of piecewise affine functions and a sequence $\{\tau_n\}$ of piecewise constant functions such that $(u_n,\tau_n)\to(u,\tau)$. By Proposition 7.7, the function G_{hom} are continuous, so we also have

$$\lim_n G_{hom}(u_n,\tau_n) = G_{hom}(u,\tau).$$

For each n, we have shown the existence of $(u_n^\varepsilon,\tau_n^\varepsilon)$ such that $u_n^\varepsilon\to u_n$ and $\tau_n^\varepsilon\to\tau_n$ in $L^p(\Omega)$ and $L^q(\Omega;\mathbb{R}^N)$ respectively and

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_n^\varepsilon,\tau_n^\varepsilon) \le G_{hom}(u_n,\tau_n),$$

so we get

$$\limsup_n \limsup_{\varepsilon\to 0} G_\varepsilon(u_n^\varepsilon,\tau_n^\varepsilon) \le G_{hom}(u,\tau).$$

From the same diagonalization argument as before, there exists some $n(\varepsilon)$ such that $n(\varepsilon)\to\infty$ as $\varepsilon\to 0$ for which, by defining $(u_\varepsilon,\tau_\varepsilon) := (u^\varepsilon_{n(\varepsilon)},\tau^\varepsilon_{n(\varepsilon)})$ we obtain

$$u_\varepsilon \to u \text{ strongly in } L^p(\Omega),\ \tau_\varepsilon\to\tau \text{ strongly in } L^q(\Omega;\mathbb{R}^N)$$

and

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_n^\varepsilon,\tau_n^\varepsilon) \le G_{hom}(u,\tau).$$

This concludes the proof of Lemma 7.11. □

Proof of Lemma 7.12. Let $(u,\tau) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ and $f \in L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div}(f) = 0$. We assume that $u_\varepsilon \to u$ strongly in $L^p(\Omega)$ and $\tau_\varepsilon \to \tau_\varepsilon$ in $\mathscr{T}$. For constant vectors $a_i, b_i, \eta_i \in \mathbb{R}^n$, consider as before functions $\tilde{\varphi}_i \in W^{1,p}_\#(Q)$ and $\tilde{g}_i \in L^q_\#(Q;\mathbb{R}^N)$ such that

$$L_{hom}(a_i, b_i + \eta_i) = \frac{1}{|Q|} \int_Q L\big(x, a_i + \nabla\tilde{\varphi}_i(x), b_i + \eta_i + \tilde{g}_i(x)\big)\,\mathrm{d}x.$$

Denote $\partial_1 L$ the subdifferential of L with respect to the middle variable and $\partial_2 L$ the subdifferential of L with respect to the last variable. From the above we have both

$$\mathrm{div}\Big(\partial_1 L\big(y, a_i + D\tilde{\varphi}_i(y), b_i + \eta_i + \tilde{g}_i(y)\big)\Big) = 0 \; a.e. \; y \in Q, \tag{7.22}$$

and

$$\int_Q \langle \partial_2 L\big(y, a_i + D\tilde{\varphi}_i(y), b_i + \eta_i + \tilde{g}_i(y)\big), g(y)\rangle \,\mathrm{d}y = 0, \tag{7.23}$$

for any $g \in L^q_\#(Q;\mathbb{R}^N)$. It follows from (7.23) that

$$\partial_2 L\big(y, a_i + \nabla\tilde{\varphi}_i(y), b_i + \eta_i + \tilde{g}_i(y)\big) = \nabla w(y) \qquad \text{a.e. } y \in Q, \tag{7.24}$$

for some $w \in W^{1,p}_\#(Q)$. It also follows from Lemma 7.4 that w can be extended by periodicity to an element in $W^{1,p}_{loc}(\mathbb{R}^N)$. Now, let $\hat{u} \in W^{1,p}(\Omega)$ be a piecewise affine functions and $\hat{\tau} \in L^q(\Omega;\mathbb{R}^N)$ be a piecewise constant function such that for some partition $\{\Omega_i\}$ of Ω we have

$$\hat{u}(x) = \langle a_i, x\rangle + \alpha_i \text{ for } x \in \Omega_i \text{ and } \hat{\tau}(x) = b_i \text{ for } x \in \Omega_i.$$

Consider now for $x \in \Omega_i$,

$$\hat{u}_\varepsilon(x) := \hat{u}(x) + \varepsilon\tilde{\varphi}_i(\tfrac{x}{\varepsilon}) \text{ and } \hat{\tau}_\varepsilon(x) := \hat{\tau}(x).$$

