Ghoussoub's self-dual variational principle

Hannes Stuke

February 2011

CONTENTS

CHAPTER 1

INTRODUCTION

Contents

1.1 Overview

This B.A thesis deals with the book "Self-dual partial differential systems and their variational principles" [Ghoussoub 2009], in which Nassif Ghoussoub proposes a principle that is able of ensuring the solvability of a certain class of (time-dependent) PDEs. This B.A thesis has several aims

1. In Chapter 2 it is shown, that Ghoussoub's Variational principle is new, because it allows a variational treatment of equations, which are not accessible by the classical Euler-Lagrange calculus.
2. The general framework of Ghoussoub, which is mainly the calculus of convex functions is introduced next.
3. With this preliminary discussion the main theorem for the time independent case is presented in Chapter 4. Thereby the equivalence of convex variation and solving PDEs is focused. In the last part of

the chapter it is shown, that it is hard to use Ghoussoub's variational principle without knowing in advance that a solution exists by other methods.

4. The question of the bounds of the possible application of the approach is answered in chapter 5, mainly in Theorem 5.1
5. The last chapter shows how to lift the stationary results to a time-dependent setting.

.

1.2 Partial differential equations

Many processes in nature can by described by the theory of partial differential equations (PDE). They have the general form (see [Evans 2010]):

$$F : \mathbb{R}^{n^k} \times R^{n^{k-1}} \times \cdots \times \mathbb{R} \times U \to \mathbb{R}$$
$$F\left(D^k u(x), D^{k-1} u(t), ..., u(x), x\right) = 0$$

where U is on open subset of $\mathbb{R}^n$. As one can see, this term is only well defined if u k-times differentiable. But in physics we often have to deal with non-differentiable processes. One more mathematical, but rather simple example is the solution of the PDE, satisfying:

Example 1.2.1.

$$\begin{aligned} \dot{x} &= \begin{cases} 1, & x < 0 \\ 0, & x = 0 \\ -1, & x > 0 \end{cases} \\ x(0) &= x_0 \end{aligned}$$

Because the solution is not differentiable at $x = 0$, there is no solution to that problem in the sense of a classical solution of PDEs, which has to be differentiable. In order to define a theory, which could deal with such problems, we have to think of a new framework, which is capable of answering three questions:

1. In what sense are the solutions to the above PDE differentiable? We are still dealing with an PDE, that is to say that we are claiming some relation between a, howsoever defined, derivative of a function and some right-hand side.

2. This arises the next point, which is crucial for a mathematical rigor framework. What is the space of functions in which we are looking for a solution?

3. But then we also have to deal with set-valued maps, thus we have to define a framework for PDEs with set-valued right-hand site. This generalization of an PDE is theory of the differential inclusions.

We will see that we are dealing with a system like:

$$x(t) \in Ax(t) + f \tag{1.1}$$
$$x(0) = x_0 \tag{1.2}$$

where x is an absolute continuous function. Thus equation (1.1) only holds for almost all t. Of course, if we want to show some existence results, we have to make assumptions on A and on the spaces in which we are looking for the solutions. All of the questions are treated in the chapters 3 and 4.

1.3 Existence methods

There are various different methods to ensure the existence of a solution of a PDE [1]. One important method is the variational formulation leading to Euler-Lagrange equations, see Chapter 2. According to this technique there arises the natural question of which PDEs are treatable as variational problem, thus could be written as Euler-Lagrange equations. This is called the inverse problem of the Euler-Lagrange variation, which is still an issue of topical of research. But there are PDEs which could be solved via this variational approach, see for instance chapter two.
Ghoussoub has shown, by evolving an idea of Auchmuty [Auchmuty 1987], that there is nevertheless a variational formulation for some of that PDEs. The connection between convex variation and PDEs will be established in chapter four for the stationary case and in chapter 6 in the dynamical case.

1.4 Notation

Throughout the whole text Ω denotes a smooth bounded, open and connected domain of R^n, if nothing else is explicitly stated. X is a reflexive Banach

[1] PDE means from now on PDE in the sense of equation (1.1)

space and X^* its dual. The space $L_X^p(\Omega), p < 1 < +\infty$ denotes the space of all Lebesgue-integrable functions $u : \Omega \mapsto X$ equipped with the norm:

$$\|u\|_{L_X^p} = \left(\int_\Omega \|u(x)\|^p dx\right)^{\frac{1}{p}}$$

Definition. Let H be an Hilbert space, s.t.

1. $X \subset H \subset X^*$ are dense.

2. The canonical injections are continuous.

Then the triple (X, H, X^*) is called an *evolution triple* (e. t). For further details see [Zeidler 1990a, Section 23]

Definition (Time dependent spaces)**.** Let $X \subset H \subset X^*$ be an *evolution triple.* Consider the space

$$\mathcal{H}_{p,q}[0,T] := \{u : u \in L_X^p[0,T], \dot{u} \in L_{X^*}^q[0,T]\}$$

equipped with the norm

$$\|u\|_{\mathcal{H}_{p,q}} := \|u\|_{L_X^p[0,T]} + \|\dot{u}\|_{L_{X^*}^q[0,T]}$$

Definition (Weak derivative)**.** Let $X \subset H \subset X^*$ be an evolution triple. Further assume $u \in L_p([0,T], X)$ $1 \leq p, q \leq \infty$. Then $\dot{u}$ denotes the *weak derivative* of u if:

$$\dot{u} : L_p([0,T], X) \to L_q([0,T], X^*), \qquad u \mapsto w$$

where $\forall \varphi \in C_0^\infty([0,T]), \forall x \in X$

$$\int_0^T \langle u(t), \dot{\varphi}(t)\rangle \, \dot{d}t = -\int_0^T \langle w(t), \varphi(t)\rangle \, dt$$

Further I will denote the partial derivative with subscript, e.g.

$$f_p(u,v) := \frac{\partial}{\partial p} f(x,p)\,|_{x=u,p=v}$$

CHAPTER 2

CLASSICAL VARIATION PRINCIPLES

Contents

I briefly recall the classical calculus of variations. The calculus of variations is a method to transform a given PDE into a problem of finding stationary points of an associated functional. I also indicate how the approach of Ghoussoub differs from the classical one. For details see [Evans 2010]

2.1 General idea

Assume you are given a PDE like $A[u] = 0$ and you can show that there exists a functional $I : X \to \mathbb{R}$, satisfying the property

$$I'[v] = A[v]$$

Then the above PDE is equivalent to v being critical point of I.

2.1.1 Euler-Lagrange equation

To be more specific consider I of the form:

$$I[u] := \int_\Omega L(x, u(x), Du(x))\, dx$$

where $u \in H_0^{1,p}(\Omega)$ and $\Omega \subset \mathbb{R}^d$ is bounded. $L : \mathbb{R}^n \times \mathbb{R} \times \Omega \mapsto \mathbb{R}$ denotes the *classical Lagrangian.*

Proposition 2.1.1. *Let the following conditions hold:*

1. $L(\cdot, u, v)$ *is measurable for all* $u \in \mathbb{R}, v \in \mathbb{R}^d$.
2. $L(x, \cdot, \cdot)$ *is almost everywhere differentiable for all* $x \in \Omega$
3. $|L(x, u, v)| \leq c_0 + c_1|u|^p + c_2|v|^p$ *for almost all* $x \in \Omega$ *and for all* $u \in \mathbb{R}, v \in \mathbb{R}^d$
4. $|\frac{\partial L}{\partial u}(x, u, v)| + \sum_i |\frac{\partial L}{\partial v^i}(x, u, v)| \leq c_3 + c_4|u|^p + c_5|v|^p$ *for all* $x \in \Omega, u \in \mathbb{R}, v \in \mathbb{R}^d$

Assume u is a stationary point of I. Then:

$$\int_\Omega \left\{ \frac{\partial L}{\partial u}(x, u(x), Du(x))\,\varphi(x) + \frac{\partial L}{\partial v_i}(x, u(x), Du(x))\,\frac{\partial \varphi(x)}{\partial x^i} \right\} dx = 0. \tag{2.1}$$

where $\varphi \in C_0^\infty(\Omega)$ *denotes a test function.*
If $u \in C^2(\Omega)$, *then*

$$\sum_i^d \frac{d}{dx}\left[\frac{\partial}{\partial v^i} L(x, u(x), Du(x))\right] - \frac{\partial}{\partial u} L(x, u(x), Du(x)) = 0 \tag{2.2}$$

This are the ***Euler-Lagrange equations***.

Proof. Proofs can be found in [Jost and Li-Jost 1999, Section 4] □

Proposition 2.1.2. *Assume further for some* $\alpha > 0, \beta \geq 0, 1 < q < \infty$

$$\forall x \in \Omega, u \in \mathbb{R}, v \in \mathbb{R}^d : \alpha\|v\|^q - \beta \leq L(x, u, v),$$

Then the functional I attains its infimum.

Proof. For the proof see [Evans 2010, Section 8.2] □

The two above Propositions provide an example for the classical Euler-Lagrange calculus for PDEs. All PDEs to be solved by that approach, must thus be writable as Euler-Lagrange equations (2.2) of a suitable Lagrange function L.

Problems of the Euler-Lagrange approach

This leads to the problems of the Euler-Lagrange approach:

1. You have to show the regularity of the stationary point, whose existence is often only easy to show with weaker assumption. So you have to choose the space of admissible functions very carefully.

2. Not all stationary points are minima of I. The existence of stationary, non-minimizing points is far more difficult to show in general.

3. As mentioned above, there is a huge class of PDEs which are not of Euler-Lagrange type [Olver 1993, Theorem 5.92],e.g. the heat equation. But it is not very easy to see, whether a given PDE is of Euler-Lagrange type or not.

I want to emphasize the last point, because many of the physical equations are derived by modeling fluxes, e.g. of heat or particles, assuming some proportionality between the change of concentration in time and in space. There are of course laws, which allow for a Lagrangian formulation, but the main difficulty remains to determine which ones do.
One answer has been found by Helmholtz, who has formulated the Helmholtz conditions, which ensure that the PDE is essentially of Euler-Lagrange type. Although there are some difficulties with this conditions, one can still show that e.g. the heat equation could not be treated with the classical calculus of variation. For a more detailed discussion see [Olver 1993]. One interesting question is, if there are other applicable variational principles, which could handle the heat equation.

2.2 Example: The heat equation

In this section the example of the heat equation is explicitly calculated.

Definition (Heat equation)**.**

$$\begin{aligned} \dot{u}(t,x) &= \Delta u(t,x) \text{ for almost every t} \qquad (2.3) \\ u(0,x) &= u_0(x) \\ u(t,x) &= 0 \text{ on } x \in \partial\Omega \end{aligned}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary. I am only looking for weak solutions, what is to say, the space of admissible functions is $\mathcal{H}_{\frac{1}{2},\frac{1}{2}}[0,T]$.

We indicate a problem with the classical Euler-Lagrange type variation and how this problem is resolved in Ghoussoub's approach. See also [Tzou and Ghoussoub 2004].

Inverse problem for the heat equation

Consider the 1-dim heat equation and a functional L for which it is the Euler-Lagrange equation. The Lagrangian has the form:

$$L\left(t, x, u\left(t, x\right), u_t\left(t, x\right), u_x\left(t, x\right)\right)$$

Consider the Euler-Lagrange equations and calculate them explicitly. Denote $q := u_t\left(t, x\right), p := u_x\left(t, x\right)$ They become:

$$0 = L_{pp}u_{xx}+L_{px}+L_{pt}+L_{qx}+L_{qt}+L_{pu}u_x+L_{qu}u_t+L_{qq}u_{tt}+2L_{pq}u_{tx}-L_u \quad (2.4)$$

If this should be the heat equation, the Lagrangian must satisfy:

$$L_{pp} = 1, L_{qq} = 0, L_{pq} = 0, L_{px} + L_{pt} + L_{qx} + L_{qt} + L_{pu}u_x + L_{qu}u_t - L_u = -u_t$$

Because of $L_{qq} = 0, L_{pq} = 0$ all possible Lagrangians L are of the form:

$$L = \lambda\left(x, t, u\right)u_t + \sigma\left(x, t, u, u_x\right)$$

By plugging in this Lagrangian in Equation 2.4 one obtains:

$$u_{xx}\sigma_{pp} + \sigma_{px} + \sigma_{pt} + \lambda_x + \lambda_t + \sigma_{pu}u_x + \lambda_u u_t - u_t\lambda_u - \sigma_u = 0$$

But there is no term involving u_t, thus there is no possibility to recover the heat equation. This shows at least the 1-dim heat equation is not of Euler-Lagrange type provided we are looking on spaces with enough differentiability.

