Introduction to the calculus of variations and Γ-convergence

Antonin Chambolle, CMAP, Ecole Polytechnique, CNRS

April 15, 2015

Contents

1 Introduction **5**

2 Characterization of the critical points **7**
2.1 Differentiability 7
2.2 First variation of a functional 8
2.3 Boundary conditions 10
2.3.1 The *prescribed* boundary conditions 10
2.3.2 The *free* boundary conditions 12
2.4 A remark on "Null Lagrangians" 13
2.5 Constrained problems, obstacles 14
2.5.1 Convex constraints 14
2.5.2 Differentiable constraints 15

3 Existence of minimizers **17**
3.1 Lower semicontinuity 17
3.2 The direct method in the calculus of variations 18
3.2.1 Basic principle 18
3.2.2 An example 19
3.3 Compactness results in infinite dimension 20
3.3.1 Ascoli-Arzelà 20
3.4 Basic lower-semicontinuity criteria 23
3.4.1 Convex functions 23
3.4.2 Non-convex cases 24
3.4.3 Necessity of the convexity 24

4 Regularity for minimizers (elliptic case) **27**
4.1 Basic regularity estimates 27
4.1.1 Obvious estimates 27
4.1.2 Higher derivability 29
4.2 The De Giorgi-Nash-Moser Thm 31
4.2.1 Towards more regularity 31
4.2.2 Proof: L^∞ control 32
4.2.3 Harnack-type estimate 36
4.2.4 Application: control of the oscillation of a solution 38

4.2.5 Conclusion: u is Hölder 38

5 Vector-valued minimization problems 41

6 Variational convergence 45
6.1 Why? 45
6.1.1 Necessity of the axioms 46
6.1.2 Properties and comments 47
6.1.3 Compactness of Γ-convergence 48
6.2 Application: discrete to continuum limits 50
6.2.1 Higher dimension 52
6.3 Application: periodic homogenization 54
6.3.1 Proof of the homogenization result 57
6.3.2 A remark in the convex case 63
6.3.3 Phase transitions, singular limits 64

Chapter 1

Introduction

The calculus of variations is the study of the minimizers or critical points of "functionals", which are functions defined in spaces of infinite dimensions, typically functional spaces.

This needs to adapt the notions of differential calculus. The stationarity of a functional $\mathcal{E}(u)$ is "simply" characterized by the equation

$$\mathcal{E}'(u) = 0 \tag{1.1}$$

which, in general, will be a partial differential equation (PDE) in u (or something more general).

A goal of the calculus of variations is to "solve" such PDEs: more precisely, to show that they actually have one or several solutions (or none...), study their properties, and possibly design numerical methods to compute these solutions or approximations.

A first important observation is that not all PDEs will be solve by a variational analysis: only the PDEs which are "variational", meaning that their equation is precisely of the form (1.1) for a particular $\mathcal{E}$.

Why is it interesting?

- it provides sometimes a very simple tool for showing existence of (weak) solutions to a problem;
- many PDEs come from problems in physics, mechanics, etc, and precisely from "variational" principles and are therefore (often minimizing) critical points of some physical energy.
- many problems in the industry (or finance, etc) are *designed* as finding the "best" state according to some criterion, and their solution is precisely a minimizer, or maximizer, of this criterion ("optimization").

Standard examples:

1. Laplace equation $-\Delta u = 0$ characterizes the critical points of the "Dirichlet energy" $\int |\nabla u|^2 dx$. In this case, since the energy is convex, critical points and minimizers are the same.

2. Horizontal elastic membrane subject to a vertical force f:

$$\min_{u=0 \text{ on } \partial\Omega} \frac{1}{2}\int_\Omega |\nabla u|^2 dx - \int_\Omega f u dx,$$

in this case equation (1.1) reads

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Variant with a nonlinear potential energy:

$$\min_{u=0 \text{ on } \partial\Omega} \frac{1}{2}\int_\Omega |\nabla u|^2 dx + \int_\Omega F(u) dx,$$

then the equation becomes $-\Delta u + F'(u) = 0$.

3. Nonlinear elasticity: $\mathcal{E}(u) = \int_\Omega W(\nabla u)dx$ where now $u : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^3$ is vectorial valued. W might depend on $\nabla u^T \nabla u$, $\det \nabla u$... Equation (1.1) becomes now a system.

4. Minimal surfaces, geodesics: $\mathcal{E}(u) = \int_\Omega \sqrt{1 + |\nabla u|^2} dx$ (area of the graph of u), $\mathcal{E}(u) = \int_0^1 |u'(t)|^2 \, dt$ with $u(0) = x_0$, $u(1) = x_1$, $u(x) \in \mathcal{M}$ a manifold for each x.

For each of these problems, the natural questions are: is there existence of a solution? Uniqueness? How can it be characterized? How can it be computed?

Chapter 2

Characterization of the critical points

This chapter addresses the issue of the meaning of equation (1.1).

2.1 Differentiability

Definition 1. *Let X be a Banach space and U an open subset. $\mathcal{E} : U \subset X \to \mathbb{R}$ is "Fréchet"-differentiable at $u \in U$ if there exists a continuous linear form $D\mathcal{E}(u) \in X^*$, called "differential", such that*

$$\lim_{v \to 0} \frac{|\mathcal{E}(u+v) - \mathcal{E}(u) - D\mathcal{E}(u) \cdot v|}{\|v\|_X} = 0$$

which can also be written (using Hardy's "o" notation)

$$\mathcal{E}(u+v) = \mathcal{E}(u) + D\mathcal{E}(u) \cdot v + o(\|v\|_X).$$

$\mathcal{E}$ is of class $C^1(U)$ if $u \mapsto D\mathcal{E}(u)$ is continuous.

Definition 2. *$\mathcal{E}$ is said to be Gâteaux-differentiable at $u \in U$ in the direction $v \in X$ if*

$$D_v\mathcal{E}(u) = \frac{d}{dt}\mathcal{E}(u+tv)_{|t=0}$$

exists. It is said to be Gâteaux-differentiable at u if it exists for all v.

Clearly, if $\mathcal{E}$ is (Fréchet-)differentiable at u, then it is Gâteaux-differentiable, and $D_v\mathcal{E}(u) = D\mathcal{E}(u) \cdot v$ for all v. The converse is not true. Examples for functions defined in $X = \mathbb{R}^2$:

1. $u(x,y) = \frac{x^3}{x^2+y^2}$ (and $u(0) = 0$): the Gâteaux derivative in direction $(a,b) \neq 0$ is the function $a^3/(a^2+b^2)$ which is not linear in (a,b)...

2. $u(x,y) = \frac{x^2 y}{x^4+y^2}\sqrt{x^2+y^2} \le \frac{1}{2}\sqrt{x^2+y^2}$ (and 0 in 0): the Gâteaux derivative is 0, however, if $(x,y) = (t,t^2) \to 0$ as $t \to 0$,

$$\frac{u(x,y)}{\sqrt{x^2+y^2}} = \frac{t^4}{2t^4} = \frac{1}{2} \not\to 0.$$

The Fréchet-differential $D\mathcal{E}(u)$ is also denoted $\mathcal{E}'(u)$, and by definition, a critical point u is a point where (1.1) holds.

2.2 First variation of a functional

In practice, to derive the stationarity conditions of an energy, and in particular the minimality conditions, it is enough to know how to compute directionals derivatives for all "admissible" v. Indeed, assume that u is a minimizer of $\mathcal{E}$ over a set $K \subset X$: then given v, provided $u + tv \in K$ for $t > 0$ small (which is the meaning of an "admissible variation" v) one always have $\mathcal{E}(u+tv) \ge \mathcal{E}(u)$ so that

$$\lim_{t\to 0, t>0} \frac{\mathcal{E}(u+tv) - \mathcal{E}(u)}{t} \ge 0.$$

Then, if $\mathcal{E}$ is differentiable at u, one recovers that $D\mathcal{E}(u)\cdot v \ge 0$. If both v and $-v$ are admissible, one recovers $D\mathcal{E}(u)\cdot v = 0$.

For a general theory, we will restrict ourselves to functionals of the form

$$\mathcal{E}(u) = \int_\Omega \mathcal{L}(x, u(x), Du(x)\, dx \tag{2.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open set, and $u : \Omega \to \mathbb{R}^m$ is a possibly vectorial-valued function (in some functional space). Here "Du" is the differential, identified to an $m\times n$ matrix with entries $\partial u^i/\partial x_\alpha$, $i = 1,\dots,m$, $\alpha = 1,\dots,n$ (but sometimes, especially for scalar functions, we will also use the Gradient ∇u which is a vector). The function

$$\begin{array}{rccl} \mathcal{L}: & \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n} & \to & \mathbb{R} \\ & (x,u,p) & \mapsto & \mathcal{L}(x,u,p) \end{array}$$

is called the "**Lagrangian**", and is assumed to be smooth.

Then, given u, v and $t > 0$ small, one has

$$\mathcal{E}(u+tv) = \mathcal{E}(u) + t\int_\Omega \sum_i \frac{\partial\mathcal{L}}{\partial u^i}(x,u,Du)v^i + \sum_{i,\alpha}\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\frac{\partial v^i}{\partial x_\alpha}dx + o(1),$$

yielding

$$\frac{d}{dt}\mathcal{E}(u+tv)_{|t=0} = \int_\Omega \sum_i \frac{\partial\mathcal{L}}{\partial u^i}(x,u,Du)v^i + \sum_{i,\alpha}\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\frac{\partial v^i}{\partial x_\alpha}dx.$$

Hence at a critical point, one should have for all v admissible:

$$\int_\Omega \sum_i \frac{\partial\mathcal{L}}{\partial u^i}(x,u,Du)v^i + \sum_{i,\alpha}\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\frac{\partial v^i}{\partial x_\alpha}dx = 0 \tag{2.2}$$

If all smooth $v = (v^1, \dots, v^m)$ with compact support are admissible, one can integrate by parts the last term, yielding the following form for equation (1.1)

$$\frac{\partial \mathcal{L}}{\partial u^i}(x, u, Du) - \sum_{\alpha=1}^{n} \frac{\partial}{\partial x_\alpha}\left(\frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du)\right) = 0$$

for all $i = 1, \dots, m$, at least in the distributional sense (in $\mathcal{D}'$ [13]).

Definition 3. *The Euler-Lagrange equation (or system) associated to $\mathcal{E}$, given by* (2.1), *is*

$$\frac{\partial \mathcal{L}}{\partial u^i}(x, u, Du) = \sum_{\alpha=1}^{n} \frac{\partial}{\partial x_\alpha}\left(\frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du)\right) = \operatorname{div}\left(\frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du)\right) \tag{2.3}$$

for all $i = 1, \dots, m$.

One observes that this equation, which characterizes the (smooth enough) critical points of $\mathcal{E}$, is in *divergence form*. In particular, non-divergence equations with non constant coefficients such as

$$\sum_{i,j} a_{i,j} \frac{\partial^2 u}{\partial x_i \partial x_j} = 0$$

are in general *not* variational.

Examples

1. $\mathcal{E}(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx + \int_\Omega F(u) dx$: here $\mathcal{L}(x, u, p) = |p|^2/2 + F(u)$ and one recovers the equation

$$\operatorname{div} \nabla u = \partial_u F(u)$$

(or $\Delta u = F'(u)$). In practice, the steps to derive these equations are as follow: assume u is a minizer of $\mathcal{E}$ in $H^1_0(\Omega)$ (that is, with Dirichlet boundary conditions $u = 0$ on $\partial\Omega$. Then, any variation $v \in H^1_0(\Omega)$ is admissible: for any $t > 0$ (or $t \in \mathbb{R}$), $u + tv \in H^1_0(\Omega)$ and one can write

$$\mathcal{E}(u + tv) \geq \mathcal{E}(u).$$

Hence,

$$0 \leq \frac{\mathcal{E}(u + tv) - \mathcal{E}(u)}{t} = \int_\Omega \nabla u \cdot \nabla v \, dx + \frac{t}{2}\int_\Omega |\nabla v|^2 dx + \int_\Omega \frac{F(u + tv) - F(u)}{t} dx.$$

The last term is

$$\int_\Omega \frac{1}{t}\int_0^t F'(u + sv) v \, dx$$

so that if F is C^1, it is easy to justify (using for instance Lebesgue's convergence theorem) that its limit as $t \to 0$ is precisely $\int_\Omega F'(u) v \, dx$. Hence,

$$\int_\Omega \nabla u \cdot \nabla v + F'(u) v \, dx = 0$$

for all $v \in H_0^1(\Omega)$ (or any $v \in C_c^\infty(\Omega)$ would be enough to conclude). This is precisely the *weak* form or *variational* form for $\Delta u = F'(u)$ (and it is standard that if u has two derivatives in L^2, the two forms are equivalent).

2. For $m = 1$, given a matrix $A = (a_{\alpha,\beta})_{1\le\alpha,\beta\le n}$, we let $\mathcal{L}(x,u,p) = \sum_{\alpha,\beta} a_{\alpha,\beta} p_\alpha p_\beta$. This is the Lagrangian for $\int_\Omega (A\nabla u)\cdot\nabla u dx$: then the equation is

$$\operatorname{div} A\nabla u = 0.$$

3. $\mathcal{L}(x,u,p) = (1/q)|p|^q$, $1 < q < +\infty$: the equation is the q-Laplace equation

$$\operatorname{div} |\nabla u|^{q-2}\nabla u = 0.$$

For $q = 1$, the Lagrangian is not regular enough and the signification of the limit equation

$$\text{“div}\, \frac{\nabla u}{|\nabla u|} = 0''$$

must be clarified when $\nabla u = 0$...

4. $\mathcal{L}(x,u,p) = \sqrt{1+|p|^2}$: the equation is the minimal surface (or graph) equation

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

2.3 Boundary conditions

The Euler-Lagrange equations we have derived up to now are not complete. Indeed, if one considers for instance the equation $\Delta u = 0$ (which characterizes the critical points of $\int_\Omega |\nabla u|^2 dx$), it has an infinite number of solutions: any harmonic function in Ω is a solution. However, if we are considering the problem

$$\min_{u\in X} \int_\Omega |\nabla u|^2 dx,$$

for $X = H^1(\Omega)$ (such that the functional is finite) or $X = H_0^1(\Omega)$ (that is, if we require in addition that u vanishes on the boundary), then most harmonic solutions are *not* a solution of the problem (neither a minimizer nor a critical point, which in this case are the same). Indeed, the minimal value for this problem is 0, reached for $u = 0$, while in general if u is harmonic, the integral will not vanish...

Remarks: if $X = H^1(\Omega)$, there are still infinitely many solutions, which ones? Given u a harmonic function with finite energy, what minimization problem does it solve?

2.3.1 The *prescribed* boundary conditions

The “Dirichlet” or prescribed boundary conditions are the conditions which come from restrictions imposed in the space of definition of the functional $\mathcal{E}$. They are thus, strictly speaking, not part of the Euler-Lagrange equation "$\mathcal{E}'(u) = 0$" (1.1), but in the condition "$u \in X$". Let us consider two examples: the first is the problem

$$\min_{u \in g + H_0^1(\Omega)} \int_\Omega |\nabla u|^2 dx, \tag{2.4}$$

where $g \in H^1(\Omega)$ (which is necessary to guarantee that there exists at least one function u with finite energy, otherwise the problem cannot have a solution!)

In this case, the equation which is solved by a minimier is

$$\begin{cases} -\Delta u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{2.5}$$

The first line in this equation comes from the stationarity of $\mathcal{E}$ at u, while the second is because we have prescribed that $u \in g + H_0^1(\Omega)$.

In this case, the admissible variations v are the functions $v \in H_0^1(\Omega)$ (or the dense subclass $v \in C_c^\infty(\Omega)$. Indeed for such functions, $u + tv$ is not modified on $\partial\Omega$. Note that the admissible variations do not satisfy $v = g$, but $v = 0$, on the boundary!

Warning: if the Dirichlet condition $u = g$ is not the trace of a H^1 function on $\partial\Omega$, it will precisely mean that for any u with $u = g \in \Omega$, $\int_\Omega |\nabla u|^2 dx = +\infty$. Hence the variational problem (2.4) cannot be solved. However, this does not mean that (2.5) is not solvable! For instance, there exists in the open planar disk B_1 a function $u \in C^\infty(B_1)$ (even analytic) such that $\Delta u = 0$ in B_1, $u(\cos\theta, \sin\theta) = \sin(\theta/2)$, $-\pi \le \theta \le \pi$. But there is no function in $H^1(B_1)$ with a jump discontinuity on the boundary ∂B_1.

A second example is illustrated by the following problem

$$\min_{u \in g + H_0^2(\Omega)} \int_\Omega |\Delta u|^2 dx$$

for a given $g \in H^2(\Omega)$. We recall that $H_0^2(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $H^2(\Omega)$, hence for the norm $u \mapsto \sqrt{\int u^2 + |Du|^2 + |D^2u|^2 dx}$.

The Euler-Lagrange equation for this problem is

$$\begin{cases} \Delta^2 u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \\ \nabla u = \nabla g & \text{on } \partial\Omega. \end{cases}$$

The latter condition follows because functions in $v \in H_0^2(\Omega)$ have a vanishing gradient on the boundary (if this boundary is smooth enough). Since the tangential gradient is obviously vanishing (because $v = 0$ on $\partial\Omega$), it is equivalent to write that the normal gradient $\partial v/\partial\nu = \nabla v \cdot \nu = 0$. Hence one could replace the last condition in the equation by

$$\frac{\partial u}{\partial \nu} = \frac{\partial g}{\partial \nu} \qquad \text{on } \partial\Omega,$$

however, this is still (in my opinion) a "Dirichlet" condition: it is a restriction on the space of minimization (which makes sense because functions with $\int |\Delta u|^2 < \infty$ have a well-defined gradient on a regular boundary), it is not coming from the first variation equation $\mathcal{E}'(u) = 0$!

2.3.2 The *free* boundary conditions

On the contrary, the "Neumann" or free boundary conditions are the conditions which do not arise from restrictions on u but from the equilibrium equation $\mathcal{E}'(u) = 0$, written in the variational sense $D\mathcal{E}(u) \cdot v = 0$, and the fact that the variations (or test functions) v are allowed to vary freely on a part of the boundary, so that this equation contains, in fact, an information on the behaviour of u on the boundary.

Examples

Let us start with a few examples: assume u is a solution of

$$\min_{u \in H^1(\Omega)} \int_\Omega |\nabla u|^2 + (u-g)^2 dx$$

Then considering variations $u + tv$, $v \in H^1(\Omega)$, t small, one easily finds that

$$\int_\Omega \nabla u \cdot \nabla v + (u-g) v \, dx = 0, \tag{2.6}$$

this for any $v \in H^1_0(\Omega)$. If we consider first $v \in C^\infty_c(\Omega)$, then one can integrate by parts the first half and find

$$-\int_\Omega \Delta u \, v \, dx + \int_\Omega (u-g) v \, dx = 0$$

and since this must be true for all $v \in C^\infty_c(\Omega)$, we must have

$$-\Delta u + u = g \qquad \text{in } \Omega.$$

Now, consider $v \in C^\infty(\bar{\Omega})$. Then, integrating by parts (2.6) again one finds

$$\int_{\partial\Omega} v \nabla u \cdot \nu \, d\sigma + \int_\Omega (-\Delta u + u - g) v \, dx = 0,$$

however we already *know* that $-\Delta u + u - g = 0$, so that this becomes simply

$$\int_{\partial\Omega} v \nabla u \cdot \nu \, d\sigma = 0.$$

Since this must hold for all smooth v, we find that $\partial v / \partial \nu = 0$. Hence the Euler-Lagrange equation is now:

$$\begin{cases} -\Delta u + u = g & \text{in } \Omega \\ \frac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases}$$

The important point is that now, the second equation does not come from restrictions on the function space where the functional is minimized. It is *part of* the condition $\mathcal{E}'(u) = 0$, and comes, on the contrary, from the fact that admissible variations v are free to vary on the boundary.

Example: consider now a (smooth) solution to the problem

$$\min_{u \in H^1(\Omega)} \int_\Omega |\nabla u|^2\, dx - 2\int_{\partial\Omega} uf\, d\sigma.$$

Can you show that (i) $-\Delta u = 0$ in Ω? (ii) $\partial u/\partial\nu = g$ on $\partial\Omega$?

General situation

For the general form (2.1), if we look for a critical point with "free" boundary conditions (that is, variations $v \in C^\infty(\bar\Omega)^m$, which do not vanish on the boundary, are admissible), we find in addition to (2.3) (which is obtained precisely considering the variations $v \in C_c^\infty(\bar\Omega)^m$) that, integrating (2.2) by parts

$$0 = \int_{\partial\Omega} v^i \sum_{\alpha=1}^n \left(\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\right)\nu_\alpha\, d\sigma \\ + \int_\Omega \left(\frac{\partial\mathcal{L}}{\partial u^i}(x,u,Du) - \sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha}\left(\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\right)\right) v^i\, dx$$

for all $i = 1,\dots,m$ (here $\nu = (\nu_\alpha)_{\alpha=1}^n$ is the outer normal to the boundary of $\partial\Omega$). Since (2.3) holds, the second integral vanishes, and since the first must then vanish for all v, we deduce that for $i = 1,\dots,m$,

$$\left(\frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u,Du)\right)_{\alpha=1}^n \cdot \nu = 0. \tag{2.7}$$

When $\partial\Omega$ or u is not regular enough to justify such computations, we say that the condition (2.7) holds "in the weak sense".

