Lectures on Elliptic Partial Differential Equations

Luigi Ambrosio, Alessandro Carlotto, Annalisa Massaccesi

Contents

Preface

This text originated from the Ph.D. course given by the first author at the Scuola Normale Superiore in the academic year 2009-2010. The second and third authors, students of S.N.S. at that time, wrote down the first version of the notes. Since then the course has been replicated several times and the text has been gradually revised and expanded by the authors, until the present version.

All the material covered in this book is by now classical, the most recent contents being related to the theory of viscosity solutions. The presentation, the sequence of topics and even many proofs have been strongly inspired by the books by Giaquinta [41] and by Caffarelli-Cabré [15], the latter for the regularity theory of viscosity solutions. As in Evans' book [35], the guiding philosophy has been not to present the results in their maximal generality and under sharp technical assumptions, but rather to introduce the reader to the basic techniques of this beautiful and fundamental theory through the discussion of model cases. For this reason, this work is meant as a textbook and not as a reference book. Besides the treatises we already mentioned, the interested reader can consult many excellent and more systematic books on elliptic partial differential equations such as [44], [45], [49], [67], [83] among others, as well as the forthcoming monograph [74] on nonlinear Calderón-Zygmund theory. Further sources, more specific to the Calculus of Variations are, for instance, [24] and [47].

Ph.D. courses at S.N.S. are often attended also by students in the last years of their undergraduate studies. For this reason the course, while advanced in many respects, is meant to be self-contained, with four short appendices on Sobolev spaces, basic Real and Harmonic Analysis, Hausdorff measures and convex functions. Taking these appendices into account, only a basic knowledge of Functional Analysis and Measure Theory (and preferably also some fluency with Sobolev spaces of functions of one independent variable) is required.

Special thanks go to Guido De Philippis and Giuseppe Mingione, for their valuable comments. We also wish to thank the many students that, over the years, pointed out typos and inaccuracies in the preliminary versions.

Main Notation

$A \subset B$	inclusion of sets, not necessarily strict;
$A \Subset B$	relatively compact inclusion, meaning that $\overline{A}$ is a compact subset of B;
$\langle \cdot, \cdot \rangle$	standard inner product in $\mathbb{R}^d$;
$\lvert \cdot \rvert$	Euclidean norm in $\mathbb{R}^d$, induced by the standard inner product;
$B_r(x)$	ball with center x and radius r (also $B_r = B_r(0)$, $B = B_1$);
$\mathscr{L}^n$	Lebesgue measure in $\mathbb{R}^n$;
ω_n	volume of the unit ball in $\mathbb{R}^n$;
$L^p(\Omega; \mathbb{R}^m)$	Lebesgue L^p space of $\mathbb{R}^m$-valued functions defined over Ω;
$\fint_\Omega u \, d\mu$	mean value, namely $\int_\Omega u \, d\mu / \mu(\Omega)$ (also denoted u_Ω);
$u_{x,s}$	mean value of the function u on the ball $B_s(x)$;
$\mathrm{Lip}(\Omega; \mathbb{R}^{\mathrm{m}})$	space of $\mathbb{R}^m$-valued functions (uniformly) Lipschitz continuous in Ω;
$C^k(\Omega; \mathbb{R}^m)$	space of $\mathbb{R}^m$-valued functions continuously k-differentiable in Ω;
$C^{k,\alpha}(\Omega; \mathbb{R}^m)$	subspace of $C^k(\Omega; \mathbb{R}^m)$ of functions with k-th derivatives α-Hölder continuous;
$C_c^\infty(\Omega; \mathbb{R}^m)$	space of $\mathbb{R}^m$-valued smooth functions compactly supported in Ω;
$\partial_{x_\alpha} u$	α-th partial derivative (weak or classical);
∇u	gradient of u (weak or classical);
$\nabla^2 u$	Hessian of u (weak or classical);
Δu	Laplacian of u (weak or classical);
$W^{k,p}(\Omega; \mathbb{R}^m)$	Sobolev $W^{k,p}$-space of $\mathbb{R}^m$-valued functions defined over Ω;
$W_0^{k,p}(\Omega; \mathbb{R}^m)$	subspace of $W^{k,p}(\Omega; \mathbb{R}^m)$ of functions with null trace at the boundary;
$H^{k,p}(\Omega; \mathbb{R}^m)$	Sobolev $H^{k,p}$-space of $\mathbb{R}^m$-valued functions defined over Ω;
$H_0^{k,p}(\Omega; \mathbb{R}^m)$	subspace of $H^{k,p}(\Omega; \mathbb{R}^m)$ of functions with null trace at the boundary;
$H^k(\Omega; \mathbb{R}^m)$	Hilbertian Sobolev $H^{k,2}$-space of $\mathbb{R}^m$-valued functions defined over Ω;
$H_0^k(\Omega; \mathbb{R}^m)$	subspace of $H^k(\Omega; \mathbb{R}^m)$ of functions with null trace at the boundary.

1 Variational aspects of some classes of elliptic problems

We shall start our discussion presenting some basic existence results of weak solutions for certain classes of linear elliptic partial differential equations, see also [6],[8] or [20].

Let us consider the generalized Poisson equation

$$\begin{cases} -\Delta u = f - \sum_\alpha \partial_{x_\alpha} F^\alpha \\ u \in H_0^1(\Omega;\mathbb{R}) \end{cases} \tag{1.1}$$

with data f, $F^\alpha \in L^2(\Omega;\mathbb{R})$ for some open, bounded and regular domain $\Omega \subset \mathbb{R}^n$. We shall convene that the word *regular* is used in this monograph to describe any domain Ω that is locally the epigraph of a C^1 function of $(n-1)$ variables, written in a suitable system of coordinates, near any boundary point.

This equation has to be intended in a weak sense, meaning the following: we look for functions $u \in H_0^1(\Omega;\mathbb{R})$ satisfying

$$\int_\Omega \langle \nabla u, \nabla \varphi \rangle \, dx = \int_\Omega (f\varphi + \sum_\alpha F^\alpha \partial_{x_\alpha} \varphi) \, dx \qquad \forall \varphi \in C_c^\infty(\Omega;\mathbb{R}). \tag{1.2}$$

Equivalently, by continuity of the bilinear form in the left-hand side of (1.2) and density of $C_c^\infty(\Omega;\mathbb{R})$, the previous condition could be requested to hold for any $\varphi \in H_0^1(\Omega;\mathbb{R})$.

In order to obtain existence we just need to apply Riesz's theorem to the associated *linear* functional $F(v) = \int_\Omega (fv + \sum_\alpha F^\alpha \partial_{x_\alpha} v) \, dx$ on the Hilbert space $H_0^1(\Omega;\mathbb{R})$ endowed with the scalar product

$$(u,v) := \int_\Omega \langle \nabla u, \nabla v \rangle \, dx, \tag{1.3}$$

which is equivalent to the usual one thanks to the Poincaré inequality (first version) proved in Theorem A.12. Let us notice that, at this stage, no regularity assumption on Ω is needed.

We can consider many variants of the previous problem, basically by introduction of one or more of the following elements:

1. more general differential operators instead of the Laplacian;
2. inhomogeneous or mixed boundary conditions;
3. systems instead of single equations.

Our purpose now is to briefly discuss each of these situations, before moving to more general existence results which are the object of the second part of this chapter.

1.1 Weak solvability results

The first variant is to consider scalar problems having the form

$$\begin{cases} -\sum_{\alpha,\beta} \partial_{x_\alpha}(A^{\alpha\beta}\partial_{x_\beta} u) = f - \sum_\alpha \partial_{x_\alpha} F^\alpha \\ \\ u \in H_0^1(\Omega;\mathbb{R}) \end{cases}$$

where, as before f, $F^\alpha \in L^2(\Omega;\mathbb{R})$, and $A \in \mathbb{R}^{n\times n}$ is a constant matrix satisfying the following requirements:

(i) $A^{\alpha\beta} \in \mathbb{R}^{n\times n}$ is symmetric, that is $A^{\alpha\beta} = A^{\beta\alpha}$

(ii) A has only positive eigenvalues, equivalently, $A \geq cI$ for some $c > 0$, in the sense of quadratic forms.

Here and in the sequel of this book we use the capital letter I to denote the identity $n \times n$ matrix. It is not difficult to show that a change of independent variables, namely letting $u(x) = v(A^{-1}x)$, transforms this problem into one of the form (1.1). For this reason it is appropriate to deal instead with the case of a non-constant matrix $A(x) \in L^\infty(\Omega;\mathbb{R}^{n\times n})$ satisfying:

(i) $A(x)$ is symmetric for $\mathscr{L}^n$-a.e. $x \in \Omega$;

(ii) there exists a positive constant c such that

$$A(x) \geq cI \quad \text{for } \mathscr{L}^n\text{-a.e. } x \in \Omega. \tag{1.4}$$

As indicated above, the previous problem has to be intended in weak sense and precisely we require

$$\int_\Omega \langle A\nabla u, \nabla\varphi\rangle \, dx = \int_\Omega (f\varphi + \sum_\alpha F^\alpha \partial_{x_\alpha}\varphi)\, dx \qquad \forall \varphi \in C_c^\infty(\Omega;\mathbb{R}). \tag{1.5}$$

By continuity and density, as explained above in a special case, this condition is equivalent to require the validity of the identity above for all $\varphi \in H_0^1(\Omega;\mathbb{R})$. In order to obtain existence we could easily modify the previous argument, using the equivalent scalar product

$$(u, v) := \int_\Omega \sum_{\alpha,\beta} A^{\alpha\beta}\partial_{x_\alpha} u \partial_{x_\beta} v \, dx \, .$$

However, in order to drop the assumption on the essential boundedness of A, we prefer here to proceed differently and introduce some ideas that belong to the so-called *direct method* of the Calculus of Variations. Let then $A : \Omega \to \mathbb{R}^{n\times n}$ be a measurable map satisying

the assumptions (i) and (ii) above, and let us consider the functional $\mathcal{F} : H^1_0(\Omega;\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\mathcal{F}(v) = \frac{1}{2}\int_\Omega \langle A\nabla v, \nabla v\rangle \, dx - \int_\Omega fv \, dx - \sum_\alpha \int_\Omega F^\alpha \partial_{x_\alpha} v \, dx. \tag{1.6}$$

First we note that, thanks to the assumption (1.4) on A, for all $\varepsilon > 0$ it holds

$$\mathcal{F}(v) \geq \frac{c}{2}\int_\Omega |\nabla v|^2 \, dx - \frac{1}{2\varepsilon}\int_\Omega (|f|^2 + \sum_\alpha |F^\alpha|^2) \, dx - \frac{\varepsilon}{2}\int_\Omega (v^2 + |\nabla v|^2) \, dx.$$

Choosing $\varepsilon < c/2$ we get

$$\mathcal{F}(v) \geq \frac{c}{4}\int_\Omega |\nabla v|^2 \, dx - \frac{1}{2\varepsilon}\int_\Omega (|f|^2 + \sum_\alpha |F^\alpha|^2) \, dx - \frac{\varepsilon}{2}\int_\Omega v^2 \, dx$$

and now, thanks to the Poincaré inequality (Theorem A.12), we can choose ε even smaller to obtain

$$\mathcal{F}(v) \geq \frac{c}{8}\int_\Omega |\nabla v|^2 \, dx - \frac{1}{2\varepsilon}\int_\Omega (|f|^2 + \sum_\alpha |F^\alpha|^2) \, dx.$$

In particular $\mathcal{F}$ is *coercive*, that is to say

$$\lim_{\|v\|_{H^1_0(\Omega;\mathbb{R})} \to +\infty} \mathcal{F}(v) = +\infty. \tag{1.7}$$

As a result, in order to look for its minimum points, it is enough to consider a closed ball of $H^1_0(\Omega;\mathbb{R})$.

Moreover $\mathcal{F}$ is *lower semicontinuous with respect to the weak convergence*. Indeed, using Fatou's lemma and the fact that $u_h \to u$ in $H^1(\Omega;\mathbb{R})$ implies $\nabla u_{h(k)} \to \nabla u$ $\mathscr{L}^n$-a.e. in Ω for a suitable subsequence $h(k)$, it is not difficult to prove that $\mathcal{F}$ is lower semicontinuous with respect to the strong convergence (we shall also prove this in Theorem 1.10, in a more general framework). In addition, $\mathcal{F}$ is convex, being the sum of a linear and a convex functional. It follows that $\mathcal{F}$ is also lower semicontinuous with respect to the weak convergence (this is a standard fact, see e.g. Corollary 3.8 in [8]).

Now, take any minimizing sequence (u_h) of $\mathcal{F}$: since $H^1_0(\Omega;\mathbb{R})$ is a reflexive space (being Hilbert), we can assume, possibly extracting a subsequence, that $u_h \rightharpoonup u$ for some $u \in H^1_0(\Omega;\mathbb{R})$. Hence, by lower semicontinuity,

$$\mathcal{F}(u) \leq \liminf_{h\to\infty} \mathcal{F}(u_h) = \inf_{H^1_0(\Omega;\mathbb{R})} \mathcal{F} \tag{1.8}$$

and we conclude that u is a (global) minimizer of $\mathcal{F}$. Actually, the functional $\mathcal{F}$ is strictly convex and so u is its unique minimizer.

If A is essentially bounded (meaning that $A(x) \in L^\infty(\Omega;\mathbb{R}^{n\times n})$) we get $d\mathcal{F}(u) = 0$, where $d\mathcal{F}$ is the differential in the Gateaux sense of $\mathcal{F}$, i.e.

$$d\mathcal{F}(u)\,[\varphi] := \lim_{\varepsilon\to 0} \frac{\mathcal{F}(u+\varepsilon\varphi) - \mathcal{F}(u)}{\varepsilon} \qquad \forall \varphi \in H^1_0(\Omega;\mathbb{R})\,.$$

A simple computation gives

$$d\mathcal{F}(u)\,[\varphi] = \int_\Omega \langle A\nabla u, \nabla\varphi\rangle\, dx - \int_\Omega f\varphi\, dx - \sum_\alpha \int_\Omega F^\alpha \partial_{x_\alpha}\varphi\, dx \tag{1.9}$$

and the desired result follows.

Even in the case when $|A| \in L^2_{\mathrm{loc}}$ we can still differentiate the functional, but a priori only along directions $\varphi \in C^\infty_c(\Omega;\mathbb{R})$, and recover that u satisfies the weak formulation of our equation.

1.2 Inhomogeneous boundary conditions

We now turn to study the following boundary value problem for $u \in H^1(\Omega;\mathbb{R})$:

$$\begin{cases} -\Delta u = f - \sum_\alpha \partial_{x_\alpha} F^\alpha & \text{in } \Omega \\ \\ u = g & \text{on } \partial\Omega \end{cases}$$

with $f,\ F^\alpha \in L^2(\Omega;\mathbb{R})$ and a suitable class of functions $g \in L^2(\partial\Omega;\mathbb{R})$.

Since the immersion $H^1(\Omega;\mathbb{R}) \hookrightarrow C^0(\overline{\Omega};\mathbb{R})$ does not hold if $n \geq 2$ (see Example 2.8), the boundary condition has to be considered in the *weak sense* described below. Unless otherwise stated, we indicate with Ω an open, bounded and regular subset of $\mathbb{R}^n$; recall that for any such domain the equality $H^{1,p}(\Omega;\mathbb{R}) = W^{1,p}(\Omega;\mathbb{R})$ holds, see Theorem A.7 in Appendix A.

Theorem 1.1. *For any $p \in [1,\infty)$ the restriction operator*

$$T : C^1(\overline{\Omega};\mathbb{R}) \to C^0(\partial\Omega;\mathbb{R}) \tag{1.10}$$

satisfies $\|Tu\|_{L^p(\partial\Omega;\mathbb{R})} \leq c(p,\Omega)\|u\|_{W^{1,p}(\Omega;\mathbb{R})}$. Therefore it can be uniquely extended to a linear and continuous operator from $H^{1,p}(\Omega;\mathbb{R})$ to $L^p(\partial\Omega;\mathbb{R})$. This operator is called trace operator *and it is still denoted by T.*

Proof. We prove the result only in the case when Ω is the subgraph of a C^1 function f inside the domain $\Omega' \times (a,b)$, where $\Omega' \subset \mathbb{R}^{n-1}$ open, $a' = \inf f > a$, showing the estimate in question for the piece of the boundary defined by

$$\Gamma := \{(x', f(x')) :\ x' \in \Omega'\}\,.$$

(Here we use the notation $x = (x', t)$ with $x' \in \Omega'$ and $t \in (a, b)$). The general case can be easily obtained by means of a partition of unity.

By the fundamental theorem of calculus, for all $t \in (0, a' - a)$ we have

$$|u(x', f(x') - t) - u(x', f(x'))|^p \leq \left| \int_{f(x')-t}^{f(x')} \partial_{x_n} u(x', r)\, dr \right|^p \leq (b-a)^{p-1} \int_a^{f(x')} |\partial_{x_n} u(x', r)|^p\, dr\,.$$

An integration with respect to x' now gives

$$\int_{\Omega'} |u(x', f(x') - t) - u(x', f(x'))|^p\, dx' \leq (b-a)^{p-1} \int_\Omega |\partial_{x_n} u|^p\, dx.$$

So, using the inequality $|r + s|^p \leq 2^{p-1}(|r|^p + |s|^p)$ and recalling the form of the area element of a graph, one gets

$$\frac{1}{\sqrt{1+L^2}} \int_\Gamma |u|^p\, d\sigma \leq 2^{p-1} \int_{\Omega'} |u(x', f(x') - t)|^p\, dx' + 2^{p-1}(b-a)^{p-1} \int_\Omega |\partial_{x_n} u|^p\, dx\,,$$

where L is the Lipschitz constant of f, namely

$$L := \sup_{x' \neq y'} \frac{|f(y') - f(x')|}{|y' - x'|}$$

which can obviously be bounded from above by the supremum of the length of the gradient ∇f since $f \in C^1(\overline{\Omega'}; \mathbb{R})$.

Now we average this estimate with respect to $t \in (0, a' - a)$ and exploit the fact that the determinant of the gradient of the map $(x', t) \mapsto (x', f(x') - t)$ is identically equal to -1, to conclude

$$\frac{1}{\sqrt{1+L^2}} \int_\Gamma |u|^p\, d\sigma \leq \frac{2^{p-1}}{a' - a} \int_\Omega |u|^p\, dx + 2^{p-1}(b-a)^{p-1} \int_\Omega |\partial_{x_n} u|^p\, dx. \quad \square$$

Because of the previous result, for $u \in H^{1,p}(\Omega; \mathbb{R})$ we will interpret the boundary condition $u|_{\partial\Omega} = g$ as

$$Tu = g. \tag{1.11}$$

It can also be easily proved that Tu is characterized by the identity

$$\int_\Omega u \partial_{x_\alpha} \varphi\, dx = -\int_\Omega \varphi \partial_{x_\alpha} u\, dx + \int_{\partial\Omega} \varphi\, Tu\, \nu_\alpha\, d\sigma \qquad \forall \varphi \in C^1(\overline{\Omega}; \mathbb{R}) \tag{1.12}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit normal vector, pointing out of Ω.

Remark 1.2. It is possible to show that the previously defined trace operator T is not surjective if $p > 1$ and that its image can be described in terms of fractional Sobolev spaces $W^{s,p}$, characterized by the finiteness of the integral

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^p}{|x - y|^{n+sp}} \, dx \, dy,$$

with $s = 1 - 1/p$ (see [2]). For instance, if Ω is a $(n+1)$-dimensional halfspace, then the image of the trace operator is exactly $W^{1-\frac{1}{p},p}(\partial\Omega; \mathbb{R})$ and a similar result holds for regular open sets. The borderline case $p = 1$ is special, and in this case E. Gagliardo proved in [39] the surjectivity of T.

We can now mimic the argument described in the previous section in order to obtain an existence result, provided the function g belongs to the image of T, thus assuming that there exists a function $\widetilde{u} \in W^{1,2}(\Omega; \mathbb{R}) = H^1(\Omega; \mathbb{R})$ such that $T\widetilde{u} = g$. Indeed, if this is the case, our problem is reduced to showing the existence of a solution for the equation

$$\begin{cases} -\Delta v = \widetilde{f} - \sum_\alpha \partial_{x_\alpha} \widetilde{F^\alpha} & \text{in } \Omega \\ \\ v \in H^1_0(\Omega; \mathbb{R}) \end{cases}$$

where $\widetilde{f} = f$ and $\widetilde{F^\alpha} = F^\alpha - \partial_{x_\alpha} \widetilde{u}$. This is precisely the first problem we have discussed above and so, denoted by v its unique solution, the function $u = v + \widetilde{u}$ satisfies both our equation and the required boundary condition.

Finally, let us discuss the so-called Neumann boundary conditions, involving the behavior of the normal derivative of u on the boundary. Like we did above, we shall denote by $\nu = (\nu_1, \ldots, \nu_n)$ the outward pointing unit normal of $\partial\Omega$. We consider a problem of the form

$$\begin{cases} -\sum_{\alpha,\beta} \partial_{x_\alpha}(A^{\alpha\beta} \partial_{x_\beta} u) + \lambda u = f - \sum_\alpha \partial_{x_\alpha} F^\alpha & \text{in } \Omega \\ \\ \sum_{\alpha,\beta} A^{\alpha\beta} \nu_\alpha \partial_{x_\beta} u = g & \text{on } \partial\Omega \end{cases}$$

with $A^{\alpha\beta}$ a real matrix and $\lambda > 0$ a fixed constant. In fact, if $\lambda = 0$, the existence of a solution to the Neumann problem is not guaranteed, as it is shown by the ordinary differential equation $u'' = 1$ with boundary conditions $u'(0_+) = 1$ and $u'(1_-) = 0$.

For the sake of brevity, we just discuss the case when $A^{\alpha\beta} = \delta^{\alpha\beta}$ and $F = 0$ so that the problem above takes the form

$$\begin{cases} -\Delta u + \lambda u = f & \text{in } \Omega \\ \\ \partial_\nu u = g & \text{on } \partial\Omega \end{cases}$$

In order to give it a precise meaning, we first observe that if u, $v \in C^1(\overline{\Omega};\mathbb{R})$ then

$$\int_\Omega \langle \nabla u, \nabla v\rangle \, dx = -\int_\Omega v\Delta u \, dx + \int_{\partial\Omega} v\partial_\nu u \, d\sigma. \tag{1.13}$$

Thus it is natural to require that, for any $v \in C^1(\overline{\Omega};\mathbb{R})$, a solution u satisfies

$$\int_\Omega [\langle \nabla u, \nabla v\rangle + \lambda uv] \, dx = \int_\Omega vf \, dx + \int_{\partial\Omega} vg \, d\sigma. \tag{1.14}$$

In order to obtain existence (and uniqueness) for this problem when $g \in L^2(\partial\Omega;\mathbb{R})$, it is enough to apply Riesz's theorem to the continuous linear functional

$$\mathcal{F}(v) = \int_\Omega vf \, dx + \int_{\partial\Omega} vg \, d\sigma$$

with respect to the bilinear form defined on $H^1(\Omega;\mathbb{R})$ by

$$a(u,v) := \int_\Omega [\langle \nabla u, \nabla v\rangle + \lambda uv] \, dx, \tag{1.15}$$

which is clearly equivalent to the standard Hilbert product on the same space since $\lambda > 0$.

1.3 Elliptic systems

In order to deal with systems, we first need to introduce an appropriate notation. We will consider functions $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ and, consequently, we will use Greek letters (say α, $\beta, \ldots$) in order to indicate the coordinates in the *starting domain* of such maps (so that α, $\beta \in \{1, 2, \ldots, n\}$), while we will use Latin letters (say $i, j, k, \ldots$) for the *target domain* (and hence i, $j \in \{1, 2, \ldots, m\}$). We will need to work with matrices with four indices like $A_{ij}^{\alpha\beta}$ (i.e. rank-four tensors) whose meaning should be clear from the context. Our first purpose now is to see whether it is possible to adapt some ellipticity condition (having the form $A \geq cI$ for some $c > 0$) to the case of vector-valued maps. The most natural idea is to define the *Legendre condition*

$$\sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta}\xi_\alpha^i\xi_\beta^j \geq c\,|\xi|^2\,, \qquad \forall \xi \in \mathbb{R}^{m\times n} \tag{1.16}$$

where $\mathbb{R}^{m\times n}$ indicates, as above, the space of $m\times n$ real matrices and c is a strictly positive constant. Let us employ such a condition in order to prove an existence and uniqueness result for the system

$$\begin{cases} -\sum_{\alpha,\beta,j} \partial_{x_\alpha}(A_{ij}^{\alpha\beta}\partial_{x_\beta}u^j) = f_i - \sum_\alpha \partial_{x_\alpha}F_i^\alpha \qquad i = 1, \ldots, m \\ \\ u \in H_0^1(\Omega;\mathbb{R}^m) \end{cases}$$

with data f_i, $F_i^\alpha \in L^2(\Omega;\mathbb{R})$.

The weak formulation of the problem is

$$\int_\Omega \sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta} \partial_{x_\beta} u^j \partial_{x_\alpha} \varphi^i \, dx = \int_\Omega \left[\sum_i f_i \varphi^i + \sum_{\alpha,i} F_i^\alpha \partial_{x_\alpha} \varphi^i \right] dx \tag{1.17}$$

to hold for every $\varphi \in C_c^1(\Omega;\mathbb{R}^m)$. Now, if the matrix $A_{ij}^{\alpha\beta}$ is symmetric with respect to the transformation $(\alpha, i) \to (\beta, j)$ (which is implied, for instance, by the symmetries in both (α, β) and (i, j)), then it defines a scalar product on $H_0^1(\Omega;\mathbb{R}^m)$ by means of the formula

$$(\varphi, \psi) = \int_\Omega \sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta} \partial_{x_\alpha} \varphi^i \partial_{x_\beta} \psi^j \, dx. \tag{1.18}$$

If, moreover, A satisfies the Legendre condition (1.16) for some $c > 0$, it is immediate to see that this scalar product is equivalent to the standard one (i.e. with $A_{ij}^{\alpha\beta} = \delta^{\alpha\beta}\delta_{ij}$) and so we are led to apply again Riesz's theorem to conclude the proof.

It should be noted that in the proof of such an existence result (and, in particular, in the scalar case) the symmetry hypothesis with respect to the transformation $(\alpha, i) \to (\beta, j)$ is not really necessary, since we can exploit the following:

Theorem 1.3 (Lax-Milgram, [68])**.** *Let H be a (real) Hilbert space and let $a : H \times H \to \mathbb{R}$ a bilinear, continuous and coercive form, the latter assumption meaning that*

$$a(u, u) \geq \lambda \, |u|^2 \quad \forall u \in H,$$

for some $\lambda > 0$. Then for any $F \in H^*$ there exists $u_F \in H$ such that $a(u_F, v) = F(v)$ for all $v \in H$.

Proof. By means of Riesz's theorem it is possible to define a linear operator $T : H \to H$ such that

$$a(u, v) = \langle Tu, v \rangle \quad \forall u, \, v \in H$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product on the Hilbert space H. Notice that such T is continuous since

$$\|Tu\|^2 = \langle Tu, Tu \rangle = a(u, Tu) \leq c \, \|u\| \, \|Tu\| \, ,$$

where c is a constant of continuity for $a(\cdot, \cdot)$ and hence one immediately derives an upper bound on the operator norm of T, namely $\|T\| \leq c$. Now we introduce the auxiliary bilinear form

$$\widetilde{a}(u, v) = \langle TT^* u, v \rangle = \langle T^* u, T^* v \rangle \, ,$$

for T^* the adjoint of T with respect to $\langle \cdot, \cdot \rangle$; we remark that $\widetilde{a}(\cdot, \cdot)$ is obviously symmetric and continuous. Moreover, thanks to the coercivity of a, we have that $\widetilde{a}$ is coercive too, because

$$\lambda \, \|u\|^2 \leq a(u, u) = \langle Tu, u \rangle = \langle u, T^* u \rangle \leq \|u\| \, \|T^* u\| = \|u\| \sqrt{\widetilde{a}(u, u)}$$

and so $\widetilde{a}(u,u) \geq \lambda^2 \|u\|^2$. Since $\widetilde{a}$ determines an equivalent scalar product on H we can apply again Riesz's theorem to obtain a vector $u'_F \in H$ such that

$$\widetilde{a}(u'_F, v) = F(v) \quad \forall\, v \in H.$$

By the definitions of T and $\widetilde{a}$ the thesis is proved by simply setting $u_F = T^* u'_F$: indeed

$$F(v) = \tilde{a}(u'_F, v) = \langle T^* u'_F, T^* v\rangle = \langle T u_F, v\rangle = a(u_F, v) \qquad \forall v \in H. \qquad \square$$

As indicated above, we now want to formulate a different notion of ellipticity for the case of vector-valued maps. To this aim, it is useful to start by analyzing the scalar case in more detail. First of all, we wish to compare the two following conditions:

(E) $A \geq \lambda I$, that is $\langle Av, v\rangle \geq \lambda |v|^2$ for all $v \in \mathbb{R}^{m\times n}$ (*ellipticity*);

(C) $a_A(u,u) = \int_\Omega \langle A\nabla u, \nabla u\rangle\, dx \geq \lambda \int_\Omega |\nabla u|^2\, dx$ for all $u \in H^1_0(\Omega;\mathbb{R}^m)$ (*coercivity*).