From the convexity of L we get

$$
\begin{aligned}
L\big(\frac{x}{\varepsilon},\nabla u_\varepsilon(x),\tau_\varepsilon(x)+f(x)\big) \geq L\big(\frac{x}{\varepsilon},\nabla \hat{u}_\varepsilon(x),\hat{\tau}_\varepsilon(x)+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big)\\
+\langle \partial_1 L\big(\frac{x}{\varepsilon},\nabla \hat{u}_\varepsilon(x),\hat{\tau}_\varepsilon(x)+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),\nabla u_\varepsilon(x)-\nabla \hat{u}_\varepsilon(x)\rangle\\
+\langle \partial_2 L\big(\frac{x}{\varepsilon},\nabla \hat{u}_\varepsilon(x),\hat{\tau}_\varepsilon(x)+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),\tau_\varepsilon(x)-\hat{\tau}_\varepsilon(x)\rangle\\
+\langle \partial_2 L\big(\frac{x}{\varepsilon},\nabla \hat{u}_\varepsilon(x),\hat{\tau}_\varepsilon(x)+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),f(x)-\eta_i-\tilde{g}_i(\frac{x}{\varepsilon})\rangle.
\end{aligned}
$$

Consider now smooth functions $\Psi_i:\Omega_i\to\mathbb{R}$ with compact support such that $0<\Psi_i<1$. Multiplying the above convexity inequality by Ψ_i, integrating over Ω_i and adding over all i, we get the following:

$$
\begin{aligned}
\int_\Omega L(\frac{x}{\varepsilon},\nabla u_\varepsilon,\tau_\varepsilon+f)\,\mathrm{d}x \geq \sum_i\int_{\Omega_i} L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big)\Psi_i(x)\,\mathrm{d}x\\
+\sum_i\int_{\Omega_i}\langle \partial_1 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),\nabla u_\varepsilon(x)-\nabla\hat{u}_\varepsilon(x)\rangle\Psi_i(x)\,\mathrm{d}x\\
+\sum_i\int_{\Omega_i}\langle \partial_2 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),\tau_\varepsilon(x)-\hat{\tau}_\varepsilon(x)\rangle\Psi_i(x)\,\mathrm{d}x\\
+\sum_i\int_{\Omega_i}\langle \partial_2 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),f(x)-\eta_i\rangle\Psi_i(x)\,\mathrm{d}x\\
+\sum_i\int_{\Omega_i}\langle \partial_2 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),-\tilde{g}_i(\frac{x}{\varepsilon})\rangle\Psi_i(x)\,\mathrm{d}x.
\end{aligned}
$$

Now we deal with each term independently. For the first term on the right hand side of the above expression we have

$$
\int_{\Omega_i} L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big)\Psi_i(x)\,\mathrm{d}x \to \int_{\Omega_i} L_{hom}(a_i,b_i+\eta_i)\Psi_i(x)\,\mathrm{d}x,
$$

by virtue of Lemma 7.3.

For the second term, by integrating by parts and by then taking into account (7.22) we obtain

$$
\begin{aligned}
&\int_{\Omega_i}\langle \partial_1 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),\nabla u_\varepsilon(x)-\nabla\hat{u}_\varepsilon(x)\rangle\Psi_i(x)\,\mathrm{d}x\\
&=-\int_{\Omega_i}\langle \partial_1 L\big(\frac{x}{\varepsilon},a_i+\nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}),b_i+\eta_i+\tilde{g}_i(\frac{x}{\varepsilon})\big),(u_\varepsilon-\hat{u}_\varepsilon)\nabla\Psi_i(x)\rangle\,\mathrm{d}x.
\end{aligned}
$$

It follows from Lemma 7.3 and Proposition 7.8 below, that if $\varepsilon \to 0$ then,

$$\int_{\Omega_i} \langle \partial_1 L\big(\frac{x}{\varepsilon}, a_i + \nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i + \eta_i + \tilde{g}_i(\frac{x}{\varepsilon})\big), (u_\varepsilon - \hat{u}_\varepsilon)\nabla\Psi_i(x)\rangle \,\mathrm{d}x \to \int_{\Omega_i} \langle \partial_1 L_{hom}(a_i, b_i + \eta_i), (u - \hat{u})\nabla\Psi_i(x)\rangle \,\mathrm{d}x$$

Integrate by parts on more time to get

$$\int_{\Omega_i} \langle \partial_1 L_{hom}(a_i, b_i + \eta_i), (u - \hat{u})\nabla\Psi_i(x)\rangle \,\mathrm{d}x = -\int_{\Omega_i} \langle \partial_1 L_{hom}(a_i, b_i + \eta_i), \nabla u - \nabla\hat{u}\rangle \Psi_i(x) \,\mathrm{d}x,$$

from which one has

$$\int_{\Omega_i} \langle \partial_1 L\big(\frac{x}{\varepsilon}, a_i + \nabla\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i + \eta_i + \tilde{g}_i(\frac{x}{\varepsilon})\big), \nabla u_\varepsilon(x) - \nabla\hat{u}_\varepsilon(x)\rangle \Psi_i(x) \,\mathrm{d}x \to \int_{\Omega_i} \langle \partial_1 L_{hom}(a_i, b_i + \eta_i), \nabla u - \nabla\hat{u}\rangle \Psi_i(x) \,\mathrm{d}x.$$