Remark: you can also check that "Robin" or "oblique" boundary conditions, for critical points of first order functionals, fall into this category (they arise when the energy has nonlinear terms on the boundary).

2.4 A remark on "Null Lagrangians"

Consider again the energy in the standard form (2.1).

Definition 4. *$\mathcal{L}$ is called a "Null Lagrangian" if for any $u : \Omega \to \mathbb{R}^m$ smooth enough,* (2.3) *is satisfied.*

It means that any smooth u is a critical point of the energy. In this case, given u, v regular enough with $u = v$ on the boundary $\partial\Omega$, one has $\mathcal{E}(u) = \mathcal{E}(v)$!

Proof. We let for $t \in [0,1]$ $f(t) = \mathcal{E}(tu + (1-t)v)$. We then compute (to simplify we denote $u_t = tu + (1-t)v$):

$$f'(t) = \int_\Omega \sum_i \frac{\partial\mathcal{L}}{\partial u^i}(x,u_t,Du_t)(u^i - v^i) + \sum_{i,\alpha} \frac{\partial\mathcal{L}}{\partial p^i_\alpha}(x,u_t,Du_t)\frac{\partial(u^i - v^i)}{\partial x_\alpha} dx.$$

Integrating by parts (we recall that $u - v = 0$ on the boundary) and using (2.3), it follows $f'(t) = 0$, hence f is constant and $f(0) = f(1)$. □

But: does it exist? (But, of course, for the trivial case $\mathcal{L} = 0$.) If $m = 1$ one can show that the only Null-Lagrangians are of the form $\mathcal{L}(x, u, p) = a \cdot p + b(x)$, a a constant vector, or more generally $\mathcal{L}(x, u, p) = \phi(x) \cdot p + u \mathrm{div}\, \phi(x) + b(x)$ (for ϕ a C^1 vector field). In this case,

$$\int_\Omega \mathcal{L}(x, u, Du)dx = \int_\Omega \mathrm{div}\,(\phi u) + b(x)dx = \int_{\partial\Omega} u\phi \cdot \nu \, d\sigma + \int_\Omega b(x)dx$$

and depends only on the boundary values of u. For $m > 1$ we will see in Chapter 5 that there are less trivial examples.

2.5 Constrained problems, obstacles

When there are constraints for a minimization problem, the Euler-Lagrange equation is not valid anymore, at least in general. Let us just illustrate this with two important examples: the convex constraints, and the differentiable constraints.

2.5.1 Convex constraints

Consider the obstacle problem

$$\min_{u \geq \psi} \int_\Omega |\nabla u|^2 dx :$$

where the function u is constrained to belong to the convex set $C = \{u \in H^1_0(\Omega)\,, u \geq \psi\}$. In this case, instead of considering a variation of the form $u + tv$ (which would raise the difficult issue of how to choose v) it is more convenient to consider another candidate $v \in C$, and variations of the form $tv + (1 - t)u \in C$ (as $tv + (1 - t)u = u + t(v - u)$, it corresponds to nothing else than choosing a variation of the form $v - u$, $v \in C$). It easily follows, for a minimizer, the equation

$$D\mathcal{E}(u) \cdot (v - u) \geq 0$$

for all $v \in C$, called a "variational inequality". In the example above, this reads

$$\int_\Omega \nabla u \cdot \nabla(v - u)dx \geq 0$$

for all v with $v \geq \psi$. Now, in the set $\{u > \psi\}$ (or at least in its interior if u is not continuous), one can consider a v of the form $u \pm t\phi$, $\phi \in C^\infty_c(\{u > \psi\})$, which also satisfies $u \pm t\phi \geq \psi$ a.e. for t small enough. It follows that

$$\int_\Omega \nabla u \cdot \nabla \phi \, dx = 0$$

for all such ϕ, in other words, $-\Delta u = 0$ as in the unconstrained case, in the set where the constraint is not active.

On the other hand, a function $v \geq u$ always satisfies $v \in C$. This means that for any $\phi \geq 0$, $\phi \in H_0^1(\Omega)$, one will have

$$\int_\Omega \nabla u \cdot \nabla \phi \, dx \geq 0.$$

In other words, $-\Delta u \geq 0$, which means u is "superharmonic". One can show (since $-\Delta u$ is a signed distribution) that it is a nonnegative Radon measure, and write the Euler-Lagrange equation, in this case

$$\begin{cases} -\Delta u = \mu & \mu \text{ nonnegative measure,} \\ u = 0 & \text{on } \partial\Omega, \\ \mu(\{u > \psi\}) = 0. \end{cases}$$

A second typical example is the following variational problem

$$\min_{\mu \in \mathcal{P}(\Omega)} \int_{\Omega \times \Omega} W(x,y) d\mu(x) d\mu(y)$$

where $\mathcal{P}(\Omega)$ is the set of probability measures on Ω and W a nonnegative, symmetric weight. If μ is minimizing, one finds that for any ν

$$\int_{\Omega \times \Omega} W(x,y) d\mu(x) d(\nu - \mu)(y) \geq 0$$

and one can deduce that there exists a constant c such that $\int_\Omega W(x,y) d\mu(y) = c$ μ-a.e.

2.5.2 Differentiable constraints

Formally, one solves a problem of the form

$$\min_{\mathcal{F}(u)=0} \mathcal{E}(u)$$

(or one searchs for a critical point). A minimizer should satisfy that $D\mathcal{E}(u) \cdot v = 0$ in all directions which is tangent to the manifold $\{u : \mathcal{F}(u) = 0\}$, that is, with $D\mathcal{F}(u) \cdot v = 0$. Hence the equation should read formally "$D\mathcal{E}(u) \parallel D\mathcal{F}(u)$". In other words, there exists a "Lagrange multiplier" $\lambda \in \mathbb{R}$ such that

$$D\mathcal{E}(u) = \lambda D\mathcal{F}(u). \tag{2.8}$$

Example: we consider the problem $\min_{\int_\Omega u^2 = 1} \int_\Omega |\nabla u|^2 dx$, for $u \in H_0^1(\Omega)$. For a general $v \in H_0^1(\Omega)$, there is no reason for which one would have $\int_\Omega (u + tv)^2 dx = 1$ for any t small enough (it is even impossible unless $v = 0$). However, one may consider a path through u on the variety $\int_\Omega u^2 dx = 1$ with tangent v at u, such as

$$t \mapsto \frac{u + tv}{\|u + tv\|_{L^2}}$$

The minimality of u yields

$$\frac{\int_\Omega |\nabla u + t\nabla v|^2 dx}{\int_\Omega |u + tv|^2 dx} \geq \int_\Omega |\nabla u|^2 dx$$

for all $v \in H_0^1$ and any t small (or not). Developping, we obtain

$$\frac{\int_\Omega |\nabla u|^2 dx + 2t \int_\Omega \nabla u \cdot \nabla v \, dx + t^2 \int_\Omega |\nabla v|^2 dx}{1 + 2t \int_\Omega uv \, dx + t^2 \int_\Omega v^2 dx} \geq \int_\Omega |\nabla u|^2 dx.$$

This means that the left-hand side of this inequality is minimized for $t = 0$, hence that its derivative at $t = 0$ should vanish. This derivative is precisely

$$\int_\Omega \nabla u \cdot \nabla v \, dx - \int_\Omega |\nabla u|^2 dx \int_\Omega uv \, dx.$$

Hence the Euler-Lagrange equation for this problem is

$$-\Delta u = \lambda u$$

and has the form (2.8), with $\lambda = \int_\Omega |\nabla u|^2 dx$.

Chapter 3

Existence of minimizers

A.k.a. "The Direct Method in the Calculus of Variations".

3.1 Lower semicontinuity

Recall that "a continuous function on a compact set reaches its bounds", i.e., if $f : C \to \mathbb{R}$ is continuous and C is compact, there exists, $x, y \in C$ such that $f(x) = \max_C f$ and $f(y) = \min_C f$. Now, if we only need to *minimize* f, is the continuity really needed? In fact, only half of it is necessary.

Definition 5. *Let $U \subset X$ be an open subset of a metric vector space X. We say that $f : U \to \bar{\mathbb{R}}$ is lower semicontinuous (lsc) at $x \in U$ if for any sequence $x_n \to x$, one has*

$$f(x) \leq \liminf_{n\to\infty} f(x_n). \tag{3.1}$$

f is lsc in U if it is lsc at any $x \in U$. We say that f is upper semicontinuous (usc) if $-f$ is lsc.

(In a general topological space, this definition defines the "sequential lower semicontinuity".) Here, $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, the extended real line.

Proposition 1. *Let $f : C \in \mathbb{R} \cup \{+\infty\}$ be a lsc function on a nonempty compact metric set C, and assume $f \not\equiv +\infty$. Then f reaches its minimum: there exists $x \in C$ such that $f(x) = \min_C f$.*

Proof. Since $\exists x \in C$ with $f(x) < +\infty$, we can consider a minimizing sequence (x_n), that is a sequence such that $\lim_n f(x_n) = \inf_C f$. Since C is compact, there exists a subsequence $(x_{n_k})_k$ converging to a point $x \in C$. Since f is lsc, one has

$$f(x) \leq \liminf_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = \inf_C f.$$

And since $x \in C$, it follows $f(x) = \min_C f$. □

Proposition 2. *Let $(f_i)_{i\in I}$ be a family of (extended) real-valued lsc functions. Then $f(x) := \sup_{i\in I} f_i(x)$ is lsc.*

Proof. Indeed, if $x_n \to x$, for any i, we have $f_i(x) \le \liminf_{n\to\infty} f_i(x_n) \le \liminf_{n\to\infty} f(x_n)$. Hence, $f(x) = \sup_i f_i(x) \le \liminf_{n\to\infty} f(x_n)$. □

Important examples of lsc functionals are defined by duality: for instance if

$$\mathcal{E}(u) = \sup\left\{\int_\Omega u\mathrm{div}\,\phi\,dx \;:\; \phi \in C_c^\infty(\Omega;\mathbb{R}^n), \|\phi\|_{L^2} \le 1\right\} \tag{3.2}$$

then by definition it is obviously lsc in $L^2(\Omega)$, thanks to Proposition 2 (or in $L^1(\Omega)$, or even, in fact, in the distributional sense—sequentially). On the other hand, one can show that for any $u \in L^2(\Omega)$,

$$\mathcal{E}(u) = \begin{cases} \|\nabla u\|_{L^2} & \text{if } u \in H^1(\Omega), \\ +\infty & \text{else.} \end{cases}$$

Proposition 3. *f is lsc if and only if for all $a \in \mathbb{R}$, the sets $\{f \le a\}$ are closed, or if and only if for all $a \in \mathbb{R}$, the sets $\{f > a\}$ are open.*

3.2 The direct method in the calculus of variations

3.2.1 Basic principle

Given a minimization problem

$$\min_{u\in X} \mathcal{E}(u),$$

the "direct method in the calculus of variations" consists in

- showing that minimizing sequences belong to a compact set;
- showing that $\mathcal{E}$ is lsc on this set.

However, an important remark here is that in infinite dimension, not all topologies coincide, and it is "enough" to find *a* topology in which minimizing sequences are compact and $\mathcal{E}$ is lsc. Observe however that

- The weaker the topology, the easier it will be for sequences to converge, and in particular it will be more likely that a minimizing sequence is compact;
- However, the stronger the topology, the easier it is for a function to be continuous or lsc, since there are less converging sequences on which the continuity or lower semicontinuity needs to be checked... (for instance, the example (3.2) is obviously continuous in the strong H^1 topology, but only lsc for the L^2 topology.)

Hence, more precisely the technique of proof follows these steps:

1. one checks that $\{u \in X \;:\; \mathcal{E}(u) < +\infty\} \ne \emptyset$ (otherwise there will not exist minimizing sequences);

2. one considers a minimizing sequence $(u_n)_n$, that is, such that $\mathcal{E}(u_n) \to \inf_X \mathcal{E}$ as $n \to \infty$. This always exists, by definition of the "inf", as soon as the previous condition is satisfied:

$$\forall n \in \mathbb{N}\, \exists u_n \text{ s.t. } \mathcal{E}(u_n) \le \max\{-n, \inf_X \mathcal{E} + 1/n\};$$

3. one looks for a topology on X for which such a sequence is [sequentially] pre-compact, that is, for which there will be subsequences (u_{n_k}) which converge to some limit u;

4. one tries to check that $\mathcal{E}$ is lsc in this topology, yielding

$$\mathcal{E}(u) \le \liminf_{k\to\infty} \mathcal{E}(u_{n_k}).$$

3.2.2 An example

We consider the problem

$$\min_{u\in X} \mathcal{E}(u) := \int_\Omega |\nabla u|^2 - 2fu\,dx \tag{3.3}$$

for a given function $f \in L^2(\Omega)$, where Ω is a bounded open subset of $\mathbb{R}^n$, and X is either $H^1_0(\Omega)$ or $H^1(\Omega)$.

The first important point is to observe that there exists u with $\mathcal{E}(u) < \infty$: for instance, $\mathcal{E}(0) = 0$. This allows to consider a minimizing sequence $(u_k)_k$, with $\mathcal{E}(u_k) \to \inf_X \mathcal{E}$.

We now want to check whether $(u_k)_k$ is (pre)compact. The standard way is to start showing bounds on the sequence. We can for instance assume that we have chosen each u_k in such a way that $\mathcal{E}(u_k) \le \mathcal{E}(0) = 0$, hence for all k,

$$\int_\Omega |\nabla u_k|^2 dx \le 2\int_\Omega f u_k\,dx \le 2\|f\|_{L^2}\|u_k\|_{L^2}.$$

If $X = H^1_0(\Omega)$, we can use the *Poincaré inequality* to deduce that

$$\|u_k\|^2_{L^2} \le C\int_\Omega |\nabla u_k|^2 dx \le 2C\|f\|_{L^2}\|u_k\|_{L^2},$$

hence $\|u_k\|_{L^2} \le 2C\|f\|_{L^2}$ and it follows that for some constant $C > 0$,

$$\|u_k\|_{H^1} \le C\|f\|_{L^2},$$

the sequence is uniformly bounded in $H^1(\Omega)$.

If $X = H^1(\Omega)$, assuming Ω is regular (for instance with Lipschitz boundary), the *Poincaré-Wirtinger* inequality establishes that there exists a constant C such that for each k, there is a constant c_k (for instance, $c_k = \int_\Omega u_k\,dx/|\Omega|$) such that

$$\begin{aligned} \|u_k - c_k\|^2_{L^2} \le C\|\nabla u_k\|^2_{L^2} \le 2C\int_\Omega f u_k\,dx \\ = 2C\int_\Omega f(u_k - c_k)\,dx + 2Cc_k\int_\Omega f\,dx \\ \le 2C\|f\|_{L^2}\|u_k - c_k\|_{L^2} + 2Cc_k\int_\Omega f\,dx \end{aligned} \tag{3.4}$$

and one cannot hope to control this norm, unless $\int_\Omega f\,dx = 0$. However, this is good news if one observes that if $\int_\Omega f\,dx \neq 0$, for any constant function $c \in \mathbb{R}$, $\mathcal{E}(c) = -2c\int_\Omega f\,dx$ and the infimum over constants is $-\infty$. Hence for $X = H^1(\Omega)$, there cannot be a solution (and minimizing sequences cannot be compact) unless $\int_\Omega f\,dx = 0$. In this case, (3.4) yields

$$\|u_k - c_k\|_{L^2} \leq 2C\|f\|_{L^2}$$

and again one can deduce a global bound in H^1 for the modified sequence $\bar{u}_k = u_k - c_k$.

How do we use these bounds? We need compactness results ensuring that bounded functions in H^1 will be compact in *some* topology.

3.3 Compactness results in infinite dimension

We will gather here three essential compactness results which are useful in the study of minimizing sequences.

3.3.1 Ascoli-Arzelà

Theorem 1 (Ascoli Arzelà). *Let X be a compact metric space and $(u_k)_k$ a sequence of uniformly bounded:*

$$\sup_{x\in X}\sup_{k\geq 0} |u_k(x)| < +\infty$$

and uniformly equicontinuous functions:

$$\forall\varepsilon > 0, \exists\eta > 0, \forall x, y, \text{dist}\,(x, y) < \eta \Rightarrow |u_k(x) - u_k(y)| \leq \varepsilon. \tag{3.5}$$

Then, there exists $u \in C^0(X)$ and a subsequence u_{k_l} such that $u_{k_l} \to u$ uniformly as $l \to \infty$.

Proof. Consider (x_n) a countable dense sequence in X. One can finds u_{k_l} such that $u_{k_l}(x_n)$ has a limit, denoted $u(x_n)$, for all n. This is a classical diagonal procedure: first extract $u_{\phi_1(k)}(x_1)$ converging to $u(x_1)$. Then, from the sequence $u_{\phi_1(k)}(x_2)$, extract $u_{\phi_2(k)}(x_2)$ which converges to $u(x_2)$. Etc... Eventually let $k_l := \phi_l(l)$.

Now, one checks easily that the uniform continuity of u_k passes to the limit:

$$\forall\varepsilon > 0, \exists\eta > 0, \forall x_n, x_m, \text{dist}\,(x_n, x_m) < \eta \Rightarrow |u(x_n) - u(x_m)| \leq \varepsilon. \tag{3.6}$$

Use this to show that u can be extended to all of X, into a well-defined continuous function u. Idea: for $x \in X$, consider $x_{n_i} \to x$, then $u(x_{n_i})$ is a Cauchy sequence thanks to (3.6) and has a limit, which we denote $u(x)$, once one has checked that this limit was not depending on the particular subsequence (x_{n_i}). Again, u satisfies (3.6) but now everywhere in X and not only on the dense subset $\{(x_n) : n \geq 1\}$.

Eventually, by compactness of X, given ε and η as in the equi-continuity statement, one can cover X by finitely many balls $B(x_{n_i}, \eta)$, $i = 1, \ldots, N$. If l is large enough, $|u_{k_l}(x_{n_i}) - u(x_{n_i})| \leq \varepsilon$ for all $i = 1, \ldots, N$. Then if $x \in X$, there is i such that $x \in B(x_{n_i}, \eta)$, and

$$|u_{k_l}(x) - u(x)| \leq |u_{k_l}(x) - u_{k_l}(x_{n_i})| + |u_{k_l}(x_{n_i}) - u(x_{n_i})| + |u(x_{n_i}) - u(x)| \leq 3\varepsilon$$

and the proof is complete. □

Application: In dimension 1, if $\Omega = (0,1)$ in (3.3), we obtain from the bound $\int_\Omega |u_k'|^2\,dx \le c < \infty$ that the u_k are uniformly equicontinuous: indeed, for each k and $x,y\in\Omega$,

$$|u_k(y) - u_k(x)| = \left| \int_{[x,y]} u_k'(s)\,ds \right| \le \sqrt{|x-y|}\sqrt{c} \tag{3.7}$$

from which (3.5) is easily deduced. (Rigorously, one should first approximate each u_k by smooth functions before writing inequality (3.7).) It follows from Ascoli's theorem that up to a subsequence, the u_k (or the modified $\bar{u}_k$) converge uniformly to a limit u.

Exercise: One can deduce that u is a minimizer because in $L^2(\Omega)$, for instance,

$$v \mapsto \begin{cases} \frac{1}{2}\int_\Omega |\nabla v|^2 dx & \text{if } u \in H^1(\Omega), \\ +\infty & \text{else.} \end{cases}$$

is lower semicontinuous, as one can check it coincides with the sup of continuous functionals:

$$\sup\left\{ \int_\Omega u\mathrm{div}\,\phi\,dx - \frac{1}{2}\int_\Omega |\phi|^2 dx \,:\, \phi \in C_c^\infty(\Omega;\mathbb{R}^n), \|\phi\|_{L^2} \le 1 \right\}$$

(using Riesz' representation theorem).

However, the use of Ascoli-Arzelà's theorem is restricted to few situations, where one can show uniform equi-continuity of a sequence of functions (mostly, in 1D, or when the sequence solves higher order problems).

The next important compactness theorem in functional analysis is Rellich's theorem:

Theorem 2 (Rellich)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded, regular*[1] *open set. Let $(u_k)_{k\ge 1}$ be a bounded sequence in $W^{1,p}(\Omega)$, $1 \le p \le +\infty$*[2]*. Then there exists $u \in L^p(\Omega)$ and (u_{k_l}) a subsequence which converges to u in $L^p(\Omega)$. If $p > 1$ one has in addition that $W^{1,p}(\Omega)$.*

Proof. The proof relies on three ingredients:

1. A continuous extension operator $T : W^{1,p}(\Omega) \to W_0^{1,p}(B)$, where B is a large ball containing Ω, such that $\|Tu\|_{W^{1,p}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\Omega)}$ and $Tu_{|\Omega} = u$. This is the reason for which the domain needs to be bounded and regular, one can "easily" build such an operator for Lipschitz domains [8]. (To simplify, we now denote u_k the sequence Tu_k.)