We stress that here and in the discussion below the matrix $A \in \mathbb{R}^{m\times n}$ is constant. It is obvious by integration that $(E) \Rightarrow (C)$ and we may wonder about the converse implication. As we will see below, this holds in the scalar case ($m = 1$, see Proposition 1.4) and fails in the vectorial case ($m > 1$, see Example 1.5).

Proposition 1.4. *Let (C) and (E) as above, $m = 1$ and let A be symmetric. Then (C) is equivalent to (E).*

Proof. Let us prove that (C) implies (E). The computations become more transparent if we work with functions having complex values, and so let us define for any $u,\ v \in H^1_0(\Omega;\mathbb{C})$ the Hermitian form

$$a_A(u,v) := \int_\Omega \left\langle A\nabla u, \overline{\nabla v}\right\rangle\, dx = \int_\Omega \sum_{\alpha,\beta} A^{\alpha\beta} \partial_{x_\alpha} u \overline{\partial_{x_\beta} v}\, dx.$$

Here $\nabla u \in \mathbb{C}^n$ stands for $\nabla \Re u + i \nabla \Im u$, where $\Re u$ and $\Im u$ are respectively the real and imaginary parts of u. A simple computation, exploiting the fact that the matrix A is symmetric (as it is the case in our setting), shows that

$$a_A(u,u) = a_A(\Re u, \Re u) + a_A(\Im u, \Im u).$$

Hence, (C) implies

$$a_A(u,u) \geq \lambda \int_\Omega |\nabla u|^2\, dx, \tag{1.19}$$

since patently $|\nabla u|^2 = |\nabla \Re u|^2 + |\nabla \Im u|^2$. Now consider a function $\varphi \in C^\infty_c(\Omega;\mathbb{R})$ and define $u_\tau(x) = \varphi(x) e^{i\tau x\cdot\xi}$. Since A is constant, we have that

$$\frac{1}{\tau^2} a_A(u_\tau, u_\tau) = \int_\Omega \varphi^2 \sum_{\alpha,\beta} A^{\alpha\beta}\xi_\alpha\xi_\beta\, dx + o_\tau = \sum_{\alpha,\beta} A^{\alpha\beta}\xi_\alpha\xi_\beta \int_\Omega \varphi^2\, dx + o_\tau$$

with $o_\tau \to 0$ as $\tau \to +\infty$, and similarly

$$\frac{1}{\tau^2}\int_\Omega |\nabla u_\tau|^2\, dx = \int_\Omega \varphi^2\, |\xi|^2\, dx + o_\tau\,.$$

Hence, exploiting our coercivity assumption and letting $\tau \to +\infty$ in (1.19) we get

$$\left(\sum_{\alpha,\beta} A^{\alpha\beta}\xi_\alpha\xi_\beta - \lambda\,|\xi|^2\right)\int_\Omega \varphi^2\, dx \geq 0 \tag{1.20}$$

which immediately implies the thesis (it is enough to choose φ not identically zero). $\square$

Actually, every single part of our discussion is still true in the case when $A^{\alpha\beta} = A^{\alpha\beta}(x)$ is, for instance, a bounded Borel function in Ω and we can conclude that (E) holds for $\mathscr{L}^n$-a.e. $x \in \Omega$: we just need to choose, in the very last step, an appropriate sequence of rescaled and normalized test functions concentrating around x_0, for any Lebesgue point x_0 of all components of A. The conclusion comes, in this situation, by invoking Lebesgue's differentiation theorem.

For the reader's convenience we recall here some basic facts concerning Lebesgue points. Given $f \in L^1_{\rm loc}(\mathbb{R}^n;\mathbb{R})$ and $x_0 \in \mathbb{R}^n$, we say that x_0 is a Lebesgue point for f if there exists $\lambda \in \mathbb{R}$ such that

$$\lim_{r\downarrow 0} \fint_{B_r(x_0)} |f(y) - \lambda|\, dy = 0. \tag{1.21}$$

In this case λ is unique and it is sometimes written

$$\lambda = \widetilde{f}(x_0) = \widetilde{\lim_{x\to x_0}} f(x). \tag{1.22}$$

Notice that both the notions of Lebesgue point and the value of $\widetilde{f}$ *are invariant* in the Lebesgue equivalence class of f. The *Lebesgue differentiation theorem* asserts that for $\mathscr{L}^n$-a.e. $x_0 \in \mathbb{R}^n$ the following two properties hold: x_0 is a Lebesgue point of f and $\widetilde{f}(x_0) = f(x_0)$. Notice however that the validity of the second property at a given x_0 *does depend* on the choice of a representative of f in the Lebesgue equivalence class.

Going back to the previous discussion, it is very interesting to note that the argument above does not give a complete equivalence when $m > 1$. On the one hand, the coercivity condition

$$a_A(u,u) \geq \lambda \int_\Omega |\nabla u|^2\, dx \qquad u \in H^1_0(\Omega;\mathbb{R}^m) \tag{1.23}$$

can be applied to test functions having the form $u_\tau(x) = \varphi(x) b e^{i\tau x\cdot a}$ with $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ and implies the *Legendre-Hadamard condition*

$$\sum_{\alpha,\beta,i,j} A^{\alpha\beta}_{ij}\xi^i_\alpha\xi^j_\beta \geq \lambda\,|\xi|^2 \qquad \text{for all } \xi = a\otimes b\,, \tag{1.24}$$

that is the *Legendre condition* restricted to rank-one matrices $\xi^i_\alpha = a_\alpha b^i$. On the other hand, explicit examples show that the Legendre-Hadamard condition is in general *strictly weaker* than the Legendre condition.

Example 1.5. When $m = n = 2$, let the tensor $A^{\alpha\beta}_{ij}$ be chosen so that

$$\sum_{\alpha,\beta,i,j} A^{\alpha\beta}_{ij}\xi^i_\alpha\xi^j_\beta = \det(\xi) + \varepsilon\,|\xi|^2 \tag{1.25}$$

with $\varepsilon \geq 0$. Since any rank one matrix has vanishing determinant, the Legendre-Hadamard condition with $\lambda = \varepsilon$ is fulfilled. On the other hand, our quadratic form, restricted to diagonal matrices with eigenvalues t and $-t$, equals

$$-t^2 + 2t^2\varepsilon.$$

It follows that even the Legendre condition with $\lambda = 0$ fails when $2\varepsilon < 1$.

Nevertheless, the Legendre-Hadamard condition is sufficient to imply coercivity:

Theorem 1.6 (Gårding, [48])**.** *Assume that the constant matrix $A^{\alpha\beta}_{ij}$ satisfies the Legendre-Hadamard condition for some positive constant λ and the symmetry condition $A^{\alpha\beta}_{ij} = A^{\beta\alpha}_{ji}$. Then $a_A(u,u) \geq \lambda \int |\nabla u|^2\, dx$ for all $u \in H^1(\mathbb{R}^n;\mathbb{R}^m)$.*

In the proof of Gårding's theorem (see [48]), we denote by $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ the Schwartz space of smooth $\mathbb{C}$-valued functions that decay at ∞ faster than any polynomial, and by $\widehat{\varphi}$ and $\widetilde{\varphi}$ the Fourier transform of φ and its inverse, respectively defined by

$$\widehat{\varphi}(\xi) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} \varphi(x)e^{-ix\cdot\xi}\, dx \tag{1.26}$$

and

$$\widetilde{\varphi}(x) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} \varphi(\xi)e^{ix\cdot\xi}\, d\xi. \tag{1.27}$$

We will also make use of the Plancherel identity:

$$\int_{\mathbb{R}^n} \widehat{\varphi}\overline{\widehat{\psi}}\, d\xi = \int_{\mathbb{R}^n} \varphi\overline{\psi}\, dx \qquad \forall \varphi,\, \psi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C}). \tag{1.28}$$

Proof. By density it is enough to prove the result when $u \in C^\infty_c(\mathbb{R}^n;\mathbb{R}^m)$ (indeed, recall that $H^1(\mathbb{R}^n;\mathbb{R}^m) = H^1_0(\mathbb{R}^n;\mathbb{R}^m)$ as was discussed after Definition A.10). In this case we use the representation

$$u(\xi) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} \varphi(x)e^{-ix\cdot\xi}\, dx\,,$$

that is $u(\xi) = \widehat{\varphi}(\xi)$ for some $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C}^m)$. Therefore, we have that

$$\partial_{\xi_\alpha} u^j(\xi) = -i\widehat{x_\alpha \varphi^j} \ ,$$

hence

$$\begin{aligned} a_A(u,u) &= \int_{\mathbb{R}^n} \sum_{\alpha,\beta,j,l} A^{\alpha\beta}_{jl} \frac{\partial u^j}{\partial \xi_\alpha} \overline{\frac{\partial u^l}{\partial \xi_\beta}} \, d\xi = -i^2 \sum_{\alpha,\beta,j,l} A^{\alpha\beta}_{jl} \int_{\mathbb{R}^n} \widehat{x_\alpha \varphi^j} \overline{\widehat{x_\beta \varphi^l}} \, d\xi \\ &= \int_{\mathbb{R}^n} \sum_{\alpha,\beta,j,l} A^{\alpha\beta}_{jl} (x_\alpha \varphi^j)(x_\beta \overline{\varphi^l}) \, dx \, , \end{aligned}$$

the last passage relying on the Plancherel identity (1.28) and the constancy of $A^{\alpha\beta}_{jl}$. Now, notice that we can apply our hypothesis to get

$$\sum_{\alpha,\beta,j,l} A^{\alpha\beta}_{jl} a_\alpha b^j a_\beta \overline{b^l} \geq \lambda |a|^2 |b|^2$$

with $a = x \in \mathbb{R}^n$ and $b = \varphi(x) \in \mathbb{C}^n$, so that

$$a_A(u,u) \geq \lambda \int_{\mathbb{R}^n} |x|^2 \, |\varphi(x)|^2 \, dx. \tag{1.29}$$

If we perform the same steps with $\delta^{\alpha\beta}\delta_{jl}$ in lieu of $A^{\alpha\beta}_{jl}$ we see at once that

$$\int_{\mathbb{R}^n} |\nabla u|^2 \, (\xi) \, d\xi = \int_{\mathbb{R}^n} |x|^2 \, |\varphi(x)|^2 \, dx. \tag{1.30}$$

Comparing (1.29) and (1.30) we conclude the proof. □

Remark 1.7. Gårding's theorem marks in some sense the difference between *pointwise* and *integral* inequalities. It is worth mentioning some related inequalities that are typically *nonlocal*, meaning that they do not arise from the integration of a pointwise inequality. An important example is *Korn's inequality* (see, for instance, [63]):

$$\int_{\mathbb{R}^n} |\nabla u|^p \, dx \leq c(n,p) \int_{\mathbb{R}^n} \left| \frac{\nabla u + (\nabla u)^t}{2} \right|^p dx \qquad \text{for all } u \in C^\infty_c(\mathbb{R}^n; \mathbb{R}^n), \tag{1.31}$$

for $p \in (1, \infty)$. A variant of this example is the *Korn-Poincaré inequality*: if Ω is an open, bounded and regular set in $\mathbb{R}^n$ and $p \in (1, \infty)$, then

$$\inf_{c \in \mathbb{R}^m, \, B^t = -B} \int_\Omega |u(x) - Bx - c|^p \, dx \leq c(\Omega, p) \int_\Omega \left| \frac{\nabla u + (\nabla u)^t}{2} \right|^p dx. \tag{1.32}$$

1.4 Necessary minimality conditions

The importance of the Legendre-Hadamard condition is also clear from a variational perspective. Indeed, let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function, that is $u \in W^{1,\infty}_{\rm loc}(\Omega;\mathbb{R}^m)$ for some open, bounded and regular set $\Omega \subset \mathbb{R}^n$, fix a Lagrangian L and define a functional

$$\mathcal{L}(u,\Omega') = \int_{\Omega'} L(x,u,\nabla u)\,dx \qquad \forall\,\Omega' \Subset \Omega.$$

We say that u is a *local minimizer* for $\mathcal{L}$ if

$$\mathcal{L}(u,\Omega') \le \mathcal{L}(v,\Omega') \qquad \text{for all } v \in W^{1,\infty}_{\rm loc}(\Omega;\mathbb{R}^m) \text{ such that } \{v \ne u\} \Subset \Omega' \Subset \Omega. \tag{1.33}$$

We will make the following standard assumptions on the Lagrangian: we assume that $L : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is Borel and, denoting the variables as (x,s,p), we assume that L is of class C^1 in (s,p) with

$$\sup_K \left(|L| + |\partial_s L| + |\partial_p L|\right) < +\infty \tag{1.34}$$

for any domain $K = \Omega' \times \{(s,p) : \ |s| + |p| \le R\}$ with $R > 0$ and $\Omega' \Subset \Omega$. In this case it is possible to show that the map

$$t \mapsto \int_{\Omega'} L(x,u+t\varphi,\nabla u + t\nabla\varphi)\,dx$$

is of class C^1 for all u, $\varphi \in W^{1,\infty}_{\rm loc}(\Omega;\mathbb{R}^m)$ and $\Omega' \Subset \Omega$, and its derivative equals

$$\int_{\Omega'} \left[\partial_s L(x,u+t\varphi,\nabla u + t\nabla\varphi)\cdot\varphi + \partial_p L(x,u+t\varphi,\nabla u + t\nabla\varphi)\cdot\nabla\varphi\right]\,dx$$

(the assumption (1.34) is needed to differentiate under the integral sign). As a consequence, if a locally Lipschitz function u is a local minimizer and $\varphi \in C^\infty_c(\Omega';\mathbb{R}^m)$, since $\mathcal{L}(u,\Omega') \le \mathcal{L}(u+t\varphi,\Omega')$ we can differentiate at $t=0$ to obtain

$$\int_{\Omega'} \Big[\sum_i \partial_{s_i} L(x,u,\nabla u)\varphi^i + \sum_{\alpha,i} \partial_{p_i^\alpha} L(x,u,\nabla u)\partial_{x_\alpha}\varphi^i\Big]\,dx = 0.$$

Hence, exploiting the arbitrariness of φ, we derive the *Euler-Lagrange equations* in the weak form:

$$\sum_\alpha \partial_{x_\alpha}\partial_{p_i^\alpha} L(x,u,\nabla u) = \partial_{s_i} L(x,u,\nabla u) \qquad i = 1,2,\dots,m.$$

Exploiting this idea, we can associate to many classes of partial differential equations appropriate energy functionals, so that the considered problem is nothing but the Euler-Lagrange equation for the corresponding functional. For instance, neglecting the boundary conditions (that can actually be taken into account by an appropriate choice of the ambient functional space), an equation having the form

$$-\Delta u = g(x, u)$$

arise from the functional

$$L(x, s, p) = \frac{1}{2}|p|^2 - \int_0^s g(x, r)\, dr.$$

Our aim is now to find another *necessary minimality condition* corresponding to the stability inequality

$$\frac{d^2}{dt^2}\mathcal{L}(u + t\varphi)\Big|_{t=0} \geq 0.$$

We need to add hypotheses on the Lagrangian L, in analogy with what has been done above: in this case we require that L is of class C^2 in (s, p) and

$$\sup_K \left(|\partial_s\partial_s L| + |\partial_s\partial_p L| + |\partial_p\partial_p L|\right) < +\infty$$

for any domain $K = \Omega' \times \{(s, p) : |s| + |p| \leq R\}$ with $\Omega' \Subset \Omega$. The necessary minimality condition is then given by

$$0 \leq \Gamma(\varphi, \varphi) = \int_\Omega \left[\langle A\nabla\varphi, \nabla\varphi\rangle + 2\langle B\nabla\varphi, \varphi\rangle + \langle C\varphi, \varphi\rangle\right] dx \qquad \forall \varphi \in C_c^\infty(\Omega; \mathbb{R}^m), \quad (1.35)$$

where the dependence on x and all indices are omitted for the notational convenience, and we have set

$$\begin{cases} A(x) = \partial_p\partial_p L(x, u(x), \nabla u(x)) \\ B(x) = \partial_p\partial_s L(x, u(x), \nabla u(x)) \\ C(x) = \partial_s\partial_s L(x, u(x), \nabla u(x)) \end{cases} \quad (1.36)$$

and we allow the symbol $\langle\cdot,\cdot\rangle$ to denote the standard Euclidean product in $\mathbb{R}^d$ where the value of d may vary as one considers different terms (in the equation (1.35) above, corresponding to the three terms we have $d = mn, d = m, d = m$ respectively).

We can finally obtain *pointwise* conditions on the local minimizer u by means of the following theorem, whose proof follows along the lines of the argument presented for Proposition 1.4 (namely the assertion that coercivity implies ellipticity when $m = 1$).

Theorem 1.8. *Consider the bilinear form on $H^1_0(\Omega;\mathbb{R}^m)$ defined by*

$$\Theta(u,v) = \int_\Omega (A\nabla u\nabla v + B\nabla u\cdot v + Cu\cdot v)\,dx\,, \tag{1.37}$$

where $A = A^{\alpha\beta}_{ij}(x)$, $B = B^{\alpha}_{ij}(x)$ and $C = C_{ij}(x)$ are Borel and L^∞ functions. Moreover, assume A to be symmetric, namely $A^{\alpha\beta}_{ij}(x) = A^{\beta\alpha}_{ji}(x)$ for $\mathscr{L}^n$-a.e. $x\in\Omega$. If $\Theta(u,u)\geq 0$ for all $u\in H^1_0(\Omega;\mathbb{R}^m)$ then $A(x)$ satisfies the Legendre-Hadamard condition with $\lambda = 0$ for $\mathscr{L}^n$-a.e. $x\in\Omega$.

Hence, in our case, we find that $\partial_p\partial_p L(x,u(x),\nabla u(x))$ satisfies the Legendre-Hadamard condition with $\lambda = 0$ for $\mathscr{L}^n$-a.e. $x\in\Omega$.

1.5 Lower semicontinuity of integral functionals

Tonelli's theorem is a first powerful tool leading to an existence result for minimizers of integral functionals of the form

$$\mathcal{L}(u) := \int_\Omega L(x,u(x),\nabla u(x))\,dx \tag{1.38}$$

in suitable functional spaces (which may allow, for instance, to impose various sorts of boundary conditions). With respect to the (somewhat more general) setting presented in the previous section, we take $\Omega' = \Omega \Subset \mathbb{R}^n$. In particular, notice that we work on a set of finite Lebesgue measure.

Before stating Tonelli's theorem [91] (see [83] for a broader discussion and contextualization), we recall a useful characterization of uniformly integrable families of functions. A comprehensive treatment of this subject can be found for instance in [97], see also [5, Theorem 1.38] or [19].

Theorem 1.9 (Dunford-Pettis, [33])**.** *Let $(X,\mathcal{A},\mu)$ be a finite measure space and let $\mathscr{F}\subset L^1((X,\mathcal{A},\mu);\mathbb{R})$. Then the following facts are equivalent:*

(i) the family $\mathscr{F}$ is sequentially relatively compact with respect to the weak-L^1 topology;

(ii) there exists $\phi:[0,\infty)\to[0,\infty]$, with $\phi(t)/t\to+\infty$ as $t\to\infty$, such that

$$\int_X \phi(|f|)\,d\mu \leq 1 \qquad \forall\, f\in\mathscr{F};$$

(iii) $\mathscr{F}$ is uniformly integrable, i.e.

$$\forall\,\varepsilon>0 \quad \exists\,\delta>0 \quad s.t. \quad \mu(A)<\delta \quad\Longrightarrow\quad \int_A |f|\,d\mu<\varepsilon \quad \forall\, f\in\mathscr{F}.$$

Theorem 1.10 (Tonelli, [91]). *Let $L : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a Lagrangian*[1] *satisfying the following properties:*

(1) L is non-negative;

(2) L is lower semicontinuous with respect to the variable s and the partial derivatives $\partial_{p_i^\alpha} L$ exist and are also continuous with respect to s;

(3) $p \mapsto L(x, s, p)$ is convex[2]*.*

Then any sequence $(u_h) \subset W^{1,1}(\Omega; \mathbb{R}^m)$ converging in $L^1(\Omega; \mathbb{R}^m)$ (the limit denoted by u), with $|\nabla u_h|$ uniformly integrable, satisfies the lower semicontinuity inequality

$$\liminf_{h\to\infty} \mathcal{L}(u_h) \geq \mathcal{L}(u). \tag{1.39}$$

Proof. We start by noticing that there is a subsequence $u_{h(k)}$ such that

$$\liminf_{h\to\infty} \mathcal{L}(u_h) = \lim_{k\to\infty} \mathcal{L}(u_{h(k)})$$

and, possibly extracting one more subsequence, we can assume that

$$u_{h(k)} \longrightarrow u \quad \mathscr{L}^n\text{-a.e. in } \Omega.$$

Thanks to Theorem 1.9 we can also assume the weak-L^1 convergence of the gradients

$$\nabla u_{h(k)} \rightharpoonup g \quad \text{in } L^1(\Omega; \mathbb{R}^{m\times n}).$$

Passing to the limit in the integration by parts formula, this immediately implies that u belongs to $W^{1,1}(\Omega; \mathbb{R}^m)$ and that $\nabla u = g$.

By virtue of Egorov's Theorem (see, for instance, [36]), for all $\varepsilon > 0$ there exists a compact subset $K_\varepsilon \subset \Omega$ such that

- $|\Omega \setminus K_\varepsilon| < \varepsilon$;
- $\partial_p L(x, u_{h(k)}(x), \nabla u(x)) \to \partial_p L(x, u(x), \nabla u(x))$ uniformly on K_ε;
- $\partial_p L(x, u(x), \nabla u(x))$ is bounded on K_ε.

[1] As usual, we denote the variables by $(x, s, p) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$.

[2] We will see that this assumption can be considerably weakened.

Because of the convexity assumption *(3)* and the non-negativity of L as per *(1)*, we can estimate

$$\begin{aligned}
&\liminf_{h\to\infty} \mathcal{L}(u_h) = \lim_{k\to\infty} \mathcal{L}(u_{h(k)}) = \lim_{k\to\infty} \int_\Omega L(x, u_{h(k)}(x), \nabla u_{h(k)}(x))\, dx \\
&\geq \liminf_{k\to\infty} \int_{K_\varepsilon} L(x, u_{h(k)}(x), \nabla u_{h(k)}(x))\, dx \\
&\geq \liminf_{k\to\infty} \int_{K_\varepsilon} \left[L(x, u_{h(k)}(x), \nabla u(x)) + \left\langle \partial_p L(x, u_{h(k)}(x), \nabla u(x)), \nabla u_{h(k)}(x) - \nabla u(x)\right\rangle\right] dx \\
&\geq \int_{K_\varepsilon} L(x, u(x), \nabla u(x))\, dx + \liminf_{k\to\infty} \int_{K_\varepsilon} \left\langle \partial_p L(x, u_{h(k)}(x), \nabla u(x)), \nabla u_{h(k)}(x) - \nabla u(x)\right\rangle dx.
\end{aligned}$$

Hence, the weak-L^1convergence $\nabla u_{h(k)} \rightharpoonup \nabla u$ ensures that

$$\liminf_{h\to\infty} \mathcal{L}(u_h) \geq \int_{K_\varepsilon} L(x, u(x), \nabla u(x))\, dx$$

and as $\varepsilon \to 0$ we achieve the desired inequality (1.39). □

Before stating the following corollary, we recall that Theorem A.13 provides the compactness of the inclusion $W^{1,1}(\Omega;\mathbb{R}) \subset L^1(\Omega;\mathbb{R})$ whenever $\Omega \subset \mathbb{R}^n$ is an open, bounded and regular set.

Corollary 1.11. *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded and regular set and let L be a Borel Lagrangian satisfying hypotheses (2), (3) as in Theorem 1.10 and*

(1') $L(x, s, p) \geq \phi(|p|) + c|s|$ for some $\phi : [0, \infty) \to [0, \infty]$ with $\lim_{t\to\infty} \phi(t)/t = \infty$, $c > 0$.

Then the problem

$$\min\left\{\mathcal{L}(u) :\ u \in W^{1,1}(\Omega;\mathbb{R}^m)\right\}$$

admits a solution.

Proof. The conclusion follows from a classical application of the direct method of Calculus of Variations, where hypothesis *(1')* provides the sequential relative compactness of sublevels $\{\mathcal{L} \leq t\}$ with respect to the so-called sequential weak-$W^{1,1}$ topology (i.e. strong convergence in L^1 of the functions and weak convergence in L^1 of their gradients) and semicontinuity is given by Theorem 1.10. □

At this point one could ask whether the convexity assumption in Theorem 1.10 is natural. The answer is negative: as the Legendre-Hadamard condition is weaker than the Legendre condition, here we are in an analogous situation and Example 1.5 fits again. Let us define a weaker, although much less transparent, convexity condition, introduced by Morrey in [75].

Definition 1.12 (Quasiconvexity)**.** *A continuous function $L : \mathbb{R}^{m\times n} \to \mathbb{R}$ is said to be* quasiconvex *at $A \in \mathbb{R}^{m\times n}$ if, for all $\Omega \subset \mathbb{R}^n$ open and bounded, it holds*

$$\fint_\Omega L(A + \nabla\varphi)\, dx \geq L(A) \qquad \forall\, \varphi \in C_c^\infty(\Omega; \mathbb{R}^m). \tag{1.40}$$

We say that L is quasiconvex if it is quasiconvex at every point $A \in \mathbb{R}^{m\times n}$.

Remark 1.13. Obviously we can replace the left-hand side in (1.40) with the quantity $\fint_{\{\nabla\varphi\neq 0\}} L(A + \nabla\varphi)\, dx$: this follows from the identity

$$\fint_\Omega L(A + \nabla\varphi)\, dx = \left(1 - \frac{|\{\nabla\varphi \neq 0\}|}{|\Omega|}\right) L(A) + \frac{|\{\nabla\varphi \neq 0\}|}{|\Omega|} \fint_{\{\nabla\varphi\neq 0\}} L(A + \nabla\varphi)\, dx.$$

This remark makes it clear that the role played by Ω in this definition is only fictitious. Let us further observe that whenever (1.40) is valid for Ω, then:

- it is valid for every $\Omega' \subset \Omega$, as follows immediately by the definition;
- it is valid for $x_0 + \lambda\Omega$, for $x_0 \in \mathbb{R}^n$ and $\lambda > 0$, by simply considering the transformation $\varphi(x) \mapsto \varphi(x_0 + \lambda x)/\lambda$.

Furthermore, a simple approximation argument gives

$$\fint_\Omega L(A + \nabla\varphi)\, dx \geq L(A) \qquad \forall\, \varphi \in \mathrm{Lip_c}(\Omega; \mathbb{R}^\mathrm{m}), \tag{1.41}$$

that is to say for all Lipschitz functions having compact support in Ω.

The definition of quasiconvexity is related to Jensen's inequality (see [50] or [84] for the proof), which we briefly recall here.

Theorem 1.14 (Jensen)**.** *Let us consider a probability measure μ on a closed convex domain $C \subset \mathbb{R}^d$, with $\int_C |p|\, d\mu(p) < +\infty$, and a convex, lower semicontinuous function $L : C \to \mathbb{R} \cup \{+\infty\}$. Then*

$$\int_C L(p)\, d\mu(p) \geq L\left(\int_C p\, d\mu(p)\right). \tag{1.42}$$

Notice that the left-hand side of such inequality always makes sense because the negative part of L has at most linear growth. Thus the integral is well-defined, attaining either a finite value or $+\infty$.