For the third term, we use (7.24) to get $\partial_2 L\big(\frac{x}{\varepsilon}, a_i + D\tilde{\varphi}_i(\frac{x}{\varepsilon}), b_i + \eta_i + \tilde{g}_i(\frac{x}{\varepsilon})\big) = \nabla w(\frac{x}{\varepsilon})$ for some $w \in W^{1,p}_{\#}(Q)$. Using an integration by parts, we obtain

$$\begin{aligned}\int_{\Omega_i} \langle \nabla w(\frac{x}{\varepsilon}), \tau_\varepsilon(x) - \hat{\tau}_\varepsilon(x)\rangle \Psi_i(x) \,\mathrm{d}x &= -\int_{\Omega_i} \varepsilon w(\frac{x}{\varepsilon}) \mathrm{div}\big(\tau_\varepsilon(x) - \hat{\tau}_\varepsilon(x)\big)\rangle \Psi_i(x) \,\mathrm{d}x \\ &\quad -\int_{\Omega_i} \varepsilon w(\frac{x}{\varepsilon}) \langle \nabla\Psi_i(x), \tau_\varepsilon(x) - \hat{\tau}_\varepsilon(x)\rangle \,\mathrm{d}x,\end{aligned}$$

which goes to 0 as $\varepsilon \to 0$ since $\tau_\varepsilon \to \tau$ in the $\mathscr{T}$-topology.

Similarly as above, the fourth term can be seen to converge to

$$\int_{\Omega_i} \langle \partial_2 L_{hom}(a_i, b_i + \eta_i), f(x) - \eta_i\rangle \Psi_i(x) \,\mathrm{d}x,$$

while for the fifth term, we first observe that the function

$$m_i(x) := \langle \partial_2 L\big(x, a_i + \nabla\tilde{\varphi}_i(x), b_i + \eta_i + \tilde{g}_i(x)\big), \tilde{g}_i(x)\rangle$$

is Q-periodic, and thus setting $(m_i)_\varepsilon(x) := m_i(\frac{x}{\varepsilon})$, it follows from Lemma 7.3 that $(m_i)_\varepsilon \rightharpoonup \overline{m}_i$ weakly in L^1, where

$$\overline{m}_i = \frac{1}{|Q|}\int_Q \langle \partial_2 L\big(y, a_i + \nabla\tilde{\varphi}_i(y), b_i + \eta_i + \tilde{g}_i(y)\big), -\tilde{g}_i(y)\rangle \,\mathrm{d}y,$$

which is equal to 0 in view of (7.23). The fifth term therefore disappears as $\varepsilon \to 0$.

Putting now all of the above together we obtain that

$$\begin{aligned}\liminf_{\varepsilon\to 0}\int_\Omega L(\frac{x}{\varepsilon}, \nabla u_\varepsilon, \tau_\varepsilon + f)\,\mathrm{d}x \quad &\geq \sum_i \int_{\Omega_i} L_{hom}(a_i, b_i + \eta_i)\Psi_i(x)\,\mathrm{d}x \\ &+ \sum_i \int_{\Omega_i} \langle \partial_1 L_{hom}(a_i, b_i + \eta_i), \nabla u(x) - \nabla\hat{u}(x)\rangle \Psi_i(x)\,\mathrm{d}x \\ &+ \sum_i \int_{\Omega_i} \langle \partial_2 L_{hom}(a_i, b_i + \eta_i), f(x) - \eta_i\rangle \Psi_i(x)\,\mathrm{d}x.\end{aligned}$$

By taking into account the estimate

$$|\partial L_{hom}(a,b)| \leq M(1 + |a|^{p-1} + |b|^{q-1}) \qquad \text{for all } a, b \in \mathbb{R}^N,$$

which follows from estimate (7.17) in Proposition 7.7, and letting $\Psi_i \uparrow 1$ on each Ω_i, it follows from the dominated convergence theorem that

$$\begin{aligned}\liminf_{\varepsilon\to 0}\int_\Omega L(\frac{x}{\varepsilon}, \nabla u_\varepsilon, \tau_\varepsilon + f)\,\mathrm{d}x \quad &\geq \int_\Omega L_{hom}\big(\nabla\hat{u}(x), \hat{\tau}(x) + \tilde{f}(x)\big)\,\mathrm{d}x \\ &+ \int_\Omega \langle \partial_1 L_{hom}\big(\nabla\hat{u}(x), \hat{\tau}(x) + \tilde{f}(x)\big), \nabla u(x) - \nabla\hat{u}(x)\rangle\,\mathrm{d}x \\ &+ \int_\Omega \langle \partial_2 L_{hom}\big(\nabla\hat{u}(x), \hat{\tau}(x) + \tilde{f}(x)\big), f(x) - \tilde{f}(x)\rangle\,\mathrm{d}x.\end{aligned}$$