2. An estimate of the L^p-distance between the smoothing $u_k^\varepsilon := \rho_\varepsilon * u_k$ of u_k and u_k itself, where $\rho_\varepsilon(z) = \varepsilon^{-n}\rho(z/\varepsilon)$ is a smooth mollifying kernel. This is obtained

[1] for instance with Lipschitz boundary.
[2] as soon as $p > n$, one can invoke Ascoli's theorem, thanks to Sobolev/Morrey's injections.

as follows:

$$\begin{aligned} u_k^\varepsilon(x) - u_k(x) &= \int_{\mathbb{R}^n} \rho_\varepsilon(z)[u_k(x-z) - u_k(x)]dz \\ &= \int_{\mathbb{R}^n} \rho_\varepsilon(z) \int_0^1 -\nabla u_k(x-tz)\cdot z\, dtdz \\ &= -\varepsilon \int_0^1 \int_{\mathbb{R}^n} \left[(t\varepsilon)^{-n}\rho\left(\frac{z'}{t\varepsilon}\right)\frac{z'}{t\varepsilon}\right]\cdot \nabla u_k(x-z')dz'dt, \end{aligned}$$

which yields (using $\|f * g\|_p \le \|f\|_1 \|g\|_p$)

$$\|u_k^\varepsilon - u_k\|_{L^p} \le \varepsilon \|z\rho(z)\|_{L^1} \|\nabla u_k\|_{L^p}.$$

3. Ascoli's theorem applied to the sequences $(u_k^{1/i})_{k\ge1}$ for all i. By a diagonal argument, one can find a subsequence u_{k_l} such that for each i, $u_{k_l}^{1/i}$ converges uniformly as $l \to \infty$. Then, one uses

$$\begin{aligned} \|u_{k_m} - u_{k_l}\|_{L^p} \le \|u_{k_m} - u_{k_m}^{1/i}\|_{L^p} + \|u_{k_m}^{1/i} - u_{k_l}^{1/i}\|_{L^p} + \|u_{k_l}^{1/i} - u_{k_l}\|_{L^p} \\ \le C/i + \|u_{k_m}^{1/i} - u_{k_l}^{1/i}\|_{L^p} \end{aligned}$$

by the previous step, showing that u_{k_m} is a Cauchy sequence in L^p, which ends the proof.

For $p > 1$, one also can show (for instance by duality) that $\|\nabla u\|_{L^p}$ is lower semicontinuous in L^p. □

Application: if in problem (3.3), the domain is bounded and regular, we immediately get compactness in L^2 of the sequence (u_k or $\bar{u}_k$). Semicontinuity is as before.

In more general situation, one can "almost always" rely on the Banach-Alaoglu-Bourbaki Thm [5, 4, 12, 11, 14]:

Theorem 3 (Banach-Alaoglu-Bourbaki)**.** *Let X be a Banach space and X^* (or X') its topological dual, that is the set of continuous linear forms, endowed with the norm $\|l\|_* = \sup_{\|x\|\le1} \langle l, x\rangle$. Then $B_{X^*} = \{l \in X^* : \|l\|_* \le 1\}$ is compact for the weak-$*$ topology $\sigma^*(X^*, X)$, defined as the weakest topology such that $l \mapsto \langle l, x\rangle$ is continuous for all $x \in X$:*

$$l_n \stackrel{*}{\rightharpoonup} l \Leftrightarrow \forall x \in X, \langle l_n, x\rangle \to \langle l, x\rangle\,.$$

Remarks 1. If X is reflexive (L^p, $W^{1,p}$ with $1 < p < +\infty$...) then a result is that the unit ball of X is weakly compact. This is not true in general (ex: $X = L^1(\Omega)$).
2. If X is separable, that is, when there exists $(x_n)_{n\ge1}$ a dense sequence in X (that is, quite often), then

$$\mathrm{dist}\,(f,g) = \sum_{n\ge1} 2^{-n} |\, \langle f-g, x_n\rangle\, |$$

is a distance on bounded sets of X^*, which induces the weak-$*$ topology. It implies that from any bounded sequence of X^*, one can extract converging subsequences.

Proof. In general, the proof relies on the fact that a product of compact spaces is compact, see for instance [14]. We can give a very simple "constructive proof" in the simplified setting where X is separable, with a dense sequence $(x_n)_{n\geq 1}$. In this simpler case, it is equivalent to show the compactness of bounded sequences $(y_k)_{k\geq 1}$ of points of X^*. Assume $c = \sup_k \|y_k\|_* < +\infty$, then since $\langle y_k, x_n\rangle$ is bounded for each n, one can extract a subsequence y_{k_l} such that $\langle y_{k_l}, x_n\rangle$ converges to a value $l(x_n)$ for each n (by a diagonal argument). Then we observe that $l(x_n) - l(x_m) = \lim_l \langle y_{k_l}, x_n - x_m\rangle \leq c\|x_n - x_m\|$, hence l can be extended by continuity to all of X. (Since if $x_{n_m} \to x$, then $l(x_{n_m})$ is a Cauchy sequence thanks to the previous inequality, and converges to a value $l(x)$ independent of the sequence (x_{n_m}).) One checks that $l(x) = \lim_l \langle y_{k_l}, x\rangle$ for all x, indeed if $x_{n_m} \to x$, then for all l, m

$$\begin{aligned}|l(x) - \langle y_{k_l}, x\rangle\,| &\leq |l(x) - l(x_{n_m})| + |l(x_{n_m}) - \langle y_{k_l}, x_{n_m}\rangle\,| + |\,\langle y_{k_l}, x_{n_m} - x\rangle\,| \\ &\leq 2c\|x_{n_m} - x\| + |l(x_{n_m}) - \langle y_{k_l}, x_{n_m}\rangle\,|\end{aligned}$$

which is arbitrarily small if m is large enough and then l large enough. One deduces that l is linear, continuous, and is the weak-$*$ limit of y_{k_l}. □

Remark: This proof also shows that $\|l\|_* \leq c$, and more precisely, one easily deduces that $\|l\|_* \leq \liminf_{l\to\infty} \|y_{m_l}\|_*$. In general, if $y_k \overset{*}{\rightharpoonup} y$, then the Banach-Steinhaus theorem ensures that $\|y\|_* \leq \liminf_k \|y_k\|_*$, see for instance [5, 4].

Application: In problem (3.3), it follows from Banach-Alaoglu-Bourbaki's theorem that u_k (or $\bar{u}_k$) is bounded in $H^1(\Omega)$, hence admits weakly converging subsequences. Semicontinuity is as before.

3.4 Basic lower-semicontinuity criteria

A lot of the theory of minimizers (see for instance [6]) is devoted to the study of criteria for the lower-semicontinuity of functionals. We give here only an elementary one (however almost "necessary" for integrands of the first order of scalar-valued functions).

3.4.1 Convex functions

Theorem 4. *Consider X a Banach space. Let $C \subset X$ be a convex set: then C is closed for the strong topology (induced by the norm) if and only if it is closed for the weak topology $\sigma(X, X^*)$ (defined by $x_n \rightharpoonup x$ if and only if $\langle l, x_n\rangle \to \langle l, x\rangle$ for all $l \in X^*$.*

Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex function: then f is lsc for the strong topology if and only if it is lsc for the weak topology.

The second statement is a consequence of the first, in the particular case where the space is $X \times \mathbb{R}$ and the convex set C is the epigraph $\{(x, t) : t \geq f(x)\}$, which is closed if and only if f is lsc.

The first statement is a consequence of the Hahn-Banach separation theorem, see [5, 4]. The idea is that if f is strongly lsc, then it can be written as a supremum of continuous linear form:

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

which is weakly lsc.

Application: returning to problem (3.3), one can check easily that $u \mapsto \mathcal{E}(u)$ is convex and continuous in the strong H^1 topology, hence it is also weakly lsc in H^1. (Observe that this point of view only requires $f \in L^2$ but no assumption on the domain Ω.)

In general, one cannot hope more. If $u_k \to u$ in L^2 or weakly in H^1, one has not in general $\int_\Omega |\nabla u_k|^2 dx \to \int_\Omega |\nabla u|^2 dx$. For instance let

$$u_\varepsilon(x) = \int_0^x \operatorname{sign}\left(\frac{s}{\varepsilon}\right) ds$$

which is a kind of sawtooth. Then $|u'_\varepsilon| = 1$ a.e., $\int_0^1 |u'_\varepsilon|^2 dx = 1$, but $u_\varepsilon \to u = 0$ uniformly, which satisfies $\int_0^1 |u'|^2 dx = 0 < 1$.

Remark: observe that the previous function u_ε is a minimizing sequence for the problem

$$\min_{u \in H^1(0,1)} \int_0^1 |1 - u'^2| + u^2 dx.$$

which is not convex in u'.

3.4.2 Non-convex cases

It is not *necessary* to be convex for a functional to be lsc: for instance

$$\mathcal{E}(u) := \tfrac{1}{2} \int_\Omega |\nabla u|^2 + F(u) dx$$

for F continuous, bounded from below, is weakly lsc in $H^1(\Omega)$. However, we now will see that for scalar integrands of the form $\int \mathcal{L}(x, u, Du)$, convexity in the last variable is essentially necessary.

3.4.3 Necessity of the convexity

Lemma 1. *Consider $f : \mathbb{R} \to \mathbb{R}$ a continuous function such that*

$$u \mapsto \int_0^1 f(u(x)) dx$$

is lsc for the weak convergence in $L^p(0,1)$, $1 < p \leq \infty$ (the weak-$*$ convergence for $p = \infty$). Then f is convex.

Proof. Let $a, b \in \mathbb{R}$ and $t \in [0,1]$ and consider a 1-periodic function $\psi(x) = a$ if $0 \leq x < t$, $\psi(x) = b$ if $t \leq x < 1$. Let then $u_k(x) = \psi(kx)$ for $k \geq 1$. Then

$$u_k \rightharpoonup ta + (1-t)b$$

in L^p. Indeed, up to a subsequence, $u_k \rightharpoonup u$ (Banach-Alaoglu-Bourbaki). Then, one can identify u by testing against smooth (or just continuous) functions: if $g \in C(0,1)$,

$$\int_0^1 u_k(x)g(x)dx = \frac{1}{k}\int_0^k \psi(y)g(\tfrac{y}{k})dy = \frac{1}{k}\sum_{l=1}^k \int_{l-1}^l \psi(y)g(\tfrac{y}{k})dy$$
$$\approx \frac{1}{k}\sum_{l=1}^k ta\,g\left(\tfrac{l-1}{k}\right) + (1-t)b\,g\left(\tfrac{l}{k}\right) \stackrel{k\to\infty}{\longrightarrow} (ta+(1-t)b)\int_0^1 g(s)ds$$

Then, by semicontinuity,

$$\int_0^1 f(ta+(1-t)b)dx \le \liminf_{k\to\infty}\int_0^1 f(u_k)dx = tf(a)+(1-t)f(b).$$

□

Lemma 2. *Let $m,n \ge 1$ and $f:\mathbb{R}^{mn}\to\mathbb{R}$ such that*

$$u \mapsto \int_0^1 f(\nabla u)dx$$

*is weakly lsc in $W^{1,p}(\Omega;\mathbb{R}^m)$, where $\Omega\subset\mathbb{R}^n$. Then **if** $n=1$ **or** $m=1$, f is convex.*

The proof is essentially the same. It FAILS for both $n>1$ and $m>1$. One always have:

Lemma 3. *Let $m,n \ge 1$ and $f:\mathbb{R}^{mn}\to\mathbb{R}$ such that*

$$u \mapsto \int_0^1 f(\nabla u)dx$$

is weakly lsc in $W^{1,p}(\Omega;\mathbb{R}^m)$, where $\Omega\subset\mathbb{R}^n$. Then f is "quasiconvex" in the sense of Morrey, that is: for any $\phi\in W_0^{1,\infty}(\Omega;\mathbb{R}^m)$ and any $A\in\mathbb{R}^{m\times n}$,

$$|\Omega|f(A) \le \int_\Omega f(A+D\phi)\,dx.$$

Proof. First, observe that it is equivalent to require that

$$|\Omega'|f(A) \le \int_{\Omega'} f(A+D\phi)\,dx$$

for any $\phi\in W^{1,\infty}(\Omega';\mathbb{R}^m)$ where Ω' is any other bounded open set of $\mathbb{R}^n$. Indeed, choosing z,λ such that $z+\lambda\Omega'\subset\Omega$, for such ϕ,

$$\phi' : x\mapsto \begin{cases}\lambda\phi\left(\frac{x-z}{\lambda}\right) & \text{if } x\in z+\lambda\Omega' \\ 0 & \text{else}\end{cases} \in W_0^{1,\infty}(\Omega;\mathbb{R}^m).$$

Then

$$\int_\Omega f(A+D\phi')dx = |\Omega\setminus(z+\lambda\Omega')|f(A) + \int_{z+\lambda\Omega'} f\left(A+D\phi\left(\tfrac{x-z}{\lambda}\right)\right)dx$$
$$= |\Omega\setminus(z+\lambda\Omega')|f(A) + \lambda^n\int_{\Omega'} f(A+D\phi)dx$$

and since $|\Omega| f(A) \le \int_\Omega f(A + D\phi')dx$, it follows

$$|\Omega'| f(A) = \lambda^{-n} |z + \lambda\Omega'| f(A)$$
$$= \lambda^{-n}(|\Omega| - |\Omega \setminus (z + \lambda\Omega')|) f(A) \le \int_{\Omega'} f(A + D\phi)dx.$$

Then, we consider $\phi \in W^{1,\infty}(Q;\mathbb{R}^m)$ where $Q = (0,1)^n$ is the unit cube. We then consider the cubes $Q_z = z + \varepsilon Q \subset \Omega$, for $z \in \mathbb{Z}^n$, and the function

$$\phi_\varepsilon := \begin{cases} \varepsilon\phi\left(\frac{x-z}{\varepsilon}\right) & \text{if } x \in z + \varepsilon Q \subset \Omega \\ 0 & \text{else} \end{cases}$$

which is in $W^{1,\infty}(\Omega)$, and goes uniformly to 0 as $\varepsilon \to 0$. Let $u(x) := Ax$, $u_\varepsilon(x) := Ax + \phi_\varepsilon(x)$: one easily shows that $u_\varepsilon \rightharpoonup u$ in $W^{1,p}(\Omega;\mathbb{R}^m)$. Hence

$$|\Omega| f(A) = \int_\Omega f(Du)dx \le \liminf_{\varepsilon \to 0} \int_\Omega f(Du_\varepsilon)dx.$$

The last integral is

$$\sum_{\substack{z \in \mathbb{Z}^n \\ z + \varepsilon Q \subset \Omega}} \int_{z+\varepsilon Q} f\left(A + D\phi\left(\frac{x-z}{\varepsilon}\right)\right) dx$$
$$= \varepsilon^n \sharp\{z \in \mathbb{Z}^n \,:\, z + \varepsilon Q \subset \Omega\} \int_Q f(A + D\phi)dx \xrightarrow{\varepsilon \to 0} |\Omega| \int_Q f(A + D\phi)dx$$

and it follows (since $|Q| = 1$)

$$|Q| f(A) \le \int_Q f(A + D\phi)dx$$

establishing the quasiconvexity of f. □

Remark: if $n = 1$ or $m = 1$, then the quasiconvexity is equivalent to the convexity (why?)

Chapter 4

Regularity for minimizers (elliptic case)

Once one has shown existence of the solution to a problem such as (3.3), a natural question is in what sense the Euler-Lagrange equation

$$-\Delta u = f$$

is valid? A priori, the meaning of this equation is that "the Distributional Laplacian" of $u \in H^1(\Omega)$ (or the distributional divergence of $\nabla u \in L^2(\Omega)$) is represented by a L^2 function. However, can we say more on the regularity of u? If f is continuous, is this equation "classical", that is, is u a C^2 function? And before this, does one have that Δu is the trace of a L^2 matrix D^2u? The regularity theory aims at finding the conditions under which the weak (variational) Euler-Lagrange equations are, in fact, true a.e., or pointwise, in a stronger sense.

The answer to the second question is easily deduced by the "translation method" of Nirenberg, which we describe in the next section.

4.1 Basic regularity estimates

4.1.1 Obvious estimates

A first remark is that given a PDE in weak/variational form, it is always wise to check first from obvious estimates which could follow from taking u, if admissible, or admissible functions of u as a test function...

Choice of a test function For instance, if $u \in H^1_0(\Omega)$ and

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx$$

for all $v \in H^1_0(\Omega)$ (Pb. (3.3)) then obviously

$$\int_\Omega |\nabla u|^2 \, dx \leq \|f\|_{L^2} \|u\|_{L^2},$$

showing that the H_0^1 norm of u is controlled by $\|f\|_2$. (More precisely, one could bound it by the "H^{-1}" norm of f, $\|f\|_{H^{-1}} = \sup_{u\in H_0^1(\Omega),\|u\|_{H_0^1}\le 1} \int_\Omega f u\,dx$.)

Nonlinear variant A variant consists in using nonlinear transformations of u: as an example, if u minimizes

$$\mathcal{E}(u) = \int_\Omega |\nabla u|^2 + (u-f)^2 dx$$

over $X = H_0^1(\Omega)$ or $H^1(\Omega)$, it satisfies for any $v \in X$

$$\int_\Omega \nabla u \cdot \nabla v + (u-f)v\,dx = 0.$$

Taking $v = u$ in the equation, one finds that

$$\|u\|_{H^1} \le \|f\|_{L^2}.$$

Taking $v = \psi(u)$ for ψ nondecreasing, with $\psi(0) = 0$, one find

$$\int_\Omega \psi'(u)|\nabla u|^2 + u\psi(u)dx = \int_\Omega f\psi(u)dx.$$

Now if $\psi(t) = (t^+)^{p-1}$, $p \ge 1$, it follows (by Hölder's inequality)

$$\int_\Omega (u^+)^p dx \le \int_\Omega f(u^+)^{p-1}dx \le \|f^+\|_{L^p}\|u^+\|_{L^p}^{p-1}$$

so that $\|u^+\|_p \le \|f^+\|_p$. Sending $p \to \infty$ one also deduces that $\sup u \le \sup f^+$.

Truncation The last statement can be derived in a simpler way: if $c \ge \sup f$ (and $c \ge 0$ in case $X = H_0^1(\Omega)$), then $v = \min\{u, c\}$ is an admissible candidate in the energy: $\mathcal{E}(u) \le \mathcal{E}(v)$.

However

$$\mathcal{E}(v) = \int_{u<c} |\nabla u|^2 + (u-f)^2 dx + \int_{u>c} (c-f)^2 dx \le \mathcal{E}(u),$$

so that v is a minimizer, which is a contradiction unless $\{u > c\}$ has zero measure. It follows that $u \le c$ a.e.

Exercise: assume $X = H^1(\Omega)$ and v satisfies, in the variational sense in Ω,

$$-\Delta v + v \ge f$$

and $\nabla v \cdot \nu = 0$ on $\partial\Omega$. Show that $v \ge u$.

4.1.2 Higher derivability

Theorem 5. *Let $u \in H^1(\Omega)$ satisfy $-\Delta u \in L^2_{\text{loc}}$. Then $u \in H^2_{\text{loc}}(\Omega)$. More generally if $-\operatorname{div} A\nabla u \in L^2$ with $A(x) \in \mathbb{R}^{n\times n}$ Lipschitz-continuous and uniformly elliptic:*

$$(A\xi) \cdot \xi \geq \gamma|\xi|^2$$

for all $\xi \in \mathbb{R}^n$, then $u \in H^2_{\text{loc}}(\Omega)$.

Proof. Assume first $u \in H^1(\mathbb{R}^n)$, and $f = -\Delta u \in L^2(\mathbb{R}^n)$: by defintion this means

$$\int_{\mathbb{R}^n} \nabla u \cdot \nabla v \, dx = \int_{\mathbb{R}^n} f v \, dx$$

for all $v \in H^1(\Omega)$. In particular, we can consider for e a unit vector and $h \in \mathbb{R}$ small the test function

$$v = -D_{-h} D_h u \in H^1(\mathbb{R}^n)$$

where $D_h \psi := (\psi(x + he) - \psi(x))/h$. It follows

$$-\int_{\mathbb{R}^n} \nabla u \cdot \nabla D_{-h} D_h u \, dx = -\int_{\mathbb{R}^n} f D_{-h} D_h u \, dx.$$

Since

$$\int_{\mathbb{R}^n} g D_h f \, dx = \frac{1}{h} \int_{\mathbb{R}^n} g(x) f(x + he) - g(x) f(x) \, dx$$
$$= \frac{1}{h} \int_{\mathbb{R}^n} g(x - he) f(x) - g(x) f(x) \, dx = -\int_{\mathbb{R}^n} D_{-h} g f \, dx$$

we obtain

$$\int_{\mathbb{R}^n} |D_h \nabla u|^2 dx \leq \|f\|_{L^2} \|D_{-h} D_h u\|_{L^2}.$$

On the other hand for any $\psi \in H^1(\mathbb{R}^n)$, by Cauchy-Schwartz inequality (and, rigorously, approximating first ψ with a smooth function and then passing to the limit),

$$\int_{\mathbb{R}^n} |D_h \psi|^2 dx = \int_{\mathbb{R}^n} \frac{1}{h^2} \Big| \int_0^h \nabla \psi(x + se) \cdot e \, ds \Big|^2 dx$$
$$\leq \int_{\mathbb{R}^n} \frac{1}{h} \int_0^h |\nabla \psi(x + se)|^2 ds dx = \|\nabla \psi\|^2_{L^2}$$

so that we can deduce

$$\int_{\mathbb{R}^n} |D_h \nabla u|^2 dx \leq \|f\|_{L^2} \|D_h \nabla u\|_{L^2}.$$

which shows that $\|D_h \nabla u\|_{L^2} \leq \|f\|_{L^2}$. Hence (Banach-Alaoglu-B.) there exists $\xi \in L^2(\mathbb{R}^n; \mathbb{R}^n)$ and $h_k \downarrow 0$, such that as $k \to \infty$,

$$D_{h_k} \nabla u \rightharpoonup \xi$$

weakly in L^2. Now if $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ it is easy to see that $\lim_k D_{-h_k}\psi = \nabla_e\psi = \sum_i \partial_i\psi e_i$, uniformly. Hence

$$\int_{\mathbb{R}^n} \xi\cdot\psi dx = \lim_k \int_{\mathbb{R}^n} D_{h_k}\nabla u\cdot\psi dx = -\lim_k \int_{\mathbb{R}^n} \nabla u\cdot D_{-h_k}\psi dx = -\int_{\mathbb{R}^n} \nabla u\cdot\nabla_e\psi dx$$

showing that $\xi = D^2u\,e$. Since this is valid for any unit vector e, it follows that $D^2u \in L^2(\mathbb{R}^n;\mathbb{R}^{n\times n})$, moreover for any e, $\|D^2ue\|_2 \le \|f\|_2 = \|\Delta u\|_2$. This shows that one can control *all* the second derivatives of u in L^2 from a control on just the trace of the Hessian.