The Variation

In order to resolve the heat equation variational, try the following functional.

$$\begin{aligned} I\left[u\right] &:= \int_\Omega \left[\frac{1}{2}\int_0^T \left(\|\nabla u\left(x,t\right)\|^2 + \|\nabla(-\Delta)^{-1} - \dot{u}\left(x,t\right)\|^2\right)dt \right. && (2.5) \\ &+ \left. \frac{1}{2}u\left(T,x\right)^2 + u_0\left(x\right)^2 + \frac{1}{2}u\left(0,x\right)^2 - 2u_0\left(x\right)u\left(0,x\right)\right]dx && (2.6) \end{aligned}$$

where $u \in \mathcal{H}_{\frac{1}{2},\frac{1}{2}}\left[0,T\right]$.

Example 2.2.1 (Euler-Lagrange View). According to Proposition 2.1.1 one has to calculate the stationary points of I:

$$\begin{aligned} 0 = I'[u] &= \int_\Omega \Bigg[\int_0^T \left(\nabla u \nabla \phi(t,x) + \nabla(-\Delta)^{-1}(-\dot{u}(x,t)) \nabla(-\Delta)^{-1}\left(-\dot{\phi}(x,t)\right)\right) dt \\ &+ u(T,x)\phi(T,x) + u(0,x)\phi(0,x) \\ &- 2u_0(x)\phi(0,x)\Bigg] dx \end{aligned}$$

For test function $\phi \in C_0^\infty(\Omega)$. We first consider $\phi(T,x) = \phi(0,x) = 0$:

$$I'[u] = \int_\Omega \int_0^T \left[-\Delta u(t,x) + (\Delta)^{-1}\ddot{u}(x,t)\right] \phi(x,t)\, dt = 0 \tag{2.7}$$

But from Equation 2.7 follows, that the equation

$$\Delta^2 u - \ddot{u} = 0 \tag{2.8}$$

has to hold. Even though solutions of the heat equation (2.3) satisfy Equation 2.8 the converse need not be true. Indeed 2.8 is a wave equation, which has some very different properties:

1. The equation is invariant under time-reversal operation, that is if $u(t,x)$ satisfies 2.8, then also $u(-t,x)$ satisfies such equation. The backwards solution are exactly of the same form. But for the heat equation time reversal changes the form of the equation in a fundamental way:

$$\dot{u} + \Delta u = 0$$

   Thus the past of the system is not of the same form than the future, one can actually not look into the past. This is certainly not the case for the wave equation.

2. The wave equation has a finite propagation speed. That is to say, the change of the solution at $(x,0)$ does not influence the other solution values immediately. For the heat equation it does, which you can see by considering the fundamental solution:

$$u(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\, dy$$

   If $g \leq 0$, then $u(x,t)$ is positive anywhere, no matter how small the and located the heat was distributed in the beginning.

3. The equation needs solutions, which satisfy stronger differentiability conditions.

4. The equation is of second order in time and thus needs an additional initial condition for $\dot{u}$ at time $t = 0$.

This shows that, the Euler-Lagrange equations do not look particularly promising for solving the initial heat equation 2.2 at least by choosing the above functional.

Example 2.2.2 (Ghoussoub's view). In Ghoussoub variational approach one can associate variational problems to PDEs, but one has not only to search for stationary points, but for functions for which the functional vanishes. It can be shown, that these functions are also global minima, thus the problem is a variational problem. That is a big problem, because above it had turned out, that the minima must satisfy the Euler-Lagrange equations, which are not the same as the heat equation. How can that problem be resolved?
The negative Laplace operator is self-adjoint with compact resolvent and strictly positive. Therefore take a spectral approach:

$$u(x,t) = \sum_{\lambda \in \sigma(-\Delta)} c_\lambda(t)\, h_\lambda(x)$$

. Insert this ansatz in the general solution of the problem (2.8):

$$0 = \ddot{u}(t,x) - \Delta^2 u(t,x) = \sum_\lambda \ddot{c}_\lambda(t)\, h_\lambda(x) - c_\lambda(t)\, \lambda^2 h_\lambda(x)$$

This leads to:

$$\ddot{c}_\lambda(t) = \lambda^2 c_\lambda(t)$$

From this follows:

$$c_\lambda(t) = A_\lambda e^{-\lambda t} + B_\lambda e^{\lambda t}$$

Inserting this in gives:

$$\begin{aligned} u(x,t) &= \sum_{\lambda \in \sigma(-\Delta)} \left(A_\lambda e^{-\lambda t} + B_\lambda e^{\lambda t}\right) h_\lambda(x) \\ &= \sum_{\lambda \in \sigma(-\Delta)} v_\lambda(x)\, T_{-\lambda}(t) + w_\lambda(x)\, T_\lambda(t) = \sum_{\lambda \in \sigma(-\Delta)} u_\lambda(x,t) \end{aligned}$$

with $v_\lambda(x) := A_\lambda h_\lambda(x)$, $w_\lambda(x) := B_\lambda h_\lambda(x)$, $T_{\pm\lambda}(t) := e^{\pm\lambda t}$ and $u_\lambda := v_\lambda(x)\, T_{-\lambda}(t) + w_\lambda(x)\, T_\lambda(t)$ In order to get the result by Ghoussoub's approach, one must ensure that the functional I from Equation 2.5 vanishes independently for each eigenfunction u_λ :

1. For the first term
$$\Delta u_\lambda = -\lambda \left(w_\lambda T_\lambda + v_\lambda T_{-\lambda}\right)$$

2. The second
$$\dot{u}_\lambda = \lambda \left(w_\lambda T_\lambda - v_\lambda T_{-\lambda}\right)$$

3. Thus
$$\Delta^{-1} \dot{u}_\lambda = v_\lambda T_{-\lambda} - w_\lambda T_\lambda$$

4. The sum in the integral becomes
$$\Delta u_\lambda u_\lambda + \Delta^{-1} \dot{u}_\lambda \dot{u}_\lambda = -2\lambda \left(w_\lambda^2 T_\lambda^2 + v_\lambda^2 T_{-\lambda}^2\right)$$

5. Thus the integral becomes
$$\int_0^T -\frac{1}{2} \Delta u_\lambda u_\lambda - \frac{1}{2} \Delta^{-1} \dot{u}_\lambda \dot{u}_\lambda dt = \frac{1}{2} w_\lambda^2 \left(e^{2\lambda T} - 1\right) - \frac{1}{2} v_\lambda^2 \left(e^{-2\lambda T} - 1\right) \quad (2.9)$$

6. Further
$$\begin{aligned} \frac{1}{2} u\left(T, x\right) &= \frac{1}{2} \left(w_\lambda T_\lambda\left(T\right) + v_\lambda T_{-\lambda}\left(T\right)\right)^2 \\ &= \frac{1}{2} w_\lambda^2 e^{2\lambda T} + w_\lambda v_\lambda + \frac{1}{2} v_\lambda^2 e^{-2\lambda T} \end{aligned}$$

7. When the functional I should vanish we need $w_\lambda = 0$, thus $B_\lambda = 0$ by comparison with (2.9)

8. Thus follows
$$u\left(t, x\right) = \sum_{\lambda \in \sigma(-\Delta)} v_\lambda\left(x\right) e^{-\lambda t}$$

9. Adding all further terms one gets $I\left(u\right) = \frac{1}{2} v_\lambda^2 - \frac{1}{2} v_\lambda^2 = 0$

The wrong solution $B_\lambda \neq 0$ is thrown away by the condition, that I must vanish. This is because we can get many stationary points of the integral, but only those, which are also zeros are a weak solution of the heat equation. The functions u satisfying $I\left[u\right] = 0$ are thus weak solutions for the heat equation (2.3)

Remark 2.2.1. One can see directly by inserting a solution of the heat equation in (2.5), that the Lagrangian vanishes:

$$\begin{aligned} I(u) &= \int_0^T \int_\Omega \|\nabla u(x,t)\|^2 dt + \frac{1}{2} u(T,x)^2 - \frac{1}{2} u(0,x)^2 \, dx \\ &= \int_0^T \int_\Omega -\frac{1}{2} \frac{d}{dt} u(t,x)^2 \, dt + \frac{1}{2} u(T,x)^2 - \frac{1}{2} u(0,x)^2 \, dx \end{aligned}$$

That zero is a minima of such functionals, is shown in its general form later.

In this section, it was shown by the example of the heat equation, that the variational calculus of Ghoussoub can handle PDEs, which are not accessible by Euler-Lagrange variation. For this it was necessary to prove, that the functional vanishes to eliminate the wrong solutions of the Euler-Lagrange equations. Why this condition is important will be explained later.

CHAPTER 3

FENCHEL CONJUGATION, LAGRANGIANS AND SELF-DUALITY

Contents

In this chapter the Legendre dual and its properties will be introduced, e.g. its connection to the inverse of operators. It is also shown how this framework can be used to generalize the technique of "completing the square", which is used to solve quadratic equations.

Throughout this chapter assume:

$$f : X \to \overline{\mathbb{R}} \text{ is proper } (\mathrm{dom}\,(f) \neq \emptyset, f > -\infty),\ \text{lsc and convex}$$

3.1 Mathematical prerequisites

The following Theorems, which could be looked up e.g. in Rockafellar [1972] are needed.

Definition (Epigraph)**.** The *epigraph* of an function $f : \mathrm{dom}\,(f) \subset X \to \mathbb{R}$, is defined as the set:

$$\operatorname{epi}(f) := \{(x, \mu) \in \operatorname{dom}(f) \times \mathbb{R} : f(x) \leq \mu\}$$

Proposition 3.1.1. *The epigraph of a convex function $g : X \to \mathbb{R}$ is a convex subset of $X \times \mathbb{R}$*

Theorem 3.1. *Let $C \subset X$ be convex and closed. Then C is the intersection of all closed half-spaces H_i containing it.*

$$H = \bigcap_{i \in I} H_i$$

Proof. Recall that to each $x^* \in X^*, \mu \in \mathbb{R}$ one can associate a closed half space by:

$$H_{(x^*,\mu)} := \{x \in X : \langle x, x^* \rangle \geq \mu\}$$

Let now $x \in X \setminus C$. From Hahn-Banach Theorem one knows, that there exists $x^* \in X^*$ s.t. $x^*(x) < \inf_{c \in C} x^*(c) =: \alpha$. We can easily see:

$$x \notin H_{(x^*,\alpha)}, C \subset H_{(x^*,\alpha)}$$

That is to say $x \notin H$. Thus $H \subset C$.
The other direction is clear, since for each $i \in I \Rightarrow C \subset H_i$

□

3.2 The Legendre-Fenchel dual

Definition. *The subdifferential*
The subdifferential of $\partial f : \operatorname{dom}(f) \to \mathfrak{P}(X^*)$ is a set valued function, which is defined as follows:

$$\partial f(x) := \{p \in X^*, \langle p, y - x \rangle \leq f(y) - f(x)\} \tag{3.1}$$

Example 3.2.1. We go back to the Example 1.2.1 and assume $x_0 > 0$. We see that in the sense of the subdifferential the function

$$x(t) = \begin{cases} x_0 - t, & t \in [0, x_0] \\ 0 & t > x_0 \end{cases}$$

is a solution to the problem, which is absolutely continuous, but not differentiable at $t = x_0$

Proof. Look at Figure 3.2. I just proof the case $t \leq x_0$, where the subdifferential is -1:

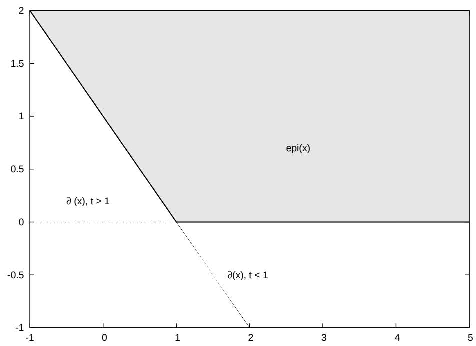

Figure 3.1: The figure shows, that the epigraph of the function x lies completely above the half space associated to the subdifferential. As one can see for is the subdifferential for $t = 1$ is multivalued. It consists of all slopes, which lie completely under the epigraph of x.