Now, let $f \in L^1(\Omega;\mathbb{R}^{m\times n})$ and consider the law μ of the map f with respect to the rescaled Lebesgue measure $\mathscr{L}^n/\mathscr{L}^n(\Omega)$, that is to say the push-forward measure defined for any Borel set $S \subset \mathbb{R}^{m\times n}$ by

$$\mu(S) = \frac{\mathscr{L}^n(f^{-1}(S))}{\mathscr{L}^n(\Omega)}.$$

Observe that the summability of $f \in L^1(\Omega;\mathbb{R}^{m\times n})$ gives $\int |p|\, d\mu < +\infty$. If $L : \mathbb{R}^{m\times n} \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous and convex, thanks to Jensen's inequality one has

$$\fint_\Omega L(f(x))\, dx = \int_{\mathbb{R}^{m\times n}} L(p)\, d\mu(p) \geq L\left(\int_{\mathbb{R}^{m\times n}} p\, d\mu(p)\right) = L\left(\fint_\Omega f\, dx\right). \tag{1.43}$$

In particular, we can prove the following assertion:

Proposition 1.15. *Any convex lower semicontinuous function $L : \mathbb{R}^{m\times n} \to \mathbb{R} \cup \{+\infty\}$ is quasiconvex.*

Proof. Since for any $\varphi \in C_c^\infty(\Omega;\mathbb{R}^m)$ one has

$$\fint_\Omega (A + \nabla\varphi(x))\, dx = A,$$

by applying (1.43) to the L^1 function $f(x) = A + \nabla\varphi(x)$ we conclude. □

Therefore, quasiconvexity should be considered as a weak version of convexity: indeed, if L is convex then the inequality (1.43) holds for all maps f, thanks to Jensen's inequality; on the other hand the condition (1.40) concerns only gradient maps (more precisely gradients of maps coinciding with an affine function on the boundary of the domain). Equivalently, quasiconvexity should be understood as defined by (1.42) for measures μ in $\mathbb{R}^{m\times n}$ generated by gradient maps (see [79] and [62]).

Remark 1.16. The following chain of implications holds:

$$\text{convexity} \implies \text{quasiconvexity} \implies \partial_p\partial_p L(A) \text{ satisfies Legendre-Hadamard with } \lambda = 0.$$

The second implication can be justified by differentiating twice the smooth map

$$t \mapsto \int_\Omega (A + t\nabla\varphi)\, dx,$$

which has a local minimum at $t = 0$ for any given choice of the test function φ (where Ω is any open and bounded set in $\mathbb{R}^n$ containing the support of φ). Indeed, with a straightforward computation one obtains

$$\int_\Omega \langle \partial_p\partial_p L(A)\nabla\varphi, \nabla\varphi\rangle\, dx \geq 0,$$

which implies the Legendre-Hadamard condition with $\lambda = 0$ (as already seen above, when we discussed how (1.23) implies (1.24)).

All these notions are equivalent when either $n = 1$ or $m = 1$. Furthermore:

- an integration by parts easily yields
$$\int_\Omega \det(A + \nabla\varphi)\, dx = \det(A)|\Omega| \qquad \forall \varphi \in C_c^\infty(\Omega; \mathbb{R}^n).$$
Hence, the determinant map $\det : \mathbb{R}^{2\times 2} \to \mathbb{R}$ (cf. Example 1.5) provides a quasiconvex function that is not convex when $n = m = 2$, and considering the determinant of a 2×2 minor allows to handle the case $\min\{m, n\} \geq 2$ as well;

- when $\max\{n, m\} \geq 3$ and $\min\{n, m\} \geq 2$, Šverak showed in [95] that there exist highly nontrivial examples, building on an algebraic result in [90], showing that the Legendre-Hadamard condition does not imply quasiconvexity (see also [24] and [32]);

- the equivalence between Legendre-Hadamard condition and quasiconvexity is still open for $n = m = 2$.

Let us recall that we introduced quasiconvexity as a "natural" hypothesis to improve Tonelli's theorem. The following Theorem 1.20 confirms this fact.

Definition 1.17 (w^*-convergence in $W^{1,\infty}$)**.** *Let us consider an open set $\Omega \subset \mathbb{R}^n$ and $f_k \in W^{1,\infty}(\Omega; \mathbb{R}^m)$. We write $f_k \to f$ in w^*-$W^{1,\infty}(\Omega; \mathbb{R}^m)$ if*

- *$f_k \to f$ uniformly in Ω;*

- *$\|\nabla f_k\|_{L^\infty(\Omega;\mathbb{R}^m)}$ is uniformly bounded.*

Proposition 1.18. *If $f_k \to f$ in w^*-$W^{1,\infty}(\Omega; \mathbb{R}^m)$, then $f \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ and $\nabla f_k \overset{*}{\rightharpoonup} \nabla f$.*

This is a direct consequence of the fact that (∇f_k) is sequentially compact in the w^*-topology of L^∞, and any weak* limit provides a weak derivative of f (hence $f \in W^{1,\infty}$, the limit is unique and the whole sequence of derivatives w^*-converges).

Before stating Morrey's lower semicontinuity theorem we give a quick proof, based on a blow-up argument, of Rademacher's differentiability theorem.

Theorem 1.19 (Rademacher)**.** *Given $\Omega \subset \mathbb{R}^n$ open, let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. If $x_0 \in \Omega$ is a Lebesgue point of the weak gradient ∇f, namely $\fint_{B_r(x_0)} |\nabla f - L|\, dx \to 0$ as $r \downarrow 0$ for some linear map $L : \mathbb{R}^n \to \mathbb{R}^m$, then f is differentiable at x_0 and the (classical) gradient at x_0 is equal to L. In particular, any locally Lipschitz function is differentiable $\mathscr{L}^n$-a.e. and its differential coincides $\mathscr{L}^n$-a.e. with the weak gradient.*

Proof. The property which we need to prove can be equivalently stated as follows:

$$f_r(y) \to L(y) \text{ uniformly on } \overline{B}_1 \text{ as } r \downarrow 0,$$

where $f_r(y) = (f(x_0 + ry) - f(x_0))/r$ are the rescaled maps. Notice that the functions f_r are uniformly Lipschitz continuous in $\overline{B}_1$ and uniformly bounded (because $f_r(0) = 0$), hence the family (f_r) is relatively compact in $C^0(\overline{B}_1; \mathbb{R}^m)$ as $r \downarrow 0$. Hence, it suffices to show that any limit point $f_0(y) = \lim_i f_{r_i}(y)$ coincides with $L(y)$. A simple change of variables shows that $\nabla f_r(y) = \nabla f(x_0 + ry)$ in B_1 (where gradients are obviously understood as weak gradients), thus

$$\fint_{B_1} |\nabla f_r - L| \, dy = \fint_{B_r(x_0)} |\nabla f - L| \, dx.$$

It follows that $\nabla f_r \to L$ in $L^1(B_1; \mathbb{R}^{m\times n})$, hence we are in position to invoke Proposition 1.18 and conclude that $f_0 \in W^{1,\infty}(B_1; \mathbb{R}^m)$ and by comparison $\nabla f_0 = L$ in L^∞, hence $\mathscr{L}^n$-a.e. in B_1. By the constancy Theorem A.5 we get $f_0(y) = L(y) + c$ for some constant c, which obviously should be 0 because $f_0(0) = \lim_i f_{r_i}(0) = 0$. □

Theorem 1.20 (Morrey, [75])**.** *Assume that $L : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to [0, \infty)$ is continuous and $u \in W^{1,\infty}(\Omega; \mathbb{R}^m)$. If $L(x, u(x), \cdot)$ is quasiconvex for $\mathscr{L}^n$-a.e. $x \in \Omega$, then the functional $\mathcal{L}$ defined by* (1.38) *is lower semicontinuous at u with respect to the w^*-$W^{1,\infty}(\Omega; \mathbb{R}^m)$ convergence. Conversely, if $\mathcal{L}$ lower semicontinuous at u with respect to the w^*-$W^{1,\infty}(\Omega; \mathbb{R}^m)$ convergence, one has that $L(x, u(x), \cdot)$ is quasiconvex at $\nabla u(x)$ for $\mathscr{L}^n$-a.e. $x \in \Omega$.*

Proof.

Part I: Necessity of quasiconvexity. It is sufficient to prove the result for any Lebesgue point $x_0 \in \Omega$ of ∇u. The main tool is a blow-up argument: if $Q = (-1/2, 1/2)^n$ is the unit cube centered at 0, $Q_r(x_0) = x_0 + rQ \subset \Omega$ and $v \in W^{1,\infty}_0(Q; \mathbb{R}^m)$, we set

$$\mathcal{L}_r(v) := \int_Q L(x_0 + ry, u(x_0 + ry) + rv(y), \nabla u(x_0 + ry) + \nabla v(y)) \, dy.$$

The formal limit as $r \downarrow 0$ of $\mathcal{L}_r$, namely

$$\mathcal{L}_0(v) := \int_Q L(x_0, u(x_0), \nabla u(x_0) + \nabla v(y)) \, dy$$

is lower semicontinuous at $v = 0$ with respect to the w^*-$W^{1,\infty}(Q; \mathbb{R}^m)$ convergence because of the following two facts:

- each $\mathcal{L}_r$ is lower semicontinuous at 0 with respect to the w^*-$W^{1,\infty}(Q;\mathbb{R}^m)$ convergence, indeed

$$\begin{aligned}\mathcal{L}_r(v) &= \frac{1}{r^n}\int_{Q_r(x_0)} L\left(x, u(x) + rv\left(\frac{x-x_0}{r}\right), \nabla u(x) + \nabla v\left(\frac{x-x_0}{r}\right)\right) dx \\ &= \frac{1}{r^n}\left(\mathcal{L}\left(u + rv\left(\frac{x-x_0}{r}\right)\right) - \int_{\Omega\setminus Q_r(x_0)} L(x, u(x), \nabla u(x))\, dx\right);\end{aligned}$$

- being x_0 a Lebesgue point for ∇u, for any bounded sequence $(v_h) \subset W_0^{1,\infty}(Q;\mathbb{R}^m)$ it is easily checked that the continuity of L gives

$$\lim_{r\to 0^+} \sup_h |\mathcal{L}_r(v_h) - \mathcal{L}_0(v_h)| = 0.$$

Let us introduce the auxiliary function

$$H(p) := L(x_0, u(x_0), \nabla u(x_0) + p).$$

Given a test function $\varphi \in C_c^\infty(Q;\mathbb{R}^m)$, we work with the Q-periodic function ψ such that $\psi|_Q = \varphi$ and the sequence of highly oscillating (being $\frac{1}{h}$-periodic) functions

$$v_h(x) := \frac{1}{h}\psi(hx)\,,$$

which obviously converge uniformly to 0 as we let $h \to \infty$. Since $\nabla v_h(x) = \nabla\psi(hx)$ we also have $v_h \overset{*}{\rightharpoonup} 0$ in $W^{1,\infty}(Q;\mathbb{R}^m)$, so that, thanks to the lower semicontinuity of $\mathcal{L}_0$ at 0, one has

$$\begin{aligned}H(0) = \mathcal{L}_0(0) \;&\le\; \liminf_{h\to\infty} \int_Q H(\nabla v_h(x))\, dx = \liminf_{h\to\infty} \frac{1}{h^n}\int_{Q_h} H(\nabla\psi(y))\, dy \\ &= \int_Q H(\nabla\psi(y))\, dy = \int_Q H(\nabla\varphi(y))\, dy,\end{aligned}$$

where $Q_h = (-h/2, h/2)^n$. This is precisely the quasiconvexity property for $L(x_0, u(x_0), \cdot)$ at $\nabla u(x_0)$.

Part II: Sufficiency of quasiconvexity. We split the proof in several steps, reducing ourselves to increasingly simpler cases. First, since any open set Ω can be monotonically approximated by bounded open sets with closure contained in Ω, we can assume that Ω is bounded and that $L \in C^0(\overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n})$. Since Ω can be written as the disjoint union of half-open cubes, by the superadditivity of the lim inf we can also assume that $\Omega = Q$ is a n-cube with side length ℓ. We further set

$$M := \sup\left\{|(x, \nabla u_h(x))| :\; x \in \overline{\Omega},\; h \in \mathbb{N}\right\},$$

where (u_h) is a sequence that is converging to u in w^*-$W^{1,\infty}(\Omega;\mathbb{R}^m)$, as encoded in Definition 1.17. Now, considering the decomposition

$$L(x,u_h(x),\nabla u_h(x)) = \big[L(x,u_h(x),\nabla u_h(x)) - L(x,u(x),\nabla u_h(x))\big] + L(x,u(x),\nabla u_h(x))$$

we see immediately that it is enough to consider Lagrangians $L_1(x,p)$ which are independent of s (just take $L_1(x,p) = L(x,u(x),p)$, quasiconvex with respect to p for $\mathscr{L}^n$-a.e. $x \in \Omega$). Here we have exploited the continuity assumption on the Lagrangian L together with Definition 1.17 to ensure that the argument of the third slot of L varies in a compact set.

The next step is to reduce ourselves to Lagrangians independent of x. To this aim, consider a modulus of continuity ω for $L_1(\cdot,p)$ that is uniform as p varies in the ball $\overline{B_M}$, and a decomposition of Q in 2^{kn} cubes Q_i with side length $\ell/2^k$, with points $x_i \in Q_i$ such that $L_1(x_i,\cdot)$ is quasiconvex. Then, adding and subtracting $L_1(x_i,\nabla u_h(x))$ and using once more the superadditivity of $\liminf$, we get

$$\liminf_{h\to\infty}\int_Q L_1(x,\nabla u_h(x))\,dx \geq \sum_i \liminf_{h\to\infty}\int_{Q_i} L_1(x_i,\nabla u_h(x))\,dx - \omega\left(\frac{\sqrt{n}\ell}{2^k}\right)\sum_i \mathscr{L}^n(Q_i).$$

Since $\sum_i \mathscr{L}^n(Q_i) = \ell^n$, if we are able to show that

$$\liminf_{h\to\infty}\int_{Q_i} L_1(x_i,\nabla u_h(x))\,dx \geq \int_{Q_i} L_1(x_i,\nabla u(x))\,dx \quad \forall\, i,$$

we obtain

$$\begin{aligned}\liminf_{h\to\infty}\int_Q L_1(x,\nabla u_h(x))\,dx &\geq \sum_i \int_{Q_i} L_1(x_i,\nabla u(x))\,dx - \omega\left(\frac{\sqrt{n}\ell}{2^k}\right)\ell^n\\ &\geq \int_Q L_1(x,\nabla u(x))\,dx - 2\omega\left(\frac{\sqrt{n}\ell}{2^k}\right)\ell^n.\end{aligned}$$

As $k\to\infty$ we recover the liminf inequality.

Hence, we are led to show the sufficiency of lower semicontinuity for quasiconvex Lagrangians of the form $L_2(p) = L_1(x_i,p)$. In the following argument we shall use the fact that continuous quasiconvex functions are locally Lipschitz. This property can be obtained noticing that bounded convex functions w are Lipschitz, with the quantitative estimate

$$\mathrm{Lip}(w,B_r(x)) \leq \frac{\sup_{B_{2r}(x)} w - \inf_{B_{2r}(x)} w}{r}$$

(see Subsection D.1) and quasiconvex functions g satisfy the Legendre-Hadamard condition, hence $g(p)$ is separately convex i.e. convex as a function of each variable p_i^α.

We consider two cases: first, the case when the limit function u is affine and then, by a blow-up argument again, the general case. Assume now that u is affine, let $A = \nabla u$ and consider a smooth function $\psi : \Omega \to [0,1]$ with compact support. Thanks to (1.41), we can apply the quasiconvexity inequality to the test function given by $\varphi = (u_h - u)\psi$ with $R(p) = L_2(p) - L_2(0)$ to get

$$\begin{aligned}
R(A) &\leq \fint_\Omega R(A + \psi\nabla(u_h - u) + (u_h - u) \otimes \nabla\psi)\, dx \\
&\leq \fint_\Omega R((1-\psi)A + \psi\nabla u_h + (u_h - u) \otimes \nabla\psi)\, dx \\
&\leq c\left(|A| \fint_\Omega (1-\psi)\, dx + \|\nabla\psi\|_{L^\infty(\Omega;\mathbb{R})} \fint_\Omega |u_h - u|\, dx\right) + \fint_\Omega R(\psi\nabla u_h) \\
&\leq c\left(|A| \fint_\Omega (1-\psi)\, dx + \sup_h \|\nabla u_h\|_{L^\infty(\Omega;\mathbb{R}^{m\times n})} \fint_\Omega (1-\psi)\, dx\right) \\
&\qquad\qquad + c\|\nabla\psi\|_{L^\infty(\Omega;\mathbb{R})} \fint_\Omega |u_h - u|\, dx + \fint_\Omega R(\nabla u_h),
\end{aligned}$$

where the last two inequalities rely on the aforementioned local Lipschitz property (for $R(\cdot)$), and we have denoted by $c > 0$ the corresponding Lipschitz constant. At this stage, we pass to the limit first as $h \to \infty$ and then as $\psi \uparrow 1$ to obtain the result.

Finally, we consider the general case, using Theorem 1.19 and a blow-up argument. Assume that the $\liminf \int_\Omega L_2(\nabla u_h)\, dx$ is in fact a limit, which we shall denote by m, and consider the family of measures $\mu_h := L_2(\nabla u_h)\mathscr{L}^n$. Being this family bounded, we can assume with no loss of generality that μ_h weakly converge, in the duality with $C_c(\Omega;\mathbb{R})$, to some positive Radon measure μ. Recall that the evaluations on compact sets K and open sets A are respectively upper and lower semicontinuous with respect to weak convergence, i.e.

$$\mu(K) \geq \limsup_{h\to\infty} \mu_h(K), \qquad \mu(A) \leq \liminf_{h\to\infty} \mu_h(A). \tag{1.44}$$

In particular $\mu(\Omega) \leq m$, so that, if we show that $\mu \geq L_2(\nabla u)\mathscr{L}^n$, we are done. By Lebesgue's differentiation theorem for measures (see Proposition C.7, which suffices for our scopes, and [72] or [30] among others for a broader contextualization), it suffices to show that

$$\liminf_{r\downarrow 0} \frac{\mu(\overline{B}_r(x_0))}{\omega_n r^n} \geq L_2(\nabla u(x_0)) \qquad \text{for } \mathscr{L}^n\text{-a.e. } x_0 \in \Omega. \tag{1.45}$$

We shall prove this property at any differentiability point x_0 of u such that L_2 is quasiconvex at $\nabla u(x_0)$. To this aim, let $r_i \to 0$ be a sequence for which the lim inf is achieved, and $\varepsilon > 0$. For any i we can choose $h_i \geq i$ so large that

$$\int_{B_{r_i}(x_0)} L_2(\nabla u_{h_i})\, dx \leq \mu(\overline{B}_{r_i}(x_0)) + \frac{r_i^n}{i}\,, \qquad \fint_{B_{r_i}(x_0)} |u_{h_i} - u|\, dx \leq \frac{r_i}{i}\,. \tag{1.46}$$

Now consider the following rescaled functions

$$v_i(y) := \frac{u_{h_i}(x_0 + r_i y) - u(x_0)}{r_i}, \qquad w_i(y) := \frac{u(x_0 + r_i y) - u(x_0)}{r_i}$$

that patently satisfy

$$\int_{B_1} L_2(\nabla v_i)\, dy \le \frac{\mu(\overline{B}_{r_i}(x_0))}{r_i^n} + \frac{1}{i}, \qquad \fint_{B_1} |v_i - w_i|\, dy \to 0\,.$$

Since $w_i(y) \to \nabla u(x_0) y$ uniformly in $\overline{B}_1$, thanks to the differentiability assumption, we obtain that v_i converge to the linear function $y \mapsto \nabla u(x_0) y$ in $L^1(B_1; \mathbb{R}^m)$. Therefore, by the previous step, we have

$$\liminf_{i\to\infty} \frac{\mu(\overline{B}_{r_i}(x_0))}{r_i^n} \ge \liminf_{i\to\infty} \int_{B_1} L_2(\nabla v_i)\, dy - \frac{1}{i} \ge \omega_n L_2(\nabla u(x_0))\,.$$

□

The previous result shows that quasiconvexity of the Lagrangian is equivalent to sequential lower semicontinuity of the integral functional in the weak*-$W^{1,\infty}$ convergence. However, in many problems of Calculus of Variations only L^α bounds, with $\alpha < \infty$, are available on the gradient. A remarkable improvement of Morrey's result is the following:

Theorem 1.21 (Acerbi-Fusco, [1])**.** *Suppose that a Borel Lagrangian $L(x, s, p)$ is continuous in (s, p) and satisfies*

$$0 \le L(x, s, p) \le c(1 + |s|^\alpha + |p|^\alpha) \qquad \forall (x, s, p) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$$

for some $\alpha > 1$ and some constant c. Suppose also that the map $p \mapsto L(x, s, p)$ is quasiconvex for all (x, s). Then F is sequentially lower semicontinuous with respect to the weak $W^{1,\alpha}(\Omega; \mathbb{R}^m)$-topology.

2 Classical regularity theory for linear problems

In this chapter, we begin studying the local behavior of (weak) solutions $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ of a system of equations given by

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\big) = f_i - \sum_{\alpha}\partial_{x_\alpha}F^\alpha_i \qquad i=1,\dots,m \tag{2.1}$$

with $A^{\alpha\beta}_{ij} \in L^\infty(\Omega;\mathbb{R})$, $f_i \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ and $F^\alpha_i \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$.

We will see how the Caccioppoli-Leray inequality can be employed, following an idea due to L. Nirenberg, to prove existence of higher-order weak derivatives of u and suitable integrability results thereof, and how to translate these estimates into actual regularity results by means of the Sobolev embedding theorems. In the last section of the chapter, we will also briefly discuss the boundary regularity of weak solutions.

2.1 Caccioppoli-Leray inequality

We start by stating the basic energy estimate for weak solutions of problems having the form (2.1). Recall from Chapter 1 that we use the symbol $|\cdot|$ to denote the pointwise Hilbert-Schmidt norm of matrices and tensors, even though some estimates would still be valid with the (smaller) operator norm. For instance, we shall set

$$|A^{\alpha\beta}_{ij}|^2 := \sum_{\alpha,\beta,i,j}(A^{\alpha\beta}_{ij})^2.$$

Theorem 2.1 (Caccioppoli-Leray inequality, [11] and [70])**.** *If the Borel coefficients $A^{\alpha\beta}_{ij}$ satisfy the Legendre condition with $\lambda > 0$ and*

$$\sup_{x\in\Omega}|A^{\alpha\beta}_{ij}(x)| \le \Lambda < +\infty,$$

then there exists a positive constant $c_{CL} = c_{CL}(\lambda,\Lambda)$ such that, for any ball $B_R(x_0) \Subset \Omega$ and any $k \in \mathbb{R}^m$, it holds

$$c_{CL}\int_{B_{R/2}(x_0)}|\nabla u|^2\,dx \le R^{-2}\int_{B_R(x_0)}|u(x)-k|^2\,dx + R^2\int_{B_R(x_0)}|f(x)|^2\,dx + \int_{B_R(x_0)}|F(x)|^2\,dx. \tag{2.2}$$

Before proceeding to the proof, let us present two important remarks.

Remark 2.2. (1) The validity of (2.2) for all $k \in \mathbb{R}^m$ depends on the translation invariance of the equation, meaning that $u+k$ is a solution if and only if u is. Moreover, the inequality has a natural scaling invariance, as well as the equation: if we think of u as an adimensional quantity, then both sides of (2.2) have dimension length^{n-2}, because $f \sim$ length^{-2} and $F \sim$ length^{-1}.

(2) The Caccioppoli-Leray inequality could already be regarded as a basic regularity result since for a general function u the gradient ∇u cannot be controlled by the variance of u! In the sequel of this chapter, we shall employ it several times to derive other, more conventional, regularity results i.e. results concerning existence and quantitative bounds for higher derivatives of weak solutions of elliptic equations and systems.

Remark 2.3 (Absorption scheme)**.** In elliptic regularity theory it often happens that one can prove, for some $\alpha < 1$, an estimate of the form

$$A \leq BA^\alpha + C.$$

The absorption scheme allows to bound A in terms of B, C and α only and works as follows: by the Young inequality

$$ab = \varepsilon a \frac{b}{\varepsilon} \leq \frac{\varepsilon^p a^p}{p} + \frac{b^q}{\varepsilon^q q} \qquad \text{(with } 1/p + 1/q = 1\text{)}$$

so that for $p = 1/\alpha$ (hence $q = 1/(1-\alpha)$) one obtains

$$A \leq BA^\alpha + C \leq \frac{\varepsilon^p A}{p} + \frac{B^q}{\varepsilon^q q} + C.$$

Now, if we choose $\varepsilon = \varepsilon(\alpha)$ sufficiently small, so that $2\varepsilon^p \leq p$, we get

$$A \leq 2\frac{B^q}{\varepsilon^q q} + 2C.$$

Let us prove Theorem 2.1.

Proof. Without loss of generality, we can consider $x_0 = 0$ and $k = 0$. As customary in regularity theory, we choose test functions depending on the solution u itself, namely we set

$$\varphi := u\eta^2$$

where $\eta \in C_c^\infty(B_R; \mathbb{R})$ is a cutoff function with $\eta \equiv 1$ in $B_{R/2}$, $0 \leq \eta \leq 1$ and $|\nabla \eta| \leq 4/R$. Since u solves (2.1), we have that

$$\int_{B_R} \langle A\nabla u, \nabla\varphi\rangle \, dx - \int_{B_R} \langle f, \varphi\rangle \, dx - \int_{B_R} \langle F, \nabla\varphi\rangle \, dx = 0. \tag{2.3}$$

Moreover

$$\nabla\varphi = \eta^2 \nabla u + 2\eta u \otimes \nabla\eta, \tag{2.4}$$

so plugging (2.4) in (2.3) we obtain

$$\begin{aligned} 0 = & \int_{B_R} \eta^2 \langle A\nabla u, \nabla u\rangle \, dx + 2\int_{B_R} \eta \langle A\nabla u, u\otimes \nabla\eta\rangle \, dx \\ & - \int_{B_R} \langle f, \varphi\rangle \, dx - \int_{B_R} \eta^2 \langle F, \nabla u\rangle \, dx - 2\int_{B_R} \eta\langle F, u\otimes\nabla\eta\rangle \, dx. \end{aligned} \tag{2.5}$$

Let us deal with each summand separately.

- By the Legendre condition, we can estimate

$$\begin{aligned} \int_{B_R} \eta^2 \langle A\nabla u, \nabla u\rangle \, dx &= \int_{B_R} \sum_{\alpha,\beta,i,j} \eta^2 A_{ij}^{\alpha\beta} \partial_{x_\alpha} u^i \partial_{x_\beta} u^j \, dx \\ &\geq \lambda \int_{B_R} \eta^2 |\nabla u|^2 \, dx. \end{aligned}$$

- We have

$$\begin{aligned} 2\int_{B_R} \eta A\nabla u \left(u\otimes\nabla\eta\right) \, dx &= 2\int_{B_R} \eta \sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta} u^j \partial_{x_\alpha} u^i \partial_{x_\beta}\eta \, dx \\ &\leq 2\int_{B_R} \eta |A||u||\nabla u||\nabla\eta| \, dx \\ &\leq \frac{8\Lambda}{R}\int_{B_R} \left(\eta|\nabla u|\right)|u| \, dx \\ &\leq \frac{4\Lambda\varepsilon}{R}\int_{B_R} \eta^2|\nabla u|^2 \, dx + \frac{4\Lambda}{R\varepsilon}\int_{B_R} |u|^2 \, dx, \end{aligned}$$

  where the first estimate is due to Cauchy-Schwarz inequality, the second one relies on the boundedness of coefficients $A_{ij}^{\alpha\beta}$ and the estimate on $|\nabla\eta|$, and the third one is based on the Young inequality (applied for $p = q = 2$).