Now, if one only knows that $\Delta u \in L^2_{\text{loc}}(\Omega)$, for B_r a ball contained in Ω we can consider a "cut-off" $\eta \in C_c^\infty(\Omega;[0,1])$ with $\eta = 1$ in B_r. Then, ηu is a well defined function in $H^1(\mathbb{R}^n)$ ($\nabla(\eta u) = \eta\nabla u + u\nabla\eta$), and $\Delta\eta u = \eta\Delta u + 2\nabla\eta\cdot\nabla u + u\Delta\eta \in L^2(\mathbb{R}^n)$. It follows that $D^2(\eta u) \in H^2(\mathbb{R}^n)$ and in particular, $D^2u \in L^2(B_r)$, with

$$\|D^2u\|_{B_r} \le C(\|u\|_{H^1(B_{r'})} + \|\Delta u\|_{L^2(B_{r'})})$$

where $r' > r$, $B_{r'} \subset \Omega$. □

Exercise: show the same result for $-\text{div}\,A\nabla u \in L^2$, $A = (a_{i,j}(x))_{i,j}$ Lipschitz-continuous and uniformly elliptic.

Extensions. More generally, one can consider integrands of the form $u \mapsto \int_\Omega F(x,\nabla u)dx$, with $F(x,p)$ convex and "uniformly elliptic" in p, meaning now that there exists $\gamma > 0$ such that

$$(D_p^2F(x,p))\xi\cdot\xi = \sum_{i,j}\frac{\partial^2F(x,p)}{\partial p_i\partial p_j}\xi_i\xi_j \ge \gamma\|\xi\|^2$$

for all $\xi \in \mathbb{R}^n$, and check whether the result is still valid. The associated equation is

$$-\text{div}\,\nabla_pF(x,\nabla u) = \cdots$$

and it is natural to wonder whether u has second derivatives in L^2 when this quantity belongs to L^2.

If the ambient space is $\mathbb{R}^n$ and F does not depend on x, this is almost the same proof. One uses that

$$\begin{aligned} D_h\left(\nabla_pF(\nabla u)\right)(x) &= \frac{\nabla_pF(\nabla u(x+he)) - \nabla_pF(\nabla u(x))}{h} \\ &= \int_0^1 D_p^2F(\nabla u(x) + s(\nabla u(x+he) - \nabla u(x)))\frac{\nabla u(x+he) - \nabla u(x)}{h}ds \\ &= \left(\int_0^1 D_p^2F(\nabla u(x) + s(\nabla u(x+he) - \nabla u(x)))ds\right)D_h\nabla u(x) \end{aligned}$$

so that

$$\begin{aligned} -\int_{\mathbb{R}^n}\nabla_pF(\nabla u)\cdot\nabla D_{-h}D_hu\,dx &= \int_{\mathbb{R}^n} D_h\left(\nabla_pF(\nabla u)\right)\cdot D_h\nabla u\,dx \\ &\ge \gamma\|D_h\nabla u\|_{L^2}^2, \end{aligned}$$

then one can proceed as before. If F depends on x, one needs to assume that

$$|F_p(x,p) - F_p(y,p)| \le L|x-y|(1+|p|)$$

for some constant L.

The localization is more tricky... A way is to take $v = -D_{-h}\eta^2 D_h u$ as a test function in the equation, where η is as before a cut-off. We obtain

$$\int_\Omega D_h \nabla_p F(\nabla u)(2\eta\nabla\eta D_h u + \eta^2 D_h \nabla u)dx = -\int_\Omega f(D_{-h}\eta^2 D_h u)dx.$$

While from the left hand side we quite easily get, as before (assuming also $D_p^2 F \le C$) a bound of the form

$$\gamma\|\eta D_h \nabla u\|_{L^2}^2 \le 2C\|\eta D_h \nabla u\|_{L^2}\|\nabla\eta D_h u\|_{L^2} + \text{terms from right hand side.}$$

Bounding the right hand side is a bit more technical, one may for instance write, denoting $g(x) = D_h u(x)$:

$$\begin{aligned} f(x)&(D_{-h}\eta^2 g)(x) \\ &= f(x)(D_{-h}\eta)(x)(\eta g)(x-he)) + f(x)\eta(x)\frac{\eta(x-he)g(x-he) - \eta(x)g(x)}{-h} \\ &= f(x)(D_{-h}\eta)(x)(\eta g)(x-he)) + f(x)\eta(x)\tfrac{1}{h}\int_0^h \nabla_e(\eta(x-se)g(x-se))ds \end{aligned} \quad (4.1)$$

showing that (remember η is smooth)

$$\|fD_{-h}\eta^2 g\|_{L^2} \le c\|f\|_{L^2}\|\eta g\|_{L^2} + c\|\eta f\|_{L^2}\|g\|_{L^2} + \|\eta f\|_{L^2}\|\eta\nabla g\|_{L^2}.$$

(here all the norms can be taken on a small neighborhood of the support of η). The proof can be finished as before...

4.2 The De Giorgi-Nash-Moser Thm

4.2.1 Towards more regularity

If $F : \mathbb{R}^n \to \mathbb{R}$ is a strongly convex function with for all p,

$$\lambda \le D^2 F(p) \le \Lambda$$

(that is,

$$\lambda|\xi|^2 \le (D^2F(p)\xi)\cdot\xi \le \Lambda|\xi|^2$$

for all $\xi \in \mathbb{R}^n$), a minimizer of

$$\int_\Omega F(\nabla u) - fu\,dx \quad (4.2)$$

in $H^1(\Omega)$ (with any kind of boundary condition), with $f \in L^2(\Omega)$, will satisfy the equation

$$\int_\Omega \nabla_p F(\nabla u)\cdot\nabla v\,dx = \int_\Omega fv\,dx$$

for all $v \in H^1_0(\Omega)$. From the previous section, it follows $v \in H^2_{\mathrm{loc}}(\Omega)$. Taking $v = \partial_k \psi$ for $\psi \in C^\infty_c(\Omega)$, and integrating by parts, one also finds

$$-\int_\Omega \sum_{i,j} \partial^2_{i,j} F(\nabla u) \partial^2_{j,k} u \partial_i \psi dx = \int_\Omega f \partial_k \psi \, dx$$

We know that $v = \partial_k u \in H^1_{\mathrm{loc}}$. The above equation shows that in the weak (Distributional) sense,

$$-\mathrm{div}\, A(x) \nabla v = \partial_k f$$

where $A(x) = D^2_p F(\nabla u(x))$. The only information on the matrix $A(x) = (a_{i,j}(x))_{1\le i,j\le n}$ is that it is measurable, and its coefficients satisfy

$$\lambda \le A(x) \le \Lambda$$

a.e. in Ω. We will now show the following theorem, concerning the case $f = 0$.

Theorem 6 (De Giorgi-Nash-Moser). *Let $v \in H^1(B_1)$ satisfy $-\mathrm{div}\, A\nabla v = 0$, that is,*

$$\int_{B_1} \sum_{i,j} a_{i,j}(x) \partial_i v(x) \partial_j \phi(x) dx = 0 \tag{4.3}$$

for all $\phi \in H^1_0(B_1)$. Then $v \in C^{0,\alpha}(B_{1/2})$.

The result is still valid if $f \neq 0$ provided $f \in L^{n+\varepsilon}$, $\varepsilon > 0$, and also if the right hand side is of the form $Df + g$ with both f and g in $L^{n+\varepsilon}$, see [9, Thm 8.22]. For u minimizing (4.2) it implies that $u \in C^{1,\alpha}$. Hence, the matrix $A(x)$, in fact, is $C^{0,\alpha}$.

Then, more regularity is obtained from the classical theory of Schauder. The equation for u becomes

$$-\mathrm{div}\, \nabla F(u) = -\sum_{i,j} \frac{\partial^2 F}{\partial p_i \partial p_j} (\nabla u) \partial^2_{i,j} u = f,$$

and if F is smooth enough this is of the form

$$A(x) : D^2 u = f$$

where A is $C^{0,\alpha}$ (since it is a smooth function of the gradient of u). Then if also f is $C^{0,\alpha}$, it follows [9] that u is $C^{2,\alpha}$. But then A is $C^{1,\alpha}$, etc.

4.2.2 Proof: L^∞ control

Definition 6. *We say that $u \in H^1(B_1)$ is a subsolution of (4.3) if for any $\phi \in H^1_0(B_1)$ with $\phi \ge 0$,*

$$\int_{B_1} a_{i,j}(x) \partial_i u(x) \partial_j \phi(x) \, dx \ \le \ 0 \tag{4.4}$$

(where we assume that repeated indices are summed from 1 to n).

Lemma 4. *Let $u \in H^1(B_1)$ be a subsolution: then*

$$\|u^+\|_{L^\infty(B_{1/2})} \le c \|u\|_{L^2(B_1)}$$

where $c = c(n, \Lambda/\lambda)$.

We prove this lemma in several steps.

Choice of a test function First we assume that $u^+ \in L^\infty$ (a more careful choice of the truncation below could deal with this assumption, we will show a different strategy).

Given $0 < r < R < 1$ we consider a cut-off function $\eta(x) = \min\{1, (R - |x|)^+/(R - r)\}$ (it is 1 in B_r, 0 out of B_R, and $|\nabla \eta| \le 1/(R-r)$). We use $\eta^2 u^+ \in H^1_0(B_1; \mathbb{R}_+)$ as a test function in the equation:

$$\int_{B_R} a_{i,j} \partial_i u \partial_j (\eta^2 u^+) dx \le 0.$$

Observe that

$$\begin{aligned} a_{i,j}\partial_i u \partial_j(\eta^2 u^+) &= a_{i,j}\partial_i u(\partial_j \eta)\eta u^+ + a_{i,j}\partial_i u(\partial_j(\eta u^+))\eta \\ &= a_{i,j}\partial_i u^+(\partial_j\eta)\eta u^+ + a_{i,j}\partial_i(\eta u^+)\partial_j(\eta u^+) - a_{i,j}\partial_i\eta(\partial_j(\eta u^+))u^+ \\ &= a_{i,j}\partial_i(\eta u^+)\partial_j(\eta u^+) - a_{i,j}\partial_i\eta(\partial_j\eta)(u^+)^2 \end{aligned}$$

where we have used the fact that where $u^+ > 0$ or $\partial_i u^+ \neq 0$, $u = u^+$, and the symmetry of $a_{i,j}$.

It follows that

$$\int_{B_R} a_{i,j}\partial_i(\eta u^+)\partial_j(\eta u^+) dx \le \int_{B_R} (u^+)^2 a_{i,j}\partial_i\eta\partial_j\eta dx,$$

hence

$$\lambda \|\nabla(\eta u^+)\|^2_{L^2(B_R)} \le \frac{\Lambda}{(R-r)^2}\|u^+\|^2_{L^2(B_R)}.$$

We now use Sobolev's inequality for the compactly supported function ηu^+: there exists $c = c(n)$ such that

$$\|\eta u^+\|_{L^{2^*}} \le c\|\nabla(\eta u^+)\|_{L^2},$$

where we recall that $2^* = 2n/(n-2)$, and it follows

$$\|u^+\|_{L^{2^*}(B_r)} \le c\sqrt{\frac{\Lambda}{\lambda}}\frac{1}{R-r}\|u^+\|_{L^2(B_R)}. \tag{4.5}$$

This shows that we get a control of u^+ with a better exponent in a smaller ball. We will now show how to "iterate" this inequality, this trick is due to Moser (J. K. Moser, 1928-1999).

"Stability" of subsolutions

Lemma 5. *Let $u \in H^1(B_1)$ be a subsolution and assume that $\Phi : \mathbb{R} \to \mathbb{R}$ is convex, nondecreasing, Lipschitz-continuous (so that $\Phi(u) \in H^1(B_1)$), then $\Phi(u)$ is also a subsolution.*

Proof. If Φ is smooth one can write, for ϕ a non-negative test function

$$\begin{aligned} \int_{B_1} a_{i,j}\partial_i(\Phi(u))\partial_j\phi dx &= \int_{B_1} a_{i,j}\partial_i u(\Phi'(u)\partial_j\phi)dx \\ &= \int_{B_1} a_{i,j}\partial_i u\partial_j(\Phi'(u)\phi) - \int_{B_1}(a_{i,j}\partial_i u\partial_j u)\Phi''(u)\phi \le 0. \end{aligned}$$

We have the right to write this because, in particular, $\Phi'(u)\phi \in H_0^1(B_1;\mathbb{R}^+)$ (and can be approximated with smooth non-negative functions with compact support). We used $\Phi', \Phi'' \geq 0$.

Hence when Φ is smooth,

$$\int_{B_1} a_{i,j}\partial_i(\Phi(u))\partial_j\phi dx \leq 0$$

now if it is not, one can approximate it by smooth functions Φ_n (by convolution) and check that $\Phi_n(u) \to \Phi(u)$ at least weakly in $H^1(B_1)$ (which is sufficient). □

In our case, u^+ is a bounded subsolution (since it is $\Phi(u)$ with $\Phi(t) = \max\{0,t\}$) and also $(u^+)^p$ for all $p \geq 1$ (observe that $\Phi(t) = \max\{0,t\}^p$ is Lipschitz on bounded sets, so that the Lemma may be applied if u^+ is bounded).

We deduce from (4.5) that

$$\|(u^+)^p\|_{L^{2^*}(B_r)} \leq c\sqrt{\frac{\Lambda}{\lambda}}\frac{1}{R-r}\|(u^+)^p\|_{L^2(B_R)}.$$

for any $0 < r < R < 1$ and any $p \geq 1$. Let now $\chi = n/(n-2)$[1] (so that $(t^p)^{2^*} = t^{2\chi p}$). The tricky point is to let

$$p_0 = 1, p_1 = \chi, \ldots, p_k = \chi^k, \ldots$$
$$R_0 = 1, R_1 = \tfrac{3}{4}, \ldots, R_k = \tfrac{1}{2} + \tfrac{1}{2^{k+1}},$$

so that in particular $1/(R_k - R_{k+1}) = 2^{k+2}$, and use recursively (4.5) with the subsolution $(u^+)^{p_k}$, $r = R_{k+1}$, $R = R_k$:

$$\|(u^+)^{p_k}\|_{L^{2^*}(B_{R_{k+1}})} \leq c\sqrt{\frac{\Lambda}{\lambda}}2^{k+2}\|(u^+)^{p_k}\|_{L^2(B_{R_k})}.$$

This is:

$$\left(\int_{B_{R_{k+1}}} (u^+)^{2p_{k+1}}dx\right)^{\frac{1}{2\chi}} \leq c\sqrt{\frac{\Lambda}{\lambda}}2^{k+2}\left(\int_{B_{R_k}} (u^+)^{2p_k}dx\right)^{\frac{1}{2}}$$

and taking the power $(1/p_k)$ on both sides, we get

$$\|u^+\|_{L^{2p_{k+1}}(B_{R_{k+1}})} \leq \left(c\sqrt{\frac{\Lambda}{\lambda}}2^{k+2}\right)^{\frac{1}{\chi^k}} \|u^+\|_{L^{2p_k}(B_{R_k})}$$
$$\leq \prod_{l=0}^{k}\left(c\sqrt{\frac{\Lambda}{\lambda}}2^{l+2}\right)^{\frac{1}{\chi^l}} \|u^+\|_{L^2(B_1)}$$

The point to understand now is whether the product can be uniformly bounded: this is easy, as taking the logarithm we obtain

$$\ln\prod_{l=0}^{k}\left(c\sqrt{\frac{\Lambda}{\lambda}}2^{l+2}\right)^{\frac{1}{\chi^l}} = \sum_{l=0}^{k}\frac{1}{\chi^l}\left(\ln 4c\sqrt{\frac{\Lambda}{\lambda}} + l\ln 2\right)$$

[1] If the dimension $n = 2$, the the proof must be modified a little, for instance taking $\chi > 1$ arbitrary.

which is easily shown to be bounded (since $1/\chi = 1 - 2/n < 1$), by

$$\frac{n}{2} \ln 4c\sqrt{\frac{\Lambda}{\lambda}} + \frac{n(n-2)}{4} \ln 2.$$

Letting

$$C = (4c)^{\frac{n}{2}} \left(\frac{\Lambda}{\lambda}\right)^{\frac{n}{4}} 2^{\frac{n(n-2)}{4}}$$

which as we see, depends only on n (also through c which is the constant in the Sobolev inequality) and on the ellipticity constant Λ/λ, we find that for all k (using also $R_{k+1} \geq 1/2$),

$$\|u^+\|_{L^{2p_{k+1}}(B_{1/2})} \leq C\|u^+\|_{L^2(B_1)}$$

and now we can send $k \to \infty$:

$$\|u^+\|_{L^\infty(B_{1/2})} \leq C\|u^+\|_{L^2(B_1)}$$

which almost shows Lemma 4. Indeed, we need to check that this is also true without the a priory assumption $u^+ \in L^\infty$.

Approximation by bounded subsolutions If we can show that any subsolution can be approximated in L^2 (or H^1-weak) by bounded subsolutions u_k, then it will be enough to pass to the limit in

$$\|u_k^+\|_{L^\infty(B_{1/2})} \leq C\|u_k^+\|_{L^2(B_1)} \tag{4.6}$$

and observe that $u \mapsto \|u\|_\infty$ is lsc with respect to the L^2 convergence (this is easy) to conclude that Lemma 4 is true.

The key observation is that a subsolution is in fact a critical point (or here, since the energy $\int_{B_1} a_{i,j}\partial_i v \partial_j v dx$ is convex, a minimizer) of the energy with a constraint $v \leq u$.

It is therefore natural to introduce u_k the solution of

$$\min \left\{ \int_{B_1} a_{i,j}\partial_i v \partial_j v \, dx : v \in H^1(B_1), v = \min\{u, k\} \text{ on } \partial B_1, v \leq \min\{u, k\} \text{ a.e.} \right\}$$

which satisfies $u_k^+ \leq k$. If $\phi \geq 0$ is a smooth, nonnegative compactly supported test function, we see that $u_k - \varepsilon\phi$ is also admissible so that for all $\varepsilon > 0$ small,

$$\begin{aligned}\int_{B_1} a_{i,j}\partial_i u_k \partial_j u_k dx &\leq \int_{B_1} a_{i,j}\partial_i (u_k - \varepsilon\phi)\partial_j (u_k - \varepsilon\phi) dx \\ &= \int_{B_1} a_{i,j}\partial_i u_k \partial_j u_k dx - 2\varepsilon \int_{B_1} a_{i,j}\partial_i u_k \partial_j \phi \, dx + \varepsilon^2 \int_{B_1} a_{i,j}\partial_i \phi \partial_j \phi \, dx.\end{aligned}$$

It follows that

$$\int_{B_1} a_{i,j}\partial_i u_k \partial_j \phi \, dx \leq 0 \tag{4.7}$$

showing that u_k is a subsolution: and hence (4.6) holds. Observe also that

$$\lambda\|\nabla u_k\|_{L^2}^2 \leq \int_{B_1} a_{i,j}\partial_i u_k \partial_j u_k dx \leq \int_{B_1 \cap \{u<k\}} a_{i,j}\partial_i u \partial_j u dx$$

(since $\min\{u,k\}$ is admissible in the minimization problem) so that u_k is uniformly bounded in $H^1(B_1)$ (we need to use also the boundary condition, and Poincaré's inequality, to make sure this is true), and up to a subsequence converges (weakly in H^1, or strongly in L^2) to some $\tilde{u} = \sup_k u_k \le u$, with $\tilde{u} = u$ on ∂B_1. By lsc of the (convex) energy

$$\int_{B_1} a_{i,j}\partial_i\tilde{u}\partial_j\tilde{u}\,dx \le \liminf_{k\to\infty}\int_{B_1} a_{i,j}\partial_i u_k\partial_j u_k dx \le \int_{B_1} a_{i,j}\partial_i u\partial_j u\,dx.$$

Passing to the limit in (4.7) we also see that

$$\int_{B_1} a_{i,j}\partial_i\tilde{u}\partial_j\phi\,dx \le 0$$

for any smooth nonnegative test function ϕ with compact support, or in fact, by density, any $\phi \in H^1_0(B_1;\mathbb{R}_+)$. In particular taking $\phi = u - \tilde{u} \in H^1_0(B_1;\mathbb{R}_+)$ in the subsolution inequation for u

$$\int_{B_1} a_{i,j}\partial_i u\partial_j u\,dx \le \int_{B_1} a_{i,j}\partial_i u\partial_j\tilde{u}\,dx.$$

Hence

$$\begin{aligned}\lambda\|\nabla(u-\tilde{u})\|^2_{L^2} &\le \int_{B_1} a_{i,j}\partial_i(u-\tilde{u})\partial_j(u-\tilde{u})dx\\ &= \int_{B_1} a_{i,j}\partial_i u\partial_j u\,dx + 2\int_{B_1} a_{i,j}\partial_i u\partial_j\tilde{u}\,dx + \int_{B_1} a_{i,j}\partial_i\tilde{u}\partial_j\tilde{u}\,dx\\ &\le 2\left(\int_{B_1} a_{i,j}\partial_i u\partial_j u\,dx - \int_{B_1} a_{i,j}\partial_i u\partial_j\tilde{u}\,dx\right) \le 0,\end{aligned}$$

showing that $\tilde{u} = u$ achieving the proof of Lemma 4.