$\tilde{t} < x_0$

$$x\left(\tilde{t}\right) - x\left(t\right) = t - \tilde{t} = -\left(\tilde{t} - t\right) \rightarrow -1 \in \partial x\left(t\right)$$

$\tilde{t} \geq x_0$

$$x\left(\tilde{t}\right) - x\left(t\right) = t \geq -\left(\tilde{t} - t\right) \rightarrow -1 \in \partial x\left(t\right)$$

□

Example 3.2.2. Consider the rather mathematical example, which is shown in Figure 3.2.2:

$$f := \begin{cases} 0 & x \leq 0 \\ 1 & x \in (0, 1] \\ +\infty & x > 1 \end{cases}$$

The subdifferential has three different cases (single-valued, set-valued and empty):

$$\partial f\left(x\right) := \begin{cases} 0 & x < 0 \\ \alpha \in \mathbb{R}_{\geq 1} & x = 1 \\ \emptyset & x > 0 \vee x \in (0, 1) \end{cases}$$

Remark 3.2.1. The set $\partial f\left(x\right)$ is geometrically interpreted, the set of all half-spaces containing $f\left(x\right)$ for which $\mathrm{epi}\left(f\right)$ is entirely on one side. This could be easily seen by definition.

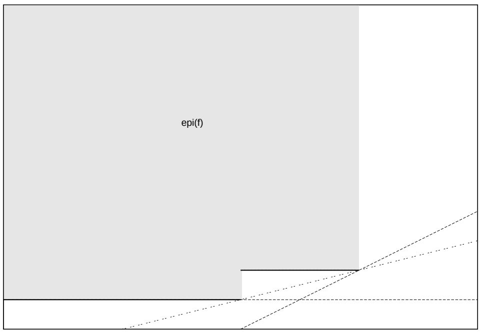

Figure 3.2: f (x). Note the three different cases explained above

Remark 3.2.2. The subdifferential is for Gateaux-differentiable functions the same as the Gateaux-differential.

Definition. *Legendre-Fenchel Dual*
The Legendre-Fenchel dual f^* of f is defined as:

$$f^* : X^* \to \mathbb{R}, \qquad x^* \mapsto \sup_{y \in X} \langle y, x^* \rangle - f(y) \tag{3.2}$$

It is a closed convex function on X^*.

Remark 3.2.3. What does the function f^* mean? Denote by F^* the set of all pairs $(x^*, \mu) \in X^* \times \mathbb{R}$, such that the associated affine function

$$A_{(x^*,\mu)}(x) := \langle x, x^* \rangle - \mu$$

is majorized by f:

$$\begin{aligned} A_{(x^*,\mu)}(x) \leq f(x) &\Rightarrow \langle x, x^* \rangle - f(x) \leq \mu \\ &\Rightarrow \sup_{x \in X} \{\langle x, x^* \rangle - f(x)\} \leq \mu \end{aligned}$$

We see that F^* is actually contained in the epigraph of f^*. Indeed $f^*(x^*)$ is the best μ, such that f majorizes the associated affine function. To rephrase this fact: The L-F-D gives for each "slope" x^* the smallest translation μ, such that the interior of the epigraph of f is entirely on one side. Thus f^* is a closed convex function and in the view of the point-wise supremum:

$$f(x) = \sup_{(x^*,\mu) \in \mathrm{epi}(f^*)} \{\langle x, x^* \rangle - \mu\} = \sup_{x^* \in X^*} \{\langle x, x^* \rangle - f^*(x^*)\}$$

Thus the closure of the epigraph is completely described by the L-F-D.

The above Theorems show, that any convex set, especially the epigraph of a convex function, can be described by the intersection of half spaces containing it. This half spaces could be described by a pair $(x^*, \mu) \in X^* \times \mathbb{R}$. Thus there are different possibilities of describing the same set. The Legendre transform could be interpreted as switching between this two types of description. Take, for example, $X = \mathbb{R}$ and $f(x) = x^2$. This case is shown in Figure 3.2. As one can the set $\operatorname{epi}(f)$ be characterized by all half spaces containing it. In fact, one only needs the best approximating half spaces $H_{\alpha,\beta}$. which are those, for which $\beta = f^*(\alpha)$ holds.

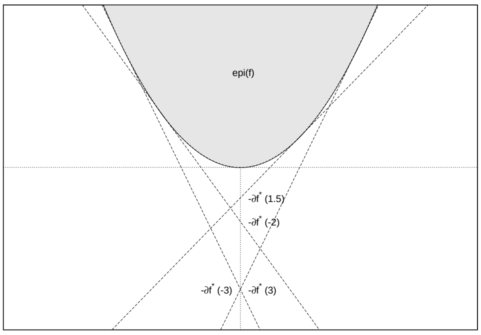

Figure 3.3: Legendre dual of x^2. One can see that in the case of $X = \mathbb{R}$, the Legendre dual is just the negative y-abscissa of the tangents

Remark 3.2.4. Note, that when the dual of

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

attains its supremum at $\overline{x}$, the function is f subdifferentiable at $\overline{x}$ and $p \in \partial f(\overline{x})$.

Proof. See Proposition 3.2.3. Note only here the geometric interpretation. If the supremum is attained, the affine function with the "slope" p and translation $f^*(p)$ really contains one element of the graph of f. This element has the subdifferential with "slope" p, since the graph lies completely one one side of the affine function. □

But this must not the case, consider the example

$$f : \mathbb{R} \to \mathbb{R} \qquad x \mapsto e^{-x}$$

One can easily see, that $f^*(0) = 0$, but there is no $\overline{x} \in \mathbb{R}$ such that $0 \in \partial f(\overline{x})$, because then must hold:

$$f(x) \geq f(\overline{x}), \qquad \forall x \in \mathbb{R}$$

This reflects the fact, that the L-F-D just represents the closure of the epigraph and not the epigraph as such.

3.2.1 Properties of the L-F-D

In this section some facts about the L-F-D are collected. Most of them can be found in [Zeidler 1985].

Proposition 3.2.1. *Involution property*
If f is proper, lsc and convex, then $f^{**} = f$

Proposition 3.2.2. *Legendre-Fenchel Inequality*
For every $(x, p) \in X \times X^*$, we have

$$f(x) + f^*(p) \geq \langle x, p \rangle \tag{3.3}$$

This is called the Legendre-Fenchel inequality

Proof.

$$\begin{aligned} f(x) + f^*(p) &= f(x) + \sup\{\langle y, p \rangle - f(y), y \in X\} \\ &\geq \langle x, p \rangle \end{aligned}$$

□

Proposition 3.2.3. *The next three statements are equivalent:*

1. $f(x) + f^*(p) = \langle x, p \rangle$
2. $p \in \partial f(x)$
3. $x \in \partial f^*(p)$

Proof. Show that (1), implies (2), (2) implies (3) and (3) implies (1)

1. Assume (1) holds. Then:

$$\langle y - x, p \rangle = -\sup\{f(x) - f(z) + \langle z, p \rangle, z \in X\} + \langle y, p \rangle \leq f(y) - f(x)$$

2. Assume (2) holds. Then:

$$\begin{aligned}\langle x, q-p\rangle &= \langle x, q\rangle + \langle y-x, p\rangle - \langle y, p\rangle \\ &\le f(y) - f(x) - \langle y, p\rangle + \langle x, q\rangle \\ &\le -f^*(p) - f(x) + \langle x, q\rangle \\ &\le -f^*(p) + \sup\{\langle z, q\rangle - f(z), z \in X\} \\ &= f^*(q) - f^*(p)\end{aligned}$$

3. Assume (3) holds. Because of Proposition 3.2.2 one only has to show $\langle x, p\rangle \ge f(x) + f^*(p)$. We know $\langle x, q-p\rangle \le f^*(q) - f^*(p) \Leftrightarrow \langle x, p\rangle \ge f^*(p) - f^*(q) + \langle x, q\rangle$. Now everything follows from Proposition 3.2.2

□

Corollary 3.1. *For F convex and lsc denote by:*

$$T : X \to X^* \qquad x \mapsto \partial f(x)$$

The Legendre-Fenchel-Duality gives another way of inverting T, in the set sense with

$$\partial f^* = T^{-1}$$

Proof. Let $p \in Tx$. Define the operator A by

$$A : X^* \to X \qquad p \mapsto \partial f^*(p)$$

By Proposition 3.2.3 $x \in Ap \Rightarrow x \in ATx$. Thus $A = T^{-1}$ □

Further properties of the L-F-D

All of this properties could be looked up in literature [Ghoussoub 2009, Proposition 2.6]

Definition. Let $f, g : X \to \mathbb{R}$ proper, lsc and convex. Define the *infimal convolution* $f \star g$ by

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

Proposition 3.2.4. *Let $f : X \to \overline{\mathbb{R}}$ be a proper function. Then the following properties hold:*

1. *f^* is proper convex and lsc*

2. *Define for* $\lambda \in \mathbb{R}_{\geq 0}$
$$(\lambda * f)(x) := \lambda f\left(\frac{x}{\lambda}\right)$$
Then holds:
$$(\lambda f)^* = \lambda * f^*$$
3. *If* $f \leq \psi \Rightarrow \psi^* \leq f^*$
4. *If* f *and* ψ *are proper, then*
$$(f \star \psi)^* = f^* + \psi^*$$
5. *If* $dom(\psi) - dom(\psi)$ *contains a neighborhood of the origin, then*
$$(f + \psi)^* = f^* \star \psi^*$$
6. *Let* $\alpha \in \mathbb{R}$*, then* $(f + \alpha)^* = f^* - \alpha$
7. *Take* $a \in X$*. Denote* $f_a(x) = f(x - a)$*. Then* $f_a^*(p) = f^*(p) + \langle a, p\rangle$

Example 3.2.3. Define $f : X \to \mathbb{R}, \alpha \geq 1$ by:
$$f(x) = \frac{1}{p}\|x\|^\alpha$$
It holds for $\beta \geq 0, \frac{1}{\alpha} + \frac{1}{\beta} = 1$
$$f^*(p) = \frac{1}{\beta}\|p\|^\beta$$

Proof. From Young's inequality for real numbers follows:
$$\langle x, p\rangle \leq \|x\|\|p\| \leq \frac{1}{\alpha}\|x\|^\alpha + \frac{1}{\beta}\|p\|^\beta$$
Further there is equality in the Young's inequality
$$\|x\|^\alpha = \|p\|^\beta$$
and equality in the first inequality if:
$$p = J(\lambda x)$$
where J denotes the duality mapping. □

3.3 Lagrangians

In this section Lagrangians on different spaces are introduced and shown how they behave under conjugation

Definition. A proper, lsc, convex function $L : X \times X^* \to \mathbb{R}$ is called Lagrangian.
To any proper, convex and lsc function $f : X \to \mathbb{R}$ one can associate the *simple Lagrangian* $L(x,p) := f(x) + f^*(p)$

Remark 3.3.1. The L-F-D of a Lagrangian L is defined as:

$$L^* : X^* \times X \to \mathbb{R}, (q,y) \mapsto \sup\{\langle x,q\rangle + \langle y,p\rangle - L(x,p), (x,p) \in X \times X^*\} \tag{3.4}$$

Definition. A Lagrangian $L : X \times X^* \to \mathbb{R}$ is called *(partially) self-dual*, if

$$L(x,0) = L^*(0,x) \qquad \forall x \in X \tag{3.5}$$

or

$$L(x,p) = L^*(p,x), \qquad \forall (x,p) \in X \times X^* \tag{3.6}$$

The set of all self-dual Lagrangians is denoted by $\mathcal{L}(X)$.

Example 3.3.1. Obviously any simple Lagrangian is self-dual. This class of Lagrangian is a very important class, which is easy to handle.

Example 3.3.2. Self duality on Hilbert spaces

On a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ with norm $\|\cdot\|$ induced by the inner product, the function $q(x) := \frac{1}{2}\|x\|^2$ is the only function for which self-duality holds:

$$q(x) = q^*(x)$$

Proof. First we prove the q is actually self-dual.

$$q^*(x) = \sup_{y\in X}\{\left\langle x - \frac{1}{2}y, y\right\rangle\} \tag{3.7}$$

This obviously attains it supremum at $y = x$ leading to $q^*(x) = q(x)$. For the other direction note that from the Legendre-Fenchel inequality for any self-dual function Γ follows:

$$2\Gamma(x) = \Gamma(x) + \Gamma^*(x) \geq \langle x,x\rangle \Rightarrow \frac{1}{2}\|x\|^2 \leq \Gamma(x) \tag{3.8}$$

But this implies by the self-duality of $\frac{1}{2}\|\cdot\|$:

$$\frac{1}{2}\|x\|^2 \geq \Gamma^*(x) = \Gamma(x) \tag{3.9}$$

Thus we have $\Gamma(x) = \frac{1}{2}\|x\|^2$ □

Remark 3.3.2. Note that self-duality for Lagrangians is slightly different. To see this take the Banach space $X = \mathbb{R}$ and the Hilbert space $H = \mathbb{R}^2 = X \times X^*$. The Lagrangian L is self dual if holds:

$$\forall (x, p) \in \mathbb{R}^2 : L^* (p, x) = L (x, p)$$

while the Hilbert space self-duality means:

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

Example 3.3.3. The Lagrangian

$$M_\lambda : X \times X^* \to \mathbb{R} \qquad (x, p) \mapsto \frac{\|x\|^2}{2\lambda} + \frac{\lambda\|p\|^2}{2}, \lambda > 0$$

This Lagrangian is self-dual.