- By the Young inequality and since $\eta^2 \leq 1$,

$$\begin{aligned} \int_{B_R} \langle f, \varphi\rangle \, dx &= \int_{B_R} \eta^2 \sum_i f_i u^i \, dx \\ &\leq \frac{1}{2R^2}\int_{B_R} |u|^2 \, dx + \frac{R^2}{2}\int_{B_R} |f|^2 \, dx. \end{aligned}$$

- Similarly, one has

$$\begin{aligned} \int_{B_R} \eta^2 \langle F, \nabla u\rangle \, dx &= \int_{B_R} \sum_{\alpha,i} \eta^2 F_i^\alpha \partial_{x_\alpha} u^i \, dx \\ &\leq \frac{\lambda}{4}\int_{B_R} \eta^2|\nabla u|^2 \, dx + \frac{1}{\lambda}\int_{B_R} |F|^2 \, dx. \end{aligned}$$

- Again by the same arguments (Schwarz inequality, estimate on $|\nabla\eta|$ and Young inequality)

$$\begin{aligned} 2\int_{B_R} \eta\langle F, u\otimes\nabla\eta\rangle\,dx &= 2\int_{B_R}\sum_{\alpha,i}\eta F_i^\alpha u^i\partial_{x_\alpha}\eta\,dx \\ &\le 4\int_{B_R}|F|^2\,dx + \frac{4}{R^2}\int_{B_R}|u|^2\,dx. \end{aligned}$$

Therefore, from (2.5) it follows that

$$\begin{aligned} \lambda\int_{B_R}\eta^2|\nabla u|^2\,dx &\le \int_{B_R}\eta^2\langle A\nabla u,\nabla u\rangle\,dx \\ &\le \left(\frac{4\Lambda\varepsilon}{R}+\frac{\lambda}{4}\right)\int_{B_R}\eta^2|\nabla u|^2\,dx + \left(\frac{4\Lambda}{R\varepsilon}+\frac{1}{2R^2}+\frac{4}{R^2}\right)\int_{B_R}|u|^2\,dx \\ &\qquad + \frac{R^2}{2}\int_{B_R}|f|^2\,dx + \left(\frac{1}{\lambda}+4\right)\int_{B_R}|F|^2\,dx. \end{aligned}$$

By choosing ε sufficiently small, in such a way that $4\Lambda\varepsilon/R = \lambda/4$, one can absorb the Dirichlet-energy term on the right-hand side, and the thesis follows after just noticing that

$$\int_{B_R}\eta^2|\nabla u|^2\,dx \ge \int_{B_{R/2}}|\nabla u|^2\,dx.$$

□

Remark 2.4 (Widman's hole-filling technique, [96])**.** There exists a sharper version of the Caccioppoli-Leray inequality, let us illustrate it in the simpler case $f = 0$, $F = 0$. Indeed, since

$$|\nabla\eta| \le \frac{4}{R}\chi_{B_R\setminus B_{R/2}},$$

following the proof of Theorem 2.1 one obtains

$$\int_{B_{R/2}}|\nabla u(x)|^2\,dx \le \frac{c}{R^2}\int_{B_R\setminus B_{R/2}}|u(x)-k|^2\,dx, \tag{2.6}$$

which holds true for some positive constant c that does not depend on R. If we choose $k := \fint_{B_R\setminus B_{R/2}} u\,dx$, the Poincaré inequality (A.28) in the regular domain $B_1\setminus\overline{B}_{1/2}$ and a scaling argument give

$$\int_{B_{R/2}}|\nabla u(x)|^2\,dx \le c\int_{B_R\setminus B_{R/2}}|\nabla u(x)|^2\,dx. \tag{2.7}$$

Adding to (2.7) the term $c\int_{B_{R/2}} |\nabla u(x)|^2\,dx$, we get

$$(c+1)\int_{B_{R/2}} |\nabla u(x)|^2\,dx \le c\int_{B_R} |\nabla u(x)|^2\,dx.$$

Setting $\theta := c/(c+1) < 1$, we obtained the decay inequality

$$\int_{B_{R/2}} |\nabla u(x)|^2\,dx \le \theta\int_{B_R} |\nabla u(x)|^2\,dx.$$

Iterating (2.6) d times, with d integer satisfying $2^d r < R \le 2^{d+1} r$, it is not difficult to get

$$\int_{B_r} |\nabla u(x)|^2\,dx \le 2^\alpha \left(\frac{r}{R}\right)^\alpha \int_{B_R} |\nabla u(x)|^2\,dx \qquad 0 < r \le R \tag{2.8}$$

with $(1/2)^\alpha = \theta$, i.e. $\alpha = \log_2(1/\theta)$. When $n = 2$, this implies that $u \in C^{0,\alpha/2}(\Omega;\mathbb{R}^m)$, as we will see.

The following is another example of "unnatural" inequality, which provides additional information on functions that satisfy it.

Definition 2.5 (Reverse Hölder's inequality)**.** *Let $\alpha \in (1,\infty)$. A non-negative function $f \in L^\alpha_{\mathrm{loc}}(\Omega;\mathbb{R})$ satisfies a* reverse Hölder's inequality *with exponent α if there exists a constant $c > 0$ such that*

$$\fint_{B_{R/2}(x)} f^\alpha\,dx \le c\left(\fint_{B_R(x)} f\,dx\right)^\alpha \qquad \forall B_R(x) \Subset \Omega.$$

We will see, in the sequel of this section, that the gradient of a solution of an elliptic equation having the form described above satisfies a reverse Hölder's inequality. Combining this fact with the Gehring's Lemma (see [40]), it is possible to obtain various regularity results. In particular, one can follow this path to obtain a full solution of Hilbert's *XIX* problem when $n = 2$, namely for functions of two variables.

2.2 Sobolev embeddings

For the sake of completeness, we now recall the Sobolev inequalities and the associated embedding theorems. Detailed proofs will be provided later on: concerning the cases $p = n$ and $p > n$, we will see them in the more general context of Morrey's theory (see Subsection 3.1). We will treat the case $p < n$ while dealing with De Giorgi's solution of Hilbert's XIX problem since slightly more general versions of the Sobolev inequality are needed there (see Subsection A.3 for the proof of this result and Subsection 3.6 for the applications of these tools to Hilbert's regularity question). These topics are treated in most classical textbooks, and we refer the reader to [2, 8, 36] among others for a broader discussion.

Theorem 2.6 (Sobolev inequalities)**.** *Let Ω be either the whole space $\mathbb{R}^n$ or an open, bounded and regular domain.*

- *If $1 \le p < n$, denoting with $p^* := \frac{np}{n-p} > p$ the Sobolev conjugate exponent (characterized by the relation $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$), we have the continuous embedding*

$$W^{1,p}(\Omega;\mathbb{R}) \hookrightarrow L^{p^*}(\Omega;\mathbb{R}).$$

- *If $p = n$, we have the inclusion of $W^{1,n}(\Omega;\mathbb{R})$ in the space $BMO(\Omega;\mathbb{R})$ of functions of bounded mean oscillation; this provides exponential integrability in bounded subsets of Ω.*[3]

- *If $p > n$, we have the continuous embedding*

$$W^{1,p}(\Omega;\mathbb{R}) \hookrightarrow C^{0,1-n/p}(\overline{\Omega};\mathbb{R}).$$

Remark 2.7. For purely notational convenience, we agree to define $p^* = \infty$ whenever $p \ge n$. This extension will be tacitly assumed in the sequel of this monograph, see for instance the proof of Theorem 3.35.

Example 2.8. Consider the ball $B_{1/2}(0) \subset \mathbb{R}^n$, with $n \ge 2$, and the radial function $u(x) := \log(|\log|x||)$. Such a function is smooth in $B_{1/2}(0) \setminus \{0\}$, its derivative being

$$\nabla u(x) = \frac{x}{|x|^2 \log|x|}.$$

For any $0 < \varepsilon < 1/2$ and for any test function $\varphi \in C^1_c\left(B_{1/2}(0);\mathbb{R}\right)$, we compute

$$\int_{B_{1/2}(0)\setminus B_\varepsilon(0)} u\partial_{x_\alpha}\varphi\,dx = -\int_{B_{1/2}(0)\setminus B_\varepsilon(0)} \varphi\partial_{x_\alpha}u\,dx - \int_{\partial B_\varepsilon(0)} u\varphi\nu_\alpha\,d\mathscr{H}^{n-1},$$

where $\nu_\alpha = \frac{x_\alpha}{|x|}$ is the α-th component of the normal ν on $\partial B_\varepsilon(0)$. For the latter summand we can estimate

$$\begin{aligned}\left|\int_{\partial B_\varepsilon(0)} u\varphi\nu_\alpha\,d\mathscr{H}^{n-1}\right| &\le \log|\log(\varepsilon)|\left|\int_{\partial B_\varepsilon(0)} \varphi\nu_\alpha\,d\mathscr{H}^{n-1}\right| \\ &\le n\omega_n\varepsilon^{n-1}\log|\log(\varepsilon)|\|\varphi\|_{L^\infty\left(B_{1/2}(0);\mathbb{R}\right)} \xrightarrow{\varepsilon\to 0} 0.\end{aligned}$$

Therefore the function u belongs to $H^1\left(B_{1/2}(0);\mathbb{R}\right)$, even though it is not continuous in the whole ball $B_{1/2}(0)$.

[3]The result is basically sharp, as the example of $(-\log|x|)^\alpha \in W^{1,n}(B_{1/2}(0);\mathbb{R})$ for $n > 1$ and $\alpha \in (0, 1-1/n)$ shows.

Remark 2.9. Combining the Poincaré inequality with the inequality

$$\left(\fint_{B_1} |v - v_{0,1}|^{p^*}\,dx\right)^{1/p^*} \le c_I\left[\left(\fint_{B_1} |v - v_{0,1}|^{p}\,dx\right)^{1/p} + \left(\fint_{B_1} |\nabla v|^{p}\,dx\right)^{1/p}\right],$$

coming from the continuity of the embedding $W^{1,p} \hookrightarrow L^{p^*}$, we get

$$\left(\fint_{B_1} |v - v_{0,1}|^{p^*}\,dx\right)^{1/p^*} \le c\left(\fint_{B_1} |\nabla v|^{p}\,dx\right)^{1/p},$$

for some constant c depending on c_I and $c_{P,II}$ ($c_{P,II}$ being the constant in the Poincaré inequality (A.28)). Here we have adopted the following notation: $v_{x,s}$ stands for the mean value of the function v over the ball $B_s(x)$.

By a standard scaling argument this gives

$$\left(\fint_{B_R} |u - u_{0,R}|^{p^*}\,dx\right)^{1/p^*} \le cR\left(\fint_{B_R} |\nabla u|^{p}\,dx\right)^{1/p}. \tag{2.9}$$

If u solves an equation of the form (2.1) with $f = F = 0$, combining (2.9) with the Caccioppoli-Leray inequality when $p^* = 2$ (that is to say $p = 2n/(n+2) < 2$), we write

$$c_{CL}^{1/2} R\left(\fint_{B_{R/2}} |\nabla u|^{2}\,dx\right)^{1/2} \le \left(\fint_{B_R} |u - u_{0,R}|^{2}\,dx\right)^{1/2} \le cR\left(\fint_{B_R} |\nabla u|^{p}\,dx\right)^{1/p}.$$

This way we proved that $|\nabla u|^p$ satisfies a reverse Hölder's inequality with exponent given by $\alpha = 2/p > 1$ and multiplicative constant $c/c_{CL}^{1/2}$, namely

$$\left(\fint_{B_{R/2}} |\nabla u|^{2}\,dx\right)^{1/2} \le \frac{c}{c_{CL}^{1/2}}\left(\fint_{B_R} |\nabla u|^{p}\,dx\right)^{1/p}.$$

Remark 2.10 (Embedding for higher order Sobolev spaces)**.** Recall first that higher order Sobolev spaces $W^{k,p}(\Omega;\mathbb{R})$, with $k \ge 1$ integer and $1 \le p \le \infty$, are recursively defined as

$$W^{k,p}(\Omega;\mathbb{R}) := \left\{u \in W^{1,p}(\Omega;\mathbb{R}) :\ \nabla u \in W^{k-1,p}(\Omega;\mathbb{R}^n)\right\}.$$

Together with the Sobolev embedding in Theorem 2.6, applied for $p > n$, another way to gain continuity when $p < n$ is to use the Sobolev spaces $W^{k,p}$, with k sufficiently large. In fact, we obtain a continuous embedding of $W^{k,p}(\Omega;\mathbb{R})$ into a space of Hölder continuous functions as long as $k > \lfloor \frac{n}{p} \rfloor$ (where $\lfloor\cdot\rfloor$ denotes the integer part). More precisely, if we let Ω denote (as above) an open, bounded and regular domain in $\mathbb{R}^n$ the following statements hold:

(i) If $kp < n$, we get $W^{k,p}(\Omega;\mathbb{R}) \hookrightarrow L^q(\Omega;\mathbb{R})$ for all $1 \le q \le p_k^*$ where p_k^* is obtained from p iterating the $*$ operation k-times, namely

$$\frac{1}{p_k^*} = \frac{1}{p} - \frac{k}{n}.$$

(ii) If $kp = n$, we get $W^{k,p}(\Omega;\mathbb{R}) \hookrightarrow L^q(\Omega;\mathbb{R})$ for all $1 \le q < \infty$;

(iii) If $kp > n$ and $k - \frac{n}{p} \notin \mathbb{N}$, we get $W^{k,p}(\Omega;\mathbb{R}) \hookrightarrow C^{l,\alpha}(\overline{\Omega};\mathbb{R})$ for $\ell = \lfloor k - \frac{n}{p} \rfloor$ and all $0 \le \alpha \le k - l - n/p$;

(iv) If $kp > n$ and $k - \frac{n}{p} = \ell + 1 \in \mathbb{N}$, we get $W^{k,p}(\Omega;\mathbb{R}) \hookrightarrow C^{l,\alpha}(\overline{\Omega};\mathbb{R})$ for all $0 \le \alpha < 1$.

2.3 A priori estimates and the Nirenberg method

Let us now consider the case that $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ is a weak solution of an elliptic equation of the form (2.1): if we do not assume existence of derivatives of higher order, or higher integrability of the weak gradient, the arguments presented in the previous remark are not immediately applicable to provide classical regularity results (i.e. differentiability) of u, even when the data on the right-hand side are actually smooth. The scope of the following discussion is precisely to present a general approach to gain better integrability, hence interior regularity, for weak solutions of linear elliptic problems.

In order to illustrate Nirenberg's method in the simplest possible setting, let us initially consider a (local) solution $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ of the Poisson equation

$$-\Delta u = f \qquad f \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R}).$$

Consistently with what has just been stated, our scope is to prove that u belongs to $H^2_{\mathrm{loc}}(\Omega;\mathbb{R})$. This will be the key step in transferring regularity from the data to the solution.

When we talk about an *a priori* estimate, we mean the following argument. Suppose that we already knew that $\partial_{x_\alpha} u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$: then it would not be difficult to check (using the fact that higher order weak derivatives commute, as this is the case for classical ones) that this function solves

$$-\Delta\left(\partial_{x_\alpha} u\right) = \partial_{x_\alpha} f$$

in a weak sense. Hence, for any ball $B_R(x_0) \Subset \Omega$, by the Caccioppoli-Leray inequality we get

$$c_{CL} \int_{B_{R/2}(x_0)} \left|\nabla\left(\partial_{x_\alpha} u\right)\right|^2 \, dx \le \frac{1}{R^2} \int_{B_R(x_0)} \left|\partial_{x_\alpha} u\right|^2 \, dx + \int_{B_R(x_0)} |f|^2 \, dx. \tag{2.10}$$

We have chosen the Poisson equation because constant coefficients differential operators commute with convolution, so in this case the *a priori* regularity assumption can be *a posteriori* removed. Indeed, we can now repeat the argument and gain estimate (2.10) for $u * \rho_\varepsilon$ in lieu of u (with $f * \rho_\varepsilon$ in lieu of f), since $u * \rho_\varepsilon$ satisfies

$$-\Delta(u * \rho_\varepsilon) = f * \rho_\varepsilon.$$

At this stage, we claim that passing to the limit as $\varepsilon \to 0$ we obtain that $u \in H^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ and that the same inequality holds for u, starting from the sole assumption $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$. Indeed, if we combine (2.10) with the standard bound

$$\int_{B_r(x_0)} |w * \rho_\varepsilon|^2 \, dx \le \int_{B_{r+\varepsilon}(x_0)} |w|^2 \, dx$$

applied to both the first derivatives of u and to the datum f, we obtain that the family $(\partial_{x_\alpha}\partial_{x_\beta} u_\varepsilon)$ is uniformly bounded in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ for any given choice of the indices α, β, and since $\nabla u_\varepsilon \to \nabla u$ in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ we conclude (just invoking Proposition A.4) that $u \in H^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ and the weak derivative $\partial_{x_\alpha}\partial_{x_\beta} u$ coincides with any subsequential limit of $\big(\partial_{x_\alpha}\partial_{x_\beta} u_\varepsilon\big)$ as one lets $\varepsilon \to 0$.

The situation is more delicate when the coefficients $A^{\alpha\beta}_{ij}$ are not constant, so that differentiating the equation would introduce extra terms on the right-hand side. Nirenberg's idea (see [78]) is to introduce discrete partial derivatives

$$\Delta_{h,\alpha} u(x) := \frac{u(x + he_\alpha) - u(x)}{h} = \frac{\tau_{h,\alpha} u - u}{h}(x).$$

Remark 2.11. Some basic properties of differentiation are still true and easy to prove:

- (sort of) Leibniz property

$$\Delta_{h,\alpha}(uv) = (\tau_{h,\alpha} u)\Delta_{h,\alpha} v + (\Delta_{h,\alpha} u)v = (\tau_{h,\alpha} v)\Delta_{h,\alpha} u + (\Delta_{h,\alpha} v)u;$$

- integration by parts (ultimately relying on the translation invariance of Lebesgue measure, which grants the identity $\int u\tau_{h,\alpha} v \, dx = \int v\tau_{-h,\alpha} u \, dx$)

$$\int_\Omega \varphi(x)\Delta_{h,\alpha} u(x) \, dx = -\int_\Omega u(x)\Delta_{-h,\alpha}\varphi(x) \, dx \quad \forall\, \varphi \in C^1_c(\Omega;\mathbb{R}),\ |h| < \mathrm{dist}(\mathrm{supp}\,(\varphi), \partial\Omega).$$

In the next lemma we characterize functions in $W^{1,p}$, with $p > 1$, in terms of uniform L^p bounds for the corresponding discrete partial derivatives.

Lemma 2.12. *Consider $u \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$, with $1 < p \le \infty$ and fix $\alpha \in \{1,\dots,n\}$. The partial derivative $\partial_{x_\alpha} u$ belongs to $L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$ if and only if the family $(\Delta_{h,\alpha} u)$ is uniformly bounded in L^p_{loc} as one lets $h \to 0$, more precisely if*

$$\forall\, \Omega' \Subset \Omega \quad \exists\, c = c(\Omega') \quad \textit{such that} \quad \left| \int_{\Omega'} (\Delta_{h,\alpha} u)\, \varphi \right| \le c \|\varphi\|_{L^{p'}(\Omega';\mathbb{R})} \quad \forall\, \varphi \in C^1_c(\Omega';\mathbb{R}),$$

with $1/p + 1/p' = 1$ and $|h| < \mathrm{dist}(\Omega', \partial\Omega)/2$.

Proof. Let us start by assuming that $\partial_{x_\alpha} u \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$ and show that $(\Delta_{h,\alpha} u)$ is uniformly bounded in $L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$. To that scope, one needs to recall the following estimate (also needed in the proof of the Rellich compactness result, Theorem A.14), which holds true for any $|h| < \mathrm{dist}(\Omega', \partial\Omega)$

$$\|\tau_{h,\alpha}\varphi - \varphi\|^p_{L^p(\Omega';\mathbb{R})} \le |h|^p \int_0^1 \int_{\Omega'_{|h|}} |\partial_{x_\alpha}\varphi(y)|^p \, dy\, ds = |h|^p \, \|\partial_{x_\alpha}\varphi\|^p_{L^p(\Omega'_{|h|};\mathbb{R})}$$

where we employed Hölder's inequality and Fubini's theorem. Hence

$$\begin{aligned} \left| \int_{\Omega'} (\Delta_{h,\alpha} u)\, \varphi \right| &\le \|\Delta_{h,\alpha} u\|_{L^p(\Omega';\mathbb{R})} \|\varphi\|_{L^{p'}(\Omega';\mathbb{R})} \\ &\le \|\partial_{x_\alpha} u\|_{L^p(\Omega'_{|h|};\mathbb{R})} \|\varphi\|_{L^{p'}(\Omega';\mathbb{R})} \le \|\partial_{x_\alpha} u\|_{L^p(\Omega'';\mathbb{R})} \|\varphi\|_{L^{p'}(\Omega';\mathbb{R})} \end{aligned}$$

where $\Omega'' := \Omega'_{\overline{h}}$ for $\overline{h} := \mathrm{dist}(\Omega', \partial\Omega)/2$, thus our thesis is proven with

$$c(\Omega') = \|\partial_{x_\alpha} u\|_{L^p(\Omega'';\mathbb{R})}.$$

For the converse implication, fix a domain $\Omega' \Subset \Omega$: we have

$$\left| \int_{\Omega'} u \partial_{x_\alpha}\varphi \, dx \right| = \left| \lim_{h\to 0} \int_{\Omega'} u \Delta_{-h,\alpha}\varphi \, dx \right| = \left| - \lim_{h\to 0} \int_{\Omega'} (\Delta_{h,\alpha} u)\, \varphi \, dx \right| \le c \|\varphi\|_{L^{p'}(\Omega';\mathbb{R})}$$

with $c = c(\Omega')$ and hence, given the duality relation between $L^p(\Omega';\mathbb{R})$ and $L^{p'}(\Omega';\mathbb{R})$, the weak derivative $\partial_{x_\alpha} u$ exists and belongs to $L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$. □

Let us see how Lemma 2.12 allows to obtain a regularity result, still in the simplified case of the Poisson equation. Suppose $f \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ and, as above, $\Delta u = f$ for some $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$. Then translation invariance and linearity of the equation imply

$$-\Delta \tau_{h,\alpha} u = \tau_{h,\alpha} f, \qquad -\Delta(\Delta_{h,\alpha} u) = \Delta_{h,\alpha} f,$$

for any $\Omega' \Subset \Omega$ and for $|h| < \mathrm{dist}(\Omega', \partial\Omega)$. Thanks to Lemma 2.12, $\Delta_{h,\alpha} f$ is bounded in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ and thus by the Caccioppoli-Leray inequality $|\nabla \Delta_{h,\alpha} u|$ is bounded in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$.

As $\Delta_{h,\alpha}(\nabla u) = \nabla \Delta_{h,\alpha} u$ is uniformly bounded in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^n)$, thanks to Lemma 2.12 again (applied componentwise) we get

$$\partial_{x_\alpha}(\nabla u) \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^n).$$

This is the model case to be kept in mind throughout this section.

After these preliminaries about Nirenberg's method, we are now ready to prove the main result concerning H^2 regularity.

Theorem 2.13. *Let Ω be an open domain in $\mathbb{R}^n$. Consider a map $A \in C^{0,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m^2\times n^2})$ such that $A(x) := A^{\alpha\beta}_{ij}(x)$ satisfies the Legendre-Hadamard condition for some continuous, positive ellipticity function $\lambda : \Omega \to \mathbb{R}$, as well as the uniform bound*

$$\sup_{x\in\Omega} |A^{\alpha\beta}_{ij}(x)| \le \Lambda < +\infty.$$

Then, for every $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ weak solution of the equation

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\big) = f_i - \sum_\alpha \partial_{x_\alpha} F^\alpha_i \qquad i = 1,\dots,m$$

with data $f \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ and $F \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$, one has that $u \in H^2_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ and for every choice of subsets $\Omega' \Subset \Omega'' \Subset \Omega$ there exists a constant $c := c(\Omega',\Omega'',A)$ such that

$$\int_{\Omega'} |\nabla^2 u|^2\,dx \le c\left(\int_{\Omega''} |u|^2\,dx + \int_{\Omega''} |f|^2\,dx + \int_{\Omega''} |\nabla F|^2\,dx\right).$$

In order to simplify the notation, in the following proof let γ denote the unit vector corresponding to a given fixed direction and consequently $\tau_h := \tau_{h,\gamma}$ and $\Delta_h := \Delta_{h,\gamma}$.

Remark 2.14. Even though the conclusion above concerns a generic domain $\Omega' \Subset \Omega$, it is enough to prove it for *balls* inside Ω. More precisely, if $2R < \mathrm{dist}(\Omega',\partial\Omega)$, we just need to prove the inequality

$$\int_{B_{R/2}(x_0)} |\nabla^2 u|^2\,dx \le c\left(\int_{B_{2R}(x_0)} |u|^2\,dx + \int_{B_{2R}(x_0)} |f|^2\,dx + \int_{B_{2R}(x_0)} |\nabla F|^2\,dx\right)$$

for any $x_0 \in \Omega'$. The general result can be easily obtained by a compactness and covering argument. In fact, a standard argument allows to prove a scaling-invariant counterpart of such an estimate, namely

$$c_N \int_{B_{R/2}(x_0)} |\nabla^2 u|^2\,dx \le \frac{1}{R^4}\int_{B_{2R}(x_0)} |u|^2\,dx + \int_{B_{2R}(x_0)} |f|^2\,dx + \frac{1}{R^2}\int_{B_{2R}(x_0)} |\nabla F|^2\,dx,$$

where $c_N = c_N(A)$ is a constant only depending on $A_{ij}^{\alpha\beta}$ and possibly on the ambient dimension. We leave the details to the reader.

Notice also that the given statement is redundant, since the term $\sum_\alpha \partial_{x_\alpha} F_i^\alpha$ can always be absorbed into f. We will see however that the optimal estimate, which does not involve derivatives of f, is obtained precisely doing the opposite, i.e. considering heuristically f as a divergence.

Proof. We assume $x_0 = 0$ and, by the previous remark, $F = 0$ (possibly renaming f). In addition, we prove the result under the stronger assumption that the Legendre condition with constant λ holds uniformly in Ω. The general case can be dealt with using Korn's technique, along the lines of the discussion we are about to present in the proof of Theorem 3.16, since for systems with constant coefficients coercivity can be obtained invoking Theorem 1.6.

First note that the given equation is equivalent, by definition, to the identity

$$\int_\Omega \langle A\nabla u, \nabla\varphi\rangle\, dx = \int_\Omega \langle f, \varphi\rangle\, dx \tag{2.11}$$

for all $\varphi \in C_c^\infty(\Omega;\mathbb{R}^m)$. If we apply it to the test function $\tau_{-h}\varphi$ with $|h| < \operatorname{dist}(supp(\varphi), \partial\Omega)$ and we do a change of variable, we find

$$\int_\Omega \langle \tau_h(A\nabla u), \nabla\varphi\rangle\, dx = \int_\Omega \langle \tau_h f, \varphi\rangle\, dx. \tag{2.12}$$

Subtracting (2.11) to equation (2.12) and dividing by h, we get (thanks to the aforementioned discrete Leibniz property)

$$\int_\Omega \langle (\tau_h A)\nabla(\Delta_h u), \nabla\varphi\rangle\, dx = \int_\Omega \langle \Delta_h f, \varphi\rangle\, dx - \int_\Omega \langle (\Delta_h A)\nabla u, \nabla\varphi\rangle\, dx,$$

which is nothing but the weak form of the equation

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big((\tau_h A)_{ij}^{\alpha\beta}\partial_{x_\beta} v^j\big) = f_i' - \sum_\alpha \partial_{x_\alpha} G_i^\alpha \qquad i = 1,\dots,m \tag{2.13}$$

for $v = \Delta_h u$ and with data $f' := \Delta_h f$ and $G := -(\Delta_h A)\nabla u$.

Now, the basic idea of the proof is to use the Caccioppoli-Leray inequality. However, a direct application of such an inequality would lead to an estimate having the L^2 norm of f' on the right-hand side, and we know from Lemma 2.12 that this norm can be uniformly bounded in h only if $f \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$. Hence, we will rather revisit its proof, trying to get estimates depending only on the L^2 norm of f (heuristically, we see f' as a divergence).

To this aim, take a cutoff function η compactly supported in B_R, with $0 \le \eta \le 1$, identically equal to 1 on $B_{R/2}$ and such that $|\nabla\eta| \le 4/R$, and consider equation (2.13)

with the test function $\varphi := \eta^2 \Delta_h u = \eta^2 v$ with $|h| < R/2$. Using Young's inequality as in the proof of Theorem 2.1, we get

$$\begin{aligned}\frac{3\lambda}{4}\int_{B_R} \eta^2 |\nabla v|^2\,dx \leq\ & \frac{4\Lambda\varepsilon}{R}\int_{B_R} \eta^2|\nabla v|^2\,dx \\ & + \left(\frac{4\Lambda}{R\varepsilon}+\frac{4}{R^2}\right)\int_{B_R}|v|^2\,dx + \int_{B_R}\eta^2 v\Delta_h f\,dx + \left(\frac{1}{\lambda}+4\right)\int_{B_R}|G|^2\,dx,\end{aligned}$$

with λ, Λ depending only on A.

As in the proof of Theorem 2.1, we choose $\varepsilon > 0$ so to absorb the term corresponding to the L^2-norm of $\eta\nabla v$ in the left-hand side of the inequality, so that, up to some constant $c > 0$ depending on (λ, Λ, R), we get

$$c\int_{B_R}\eta^2|\nabla v|^2\,dx \leq \int_{B_R}|v|^2\,dx + \int_{B_R}\eta^2 v\Delta_h f\,dx + \int_{B_R}|G|^2\,dx. \tag{2.14}$$

We now study each term of (2.14) separately. Firstly

$$\int_{B_R}|v|^2\,dx \leq \int_{B_{R+h}}|\nabla u|^2\,dx$$

by means of the estimate on incremental ratios as mentioned in the proof of Lemma 2.12. The right-hand side can in turn be estimated using the classical Caccioppoli-Leray inequality for u between the balls $B_{3R/2}$ and B_{2R}, and it gives an upper bound of the desired form.