where $\tilde{f} \in L^q(\Omega;\mathbb{R}^N)$ is a function defined by $\tilde{f}(x) = \eta_i$ on Ω_i. The above is valid for arbitrary piecewise affine function $\hat{u}$, and piecewise constant functions $\hat{\tau}, \tilde{f}$. We can then let $\hat{u} \to u$ in $W^{1,p}(\Omega)$ and $\hat{\tau} \to \tau$ and $\tilde{f} \to f$ in $L^q(\Omega;\mathbb{R}^N)$ to obtain

$$\liminf_{\varepsilon\to 0}\int_\Omega L(\frac{x}{\varepsilon}, \nabla u_\varepsilon, \tau_\varepsilon + f)\,\mathrm{d}x \geq \int_\Omega L_{hom}\big(\nabla u(x), \tau(x) + f(x)\big)\,\mathrm{d}x.$$

This completes the proof. □

Before proceeding to the next subsection, we note the following slight extension of Lemma 7.11, which will be needed for Proposition 7.15 below. We note that the proof is known when G_ε is independent of the second variable, and here we show that the same proof with minor modification works for general Lagrangians just as in Theorem 7.9.

Lemma 7.13 *Let G_ε and G_{hom} be as in Theorem 7.9. Then, for any $(u,\tau) \in W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$, there exist a sequence $(u_\varepsilon,\tau_\varepsilon)$ such that $u - u_\varepsilon \rightharpoonup 0$ weakly in $W^{1,p}(\Omega)$ and $\tau_\varepsilon \to \tau$ in the $\mathscr{T}$-topology. Furthermore, $u - u_\varepsilon \in W_0^{1,p}(\Omega)$ and for this sequence:*

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) \leq G(u,\tau).$$

Proof. From Theorem 7.9, there exist a sequence $(\tilde{u}_\varepsilon,\tau_\varepsilon)$ with $\tilde{u}_\varepsilon \to u$ in $L^p(\Omega)$ and $\tau_\varepsilon \to \tau$ in the $\mathscr{T}$-topology, such that

$$G_{hom}(u,\tau) = \lim_{\varepsilon\to 0} G_\varepsilon(\tilde{u}_\varepsilon,\tau_\varepsilon).$$

Up to a subsequence one may assume that

$$\tilde{u}_\varepsilon \rightharpoonup u \text{ weakly in } W^{1,p}(\Omega).$$

Pick any $\varphi \in W_0^{1,p}(\Omega)$ with $\varphi > 0$ in Ω. Define:

$$u_\varepsilon(x) := \begin{cases} \tilde{u}_\varepsilon & u(x)-\varphi(x) \leq \tilde{u}_\varepsilon \leq u(x)+\varphi(x) \\ u(x)-\varphi(x) & \tilde{u}_\varepsilon(x) < u(x)-\varphi(x) \\ u(x)+\varphi(x) & u(x)+\varphi(x) < \tilde{u}_\varepsilon(x) \end{cases}.$$

Note that $u_\varepsilon - u \in W_0^{1,p}(\Omega)$ and since $\tilde{u}_\varepsilon \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$, so $u_\varepsilon \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$. Note that $L + C_0 \geq 0$. For any $f \in L^q(\Omega;\mathbb{R}^N)$ with $\operatorname{div}(f) = 0$ we have

$$\begin{aligned} G_\varepsilon(u_\varepsilon,\tau_\varepsilon) + C_0|\Omega| \;\; \leq \;\; & \int_{\{u_\varepsilon\neq\tilde{u}_\varepsilon\}} \Big[L\big(\frac{x}{\varepsilon},\nabla u_\varepsilon(x), f(x)+\tau_\varepsilon(x)\big)+C_0\Big]\,dx \\ & + \int_{\{u_\varepsilon=\tilde{u}_\varepsilon\}} \Big[L\big(\frac{x}{\varepsilon},\nabla u_\varepsilon(x), f(x)+\tau_\varepsilon(x)\big)+C_0\Big]\,dx. \end{aligned}$$

For x in the set $\{u_\varepsilon \neq \tilde{u}_\varepsilon\}$, the norm of $\nabla \tilde{u}_\varepsilon(x)$ is controlled by the norm of $|\nabla u(x)| + |\nabla \varphi(x)|$. It follows that

$$G_\varepsilon(u_\varepsilon, \tau_\varepsilon) + C_0|\Omega| \leq$$

$$\int_{\{u_\varepsilon \neq \tilde{u}_\varepsilon\}} \Big[C_1\Big((|\nabla u(x)| + |\nabla \varphi(x)|)^p + |\tau_\varepsilon(x) + f(x)|^q + 1\Big) + C_0\Big]\, dx$$
$$+ \int_\Omega \Big[L\big(\frac{x}{\varepsilon}, \nabla \tilde{u}_\varepsilon(x), f(x) + \tau_\varepsilon(x)\big) + C_0\Big]\, dx.$$