Proof. I show, that M_λ is a simple self-dual Lagrangian generated by:

$$f (x) := \frac{\|x\|^2}{2\lambda}$$

Note at first, that f proper, convex and lsc. Thus it must only be proven, that $f^* (p) = \frac{\lambda\|p\|^2}{2}$. Take $x \in X$:

$$\begin{aligned} 0 &\leq \left(\|\sqrt{\frac{1}{2\lambda}} x\| - \|\sqrt{\frac{\lambda}{2}} p\| \right)^2 \\ &\Rightarrow \|x\|\|p\| - \|\sqrt{\frac{\lambda}{2}} p\|^2 \leq \|\sqrt{\frac{1}{2\lambda}} x\|^2 \\ &\Rightarrow \langle x, p \rangle - \frac{\lambda}{2}\|p\|^2 \leq \frac{1}{2\lambda}\|x\|^2 \end{aligned}$$

It was proven, that the half-space $H_{p,\beta}$ with $\beta = \frac{\lambda\|p\|^2}{2}$ lies completely below the epigraph of f. But for the L-F-D one also has to show, that the value β is optimal, in the sense, that for any smaller translation $\alpha \in \mathbb{R}_{\leq \beta}$ the half space $H_{p,\alpha}$ intersects the interior of the epigraph of f. This can easily be seen by inserting $x := J (\lambda p)$ in the definition of the L-F-D. Thus we have $\left(\frac{\|x\|^2}{2\lambda} \right)^* (p) = \frac{\lambda\|p\|^2}{2}$ and by Proposition 3.2.1 the self-duality of M_λ. □

3.3.1 Path spaces $L^p_X ([0, T])$

In order to solve problems, which involve time dependence, we have to lift the above definitions to path spaces. I want to emphasize, that the main ideas remain valid, and only the Banach space, on which we have developed the above calculus has to adopted to involve time dependence.

Legendre transform on path spaces

Definition. For any $p, q \in \mathbb{R}, \frac{1}{p} + \frac{1}{q} = 1$ and any time-dependent Lagrangian L one can associate a path-Lagrangian $\mathfrak{L} : L^p_X \times L^q_{X^*} \mapsto \mathbb{R}$

$$\mathfrak{L}(u, p) = \int_0^T L(t, u(t), p(t))\, dt$$

The Fenchel dual of $\mathfrak{L}$ is defined as

$$\mathfrak{L}^*(q, v) = \sup \left\{ \int_0^T \left[\langle q(t), u(t) \rangle + \langle p(t), v(t) \rangle - L(t, u(t), p(t)) \right] dt \right. \\ \left. (u, p) \in L^p_X \times L^q_{X^*} \right\}$$

Proposition 3.3.1. *Let L be a self-dual time-dependent Lagrangian. Then $\mathfrak{L}^*(p, u) = \int_0^T L^*(t, p(t), u(t))\, dt$ is self-dual.*

Proof. The proof could be found in [Ghoussoub 2009]. □

3.3.2 Evolution triples

In this section it is shown how to lift Lagrangians in the case of evolution triples

Lemma 3.1. *Let (X, H, X^*) be an e. t. and L self-dual. Assume:*

1. *$\forall x \in X : L(x, \cdot)$ is continuous*

2. *$\exists x_0 \in X : p \to L(x_0, p)$ is bounded on bounded sets of X^*.*
   Then the lifted Lagrangian

   $$M(x, p) = \begin{cases} L(x, p) & x \in X \\ +\infty & x \in H \setminus X \end{cases}$$

   is self-dual on $H \times H$.

Proof. The proof can be found in [Ghoussoub 2009, Lemma 3.4]. □

In this chapter it was shown, that in the case of convex funtionals, there exists a fundamental connection between the subdifferential and a variational problem due to the L-F-D, in which subdifferentiability is the equality case. With this mathematical concept it is possible to tackle the main part of Ghoussoub's principle, which exploits the above facts in a very tricky way. This will be done in the next chapter.

CHAPTER 4

THE VARIATIONAL PRINCIPLE OF GHOUSSOUB

Contents

In this chapter I introduce the variational principle of Ghoussoub, which ensures the existence of a zero for self-dual Lagrangians L. First of all, I will show the connection between the calculus of variation and the solvability of PDEs by referring to a trick a Auchmuty. Then I introduce Ghoussoub's principle in the case, that the Lagrangian is simple to show what the trick behind it is. After that, I state the main Theorem of Ghoussoub for general self-dual Lagrangians.

4.0.3 Connection to PDEs

Assume you could write $A := \partial f$. Then the basic differential inclusion PDE (1.1), becomes

$$F(u) + p \in \partial f(u) \tag{4.1}$$

where $F \in \mathcal{C}(X, X^*)$ and $f : X \to \mathbb{R}$ is a proper function.
Instead of showing solvability directly, Auchmuty has proposed a method [Auchmuty 1987], to transform this problem into a problem of the calculus

of variation. In order to do that, he has essentially applied the following trick:

1. From the Fenchel inequality, one knows

$$f(u) + f^*(p) \geq \langle u, p \rangle \text{ with equality if and only if } p \in \partial f(u) \tag{4.2}$$

2. Define $I(u) = f(u) + f^*(F(u) + p) - \langle u, F(u) + p \rangle$

3. From 1. follows, that u solves equation (4.1) if and only if $\alpha := \inf_{u \in X} I(u) = 0$ and the infimum is attained by some element $\overline{u} \in X$.

   Proof. Assume $\overline{u} \in X$, then

$$\begin{aligned} & I(\overline{u}) = 0 \\ \overset{\text{def}}{\Leftrightarrow} \quad & f(\overline{u}) + f^*(F(\overline{u}) + p) - \langle \overline{u}, F(\overline{u}) + p \rangle = 0 \\ \overset{\text{eq. in 1.}}{\Leftrightarrow} \quad & F(\overline{u}) + p \in \partial f(\overline{u}) \end{aligned}$$

□

Thus solving the very general equation (4.1) is equivalent to proving

$$0 = \alpha = \min_{u \in X} I(u) \tag{4.3}$$

This latter problem was studied by Optimization Theory. One key result is, that the problem (4.3) has a minimum provided f is convex and lower semi-continuous [1] , and I has some boundedness conditions. But the minimum, even if it exists, need not be equal to zero. The framework of Ghoussoub tries to ensure such property for a vast class of PDEs: the ones which allows the formulation with convex self-dual Lagrangians.

Example 4.0.4. The simple Lagrangian associated to the above equation is:

$$L(x, p) := \begin{cases} f(x) + f^*(p) & (x, p) \in \operatorname{dom}(f) \times (\operatorname{im}(F + p) \cap \operatorname{dom}(f^*)) \\ +\infty & \text{otherwise} \end{cases}$$

where $f : \operatorname{dom}(f) \subset X \to \mathbb{R}$ is proper and convex. One has then, of course, to take the infimum over both spaces.

[1] Henceforth, this will be abbreviated with lsc

Example 4.0.5 (Poisson equation). This example gives an intuition of what the main idea is related to. Assume you want to solve the Poisson equation:

$$\Delta u = h, \qquad u \in H_0^1(\Omega), h \in L^2(\Omega)$$

on a bounded smooth domain $\Omega \subset \mathbb{R}^n$. To apply Auchmuty's approach choose:

$$F = 0$$
$$f(u) = \frac{1}{2}\langle -\Delta u, u\rangle_{L^2} = \langle u, u\rangle_\Delta = \frac{1}{2}\|u\|_\Delta^2$$

This functional defines a norm, because of the positivity and the injectivity of the Laplace operator. The L-F-D dual f^* of f is:

$$f^*(p) = \frac{1}{2}\left\langle -\Delta^{-1}p, p\right\rangle_{L^2}$$

This can be shown by calculation. Denote $p = -\Delta z$:

$$0 \leq \frac{1}{2}\|z - x\|_\Delta = \frac{1}{2}\left(\langle z, z\rangle_\Delta - 2\langle x, z\rangle_\Delta + \langle x, x\rangle_\Delta\right) = f^*(p) - \langle x, p\rangle + f(x)$$

What does I look like?

$$\begin{aligned} I(u) &= \frac{1}{2}\|u\|_\Delta^2 + \|(-\Delta)^{-1}(-h)\|_\Delta^2 - \left\langle u, (-\Delta)^{-1}(-h)\right\rangle_\Delta \\ &= \frac{1}{2}\|u - (-\Delta)^{-1}(-h)\|_\Delta^2 \end{aligned}$$

But this is equivalent to the Poisson equation. Thus Auchmuty has just "completed the square".

4.1 Ghoussoub variational principle - the simple setting

In this section I will introduce the variational principle of Ghoussoub. The main idea can be traced back to Auchmuty [Auchmuty 1987].

4.1.1 The Variational Principle

Definition. A function $f : X \mapsto \overline{\mathbb{R}}$ is called *weakly coercive* iff

$$f(u) \to \infty \text{ as } \|u\| \to \infty$$

Assume you have a convex lsc function $f : X \mapsto \overline{R}$. We know, from [Zeidler 1995, Theorem 2.D], that this function has a minimum provided it is weakly coercive. Having a minimum is equivalent to

$$\exists \overline{x} \in X : 0 \in \partial f(\overline{x})$$

Proposition 4.1.1. *Assume $f^* : X^* \mapsto \overline{\mathbb{R}}$ is bounded on a neighborhood V of zero. Then f is weakly coercive.*

Proof. Let $x \in X$. Take $M := \sup_{v \in V} f^*(v)$. Choose $\delta \geq 0$ such that $p := \delta J(x) \in V$. $\forall \lambda \geq 0$ holds:

$$f(\lambda x) \geq \langle \lambda x, p \rangle - f^*(p) \geq \delta \lambda \|x\|^2 - M$$

□

But this implies for $\lambda \to \infty : \varphi(\lambda x) \to \infty$. Since x was arbitrary, φ is weakly coercive.

Remark 4.1.1. Under the above conditions, one can define the following functional:

$$I : X \to \mathbb{R}, x \mapsto f(x) + f^*(0)$$

and I has a zero.

Proof. This is because:

$$f^*(0) = \sup_{x \in X} -f(x) = -\inf_{x \in X} f(x) = -f(\overline{x})$$

□

4.1.2 Stationary existence results

Ghoussoub has to show, that there exists an useful framework, which can ensure, that I really attains its minimum and that this minimum is zero. In order to do so, he uses some basic results from the theory of convex optimization. I will henceforth consider the case, where one wants to solve the PDE

$$0 \in \partial f$$

with f being proper, lsc and convex. In this class are all monotone, linear, symmetric operators, because they are the subdifferential of the functional:

$$f : \mathrm{dom}(f) \subset X \to \mathbb{R} : x \mapsto \frac{1}{2} \langle x, Ax \rangle$$

For those PDEs the associated Lagrangian $L(x, p)$ is just the simple self-dual Lagrangian of f.

Theorem 4.1 (Existence in the simple stationary case)**.** *Let $L(x,p)$ be the simple self-dual Lagrangian of f and assume further that there exists $x_0 \in X$, such that $p \mapsto L^*(x_0,p)$ is bounded on a ball around $0 \in X^*$. Then there exists $\overline{x} \in X$ such that:*

$$0 \in \partial f(\overline{x})$$

There are two ways to proof this theorem. The first one emphasizes the PDE site of the problem and the second the convex optimization view.

Proof. PDE style One famous result of Rockafellar shows that the subderivative is an maximal monotone operator T:

$$\partial f(x) = Tx$$

Thus the PDE which is solved by that Lagrangian is:

$$Tx = 0$$

Now we have $\partial f^* = (\partial f)^{-1} = T^{-1}$. If one could show $0 \in \operatorname{dom}(T^{-1})$ we are done. Thus there is only left to show, that f^* is subdifferentiable at zero. But this is guaranteed by the fact, that f^* is bounded on bounded neighborhoods of 0 and thus is subdifferentiable. □

Proof. Optimization style Just apply Proposition 4.1.1 to get, that f attains its minimum and the functional I attains its infimum zero. □

In the view of the optimization problem, the main difficulty is to show, that f attains its minimum, while in the view of partial differential equations the main problem is to show that $0 \in \operatorname{dom}(T^{-1})$.
Both statements are simple corollary of the assumption, that f^* is bounded. In the view to the PDE view Ghoussoub is just "completing the square" according to Auchmuty's, see Example 4.0.5. That is to say, that he has just shown that the norm of the PDE is zero, if it has a solution. But that the PDE has a solution is a well hidden, simple corollary of the boundedness assumption.