Concerning the term $\int \eta^2 v\Delta_h f\,dx$, by means of discrete integration by parts and the Young inequality, we can write

$$\left|\int_{B_R}\eta^2 v\Delta_h f\,dx\right| \leq \tilde{\varepsilon}\int_{B_R}|\Delta_{-h}(\eta^2 v)|^2\,dx + \frac{1}{\tilde{\varepsilon}}\int_{B_R}|f|^2\,dx\,. \tag{2.15}$$

The first term in the right-hand side of (2.15) can be estimated (since $|\nabla\eta^2|^2 \leq 64/R^2$) with

$$\int_{B_{R+h}}|\nabla(\eta^2 v)|^2\,dx \leq 2\int_{B_{R+h}}\eta^4|\nabla v|^2\,dx + \frac{128}{R^2}\int_{B_{R+h}}|v|^2\,dx,$$

so that, choosing $\tilde{\varepsilon}$ sufficiently small and using the inequality $\eta^4 \leq \eta^2$, we can absorb the first term and use once more the Caccioppoli-Leray inequality to estimate $\int_{B_{R+h}}|v|^2\,dx$. The term involving the L^2-norm of G can be estimated in the very same way, using this time also the local Lipschitz assumption on A to bound $\Delta_h A$, so that finally we put together all the corresponding estimates to obtain the thesis (the conclusion comes from Lemma 2.12 and then letting $h \to 0$ in the estimate involving $v = \Delta_h u$). □

Remark 2.15. It should be clear from the proof that the previous result only concerns *interior* regularity and cannot be used in order to get information about the behavior of the function u near the boundary $\partial\Omega$. In other terms, we can not guarantee that the constant c remains bounded as Ω' invades Ω (so that $R \to 0$), even if global regularity assumptions on A, u, f and F are made. The issue of boundary regularity requires different techniques, that will be described in Section 2.5.

Remark 2.16. Given a problem having the form considered in the statement of Theorem 2.13 if we make stronger assumption on the data we can iterate the method above to get stronger conclusions. For instance, if $f \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m), F \in H^2_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$ and $A \in C^{1,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m^2\times n^2})$ then one can apply this argument to any first partial derivative of u to gain that in fact $u \in H^3_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$. It follows that if f, F and A are smooth (C^∞), then so will be the solution u.

2.4 Decay estimates for systems with constant coefficients

Our next target towards the development of a general regularity theory is to derive some refined decay estimates for solutions of homogeneous equations defined by differential operators with constant coefficients. These will come into play in Chapter 3 when presenting Schauder theory.

Lemma 2.17. *Let $A = A^{\alpha\beta}_{ij}$ be a constant matrix satisfying the Legendre-Hadamard condition for some $\lambda > 0$, let $\Lambda = |A|$ and let $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ satisfy the system*

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\right) = 0 \qquad i = 1,\dots,m.$$

Then, these two inequalities hold true for any $B_r(x_0) \subset B_R(x_0) \Subset \Omega$:

$$\int_{B_r(x_0)} |u|^2\,dx \le c_D\left(\frac{r}{R}\right)^n \int_{B_R(x_0)} |u|^2\,dx \tag{2.16}$$

$$\int_{B_r(x_0)} |u - u_{x_0,r}|^2\,dx \le c_E\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u - u_{x_0,R}|^2\,dx \tag{2.17}$$

with $c_D = c_D(n,\lambda,\Lambda), c_E = c_E(n,\lambda,\Lambda)$ depending only on n, λ and Λ.

We remind the reader that $u_{x_0,s}$ stands for the mean value of the function u on the ball $B_s(x_0)$.

Proof. For the sake of readability, we split the proof in two steps.

Part I: proof of (2.16)**.** By a standard rescaling argument, it is enough to study the case $R = 1$. In the sequel of this proof, let k be the smallest integer such that $k > n/2$ (and thus $H^k \hookrightarrow C^0$). First of all, by the Caccioppoli-Leray inequality, we have that

$$c_{CL} \int_{B_{1/2}(x_0)} |\nabla u|^2 \, dx \leq \int_{B_1(x_0)} |u|^2 \, dx$$

for some positive constant $c_{CL} = c_{CL}(\lambda, \Lambda)$. Now, for any $\alpha \in \{1, 2, \ldots, n\}$, we know that $\partial_{x_\alpha} u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ by Theorem 2.13, and since the matrix A has constant coefficients, this function solves the same linear equation. Hence, we can iterate the argument in order to get an estimate having the form

$$\int_{B_{2^{-k}}(x_0)} \sum_{d \leq k} |\nabla^d u|^2 \leq c_k \int_{B_1(x_0)} |u|^2 \, dx$$

for some $c_k = c_k(n, \lambda, \Lambda) > 0$, where we have denoted by $\nabla^d u$ the tensor collecting all partial derivatives of order k of the function u (thus consisting of $n^k m$ components). Consequently, thanks to our choice of the integer k, we can find a constant $\kappa = \kappa(n, \lambda, \Lambda)$ such that

$$\sup_{B_{2^{-k}}(x_0)} |u|^2 \, dx \leq \kappa \int_{B_1(x_0)} |u|^2 \, dx.$$

In order to conclude the proof, it is appropriate to distinguish two cases. If $r \leq 2^{-k}$, then

$$\int_{B_r(x_0)} |u|^2 \, dx \leq \omega_n r^n \sup_{B_{2^{-k}}(x_0)} |u|^2 \leq \kappa \omega_n r^n \int_{B_1(x_0)} |u|^2 \, dx,$$

where ω_n denotes the Lebesgue measure of the unit ball in $\mathbb{R}^n$. Hence, for this case we have the thesis, provided we simply set $c_D = \kappa(n, \lambda, \Lambda)\omega_n$. If instead $r \in (2^{-k}, 1)$, then we just need to rely on the trivial inequality $\int_{B_r(x_0)} |u|^2 \, dx \leq \int_{B_1(x_0)} |u|^2 \, dx$ and ensure that $c_D \geq 2^{kn}$.

We can now prove the second inequality, that concerns the notion of variance of the function u on a ball.

Part II: proof of (2.17)**.** Again, it is useful to study two cases separately. If $r \leq R/2$, then by the Poincaré inequality (see Theorem A.16) there exists a constant $c_{P,II}$ such that

$$\int_{B_r(x_0)} |u - u_{x_0,r}|^2 \, dx \leq c_{P,II} r^2 \int_{B_r(x_0)} |\nabla u|^2 \, dx$$

and so

$$\begin{aligned}\int_{B_r(x_0)} |u-u_{x_0,r}|^2\,dx &\le c_{P,II} r^2 c_D \left(\frac{2r}{R}\right)^n \int_{B_{R/2}(x_0)} |\nabla u|^2\,dx \\ &\le \frac{c_{P,II}}{c_{CL}} c_D \left(\frac{2r}{R}\right)^{n+2} \int_{B_R(x_0)} |u-u_{x_0,R}|^2\,dx\end{aligned}$$

respectively by the previous result applied to the gradient ∇u and by the Caccioppoli-Leray inequality. For the case $R/2 < r \le R$ we need to use the following fact (see Remark 2.18 below): the mean value $u_{x_0,r}$ is the unique minimizer of

$$m \longmapsto \int_{B_r(x_0)} |u-m|^2\,dx\,. \tag{2.18}$$

If we give this for granted, the conclusion is easy because

$$\int_{B_r(x_0)} |u-u_{x_0,r}|^2\,dx \le \int_{B_r(x_0)} |u-u_{x_0,R}|^2\,dx \le 2^{n+2}\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u-u_{x_0,R}|^2\,dx$$

so that the claimed inequality holds true provided we require $c_E \ge 2^{n+2}\cdot\max\left\{\frac{c_{P,II}}{c_{CL}}c_D; 1\right\}$. □

Remark 2.18. Let us go back to the study of

$$\inf_{m\in\mathbb{R}} \int_\Omega |u-m|^p\,dx \tag{2.19}$$

for $1 \le p < \infty$ and $u \in L^p(\Omega;\mathbb{R})$ where Ω is any open, bounded domain in $\mathbb{R}^n$. As we pointed out above, this problem is easily solved, when $p = 2$, by the mean value u_Ω: it suffices to notice that

$$\int_\Omega |u-m|^2\,dx = \int_\Omega |u|^2\,dx - 2m\int_\Omega u\,dx + m^2\mathscr{L}^n(\Omega)\,.$$

Nevertheless, this is not true in general when $p \ne 2$. Of course

$$\inf_{m\in\mathbb{R}} \int_\Omega |u-m|^p\,dx \le \int_\Omega |u-u_\Omega|^p\,dx$$

but we also claim that, for any $m \in \mathbb{R}$, we have

$$\int_\Omega |u-u_\Omega|^p\,dx \le 2^p \int_\Omega |u-m|^p\,dx\,. \tag{2.20}$$

Since the problem is clearly translation-invariant, it is sufficient to prove inequality (2.20) for $m = 0$. But in this case

$$\int_\Omega |u - u_\Omega|^p \, dx \le 2^{p-1} \int_\Omega |u|^p \, dx + 2^{p-1} \int_\Omega |u_\Omega|^p \, dx \le 2^p \int_\Omega |u|^p \, dx \,,$$

thanks to the elementary inequality $|a+b|^p \le 2^{p-1}\big(|a|^p + |b|^p\big)$ and to the fact that

$$\int_\Omega |u_\Omega|^p \, dx \le \int_\Omega |u|^p \, dx$$

which is a standard consequence of the Hölder (or Jensen) inequality. Incidentally, let us remark that for any given $p > 1$ the minimum for (2.19) is unique by strict convexity of the function $t \mapsto |t|^p$.

2.5 Regularity up to the boundary

Following the same conceptual scheme of Section 2.3, let us first consider a simple special case. Suppose we have to deal with the Poisson equation

$$-\Delta u = f \qquad \text{in } \Omega \tag{2.21}$$

where $\Omega := (-a, a)^{n-1} \times (0, a)$ is a rectangle in $\mathbb{R}^n$ with sides parallel to the coordinate axes. We consider the equation in $H^1(\Omega; \mathbb{R})$, with Dirichlet boundary condition $u = 0$ on $\{x_n = 0\}$. Let us use coordinates $x = (x', x_n)$ with $x' \in \mathbb{R}^{n-1}$ and assume $f \in L^2(\Omega; \mathbb{R})$.

The rectangle $\Omega' = (-a/2, a/2)^{n-1} \times (0, a/2)$ is not relatively compact in Ω, nevertheless via Nirenberg's method one can prove estimates of the form

$$\int_{\Omega'} \left|\partial_{x_\gamma} \nabla u\right|^2 \, dx \le \frac{c}{a^2} \int_\Omega |\nabla u|^2 \, dx + c \int_\Omega |f|^2$$

for $\gamma = 1, 2, \ldots, n-1$, with $c > 0$ a purely dimensional constant. Indeed, this is achieved by choosing test functions of the form $\varphi = \eta^2 \Delta_{h,\gamma} u$, where the support of η is allowed to touch the hyperplane $\{x_n = 0\}$ (because of the homogeneous Dirichlet boundary condition on u). Equation (2.21) can then be rewritten as

$$-\frac{\partial^2 u}{\partial x_n^2} = \Delta_{x'} u + f$$

and here the right-hand side $\Delta_{x'} u + f$ belongs to $L^2(\Omega'; \mathbb{R})$. We conclude that also the second derivative in the x_n direction is in $L^2(\Omega'; \mathbb{R})$, hence $u \in H^2(\Omega'; \mathbb{R})$. Notice that this argument only requires only the validity of the homogeneous Dirichlet condition on the portion $\{x_n = 0\}$ of the boundary of Ω. In addition, this homogeneous Dirichlet condition

also ensures that all functions $\partial_{x_\gamma} u$, with $\gamma = 1, \dots, n-1$, have null trace on $\{x_n = 0\}$, which is crucial for the iteration of this argument with higher-order derivatives (see the statement of Theorem 2.20 below).

Now we want to use this idea in order to study the regularity up to the boundary for linear elliptic problems having the general form

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A_{ij}^{\alpha\beta}\partial_{x_\beta} u^j\big) = f_i - \sum_{\alpha} \partial_{x_\alpha} F_i^\alpha \qquad i = 1, \dots, m \tag{2.22}$$

for $u \in H_0^1(\Omega; \mathbb{R}^m)$. We make the following hypotheses:

(a) $f \in L^2(\Omega; \mathbb{R}^m)$;

(b) $F \in H^1(\Omega; \mathbb{R}^{m\times n})$;

(c) $A \in C^{0,1}(\Omega; \mathbb{R}^{m^2\times n^2})$;

(d) $A(x)$ satisfies the Legendre-Hadamard condition uniformly in Ω;

(e) Ω has a C^2 boundary, namely the domain is locally the epigraph of a C^2 function, up to a rigid motion.

Theorem 2.19. *Under the previous assumptions, the function u belongs to $H^2(\Omega; \mathbb{R}^m)$ and*

$$\|u\|_{H^2(\Omega;\mathbb{R}^m)} \le c\big(\|f\|_{L^2(\Omega;\mathbb{R}^m)} + \|F\|_{H^1(\Omega;\mathbb{R}^{m\times n})}\big),$$

for some constant $c = c(\Omega, A, n)$.

Proof. Since we already have the interior regularity result at our disposal, it suffices to show that for any $x_0 \in \partial\Omega$ there exists a neighborhood Ω' of x_0 in Ω such that $u \in H^2(\Omega'; \mathbb{R}^m)$, with suitable estimates. Without loss of generality we shall assume $x_0 = 0$.

Part I: flat boundary. Assume first we can find a neighborhood Ω_0 of $x_0 = 0$ in $\mathbb{R}^n$ such that $\Omega \cap \Omega_0$ is a rectangle of the form $(-a, a)^{n-1} \times (0, a)$ for some $a > 0$. By applying Nirenberg's method as described above in the special case of the constant coefficients operator $-\Delta$ we get $\partial_{x_\gamma} u^i \in H^1(\Omega'; \mathbb{R})$ for $\gamma = 1, 2, \dots, n-1$ and $i = 1, 2, \dots, m$, where we have set $\Omega' = (-a/2, a/2)^{n-1} \times (0, a/2)$.

At this stage, equation (2.22) readily implies that $\sum_j \partial_{x_n}(A_{ij}^{nn}\partial_{x_n} u^j) \in L^2(\Omega'; \mathbb{R})$ for any $i \in \{1, 2, \dots, m\}$. In turn this gives, by the Leibniz rule, that the function $\sum_j A_{ij}^{nn}\partial^2_{x_n x_n} u^j$ belongs to $L^2(\Omega'; \mathbb{R})$ for any choice of i; this argument is formal because one of the factors is only a distribution (a priori not yet a function). To make this step

rigorous, we work with the difference quotients in the x_n direction and employ the discrete Leibniz rule: since by Lemma 2.12 the difference quotients $\sum_j \Delta_h(A^{nn}_{ij}\partial_{x_n}u^j)$ have uniformly bounded L^2 norm in $\Omega'_h = \{x \in \Omega' : \mathrm{dist}(x, \partial\Omega') > h\}$, we obtain that the same conclusion holds true for $\sum_j A^{nn}_{ij}\Delta_h\partial_{x_n}u^j$. However, the matrix A^{nn}_{ij} is invertible with $\det(A^{nn}_{ij}) \geq \lambda^m$ (as a consequence of the Legendre-Hadamard condition applied to $\xi = a \otimes b$ for $a_\alpha = \delta^n_\alpha$) and thus we get

$$\limsup_{h\to 0^+} \int_{\Omega'_h} |\Delta_h\partial_{x_n}u^j|^2\, dx < +\infty$$

which gives $\partial^2_{x_nx_n}u^j \in L^2(\Omega';\mathbb{R})$, again by invoking Lemma 2.12.

Part II: flattening the boundary. By assumption, there exist a defining function $h \in C^2(\mathbb{R}^{n-1};\mathbb{R})$ and a neighborhood Ω_0 of the origin in $\mathbb{R}^n$ such that (up to a rigid motion, so that the hyperplane $\{x_n = 0\}$ is tangent to $\partial\Omega$ at the origin)

$$\Omega \cap \Omega_0 = \{x \in \Omega_0 :\ x_n > h(x')\}\,.$$

As a result, we can define the change of variables $x'_n = x_n - h(x')$ and the function $H(x', x_n) = (x', x_n - h(x'))$ mapping $\Omega \cap \Omega_0$ onto $H(\Omega \cap \Omega_0)$ (which contains a rectangle of the form $\tilde{\Omega} = (-a, a)^{n-1} \times (0, a)$). We further set

$$\tilde{\Omega}' = (-a/2, a/2)^{n-1} \times (0, a/2), \quad \Omega' := H^{-1}(\tilde{\Omega}').$$

It is clear that $H : \Omega' \to \tilde{\Omega}'$ is invertible and, denoted its inverse by G, both H and G are C^2 functions. Moreover ∇H is a triangular matrix with $\det(\nabla H) = 1$. Furthermore, the maps G and H induce isomorphisms between both H^1 and H^2 spaces (via change of variables in the definition of weak derivatives) and thus it suffices to show that the function $v(y) = u(G(y))$ belongs to $H^2(\tilde{\Omega};\mathbb{R}^m)$. To this aim, we need to check that v solves in $\tilde{\Omega}$ the problem

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(\tilde{A}^{\alpha\beta}_{ij}\partial_{x_\beta}v^j\big) = \tilde{f}_i - \partial_{x_\alpha}\tilde{F}^\alpha_i \qquad i = 1, \dots, m \tag{2.23}$$

subject to the boundary condition $v = 0$ on $\{y'_n = 0\} \cap \tilde{\Omega}$, to be interpreted in the usual weak sense of traces, and where it has been set

$$\widetilde{f}_i(y) = f_i(G(y)), \quad \widetilde{F}^\alpha_i(u) = \sum_\gamma F^\gamma_i \partial_{x_\gamma} H^\alpha(G(y)), \quad \widetilde{A}^{\alpha\beta}_{ij} = \sum_{\alpha',\beta'} \partial_{x_{\alpha'}} H^\alpha A^{\alpha'\beta'}_{ij} \partial_{x_{\beta'}} H^\beta(G(y)). \tag{2.24}$$

These formulae can be easily derived by means of an elementary computation, starting from the weak formulation of the problem and applying a change of variables in order to express the different integrals in terms of the new coordinates. For instance

$$\sum_i \int_{\Omega'} f_i(x)\varphi^i(x)\,dx = \sum_i \int_{\tilde{\Omega}'} f_i(G(y))\varphi^i(G(y))\det(\nabla G(y))\,dy,$$

but then we recall that $\det(\nabla G) = 1$ and we can set $\psi(y) = \varphi(G(y))$ so that the previous equation takes the form

$$\sum_i \int_{\Omega'} f_i(x)\varphi^i(x)\,dx = \sum_i \int_{\tilde{\Omega}'} \widetilde{f}_i(y)\psi^i(y)\,dy.$$

The computation for $\widetilde{F}$ or $\widetilde{A}$ is less trivial, but there is no conceptual difficulty. We just see the first one:

$$\begin{aligned}\sum_{\gamma,i} \int_{\Omega'} F_i^\gamma(x)\partial_{x_\gamma}\varphi^i(x)\,dx &= \sum_{\gamma,i} \int_{\tilde{\Omega}'} F_i^\gamma(G(y))\partial_{x_\gamma}\varphi^i(G(y))\det(\nabla G(y))\,dy \\ &= \sum_{\alpha,\gamma,i} \int_{\tilde{\Omega}'} F_i^\gamma(G(y))\partial_{y_\alpha}\psi^i(y)\partial_{x_\gamma}H^\alpha(G(y))\,dy\end{aligned}$$

which leads to the conclusion. Note that here and above the arbitrary test function φ has been replaced by the arbitrary test function ψ. However, we should be concerned whether the conditions we have imposed on A (in particular, the Legendre-Hadamard condition) still hold true for $\widetilde{A}$. This is indeed the case and we can verify it directly by means of the expression of $\widetilde{A}$ given in equation (2.24). For any $\tilde{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$

$$\begin{aligned}\sum_{\alpha,\beta,i,j} \widetilde{A}_{ij}^{\alpha\beta}(y)\tilde{a}_\alpha\tilde{a}_\beta b^i b^j &= \sum_{\alpha,\beta,i,j}\sum_{\alpha',\beta'} A_{ij}^{\alpha'\beta'}(G(y))\partial_{x_{\alpha'}}H^\alpha(G(y))\tilde{a}_\alpha\partial_{x_{\beta'}}H^\beta(G(y))\tilde{a}_\beta b^i b^j \\ &\geq \sum_{\alpha',\beta',i,j}\sum_{\alpha,\beta} A_{ij}^{\alpha'\beta'}(G(y))\partial_{x_{\alpha'}}H^\alpha(G(y))\tilde{a}_\alpha\partial_{x_{\beta'}}H^\beta(G(y))\tilde{a}_\beta b^i b^j \\ &\geq \lambda\sum_{\alpha'}|\sum_\alpha \partial_{x_{\alpha'}}H^\alpha(G(y))\tilde{a}_\alpha|^2\sum_i |b^i|^2.\end{aligned}$$

and the conclusion comes from the invertibility of the Jacobian of the map H at $G(y)$. Thus, $\widetilde{A}$ satisfies the Legendre-Hadamard condition for an appropriate constant $\lambda' > 0$ depending on λ and H, and of course $\widetilde{A} \in C^{0,1}(\tilde{\Omega}';\mathbb{R})$.

Through this transformation of the domain, we can finally apply the argument presented in the first part of the proof applied here to the domains $(\tilde{\Omega}, \tilde{\Omega}')$ and find that $v \in H^2(\tilde{\Omega}';\mathbb{R})$, with suitable integral bounds. Coming back to the original variables we obtain the H^2 regularity of u in Ω'' and the claimed estimates. □

If both the boundary of the domain and the data are sufficiently regular, this method can be iterated to get the following theorem.

Theorem 2.20. *Assume, in addition to the hypotheses above, that* $f \in H^k(\Omega;\mathbb{R}^m)$ *and also* $F \in H^{k+1}(\Omega;\mathbb{R}^{m\times n})$, $A \in C^{k,1}(\Omega,\mathbb{R}^{m^2\times n^2})$ *with* Ω *such that* $\partial\Omega \in C^{k+2}$. *Then* $u \in H^{k+2}(\Omega;\mathbb{R}^m)$.

We are not going to present the detailed proof of the previous result, but the basic ideas to get started consist in noticing that (under those assumptions) all tangential derivatives to u also belong to $H^1_0(\Omega;\mathbb{R}^m)$ and differentiating the equation with respect to each fixed tangential direction to derive a system having the form

$$-\sum_{\alpha,\beta,j}\partial_{x_\alpha}\left(A^{\alpha\beta}_{ij}\partial_{x_\beta}(\partial_{x_\gamma}u^j)\right) = \partial_{x_\gamma}f_i - \sum_{\alpha}\left(\partial_{x_\alpha}(\partial_{x_\gamma}F^\alpha_i) + \sum_{\beta,j}\partial_{x_\alpha}\left(\partial_{x_\gamma}A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\right)\right).$$

3 Interior regularity for nonlinear equations

So far, we have just dealt with linear problems and the wealth of different situations was only based on the possibility of varying the elliptic operator, the boundary conditions and the number of dimensions involved in the equations. We will see now that Nirenberg's technique is also particularly appropriate in dealing with nonlinear partial differential equations, as those arising from Euler-Lagrange equations of non-quadratic functionals.

Consider a function $L \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$ and assume the following:

(i) there exists a constant $C > 0$ such that $|\nabla^2 L(\xi)| \le C$ for any $\xi \in \mathbb{R}^{m\times n}$;

(ii) L satisfies a uniform Legendre condition, namely

$$\sum_{\alpha,\beta,i,j} \partial_{p_i^\alpha}\partial_{p_j^\beta} L(p)\xi_i^\alpha \xi_j^\beta \ge \lambda\,|\xi|^2\,, \quad \forall \xi \in \mathbb{R}^{m\times n},$$

for some $\lambda > 0$ independent of $p \in \mathbb{R}^{m\times n}$.

For notational convenience, set $B_i^\alpha := \frac{\partial L}{\partial p_i^\alpha}$ and $A_{ij}^{\alpha\beta} := \frac{\partial^2 L}{\partial p_i^\alpha \partial p_j^\beta}$ and notice that $A_{ij}^{\alpha\beta}$ is symmetric with respect to the transformation $(\alpha, i) \to (\beta, j)$.

Let $\Omega \subset \mathbb{R}^n$ be an open domain and let $u \in H^1_{\rm loc}(\Omega;\mathbb{R}^m)$ be a local minimizer of the functional

$$w \longmapsto \mathcal{L}(w) := \int_\Omega L(\nabla w)\,dx\,.$$

We wish to discuss the implication

$$L \in C^\infty \Rightarrow u \in C^\infty$$

which is strictly related to Hilbert's *XIX* problem (initially posed for functions of two variables and in the category of analytic functions, see [52]). In the sequel of this chapter we will first treat the case $n = 2$ and much later the case $n \ge 3$, which is significantly harder.

Recall that u is a local minimizer for $\mathcal{L}$ when

$$u' \in H^1_{\rm loc}(\Omega;\mathbb{R}^m),\ {\rm supp}(u-u') \subset \Omega' \Subset \Omega \quad \Longrightarrow \quad \int_{\Omega'} L(\nabla u')\,dx \ge \int_{\Omega'} L(\nabla u)\,dx\,.$$

If this is the case, we have already seen how the Euler-Lagrange equation can be obtained: considering perturbations of the form $u' = u + t\varphi$, with $\varphi \in C_c^\infty(\Omega;\mathbb{R}^m)$, one can prove (using the fact that the regularity assumptions on L allow differentiation under the integral sign) that

$$0 = \frac{d}{dt}\left[\int_\Omega L(\nabla u + t\nabla\varphi)\,dx\right]_{t=0} = \sum_{\alpha,i}\int_\Omega B_i^\alpha(\nabla u)\frac{\partial \varphi^i}{\partial x_\alpha}\,dx\,.$$

Now, suppose γ is a fixed coordinate direction (and let e_γ be the corresponding unit vector) and $h > 0$ a small positive increment: if we apply the previous argument to a test function having the form $\tau_{-h,\gamma}\varphi$, we get

$$\sum_{\alpha,i} \int_\Omega \tau_{h,\gamma}(B_i^\alpha(\nabla u)) \frac{\partial \varphi^i}{\partial x_\alpha}\, dx = 0$$

and consequently, subtracting this equation to the previous one and dividing by h

$$\sum_{\alpha,i} \int_\Omega \Delta_{h,\gamma}(B_i^\alpha(\nabla u)) \frac{\partial \varphi^i}{\partial x_\alpha}\, dx = 0\,.$$

However, as a consequence of the regularity of L, for any $\alpha = 1, \dots, n$ and any $i = 1, \dots, m$ we can then write

$$\begin{aligned} B_i^\alpha(\nabla u(x + he_\gamma)) - B_i^\alpha(\nabla u(x)) &= \int_0^1 \frac{d}{dt} B_i^\alpha(t\nabla u(x + he_\gamma) + (1-t)\nabla u(x))\, dt \\ = \sum_{\beta,j} \left[\int_0^1 A_{ij}^{\alpha\beta}(t\nabla u(x + he_\gamma) + (1-t)\nabla u(x))\, dt \right] &\left[\frac{\partial u^j}{\partial x_\beta}(x + he_\gamma) - \frac{\partial u^j}{\partial x_\beta}(x) \right] \end{aligned}$$

and setting

$$\widetilde{A}_{ij,h}^{\alpha\beta}(x) := \int_0^1 A_{ij}^{\alpha\beta}(t\nabla u(x + he_\gamma) + (1-t)\nabla u(x))\, dt$$

we rephrase the previous condition as

$$\sum_{\alpha,\beta,i,j} \int_\Omega \widetilde{A}_{ij,h}^{\alpha\beta}(x) \frac{\partial \Delta_{h,\gamma} u^j}{\partial x_\beta}(x) \frac{\partial \varphi^i}{\partial x_\alpha}(x)\, dx = 0\,.$$

Hence, the function $w = \Delta_{h,\gamma} u$ solves the equation

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(\widetilde{A}_{ij,h}^{\alpha\beta} \partial_{x_\beta} w^j\big) = 0 \qquad\qquad i = 1, \dots, m. \tag{3.1}$$

It is obvious from the definition that $\widetilde{A}_{ij,h}^{\alpha\beta}(x)$ satisfies both the Legendre condition for the given constant $\lambda > 0$ and a uniform upper bound on the L^∞ norm. Therefore we can apply the Caccioppoli-Leray inequality to the problem (3.1) to obtain a positive constant c, not depending on h, such that

$$\int_{B_R(x_0)} |\nabla(\Delta_{h,\gamma} u)|^2\, dx \le \frac{c}{R^2} \int_{B_{2R}(x_0)} |\Delta_{h,\gamma} u|^2\, dx \le \frac{c}{R^2} \int_{B_{2R+h}(x_0)} |\nabla u|^2\, dx$$

for any $B_R(x_0) \subset B_{2R}(x_0) \Subset \Omega$. As a result, by Lemma 2.12, we deduce that

$$u \in H^2_{\mathrm{loc}}(\Omega; \mathbb{R}^m), \tag{3.2}$$

and in fact the Hessian $\nabla^2 u$ satisfies an L^2 integral bound of the type above. Moreover, we have that

(i) $\Delta_{h,\gamma} u \to \partial_{x_\gamma} u$ in $L^2_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ (this is clearly true if u is regular and then the result for $u \in H^2_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ follows via a standard approximation argument; furthermore we can exploit the fact that the operators $\Delta_{h,\gamma}$ are uniformly bounded, in the sense encoded in the previous estimate, to gain weak H^1 subsequential convergence);

(ii) as a result of part (i), $w = \partial_{x_\gamma} u$ satisfies the equation

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha} \big(A^{\alpha\beta}_{ij}(\nabla u) \partial_{x_\beta} w^j \big) = 0 \qquad i = 1, \dots, m, \tag{3.3}$$

in the standard weak sense. In fact, notice that

$$\int_0^1 A^{\alpha\beta}_{ij}(t\nabla u(x + he_\gamma) + (1-t)\nabla u(x))\, dt \overset{h\to 0}{\to} A^{\alpha\beta}_{ij}(\nabla u(x))$$

in L^p_{loc} for any $1 \le p < \infty$, as an easy consequence of the continuity of translations in L^p_{loc} and the continuity of A.