Take now the infimum over all $f \in L^q(\Omega; \mathbb{R}^N)$ with $\mathrm{div}(f) = 0$ and subtract the latter by $C_0|\Omega|$. Since $|\{u_\varepsilon \neq \tilde{u}_\varepsilon\}| \to 0$ and $G_{hom}(u, \tau) = \lim_\varepsilon G_\varepsilon(\tilde{u}_\varepsilon, \tau_\varepsilon)$, we obtain

$$\limsup_{\varepsilon \to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \leq G(u, \tau).$$

7.3.2 Variational homogenization of maximal monotone operators on $W_0^{1,p}(\Omega)$

In this section we establish a homogenization result for selfdual Lagrangians on $W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$ where $\frac{1}{p} + \frac{1}{q} = 1$ and then proceed to prove Theorem 7.1.

Theorem 7.14 *Let Ω be a regular bounded domain, Q be an open non-degenerate parallelogram in $\mathbb{R}^n$, and $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be an Ω-dependent selfdual Lagrangian such that:*
(1) For each $a, b \in \mathbb{R}^N$ the function $x \to L(x, a, b)$ is Q-periodic,
(2) For some constants $C_0, C_1 \geq 0$, we have for every $x \in \mathbb{R}^N$,

$$C_0(|a|^p + |b|^q) \leq L(x, a, b) \leq C_1(|a|^p + |b|^q + 1), \tag{7.25}$$

where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $u_n^* \to u^*$ strongly in $W^{-1,q}(\Omega)$ and let u_n be solutions and τ_n be momenta for the Dirichlet boundary value problems

$$\begin{cases} \tau_n(x) \in \bar{\partial} L(\frac{x}{\varepsilon_n}, \nabla u_n(x)) & a.e. \quad x \in \Omega \\ -\mathrm{div}(\tau_n(x)) = u_n^*(x) & x \in \Omega \\ u_n \in W_0^{1,p}(\Omega). \end{cases} \tag{7.26}$$

Then, up to a subsequence,

$$u_n \to u \text{ weakly in } W_0^{1,p}(\Omega) \text{ and } \tau_n \to \tau \text{ weakly in } L^q(\Omega;\mathbb{R}^N),$$

where u is a solution and τ is a momentum of the homogenized problem

$$\begin{cases} \tau(x) \in \bar{\partial} L_{hom}(\nabla u(x)) & a.e. \quad x \in \Omega \\ -\mathrm{div}(\tau(x)) = u^*(x) & x \in \Omega \\ u \in W_0^{1,p}(\Omega), \end{cases} \tag{7.27}$$

where L_{hom} is the selfdual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$ defined by

$$L_{hom}(a,b) := \min_{\substack{\varphi \in W_{\#}^{1,p}(Q) \\ g \in L_{\#}^q(Q;\mathbb{R}^N)}} \frac{1}{|Q|} \int_Q L\big(x, a + D\varphi(x), b + g(x)\big)\, \mathrm{d}x. \tag{7.28}$$

This will follow from the following proposition.

Proposition 7.15 *Let Ω, Q and L be as in Theorem 7.14, and let $\{F_\varepsilon; \varepsilon > 0\}$ be the family of selfdual Lagrangians on $W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$ defined by*

$$F_\varepsilon(u,u^*) := \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ -\mathrm{div} f = u^*}} \int_\Omega L\big(\frac{x}{\varepsilon}, \nabla u(x), f(x)\big)\, \mathrm{d}x.$$

Then, there exists a selfdual Lagrangian F_{hom} on $\mathbb{R}^N \times \mathbb{R}^N$ that is a Γ-limit of $\{F_\varepsilon; \varepsilon > 0\}$ on $W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$. It is given by the formula

$$F_{hom}(u,u^*) := \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ -\mathrm{div} f = u^*}} \int_\Omega L_{hom}\big(\nabla u(x), f(x)\big)\, \mathrm{d}x,$$

where L_{hom} is the selfdual Lagrangian on $\mathbb{R}^N \times \mathbb{R}^N$ defined by (7.28), and which satisfies for all $(a,b) \in \mathbb{R}^N \times \mathbb{R}^N$

$$C_0(|a|^p + |b|^q - 1) \leq L_{hom}(a,b) \leq C_1(|a|^p + |b|^q + 1).$$

Proof. Note first that the selfduality and uniform bounds of L_{hom} follow from Proposition 7.7. It also follows from Proposition 3.22 that both F_ε and F_{hom} are self-dual Lagrangians on $W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$. Given $(u,u^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$, we now show the existence of a sequence $\{(u_\varepsilon, u_\varepsilon^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)\}$ with $u_\varepsilon \to u$ weakly in $W_0^{1,p}(\Omega)$ and $u_\varepsilon^* \to u^*$ strongly in $W^{-1,q}(\Omega)$ and such that

$$\limsup_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon, u_\varepsilon^*) \leq F_{hom}(u,u^*). \tag{7.29}$$