4.2 The maximal monotone case

In this section the variational principle is brought in the more general context of maximal monotone operators.

Convex analysis preliminaries

Proposition 4.2.1. *Let* $L : X \times X^* \to \mathbb{R}$ *be a convex lsc Lagrangian. Define*

$$h : X^* \to \mathbb{R}, p \mapsto \inf_{x \in X} L(x, p)$$

If h *is subdifferentiable at zero, then:*

$$\exists \overline{x} \in X : -h(0) = L^*(0, \overline{x})$$

Proof. Assume $\overline{x} \in \partial h(0)$

$$h(p) - h(0) \geq \langle p, \overline{x} \rangle \Leftrightarrow L(x, p) \geq \langle \overline{x}, p \rangle + \langle x, 0 \rangle + h(0)$$

The case $p = 0$ shows, that $h(0)$ is optimal.
Consider the definition of the L-F-D of L:

$$L(x, p) \geq \langle \overline{x}, p \rangle - L^*(0, \overline{x})$$

and L^* is optimal.

By comparison one actually has:

$$-h(0) = L^*(0, \overline{x})$$

□

Remark 4.2.1. Geometric interpretation
This proof has a very simple geometric interpretation. If h is subdifferentiable, the infimum of $L(\cdot, p)$ lies for each p completely above the half-space $H_{\overline{x}, -h(0)}$. Thus the best approximating translation of the hyperplane with 'slope' $\overline{x}$ is $-h(0)$. But since $L^*(\overline{x}, 0)$ is the negative of the best approximating translation, we get the equality.

Definition (Symmetric vector field)**.** Let L be a Lagrangian on $X \times X^*$. Let $\overline{\partial} L(x)$ denote the *symmetric vector field*, which is defined as

$$\overline{\partial} L(x) = \{p, (p, x) \in \partial L(x, p)\}$$

Example 4.2.1. Consider a simple self-dual Lagrangian L. The symmetric vector field is just:

$$\overline{\partial} L(x) = \partial f(x)$$

Theorem 4.1 (Existence in stationary case)**.** *Let L be a (partially) self-dual Lagrangian and assume further that there exists $x_0 \in X$, such that $p \mapsto L(x_0, p)$ is bounded on a ball around $0 \in X^*$. Then there exists $\overline{x} \in X$ such that*

$$L(\overline{x}, 0) = \inf_{x \in X} L(x, 0) = 0 \tag{4.4}$$

$$0 = \overline{\partial} L(\overline{x}) \tag{4.5}$$

$$\tag{4.6}$$

Remark 4.2.2. For a closed (l.s.c) convex function, like L, it is known, that the above imposed boundedness condition is equivalent to $p \mapsto L(x_0, p)$ being continuous at a neighborhood of 0. [Roberts and Varberg 1973, see]
Due to self-duality, this is equivalent to $p \mapsto L^*(p, x_0)$ being continuous at 0.

Proof. First I give a short sketch

1. Show, the equivalence of Equation 4.4 and Equation 4.5. To show this, employ the L-F-D.

2. Show, that the inifimum is attained at some $\overline{x}$. Use the self-duality to show, that the infimum is zero.

Define the functional I

$$I(x) := L(x, 0)$$

1. First of all show, that $I \geq 0$

   Proof.

$$\begin{aligned} 2I(x) &= 2L(x,0) = L(x,0) + L^*(0,x) \\ &\geq \langle (x,0), (0,x) \rangle \\ &= 0 \end{aligned}$$

□

Next show:

$$I(\overline{x}) = 0 \Leftrightarrow 0 \in \overline{\partial} L(x)$$

Proof. This follows from the above proof in the case of equality and the properties of the L-F-D. □

We are now left with an variational problem.

2. Define the functional:

$$h(p) := \inf_{x \in X} L(x, p)$$

From the first step follows:

$$0 \leq h(0)$$

Since $L(x_0, \cdot)$ is bounded above for a neighborhood of zero, h is bounded above in a neighborhood of zero and is thus subdifferentiable at zero. Employ Proposition 4.2.1 to get:

$$-h(0) = L^*(0, \overline{x})$$

for some $\overline{x} \in X$. But from (partially) self-duality of L and Step 1 follows:

$$0 \leq h(0) = -L^*(0, \overline{x}) = -L(\overline{x}, 0) \leq 0$$

Thus is follows:

$$L(\overline{x}, 0) = 0$$

The infimum is attained and zero.

We have now proved the central existence result of Ghoussoub. For further details, see [Ghoussoub 2009, Theorem 6.1]

□

Remark 4.2.3. Almost the same result was stated by Rockafellar 1967 (e.g. [Zeidler 1985, Theorem 52.A and Corollary 52.2], but with one main difference. It could not be shown, that the solution to the minimum problems are in the symmetric vector field.

4.3 What is Ghoussoub doing?

Assume henceforth, that the condition of Theorem 4.1 hold. According to Remark 4.2.3, this means that the Slater conditions hold, what is equivalent to ([Zeidler 1985, see]

$$\exists u, v \in X : L(u, 0) + L^*(0, v) = L(u, 0) + L(v, 0) = 0$$

But since $\forall x \in X : L(x, 0) \geq 0$ it follows:

$$L(u, 0) = 0 \text{ and } L(v, 0) = 0$$

Thus $0 \in \partial\overline{L}(u,0)$ and $0 \in \overline{\partial}L(v,0)$. Ghoussoub just uses a very old Theorem of Rockafellar and hides the Slater conditions in the boundedness condition.
In the next section, I show, that one do not even need this result of Rockafellar. There is a far more simple way to ensure the existence of solution of such PDEs. The problem is with the assumption Ghoussoub makes, because self-duality is not very intuitive.

Simple proof

As stated Ghoussoub's main theorem for the treatable problems shows the equivalence of maximal monotone operators and the symmetric derivative set self-dual Lagrangians. See Theorem 5.1. In this section is shown, that the boundedness assumption of Theorem 4.1 in a very simple way equivalence to the solvability of PDEs. This has been already shown for the simple case via Theorem 4.1.

Proposition 4.3.1. *Let $T : X \to X^*$ be maximal monotone. Then the function:*

$$H(x,p) := \begin{cases} \langle x,p\rangle & (x,p) \in \mathit{graph}(T) \\ +\infty & \mathit{otherwise} \end{cases}$$

is convex (graph(T)) and proper

Proof. Take $0 < \lambda < 1$ Assume $(x_0,p_0),(x_1,p_1) \in \operatorname{graph}(T)$

$$\begin{aligned} & H(\lambda x_0 + (1-\lambda)x_1, \lambda p_0 + (1-\lambda)p_1) - \lambda H(x_0,p_0) - (1-\lambda)H(x_1,p_1) \\ & = \lambda^2 \langle x_0,p_0\rangle + (1-\lambda)^2 \langle x_1,p_1\rangle - \lambda(1-\lambda)(\langle x_1,p_0\rangle + \langle x_1,p_0\rangle) \\ & - \lambda\langle x_0,p_0\rangle - (1-\lambda)\langle x_1,p_1\rangle \\ & = \left(\lambda^2 - \lambda\right)(\langle x_1,p_1\rangle + \langle x_0,p_0\rangle - \langle x_1,p_0\rangle - \langle x_0,p_1\rangle) \\ & = \left(\lambda^2 - \lambda\right)(\langle x_0 - x_1,p_0\rangle - \langle x_0 - x_1,p_1\rangle) \leq 0 \end{aligned}$$

The properness follows because the graph of a maximal monotone operator is never empty. □

Proposition 4.3.2. *If $(x,p) \in \mathit{graph}(T) : (p,x) \in \partial H(x,p)$, otherwise $\partial H(x,p) = \emptyset$*

Proof. If $(x,p) \notin \operatorname{graph}(T)$ everything is clear, since $H(x,p) = +\infty$.
Assume thus $(x,p) \in \operatorname{graph}(T)$. We have to show:

$$\forall (u,v) \in X \times X^* : H(u,v) - H(x,p) \geq \langle p, u - x\rangle - \langle x, v - p\rangle$$

$$\begin{aligned} & H(u,v) - H(x,p) - \langle x, v-p\rangle - \langle p, u-x\rangle \\ &= \langle u,v\rangle - \langle x,p\rangle - \langle x, v-p\rangle - \langle p, u-x\rangle \\ &= \langle v-p, u-x\rangle \geq 0 \end{aligned}$$

□

Example 4.3.1. Assume T is single-valued, invertible, linear and monotone. Then $H(x,p) = \frac{1}{2}\langle Tx, x\rangle + \frac{1}{2}\langle T^{-1}p, p\rangle$ if $(x,p) \in \operatorname{graph}(T)$ and thus $(Tx, T^{-1}p) \in \partial H(x,p)$.
If T is further symmetric, it is easy to show that:

$$f(x) := \begin{cases} \frac{1}{2}\langle x, Tx\rangle & x \in \operatorname{dom}(T) \\ +\infty & \text{otherwise} \end{cases}$$

$$f^*(p) := \begin{cases} \frac{1}{2}\langle T^{-1}p, p\rangle & p \in \operatorname{im}(T) \\ \infty & \text{otherwise} \end{cases}$$

If f is lsc, this is the simple Lagrangian case, which has been treated above.

Proof. Assume $x \in \operatorname{dom}(T)$ and $p = Tz$. Then:

$$\begin{aligned} f^*(p) + f(x) - \langle x,p\rangle &= \frac{1}{2}\langle x, Tx - Tz\rangle + \frac{1}{2}\langle z - x, Tz\rangle \\ &= \frac{1}{2}\langle z - x, Tz - Tx\rangle \\ &\geq 0 \end{aligned}$$

In all other cases, this inequality holds trivially. Note that, for $p = Tx$ equality holds. □

Ghoussoub assigns now to every maximal monotone operator, the self-dual Lagrangian L_s (see [Ghoussoub 2009, Theorem 5.1])

$$\begin{aligned} L_s(x,p) := \inf \Big\{ & \frac{1}{2}H^*(x_1,p_1) + \frac{1}{2}H^{**}(p_2,x_2) + \frac{1}{8}\|x_1 - x_2\|^2 + \frac{1}{8}\|p_1 - p_2\|^2 \\ & , (p,x) = \frac{1}{2}((x_1,p_1) + (x_2,p_2)) \Big\} \end{aligned}$$

where H is defined as above.

Proposition 4.3.3. *Suppose $p \mapsto L_s(x_0, p)$ is bounded on some neighborhood V of zero. There is a neighborhood $U \subset V$ s.t. $H^*(x_0, U)$ is bounded.*

Proof. We begin with the definition of L_s.

$$L_s(x,p) := \inf\left\{\frac{1}{2}H^*(x_1,p_1) + \frac{1}{2}H^{**}(p_2,x_2) + \frac{1}{8}\|x_1 - x_2\|^2 + \frac{1}{8}\|p_1 - p_2\|^2 \right.$$
$$\left. ,(p,x) = \frac{1}{2}\left((x_1,p_1) + (x_2,p_2)\right)\right\}$$

From the assumption follows, that $\exists M > 0$ s.t. $\|L_s(x_0,V)\| \leq M$.

Step 1 L_s attains its infimum for every $\forall (x_0,p) \in \{x_0\} \times V$ in a bounded region.