4.2.3 Harnack-type estimate

This part is a bit complicated, and uses nonlinear test functions. It shows:

Theorem 7. *Let $u \in H^1(B_2)$ a nonnegative supersolution[2] of (4.3), and assume there exists $\alpha \in (0,1)$ with $|\{u \ge 1\} \cap B_1| \ge \alpha|B_1|$. Then there exists $C = C(\alpha, n, \Lambda/\lambda)$ such that*

$$\inf_{B_{1/2}} u \ge C.$$

This is almost a "Harnack inequality", which would show that $\inf_{B_{1/2}} u \ge C$ if $\max u \ge 1$ (see next section).

Proof. First it is enough to show the result for $u + \delta$, $\delta > 0$, and then pass to the limit $\delta \downarrow 0$. Hence we assume $u \ge \delta > 0$.

[2] meaning that $-u$ is a subsolution.

Then we consider $v = (\ln u)^- = \max\{0, -\ln u\} \le -\ln\delta$. Since $-u$ is a subsolution and $t \mapsto \max\{0, -\ln -t\}$ is convex, nondecreasing, and Lipschitz in $(-\infty, -\delta)$, then v also is a subsolution (Lemma 5). Hence from Lemma 4, v is bounded from above and

$$\sup_{B_{1/2}} v \le c\|v\|_{L^2(B_1)}.$$

(Here and in the following, "sup" and "inf" are the essential supremum and infimum, that is, up to Lebesgue-negligible sets.)

We recall that a variant of Poincaré's inequality states that there exists a constant C depending only on $\alpha > 0$ such that for $w \in H^1(B_1)$, if

$$|\{w = 0\}| \ge \alpha|B_1|$$

then $\|w\|_{L^2} \le C\|\nabla w\|_{L^2}$. This is easily (and classicaly) proved by contradiction: if not, there exists w_n with $|\{w_n = 0\}| \ge \alpha|B_1|$ and $n\|\nabla w_n\|_{L^2} \le \|w_n\|_{L^2}$, etc... **Exercise:** show then that one may assume $\|w_n\|_{L^2} = 1$, and that up to subsequences, w_n converges to a constant function. Find a contradiction.

Hence, we have

$$\sup_{B_{1/2}} v \le c\|v\|_{L^2(B_1)} \le C\|\nabla v\|_{L^2(B_1)} \tag{4.8}$$

where now C depends on α, n, Λ/λ. The "miraculous" trick is that the right-hand side can be bounded "universally". To see this, we use the fact that u is a supersolution in B_2 and consider a test function $\phi = \eta^2/u$ where $\eta \in C_c^\infty(B_2; [0, 1])$, $\eta = 1$ in B_1 (since $u \ge \delta$, $\phi \in H_0^1(B_2; \mathbb{R}_+)$). It follows

$$\int_{B_2} a_{i,j}\partial_i u \left(-\frac{\eta^2}{u^2}\partial_j u + 2\frac{\eta\partial_j\eta}{u}\right) dx \ge 0$$

Hence, since $\nabla u/u = \nabla \ln u$,

$$\begin{aligned}\int_{B_2} \eta^2 a_{i,j}\partial_i \ln u\partial_j \ln u\, dx &\le 2\int_{B_2} \eta a_{i,j}\partial_i \ln u\partial_j\eta\, dx \\ &\le 2\sqrt{\int_{B_2} \eta^2 a_{i,j}\partial_i \ln u\partial_j \ln u\, dx}\sqrt{\int_{B_2} a_{i,j}\partial_i\eta\partial_j\eta\, dx}\end{aligned}$$

because of Cauchy-Schwartz inequality (applied to the nonnegative quadratic form $\int_{B_2} a_{i,j}\xi_i\xi_j dx$), and it follows

$$\int_{B_2} \eta^2 a_{i,j}\partial_i \ln u\partial_j \ln u\, dx \le 4\int_{B_2} a_{i,j}\partial_i\eta\partial_j\eta\, dx.$$

Choosing well η, the right-hand side can be made as small as $4\Lambda|B_2 \setminus B_1|$, in any case it is a constant independent on u: it follows

$$\lambda\|\nabla v\|^2_{L^2(B_1)} \le \int_{B_1} a_{i,j}\partial_i \ln u\partial_j \ln u\, dx \le C(n)\Lambda$$

and we deduce from (4.8) that

$$\sup_{B_{1/2}} v \le C$$

with a constant depending only on $n, \alpha, \Lambda/\lambda$. In other words $\inf_{B_{1/2}} u \ge e^{-C} > 0$, which is what we needed to prove. □

4.2.4 Application: control of the oscillation of a solution

Theorem 8. *Let u be a solution in B_4. Then there exists a constant $\gamma = \gamma(n, \Lambda/\lambda) \in (0,1)$ such that*

$$\operatorname{osc}_{B_{1/2}} u \le \gamma \operatorname{osc}_{B_2}(u) :$$

the oscillation of u decreases as we move towards the center of the ball.

Proof. This is a relatively simple consequence of the previous results. By Lemma 4, u is bounded in B_2. Letting $\alpha = \sup_{B_2} u$, $\beta = \inf_{B_2} u$, $\alpha' = \sup_{B_{1/2}} u$, $\beta' = \inf_{B_{1/2}} u$, we have $u - \beta \ge 0$ and $\alpha - u \ge 0$ in B_2. At each point, either $u - \beta \ge (\alpha - \beta)/2$, or $u - \beta < (\alpha - \beta)/2$ so that $\alpha - u > (\alpha - \beta)/2$. Hence one of the two sets

$$\left\{ 2\frac{u - \beta}{\alpha - \beta} \ge 1 \right\} \cap B_1 \quad \text{or} \quad \left\{ 2\frac{\alpha - u}{\alpha - \beta} > 1 \right\} \cap B_1$$

has a measure larger than $|B_1|/2$. Assume for instance it is the first one, then, since clearly, $2(u - \beta)/(\alpha - \beta)$ (which is a solution of (4.3)) is a supersolution in B_2, there is a constant C which depends only on n and Λ/λ such that from Theorem 7,

$$2\frac{u - \beta}{\alpha - \beta} \ge C$$

in $B_{1/2}$. It follows that $\beta' - \beta \ge (C/2)(\alpha - \beta)$, so that $\alpha' - \beta' \le \alpha - \beta' \le (1 - C/2)(\alpha - \beta)$. This shows the results with $\gamma = (1 - C/2) \in (1/2, 1)$. □

Remark: you can show easily that γ is scale-invariant: if u is a solution in $B(x, 4r)$, then $\operatorname{osc}_{B(x,r/2)} u \le \gamma \operatorname{osc}_{B(x,2r)} u$. (Exercise...)

4.2.5 Conclusion: u is Hölder

Theorem 9. *Let u be a solution in B_4. Then there exists $\alpha > 0$ such that*

$$\sup_{B_1} |u| + \sup_{x,y \in B_1} \frac{|u(x) - u(y)|}{|x - y|^\alpha} \le C\|u\|_{L^2(B_4)},$$

C and α depending on n and Λ/λ.

Exercise: show then that if u is a solution in Ω, for all $\Omega' \subset\subset \Omega$ (*i.e.*, such that $\overline{\Omega'}$ is a compact subset of Ω), then $\|u\|_{C^{0,\alpha}(\Omega)} \le C\|u\|_{L^2(\Omega)}$ for some constant C.

Proof. We prove the theorem: first, the fact that $|u|$ is bounded by $C\|u\|_{L^2(B_4)}$ in B_1 follows from Lemma 4. Then, consider $x, y \in B_1$ and let $r = |x - y| \le 2$, $z = (x + y)/2$. By definition $|u(x) - u(y)| \le \operatorname{osc}_{B(z,r/2)} u$. Hence if r is small enough (so that $B(z, 4r) \subset B_4$),

$$|u(x) - u(y)| \le \gamma \operatorname{osc}_{B(z,2r)} u$$

and by induction,

$$|u(x) - u(y)| \le \gamma^2 \operatorname{osc}_{B(z,8r)} u \le \cdots \le \gamma^k \operatorname{osc}_{B(z,4^k r/2)} u \le C\gamma^k \|u\|_{L^2(B_4)}$$

as long as $B(z, 4^k r/2) \subset B_2$, hence as long as $1 + 4^k r/2 \leq 2$ (this is true even for $k = 0$, in fact). Choosing the largest $k \geq 0$ such that this is true, that is, such that $4^k \leq 2/r$ and $4^k \geq 1/(2r)$, we find that (be careful that $\ln \gamma < 0$)

$$\gamma^k = 4^{k \frac{\ln \gamma}{\ln 4}} = \left(\frac{1}{4^k} \right)^{\left| \frac{\ln \gamma}{\ln 4} \right|} \leq (2r)^{\left| \frac{\ln \gamma}{\ln 4} \right|}$$

so that, letting $\alpha = -\ln \gamma / \ln 4 \in (0, 1/2)$ (since $\gamma > 1/2$),

$$|u(x) - u(y)| \leq C|x - y|^\alpha \|u\|_{L^2(B_4)}$$

which was the thesis we wanted to prove. □

Chapter 5

Vector-valued minimization problems

[maybe this part will be skipped as too close to S. Serfaty's lecture, if time permits I'll write some notes on the subject, otherwise the main references are [6, 3, 10]]

A few results: we recall the definition

Definition 7. *$f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is "quasiconvex" in the sense of Morrey if and only if, given Ω a bounded open set in $\mathbb{R}^n$, for any $\phi \in W_0^{1,\infty}(\Omega;\mathbb{R}^m)$ and any $A \in \mathbb{R}^{m\times n}$,*

$$|\Omega| f(A) \leq \int_\Omega f(A + D\phi)\, dx.$$

The definition is independent on the choice of Ω.

Some properties:

Proposition 4. *Let f be as in definition 7, and assume it is locally bounded (which means precisely that $\sup_{|A|\leq R} f(A) < +\infty$ for any $R > 0$). Then f is rank-one convex:*

$$f(tA + (1-t)B) \leq tf(A) + (1-t)f(B)$$

for $t \in [0,1]$, as soon as $\mathrm{rk}(B - A) = 1$, that is $B - A = a \otimes \nu$ for some $a \in \mathbb{R}^m$, $\nu \in \mathbb{R}^n$. In particular, if $m = 1$ or $n = 1$, f is convex.

Proof. We only give the idea of the proof: without loss of generality (just rotating the axes) one can assume $\nu = e_1$ and $\Omega = (0,1)^n$, and $t = 1/2$. We let $C = (A + B)/2$. Then we consider for $\varepsilon > 0$ the function $u_\varepsilon(x) = Ax+\varepsilon(k+1)a$ if $2k\varepsilon \leq x_1 \leq (2k+1)\varepsilon$, $u_\varepsilon(x) = Bx-\varepsilon ka$ if $(2k+1)\varepsilon \leq x_1 \leq (2k+2)\varepsilon$, and we check that since $B-A = a\otimes e_1$, it is continuous (if $x_1 = (2k+1)\varepsilon$, $(B-A)x-\varepsilon ka-\varepsilon(k+1)a = a(2k+1)\varepsilon-\varepsilon(2k+1)a = 0$, etc). Then, one checks that $u_\varepsilon \to Cx$ uniformly. Given a cut-off $\eta \in C_c^\infty((0,1)^n)$ (with $0 \leq \eta \leq 1$, $\eta = 1$ on $Q^\delta = (\delta, 1-\delta)^n$), and letting $\phi(x) = \eta(x)(u_\varepsilon(x) - Cx)$, one has by definition

$$f(C) \le \int_Q f(C + D\phi)dx = \int_Q f(\eta Du_\varepsilon + (1-\eta)C + (u_\varepsilon - Cx)\otimes \nabla\eta)dx$$
$$\le \int_{Q^\delta} f(Du_\varepsilon)dx + \int_{Q\setminus Q^\delta} f(\eta Du_\varepsilon + (1-\eta)C + (u_\varepsilon - Cx)\otimes \nabla\eta)dx$$

Now, if ε is small enough, then since $u_\varepsilon \to C$ uniformly, for any x

$$\eta Du_\varepsilon(x) + (1-\eta)C(x) + (u_\varepsilon(x) - Cx)\otimes\nabla\eta(x) \in \{M : \mathrm{dist}\,(M, [A,B]) \le 1\}$$

(or any neighborhood of $[A, B]$) so that by assumption, f is less than some bound M, hence

$$f(C) \le \int_{Q^\delta} f(Du_\varepsilon)dx + |Q \setminus Q^\delta| M.$$

Since $Du_\varepsilon = A$ in about half of the domain and $Du_\varepsilon = B$ in the other half, we find, sending $\varepsilon \to 0$, that $f(C) \le (f(A) + f(B))/2$ up to an error which goes to zero as $\delta \to 0$. □

It is well known that a convex function which is locally bounded is also locally Lipschitz. When the function has a polynomial growth, this can be made more precise.

Proposition 5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex "with growth p": there exist $\alpha, \beta > 0$ such that*

$$\alpha(|A|^p - 1) \le f(A) \le \beta(|A|^p + 1)$$

for any $A \in \mathbb{R}^n$. Then there is a constant $C > 0$ such that

$$|f(A) - f(B)| \le C(1 + |A|^{p-1} + |B|^{p-1})|A - B|.$$

Proof. Let $A, B \in \mathbb{R}^n$ and set $R = \max\{|A|, |B|, 1\}$. If $A \ne B$, there exists a unique C with $|C-A| = R$ and $B \in [A, C]$. Observe that $|C| \le 2R$. Letting $t = |B-A|/|C-A| = |B - A|/R$, we have $B = tC + (1-t)A$, hence

$$f(B) \le (1-t)f(A) + tf(C)$$

so that

$$f(B)-f(A) \le |B-A|\frac{f(C) - f(A)}{R} \le |B-A|\frac{\beta(1 + (2R)^p) - \alpha}{R} \le |B-A|(1+2^p R^{p-1})$$

which ends the proof, since $R^{p-1} \le |A|^{p-1} + |B|^{p-1}$ unless this sum is less than 2. □

Corollary 1. *Assume f is quasiconvex, locally bounded, with growth p. Then again, there is a constant $C > 0$ such that*

$$|f(A) - f(B)| \le C(1 + |A|^{p-1} + |B|^{p-1})|A - B|. \tag{5.1}$$

Proof. It is enough to apply the previous results to f in rank-1 directions. Given $A = (a_1, \dots, a_n), B = (b_1, \dots, b_n) \in \mathbb{R}^{m\times n}$ two matrices, with $a_i, b_i \in \mathbb{R}^m$, one has

$$\begin{aligned} A - B &= (A - (b_1, a_2, \dots, a_n)) \\ &\quad + ((b_1, a_2, \dots, a_n) - (b_1, b_2, a_3, \dots, a_n)) + \cdots + ((b_1, \dots, b_{n-1}, a_n) - B). \end{aligned}$$

Since f is rank-one convex with p growth, one has

$$f(A) - f(b_1, a_2, \dots, a_n) \leq C(1 + |A|^{p-1} + |B|^{p-1})|b_1 - a_1|, \quad \text{etc}$$

Summing, we obtain

$$f(A) - f(B) \leq C(1 + |A|^{p-1} + |B|^{p-1}) \sum_{i=1}^{n} |b_i - a_i| \leq C\sqrt{n}(1 + |A|^{p-1} + |B|^{p-1})|B - A|.$$

□

Chapter 6

Variational convergence

6.1 Why?

Let $\mathcal{E}_n : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ be functionals and assume we want to understand the "limit" of the problems

$$\min_{x \in X} \mathcal{E}_n(x) \tag{6.1}$$

In other words: if $(x_n)_n$ is the sequence of minimizers of (6.1), how does it behave as $n \to \infty$?

The best would be to define a notion of convergence for the $\mathcal{E}_n$ which would yield that "$\mathcal{E}_n \to \mathcal{E}$" implies that $x_n \to x$ which minimizes $\mathcal{E}$. One would also like to have a bit of "stability", and be coherent with the values of the energy, that is:

(A) if $\mathcal{G}$ is continuous and $\mathcal{E}_n \to \mathcal{E}$, then $\mathcal{E}_n + \mathcal{G} \to \mathcal{E} + \mathcal{G}$,

(B) if $\mathcal{E}_n \to \mathcal{E}$, $\inf_X \mathcal{E}_n \to \inf_X \mathcal{E}$.

Let us introduce the following definition [7, 2]:

Definition 8. *We say that the sequence of functionals $(\mathcal{E}_n)_n : X \to \mathbb{R} \cup -\infty, +\infty$, where X is a metric space, Γ-converges to $\mathcal{E}$ in X if for any $x \in \mathcal{E}$,*

i. when $x_n \to X$, then $\mathcal{E}(x) \leq \liminf_n \mathcal{E}_n(x_n)$;

ii. there exists $x_n \to x$ such that $\limsup_n \mathcal{E}_n(x_n) \leq \mathcal{E}(x)$.

Another name for this notion is *epi-convergence* [1]. The sequence $(x_n)_n$ in (ii) is called a *recovery sequence.*

Now it is very easy to check that this will satisfy (A) and (B) above. In fact, (A) is obvious from the definition (since in both axioms i. and ii., one will have $\mathcal{G}(x_n) \to \mathcal{G}(x)$). For (B), we have:

Theorem 10. *Let $\mathcal{G}$ be continuous and $\mathcal{E}_n$ which Γ-converges to E. Assume that for each n, x_n is a minimizer of $\mathcal{E}_n + \mathcal{G}$, or more generally that there exists $\varepsilon_n \downarrow 0$ such that for any $x \in X$*

$$\mathcal{E}_n(x_n) + \mathcal{G}(x_n) \leq \mathcal{E}_n(x) + \mathcal{G}(x).$$

Then if $\bar{x}$ is a cluster point of $(x_n)_n$ (that is, $x_{n_k} \to \bar{x}$ for some subsequence), $\bar{x}$ is a minimizer of $\mathcal{E} + \mathcal{G}$.

Proof. This is obvious: if $y \in X$, by (ii) one finds $y_n \to y$ such that $\limsup_n \mathcal{E}_n(y_n) \leq \mathcal{E}(y)$. One has that

$$\mathcal{E}(x) \leq \liminf_k \mathcal{E}_{n_k}(x_{n_k})$$

(using (i), possibly for the sequence $x'_n = x$ if $n \neq n_k$ for some k, and x_{n_k} else). Hence

$$\begin{aligned}\mathcal{E}(x) + \mathcal{G}(x) &\leq \liminf_k \mathcal{E}_{n_k}(x_{n_k}) + \lim_k G(x_{n_k}) \\ &\leq \liminf_k \mathcal{E}_{n_k}(x_{n_k}) + G(x_{n_k}) \leq \liminf_k \mathcal{E}_{n_k}(y_{n_k}) + G(y_{n_k}) + \varepsilon_{n_k} \\ &\leq \limsup_k \mathcal{E}_{n_k}(y_{n_k}) + \lim_k G(y_{n_k}) + \varepsilon_{n_k} \leq \mathcal{E}(y) + \mathcal{G}(y).\end{aligned}$$

□

Remark: one also easily deduces, in this case, that $\mathcal{E}(x_{n_k}) + \mathcal{G}(x_{n_k}) \to \min_X \mathcal{E} + \mathcal{G}$.

6.1.1 Necessity of the axioms

Hence, it works. It turns out that formally one can check that conditions (i) and (ii) are necessary in order for (A) and (B) to hold. We assume for simplicity that E_n, E are bounded from below (for instance, nonnegative), coercive and lsc.

Consider $\bar{x} \in X$ and $x_n \to \bar{x}$, for $\varepsilon > 0$ small one could define

$$\mathcal{E}_n(x) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x, \bar{x})$$

Assume this has a minimizer x_n^ε, and x_n^ε converges to some x^ε. Then the stability (A) ensures that x^ε should minimize $x \mapsto \mathcal{E}(x) + \mathrm{dist}\,(x, \bar{x})/(2\varepsilon)$, while (B) yields

$$\lim_n \mathcal{E}_n(x_n^\varepsilon) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x_n^\varepsilon, \bar{x}) = \mathcal{E}(x^\varepsilon) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x^\varepsilon, \bar{x}).$$

Now as $\varepsilon \to 0$, $x^\varepsilon \to \bar{x}$ by construction (unless $\mathcal{E}$ is infinite in a neighborhood of $\bar{x}$), and (we also assume that $\mathcal{E}$ is lsc), $\mathcal{E}(x) \leq \liminf_{\varepsilon \to 0} \mathcal{E}(x^\varepsilon)$. Since

$$\begin{aligned}\mathcal{E}(x^\varepsilon) &\leq \lim_n \mathcal{E}_n(x_n^\varepsilon) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x_n^\varepsilon, \bar{x}) \\ &\leq \liminf_n E_n(x_n) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x_n, \bar{x}) \quad \text{[by minimality of } x_n^\varepsilon] \\ &= \liminf_n E_n(x_n)\end{aligned}$$

we deduce that (i) must hold.