For that we consider $\{G_\varepsilon; \varepsilon > 0\}$ be a family of functionals on $W^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ defined by

$$G_\varepsilon(u,\tau) := \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L\big(\frac{x}{\varepsilon}, \nabla u(x), \tau(x)+f(x)\big)\,\mathrm{d}x,$$

and

$$G_{hom}(u,\tau) := \inf_{\substack{f\in L^q(\Omega;\mathbb{R}^N)\\ \operatorname{div} f=0}} \int_\Omega L_{hom}\big(\nabla u(x), \tau(x)+f(x)\big)\,\mathrm{d}x,$$

Take $\tau \in L^q(\Omega;\mathbb{R}^N)$ such that $\operatorname{div}(\tau) = u^*$. It follows from Lemma 7.11 and Lemma 7.13 that there exists $(u_\varepsilon, \tau_\varepsilon) \in W_0^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N)$ such that $u_\varepsilon \to u$ strongly in $L^p(\Omega)$ and $\tau_\varepsilon \to \tau$ strongly in $L^q(\Omega;\mathbb{R}^N)$ and

$$\limsup_{\varepsilon\to 0} G_\varepsilon(u_\varepsilon, \tau_\varepsilon) \leq G_{hom}(u,\tau).$$

The sequence u_ε is bounded in $W_0^{1,p}(\Omega)$, so we may assume $u_\varepsilon \to u$ weakly in $W_0^{1,p}(\Omega)$. Since $\tau_\varepsilon \to \tau$ strongly in $L^q(\Omega;\mathbb{R}^N)$, it follows that $u_\varepsilon^* := \operatorname{div}(\tau_\varepsilon) \to \operatorname{div}(\tau) = u^*$ strongly in $W^{-1,q}(\Omega)$. Thus, the inequality (7.29) follows by noticing that $G_\varepsilon(u_\varepsilon,\tau_\varepsilon) = F_\varepsilon(u_\varepsilon, u_\varepsilon^*)$ and $G_{hom}(u,\tau) = F_{hom}(u,u^*)$.

We shall now show that if $(u,u^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$ and $u_\varepsilon \to u$ weakly in $W_0^{1,p}(\Omega)$ and $u_\varepsilon^* \to u^*$ strongly in $W^{-1,q}(\Omega)$ then

$$F_{hom}(u,u^*) \leq \liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon, u_\varepsilon^*). \tag{7.30}$$

Take an arbitrary element in $(v,v^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$. From the above, there exists $(v_\varepsilon, v_\varepsilon^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$ with $\tilde{v}_\varepsilon \to v$ weakly in $W_0^{1,p}(\Omega)$ and $v_\varepsilon^* \to v^*$

strongly in $W^{-1,q}(\Omega)$ and such that

$$\limsup_{\varepsilon\to 0} F_\varepsilon(v_\varepsilon, v_\varepsilon^*) \le F_{hom}(v, v^*).$$

By the self duality of F_ε we have

$$\begin{aligned} F_\varepsilon(u_\varepsilon, u_\varepsilon^*) = F_\varepsilon^*(u_\varepsilon^*, u_\varepsilon) &= \sup\{\langle u_\varepsilon, w^*\rangle + \langle u_\varepsilon^*, w\rangle - F_\varepsilon(w, w^*) \; : \\ &\qquad (w, w^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)\} \\ &\ge \langle u_\varepsilon, v_\varepsilon^*\rangle + \langle u_\varepsilon^*, v_\varepsilon\rangle - F_\varepsilon(v_\varepsilon, v_\varepsilon^*), \end{aligned}$$

from which we get

$$\begin{aligned} \liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon, u_\varepsilon^*) &\ge \liminf_{\varepsilon\to 0}\{\langle u_\varepsilon, v_\varepsilon^*\rangle + \langle u_\varepsilon^*, v_\varepsilon\rangle - F_\varepsilon(v_\varepsilon, v_\varepsilon^*)\} \\ &= \langle u, v^*\rangle + \langle u^*, v\rangle - \limsup_{\varepsilon\to 0} F_\varepsilon(v_\varepsilon, v_\varepsilon^*) \\ &\ge \langle u, v^*\rangle + \langle u^*, v\rangle - F_{hom}(v, v^*). \end{aligned}$$

Since the above holds for an arbitrary $(v, v^*) \in W_0^{1,p}(\Omega) \times W^{-1,q}(\Omega)$, we obtain

$$F_{hom}^*(u^*, u) \le \liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon, u_\varepsilon^*).$$

Taking into consideration that F_{hom} is selfdual we obtain

$$F_{hom}(u, u^*) = F_{hom}^*(u^*, u) \le \liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon, u_\varepsilon^*),$$

as desired. □

Proof of Theorem 7.14. Since (u_n, τ_n) are solutions of (7.26), it follows that

$$\begin{aligned} 0 &= \int_\Omega L(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x))\,\mathrm{d}x - \int_\Omega \langle \nabla u_n(x), \tau_n(x)\rangle_{\mathbb{R}^N}\,dx \\ &= \int_\Omega L(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x))\,dx - \int_\Omega u_n(x) u_n^*(x)\,dx. \end{aligned} \tag{7.31}$$