Proof. Note for all $\forall x_1,x_2,p_1,p_2$ with $p_2 = 2p - p_1, x_2 = 2x - x_1$ holds, because of L-F-D:

$$\frac{1}{2}H^*(x_1,p_1) + \frac{1}{2}H^{**}(p_2,x_2) + \frac{1}{8}\|x_1 - x_2\|^2 + \frac{1}{8}\|p_1 - p_2\|^2$$
$$\geq \frac{1}{2}\left(\langle x_1,p_2\rangle + \langle x_2,p_1\rangle + \|x - x_1\|^2 + \|p - p_1\|\right)$$
$$= \frac{1}{2}\left(2\langle x,p_1\rangle - \langle x_1,p_1\rangle + 2\langle x_1,p\rangle - \langle x_1,p_1\rangle + \|x - x_1\|^2 + \|p - p_1\|^2\right)$$
$$+ \|x\|^2 - 2\langle x_1,x\rangle + \|x_1\|^2 + \|p\|^2 - 2\langle p_1,p\rangle + \|p_1\|^2$$

For large enough p_1,x_1, we neglect the linear terms, because they are certainly dominated by the positive quadratic norm. Apply the Cauchy-Schwartz inequality to the remaining inner products. We thus get for large enough p_1,x_1

$$\frac{1}{2}H^*(x_1,p_1) + \frac{1}{2}H^{**}(p_2,x_2) + \frac{1}{8}\|x_1 - x_2\|^2 + \frac{1}{8}\|p_1 - p_2\|^2$$
$$\geq \frac{1}{2}\left(-2\|x_1\|\|p_1\| + \|x_1\|^2 + \|p_1\|^2\right)$$
$$= \frac{1}{2}\left(\|x_1\| - \|p_1\|\right)^2$$

But this will get bigger, than M. □

Step 2 We have thus for every $p \in V$, $(x_1,p_1),(x_2,p_2) \in X \times X^*$ with bounded norms, s.t. $\frac{1}{2}(x_1 + x_2) = x_0$ and $\frac{1}{2}(p_1 + p_2) = p$. But since we have $M \geq H^{**}(p_2,x_2) \geq H^*(x_2,p_2)$ and H^* is convex we know that H^* is above bounded on the convex hull of all such points. But this convex hull contains a neighborhood $U \subset X^*$ of zero

□

Thus instead of assuming, that $L(x_0, \cdot)$ is bounded in Theorem 4.1 assume directly, that $H^*(x_0, \cdot)$ on a neighborhood.

Proposition 4.3.4. *H is lsc on the graph of T.*

Proof. Let $(x,p) \in \operatorname{graph}(T)$. By [Zeidler 1985, Proposition 38.7], it is sufficient to show: $\forall \epsilon > 0, \exists \delta > 0$ s.t.

$$\| (u,v) - (x,p) \| \leq \delta \Rightarrow H(x,p) - \epsilon \leq H(u,v)$$

We know that H is subdifferentiable on the graph from Proposition 4.3.6. Let $(p,x) \in \partial H(x,p)$. Choose $\delta \leq \frac{\epsilon}{2(\|x\|+\|p\|)}$ Now, there is for every $(u,v) \in X \times X^*$.

$$H(u,v) \geq H(x,p) + \langle p, u-x \rangle + \langle x, v-p \rangle \geq H(x,p) - \|p\|\|u-x\| - \|x\|\|v-p\|$$
$$\geq H(x,p) - \epsilon$$

We thus have $H = H^{**}$ on the graph of T. From [Precupanu and Barbu 1986] follows, that H^{**} can just differ from H on the boundary. □

Proposition 4.3.5.

$$H^{**}(x,0) = \begin{cases} 0 & (x,0) \in \overline{graph(T)} \\ +\infty & (x,0) \in \overline{graph(T)}^c \end{cases}$$

Proof.

1. Assume $(x,0) \in \overline{\operatorname{graph}(T)}^c$. Then exists a neighborhood $U \subset \operatorname{graph}(T)^c$ and $H(U) = +\infty$. Thus for every $x_n \to x, p_n \to 0$: $\liminf H(x_n, p_n) = +\infty$.
2. Assume $(x,0) \in \overline{\operatorname{graph}(T)}$. There are $x_n \to x, p_n \to 0, (x_n, p_n) \in \operatorname{graph}(T)$. Thus $H^{**}(x,0) = \liminf H(x_n,p_n) = \liminf \langle x_n, p_n \rangle = 0$

□

Proposition 4.3.6. *If $(x,0) \in dom(H^{**})$, then is H^{**} subdifferentiable at $(x,0)$*

Proof. We show $(0,x) \in \partial H^{**}(x,0)$ We just have to consider $(u,v) \in \operatorname{dom}(H^{**})$

$$\begin{aligned} H^{**}(u,v) - H^{**}(x,0) - \langle x, v \rangle &= \langle u-x, v \rangle = \liminf \langle u_n - x_n, v_n \rangle \\ &= \liminf \langle u_n - x_n, v_n - p_n \rangle + \langle u_n, p_n \rangle - \langle x_n, p_n \rangle \\ &\geq 0 \end{aligned}$$

□

Proposition 4.3.7. *If $p \mapsto L^*(x_0, p)$ is bounded above on a neighborhood of zero, then*

1. *$h(p) := \inf_{x \in X} H^*(x, p)$ is subdifferentiable at zero.*
2. *$\exists \overline{x} \in X$ s.t. $H^{**}(\overline{x}, 0)$ bounded above.*

Proof. 1. Note that, $h(p) \leq H^*(x_0, p)$ is bounded above on a neighborhood of zero. Further follows from Theorem 5.4 [Rockafellar 1972], that h is convex and from Proposition 47.5 in [Zeidler 1985] that it is continuous at 0 and from Theorem 47.A [Zeidler 1985], that h is subdifferentiable.

2. From subdifferentiability of h follows:

$$h(p) \geq h(0) + \langle \overline{x}, p \rangle = \alpha + f(p)$$

But then follows from Proposition 2.5.5 in [Ghoussoub 2009]:

$$h^*(x) \leq f^*(x) - \alpha$$

We have

$$h^*(x) = \sup_{v \in X^*} \{\langle x, v \rangle - h(v)\} = \sup_{(u,v) \in X \times X^*} \{\langle x, v \rangle - L^*(u, v)\} = L^{**}(0, x)$$

$$f^*(x) = \sup_{p \in X^*} \{\langle x, p \rangle - \langle \overline{x}, p \rangle\} = \begin{cases} 0 & x = \overline{x} \\ +\infty & x \neq \overline{x} \end{cases}$$

We thus know that:

$$H^{**}(0, \overline{x}) \leq -\alpha$$

Note that:

$$0 \leq \inf_{x \in X} L^*(x, 0) \leq L^*(x_0, 0) < +\infty$$

□

Corollary 4.1. *There is $(\overline{x}, 0) \in X \times X^*$, s.t. $H^{**}(\overline{x}, 0) = 0$.*

Proof. From Proposition 4.3.7 follows, that there exists a neighborhood $V \subset X^*$ and some $x_0 \in X$ s.t. $L^{**}(x_0, 0)$ is bounded. Thus it follows from Proposition 4.3.5 that $H^{**}(x_0, 0) = 0$ □

Proposition 4.3.8. *Assume $H^{**}(x, 0) = 0$, then $(x, 0) \in graph(T)$*

Proof. From Proposition 4.3.6 follows, that H^{**} is subdifferentiable at $(x, 0)$. Assume $(u, v) \in \text{graph}(T)$. We know from 4.3.4, that $H^{**}(u, v) = H(u, v) = \langle u, v \rangle$

$$0 \leq H^{**}(u, v) - H^{**}(x, 0) - \langle x, v \rangle = \langle u - x, v - 0 \rangle$$

But since T is maximal monotone it follows, that $(x, 0) \in \text{graph}(T)$ □

CHAPTER 5

SELF-DUAL LAGRANGIANS

Contents

In the last chapters the concept of self-dual Lagrangians has been introduced and a variational principle showing, that they can be useful for solving PDEs. This was shown explicitly for the heat equation. Despite, the notion of self-duality seems rather artificial, and it is a worth looking at the range of PDEs, which can be tackled with that approach. In this chapter some amazing permanence properties for self-dual Lagrangians are collected allowing to construct a whole family of self-dual Lagrangians, if one has found one. Further it is shown, that in fact, the PDEs, which can be brought into variational form are those involving maximal monotone operators.

5.1 Operation on Lagrangians

There are many operations which keep self-duality.

Definition. Lets define some operations

Scaling Define for $L \in \mathcal{L}(X)\, X$ the *scaled Lagrangian* $\lambda \cdot L, \lambda \geq 0$, by

$$(\lambda \cdot L)(x, p) = \lambda^2 L\left(\frac{x}{\lambda}, \frac{p}{\lambda}\right)$$

Convolution If $L, M \in \mathcal{L}(X)$, define the convolution $L \star M$ by

$$(L \star M)(x,p) := \inf\{L(z,p) + M(x-z,p), z \in X\}$$

Addition If $L, M \in \mathcal{L}(X)$, define the sum $L \oplus M$ by

$$(L \oplus M)(x,p) := \inf\{L(x,r) + M(x,p-r), r \in X^*\}$$

Right operator shift If $L \in \mathcal{L}(X)$ and $\Gamma : X \mapsto X^*$ bounded linear operator, define the Lagrangian L_Γ by

$$L_\Gamma(x,p) := L(x, p - \Gamma x)$$

Left operator shift If ${}_\Gamma L \in \mathcal{L}(X)$ and $\Gamma : X \mapsto X^*$ an invertible operator, define the Lagrangian ${}_\Gamma L$ by

$${}_\Gamma L(x,p) := L\left(x - \Gamma^{-1} p, p\right)$$

Translation Let $L \in \mathcal{L}(X)$. Define for $y \in X$ the *translated Lagrangian* by:

$$(T^v \cdot L)(x,p) = L(x, p+v) - \langle x, v\rangle$$

Proposition 5.1.1. *Permanence properties of self-duality*
This facts are proved in [Ghoussoub 2009, Section 3.2]

1. *If $L \in \mathcal{L}(X)$, then $\lambda \cdot L \in \mathcal{L}(X)$*
2. *If $L \in \mathcal{L}(X)$, then $T^v L \in \mathcal{L}(X)$ for any $v \in X^*$*
3. *If Γ^* is skew-adjoint ($\Gamma^* = -\Gamma$), then $L_\Gamma \in \mathcal{L}(X)$*
4. *If Γ^* and Γ^{-1} are skew-adjoint, then ${}_\Gamma L \in \mathcal{L}(X)$*
5. *If $L, M \in \mathcal{L}(X)$ and $dom(L)_1 - dom(M)_1$ contains a neighborhood of the origin, then also $L \oplus M, L \star M \in \mathcal{L}(X)$*

Corollary 5.1. *Characterization of maximal monotone linear operators*
Consider a bounded positive linear operator $A : X \mapsto X^*$. Denote by J the duality mapping. The mapping

$$A + J$$

is surjective.

Proof. See also [Ghoussoub 2009, Corollary 6.2, 6.3]
By definition $J(x) = \partial q(x)$ with $q(x) := \frac{1}{2}\|x\|^2$. Define the operator:

$$\Gamma : X \to X^* \qquad x \mapsto \frac{1}{2}(Ax - A^*x)$$

This operator is clearly skew-adjoint and bounded. For any $f \in X^*$ consider the functional:

$$\psi(x) := q(x) + \frac{1}{2}\langle Ax, x\rangle - \langle x, f\rangle$$

Since ψ is obviously convex, proper and lsc. Note also, that ψ is coercive. Define the associated Lagrangian:

$$L(x, p) = \psi(x) + \psi^*(p)$$

From Proposition 5.1.1(3) follows, that L_Γ is also self-dual. In order to apply Theorem 4.1 show, that $L(x, 0)$ is coercive: Note first, that

$$(q(x) - \langle x, f\rangle)^*(p) = q(p - f)$$

by [Ghoussoub 2009, Proposition 2.6]
Denote $\Phi(x) := \frac{1}{2}\langle Ax, x\rangle$. Next show, that $\Phi^*(p) \geq 0$. Assume the opposite for $p \in X^*$, then

$$0 = \Phi(0) \geq -\Phi^*(p) > 0$$

Since the L-F-D of the sum of two functionals is the infimal convolution holds (Proposition 3.2.4(5)):

$$\psi^*(p) \geq 0$$

and $L_\Gamma(x, 0)$ is coercive. Apply now 4.1 to L_Γ to get an element $\overline{x}$ such that:

$$0 = L_\Gamma(\overline{x}, 0) = L(\overline{x}, -\Gamma\overline{x}) = L^*(-\Gamma\overline{x}, \overline{x}) = -2\langle\Gamma\overline{x}, \overline{x}\rangle$$

The last equality holds, since Γ is skew-adjoint. From this follows by L-F-D:

$$-\Gamma\overline{x} \in \partial\psi(\overline{x})$$

But this is nothing else then:

$$\frac{1}{2}(A^* - A) \in J(x) + \frac{1}{4}\partial(\langle A^*x, x\rangle + \langle Ax, x\rangle) - f$$

Thus:

$$f \in (A + J)(x)$$

□

Lemma 5.1. *Let* $L(x,p)$ *be a self-dual Lagrangian. Then* $v \in \overline{\partial} L(x)$ *if and only if* $0 \in \overline{\partial} T^v L(x,0)$

Proof. **Step 1** $\Rightarrow$. Assume $(0,x) \in \partial T^v L(x,0)$ Then $(0,x) \in \partial T^v L(x,p) = \partial L(x,p) - (p,0)$

1.