In order to solve Hilbert's *XIX* problem, we would like to apply a classical result by Schauder (see [85]) saying that, if w is a weak solution of a problem in divergence form, namely if

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha} \big(B^{\alpha\beta}_{ij} \partial_{x_\beta} w^j \big) = 0 \qquad i = 1, \dots, m$$

then

$$B \in C^{0,\alpha}(\Omega; \mathbb{R}^{m^2 \times n^2}) \Rightarrow w \in C^{1,\alpha}(\Omega; \mathbb{R}^m),$$

and so $u \in C^{2,\alpha}(\Omega; \mathbb{R}^m)$ and $B \in C^{1,\alpha}(\Omega; \mathbb{R}^{m^2 \times n^2})$. At that stage, this differentiation argument can be iterated infinitely many times as long as $L \in C^\infty$. But to do so we first need to improve the regularity of $B(x) = A(\nabla u(x))$. As a matter of fact, at this point we only know that $A(\nabla u) \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^{m^2 \times n^2})$, while for this argument to work we would rather need $A(\nabla u) \in C^{0,\alpha}(\Omega; \mathbb{R}^{m^2 \times n^2})$. When $n = 2$ we can apply Widman's technique (based on equation (2.8)) to the problem (3.3) to obtain Hölder regularity of ∇u, both in the scalar and in the vectorial case. The situation is much harder in the case $n > 2$, which requires deep new ideas. The celebrated theory by De Giorgi-Nash-Moser solves the problem in the scalar case, while for vector-valued maps new difficulties arise.

3.1 Hölder, Morrey and Campanato spaces

In this section we introduce the Hölder spaces $C^{0,\alpha}$, the Morrey spaces $L^{p,\lambda}$ and the Campanato spaces $\mathcal{L}^{p,\lambda}$. All these spaces, besides the standard Lebesgue spaces, play an important role in elliptic regularity theory.

Definition 3.1 (Hölder spaces)**.** *Given $A \subset \mathbb{R}^n$, $u : A \to \mathbb{R}^m$ and $\alpha \in (0,1]$ we define the α-Hölder semi-norm on A as*

$$[u]_{C^{0,\alpha}(A;\mathbb{R}^m)} := \sup_{x \neq y \in A} \frac{|u(x) - u(y)|}{|x-y|^\alpha}.$$

We say that u is α-Hölder in A, and write $u \in C^{0,\alpha}(A;\mathbb{R}^m)$, if $[u]_{C^{0,\alpha}(A;\mathbb{R}^m)} < +\infty$.

If $\Omega \subset \mathbb{R}^n$ is open, we say that $u : \Omega \to \mathbb{R}^m$ is locally α-Hölder if for any $x \in \Omega$ there exists a neighborhood $U_x \Subset \Omega$ such that $[u]_{C^{0,\alpha}(U_x;\mathbb{R}^m)} < +\infty$. The corresponding vector space of functions is denoted by $C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$.

If $k \in \mathbb{N}$, the space of functions of class $C^k(\Omega;\mathbb{R}^m)$ with all $i-$th derivatives $\nabla^i u$ with $|i| \leq k$ in $C^{0,\alpha}(\Omega;\mathbb{R}^m)$ will be denoted by $C^{k,\alpha}(\Omega;\mathbb{R}^m)$.

Remark 3.2. The spaces $C^{k,\alpha}(\Omega;\mathbb{R}^m)$ are Banach spaces when endowed with the norm

$$\|u\|_{C^{k,\alpha}(\Omega;\mathbb{R}^m)} = \|u\|_{C^k(\Omega;\mathbb{R}^m)} + [\nabla^k u]_{C^{0,\alpha}(\Omega;\mathbb{R}^m)}.$$

Definition 3.3 (Morrey spaces)**.** *Assume $\Omega \subset \mathbb{R}^n$ open, $\lambda \geq 0$ and $1 \leq p < \infty$. We say that a function $f \in L^p(\Omega;\mathbb{R})$ belongs to $L^{p,\lambda}(\Omega;\mathbb{R})$ if*

$$\sup_{0<r<d_\Omega,\, x_0 \in \Omega} r^{-\lambda} \int_{\Omega(x_0,r)} |f|^p \, dx < +\infty$$

where $\Omega(x_0,r) := \Omega \cap B_r(x_0)$ and d_Ω is the diameter of Ω. It is easy to verify that

$$\|f\|_{L^{p,\lambda}(\Omega;\mathbb{R})} := \left(\sup_{0<r<d_\Omega,\, x_0 \in \Omega} r^{-\lambda} \int_{\Omega(x_0,r)} |f|^p \, dx \right)^{\frac{1}{p}}$$

is a norm on $L^{p,\lambda}(\Omega;\mathbb{R})$.

Remark 3.4. We mention here some of the basic properties of the Morrey spaces $L^{p,\lambda}$:

(i) $L^{p,\lambda}(\Omega;\mathbb{R})$ are Banach spaces;

(ii) $L^{p,0}(\Omega;\mathbb{R}) = L^p(\Omega;\mathbb{R})$;

(iii) $L^{p,\lambda}(\Omega;\mathbb{R}) = \{0\}$ if $\lambda > n$;

(iv) $L^{p,n}(\Omega;\mathbb{R}) = L^\infty(\Omega;\mathbb{R})$;

(v) $L^{q,\mu}(\Omega;\mathbb{R}) \subset L^{p,\lambda}(\Omega;\mathbb{R})$ if Ω is bounded, $q \geq p$ and $(n-\lambda)/p \geq (n-\mu)/q$.

Note that the condition $(n-\lambda)/p \geq (n-\mu)/q$ can also be expressed by requiring $\lambda \leq \lambda_c$ with the critical value λ_c defined by the equation $(n-\lambda_c)/p = (n-\mu)/q$. The proof of the first result is standard, the second statement is trivial, while the third and fourth ones are straightforward applications of Lebesgue Differentiation Theorem (see Section B.3). Finally, the last one relies on the Hölder inequality:

$$\begin{aligned}\left(\int_{\Omega(x,r)} |f|^p \, dx\right) &\leq \left(\int_{\Omega(x,r)} |f|^q \, dx\right)^{\frac{p}{q}} (\omega_n r^n)^{1-\frac{p}{q}} \\ &= c\,\|f\|^p_{L^{q,\mu}}\, r^{\mu\frac{p}{q}+n\left(1-\frac{p}{q}\right)} = c\,\|f\|^p_{L^{q,\mu}}\, r^{\lambda_c},\end{aligned}$$

for some constant $c = c(n,p,q)$.

Definition 3.5 (Campanato spaces)**.** *Assume $\Omega \subset \mathbb{R}^n$ open, $\lambda \geq 0$, $1 \leq p < \infty$. We say that a function $f \in L^p(\Omega;\mathbb{R})$ belongs to the Campanato space $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})$ if*

$$\sup_{x_0\in\Omega,\ 0<r<d_\Omega} r^{-\lambda} \int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx < +\infty, \tag{3.4}$$

where, as before, d_Ω is the diameter of Ω and

$$f_{x_0,r} := \fint_{\Omega(x_0,r)} f(x)\, dx. \tag{3.5}$$

It is easy to verify that

$$[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} := \left(\sup_{x_0\in\Omega,\ 0<r<d_\Omega} r^{-\lambda} \int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx\right)^{1/p}$$

is a seminorm on $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})$.

The mean $f_{x_0,r}$ defined in (3.5) might not be optimal in the calculation of the sort of p-variance in (3.4), anyway it gives equivalent results, thanks to inequality (2.20).

Remark 3.6. As in Remark 3.4, we briefly highlight the main properties of Campanato spaces.

(i) As defined above, $\|\cdot\|_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}$ is merely a seminorm because constants have null $\mathcal{L}^{p,\lambda}$ norm. If Ω is connected, then $\mathcal{L}^{p,\lambda}$ modulo constants is a Banach space.

(ii) $\mathcal{L}^{q,\mu} \subset \mathcal{L}^{p,\lambda}$ when Ω is bounded, $q \geq p$ and $(n-\lambda)/p \geq (n-\mu)/q$.

(iii) $C^{0,\alpha} \subset \mathcal{L}^{p,n+\alpha p}$, because we can take the average in x' of the pointwise estimate $|f(x') - f(x)| \le [f]_{C^{0,\alpha}(\Omega;\mathbb{R})}(2r)^\alpha$, which holds true for every $x, x' \in B_r(x_0)$, thereby getting the same estimate with $f_{x_0,r}$ in lieu of $f(x')$. By integration on $\Omega(x_0, r)$ we get

$$\int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx \le [f]^p_{C^{0,\alpha}(\Omega)}(2r)^{\alpha p} \mathscr{L}^n \left(B_r(x_0)\right) = [f]^p_{C^{0,\alpha}(\Omega;\mathbb{R})} 2^{\alpha p} \omega_n r^{n+\alpha p} \ .$$

We will see that the converse statement holds (namely functions in these Campanato spaces have Hölder-continuous representatives in their Lebesgue equivalence class), and this turns out to be very useful since it allows replacing the pointwise definition of Hölder spaces with an integral one.

Actually, Campanato spaces are interesting only when $\lambda \ge n$, precisely because of their relationship with Hölder spaces. On the contrary, if $\lambda < n$, Morrey spaces and Campanato spaces are basically equivalent. In the proof of this and other results we need a mild regularity assumption on Ω, namely the existence of a constant $c_* > 0$ satisfying

$$\mathscr{L}^n \left(\Omega \cap B_r(x_0)\right) \ge c_* r^n \quad \forall x_0 \in \overline{\Omega}, \ \forall r \in (0, d_\Omega) \, . \tag{3.6}$$

For instance, this assumption is satisfied by bounded regular domains (or even bounded domains that are locally epigraphs of of Lipschitz functions), while it rules out domains with cusps.

Theorem 3.7. *Let $\Omega \subset \mathbb{R}^n$ be an open bounded set satisfying* (3.6) *and let $0 \le \lambda < n$. Then the spaces $L^{p,\lambda}(\Omega;\mathbb{R})$ and $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})$ are equivalent, namely*

$$\| \cdot \|_{L^{p,\lambda}(\Omega;\mathbb{R})} \simeq [\cdot]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} + \| \cdot \|_{L^p(\Omega;\mathbb{R})}.$$

Proof. Throughout the proof we let c denote a positive constant which is only allowed to depend on c_* in (3.6) and on n, p, λ, and which may vary from line to line or even within the same line.

Without using the hypothesis on λ, one can easily prove that $L^{p,\lambda} \subset \mathcal{L}^{p,\lambda}$: indeed Jensen's inequality ensures

$$\int_{\Omega(x_0,r)} |f_{x_0,r}|^p \, dx \le \int_{\Omega(x_0,r)} |f(x)|^p \, dx$$

and thus we can estimate

$$\begin{aligned} \int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx &\le 2^{p-1} \left(\int_{\Omega(x_0,r)} |f(x)|^p \, dx + \int_{\Omega(x_0,r)} |f_{x_0,r}|^p \, dx \right) \\ &\le 2^p \int_{\Omega(x_0,r)} |f(x)|^p \, dx \ . \end{aligned}$$

Conversely, we would like to bound $r^{-\lambda}\int_{\Omega(x_0,r)}|f(x)|^p\,dx$ in terms of $[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}+\|f\|_{L^p(\Omega;\mathbb{R})}$ for every $0<r<d_\Omega$ and every $x_0\in\Omega$. As a first step, by the triangle inequality we have

$$\int_{\Omega(x_0,r)}|f(x)|^p\,dx\le 2^{p-1}\int_{\Omega(x_0,r)}|f(x)-f_{x_0,r}|^p\,dx+cr^n|f_{x_0,r}|^p\le c\left(r^\lambda[f]^p_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}+r^n|f_{x_0,r}|^p\right),$$

so that we only need to obtain a suitable upper bound for the summand $|f_{x_0,r}|^p$.

To that goal, let us consider the following inequality involving means on concentric balls: when $x_0\in\Omega$ is fixed and $0<r<\rho<d_\Omega$, it holds

$$\begin{aligned}
c_*r^n|f_{x_0,r}-f_{x_0,\rho}|^p &\le \int_{\Omega(x_0,r)}|f_{x_0,r}-f_{x_0,\rho}|^p\,dx\\
&\le 2^{p-1}\left(\int_{\Omega(x_0,r)}|f_{x_0,r}-f(x)|^p\,dx+\int_{\Omega(x_0,r)}|f(x)-f_{x_0,\rho}|^p\,dx\right)\\
&\le 2^{p-1}[f]^p_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}\left(r^\lambda+\rho^\lambda\right)\le 2^p[f]^p_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}\rho^\lambda,
\end{aligned}$$

thus we obtained that

$$|f_{x_0,r}-f_{x_0,\rho}|\le c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}r^{-\frac{n}{p}}\rho^{\frac{\lambda}{p}}=c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}\left(\frac{\rho}{r}\right)^{\frac{n}{p}}\rho^{\frac{\lambda-n}{p}}. \tag{3.7}$$

Now fix a radius $R>0$: if $r=2^{-(k+1)}R$ and $\rho=2^{-k}R$, inequality (3.7) means that

$$|f_{x_0,R/2^{k+1}}-f_{x_0,R/2^k}|\le c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}\left(\frac{R}{2^k}\right)^{\frac{\lambda-n}{p}}, \tag{3.8}$$

and, adding up these summands for $k=0,\dots,N-1$, we get that

$$|f_{x_0,R/2^N}-f_{x_0,R}|\le c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}R^{\frac{\lambda-n}{p}}\left(\frac{2^{N\frac{n-\lambda}{p}}-1}{2^{\frac{n-\lambda}{p}}-1}\right)\le c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}\left(\frac{R}{2^N}\right)^{\frac{\lambda-n}{p}} \tag{3.9}$$

thanks to the fact that $\lambda<n$.

Let us go back to our purpose of estimating $|f_{x_0,r}|^p$: we choose $R\in(d_\Omega/2,d_\Omega)$ and $N\in\mathbb{N}$ such that $r=R/2^N$. Again, by the triangle inequality

$$|f_{x_0,r}|^p\le 2^{p-1}\left(|f_{x_0,r}-f_{x_0,R}|^p+|f_{x_0,R}|^p\right)$$

and since

$$|f_{x_0,R}|^p\le cd_\Omega^{-n}\|f\|^p_{L^p(\Omega;\mathbb{R})},$$

the only thing left to conclude is to apply inequality (3.9) in this case:

$$|f_{x_0,r}-f_{x_0,R}|^p\le c[f]^p_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}r^{\lambda-n}.$$

Therefore we can conclude that

$$r^{-\lambda}\int_{\Omega(x_0,r)}|f|^p \le c\left([f]^p_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})}+d_\Omega^{-\lambda}\|f\|^p_{L^p(\Omega;\mathbb{R})}\right).$$

□

Remark 3.8. When the dimension of the domain is n, the Campanato space $\mathcal{L}^{1,n}$ is very important in harmonic analysis and elliptic regularity theory: after John-Nirenberg seminal paper [61], it is customary to denote this space BMO (for *bounded mean oscillation*). It consists of all functions $f:\Omega\to\mathbb{R}$ such that there exists a constant $c>0$ for which the inequality

$$\int_{\Omega(x_0,r)}|f(x)-f_{x_0,r}|\,dx\le cr^n\quad\forall\, r\in(0,d_\Omega),\ \forall\, x_0\in\Omega$$

is satisfied. Notice that $L^\infty(\Omega;\mathbb{R})\subsetneq BMO(\Omega;\mathbb{R})$ as one can see, for example, by considering $\Omega=(0,1)$ and $f(x)=\log x$. For any a, $r>0$ it is easy to check that

$$\int_a^{a+r}|\log t-\log(a+r)|\,dt=\int_a^{a+r}(\log(a+r)-\log t)\,dt=r+a\log\left(\frac{a}{a+r}\right)\le r,$$

viewing the interval $(a,a+r)$ as the one-dimensional ball of center $a+r/2$ and radius $r/2$. Hence, we conclude $\log x\in BMO(\Omega;\mathbb{R})$. Notice that in this calculation we replaced the mean $\fint_a^{a+r}\log s\,ds$ with $\log(a+r)$, but, up to a multiplicative factor 2, this is irrelevant with respect to the conclusion. On the contrary, it is obvious that $\log x\notin L^\infty(\Omega;\mathbb{R})$ so the inclusion in question is indeed strict.

Theorem 3.9 (Campanato, [17])**.** *With the notation above, when $n<\lambda\le n+p$ the Campanato spaces $\mathcal{L}^{p,\lambda}$ are equivalent to the Hölder spaces $C^{0,\alpha}$ with $\alpha=(\lambda-n)/p$. Moreover, if Ω is connected and $\lambda>n+p$, then $\mathcal{L}^{p,\lambda}$ is equivalent to the set of constants.*

Proof. As in the proof of Theorem 3.7, the letter c denotes a generic constant depending on the exponents, the space has dimension n and the constant c_* satisfies (3.6).

Let $\lambda=n+\alpha p$. We already observed in Remark 3.6 that $C^{0,\alpha}\subset\mathcal{L}^{p,\lambda}$, so we need to prove the converse inclusion: given a function $f\in\mathcal{L}^{p,\lambda}$, we are looking for a representative $\tilde f$ in the Lebesgue equivalence class of f which belongs to $C^{0,\alpha}$.

Recalling inequality (3.8) with fixed radius $R>0$ and $x\in\Omega$, we obtain that every sequence of the form $(f_{x,R/2^k})$ has the Cauchy property. Hence, we define

$$\tilde f(x):=\lim_{k\to\infty}\fint_{\Omega(x,R/2^k)}f(y)\,dy.$$

Clearly, by the very definition of $\tilde{f}(x)$ we have

$$\fint_{\Omega(x,R/2^k)} |f(y) - f_{x,R/2^k}|^p \, dy \longrightarrow 0 \quad \Longrightarrow \quad \fint_{\Omega(x,R/2^k)} |f(y) - \tilde{f}(x)|^p \, dy \longrightarrow 0, \tag{3.10}$$

but since $c_* r^n \leq \mathscr{L}^n(\Omega(x,r)) \leq \omega_n r^n$, for $r \in \big(R/2^{k+1}, R/2^k\big)$ we have

$$\fint_{\Omega(x,r)} |f(y) - \tilde{f}(x)|^p \, dy \leq \frac{2^n \omega_n}{c_*} \fint_{\Omega(x,R/2^k)} |f(y) - \tilde{f}(x)|^p \, dy,$$

so that (3.10) implies that

$$\fint_{\Omega(x,r)} |f(y) - \tilde{f}(x)|^p \, dy \longrightarrow 0 \qquad \text{as } r \downarrow 0\,.$$

In particular, notice that $\tilde{f}$ does not depend on the chosen initial radius R. Let us prove that

$$\tilde{f} \in C^{0,\alpha}(\Omega;\mathbb{R}).$$

We employ again an inequality from the proof of Theorem 3.7: letting $N \to \infty$ in the first inequality of (3.9), we get that

$$|\tilde{f}(x) - f_{x,R}| \leq c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} R^\alpha$$

with $\alpha = (\lambda - n)/p$; as a result, given $x, y \in \Omega$ and choosing $R = 2|x-y|$,

$$|\tilde{f}(x) - \tilde{f}(y)| \leq |\tilde{f}(x) - f_{x,R}| + |f_{x,R} - f_{y,R}| + |f_{y,R} - \tilde{f}(y)| \leq c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} |x-y|^\alpha + |f_{x,R} - f_{y,R}|.$$

Therefore, the theorem will be proved once we estimate $|f_{x,R} - f_{y,R}|$. To this aim, we simply exploit the inclusion $\Omega(y, R/2) \subset \Omega(x,R)$ and the very definition of $\mathcal{L}^{p,\lambda}$ to get

$$\begin{aligned} c_* 2^{-n} R^n |f_{x,R} - f_{y,R}|^p &\leq \int_{\Omega(y,R/2)} |f_{x,R} - f_{y,R}|^p \, dz \\ &\leq 2^{p-1} \left(\int_{\Omega(x,R)} |f(z) - f_{x,R}|^p \, dz + \int_{\Omega(y,R)} |f(z) - f_{y,R}|^p \, dz \right) \\ &\leq 2^p [f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} R^\lambda. \end{aligned}$$

It follows that

$$|f_{x,R} - f_{y,R}| \leq c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} R^{\frac{\lambda-n}{p}} \leq c[f]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R})} |x-y|^\alpha$$

which implies the claim.

The second assertion in the statement of the theorem is a well-known fact about Hölder functions with exponent larger than one, and therefore the proof is complete. □

The following inclusions readily follow by the Hölder and the Poincaré inequalities, respectively.

Proposition 3.10 (Inclusions between Lebesgue and Morrey spaces, Morrey and Campanato spaces)**.** *For all $p \in (1,\infty)$, $L^p_{\rm loc}(\Omega;\mathbb{R}) \subset L^{1,\frac{n}{p'}}_{\rm loc}(\Omega;\mathbb{R})$. In addition,*

$$|\nabla u| \in L^{p,\lambda}_{\rm loc}(\Omega;\mathbb{R}) \quad \Longrightarrow \quad u \in \mathcal{L}^{p,\lambda+p}_{\rm loc}(\Omega;\mathbb{R}) \ . \tag{3.11}$$

Corollary 3.11 (Sobolev embedding for $p > n$)**.** *If $p > n$, then $W^{1,p}_{\rm loc}(\Omega;\mathbb{R}) \subset C^{0,\alpha}_{\rm loc}(\Omega;\mathbb{R})$, with $\alpha = 1 - n/p$. If Ω is bounded and regular, then $W^{1,p}(\Omega;\mathbb{R}) \subset C^{0,\alpha}(\Omega;\mathbb{R})$.*

Proof. By the previous proposition we get

$$u \in W^{1,p}_{\rm loc}(\Omega;\mathbb{R}) \quad \Longrightarrow \quad |\nabla u| \in L^p_{\rm loc}(\Omega;\mathbb{R}) \quad \Longrightarrow \quad \mathcal{L}^{p,p}_{\rm loc}(\Omega;\mathbb{R}). \tag{3.12}$$

At that stage, by virtue of Theorem 3.9 we actually obtain $u \in C^{0,\alpha}_{\rm loc}(\Omega;\mathbb{R})$ for $\alpha = 1-n/p$ as in the statement above. If Ω is bounded and regular we apply this inclusion to a $W^{1,p}$ extension of u to obtain the global $C^{0,\alpha}$ regularity. □

3.2 Hilbert's *XIX* problem and its solution in the two-dimensional case

Let $\Omega \subset \mathbb{R}^n$ open, let $L \in C^2(\Omega;\mathbb{R}^{m\times n})$ and let us consider a local minimizer u of the functional

$$v \mapsto \int_\Omega L(\nabla v)\, dx \tag{3.13}$$

in the sense discussed in Section 1.4. We assume that $\nabla^2 L(p)$ satisfies the Legendre condition (1.16) with $\lambda > 0$ independent of p and is uniformly bounded.

In this case, we have seen that u satisfies the Euler-Lagrange equations, for (3.13) they are

$$\sum_\alpha \partial_{x_\alpha}\left(\partial_{p_i^\alpha} L(\nabla u)\right) = 0 \qquad i = 1,\dots,m\,. \tag{3.14}$$

We have also seen in the introductory discussion of Section 3 how, differentiating (3.14) along the direction x_γ, one can obtain

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(\partial_{p_i^\alpha}\partial_{p_j^\beta} L(\nabla u)\partial_{x_\beta x_\gamma} u^j\right) = 0 \qquad i = 1,\dots,m\,. \tag{3.15}$$

In the spirit of Hilbert's *XIX* problem, we are interested in the regularity properties of u. Fix $\gamma \in \{1,\dots,n\}$ and let us set

$$\begin{aligned} w(x) &:= \partial_{x_\gamma} u(x) \in L^2(\Omega;\mathbb{R}^m)\ , \\ A(x) &:= \nabla^2 L(\nabla u(x))\ , \end{aligned}$$

thus (3.15) can be written as an elliptic problem in divergence form, namely

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha} \left(A_{ij}^{\alpha\beta}(x) \partial_{x_\beta} w^j \right) = 0 \qquad i = 1, \dots, m\,. \tag{3.16}$$

Since $w \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ by virtue of our earlier discussion, we can use the Caccioppoli-Leray inequality for w, in the sharp version of Remark 2.4. As explained at that stage, combining it with the Poincaré inequality (choosing k equal to the mean value of w on the ball $B_R(x_0) \setminus B_{R/2}(x_0)$), we obtain

$$\int_{B_{R/2}(x_0)} |\nabla w|^2 \, dx \le cR^{-2} \int_{B_R(x_0)\setminus B_{R/2}(x_0)} |w-k|^2 \, dx \le c \int_{B_R(x_0)\setminus B_{R/2}(x_0)} |\nabla w|^2 \, dx,$$

thus, adding $c \int_{B_{R/2}(x_0)} |\nabla w|^2 \, dx$ to both sides,one gets

$$\int_{B_{R/2}(x_0)} |\nabla w|^2 \, dx \le \frac{c}{c+1} \int_{B_R(x_0)} |\nabla w|^2 \, dx.$$

Now, if $\theta := \frac{c}{c+1} < 1$ and $\alpha = -\log_2 \theta$, we can write the previous inequality as

$$\int_{B_{R/2}(x_0)} |\nabla w|^2 \, dx \le \left(\frac{1}{2}\right)^\alpha \int_{B_R(x_0)} |\nabla w|^2 \, dx. \tag{3.17}$$

In order to get a power decay inequality from (3.17), we state this basic iteration lemma.

Lemma 3.12. *Consider a non-decreasing function $f : (0, R_0] \to [0, +\infty)$ satisfying*

$$f\left(\frac{\rho}{2}\right) \le \left(\frac{1}{2}\right)^\alpha f(\rho) \qquad \forall\, \rho \le R_0$$

for some $\alpha > 0$. Then

$$f(r) \le 2^\alpha \left(\frac{r}{R}\right)^\alpha f(R) \qquad \forall\, 0 < r \le R \le R_0.$$

Proof. Fix $r < R \le R_0$ and choose a number $N \in \mathbb{N}$ such that

$$\frac{R}{2^{N+1}} < r \le \frac{R}{2^N}.$$

It is clear from the iteration of the hypothesis that

$$f\left(\frac{R}{2^N}\right) \le \left(\frac{1}{2}\right)^{\alpha N} f(R),$$

thus, by monotonicity,

$$f(r) \le f\left(2^{-N} R\right) \le 2^{-\alpha N} f(R) = 2^\alpha 2^{-\alpha(N+1)} f(R) < 2^\alpha (r/R)^\alpha f(R).$$

□

Thanks to Lemma 3.12, we can derive from (3.17) the decay estimate

$$\int_{B_\rho(x_0)} |\nabla w|^2\,dx \le c\left(\frac{\rho}{R}\right)^\alpha \int_{B_R(x_0)} |\nabla w|^2\,dx \qquad \forall\, 0<\rho\le R\ ,$$

therefore $|\nabla w| \in L^{2,\alpha}_{\rm loc}(\Omega;\mathbb{R})$. So, invoking Proposition 3.10, this gives $w \in \mathcal{L}^{2,\alpha+2}_{\rm loc}(\Omega;\mathbb{R}^m)$. All these facts are true in any number n of spatial dimensions, but when $n=2$ we can apply Theorem 3.9 to get

$$w \in C^{0,\alpha/2}_{\rm loc}(\Omega;\mathbb{R}^m).$$

Since γ is arbitrary, it follows that $u \in C^{1,\alpha/2}_{\rm loc}(\Omega;\mathbb{R}^m)$ and $A = \nabla^2 L(\nabla u) \in C^{0,\alpha/2}_{\rm loc}(\Omega;\mathbb{R}^{m^2\times n^2})$.