Due to the coercivity assumption on L and the strong convergence of u_n^*, it follows that the sequence u_n is bounded in $W_0^{1,p}(\Omega)$ and τ_n is bounded in $L^q(\Omega;\mathbb{R}^N)$. Thus, up to a subsequence, $u_n \to u$ weakly in $W_0^{1,p}(\Omega)$ and $\tau_n \to \tau$ weakly in $L^q(\Omega;\mathbb{R}^N)$. We also have $\mathrm{div}(\tau_n) = u_n^* \to u^* = \mathrm{div}(\tau)$ strongly in $W^{-1,q}(\Omega)$, from which we indeed have $\tau_n \to \tau$ in the $\mathscr{T}$-topology (introduced in Theorem 7.9).

Taking $f \in L^q(\Omega;\mathbb{R}^N)$ with $\mathrm{div} f = 0$, it follows from (7.31) that

$$\begin{aligned}
\int_\Omega L(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x))\,\mathrm{d}x &= \int_\Omega u_n(x) u_n^*(x)\,dx \\
&= -\int_\Omega u_n(x)\mathrm{div}(\tau_n + f)\,dx \\
&= \int_\Omega \langle \nabla u_n(x), \tau_n + f\rangle_{\mathbb{R}^N}\,dx \\
&\leq \int_\Omega L\big(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n + f(x)\big)\,dx.
\end{aligned}$$

This indeed shows that

$$\int_\Omega L\big(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x)\big)\,\mathrm{d}x = \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ \mathrm{div} f = 0}} \int_\Omega L\big(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x) + f(x)\big)\,\mathrm{d}x.$$

Let

$$G_{\varepsilon_n}(v, \hat{\tau}) := \inf_{\substack{f \in L^q(\Omega;\mathbb{R}^N) \\ \mathrm{div} f = 0}} \int_\Omega L\big(\frac{x}{\varepsilon}, \nabla v(x), \hat{\tau}(x) + f(x)\big)\,\mathrm{d}x.$$

It then follows that $\int_\Omega L\big(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x)\big)\,\mathrm{d}x = G_{\varepsilon_n}(u_n, \tau_n)$. Define $H : W_0^{1,p}(\Omega) \times L^q(\Omega;\mathbb{R}^N) \to \mathbb{R}$ by $H(v, \tilde{\tau}) = \int_\Omega \langle \nabla v(x), \tilde{\tau}(x)\rangle_{\mathbb{R}^N}\,dx$. Note that H is continuous if we consider the weak topology of $W_0^{1,p}(\Omega)$ and the $\mathscr{T}$-topopogy for $L^q(\Omega;\mathbb{R}^N)$. It then follows from Lemma 7.12 that

$$\begin{aligned}
\int_\Omega L_{hom}\big(\nabla u(x), \tau(x)\big)\,\mathrm{d}x - H(u,\tau) &\leq \liminf_{\varepsilon_n \to 0} \big[G_{\varepsilon_n}(u_n, \tau_n) - H(u_n, \tau_n)\big] \\
&= \liminf_{\varepsilon_n \to 0} \Big[\int_\Omega L(\frac{x}{\varepsilon_n}, \nabla u_n(x), \tau_n(x))\,\mathrm{d}x - \\
&\qquad \int_\Omega u_n(x)\mathrm{div}(\tau_n(x))\,dx\Big] \\
&= 0.
\end{aligned}$$

On the other hand, we have that

$$\begin{aligned}
\int_{\Omega} L_{hom}\big(\nabla u(x), \tau(x)\big)\, \mathrm{d}x - H(u,\tau) &= \\
\int_{\Omega} \Big[L_{hom}\big(\nabla u(x), \tau(x)\big) - \langle \nabla u(x), \tau(x)\rangle_{\mathbb{R}^N} \Big]\, dx &\geq 0.
\end{aligned}$$

which means that the latter is indeed zero, i.e.,

$$\int_{\Omega} \Big[L_{hom}\big(\nabla u(x), \tau(x)\big) - \langle \nabla u(x), \tau(x)\rangle_{\mathbb{R}^N} \Big] dx = 0.$$

Since the integrand is itself non-negative we have

$$L_{hom}\big(\nabla u(x), \tau(x)\big) - \langle \nabla u(x), \tau(x)\rangle_{\mathbb{R}^N} = 0 \quad \text{a.e. } x \in \Omega,$$

which together with $-\mathrm{div}(\tau(x)) = u^*(x)$, yields

$$\begin{cases}
\tau(x) \in \bar{\partial} L_{hom}(\nabla u(x)), & a.e. x \in \Omega, \\
-\mathrm{div}(\tau(x)) = u^*(x), & x \in \Omega, \\
u \in W_0^{1,p}(\Omega).
\end{cases}$$

□

Bibliography

[AB93] Hedy Attouch and Gerald Beer. On the convergence of subdifferentials of convex functions. *Arch. Math.*, 60:389–400, 1993.