On the other hand, if $\mathcal{E}(\bar{x}) < +\infty$, by minimality one also have that

$$\mathcal{E}(x^\varepsilon) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x^\varepsilon, \bar{x}) \leq \mathcal{E}(\bar{x})$$

so that

$$\lim_n \mathcal{E}_n(x_n^\varepsilon) + \frac{1}{2\varepsilon}\mathrm{dist}\,(x_n^\varepsilon, \bar{x}) \leq \mathcal{E}(\bar{x}).$$

Hence given $\varepsilon > 0$, for n large enough $\mathcal{E}_n(x_n^\varepsilon) \leq \mathcal{E}(\bar{x}) + \varepsilon$ and $\mathrm{dist}\,(x_n^\varepsilon, \bar{x}) \leq 2\varepsilon(\mathcal{E}(\bar{x}) + \varepsilon)$. One can therefore build a global sequence $x'_n \to \bar{x}$ with $\limsup_n \mathcal{E}_n(x'_n) \leq \mathcal{E}(\bar{x})$.

6.1.2 Properties and comments

Proposition 6. *Assume* $\mathcal{E}_n \xrightarrow{\Gamma} \mathcal{E}$. *Then* $\mathcal{E}$ *is lsc.*

Proof. If $x^k \to x$, for each k using (ii) one has $x_n^k \to x^k$ with $\limsup_n \mathcal{E}_n(x_n^k) \le \mathcal{E}(x^k)$. Then (letting $N(0) = 0$), for each k, there exists $N(k) > N(k-1)$ such that

$$n \ge N(k) \Rightarrow \mathcal{E}_n(x_n^k) \le \mathcal{E}(x^k) + \frac{1}{k}, \quad \text{dist}\,(x_n^k, x^k) \le \frac{1}{k}.$$

Let for each $n \ge 1$, $y_n := x_n^k$ if $N(k) \le n \le N(k+1)$. Then, for such n,

$$\text{dist}\,(y_n, x) \le \frac{1}{k} + \text{dist}\,(x^k, x)$$

so that $y_n \to x$ as $n \to \infty$. Hence by (i),

$$\mathcal{E}(x) \le \liminf_n \mathcal{E}_n(y_n) \le \liminf_{n, N(k) \le n \le N(k+1)} \mathcal{E}(x^k) + \frac{1}{k} = \liminf_k \mathcal{E}(x^k)$$

and $\mathcal{E}$ is lsc. □

In particular, one can show that

Proposition 7. *Let* $\mathcal{E}$ *be a functional and let* $\mathcal{E}_n := \mathcal{E}$ *for each* n. *Then* $(\mathcal{E}_n)_n$ Γ-*converges to* $\text{sc}^-\mathcal{E}$, *the lower-semicontinuous envelope of* $\mathcal{E}$.

More precisely, one can introduce the following definitions:

Definition 9. *Given* $\mathcal{E}_n$ *as in Def. 8, one defines the "*Γ*-lim inf" and "*Γ*-lim sup" of* $\mathcal{E}_n$ *by letting:*

$$(\Gamma - \liminf_n \mathcal{E}_n)(x) = \inf \left\{ \liminf_n \mathcal{E}_n(x_n) \ : \ x_n \to x \right\}$$

$$(\Gamma - \limsup_n \mathcal{E}_n)(x) = \inf \left\{ \limsup_n \mathcal{E}_n(x_n) \ : \ x_n \to x \right\}.$$

Then, one checks easily that these functionals are always well-defined (while it is not true for the Γ-limit), and that

- both $\Gamma - \liminf_n \mathcal{E}_n$ and $\Gamma - \limsup_n \mathcal{E}_n$ are lsc;
- $\Gamma - \liminf_n \mathcal{E}_n \le \Gamma - \limsup_n \mathcal{E}_n$;
- $\mathcal{E}_n$ Γ-converges if and only if $\Gamma - \limsup_n \mathcal{E}_n \le \Gamma - \liminf_n \mathcal{E}_n$, and in this case the Γ-limit is precisely the common value of these two functionals.

Observe that checking (i) in Def. 8 exactly means that $\mathcal{E}(x) \le (\Gamma - \liminf_n \mathcal{E}_n)(x)$, while (ii) means that $(\Gamma - \limsup_n \mathcal{E}_n)(x) \le \mathcal{E}(x)$.

Comments: The Γ-convergence is different from the pointwise convergence. Example, in $\mathbb{R}$: $f_n(x) = 0$ if $|x| \geq 1/n$, $1 - |2nx + 1|$ if $-1/n \leq x \leq 0$, $|2nx - 1| - 1$ if $0 \leq x \leq 1/n$. Then $f_n \to 0$ pointwise, while f_n Γ-converges to $f(x) = 0$ if $x \neq 0$, $f(0) = -1$.

It is different from the uniform convergence, for instance, given any $\mathcal{E}$, $F + 1/n \to F$ uniformly even when F is not lsc.

Important: observe that $\mathcal{E}_n \to \mathcal{E}$ does not imply $-\mathcal{E}_n \to -\mathcal{E}$! In fact, one has:

Proposition 8.

$$\begin{cases} \mathcal{E}_n \xrightarrow{\Gamma} \mathcal{E} \\ -\mathcal{E}_n \xrightarrow{\Gamma} -\mathcal{E} \end{cases} \quad \Leftrightarrow \quad \mathcal{E}_n \text{ “continuously converges” to } \mathcal{E}$$

where $\mathcal{E}_n$ is said to continuously converge to $\mathcal{E}$ if for any $x_n \to x$, one has $\mathcal{E}_n(x_n) \to \mathcal{E}(x)$ (this is stronger than the uniform convergence on compact sets).

Proof. The implication $\Leftarrow$ is obvious, for the other direction we just write

$$\mathcal{E}(x) \leq \liminf_n \mathcal{E}_n(x_n) \leq \limsup_n \mathcal{E}_n(x_n) = -\liminf_n(-\mathcal{E}_n(x_n)) \leq -(-\mathcal{E}(x)) = \mathcal{E}(x).$$

□

6.1.3 Compactness of Γ-convergence

Theorem 11. *Let (X, d) be a separable metric space and let $(\mathcal{E}_n)_n$ be a sequence of functionals from $X \to \mathbb{R} \cup \{-\infty, +\infty\}$. Then, there exists a functional $\mathcal{E} : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ and a subsequence $(\mathcal{E}_{n_k})_k$ such that $\mathcal{E}_{n_k}$ Γ-converges to $\mathcal{E}$.*

Remark: one may have, here, that $\mathcal{E} \equiv -\infty$ or $\mathcal{E} \equiv +\infty$.

Before, let us introduce the following definition:

Definition 10. *Consider $(A_n)_n$ a sequence of subsets of X. We define the inferior and superior Kuratowski limits of $(A_n)_n$ as*

$$\operatorname*{Li}_{n\to\infty} A_n = \left\{x \in X \,:\, \limsup_n \operatorname{dist}(A_n, x) = 0\right\} = \{x \in X \,:\, \exists N, \forall U \in \mathcal{V}(x), n \geq N \Rightarrow U \cap A_n \neq \emptyset\},$$

$$\operatorname*{Ls}_{n\to\infty} A_n = \left\{x \in X \,:\, \liminf_n \operatorname{dist}(A_n, x) = 0\right\} = \{x \in X \,:\, \forall U \in \mathcal{V}(x), \sharp\{n \in \mathbb{N} \,:\, U \cap A_n \neq \emptyset\} = \infty\}.$$

We say that A_n converges to A in the sense of Kuratowski if and only if $A = \mathrm{Li}_n A_n = \mathrm{Ls}_n A_n$.

Here $\mathcal{V}(x)$ is the set of the neighborhoods of x. Observe that $\mathrm{Li}_n A_n \subseteq \mathrm{Ls}_n A_n$, and that

$$\operatorname*{Li}_n A_n = \{x : \exists x_n \to x, x_n \in A_n \forall n\}, \quad \operatorname*{Ls}_n A_n = \{x : \exists x_{n_k} \to x, x_{n_k} \in A_{n_k} \forall k\}.$$

Let us show first the following

Theorem 12. *Given $(A_n)_n$ a sequence of subsets of X (a separable metric space). Then there exists a subsequence $(A_{n_k})_k$ and a closed set $A \subset X$ such that $A_{n_k} \to A$ in the sense of Kuratowski.*

Proof. Let $d_n(x) := \min\{\mathrm{dist}\,(x, A_n), 1\}$. One easily shows that it is 1-Lipschitz. As usual, one uses a diagonal argument to extract a subsequence such that $d_{n_k}(x_i) \to \delta(x_i)$ as $k \to \infty$, where $(x_i)_i$ is a dense sequence in X. As usual, being δ 1-Lipschitz, it is extended in a unique way in a Lipschitz function over X and one can easily show that $d_{n_k}(x) \to \delta(x)$ for any $x \in X$ (see the proof of Ascoli's theorem).

We let $A = \{\delta = 0\}$. Then, if $x \in A$, $d_{n_k}(x) \to 0$ hence there exists $x_{n_k} \in A_{n_k}$ with $x_{n_k} \to x$. This shows that $A \subseteq \mathrm{Li}_k A_{n_k}$. On the other hand, if $x \in \mathrm{Ls}_k A_{n_k}$, by definition there exists a subsequence $x_{n_{k_l}} \to x$ with $x_{n_{k_l}} \in A_{n_{k_l}}$. But then, $d_{n_{k_l}}(x) \leq \mathrm{dist}\,(x, A_{n_{k_l}}) \to 0$, hence $\delta(x) = 0$ and $x \in A$: $\mathrm{Ls}_k A_{n_k} \subseteq A$. Hence $A_{n_k} \to A$ in the Kuratowski sense. □

Proof of Thm 11. To prove Theorem 11, we now define the sets $E_n = \mathrm{epi}\,\mathcal{E}_n = \{(x,t) \in X \times \mathbb{R}, \mathcal{E}(x) \leq t\}$. Then there is a subsequence and a closed set $E \subset X \times \mathbb{R}$ such that $E_{n_k} \to E$ in the Kuratowski sense. We first observe that E is the epigraph of an lsc function: if $(x,t) \in E$, then it is easy to check that $(x,s) \in E$ for any $s \geq t$ (there is $(x_{n_k}, t_{n_k}) \in E_{n_k}$ converging to (x,t), but then, obviously, $(x_{n_k}, t_{n_k} + s - t) \in E_{n_k}$ for all k...)

Hence there is $\mathcal{E}$ such that $\mathrm{epi}\,\mathcal{E} = E$ ($\mathcal{E}(x) = \inf\{s : (x,t) \in E\} \in [-\infty, +\infty]$). To simplify we now denote the converging subsequence $(E_{n_k})_k$ simply $(E_n)_n$. Now, consider $x_n \to x$. Observe that $(x_n, \mathcal{E}_n(x_n)) \in E_n$. If $\liminf_n \mathcal{E}_n(x_n) < +\infty$ there exists a subsequence such that $\lim_k \mathcal{E}_{n_k}(x_{n_k}) = \liminf_n \mathcal{E}_n(x_n)$. Hence, $(x_{n_k}, \mathcal{E}_{n_k}(x_{n_k})) \to (x, \liminf_n \mathcal{E}_n(x_n))$ so that

$$(x, \liminf_n \mathcal{E}_n(x_n)) \in \operatorname*{Ls}_n E_n = E.$$

By definition of $\mathcal{E}$, it follows that $\mathcal{E}(x) \leq \liminf_n \mathcal{E}_n(x_n)$, that is (i). (Observe that the epigraph of the Γ-liminf is thus the upper Kuratowski limit of the epigraphs.)

On the other hand, given $x \in X$, if $\mathcal{E}(x) < +\infty$ then for any $t \in \mathbb{R}$ with $t \geq \mathcal{E}(x)$ there is $(x_n, t_n) \in E_n$ such that $x_n \to x$, $t_n \to t$ so that $\limsup_n \mathcal{E}_n(x_n) \leq t$. If $\mathcal{E}(x) > -\infty$, choosing $t = \mathcal{E}(x)$ shows (ii). Else one must take a sequence $t_k \downarrow -\infty$ and use a diagonal sequence to show (ii). □

WARNING: all these beautiful results are *absolutely useless* (a) if one does not know how to bound globally $\mathcal{E}_n$, so that $\mathcal{E}$ is not infinite everywhere; (b) if one has a way to show that minimizing sequences will have converging subsequence, so that the Γ-convergence gives an information about the limit.

Definition 11. *We say that the sequence* $(\mathcal{E}_n)_n$ *which* Γ*-converges to* $\mathcal{E}$ *also satisfies the compactness property if for any sequence* $(x_n)_n$ *such that* $\sup_n \mathcal{E}_n(x_n) < +\infty$*, there exists a subsequence and* $x \in X$ *such that* $x_{n_k} \to x$ *as* $k \to \infty$*.*

In this case, clearly, if (x_n) is a minimizer (or almost minimizing in the sense of Theorem 10), and if one can show that the values $\mathcal{E}_n(x_n)$ are bounded, then x_n will have subsequences converging to minimizers of $\mathcal{E}$.

Let us know illustrate on a few examples this notion. We will consider:

1. Discrete-to-continuum limits;
2. Periodic Homogenization;
3. "Singular limits" (phase transition);
4. ?

6.2 Application: discrete to continuum limits

Here we will illustrate on a very basic examples a concept which can become much more complex in some contexts.

Let us just consider the function, defined for $y \in \mathbb{R}^{N+1}$ with $y_0 = y_N = 0$,

$$\mathcal{E}_N(y) = \sum_{i=0}^{N-1} \frac{kh}{2} \left(\frac{y_{i+1} - y_i}{h} \right)^2 + h \sum_{i=0}^{N} y_i g_i^N$$

where we have let $h = 1/N$, and $g_i^N = g(i/N)$ for a given continuous function $g : [0,1] \to \mathbb{R}$.

Defining

$$\mathcal{E}(y) = \frac{k}{2} \int_0^1 (y'(x))^2 dx + \int_0^1 y(x) g(x) dx$$

one expects that $\mathcal{E}_N \overrightarrow{\mathcal{E}}$. But in what sense is this true, since the functions do not "live" in the same space?

This is a very common issue in Γ-convergence problems. The solution is to embeds all variables in a common space (this way is in general not unique).

For instance, here, knowing that (Rellich) $\mathcal{E}$ (and hopefully the $\mathcal{E}_n$) is coercive in L^2 (which will help showing compactness), a good choice is to take $X = L^2(0,1)$. For $N > 0$ fixed, to a vector $y^N = (y_i^N)_{i=0}^N \in \mathbb{R}^{N+1}$ (with $y_0^N = y_N^N = 0$), one can associate

$$\bar{y}^N(x) = \sum_{i=1}^{N} y_i^N \chi_{((i-1)h, ih)}(x) \in L^2(0,1)$$

or

$$\hat{y}^N(x) = \sum_{i=1}^{N-1} y_i^N \Delta(Nx - i) \in H_0^1(0,1)$$

where $\Delta(x) = (1 - |x|)^+$ ("P1" approximation or spline of order 1). The advantage of the latter is that $\hat{y}^N \in H_0^1(0,1)$. If we can show that it is bounded in H_0^1, then it will be compact in L^2. However observe that in this case, both approximations will converge to the same limit, indeed

$$\int_0^1 |\hat{y}^N(x) - \bar{y}^N(x)|^2 dx = \sum_{i=1}^N \int_0^h |y_i^N - (y_i^N \tfrac{t}{h} + (1-\tfrac{t}{h})y_{i-1}^N)|^2 dx$$
$$= \sum_{i=1}^N h \int_0^1 (1-s)^2 |y_i^N - y_{i-1}^N|^2 dt$$
$$= \frac{h^2}{3} \sum_{i=1}^N h \left(\frac{y_i^N - y_{i-1}^N}{h}\right)^2 = \frac{h^2}{3} \int_0^1 |\hat{y}^{N\prime}(x)|^2 dx. \quad (6.2)$$

In order to talk about a Γ-limit, we extend the functionals as follows, letting for $u \in L^2(0,1)$

$$\bar{\mathcal{E}}_N(u) = \begin{cases} \mathcal{E}_N(y^N) & \text{if there exists } y^N \in \mathbb{R}^{N+1} \text{ with } y^0 = y^N = 0 \text{ and } u = \hat{y}^N, \\ +\infty \text{ else.} \end{cases}$$

and

$$\bar{\mathcal{E}}(u) = \begin{cases} \mathcal{E}(u) & \text{if } u \in H_0^1(0,1), \\ +\infty \text{ else.} \end{cases}$$

Then the result is as follows:

Proposition 9. *$\bar{\mathcal{E}}_N$ Γ-converges to $\bar{\mathcal{E}}$ in $L^2(0,1)$ as $N \to \infty$. Moreover, if $\sup_k \bar{\mathcal{E}}_{N_k}(u_{N_k}) < +\infty$ then (u_{N_k}) is bounded in $H_0^1(0,1)$, hence precompact in $L^2(0,1)$ (Rellich).*

Proof. First, the compactness is obvious, as one checks that

$$\sum_{i=0}^{N-1} \frac{kh}{2} \left(\frac{y_{i+1}^N - y_i^N}{h}\right)^2 = \frac{k}{2} \int_0^1 (\hat{y}^{N\prime}(x))^2 dx,$$

while

$$h \sum_{i=0}^N y_i g_i^N \le \max_{[0,1]} |g| \int_0^1 |\bar{y}^N| dx \le C \|\bar{y}^N\|_{L^2} \le C \left(\|\hat{y}^N\|_{L^2} + \frac{h}{\sqrt{3}} \|\hat{y}^{N\prime}\|_{L^2} \right).$$

We can conclude, using Poincaré's inequality, that if $\sup_N \mathcal{E}_N(y^N) < +\infty$, then $\hat{y}^N$ is bounded in $H_0^1(0,1)$, hence it has subsequences which converge in $L^2(0,1)$. (And, as well, (y^N).) Observe that in this case, also Ascoli Arzelà applies and the convergence is, in fact, uniform.

Let us show (i): let $u_N \to u$ in $L^2(0,1)$. If $\liminf_N \bar{\mathcal{E}}_N(u_N) = +\infty$ there is nothing to prove. Otherwise, there is a subsequence such that $\bar{\mathcal{E}}_{N_k}(u_{N_k}) < +\infty$ and $\lim_k \bar{\mathcal{E}}_{N_k}(u_{N_k}) = \liminf_N \bar{\mathcal{E}}_N(u_N)$. For each k, by definition there is $y^{N_k} \in \mathbb{R}^{N_k+1}$ such that $u_{N_k} = \hat{y}^{N_k}$ and $\bar{\mathcal{E}}_{N_k}(u_{N_k}) = \mathcal{E}(y^{N_k})$. Then, by the previous remark, both $\hat{y}^{N_k}$ and $\bar{y}^{N_k}$ converge to u in $L^2(0,1)$. One has that

$$\sum_{i=0}^{N_k-1} \frac{kh}{2} \left(\frac{y_{i+1}^{N_k} - y_i^{N_k}}{h}\right)^2 = \frac{k}{2} \int_0^1 (\hat{y}^{N_k\prime}(x))^2 dx$$

so that in particular

$$\frac{k}{2}\int_0^1 |u'(x)|^2 dx \le \liminf_k \sum_{i=0}^{N_k-1} \frac{kh}{2}\left(\frac{y_{i+1}^{N_k}-y_i^{N_k}}{h}\right)^2.$$

On the other hand,

$$h\sum_{i=0}^{N} y_i^N g_i^N = \int_0^1 \bar{y}^N(x)\bar{g}^N(x)dx$$

and since both $\bar{y}^N$ and $\bar{g}^N$ converge uniformly, respectively to u and g, it converges to $\int_0^1 y(x)g(x)dx$. This shows (i).

For (ii), the proof is here elementary. Indeed, if $u \in L^2(0,1)$ with $\bar{\mathcal{E}}(u) < +\infty$ (otherwise there is nothing to prove), so that $u \in H_0^1(0,1)$, u is also continuous and one can simply let for any N $y_i^N = u(i/N)$. One then checks that $u_N := \hat{y}^N \to u$ uniformly (and in $L^2(0,1)$), as well as $\bar{y}^N$. Now, one has that

$$h\sum_{i=0}^{N-1}\left(\frac{y_{i+1}^N - y_i^N}{h}\right)^2 = \frac{1}{h}\sum_{i=0}^{N-1}\left(\int_{ih}^{(i+1)h} \hat{y}^{N\prime}(x)dx\right)^2$$
$$\le \sum_{i=0}^{N-1}\int_{ih}^{(i+1)h}\left(\hat{y}'^N(x)\right)^2 dx = \int_0^1 |u_N'|^2 dx$$

while (by uniform convergence)

$$h\sum_{i=0}^{N} y_i^N g_i^N = \int_0^1 \bar{y}^N(x)\bar{g}^N(x)dx \to \int_0^1 u(x)g(x)dx.$$

□

As a consequence, minimizers of $\bar{\mathcal{E}}_N$ will converge to the (unique) minimizer of $\bar{\mathcal{E}}$. Observe however that for this problem, standard numerical analysis will also provide quite precise error bounds for this convergence.