The Schauder theory that we will consider in the next section will allow us to conclude that

$$u \in C^{2,\alpha/2}_{\rm loc}(\Omega;\mathbb{R}^m).$$

More specifically, this conclusion is gained by simply applying Theorem 3.17 to each function of the form $\partial_{x_\gamma} u$, which solves equation (3.16)). As long as L is sufficiently regular, the iteration of this argument solves Hilbert's *XIX* problem in the C^∞ category.

We close this section with a more technical but useful iteration lemma in the same spirit of Lemma 3.12.

Lemma 3.13. *Consider a non-decreasing function $f:(0,R_0]\to[0,+\infty)$ which satisfies for some coefficients $a>0$, $b\ge 0$ and $\varepsilon\in[0,1)$ and exponents $\alpha>\beta\ge 0$ the following inequality*

$$f(\rho) \le a\left[\left(\frac{\rho}{R}\right)^\alpha + \varepsilon\right] f(R) + bR^\beta \qquad \forall\, 0<\rho\le R\le R_0. \tag{3.18}$$

If

$$\varepsilon \le \left(\frac{1}{2a}\right)^{\frac{\alpha}{\alpha-\gamma}} \tag{3.19}$$

for some $\gamma\in(\beta,\alpha)$, then there exists a non-negative constant $c=c(\alpha,\beta,\gamma,a)$

$$f(\rho) \le c\left[\left(\frac{\rho}{R}\right)^\gamma f(R) + b\rho^\beta\right] \qquad \forall\, 0<\rho\le R\le R_0. \tag{3.20}$$

Proof. Without loss of generality, since $\varepsilon\in(0,1)$ we assume $a>1/2$. We choose $\tau\in(0,1)$ such that

$$2a\tau^\alpha = \tau^\gamma, \tag{3.21}$$

thus (3.19) is equivalent to the inequality

$$\varepsilon \le \tau^\alpha. \tag{3.22}$$

The following basic estimate uses the hypothesis (3.18) jointly with (3.21) and (3.22):

$$\begin{aligned} f(\tau R) &\leq a(\tau^\alpha + \varepsilon) f(R) + bR^\beta \\ &\leq 2a\tau^\alpha f(R) + bR^\beta = \tau^\gamma f(R) + bR^\beta. \end{aligned} \tag{3.23}$$

The iteration of (3.23) easily gives

$$\begin{aligned} f(\tau^2 R) \leq \tau^\gamma f(\tau R) + b\tau^\beta R^\beta &\leq \tau^{2\gamma} f(R) + \tau^\gamma b R^\beta + b\tau^\beta R^\beta \\ &= \tau^{2\gamma} f(R) + bR^\beta \tau^\beta (1 + \tau^{\gamma-\beta}) \ . \end{aligned}$$

Thus, it can be easily proven by induction that

$$f(\tau^N R) \leq \tau^{N\gamma} f(R) + bR^\beta \tau^{(N-1)\beta} \sum_{k=0}^{N-1} \tau^{k(\gamma-\beta)} = \tau^{N\gamma} f(R) + bR^\beta \tau^{(N-1)\beta} \left(\frac{1 - \tau^{N(\gamma-\beta)}}{1 - \tau^{(\gamma-\beta)}} \right).$$

So, given $0 < \rho \leq R \leq R_0$, if N satisfies

$$\tau^{N+1} R < \rho \leq \tau^N R,$$

we can conclude our proof by simply choosing the constant $c = c(\alpha, \beta, \gamma, a)$ in such a way that the last line in the following chain of inequalities holds:

$$\begin{aligned} f(\rho) &\leq f(\tau^N R) \leq \tau^{N\gamma} f(R) + \frac{bR^\beta \tau^{(N-1)\beta}}{1 - \tau^{(\gamma-\beta)}} \\ &= \tau^{-\gamma} \left(\tau^{(N+1)\gamma} f(R) \right) + \frac{\tau^{-2\beta}}{1 - \tau^{(\gamma-\beta)}} \left(bR^\beta \tau^{(N+1)\beta} \right) \\ &< \tau^{-\gamma} \left(\left(\frac{\rho}{R} \right)^\gamma f(R) \right) + \frac{\tau^{-2\beta}}{1 - \tau^{(\gamma-\beta)}} \left(b\rho^\beta \right) \\ &\leq c \left(\left(\frac{\rho}{R} \right)^\gamma f(R) + b\rho^\beta \right). \end{aligned}$$

□

Remark 3.14. The fundamental gain in Lemma 3.13 is the passage from R^β to ρ^β and the loss of the additive term involving ε, provided that ε is small enough. These improvements can be obtained at the price of passing from the power α to the (marginally worse) power $\gamma < \alpha$.

3.3 Back to linear problems: Schauder theory

For the sake of simplicity, we present Schauder theory in a local form, namely with the specific goal of obtaining interior estimates for solutions to suitable second-order elliptic

equations. We refer the reader to our references, for instance [45] for a treatment of the boundary (and thus global) estimates, which rely on some ideas that are analogous to those used in Section 2.5). We shall first describe a model result for constant coefficients operators, and then we will consider the case of Hölder-continuous coefficients.

We consider the problem, in divergence form, given by

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(A_{ij}^{\alpha\beta}\partial_{x_\beta}u^j\right) = -\sum_{\alpha}\partial_{x_\alpha}F_i^\alpha \qquad i=1,\dots,m \tag{3.24}$$

to be solved for $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$, for data $F \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$. We prove the following local regularity result:

Theorem 3.15. *If $A_{ij}^{\alpha\beta}$ are constant and satisfy the Legendre-Hadamard condition for some $\lambda > 0$, then for all $\mu < n+2$ it holds*

$$F \in \mathcal{L}^{2,\mu}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}) \qquad \Longrightarrow \qquad \nabla u \in \mathcal{L}^{2,\mu}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}).$$

In particular, if $\mu > n$ and $\alpha = (\mu-n)/2$, then, by Theorem 3.9, one has

$$F \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}) \quad \Longrightarrow \quad \nabla u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}).$$

Proof. In this proof, we let $c = c(n,\lambda,|A|)$ denote a positive constant whose value can change from line to line and even within the same line. Let us fix a ball $B_R(x_0) \Subset \Omega$ and compare u with the solution v of the homogeneous problem

$$\begin{cases} \sum_{\alpha,\beta,j}\partial_{x_\alpha}\left(A_{ij}^{\alpha\beta}\partial_{x_\beta}v^j\right) = 0 & \text{in } B_R(x_0) \\ v = u & \text{on } \partial B_R(x_0) \end{cases} \tag{3.25}$$

where the boundary condition is understood, as discussed in Subsection 1.2, in terms of trace operator.

Notice that $v \in H^2_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ by virtue of interior H^2 estimates for u. As a result, since ∇v belongs to $H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$ and its components $\partial_{x_\gamma}v$ solve the same problem (because we have supposed to have constant coefficients), we can use the decay estimate (2.17) for balls with radii $0 < r < R' < R$ and then pass to the limit $R' \uparrow R$. Hence, if $0 < \rho < R$, (2.17) provides the following inequality:

$$\int_{B_\rho(x_0)} |\nabla v(x) - (\nabla v)_{x_0,\rho}|^2\,dx \le c\left(\frac{\rho}{R}\right)^{n+2}\int_{B_R(x_0)} |\nabla v(x) - (\nabla v)_{x_0,R}|^2\,dx. \tag{3.26}$$

Now we try to employ (3.26) to get some estimate for u, the original solution of the inhomogeneous problem (3.24). We can write $u = w + v$, where $w \in H^1_0(B_R(x_0);\mathbb{R}^m)$, and

using $\nabla u = \nabla v + \nabla w$, then the variance-minimizing property of the mean and (3.26), eventually $\nabla v = \nabla u - \nabla w$ and $(\nabla w)_{x_0,R} = 0$, we can write

$$\begin{aligned}
&\int_{B_\rho(x_0)} |\nabla u(x) - (\nabla u)_{x_0,\rho}|^2 \, dx \\
\leq\ & 2\left(\int_{B_\rho(x_0)} |\nabla w(x) - (\nabla w)_{x_0,\rho}|^2 \, dx + \int_{B_\rho(x_0)} |\nabla v(x) - (\nabla v)_{x_0,\rho}|^2 \, dx\right) \\
\leq\ & 2\int_{B_\rho(x_0)} |\nabla w(x) - (\nabla w)_{x_0,R}|^2 \, dx + c\left(\frac{\rho}{R}\right)^{n+2} \int_{B_R(x_0)} |\nabla v(x) - (\nabla v)_{x_0,R}|^2 \, dx \\
\leq\ & c\int_{B_R(x_0)} |\nabla w(x)|^2 \, dx + c\left(\frac{\rho}{R}\right)^{n+2} \int_{B_R(x_0)} |\nabla u(x) - (\nabla u)_{x_0,R}|^2 \, dx.
\end{aligned}$$

The auxiliary function

$$f(\rho) := \int_{B_\rho(x_0)} |\nabla u(x) - (\nabla u)_{x_0,\rho}|^2 \, dx$$

is non-decreasing because of the minimality property of the mean $(\nabla u)_{x_0,\rho}$, when one minimizes $m \mapsto \int_{B_\rho(x_0)} |\nabla u(x) - m|^2 \, dx$. In order to check that f satisfies the hypothesis of Lemma 3.13, we have to estimate $\int_{B_R(x_0)} |\nabla w|^2 \, dx$. We can consider w as a function in $H^1(\mathbb{R}^n; \mathbb{R}^m)$ (null out of $B_R(x_0)$) so, by Gårding inequality (Theorem 1.6) we obtain

$$\begin{aligned}
\int_{B_R(x_0)} |\nabla w(x)|^2 \, dx &\leq c\int_{B_R(x_0)} \langle A\nabla w(x), \nabla w(x)\rangle \, dx \\
&= c\int_{B_R(x_0)} \langle F(x), \nabla w(x)\rangle \, dx \\
&= c\int_{B_R(x_0)} \langle F(x) - F_{x_0,R}, \nabla w(x)\rangle \, dx. \qquad (3.27)
\end{aligned}$$

Applying Young's inequality to (3.27) and then absorbing a small multiple of $\int_{B_R(x_0)} |\nabla w|^2 \, dx$ in the left-hand side of (3.27), we get

$$\int_{B_R(x_0)} |\nabla w(x)|^2 \, dx \leq c\int_{B_R(x_0)} |F(x) - F_{x_0,R}|^2 \, dx \leq c\|F\|_{\mathcal{L}^{2,\mu}(B_{R_0}(x_0);\mathbb{R}^{m\times n})} R^\mu,$$

where we have fixed a larger reference ball $B_{R_0}(x_0) \Subset \Omega$ and we only consider $R \leq R_0$.

Therefore we obtained the decay inequality of Lemma 3.13 for f with $\alpha = n+2$, $\beta = \mu$ and $\varepsilon = 0$, then

$$f(\rho) \leq c\left(\frac{\rho}{R}\right)^\mu f(R) + c\rho^\mu,$$

that is $\nabla u \in \mathcal{L}^{2,\mu}_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n})$. □

In the next theorem we consider the case of variable, but continuous, coefficients, proving in this case a $L^{2,\mu}$ regularity of $|\nabla u|$ with $\mu < n$; as we have seen, at that stage the Poincaré inequality then provides Hölder regularity of u if $\mu + 2 > n$.

Theorem 3.16. *Considering again* (3.24)*, suppose that the coefficients* $A_{ij}^{\alpha\beta}$ *are continuous in* Ω *and* A *satisfies a (locally) uniform Legendre-Hadamard condition for some* $\lambda > 0$*. If* $F \in L^{2,\mu}_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n})$ *with* $\mu < n$*, then* $|\nabla u| \in L^{2,\mu}_{\mathrm{loc}}(\Omega; \mathbb{R})$*.*

Since $\mu < n$, Campanato spaces and Morrey spaces coincide, so that we decided to phrase the result above using Morrey spaces for the sole sake of simplicity.

Proof. This proof relies on Korn's technique, whose basic idea is that of *freezing the coefficients* in a sense that we are about to explain. We use the same convention on c of the previous proof, namely $c = c(n, \lambda, \sup |A|)$ and allowed to vary in each inequality. Fix a point $x_0 \in \Omega$ and define

$$\tilde{F}(x) := F(x) + (A(x_0) - A(x))\, \nabla u(x),$$

so that the solution u of (3.24) also solves

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}(A_{ij}^{\alpha\beta}(x_0)\partial_{x_\beta} u^j(x)) = -\sum_{\alpha} \partial_{x_\alpha} \tilde{F}_i^{\alpha}(x) \qquad i = 1, \dots, m\,.$$

Write $u = v + w$, where v solves the homogeneous equation (3.25) in $B_R(x_0)$ and with the condition $v = u$ on $\partial B_R(x_0)$ with frozen coefficients given by $A(x_0)$. Using (2.16) for ∇v and arguing as in the previous proof we obtain

$$\begin{aligned}\int_{B_\rho(x_0)} |\nabla u(x)|^2\, dx &\le c\left(\frac{\rho}{R}\right)^n \int_{B_R(x_0)} |\nabla v(x)|^2\, dx + c\int_{B_R(x_0)} |\nabla w(x)|^2\, dx \\ &\le c\left(\frac{\rho}{R}\right)^n \int_{B_R(x_0)} |\nabla u(x)|^2\, dx + c\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2\, dx.\end{aligned}$$

Thanks to the continuity property of A, there exists a (local) modulus of continuity ω in whose terms we can write the estimate

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2\, dx \le 2\int_{B_R(x_0)} |F(x) - F_{x_0,R}|^2\, dx + 2\omega^2(R)\int_{B_R(x_0)} |\nabla u(x)|^2\, dx. \tag{3.28}$$

As a result, since $F \in L^{2,\mu}_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n})$,

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2\, dx \le 2\|F\|_{L^{2,\mu}(B_R(x_0);\mathbb{R}^{m\times n})} R^\mu + 2\omega^2(R)\int_{B_R(x_0)} |\nabla u(x)|^2\, dx.$$

We are then ready to use Lemma 3.13, for a suitable choice of a and b, with

$$f(\rho) := \int_{B_\rho(x_0)} |\nabla u(x)|^2 \, dx, \ \alpha = n, \ \beta = \mu < n \text{ and } \varepsilon = \omega^2(R)/a.$$

Thereby, we can conclude that if R is under a threshold depending only on c, α, β and ω we have

$$f(\rho) \le c \left(\frac{\rho}{R}\right)^\mu f(R) + c\rho^\mu,$$

which means $|\nabla u| \in L^{2,\mu}_{\mathrm{loc}}(\Omega;\mathbb{R})$, as desired. □

Using this preliminary result, we can now prove Theorem 3.17 Schauder's estimate for solutions of elliptic problems in divergence form. Let us also remark that for equations in non-divergence form the corresponding result is (restricting for simplicity to the scalar case)

$$\sum_{\alpha,\beta} A^{\alpha\beta}(x) \frac{\partial^2 u}{\partial x_\alpha \partial x_\beta} \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}) \qquad \Longrightarrow \qquad u \in C^{2,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}), \tag{3.29}$$

if A is of class $C^{0,\alpha}$ (see, for instance, [45, Section 6.1]).

The proof of Theorem 3.17 follows along similar lines, i.e. starting with second derivative decay estimates for constant coefficients operators, and then freezing the coefficients. Notice also that both (3.29) and the conclusion of Theorem 3.17 below are easily seen to be optimal, considering 1-dimensional ordinary differential equations of the form $au'' = f$ or $(au')' = f'$.

Theorem 3.17 (Schauder, [85])**.** *Suppose that the coefficients $A^{\alpha\beta}_{ij}(x)$ of equation* (3.24) *belong to $C^{0,\alpha}(\Omega;\mathbb{R})$ and A satisfies a (locally) uniform Legendre-Hadamard in Ω for some $\lambda > 0$. Then the following implication holds*

$$F \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}) \quad \Longrightarrow \quad \nabla u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}),$$

that is to say

$$F \in \mathcal{L}^{2,n+2\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}) \quad \Longrightarrow \quad \nabla u \in \mathcal{L}^{2,n+2\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n}).$$

Proof. With the same idea of freezing coefficients (and the same notation, too), we estimate by (2.17)

$$\int_{B_\rho(x_0)} |\nabla u(x) - (\nabla u)_{x_0,\rho}|^2 \, dx \le c \left(\frac{\rho}{R}\right)^{n+2} \int_{B_R(x_0)} |\nabla u(x) - (\nabla u)_{x_0,R}|^2 \, dx + c \int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2 \, dx. \tag{3.30}$$

Additionally, the Hölder property of A allows to rewrite (3.28) as

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2 \, dx \leq 2 \int_{B_R(x_0)} |F(x) - F_{x_0,R}|^2 \, dx + cR^{2\alpha} \int_{B_R(x_0)} |\nabla u(x)|^2 \, dx. \quad (3.31)$$

Since $F \in C^{0,\alpha}_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n})$, we obtain

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2 \, dx \leq cR^{n+2\alpha} + cR^{2\alpha} \int_{B_R(x_0)} |\nabla u(x)|^2 \, dx.$$

Theorem 3.16 applied with $\mu = n - \alpha < n$ tells us that $|\nabla u| \in L^{2,\mu}(\Omega; \mathbb{R})$, thus

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2 \, dx \leq cR^{n+2\alpha} + cR^{n+\alpha}. \quad (3.32)$$

Combining (3.32) with (3.30) and applying Lemma 3.13 with exponents $n+2$ and $n+\alpha$, we get $\nabla u \in \mathcal{L}^{2,n+\alpha}(\Omega; \mathbb{R}^{m\times n})$, so that $\nabla u \in C^{0,\alpha/2}(\Omega; \mathbb{R}^{m\times n})$, in particular $|\nabla u|$ is locally bounded. Using this information we can improve (3.32) as follows:

$$\int_{B_R(x_0)} |\tilde{F}(x) - \tilde{F}_{x_0,R}|^2 \, dx \leq cR^{n+2\alpha}.$$

Now we reach the conclusion, again by Lemma 3.13 with exponents $n+2$ and $n+2\alpha$. □

3.4 The space of BMO functions

Given a cube $Q \subset \mathbb{R}^n$, we define

$$BMO(Q; \mathbb{R}) := \left\{ u \in L^1(Q; \mathbb{R}) \ : \ \sup_{Q' \subset Q} \fint_{Q'} |u - u_{Q'}| \, dx < +\infty \right\},$$

where $u_{Q'}$ denotes the mean value of u on Q' and Q' varies in the class of all open cubes contained in Q whose sides are parallel to those of Q. We also define the seminorm $\|u\|_{BMO}$ as the supremum above. An elementary argument replacing balls with concentric cubes shows that $BMO \sim \mathcal{L}^{1,n}$, that is to say: the two spaces consist of the same elements and the corresponding semi-norms are equivalent, see also Theorem 3.26 below. Here we recall a special case of the inclusion that was proven in Proposition 3.10.

Theorem 3.18. *For any cube $Q \subset \mathbb{R}^n$ the following inclusion holds:*

$$W^{1,n}(Q; \mathbb{R}) \hookrightarrow BMO(Q; \mathbb{R}).$$

Proof. First, we remind the reader that $W^{1,n}(Q;\mathbb{R})$ continuously embed into the space $\{u:\ |\nabla u| \in L^{1,n-1}(Q)\}$, as an immediate consequence of the Hölder inequality. Then, by the Poincaré inequality, there exists a dimensional constant $c = c(n) > 0$ such that for any square $Q' \subset Q$ with sides of length h

$$\int_{Q'} |u - u_{Q'}|\ dx \le ch \int_{Q'} |\nabla u|\ dx \le c\,\|\nabla u\|_{L^{1,n-1}}\, h^n$$

and thus the conclusion readily follows. □

However, it should be clear that the previous inclusion is far from being an equality as elementary examples show. In that respect, we shall now extend to functions of n variables the example presented in Remark 3.8. To that scope, we first state a simple sufficient (and necessary, as we will see) condition for a function to belong to BMO.

Proposition 3.19. *Let $u : Q \to \mathbb{R}$ be a measurable function such that, for some $b > 0$, $c \ge 0$, the following property holds:*

$$\forall\, C \subset Q \text{ cube}, \quad \exists\, a_C \in \mathbb{R} \text{ s.t. } \mathscr{L}^n\big(C \cap \{|u - a_C| > \sigma\}\big) \le c e^{-b\sigma} \mathscr{L}^n(C)\ \ \forall \sigma \ge 0. \tag{3.33}$$

Then $u \in BMO(Q;\mathbb{R})$.

The proof of the proposition above is straightforward, since

$$\frac{1}{2}\int_C |u - u_C|\, dx \le \int_C |u - a_C|\, dx = \int_0^\infty \mathscr{L}^n\big(C \cap \{|u - a_C| > \sigma\}\big)\, d\sigma \le \frac{c}{b}\mathscr{L}^n(C).$$

Example 3.20. Thanks to Proposition 3.19 we can check that $\log|x| \in BMO\big((0,1)^n;\mathbb{R}\big)$. Indeed, $\log|x|$ satisfies (3.33) (the choice of the parameters b and c will be specified later). To see this, fix a cube $C \subset (0,1)^n$, and let h denote the length of its side. Then, we define

$$\xi := \max_{x\in C} |x|, \qquad \eta := \min_{x\in C} |x|, \qquad a_C := \log \xi,$$

so that

$$a_C - u = \log\left(\frac{\xi}{|x|}\right) \ge 0.$$

We estimate the Lebesgue measure of $C \cap \{\xi > |x|e^\sigma\} = C \cap \{a_C - u > \sigma\}$: naturally we can assume that $\xi \ge \eta e^\sigma$, otherwise there is nothing to prove, so

$$\xi e^{-\sigma} \ge \eta \ge \xi - \operatorname{diam}(C) \ge \xi - \sqrt{n}h,$$

then it follows that

$$\xi \le \frac{\sqrt{n}h}{1 - e^{-\sigma}}.$$

Finally

$$\frac{1}{h^n}\mathscr{L}^n\left(C\cap\{|u-a_C|>\sigma\}\right)\le\frac{1}{h^n}\mathscr{L}^n\big(B_{\xi e^{-\sigma}}\big)\le\frac{(\sqrt{n})^n\omega_n}{(1-e^{-\sigma})^n}e^{-n\sigma},$$

so that distinguishing the cases $\sigma\le 1$ and $\sigma>1$ we see that (3.33) holds with $b=n$ and $c=\max\{e^n,(\sqrt{n})^n\omega_n\left(1-e^{-1}\right)^{-n}\}$.

The following theorem by John and Nirenberg was first proved in [61].

Theorem 3.21 (John-Nirenberg (first version), [61])**.** *There exist constants c_1, c_2 depending only on the dimension n such that*

$$\mathscr{L}^n\left(\{|u-u_Q|>t\}\right)\le c_1e^{-c_2t/\|u\|_{BMO}}\mathscr{L}^n(Q)\qquad\forall u\in BMO(Q;\mathbb{R})\setminus\{0\}\,. \tag{3.34}$$

Remark 3.22. In the argument we present here, we will find explicitly $c_1=e$ and $c_2=1/(2^ne)$. However, these constants are not sharp.

Remark 3.23. From Theorem 3.21 we can infer that (3.33) is not only a sufficient, but also a necessary condition for u being an element of $BMO(Q;\mathbb{R})$. Indeed, Theorem 3.21 applies to every subcube $C\subset Q$ and the BMO seminorm $\|u\|_{BMO(C;\mathbb{R})}$ is obviously bounded by $\|u\|_{BMO(Q;\mathbb{R})}$.

Before presenting the proof, we discuss here two very important consequences of this result.

Corollary 3.24 (Exponential integrability of BMO functions)**.** *For any $0<c<c_2$ there exists $K(c,c_1,c_2)$ such that*

$$\fint_Q e^{c|u-u_Q|/\|u\|_{BMO}}\,dx\le K(c,c_1,c_2)\qquad\forall u\in BMO(Q;\mathbb{R})\setminus\{0\}.$$

Proof. The conclusion follows from a direct computation:

$$\int_Q e^{c|u-u_Q|}\,dx=c\int_0^\infty e^{ct}\mathscr{L}^n\left(\{|u-u_Q|>t\}\right)\,dt\le cc_1\int_0^\infty e^{(c-c_2)t}\,dt=\frac{cc_1}{c_2-c},$$

where we assumed for simplicity $\|u\|_{BMO(Q)}=1$, $\mathscr{L}^n(Q)=1$ and we used the John-Nirenberg inequality. The general case is then obtained via standard scaling arguments. □

Remark 3.25 (Better integrability of $W^{1,n}$ functions)**.** The previous theorem ensures that functions in the class BMO (and hence also in $W^{1,n}$) have exponential integrability properties. This result can be partly refined by the celebrated *Moser-Trudinger inequality*, that we quote here without proof. The result first appeared in [92], and short after a different proof was obtained by [76] which allows determining the corresponding sharp

constant. For any $n > 1$ set $\alpha_n := n\omega_{n-1}^{1/(n-1)}$ and consider a bounded domain Ω in $\mathbb{R}^n$, with $n > 1$. Then

$$\sup\left\{\int_\Omega \exp\left(\alpha_n\left(\frac{|u|}{\|u\|_{W^{1,n}}}\right)^{n/(n-1)}\right)dx \,:\, u \in W_0^{1,n}(\Omega;\mathbb{R}) \setminus \{0\}\right\} < +\infty.$$

Getting back to our discussion, it is clear that we can also exploit Theorem 3.21 to gain better integrability (in L^p spaces) of BMO functions, as is specified by the following statement.

Theorem 3.26. *For all $p \in [1,\infty)$ there exists a constant $c = c(n,p)$ such that*

$$\left(\fint_Q |u - u_Q|^p \, dx\right)^{\frac{1}{p}} \le c\|u\|_{BMO} \qquad \forall u \in BMO(Q;\mathbb{R}).$$

As a result, the following equivalences hold:

$$\mathcal{L}^{p,n}(Q;\mathbb{R}) \sim BMO(Q;\mathbb{R}) \sim \mathcal{L}^{1,n}(Q;\mathbb{R}), \tag{3.35}$$

meaning that the three sets contain the same elements and the corresponding semi-norms are all equivalent.

The proof of Theorem 3.26 relies on a simple computation, similar to the one presented before in order to get exponential integrability. Indeed, assuming $\|u\|_{BMO} = 1$, (3.34) gives

$$\fint_Q |u - u_Q|^p \, dx = p\int_0^\infty \mathscr{L}^n\big(\{|u - u_Q| > s\}\big)s^{p-1}\, ds \le c_1 p \int_0^\infty e^{-c_2 s} s^{p-1}\, ds.$$

We can now proceed with the proof of the John-Nirenberg inequality (3.34).

Proof. By homogeneity, we can assume without loss of generality that $\|u\|_{BMO} = 1$. Let $\alpha > 1$ be a parameter, to be specified later. We claim that it is possible to define, for any $k \ge 1$, a countable family of open, pairwise disjoint subcubes $\left\{Q_i^k\right\}_{i\in I_k}$ contained in Q such that

(i) $|u(x) - u_Q| \le 2^n \alpha k$ $\mathscr{L}^n$-a.e. on $Q \setminus \cup_{i\in I_k} Q_i^k$;

(ii) $\sum_{i\in I_k} \mathscr{L}^n(Q_i^k) \le \alpha^{-k}\mathscr{L}^n(Q)$.