2. $\Leftarrow$. Assume $(p,x) \in \partial L(x,p)$ Then $(0,x) \in \partial L(x,p) - (p,0) = \partial L(x,p) - (p,0) = \partial T^v L(x,0)$

□

This Lemma shows, that it is enough only to consider the case:

$$0 \in \overline{\partial} L(x)$$

5.2 Accessible PDEs

As for the classical calculus of variation, one also has to rise the question for the calculus of Ghoussoub of what kind of PDEs could be described be that approach. I mean whenever you have a PDE, which you want to solve, you would like to know if it could be worth to try to find an associated Lagrangian.

Definition. *Maximal monotone extension* Consider a monotone set $A \subset X \times X^*$. The maximal monotone extension $M_A \subset X \times X^*$ is the set with the properties:

1. $A \subset M_A$
2. M_A is monotone
3. If $\overline{M}_A$ is a monotone set containing A, then $\overline{M}_A \subset M_A$

Theorem 5.1 (Equivalence of self-dual Lagrangian and maximal monotone operators)**.** *For every self-dual Lagrangian* L *on* $X \times X^*$ *there exists an maximal monotone operator* T*, such that* $Tx = \overline{\partial} L(x)$*. And vice versa, for every maximal monotone operator* $T : X \mapsto X^*$ *there exists a self-dual Lagrangian* L *on* $X \times X^*$ *with* $\overline{\partial} L(x) = Tx$

Remark 5.2.1. One can also show, for every monotone operator $T : X \mapsto X^*$ there exits a self-dual Lagrangian L, such that $\overline{\partial} L$ is the maximal monotone extension of T

Proof. [see Ghoussoub 2009, Theorem 5.1] □

In this chapter the scope of self-dual Lagrangians was shown and it was stated, for which type of PDEs it is possible to search for an according Lagrangian for which the variational principle is applicable, those PDEs involving monotone operators. Let me further emphasize, once you have found a self-dual Lagrangian, there are many operations under which self-duality remains, but the solved equation differs.

CHAPTER 6

TIME DEPENDENT LAGRANGIANS

Contents

In this section, I will introduce the variational principle of Ghoussoub for time-dependent Lagrangians leading to two additional existence Theorems. Since the stationary existence theorem holds for general reflexive Banach spaces, it can be lifted to the time-dependent case just by choosing suitable Banach spaces with the according duality pairing. For these Ghoussoub uses the notion of the evolution triple.

6.1 Second existence result

6.1.1 Path space

For further details see [Ghoussoub 2009, Section 3] Let H denote a Hilbert space. Consider the space:

$$A_H^2 := \left\{u : [0,T] \to H, \dot{u} \in L_H^2\right\}$$

with the norm:

$$\|u\|_{A_H^2} := \left(\|u(0)_H\| + \int_0^T \|\dot{u}\|^2 dt \right)^{\frac{1}{2}}$$

Identify A_H^2 with $H \times L_H^2$ by:

$$u \mapsto (u(0), \dot{u})$$

Identify now $(A_H^2)^*$ also with $H \times L_H^2$ by the formula:

$$\langle u, (p_1, p_0) \rangle = \langle u(0), p_1 \rangle_H + \int_0^T \langle \dot{u}(t), p_0(t) \rangle \, dt$$

with $(p_1, p_0) \in H \times L_H^2$.

Theorem 6.1. *Let L be a time-dependent Lagrangian on a Hilbert space and ℓ a proper convex lsc function on $H \times H$. Consider the Lagrangian $\hat{L}$ on $A_H^2 \times (A_H^2)^* = A_H^2 \times (H \times L_H^2)$ defined by:*

$$\hat{L}(u, p) = \int_0^T L(t, u(t) - p_0(t), -\dot{u}(t)) \, dt + \ell(u(0) - p_1, u(T))$$

Then holds for any $(v, q) \in A_H^2 \times (0, L_H^2)$:

$$\hat{L}^*(q, v) = \int_0^T L^*(t, -\dot{v}(t), v(t) - q_0(t)) \, dt + \ell^*(-v(0), v(T))$$

Proof. See [Ghoussoub 2009, Theorem 2.4] □

Remark 6.1.1. Theorem 6.1 shows, that $\hat{L}$ is partially self-dual on the space $A_H^2 \times (A_H^2)^*$ provided L is self-dual and $\ell^*(x, p) = \ell(-x, p)$

Theorem 6.2. *Let $L : [0, T] \times H \times H \mapsto \mathbb{R}$ be a time-dependent self-dual Lagrangian and $\ell : H \times H \mapsto \mathbb{R}$ a convex lsc function satisfying $\ell(x, p) = \ell^*(-x, p)$. Assume further:*

A1 $\int_0^T L(t, x(t), 0) \, dt \leq C_1 \left(1 + \|x\|_{L_H^2}^2\right) \forall x \in {L_H^2}^2$

A2 $\ell(a, 0) \leq C_2 (1 + \|a\|_H^2)$

Then there exists $v \in \mathcal{H}_{2,2}[0, T]$, such that $(v, \dot{v}) \in dom(L)$ for almost all $t \in [0, T]$

1. $-\dot{v}(t) \in \overline{\partial} L(t, v(t))$ *(PDE)*
2. *v minimizes:* $\int_0^T L(t, v(t)(t), -\dot{v}(t))\, dt + \ell(v(0), v(T)) = 0$ *(Variational Problem)*

Proof. The main idea is to choose suitable spaces in a way, that one can apply Theorem 4.1. That includes showing partially self-duality of L and ensuring that the boundedness conditions hold.

1. The Lagrangian

$$\hat{L}(u, p) := \int_0^T L(t, u(t) - p_0(t), -\dot{u}(t))\, dt + \ell(u(0) - p_1, u(T))$$

is partially self-dual on A_H^2 by Remark 6.1.1 and the assumptions on ℓ.

2. Show that $\hat{L}(0, p)$ is bounded above.

$$\hat{L}(0, p) = \int_0^T L(t, -p_0(t), 0)\, dt + \ell(-p_1, 0) \leq C\left(1 + \|p_0{}^2_{L_H^2}\|\right) + \|p_1\|_H^2$$

by A1 and A2.

3. Thus one can apply Theorem 4.1 to the Lagrangian $\hat{L}$ to $v \in A_H^2$ such that:

$$0 = \hat{L}(v, 0) = \int_0^T L(t, v(t), -\dot{v}(t) + \langle v(t), \dot{v}(t) \rangle)\, dt$$
$$+ \ell(v(0), v(T)) - \frac{1}{2}\left(\|v(T)\|^2 - \|v(0)\|^2\right) \quad (6.1)$$

4. By self-duality of L:

$$L(t, x(t), p(t)) \geq \langle x(t), p(t) \rangle$$

and $\ell(x, p) = \ell^*(-x, p)$ together with L-F-D:

$$\ell(v(0), v(T)) \geq \frac{1}{2}\left(\|v(T)\|^2 - \|v(0)\|^2\right) \quad (6.2)$$

Equation 6.1 , together with the boundedness conditions 6.2 leads to:

$$L(t, v(t), -\dot{v}(t)) = \langle v(t), -\dot{v}(t) \rangle$$

Thus by L-F-D:

$$-\dot{v}(t) \in \overline{\partial} L(t, v(t))$$

For further details, see [Ghoussoub 2009, Theorem 7.1] □

Proposition 6.1.1. *For any $v_0 \in H$ choose $\ell : H \times H \to \mathbb{R}$ by:*

$$\ell(a,b) := \frac{1}{2}\|a\|^2 - 2\langle v_0, a\rangle + \|v_0\|^2 + \frac{1}{2}\|b\|^2$$

Then:

$$\ell^*(a,b) = \ell(-a,b)$$

Proof. Just check the basic properties of Proposition 3.2.4(6,7). Write ℓ in the following form:

$$\ell(a,b) = f^*_{2v_0}(a) + f(b) + \alpha$$

with $f(a) = \frac{1}{2}\|a\|^2$ and $\alpha := \|v_0\|^2$. Then:

$$\begin{aligned}
\ell^*(a,b) &= f_{2v_0}(a) + f(b) - \alpha \\
&= \frac{1}{2}\|a - 2v_0\|^2 - \|v_0\|^2 + f(b) \\
&= f(a) - 2\langle a, v_0\rangle + \|v_0\|^2 + f(b) \\
&= f^*_{2v_0}(-a) + f(b) + \alpha = \ell(-a,b)
\end{aligned}$$

□

Corollary 6.1. *Under Assumptions of Theorem 6.2 choose for any $v_0 \in H$ ℓ according to Proposition 6.1.1 to an element $v \in A^2_H$ such that:*

$$\begin{aligned}
v(0) &= v_0 \\
-\dot{v}(t) &\in \overline{\partial} L(t, v(t))
\end{aligned}$$

6.2 Third existence result

Until now, the variational principle is only applicable in the case, that L is defined on a Hilbert space. Often, like for the Laplace-Operator, we are interested in mappings from a Banach space into its dual. Those cases are not treatable with the above results. But there is a way to solve this problem, in the case that (X, H, X^*) is an evolution triple. With Lemma 3.1 it is always possible to lift a self-dual Lagrangian L on $X \times X^*$ to a Lagrangian on $H \times H$, but the boundedness assumptions are lost. Thus the main existence Theorems are not applicable to that Lagrangian. Since the domain of L is dense in the domain of the lifted Lagrangian, we can try to smooth it to satisfy the boundedness conditions. If we could show, that the smoothed solution converges to a solution of the original problem, we are done. This is done in this section.

Corollary 6.2. *Regularization Let $L : X \times X^* \mapsto \mathbb{R}$ be a self-dual Lagrangian. Consider the Lagrangians. Consider the Lagrangian M_λ as defined in Example 3.3.3. Define the regularized Lagrangian L_λ*

$$L_\lambda := L \star M_\lambda$$

This Lagrangian is also self-dual and bounded on bounded sets in the first variable.

Proof. See [Ghoussoub 2009, Section 3.5]
Self-duality is due to Theorem 5.1.1. ☐

Remark 6.2.1. The Lagrangian M_λ fulfills the boundedness conditions of Theorem 6.2 for any self-dual Lagrangian L. Note, that this regularization is closely related to the Yosida approximation.

Since we are dealing with weak solutions the above mentioned convergence is always meant in the weak sense. It is a very common result, that strong boundedness implies weak convergence. The next Lemma tries to establish coercivity and boundedness results for self-dual Lagrangians, mainly showing that boundedness implies coercivity and vice versa. This fact was already exploited in Proposition 4.1.1 and the idea remains the same.

Lemma 6.1. *Let L be self-dual on $X \times X^*$. Assume there exists $C_1, C_2 > 0, r_1 \geq r_2 > 1$ such that:*

$$C_1 \left(\|x\|_X^{r_2} - 1\right) \leq L\left(x, 0\right) \leq C_2 \left(1 + \|x\|_X^{r_1}\right)$$

Then exists $D_1, D_2 > 0, s_1, s_2, \frac{1}{r_i} + \frac{1}{s_i} = 1$ such that:

$$D_1 \left(\|p\|_{X^*}^{s_1} + \|x\|_X^{r_2} - 1\right) \leq L\left(x, p\right) \leq D_2 \left(1 + \|p\|_{X^*}^{s_2} + \|x\|_X^{r_1}\right)$$

Proof. See also [Ghoussoub 2009, Lemma 3.5]

1. Boundedness implies coercivity
   By (partially) self-duality of L we get:

   $$L\left(x, p\right) \geq \langle y, p\rangle - L^*\left(0, y\right) = \langle y, p\rangle - L\left(y, 0\right) \geq \langle y, p\rangle - C_2\left(1 + \|y\|_X^{r_1}\right)$$

   Denote:

   $$f\left(x\right) := C_2\left(1 + \|y\|_X^{r_1}\right)$$

   Note by Example 3.2.3 and Proposition 3.2.4 follows:

   $$f^*\left(p\right) = C\|p\|_{X^*}^{s_1} - D$$

   for some $C, D \geq 0$ and by taking the supremum over all $y \in Y$ and L-F-D follows:

   $$L\left(x, p\right) \geq C\|p\|_{X^*}^{s_1} - D$$

□

Remark 6.2.2. The fact, that coercivity of $L(x,0)$ implies boundness for self-dual Lagrangians has also a very simple geometric interpretation, since the L-F-D is the best translation for a half-plane below the epigraph of the Lagrangian. If it is coercive, the translation of the half-spaces growth fast enough to ensure that the function itself remains bounded even for very huge "slopes".