6.2.1 Higher dimension

Here, we have used to build the "recovery sequence" the fact that u was in $H^1(0,1)$, hence continuous. How would we do this in higher dimension? Consider for instance, for $U = (u_{i,j})_{0\le i,j\le N}$ with $u_{0,j} = u_{N,j} = u_{i,0} = u_{i,N} = 0$ for all i,j, the energy (with still, $h = 1/N$)

$$\mathcal{E}_h(U) = h^2\sum_{i,j}\frac{|u_{i+1,j}-u_{i,j}|^2}{h^2} + \frac{|u_{i,j+1}-u_{i,j}|^2}{h^2}$$

which in the same way as before, will Γ-converge, in $L^2((0,1)^2)$ to

$$\mathcal{E}(u) = \int_{(0,1)^2} |\nabla u|^2 dx$$

(for $u \in H_0^1((0,1)^2)$, $+\infty$ else).

For the "liminf" (i) we do the same as in one dimension, considering the bilinear "Q1" approximation:

$$\hat{u}^N(x,y) = \sum_{i,j} u^N_{i,j} \Delta\left(\frac{x-ih}{h}\right)\Delta\left(\frac{y-jh}{h}\right).$$

One checks that

$$\frac{\partial \hat{u}^N}{\partial x} = \sum_{i,j}\left(\frac{u^N_{i+1,j}-u^N_{i,j}}{h}\right)\chi_{(ih,(i+1)h)}(x)\Delta\left(\frac{y-jh}{h}\right).$$

so that for fixed y,

$$\begin{aligned}\int_0^1 \left(\frac{\partial \hat{u}^N}{\partial x}\right)^2 dx &= \sum_{i=0}^{N-1}\sum_{j=1}^{N-1} h\left(\frac{u^N_{i+1,j}-u^N_{i,j}}{h}\Delta\left(\frac{y-jh}{h}\right)\right)^2 \\ &\le h\sum_{i=0}^{N-1}\sum_{j=1}^{N-1}\left(\frac{u^N_{i+1,j}-u^N_{i,j}}{h}\right)^2\Delta\left(\frac{y-jh}{h}\right)\end{aligned}$$

where we have used the convexity of $X \mapsto X^2$. Then

$$\begin{aligned}\int_0^1\int_0^1 \left(\frac{\partial \hat{u}^N}{\partial x}\right)^2 dx\,dy &\le h\sum_{i=0}^{N-1}\sum_{j=1}^{N-1}\left(\frac{u^N_{i+1,j}-u^N_{i,j}}{h}\right)^2\int_0^1\Delta\left(\frac{y-jh}{h}\right)dy \\ &= h^2\sum_{i=0}^{N-1}\sum_{j=1}^{N-1}\left(\frac{u^N_{i+1,j}-u^N_{i,j}}{h}\right)^2.\end{aligned}$$

Hence if $\hat{u}^N \to u$, one will have

$$\mathcal{E}(u) \le \liminf_N \int_{[0,1]^2} |\nabla u|^2 dx\,dy \le \liminf_N \mathcal{E}_N(u^N)$$

But: what about the limsup (ii)? We cannot, here, reproduce the same proof, since given $u \in H^1_0(0,1)$, $u(ih,jh)$ does not make any sense! We use a very classical important trick: first, we approximate an arbitray $u \in H^1_0((0,1)^2)$ with *simpler* functions u_k (here, in particular, smooth, or at least continuous) such that

$$\limsup_{k\to\infty} \mathcal{E}(u_k) \le \mathcal{E}(u).$$

Here this is easy, as we know that $C^\infty_c((0,1)^2)$ is dense in $H^1_0((0,1)^2)$.

For this smooth u_k, we do as in one dimension, approximating u_k with the discrete matrix $((u^N_k)_{i,j})_{i,j}$ given by $(u^N_k)_{i,j} = u_k(ih,jh)$. Then, as before it is quite easy to show that

$$\limsup_{N\to\infty} \mathcal{E}_N(u^N_k) \le \mathcal{E}(u_k).$$

The conclusion uses a diagonal procedure: we know that for all l, there exists $k(l) > k(l-1)$ such that

$$k \ge k(l) \Rightarrow \begin{cases} \|u_k - u\|_{L^2} \le \frac{1}{l} \\ \mathcal{E}(u_k) \le \mathcal{E}(u) + \frac{1}{l}. \end{cases}$$

Then, we find $N(l) \geq N(l-1)$ such that if $N \geq N(l)$,

$$N \geq N(l) \Rightarrow \begin{cases} \|u^N_{k(l)} - u_{k(l)}\|_{L^2} \leq \frac{1}{l} \\ \mathcal{E}(u^N_{k(l)}) \leq \mathcal{E}(u_{k(l)}) + \frac{1}{l}, \end{cases}$$

so that

$$N \geq N(l) \Rightarrow \begin{cases} \|u^N_{k(l)} - u\|_{L^2} \leq \frac{2}{l} \\ \mathcal{E}(u^N_{k(l)}) \leq \mathcal{E}(u) + \frac{2}{l}. \end{cases}$$

Eventually, we let $u^N = u^N_{k(l)}$ whenever $N(l) \leq N < N(l+1)$.

6.3 Application: periodic homogenization

(See [3].) We consider for $u \in W^{1,p}(\Omega;\mathbb{R}^m)$ ($\Omega \subset \mathbb{R}^d$ bounded open set) the functions

$$\mathcal{E}_\varepsilon(u) := \int_\Omega f\left(\frac{x}{\varepsilon}, Du\right) dx.$$

Here $f : \mathbb{R}^d \times \mathbb{R}^{m\times d} \to [0,+\infty)$ is a function which is periodic in the first variable: $f(y+k, A) = f(y, A)$ for all $y \in \mathbb{R}^d, A \in \mathbb{R}^{m\times d}$ and for all $k \in \mathbb{Z}^d$. It satisfies in addition the growth condition

$$\alpha(|A|^p - 1) \leq f(y, A) \leq \beta(|A|^p + 1)$$

for all y, A, where $p > 1$ is given.

In particular it follows

$$\alpha \int_\Omega |Du|^p - 1 dx \leq \mathcal{E}_\varepsilon(u) \leq \beta \int_\Omega 1 + |Du|^p dx.$$

By compactness, there exists a sequence $(\varepsilon_k) \downarrow 0$ and $\mathcal{E} : W^{1,p}(\Omega) \to \mathbb{R}_+ \cup \{+\infty\}$ such that $\mathcal{E}_{\varepsilon_k}$ Γ-converges to $\mathcal{E}$ in $L^p(\Omega)$ as $k \to \infty$, where for all u,

$$\alpha \int_\Omega |Du|^p - 1 dx \leq \mathcal{E}(u) \leq \beta \int_\Omega 1 + |Du|^p dx.$$

It turns out that the following result is true

Theorem 13. *The Γ-limit $\mathcal{E}$ in $L^p(\Omega;\mathbb{R}^m)$ is given, for $u \in W^{1,p}(\Omega)$, by*

$$\mathcal{E}(u) = \int_\Omega f_{hom}(Du)\, dx$$

where

$$f_{hom}(A) := \lim_{T\to\infty} \frac{1}{T^d} \inf\left\{ \int_{(0,T)^d} f(x, A + D\phi(x))dx \,:\, \phi \in W^{1,p}_0((0,T)^d;\mathbb{R}^m) \right\}. \tag{6.3}$$

The result is true, still, if one has a Dirichlet condition on part of the domain (observe that the statement for the Γ-limsup depends on the Dirichlet condition and could not have been true anymore), moreover in this case the compactness of bounded sequences is immediate (if Ω is regular) thanks to Rellich's theorem.

Remark: it is not clear that the limit defining f_{hom} exists. This is a standard fact. First, we can define

$$f_T(A) = \frac{1}{T^d} \inf \left\{ \int_{(0,T)^d} f(x, A + D\phi(x))dx \,:\, \phi \in W_0^{1,p}((O,T)^d;\mathbb{R}^m) \right\}.$$

Then, observe that at least if T is an integer, one easily sees that $f_{kT}(A) \le f_T(A)$ for all $k \in \mathbb{N}$, $k \ge 1$. Indeed if ϕ is almost optimal in the definition of f_T, one can use in $(0,kT)^d$ the function $\phi'(x) = \phi(y)$ whenever $x = lT + y$, $y \in (0,T)^d$, $l \in \mathbb{Z}^d$ ($l_i = \lfloor x_i/T \rfloor$, $i = 1,\ldots,d$), showing that $f_{kT}(A) \le \frac{1}{T^d}\int_{(0,T)^d} f(x, A + D\phi)dx$.

Moreover if S is large, choosing k_S such that $k_S T \le S < (k_S+1)T$, one sees that for $\phi \in W_0^{1,p}((0,k_ST)^d;\mathbb{R}^m)$,

$$f_S(A) \le \frac{(k_ST)^d}{S^d}\frac{1}{(k_ST)^d}\int_{(0,k_ST)^d} f(x, A+D\phi(x))dx + \frac{|(0,S)^d \setminus (0,k_ST)^d|\beta(1+|A|^p)}{S^d}$$

hence

$$f_S(A) \le f_{k_ST}(A) + C\left(1 - \frac{(k_ST)^d}{S^d}\right)(1+|A|^p) \le f_T(A) + C\left(1 - \frac{(k_ST)^d}{S^d}\right)(1+|A|^p).$$

Moreover since $(k_S+1)T \ge S$, $1 - (k_ST)^d/S^d \le 1 - (1-T/S)^d \to 1$ as $S \to \infty$, so that $\limsup_{S\to\infty} f_S(A) \le f_T(A)$. It follows (taking the liminf as $T \to \infty$) that the limit exists (at least along integers, but extending this to all real numbers is really easy since for large values of T $f_T(A)$ varies very little with T), and that it is, in fact, $\inf_T f_T(A)$.

Lemma 6. *f_{hom} is continuous.*

Proof. Fix $T > 2$, let A, B two matrices and t small. Then. Let $\phi \in W_0^{1,p}((1,T-1)^d;\mathbb{R}^m)$, $\eta \in C_c^\infty((0,T)^d;[0,1])$ be a cut-off function such that $\eta \equiv 1$ in $(1,T-1)^d$ and $|\nabla\eta| \le 2$.

$$\begin{aligned}
f_{hom}(A+tB) \le f_T(A+tB) &\le \frac{1}{T^d}\int_{(0,T)^d)} f(x, A+tB+D(\phi(x) - t\eta(x)Bx))dx \\
&= \frac{1}{T^d}\int_{(0,T)^d)} f(x, A + D\phi(x) + t(1-\eta)B - t(Bx)\otimes\nabla\eta)dx \\
\le \frac{1}{T^d}\int_{(1,T-1)^d)} & f(x, A+D\phi(x))dx + C\frac{1}{T^d}\int_{(0,T)^d\setminus(1,T-1)^d} |A+tB|^p + t^p|B|^p|x|^p dx \\
&\le \left(1-\tfrac{2}{T}\right)^d \frac{1}{(T-2)^d}\int_{(1,T-1)^d)} f(x, A+D\phi(x))dx \\
&\qquad + C\left(1 - \left(1-\tfrac{2}{T}\right)^d\right)\left(|A|^p + |tB|^p + t^p\sqrt{d}^pT^p|B|^p\right).
\end{aligned}$$

Given $\varepsilon > 0$, one can assume that T is large enough and ϕ is such that

$$\frac{1}{(T-2)^d}\int_{(1,T-1)^d)} f(x, A+D\phi(x))dx \le f_{hom}(A) + \varepsilon,$$

$$C\left(1 - \left(1-\tfrac{2}{T}\right)^d\right)(|A|^p + |B|^p) \le \varepsilon,$$

and it follows (if $t \in [0,1]$)

$$f_{hom}(A+tB) \le f_{hom}(A) + 2\varepsilon + \sqrt{d}^p t^p T^p \varepsilon \le f_{hom}(A) + 3\varepsilon$$

if t is then chosen small enough. □

Lemma 7. *f_{hom} is convex in along the directions $A + \mathbb{R}(a \otimes e_i)$ where $(e_i)_{i=1}^d$ is the canonical basis of $\mathbb{R}^d$.*

Remark: as a consequence of Thm. 13, f_{hom} will also be quasiconvex hence rank-one convex.

Proof. We give the main ideas. Let A, B with $B - A = a \otimes e_1$ and let $C = (A+B)/2$. Let $T > 0$ an integer and ϕ_A, ϕ_B such that $f_{hom}(A) \approx f_T(A)$ and $f_{hom}(B) \approx f_T(B)$. Consider ϕ_A, ϕ_B almost realizing the min in $f_T(A), f_T(B)$:

$$f_{hom}(A) \approx \frac{1}{T^d}\int_{(0,T)^d} f(x, A + D\phi_A)dx, \quad f_{hom}(B) \approx \frac{1}{T^d}\int_{(0,T)^d} f(x, B + D\phi_B)dx.$$

We define then (in $\mathbb{R}^d$) a function ϕ by letting $\phi(x) = Ax + \phi_A(x) + (k+1)Ta - Cx$ if $2kT \le x_1 \le (2k+1)T$ and $\phi(x) = Bx + \phi_B(x) - kTa - Cx$ if $(2k+1)T \le x_1 \le 2(k+1)T$. Then it is continuous accross the hyperplanes $x_1 = kT$, hence locally in $W^{1,p}$. For $S = 2nT$, $n \in \mathbb{N}$, large enough, we introduce the cut-off $\eta_S = 1$ on $(1, S-1)^d$, 0 on $\partial(0,S)^d$ with $|\nabla \eta_S| = 1$ on $(0,S)^d \setminus (1, S-1)^d$. Then we find

$$f_{hom}(C) \le \lim_{S\to\infty} \frac{1}{S^d}\int_{(0,S)^d} f(x, C + D(\phi\eta_S))dx.$$

Then, we have by construction that $C + D(\phi\eta_S) = A + D\phi_A$ if $x \in (1, S-1)^d$ and $2kT \le x_1 \le (2k+1)T$, $B + D\phi_B$ if $x \in (1, S-1)^d$ and $(2k+1)T \le x_1 \le 2(k+1)T$, while otherwise $C + D(\phi\eta_S) = C + \eta_S D\phi + \phi \otimes \nabla\eta_S$ and one checks easily that

$$\int_{(0,S)^d \setminus (1,S-1)^d} f(x, C + D(\phi\eta_S))dx \le cS^{d-1}$$

for a constant c depending only on A, B, ϕ_A, ϕ_B. It follows

$$\frac{1}{S^d}\int_{(0,S)^d} f(x, C + D(\phi\eta_S))dx \lessapprox \frac{f_{hom}(A) + f_{hom}(B)}{2} + \frac{c}{S}$$

and letting $S \to \infty$ we find that f_{hom} is convex along the line (A, B). □

Corollary 2. *Estimate* (5.1) *holds for f_{hom}.*

Proof. It is enough to observe that the proof of this estimate uses only the convexity in the directions $a \otimes e_i$ for (e_i) the canonical basis. □

6.3.1 Proof of the homogenization result

limsup

We start with the limsup inequality which is simpler. Given $u \in W^{1,p}(\Omega)$ we know that there exists a sequence of piecewise affine functions u_n which converge to u in $W^{1,p}(\Omega)$ (assuming the boundary of Ω is Lipschitz) ("P1" finite elements approximation). Thanks to estimate (5.1) (Corollary 2), one can show that

$$\lim_{n\to\infty}\int_\Omega f_{hom}(Du_n)dx = \int_\Omega f_{hom}(Du)dx$$

(of more generally, $u \mapsto \int_\Omega f_{hom}(Du)dx$ is continuous in $W^{1,p}(\Omega)$). We leave the proof as an exercise, it suffices to bound $|f_{hom}(Du) - f_{hom}(Du_n)|$ using estimate (5.1).

We then show that when u is piecewise affine, there exists u_ε which converges to u such that $\limsup_\varepsilon \mathcal{E}_\varepsilon(u) \le \mathcal{E}(u)$. A diagonal trick, exactly as in Section 6.2.1, allows to conclude.

To build u_ε one proceeds independently in each subdomain where u is affine. Hence one considers an open domain $O \subset \mathbb{R}^d$ and $u = Ax$ in O. Then given $\delta > 0$ one lets

$$O^\delta = \bigcup_{\substack{z\in\delta\mathbb{Z}^d\\ z+(0,\delta)^d\subset O}} z + (0,\delta)^d$$

and one observes that $|O \setminus O^\delta| \to 0$ as $\delta \downarrow 0$. Given $\varepsilon > 0$ small ($\varepsilon << \delta$) we let $T = \lfloor \frac{\delta}{T} \rfloor - 1$. If $\varepsilon > 0$ is small enough we know that we can find $\phi_\varepsilon \in W_0^{1,p}((0,T)^d;\mathbb{R}^m)$ such that

$$\frac{1}{T^d}\int_{(0,T)^d} f(y, A + D\phi_\varepsilon(y))dy \le f_{hom}(A) + \varepsilon.$$

We then let

$$u_\varepsilon(x) := \begin{cases} Ax & \text{if } x \in O \setminus O^\delta, \\ Ax + \varepsilon\phi_\varepsilon\left(\frac{x}{\varepsilon} - \lfloor\frac{z}{\varepsilon}\rfloor\right) & \text{if } x \in z + (0,\delta)^d \subset \Omega\,, z \in \delta\mathbb{Z}^d. \end{cases}$$

In each cube $z + (0,\delta)^d$, using the periodicity of f,

$$\begin{aligned}\int_{z+(0,\delta)^d} f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx &= \int_{z+(0,\delta)^d} f\left(\tfrac{x}{\varepsilon} - \lfloor\tfrac{z}{\varepsilon}\rfloor, A + D\phi_\varepsilon\left(\tfrac{x}{\varepsilon} - \lfloor\tfrac{z}{\varepsilon}\rfloor\right)\right)dx \\ &= \varepsilon^d \int_{(0,\frac{\delta}{\varepsilon})^d} f(y+\theta_\varepsilon, A + D\phi_\varepsilon(y+\theta_\varepsilon))dy\end{aligned}$$

where we have done the change of variables $y = (x-z)/\varepsilon$ and let $\theta_\varepsilon = z/\varepsilon - \lfloor z/\varepsilon \rfloor \in [0,1]$. By the choice of ϕ_ε, and using that $\theta_\varepsilon + (0,T)^d \subset (0,\delta/\varepsilon)^d$,

$$\int_{(0,\frac{\delta}{\varepsilon})^d} f(y+\theta_\varepsilon, A + D\phi_\varepsilon(y+\theta_\varepsilon))dy \le C\frac{\delta}{\varepsilon}^{d-1}|A|^p + T^d(f_{hom}(A)+\varepsilon).$$

Hence

$$\int_{z+(0,\delta)^d} f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx \le C\delta^{d-1}\varepsilon|A|^p + \delta^d\left(\frac{T}{\delta/\varepsilon}\right)^d (f_{hom}(A)+\varepsilon),$$

and, using $|O^\delta| = \delta^d \sharp\{z \in \mathbb{Z}^d \,:\, z + (0,\delta)^d \subset O\}$,

$$\int_O f\left(\tfrac{x}{\varepsilon}, Du_\varepsilon\right) dx \le C|O \setminus O^\delta||A|^p + C|O^\delta|\frac{\varepsilon}{\delta}|A|^p + |O^\delta|\left(\frac{T}{\delta/\varepsilon}\right)^d (f_{hom}(A)+\varepsilon),$$

so that

$$\limsup_{\varepsilon\to 0}\int_O f\left(\tfrac{x}{\varepsilon}, Du_\varepsilon\right) dx \le |O^\delta| f_{hom}(A) + C|O \setminus O^\delta||A|^p \overset{\delta\to 0}{\longrightarrow} |O| f_{hom}(A).$$

liminf

Let us consider first an easier case where we consider functions u_ε such that $u_\varepsilon \to u = Ax$ in $L^p(B;\mathbb{R}^m)$, where B is a ball (centered in 0), and we wish to show that

$$|B| f_{hom}(A) \le \liminf_{\varepsilon\to 0}\int_B f(\tfrac{x}{\varepsilon}, u_\varepsilon(x))dx. \tag{6.4}$$

The right-hand side, after a change of variable, is

$$\varepsilon^d \int_{B/\varepsilon} f(y, Dv_\varepsilon(u))dy$$

where we have let $v_\varepsilon(y) = u_\varepsilon(\varepsilon y)/\varepsilon$. We observe that if the liminf in (6.4) is finite there is a subsequence such that $\int_B f(x/\varepsilon_k, Du_{\varepsilon_k})dx$ is finite and goes to the liminf. In particular, Du_{ε_k} is globally bounded in $L^p(B)$, and therefore converges weakly to A. (As usual, since $u_{\varepsilon_k} \to Ax$, one writes for any smooth ψ with compact support

$$\int_B A\psi dx = -\int_B \operatorname{div}\psi Ax\, dx = -\lim_k \int_B \operatorname{div}\psi u_{\varepsilon_k} dx = \lim_k \int_B Du_{\varepsilon_k}\psi dx$$

showing that the weak limit of Du_{ε_k} cannot be anything but A.) In what follows, to simplify, we denote (u_ε) the subsequence.

In a first stage, we assume that we work in the unit cube $Q = (0,1)^d$ rather than in the ball B and we observe that if $T = 1/\varepsilon$,

$$\varepsilon^d \int_{Q/\varepsilon} f(y, Dv_\varepsilon)dy = \frac{1}{T^d}\int_{(0,T)^d} f(y, A + (Dv_\varepsilon - A))dy \gtrapprox f_{hom}(A)$$

(the inequality would be true if we had $v_\varepsilon - Ax = 0$ on $\partial(0,T)^d$).