[Att84] Hedy Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), 1984.

[Bar84] V. Barbu. *Optimal control of variational inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[BC94] A. Braides and V. Chiadò Piat. Remarks on the homogenization of connected media. *Nonlinear Anal.*, 22(4):391–407, 1994.

[BCD92] A. Braides, V. Chiadò Piat, and A. Defranceschi. Homogenization of almost periodic monotone operators. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 9(4):399–432, 1992.

[Bee93] Gerald Beer. *Topologies on closed and closed convex sets*, volume 268 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993.

[BK95] V. Barbu and K. Kunisch. Identification of nonlinear parabolic equations. *Control Theory Adv. Tech.*, 10(4):1959–1980, 1995.

[BK96] V. Barbu and K. Kunisch. Identification of nonlinear elliptic equations. *Appl. Math. Optim.*, 33(2):139–167, 1996.

[BP86] V. Barbu and Th. Precupanu. *Convexity and optimization in Banach spaces*, volume 10 of *Mathematics and its Applications (East European Series)*. D. Reidel Publishing Co., Dordrecht, 1986. Translation from the Romanian 2nd edition.

[BR65] A. Brøndsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.*, 16:605–611, 1965.

[BS99] Regina Sandra Burachik and B. F. Svaiter. ε-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Anal.*, 7(2):117–132, 1999.

[BWY10] H.H. Bauschke, X. Wang, and L. Yao. Autoconjugate representers for linear monotone operators. *Mathematical Programming (Series B)*, 123:5–24, 2010.

[Bys01] J. Byström. Correctors for some nonlinear monotone operators. *J. Nonlinear Math. Phys.*, 8(1):8–30, 2001.

[Dal93] G. Dal Maso. *An introduction to* Γ*-convergence*, volume 8 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhauser Boston, Inc., 1993.

[De 91] R. De Arcangelis. A general homogenization result for almost periodic functionals. *J. Math. Anal. Appl.*, 156(2):358–380, 1991.

[DMS08] Alain Damlamian, Nicolas Meunier, and Jean Van Schaftingen. Periodic homogenization for convex functionals using Mosco convergence. *Ric. Mat.*, 57(2):209–249, 2008.

[DMV07] A. Damlamian, N. Meunier, and J. Van Schaftingen. Periodic homogenization of monotone multivalued operators. *Nonlinear Anal.*, 67(12):3217–3239, 2007.

[ET99] Ivar Ekeland and Roger Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[Fit88] Simon Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.

[FM87] N. Fusco and G. Moscariello. On the homogenization of quasilinear divergence structure operators. *Ann. Mat. Pura Appl.*, 146(4):1–13, 1987.

[FMT09] G. A. Francfort, F. Murat, and L. Tartar. Homogenization of monotone operators in divergence form with x-dependent multivalued graphs. *Ann. di Matematica*, 2009.

[Gho07] Nassif Ghoussoub. Anti-selfdual lagrangians: Variational resolutions of non self-adjoint equations and dissipative evolutions. *AIHP-Analyse non linéaire*, 24:171–205, 2007.

[Gho08a] Nassif Ghoussoub. *Self-dual Partial Differential Systems and Their Variational Principles*. Springer Monographs in Mathematics. Springer, 2008.

[Gho08b] Nassif Ghoussoub. A variational theory for monotone vector fields. *J. Fixed Point Theory Appl.*, 4(1):107–135, 2008.

[GMZ10] Nassif Ghoussoub, Abbas Moameni, and Ramon Zárate Sáiz. Homogenization of maximal monotone vector fields via selfdual variational calculus, 2010.

[PD90] V. C. Piat and A. Defranceschi. Homogenization of monotone operators. *Nonlinear Anal.*, 14(9):717–732, 1990.

[PDD90] V. Chiadò Piat, G. Dal Maso, and A. Defranceschi. G-convergence of monotone operators. *Ann. Inst. H. Poincar Anal. Non Linaire*, 7(3):123–160, 1990.

[Phe93] R.R. Phelps. *Convex functions, monotone operators and differentiability*, volume 1364 of *Lecture Notes in Math.* Springer Verlag, 2nd edition, 1993.

[Roc70] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[Suq82] P. Suquet. *Plasticite et Homogeneisation*. PhD thesis, Univ. Paris VI, 1982.

[SZ04] S. Simons and C. Zălinescu. A new proof for rockafellar's characterization of maximal monotone operators. *Proc. Amer. Math. Soc.*, 132(10):2969–2972 (electronic), 2004.

[VG86] P. A. Raviart V. Girault. *Finite element methods for Navier-Stokes equations. Theory and algorithms*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986.