Theorem 6.1 (Evolution triple case)**.** *Let $X \subset H \subset X^*$ be an evolution triple and consider a time-dependent Lagrangian L on $[0,T] \times X \times X^*$ and a convex lsc function ℓ on $H \times H$, which satisfy the following conditions:*

A1 *For some $p \geq 2$, $m, n > 1$ and $C_1, C_2 > 0$ holds*

$$C_1 \left(\|x\|^m_{L^p_X} - 1 \right) < \int_0^T L(t, x(t), 0)\, dt \leq C_2 \left(1 + \|x\|^n_{L^p_X} \right), \qquad \forall x \in L^p_X$$

A2 *For some $C_3 > 0$ holds:*

$$\ell(a,b) \leq C_3 \left(1 + \|a\|^2_H + \|b\|^2_H \right), \qquad \forall a,b \in H$$

A3 $\ell^*(a,b) = \ell(-a,b)$

Then the following holds for an unique $v \in \mathcal{H}_{p,q}$:

$$\begin{aligned} I(x) &= \int_0^T L(t, x(t), -\dot{x}(t))\, dt + \ell(x(0), x(T)) \\ & \textit{is self-dual on } \mathcal{H}_{p,q} \\ -\dot{v}(t) &\in \overline{\partial} L(t, v(t)) \\ (-v(0), v(T)) &\in \partial \ell(v(0), v(T)) \end{aligned}$$

Proof. Outline of the proof:

1. Lift the Lagrangian on H
2. Consider a regularized problem and show solvability
3. Establish boundness of the solutions of the regularized problem, because boundness implies weak convergence
4. Show that the weak limit of the regularized problem is indeed a solution of the original problem

1. According to Lemma 3.1 lift the Lagrangian L to a Lagrangian $\overline{L}$ on H.

2. Consider the regularization of $\overline{L}$:

$$\overline{L}_\lambda := \overline{L} \star M_\lambda$$

According to Corollary 6.2 this Lagrangian satisfies the Assumption of Theorem 6.2. Thus there exists $v_\lambda \in A_H^2$ such that:

$$\int_0^T L_\lambda \left(t, v_\lambda (t), -\dot{v}_\lambda (t)\right) dt + \ell \left(v_\lambda (0), v_\lambda (T)\right) = 0$$

3. Because of Lemma 6.1 L is coercive in both variables. Since it is lsc and convex by Assumption, there exists a minimizer $i\left(v_\lambda\right)$ in the definition of $\overline{L}_\lambda$:

$$\begin{aligned} \overline{L}_\lambda (x,p) &= \left(\overline{L} \star M_\lambda\right)(x,p) = \inf_{z \in H} \left\{ L(t,z,p) + \frac{\|x-z\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2 \right\} \\ &= L\left(t, i\left(v_\lambda\right), p\right) + \frac{\|x - i\left(v_\lambda\right)\|^2}{2\lambda} + \frac{\lambda}{2}\|p\|^2 \end{aligned}$$

The above equations give:

$$\int_0^T L\left(t, i\left(v_\lambda\right), -\dot{v}_\lambda\right) + \frac{\|v_\lambda - i\left(v_\lambda\right)\|^2}{2\lambda} + \frac{\lambda}{2}\|\dot{v}_\lambda\|^2 dt + \ell\left(v_\lambda(0), v_\lambda(T)\right) = 0 \tag{6.3}$$

From Proposition 3.2.4(2,3), Assumption A2 and Assumption A3

$$\ell(a,b) = \ell^*(-a,b) \geq D\left(1 + \|a\|_H^2 + \|b\|_H^2\right)$$

Thus ℓ is also coercive and $v_\lambda(0)$ and $v_\lambda(T)$ are bounded in H in order to satisfy equation (6.3). Because of the coercivity conditions established in Lemma 6.1, $i\left(v_\lambda\right)$ has to be bounded in L_X^p and $\dot{v}_\lambda$ is bounded in L^q. But then v_λ is bounded in L_H^2.

4. From the boundness and coercivity follows in equation 6.3:

$$\int_0^T \|v_\lambda - i\left(v_\lambda\right)\|^2 dt \leq 2\lambda C \tag{6.4}$$

for some $C \geq 0$. Because $i(v_\lambda)$ is bounded, there exists a subsequence converging weakly to v. Since $L^2 \subset L^q$, we can consider $u \in L^2$. Then holds:

$$\begin{aligned} |\langle u, (v_\lambda - v)\rangle| &\leq |\langle u, v_\lambda - i(v_\lambda)\rangle| + |\langle u, i(v_\lambda) - v\rangle| \\ &\leq \|u\|\|v_\lambda - i(v_\lambda)\| + |\langle u, i(v_\lambda) - v\rangle| \end{aligned}$$

But this goes to zero for λ to zero by Equation 6.4 and weak convergence of $i(v_\lambda)$. By Proposition 23.19 in [Zeidler 1990b] follows, that $\dot{v}_\lambda$ converges weakly to $\dot{v}$. Since L is lower semi continuous follows:

$$\int_0^T L(t, v(t), -\dot{v}(t))\, dt + \ell(v(0), v(T)) \leq 0$$

Because of self-duality as in Theorem 6.1 this inequality is equivalent to the functional being zero. The rest follows as in Theorem 6.2

See [Ghoussoub 2009, Chapter 7] □

These are Ghoussoub's three main existence theorems for partial differential equations.

Example: Heat equation

To use Ghoussoub's approach for the heat equation, in chapter two there somehow magically arose some kind of strange Lagrangian. Now it can be shown, where it steems from. It could be constructed by the simple self-dual Lagrangian, because the Laplace operator is strongly positive. Starting point is the Dirichlet functional

$$\phi(u) = \frac{1}{2}\int_\Omega |\nabla u(x)|^2 dx, \qquad u \in H_0^1(\Omega) \tag{6.5}$$

1. The L-F-D of ϕ is

   $$\phi^*(p) = \frac{1}{2}\int_\Omega |\nabla(-\Delta)^{-1}(p(x))|^2 dx, \qquad p \in H_0^{-1}(\Omega)$$

   as shown in Example 4.0.5.

2. Consider now the simple self-dual Lagrangian:

   $$L_{X\times X^*} : [0,T] \times H_0^1(\Omega) \times H_0^{-1}(\Omega) \to \mathbb{R}, \qquad (t,u,p) \mapsto \phi(u) + \phi^*(p)$$

Since $H_0^1(\Omega) \subset L^2(\Omega) \subset H_0^{-1}(\Omega)$ is an evolution triple and one can apply Theorem 6.1 to get the desired result. This difference between the stationary Poisson equation and the time-dependent heat equation is marginal. The only difference is the chosen Banach space and even the Lagrangian for the problem remains.

6.3 Summary

I have introduced the variational principle of Ghoussoub by showing its usefulness to solve PDEs. Important is the application of the convex calculus especially the Legendre Fenchel duality, which has also been introduced. Despite of the amazing results available by Ghoussoubs calculus, the main idea is quite simple and probably rather difficult to use for novel existence results. I have shown, that the conditions of Ghoussoub allow a much more simple framework to ensure the existence of the solution to an maximal monotone operator equation. The main point, which is most explicit, in the case of the simple self-dual Lagrangian is, the assurance of the boundedness condition. In order to do that, in most cases, you must already have shown, that the equation, which you want to solve already has a solution. Just consider the case of the heat equation. To show, that the boundedness conditions hold, you have to show, that

$$f(p) = \frac{1}{2}\| \left(-\Delta^{-1}\right) p\|^2$$

is continuous and bounded at zero, but that is the same as showing the solvability of the heat equation directly.

6.4 Outlook

Numerical mathematics

Those principles could turn out to be useful for numerical mathematics, because it allows a different numerical treatment of evolutionary PDEs. The simplest numerical approach is numerical integration, which is very bad e.g. for long times.
With this variational principle you can try a different approach to the numerics of evolutionary PDEs. To do this, take a parameter dependent function

$$\phi : \mathbb{R} \times \Omega \times \Lambda \mapsto \mathbb{R}$$

numerically and try to minimize the Lagrangian associated to the problem, you want to solve, numerically. This means solving:

$$\inf_{\lambda \in \Lambda} L\left(\phi\left(\cdot, \cdot, \lambda\right)\right)$$

You do not have to iterate to get the results for all times, because the minimum is already the approximated solutions for all times. Of course there must some error estimates established.

Non convex case

In order to apply the theory in the non-convex case, one needs a generalized notion of subdifferentiability and of conjugation. There have been many approaches to generalize the subdifferential of convex functions to non-convex settings. See for example [Auchmuty 1983, Buliga et al. 2009, Denkowski et al. 2003, Rockafellar and Wets 2004].

Physical interpretation

From a physical viewpoint self-duality needs to be interpreted. Ghoussoub has shown, that self-duality is also valid in other fundamental equations like some versions of the Yang-Mills equations or the Navier-Stokes equations. The physical content of the Lagrangians is only known in some cases. From the point of view of quantum physics, one attempt was done in [Majid 2000], in which self-duality is a interpreted as fundamental physical law. However, what the real physical meaning of self-duality is, is up to my knowledge still unknown.

BIBLIOGRAPHY

Giles Auchmuty. Duality for non-convex variational principles. *Journal of Differential Equations*, 1983.

Giles Auchmuty. Variational principles for operator equations and initial values problems. *Nonlinear Analysis, Theory, Methods and Applications*, 1987.

Heinz H. Bauschke, Xianfu Wang, and Liangjin Yao. Autoconjugate representers for linear monotone operators. Mathematical Programming manuscript.

Yves Lucet Bauschke, Rafal Goebel and Xianfu Wang. The proximal average: Basic theory. 2007.

Jonathan M Borwein. Maximal monotonicity via convex analysis.

A Brondsted and R. T. Rockafellar. On the subdifferentiablity of convex functions. *Proceedings of the American Mathematical Society*, 1965.

Marius Buliga, Gery de Saxce, and Claude Vallee. Bipotentials for non-monotone multivalued-operators: Fundamental results and applications. *Acta Appl Math*, 2009.

Regina Sandra Burachik and B.F. Svaiter. Maximal monotone operators, convex functions and a special family of enlargements.

Soon-Yeong Chung. An example of nonuniqueness of the chauhy problem for the heat equation. *Commun. in Partial Differential Equations*, 19(7): 1257–1261, 1994.

Zdzislaw Denkowski, Stanislaw Migorski, and Nikolaos Socrates Papageorgiou. *An Introduction to Nonlinear Analysis: Theory*. Springer, 2003.

Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

Nassif Ghoussoub. Anti-selfdual lagrangians: Variational resolutions of non-selfadjoint equations and dissipative evolutions. 2005.

Nassif Ghoussoub. *Self-dual partial differential systems and their variational principles*. Springer Monographs in Mathematics, 2009.

Nassif Ghoussoub and Leo Tzou. Anti-selfdual lagrangians ii: Unbounded non self-adjoint operators and evolution equations. 2005.

Nassif Ghoussoub, Abbas Moameni, and Ramon Zarate Saiz. Homogenization of maximal monotone vector fields via selfdual variational calculus. 2010.

Juergen Jost and Xianqing Li-Jost. *Calculus of Variation*. Cambridge University Press, 1999.

Shahn Majid. *Foundations of Quantum Group Theory*. Cambridge University Press, 2000.

Peter J. Olver. *Applications of Lie Groups to Differential Equations*. Springer, 1993.

Th. Precupanu and Viorel Barbu. *Convexity and Optimization in Banach Spaces*. Springer Netherlends, 1986.

Arthur Wayne Roberts and Dale Varberg. *Convex Functions*. Academic Press Inc, December 1973.

R. Tyrrell Rockafellar. *Convex Analysis*. Princton University Press, 1972.

R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*. Springer, 2004.

S. Simons and C. Zalinescu. A new proof for rockafellar's characterization of maximal monotone operators. *Proceedings of the American Mathematical Society*, 2008.

Georgi V. Smirnov. *Introduction to the Theory of Differential Inclusions*, volume 41 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

Leo Tzou and Nassif Ghoussoub. A variational principle for gradient flows. *Mathematische Annalen*, 330(3):519–549, 2004.

B. D. Vujanovic and S. E. Jones. *Variational Methods in Nonconservative Phenomena*. Academic Press, 1989.

Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications III*. Springer-Verlag, 1985.

Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications I*. Springer-Verlag, 1990a.

Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications II*. Springer-Verlag, 1990b.

Eberhard Zeidler. *Applied Functional Analysis: Main Principles and Their Applications*. Springer, 1995.