To make this work we consider $\delta > 0$ and a cut-off $\eta \in C_c^\infty(Q;[0,1])$ with $|\nabla\eta| \le C/\delta$ and $\eta \equiv 1$ on $Q^\delta = (\delta, 1-\delta)^d$. Then, we consider $u_\varepsilon^\delta = Ax + \eta(u_\varepsilon - Ax)$. If we replace u_ε with u_ε^δ, the inequality above becomes true and we indeed can assert that

$$f_{hom}(A) \le \int_Q f(\tfrac{x}{\varepsilon}, Du_\varepsilon^\delta(x))dx.$$

We need to understand the error between this and the energy of u_ε. A basic computation yields, using $Du_\varepsilon^\delta = A + (u_\varepsilon - Ax)\otimes\nabla\eta + \eta(Du_\varepsilon - A)$,

$$\begin{aligned}\int_Q f(\tfrac{x}{\varepsilon}, Du_\varepsilon^\delta(x))dx &\le \int_Q f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx \\ &\le C|Q\setminus Q^\delta|(1+|A|^p) + \frac{C}{\delta^p}\int_{Q\setminus Q^\delta}|u_\varepsilon - Ax|^p\, dx + C\int_{Q\setminus Q^\delta}|Du_\varepsilon|^p dx.\end{aligned}$$

Clearly, the first term in the right-hand side can be made arbitrarily small choosing δ small, and the second term goes to zero as $\varepsilon \to 0$ since we have assumed that $u_\varepsilon \to Ax$ strongly in $L^p(Q)$.

However, there is no clear way to bound the last term. A solution (due to Marcellini, see [6]) consists in considering not one but N cut-offs $\eta_1, \dots, \eta_N$, for a given integer N, such that $\eta_i \in C_c^\infty(Q^{\frac{(i-1)\delta}{N}})$, $\eta_i \equiv 1$ on $Q^{\frac{i\delta}{N}}$, $|\nabla \eta_i| \le CN/\delta$. We consider for ε the functions $u_\varepsilon^i = Ax + \eta_i(u_\varepsilon - Ax)$, and then the same computation as before shows that for each i, using $Du_\varepsilon^i = Du_\varepsilon + (1-\eta_i)(Du_\varepsilon - A) + (u_\varepsilon - Ax) \otimes \nabla \eta_i$,

$$\begin{aligned} f_{hom}(A) &\le \frac{1}{|Q^{\frac{(i-1)\delta}{N}}|} \int_{Q^{\frac{(i-1)\delta}{N}}} f(\tfrac{x}{\varepsilon}, Du_\varepsilon^i(x))dx \\ &\le \frac{1}{(1-2\frac{(i-1)\delta}{N})^d} \int_{Q^{\frac{(i-1)\delta}{N}}} f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx \\ &+ \frac{CN^p}{\delta^p} \int_{Q^{\frac{i\delta}{N}} \setminus Q^{\frac{(i-1)\delta}{N}}} |u_\varepsilon - Ax|^p dx + \int_{Q^{\frac{i\delta}{N}} \setminus Q^{\frac{(i-1)\delta}{N}}} |Du_\varepsilon - A|^p dx. \end{aligned}$$

Now, since Du_ε is bounded in L^p, one has for a constant C that

$$\sum_{i=1}^N \int_{Q^{\frac{i\delta}{N}} \setminus Q^{\frac{(i-1)\delta}{N}}} |Du_\varepsilon - A|^p dx \le \int_Q |Du_\varepsilon - A|^p dx \le C$$

so that there exists, for each $\varepsilon > 0$, an index $i(\varepsilon)$ for which

$$\int_{Q^{\frac{i\delta}{N}} \setminus Q^{\frac{(i-1)\delta}{N}}} |Du_\varepsilon - A|^p dx \le \frac{C}{N}.$$

Choosing for each ε this "good" index i, we find that

$$f_{hom}(A) \le \frac{1}{(1-2\frac{\delta}{N})^d} \int_Q f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx + \frac{CN^p}{\delta^p} \int_Q |u_\varepsilon - Ax|^p dx + \frac{C}{N}$$

so that (using $u_\varepsilon \to Ax$),

$$f_{hom}(A) \le \frac{1}{(1-2\frac{\delta}{N})^d} \liminf_{\varepsilon \to 0} \int_Q f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx + \frac{C}{N}.$$

It remains to send first $N \to \infty$, then $\delta \to 0$, to conclude.

Next step: from Q to B: the idea is elementary, it consists in covering B with smaller and smaller disjoint cubes, and use the previous result: if $(Q_i)_{i=1}^K$ are disjoint cubes in B one has

$$\begin{aligned} |B| f_{hom}(A) &= |B \setminus \bigcup_{i=1}^K Q_i| f_{hom}(A) + \sum_{i=1}^K |Q_i| f_{hom}(A) \\ &\le |B \setminus \bigcup_{i=1}^K Q_i| f_{hom}(A) + \sum_{i=1}^K \liminf_{\varepsilon \to 0} \int_{Q_i} f(\tfrac{x}{\varepsilon}, Du_\varepsilon)dx \\ &\le |B \setminus \bigcup_{i=1}^K Q_i| f_{hom}(A) + \liminf_{\varepsilon \to 0} \int_B f(\tfrac{x}{\varepsilon}, Du_\varepsilon)dx. \end{aligned}$$

Taking more and more many cubes, one can make the first term in the right-hand side arbitrarily small.

Localization? Now, we want to consider an arbitrary $u_\varepsilon \to u$ in $L^p(\Omega;\mathbb{R}^m)$, and show that

$$\int_\Omega f_{hom}(Du(x))dx \leq \liminf_{\varepsilon\to 0} \int_\Omega f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx.$$

As before we assume that the right-hand side is not $+\infty$ and we consider a subsequence such that

$$\lim_{k\to\infty} \int_\Omega f(\tfrac{x}{\varepsilon_k}, Du_{\varepsilon_k}(x))dx = \liminf_{\varepsilon\to 0} \int_\Omega f(\tfrac{x}{\varepsilon}, Du_\varepsilon(x))dx.$$

The idea is to "localize" the problem near each point, use $Du_{\varepsilon_k} \simeq Du(x)$ near each point x and use the previous result. The classical method is due initially to Fonseca and Müller. The first observation is that $\mu_k = f(\frac{x}{\varepsilon_k}, Du_{\varepsilon_k}(x))dx$ is a sequence of bounded measures which are uniformly bounded, so that we may assume that in the sense of measures, $\mu_k \stackrel{*}{\rightharpoonup} \mu$ where μ is a nonnegative measure, which means that for any $\phi \in C_c^0(\Omega)$,

$$\lim_{k\to\infty} \int_\Omega \phi(x)d\mu_k(x) = \int_\Omega \phi(x)d\mu(x).$$

We know that

$$\mu(\Omega) \leq \liminf_{k\to\infty} \mu_k(\Omega)$$

so that the Γ-liminf inequality (ii) will hold if we can show that $\mu \geq f_{hom}(Du(x))dx$.

By Radon-Nikodym's theorem, we know that $\mu = g(x)dx + \nu$ with $\nu \perp dx$, $\nu \geq 0$, and

$$g(x) = \lim_{\rho\to 0} \frac{\mu(B_\rho(x))}{|B_\rho|}$$

(Lebesgue)-almost-everywhere. Hence we need to show that $g(x) \geq f_{hom}(x, Du(x))$, dx-a.e..

An important result is the following:

Lemma 8. *Given $x \in \Omega$, for a.e. $\rho > 0$ one has*

$$\lim_{k\to\infty} \mu_k(B_\rho(x)) = \mu(B_\rho(x)).$$

Proof. If $B_\rho(x)$ is the open ball of center x and radius ρ, then

$$\mu(B_\rho(x)) \leq \liminf_{k\to\infty} \mu_k(B_\rho(x)).$$

An obvious reason is that given ν a nonnegative measure,

$$\nu(B_\rho(x)) = \sup\left\{\int_{B_\rho(x)} \phi(x)d\nu \,:\, \phi \in C_c^0(B_\rho(x);[0,1])\right\}$$

and the supremum on the right-hand side is weakly lsc with respect to ν, as a sup of continuous functions with respect to weak-star convergences of measures.

For the same reason $\nu \mapsto \nu(\overline{B_\rho(x)})$ is usc, since it is an infimum of continuous functions:

$$\nu(\overline{B_\rho(x)}) = \inf\left\{\int_\Omega \phi(x)d\nu \,:\, \phi \in C_c^0(\Omega)\,, \phi \geq \chi_{B_\rho(x)}\right\}.$$

Hence, if ρ is such that $\mu(\partial B_\rho(x)) = 0$ (which is true for all but, at most, a countable number of radii since all these sets are disjoint when ρ varies)[1], one has

$$\mu(B_\rho(x)) \;\le\; \liminf_{k\to\infty} \mu_k(B_\rho(x)) \;\le\; \limsup_{k\to\infty} \mu_k(\overline{B_\rho(x)}) \;\le\; \mu(\overline{B_\rho(x)}) \;=\; \mu(B_\rho(x)),$$

showing the Lemma. □

Now, for each $n \ge 1$, one can find $\rho_n \le \rho_{n-1}$, not in the exceptional (countable) set of the previous Lemma, such that

$$\left|\frac{\mu(B_{\rho_n}(x))}{|B_{\rho_n}|} - g(x)\right| \le \frac{1}{n}$$

and then one can find $\varepsilon_{k_n} \le \varepsilon_{k_{n-1}}$ such that

$$\left|\frac{\mu_{k_n}(B_{\rho_n}(x))}{|B_{\rho_n}|} - \frac{\mu(B_{\rho_n}(x))}{|B_{\rho_n}|}\right| \le \frac{1}{n}.$$

We then denote, to simplify, ε_{k_n} with ε_n, and we let $\varepsilon'_n = \varepsilon_n/\rho_n$. Since ε_n is chosen "after ρ_n" (for a fixed, given value of ρ_n), we should have (and we can assume that) $\varepsilon'_n \to 0$ as $n \to \infty$.

Then, the change of variable $z = x + \rho_n y$ yields

$$\begin{aligned}\frac{\mu_{k_n}(B_{\rho_n}(x))}{|B_{\rho_n}|} &= \frac{1}{|B_{\rho_n}|}\int_{\rho_n}(x) f(\tfrac{z}{\varepsilon_n}, Du_{\varepsilon_n}(z))dz \\ &= \frac{1}{|B_1|}\int_{B_1} f(\tfrac{x}{\varepsilon_n} + \tfrac{y}{\varepsilon'_n}, Du_{\varepsilon_n}(x+\rho_n y))dy \\ &= \frac{1}{|B_1|}\int_{B_1} f(\theta_n + \tfrac{y}{\varepsilon'_n}, Du_{\varepsilon_n}(x+\rho_n y))dy = \frac{1}{|B_1|}\int_{B_1} f(\tfrac{y+\varepsilon'_n\theta_n}{\varepsilon'_n}, D\tilde{u}_n(y))dy\end{aligned}$$

where $\theta_n = x/\varepsilon_n - \lfloor x/\varepsilon_n \rfloor \in [0,1]^d$ and $\tilde{u}_n(y) := (u_{\varepsilon_n}(x+\rho_n y) - u_{\varepsilon_n}(x))/\rho_n$.

Assume we can show that $\tilde{u}_n \to Du(x)\cdot y$ in $L^p(B_1)$ as $n \to \infty$. Then we are back to the previously studied case of a sequence of functions converging to an affine function in a ball (modulo the varying translations $\varepsilon'_n\theta_n$ which can be easily handled) and it will follow that

$$f_{hom}(Du(x)) \le \liminf_{n\to\infty} \frac{1}{|B_1|}\int_{B_1} f(\tfrac{y+\varepsilon'_n\theta_n}{\varepsilon'_n}, D\tilde{u}_n(y))dy.$$

Putting everything together, we will deduce $f_{hom}(Du(x)) \le g(x)$, which was to prove. It remains to show that $\tilde{u}_n \to Du(x)\cdot y$. This will be true for almost all x, and relies on fine properties of $W^{1,p}$ functions.

A first observation is that for any $\varepsilon, \rho > 0$,

$$\begin{aligned}&\left\|\frac{u_\varepsilon(x+\rho\,\cdot) - u_\varepsilon(x)}{\rho} - \frac{u(x+\rho\,\cdot) - u(x)}{\rho}\right\|_{L^p(B_1)} \\ &\qquad\le \frac{1}{\rho}\left(\int_{B_1} |u_\varepsilon(x+\rho y) - u(x+\rho y)|^p dy\right)^{\frac{1}{p}} + \frac{1}{\rho}|B_1|^{\frac{1}{p}}|u_\varepsilon(x) - u(x)|.\end{aligned}$$

[1] A simple method to check such a fact is to consider the sets $I_i = \{\rho : \mu(\partial B_\rho) \ge 1/i\}$ and observe that $\mu(\Omega) \ge \mu(\bigcup_{\rho\in I_i}\partial B_\rho) \ge (1/i)\sharp I_i$ so that $\sharp I_i \le Ci$. Hence $\{\rho : \mu(\partial B_\rho) > 0\} = \bigcup_i I_i$ is countable.

Since we could have assumed that $u_\varepsilon \to u$ a.e. (extracting an appropriate subsequence), the right-hand side goes to zero as $\varepsilon \to 0$: therefore when we have chosen ε_{k_n} above, for fixed ρ_n, we could have required in addition that

$$\left\| \tilde{u}_n - \frac{u(x+\rho_n \cdot) - u(x)}{\rho_n} \right\|_{L^p(B_1)} \le \frac{1}{n}. \tag{6.5}$$

It remains to see that we could also have chosen ρ_n so that

$$\left\| \frac{u(x+\rho_n y) - u(x)}{\rho_n} - Du(x)\cdot y \right\|_{L^p_y(B_1)} \le \frac{1}{n}. \tag{6.6}$$

For fixed $\rho > 0$ (and assuming first that $u \in C^1$), one has

$$\begin{aligned}
&\int_{B_1} \left| \frac{u(x+\rho y) - u(x)}{\rho} - Du(x)\cdot y \right|^p dy \\
&\qquad = \int_{B_1} \left| \frac{1}{\rho} \int_0^\rho (Du(x+sy) - Du(x))\cdot y ds \right|^p dy \\
&\qquad \le \int_{B_1} \frac{1}{\rho} \int_0^\rho |(Du(x+sy) - Du(x))\cdot y|^p \, ds dy \\
&\qquad\qquad \le \frac{1}{\rho} \int_0^\rho \left[\frac{1}{s^d} \int_{B_s(x)} |Du(z) - Du(x)|^p dz \right] ds
\end{aligned}$$

This quantity goes a.e. to zero as $\rho \to 0$, since we have the well known following result:

Theorem 14. *Let $h \in L^p(\Omega)$. Then for a.e. $x \in \Omega$,*

$$\lim_{s\to 0} \frac{1}{|B_s|} \int_{B_s(x)} |h(z) - h(x)|^p dz = 0$$

Remark: x is called a "Lebesgue point" if there exists $\tilde{h}(x)$ such that $\frac{1}{|B_s|}\int_{B_s(x)} |h(z) - \tilde{h}(x)|^p dz \to 0$, and $\tilde{h} = h$ a.e. is called the "precise representative" of h.

Proof. The result is a consequence of Radon-Nikodym's theorem: if $f \in L^1_{loc}$, since $f(x)dx$ is absolutely continuous with respect to Lebesgue's measure, one has $f(x)dx = \tilde{f}(x)dx$ for $\tilde{f}$ given by the limit, which exists almost everywhere:

$$\tilde{f}(x) = \lim_{s\to 0} \frac{1}{|B_s|} \int_{B_s(x)} f(z)dz.$$

But clearly, it follows that $f = \tilde{f}$ a.e., in particular, almost everywhere

$$f(x) = \lim_{s\to 0} \frac{1}{|B_s|} \int_{B_s(x)} f(z)dz.$$

We apply this to the function $|h(x) - t|^p \in L^1_{loc}$, for all $t \in \mathbb{Q}$. Then, for all t, there exists a set E^t with $|E^t| = 0$ such that for $x \notin E^t$,

$$|h(x) - t|^p = \lim_{s\to 0} \frac{1}{|B_s|} \int_{B_s(x)} |h(z) - t|^p dz.$$

Then we let $E = \bigcup_{t\in\mathbb{Q}} E^t$, so that $|E| = 0$. One has if $x \notin E$ and $t \in \mathbb{Q}$,

$$\left(\frac{1}{|B_s|}\int_{B_s(x)} |h(z) - h(z)|^p dz\right)^{1/p} \le \left(\frac{1}{|B_s|}\int_{B_s(x)} |h(z) - t|^p dz\right)^{1/p} + |t - h(x)|$$

so that

$$\limsup_{s\to 0} \left(\frac{1}{|B_s|}\int_{B_s(x)} |h(z) - h(z)|^p dz\right)^{1/p} \le 2|h(x) - t|$$

Since t is arbitrary, this value can be sent to zero and the thesis follows.

Eventually, (ii) is proved, since we have now all the elements to see that $g(x) \ge f_{hom}(Du(x))$ a.e. □

6.3.2 A remark in the convex case

We assume that $f(y,\cdot)$ is convex. In this case, f_{hom} can be represented by a "cell formula" (see [3]). Let us define

$$\hat{f}_{hom}(A) = \inf\left\{\int_{(0,1)^d} f(y, A + Du(y))dy \,:\, u \in W^{1,p}_\sharp((0,1)^d;\mathbb{R}^m)\right\}$$

where $W^{1,p}_\sharp((0,1)^d)$ is the set of periodic functions in $W^{1,p}$ ($W^{1,p}(\mathbb{R}^d/\mathbb{Z}^d)$). Observe that in the convex case, the infimum defining $\hat{f}_{hom}$ is in fact a min and we can denote u_A a solution. The function u_A is called a *corrector.*

Theorem 15. $\hat{f}_{hom} = f_{hom}$.

Sketch of proof. Let T be a large integer. If η is a cut-off in $(0,T)^d$, we have

$$\begin{aligned} f_{hom}(A) &\le \frac{1}{T^d}\int_{(0,T)^d} f(y, A + D(\eta u_A))dy \\ &= \frac{1}{T^d}\int_{(0,T)^d} f(y, A + \eta Du_A + u_A\otimes\nabla\eta))dy \approx \hat{f}_{hom}(A) \end{aligned}$$

up to an error which vanishes when the cut-off tends to 1.

On the other hand, let T be a large integer and $v \in W^{1,p}_0((0,T)^d;\mathbb{R}^m)$ such that

$$f_{hom}(A) \approx \frac{1}{T^d}\int_{(0,T)^d} f(y, A + Dv)dy$$

We can consider v as a $T\mathbb{Z}^d$-periodic function (letting $v(y + kT) = v(y)$ for $y \in (0,T)^d$ and $k \in \mathbb{Z}^d$). Then, we let

$$w(x) = \frac{1}{T^d}\sum_{k\in\mathbb{Z}^d\cap[0,T[^d} v(x + k)$$

which is now $\mathbb{Z}^d$-periodic, so that

$$\hat{f}_{hom}(A) \le \int_{(0,1)^d} f(y, A + Dw(y))dy.$$

Now,

$$f(y, A + Dw(y)) = f\Big(y, A + \frac{1}{T^d} \sum_{k \in \mathbb{Z}^d \cap [0,T[^d} Dv(y+k)\Big)$$
$$\leq \frac{1}{T^d} \sum_{k \in \mathbb{Z}^d \cap [0,T[^d} f(y, A + Dv(y+k))$$

by convexity of f. We deduce

$$\hat{f}_{hom}(A) \leq \frac{1}{T^d} \sum_{k \in \mathbb{Z}^d \cap [0,T[^d} \int_{(0,1)^d} f(y, A + Dv(y+k)) = \frac{1}{T^d} \int_{(0,T)^d} f(y, A + Dv(y))dy.$$

□

Remark: the term "correctors" derives from the fact that, in the Γ-limsup, a way to correct the affine function Ax in order to recover the right energy is to approximate it with $Ax + \varepsilon u_A(x/\varepsilon)$.

Remark: What has been proved in case $f(x, A)$ depends only on A?

6.3.3 Phase transitions, singular limits

The Cahn-Hilliard energy in dimension 1

Higher dimension

Bibliography

[1] H. Attouch. Epi-convergence and duality. Convergence of sequences of marginal and Lagrangians functions. Applications to homogenization problems in mechanics. In *Optimization and related fields (Erice, 1984)*, volume 1190 of *Lecture Notes in Math.*, pages 21–56. Springer, Berlin, 1986.

[2] Andrea Braides. Γ*-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[3] Andrea Braides and Anneliese Defranceschi. *Homogenization of multiple integrals*, volume 12 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1998.

[4] Haïm Brézis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[5] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[6] B. Dacorogna. *Direct Methods in the Calculus of Variations*. Springer Verlag, Berlin, Heidelberg, 1989.

[7] Gianni Dal Maso. *An introduction to* Γ*-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston, Inc., Boston, MA, 1993.

[8] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, FL, 1992.

[9] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[10] Charles B. Morrey, Jr. *Multiple integrals in the calculus of variations*. Classics in Mathematics. Springer-Verlag, Berlin, 2008. Reprint of the 1966 edition [MR0202511].

[11] Walter Rudin. *Analyse réelle et complexe*. Masson, Paris, 1980. Translated from the first English edition by N. Dhombres and F. Hoffman, Third printing.

[12] Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[13] L. Schwartz. *Théorie des Distributions.* Hermann, 1966. (2 volumes).

[14] Laurent Schwartz. *Analyse.* Hermann, Paris, 1970. Deuxième partie: Topologie générale et analyse fonctionnelle, Collection Enseignement des Sciences, No. 11.