The combination of linear growth in (i) and geometric decay in (ii) leads to the exponential decay of the Lebesgue measure of the superlevels: indeed, choose k such that $2^n\alpha k \le t < 2^n\alpha(k+1)$, then

$$\mathscr{L}^n\left(\{|u - u_Q| > t\}\right) \le \mathscr{L}^n\left(\{|u - u_Q| > 2^n\alpha k\}\right) \le \alpha^{-k}\mathscr{L}^n(Q)$$

by the combined use of the previous properties. Now we want $\alpha^{-k} \leq c_1 e^{-c_2 t}$ for all $t \in [2^n \alpha k, 2^n \alpha(k+1))$, which is certainly verified if we impose

$$\alpha^{-k} = c_1 e^{-c_2 2^n \alpha(k+1)}.$$

To that scope, we first express c_1, c_2 in terms of α requiring

$$e^{c_2 2^n \alpha} = \alpha, \qquad c_1 e^{-c_2 2^n \alpha} = 1$$

so that $c_2 = \log \alpha/(2^n \alpha), c_1 = \alpha$ and then we maximize $c_2(\alpha)$ with respect to $\alpha > 1$ to find

$$\alpha = e, \quad c_1 = e, \quad c_2 = \frac{1}{2^n e}.$$

Now we just need to prove the claim. If $k = 1$ we simply apply the Calderón-Zygmund decomposition (see Subsection B.4) to $f = |u - u_Q|$ for the level α and get a collection $\{Q_i^1\}_{i \in I_1}$. We have to verify that the required conditions are verified. Condition (ii) follows by Remark B.19, while (i) is obvious since $|u(x) - u_Q| \leq \alpha$ $\mathscr{L}^n$-a.e. out of the union of Q_i^1 by construction. But, since $\|u\|_{BMO} = 1$, we also know that

$$\forall i \in I_1 \qquad \fint_{Q_i^1} |u - u_{Q_i^1}| \, dx \leq 1 < \alpha,$$

hence we can iterate the construction, by applying the Calderón-Zygmund decomposition to each of the functions $|u - u_{Q_i^1}|$ with respect to the corresponding cubes Q_i^1. In this way, we find a family of disjoint cubes $\left\{Q_{i,l}^2\right\}$, each contained in one of the previous ones, that is $Q_{i,l}^2 \subset Q_i^1$ for every i, l. Moreover Remark B.19 and the inductive assumption give (for $k = 2$)

$$\sum_{i,l} \mathscr{L}^n(Q_{i,l}^2) \leq \sum_i \frac{1}{\alpha} \int_{Q_i^1} |u - u_{Q_i^1}| \, dx \leq \sum_i \frac{1}{\alpha} \mathscr{L}^n(Q_i^1) \leq \frac{1}{\alpha^2} \mathscr{L}^n(Q),$$

which is (ii). In order to get (i), notice that

$$Q \setminus \bigcup_{i,l} Q_{i,l}^2 \subset \left(Q \setminus \bigcup_i Q_i^1 \right) \cup \left(\bigcup_i (Q_i^1 \setminus \bigcup_l Q_{i,l}^2) \right)$$

so for the first set in the inclusion the thesis is obvious by the case $k = 1$. For the second one, we first observe that

$$|u_Q - u_{Q_i^1}| \leq \fint_{Q_i^1} |u_Q - u| \, dx \leq 2^n \alpha$$

and consequently, since $|u - u_{Q_i^1}| \leq \alpha$ on $Q_i^1 \setminus \cup_l Q_{i,l}^2$ we get

$$|u(x) - u_Q| \leq |u(x) - u_{Q_i^1}| + |u_{Q_i^1} - u_Q| \leq \alpha + 2^n \alpha \leq 2^n \cdot 2\alpha.$$

With minor changes, we can deal with the general case $k > 1$, provided the family of subcubes satisfies also the condition

(iii) $|u_Q - u_{Q_k^i}| \leq 2^n \alpha k$ for every $i \in I_k$.

To check conditions (i), (ii) and (iii) at each inductive step $k > 1$ is what we need to conclude the argument and the proof. □

Theorem 3.21 can be extended considering L^p norms, so that the case of BMO maps could be recovered as a suitable limit, as $p \to \infty$, of the following more general result.

Theorem 3.27 (John-Nirenberg (second version), [61])**.** *For any $p \in [1, \infty)$ and a function $u \in L^p(Q; \mathbb{R})$ define*

$$K_p^p(u) := \sup \left\{ \sum_i \mathscr{L}^n(Q_i) \left(\fint_{Q_i} |u(x) - u_{Q_i}| \, dx \right)^p \ : \ \{Q_i\} \text{ partition of } Q \right\}.$$

Then there exists a constant $c = c(p, n)$ such that

$$\|u - u_Q\|_{L^p_w} \leq c K_p(u).$$

Since

$$\forall\, C \subset Q \text{ cube}, \quad \mathscr{L}^n(C)^{\frac{1}{p}} \fint_C |u - u_C| \, dx \leq K_p(u) \leq \|u\|_{BMO}$$

it is not hard to see that indeed $K_p(u) \to \|u\|_{BMO}$ as $p \to \infty$.

The proof of Theorem 3.27 is basically the same as Theorem 3.21, the goal being to prove the polynomial decay

$$\mathscr{L}^n(\{|u - u_Q| > t\}) \leq \frac{c^p}{t^p} K_p^p(u) \qquad t > 0,$$

instead of an exponential decay.

The following important result improves the classical interpolation theorems in L^p spaces, replacing in the target the L^∞ norm with the BMO norm. This is crucial for the application to elliptic equations, as we will see in the next section.

Theorem 3.28 (Stampacchia's interpolation, [87])**.** *Let $Q, Q_* \subset \mathbb{R}^n$ be cubes, define $D = L^\infty(Q; \mathbb{R}^s)$ and take $p \in [1, \infty)$. Consider a linear operator $T : D \to BMO(Q_*; \mathbb{R})$, that is continuous with respect to the norms $(L^\infty(Q; \mathbb{R}^s), BMO(Q_*; \mathbb{R}))$ and $(L^p(Q; \mathbb{R}^s), L^p(Q_*; \mathbb{R}))$. Then for every $r \in [p, \infty)$ the operator T is continuous with respect to the $(L^r(Q; \mathbb{R}^s), L^r(Q_*; \mathbb{R}))$ topologies.*

Proof. First of all, notice that it suffices to present the proof for $s = 1$: indeed, the general case when $u = (u^1, \dots, u^s)$ follows by applying this result to the components

$$T^j u := T(0, \dots, 0, u^j, 0, \dots, 0), \quad j = 1, \dots, s$$

(where u^j is in the j-th slot of T) thanks to the linearity of the operator T in question.

Therefore, we shall assume from now onwards, and throughout the proof, that $s = 1$. We fix a partition $\{Q_i\}$ of Q_* and we regularize the operator T with respect to $\{Q_i\}$ by defining

$$\tilde{T}(u)(x) := \fint_{Q_i} |Tu(y) - (Tu)_{Q_i}| \, dy \qquad \forall\, x \in Q_i \,,$$

where $(Tu)_{Q_i}$ denotes the mean $\fint_{Q_i} Tu$. Here and below we do not write the dependence of $\tilde{T}$ from $\{Q_i\}$ for the sake of brevity. We claim that $\tilde{T}$ satisfies the assumptions of Marcinkiewicz theorem (see Theorem B.12). Indeed

(1) $\tilde{T}$ is obviously 1-subadditive;

(2) $L^\infty \to L^\infty$ continuity holds by the inequality[4]

$$\|\tilde{T}u\|_{L^\infty} = \sup_i \fint_{Q_i} |Tu(y) - (Tu)_{Q_i}| \, dy \le \|Tu\|_{BMO} \le \|T\|_{\mathcal{L}(BMO;L^\infty)} \|u\|_{L^\infty} \,;$$

(3) $L^p \to L^p$ continuity holds too, since by Jensen's inequality

$$\begin{aligned}
\|\tilde{T}u\|_{L^p}^p &= \sum_i \mathscr{L}^n(Q_i) \left(\fint_{Q_i} |Tu(y) - (Tu)_{Q_i}| \, dy \right)^p \\
&\le \sum_i \int_{Q_i} |Tu(y) - (Tu)_{Q_i}|^p \, dy \\
&\le 2^{p-1} \sum_i \int_{Q_i} \left(|Tu(y)|^p + |(Tu)_{Q_i}|^p \right) dy \le 2^p \|Tu\|_{L^p}^p \le 2^p \|T\|_{\mathcal{L}(L^p;L^p)}^p \|u\|_{L^p}^p .
\end{aligned}$$

Thanks to Marcinkiewicz theorem B.12 the operator

$$\tilde{T} : D \subset L^r(Q;\mathbb{R}) \longrightarrow L^r(Q_*;\mathbb{R}) \tag{3.36}$$

is continuous for every $r \in [p, \infty]$, and the corresponding operator norm $\|\tilde{T}\|_{\mathcal{L}(L^r;L^r)}$ can be bounded independently of the chosen partition $\{Q_i\}$. We now need to exploit this

[4] Given two Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ and a linear, continuous operator $T : X \to Y$, we denote by $\|\cdot\|_{\mathcal{L}(X;Y)}$ the operator norm $\|T\|_{\mathcal{L}(X,Y)} := \sup_{\|x\|_X \le 1} \|Tx\|_Y$.

preliminary result to gain boundedness of the operator T. In order to extract useful information from Theorem 3.27, for $r \in [p, \infty)$, we estimate

$$\begin{aligned} K_r^r(Tu) &= \sup_{\{Q_i\}} \sum_i \mathscr{L}^n(Q_i) \left(\fint_{Q_i} |Tu(y) - (Tu)_{Q_i}| \, dy \right)^r \\ &= \sup_{\{Q_i\}} \|\tilde{T}_{\{Q_i\}} u\|_{L^r}^r \leq \|\tilde{T}\|_{\mathcal{L}(L^r;L^r)}^r \|u\|_{L^r}^r, \end{aligned}$$

where we used the continuity property of $\tilde{T} : L^r(Q;\mathbb{R}) \to L^r(Q_*;\mathbb{R})$ stated in (3.36). Therefore, by Theorem 3.27, we get a constant $c = c(r, n, T)$

$$\|Tu - (Tu)_Q\|_{L_w^r} \leq c\|u\|_{L^r} \qquad \forall u \in D,$$

where L_w^r is the Marcinkiewicz space recalled in Definition B.3. Since the operator $u \mapsto (Tu)_Q$ obviously satisfies a similar L_w^r estimate (because, by Jensen's inequality, it is a continuous operator $L^r \to L^r$) and the norm L_w^r is 2-subadditive, we conclude that $\|Tu\|_{L_w^r} \leq c\|u\|_{L^r}$ for all $u \in D$. Again, thanks to Marcinkiewicz theorem, with exponents p and r, we have that the continuity $L^{r'} \to L^{r'}$ holds for every $r' \in [p, r)$. Since r is arbitrary, we got our conclusion. □

3.5 Regularity in L^p spaces

We are now ready to employ the harmonic analysis tools seen in the previous section to the study of regularity in L^p spaces for elliptic problems, by first considering the case of constant coefficients and then dealing with continuous ones. Suppose that $\Omega \subset \mathbb{R}^n$ is an open, bounded and regular set (even though all main results could also be extended to the case of Lipschitz boundary), assume that the coefficients $A_{ij}^{\alpha\beta}$ satisfy the Legendre-Hadamard condition with $\lambda > 0$ and are bounded from above (in Hilbert-Schmidt norm) by Λ, and consider the equation given by

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A_{ij}^{\alpha\beta} \partial_{x_\beta} u^j\big) = -\sum_\alpha \partial_{x_\alpha} F_i^\alpha \qquad i = 1, \dots, m \tag{3.37}$$

where a priori we require, as usual, $u \in H_0^1(\Omega;\mathbb{R}^m)$.

Keeping in mind the statement of Theorem 3.28, we define

$$TF := \nabla u$$

and wish to interpolate between estimates that have already been obtained along the course of our discussion. To that aim, we shall assume that the results presented in the previous section can actually be extended to more general domains than squares and in

particular to bounded and regular domains. The necessary modifications, of notational character, are left to the reader.

That being said, one can exploit the Caccioppoli-Leray inequality to prove that the linear map $T : L^2(\Omega;\mathbb{R}^{m\times n}) \to L^2(\Omega;\mathbb{R}^{m\times n})$ is continuous (the proof goes along the lines of the argument presented for Theorem 2.19, which deals with higher-order integrability). On the other hand, thanks to the work of Campanato (specifically, see Theorem 3.15), we also obtained the continuity of $T : \mathcal{L}^{2,\lambda} \to \mathcal{L}^{2,\lambda}$ when $0 \le \lambda < n+2$, thus choosing $\lambda = n$ and using the equivalence (3.35) we see that T is actually continuous from BMO to BMO, and in particular we can regard it as a linear continuous operator

$$T : L^\infty(\Omega;\mathbb{R}^{m\times n}) \longrightarrow BMO(\Omega;\mathbb{R}^{m\times n}). \tag{3.38}$$

Therefore, we can apply Theorem 3.28, with $s = mn$, to each component of the map T i.e. to the linear functionals

$$T^i_\alpha F := (\nabla u)^i_\alpha = \partial_{x_\alpha} u^i$$

and we finally get that

$$T : D \to L^p(\Omega;\mathbb{R}^{m\times n}) \tag{3.39}$$

is (L^p, L^p)-continuous if $p \in [2,\infty)$. Since the (unique) extension of T to the whole of L^p still maps F into ∇u, with u solution to (3.37), we have proved the following result:

Theorem 3.29. *For all $p \in [2,\infty)$ the operator $F \mapsto \nabla u$ in (3.37) maps $L^p(\Omega;\mathbb{R}^{m\times n})$ into $L^p(\Omega;\mathbb{R}^{m\times n})$ continuously.*

Let us explicitly remark the importance of weakening the norm in the target space in (3.38): we passed from L^∞ (for which, as we will see, no estimate is possible) to BMO.

Our intention is now to extend the previous result for $p \in (1,2)$, by means of a duality argument.

Lemma 3.30 (Helmholtz decomposition, [51][5])**.** *If $p \ge 2$ and B is a matrix satisfying the Legendre-Hadamard inequality with constant $\lambda > 0$, and uniformly bounded from above (in Hilbert-Schmidt norm) by Λ, a map $G \in L^p(\Omega;\mathbb{R}^{m\times n})$ can always be written as a sum*

$$G = B\nabla\phi + \tilde{G}, \tag{3.40}$$

where $\phi \in H^{1,p}_0(\Omega;\mathbb{R}^m)$, $\tilde{G}$ is divergence-free

$$\sum_\alpha \partial_{x_\alpha}(\tilde{G}^\alpha_i) = 0 \qquad \text{in } \Omega, \text{ for all } i = 1,\dots,m$$

and the following inequality holds:

$$\|\nabla\phi\|_{L^p} \le c_H \|G\|_{L^p}, \tag{3.41}$$

for some constant $c_H = c_H(n,\Omega,\lambda,\Lambda) > 0$.

[5]The classical decomposition result is named after Helmholtz [51], nonetheless we use here a more recent version for L^p functions, see [25].

Proof. It is sufficient to solve in $H^1_0(\Omega;\mathbb{R}^m)$ the equation $\mathrm{div}(B\nabla\phi) = \mathrm{div}(G)$, namely

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}(B^{\alpha\beta}_{ij}(\nabla\phi)^j_\beta) = -\sum_\alpha \partial_{x_\alpha}(G^\alpha_i) \qquad i = 1,\dots,m$$

and set $\tilde{G} := G - B\nabla\phi$. The estimate (3.41) is then just a consequence of Theorem 3.29. □

Fix $q \in (1,2)$, so that its conjugate exponent $p = q'$ is larger than 2, and further set $D := L^2(\Omega;\mathbb{R}^{m\times n})$. Our aim is to prove that $T : L^2(\Omega;\mathbb{R}^{m\times n}) \to L^q(\Omega;\mathbb{R}^{m\times n})$ is (L^q, L^q)-continuous: we are going to show that for every $F \in D$, TF belongs to $(L^p)' \sim L^q$ with a uniform estimate. In the chain of inequalities that follows we are using A^*, that is the adjoint matrix of A, which certainly also satisfies the Legendre-Hadamard property if A does. Lemma 3.30 is exploited in order to decompose the generic function $G \in L^p(\Omega;\mathbb{R}^{m\times n})$ as in (3.40), so that we can write

$$\begin{aligned}
\sup_{\|G\|_{L^p}\le 1} (TF, G)_{L^2} &= \sup_{\|G\|_{L^p}\le 1} \int_\Omega \langle TF(x), G(x)\rangle\, dx \\
&= \sup_{\|G\|_{L^p}\le 1} \int_\Omega \langle \nabla u(x), A^*\nabla\phi(x) + \tilde{G}(x)\rangle\, dx \\
&\le \sup_{\|\nabla\phi\|_{L^p}\le c_H} \int_\Omega \langle A\nabla u(x), \nabla\phi(x)\rangle\, dx \\
&= \sup_{\|\nabla\phi\|_{L^p}\le c_H} \int_\Omega \langle F(x), \nabla\phi(x)\rangle\, dx \le c_H\|F\|_{L^q}.
\end{aligned}$$

If we approximate $F \in L^q(\Omega;\mathbb{R}^{m\times n})$ in the L^q topology by functions $F_k \in L^2(\Omega;\mathbb{R}^{m\times n})$ we can use the (L^q, L^q)-continuity to prove existence of weak solutions to the equation in $H^{1,q}_0$, when the data are just in L^q. Notice that the solutions obtained in this way have no variational character anymore, since their energy $\int_\Omega \langle A\nabla u, \nabla u\rangle\, dx$ need not even be finite (for this reason they are sometimes called *very weak* solutions). As a consequence, the uniqueness of these solutions needs an *ad hoc* argument, which is once more based on the Helmholtz decomposition.

Theorem 3.31. *For all $q \in (1,2)$ there exists a continuous operator $R : L^q(\Omega;\mathbb{R}^{m\times n}) \to H^{1,q}_0(\Omega;\mathbb{R}^m)$ mapping the datum F to the unique weak solution u to equation* (3.37).

Proof. We already illustrated the construction of a solution u, by a density argument and uniform L^q bounds. To show uniqueness, it suffices to show that if $u \in H^{1,q}_0(\Omega;\mathbb{R}^m)$ the equations

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\big) = 0, \quad i = 1,\dots,m$$

imply $u = 0$ identically in Ω. To this aim, we define $G = |\nabla u|^{q-2}\nabla u \in L^p(\Omega;\mathbb{R}^{m\times n})$ and apply the Helmholtz decomposition, thereby writing $G = A^*\nabla\phi + \tilde{G}$ with $\phi \in H^{1,p}_0(\Omega;\mathbb{R}^m)$ and $\tilde{G} \in L^p(\Omega;\mathbb{R}^{m\times n})$ divergence-free. By a density argument with respect to both u and ϕ (notice that the exponents are dual) we obtain

$$\int_\Omega \langle \tilde{G}, \nabla u\rangle\, dx = 0, \qquad \int_\Omega \langle A\nabla u, \nabla\phi\rangle\, dx = 0$$

and hence

$$\int_\Omega |\nabla u|^q\, dx = \int_\Omega (G\nabla u)\, dx = \int_\Omega \langle A^*\nabla\phi, \nabla u\rangle\, dx = \int_\Omega \langle A\nabla u, \nabla\phi\rangle\, dx = 0.$$

□

Remark 3.32 (General Helmholtz decomposition)**.** Thanks to Theorem 3.31, the Helmholtz decomposition showed above can actually be obtained for every $p \in (1,\infty)$.

Remark 3.33 ($W^{2,p}$ estimates)**.** By differentiating the equation and multiplying by cutoff functions, we easily see that Theorem 3.29 and Theorem 3.31 yield for any $p \in (1,\infty)$ the implication

$$\begin{cases} -\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A^{\alpha\beta}_{ij}\partial_{x_\beta}u^j\big) = f_i & i = 1,\dots,m \\ |\nabla u| \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R}),\ f \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R}^m) \end{cases} \qquad \Longrightarrow \qquad u \in W^{2,p}_{\mathrm{loc}}(\Omega;\mathbb{R}^m).$$

Remark 3.34 (No L^∞ bound is possible)**.** As it was claimed above, let us show here that T does not map L^∞ into L^∞, with $\Omega = B_1 \subset \mathbb{R}^n$. First we prove that this phenomenon occurs if T is known to be discontinuous, then we prove that T is indeed discontinuous.

To check the first claim, let $(\overline{\Omega}_k)$ be a countable family of pairwise disjoint closed balls contained in Ω: by a scaling argument we can find (since also the rescaled operators of T on Ω_k are discontinuous) functions $F_k \in L^\infty(\Omega_k;\mathbb{R}^{m\times n})$ with $\|F_k\|_\infty = 1$ and solutions $u_k \in H^1_0(\Omega_k;\mathbb{R}^m)$ to the equation (3.37) with $\|\nabla u_k\|_\infty \geq k$. Then it is easily shown (for instance by approximation with finite families of balls) that the function

$$u(x) := \begin{cases} u_k(x) & \text{if } x \in \Omega_k \\ 0 & \text{if } x \in \Omega \setminus \cup_k \Omega_k \end{cases}$$

belongs to $H^1_0(\Omega;\mathbb{R}^m)$, solves the equation with datum

$$F(x) := \begin{cases} F_k(x) & \text{if } x \in \Omega_k \\ 0 & \text{if } x \in \Omega \setminus \cup_k \Omega_k, \end{cases}$$

but its gradient is patently not bounded.

So, it remains to prove that T is necessarily discontinuous, which we will do restricting our discussion to the scalar case for the sake of simplicity. By the same duality argument used before, if T were continuous we would get an estimate of the form

$$\|\nabla u\|_{L^1} \leq c\|F\|_{L^1}$$

whenever $u \in H_0^1(\Omega;\mathbb{R}^m)$ solves equation (3.37) for $m = 1$.

Hence, a standard approximation argument (based on convolution of the right-hand side, and Rellich compactness theorem) would imply the existence, for any vector-valued measure μ in Ω, of solutions of *bounded variation*, i.e. functions $u \in L^1(\Omega;\mathbb{R})$, whose distributional gradient, denoted here by $Du = (D_1u, \dots, D_nu)$, is a vector-valued *measure* satisfying

$$\sum_{\alpha,\beta} \int_\Omega A^{\alpha\beta}\partial_{x_\alpha}\phi \, dD_\beta u = \sum_\alpha \int_\Omega \partial_{x_\alpha}\phi \, d\mu^\alpha \qquad \forall \phi \in C_c^\infty(\Omega;\mathbb{R}). \tag{3.42}$$

and

$$|Du|(\Omega) \leq c|\mu|(\Omega), \tag{3.43}$$

where $|\mu|$ (resp. $|Du|$) denotes the total variation of the measure μ (resp. Du). We refer the reader to Chapter 3 of [5], see in particular Proposition 3.13 and equation (3.11), for further details. On the other hand, we claim that the inequality (3.43) cannot be true. Indeed, when $n = 2$ and $m = 1$, consider the identity matrix $A^{\alpha\beta} := \delta^{\alpha\beta}$ and the corresponding Laplace equation

$$-\Delta v = \delta_0, \tag{3.44}$$

where δ_0 is the Dirac measure supported in 0. The well-known fundamental solution of (3.44) is

$$v(x) = -\frac{\log|x|}{2\pi} \in C^\infty(\mathbb{R}^2 \setminus \{0\};\mathbb{R}),$$

so that $v \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{R})$ for any $p < 2$, with $\nabla v(x) = -(2\pi)^{-1}x/|x|^2$, and (understanding the second derivative in the pointwise sense) $|\nabla^2 v| \notin L^1(\Omega;\mathbb{R})$, since

$$\nabla^2 v(x) = -\frac{1}{2\pi|x|^2}\left(I - 2\frac{x\otimes x}{|x|^2}\right).$$

Now, for any $\eta \in C_c^\infty(\Omega;\mathbb{R})$ with $\eta \equiv 1$ on $B_{1/2}$ we have

$$-\Delta(\partial_{x_\alpha}(v\eta)) = -\partial_{x_\alpha}(-\delta_0 + v\Delta\eta + 2\langle\nabla v, \nabla\eta\rangle)$$

so if we introduce the vector measure μ whose components are defined by

$$\mu^1 = -\delta_0 + v\Delta\eta + 2\langle\nabla v, \nabla\eta\rangle, \quad \mu^2 = 0,$$

we have that the function $w = \partial_{x_1}(\eta v) \in L^1(\mathbb{R}^2;\mathbb{R})$ is a distributional solution in $\mathbb{R}^2$ to the equation

$$-\Delta w = -\sum_\alpha \partial_{x_\alpha}\mu^\alpha.$$

It follows that $\tilde{u} = w - u$, where u is as in (3.42) for this datum μ, is a distributional solution to the Laplace equation in B_1, and therefore standard properties of harmonic functions (for instance the mean value property and a convolution argument applied to $\tilde{u}$) imply that $\tilde{u}$ is equivalent in B_1 to a smooth function. By the properties of u and $\tilde{u}$, it follows that the distributional derivative of $w = \tilde{u} - u$ is locally representable in Ω by a measure with finite total variation. By our choice of η, this implies the same conclusion for $\partial_{x_1} v = w$ in $B_{1/2}$, and a similar argument gives the analogous result for $\partial_{x_2} v$. Since $|\nabla^2 v|$ is not summable in $B_{1/2}$, we have reached a contradiction.

Now we move from the case of constant to that of continuous coefficients, using Korn's technique (similarly to what was previously discussed in the proof of Theorem 3.16).

Theorem 3.35. *Given an open set $\Omega \subset \mathbb{R}^n$ let $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ be a solution to the equation*

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big(A^{\alpha\beta}_{ij}(x)\partial_{x_\beta}u^j\big) = f_i - \sum_\alpha \partial_{x_\alpha}F^\alpha_i \qquad i = 1,\dots,m$$

with continuous coefficients $A^{\alpha\beta}_{ij}$ which satisfy a uniform Legendre-Hadamard condition for some $\lambda > 0$. Given positive real numbers $p \in (1,\infty)$ and q such that its Sobolev conjugate exponent $q^* = qn/(n-q)$ equals p, let us suppose that $F \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$ and $f \in L^q_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$. Then $|\nabla u| \in L^p_{\mathrm{loc}}(\Omega;\mathbb{R})$.

Proof. For the sake of simplicity, we limit ourselves to give the proof for $p \geq 2$. The corresponding estimates for $p \in (1,2)$ can be recovered by means of a duality argument along the lines of the discussion presented above.

Given $s \geq 2$, we will now show that

$$|\nabla u| \in L^{s\wedge p}_{\mathrm{loc}}(\Omega;\mathbb{R}) \quad \Longrightarrow \quad |\nabla u| \in L^{s^*\wedge p}_{\mathrm{loc}}(\Omega;\mathbb{R}), \tag{3.45}$$

where we shall write $a \wedge b$ in lieu of $\min\{a,b\}$ for the sake of notational convenience. Proving such an implication is actually sufficient to complete the proof because we do know that $|\nabla u| \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R})$ (which would be the case $s = 2$) and in finitely many iterations s^* becomes larger than p, thereby proving the claimed assertion. To avoid ambiguities, let us remind the reader that it is tacitly understood that $s^* = \infty$ if $s \geq n$ as was pointed out in Remark 2.7.

Fix a point $x_0 \in \Omega$ and a radius $R > 0$ such that $B_R(x_0) \Subset \Omega$: we choose a cutoff function $\eta \in C^\infty_c(B_R(x_0);\mathbb{R})$, with $0 \leq \eta \leq 1$ and $\eta \equiv 1$ in $B_{R/2}(x_0)$. We claim that ηu belongs to $H^{1,s^*\wedge p}_0(B_R(x_0);\mathbb{R}^m)$ if $R \leq R_0$, where R_0 is a small positive constant to

be determined later depending on the modulus of continuity of A at x_0. This implies, in particular, that $|\nabla u| \in L^{s^*\wedge p}(B_{R/2}(x_0);\mathbb{R})$ and thus completes the proof of the aforementioned implication. With that goal, we proceed in three steps.

Step 1. We start by localizing the equation. Standard algebraic computations imply the following chain of equalities:

$$\begin{aligned}
&\sum_{\alpha,\beta,j}\int_{B_R(x_0)} A_{ij}^{\alpha\beta}(x)\partial_{x_\beta}(\eta u^j)(x)\partial_{x_\alpha}\varphi(x)\,dx \\
&\qquad = \sum_{\alpha,\beta,j}\int_{B_R(x_0)} A_{ij}^{\alpha\beta}(x)\left(\eta(x)\partial_{x_\beta}u^j(x)+u^j(x)\partial_{x_\beta}\eta(x)\right)\partial_{x_\alpha}\varphi(x)\,dx \\
&\qquad = \sum_{\alpha,\beta,j}\int_{B_R(x_0)} A_{ij}^{\alpha\beta}(x)(\partial_{x_\beta}u^j(x)\partial_{x_\alpha}(\eta\varphi)(x)+u^j(x)\partial_{x_\beta}\eta(x)\partial_{x_\alpha}\varphi(x) \\
&\qquad\qquad\qquad - \partial_{x_\beta}u^j(x)\partial_{x_\alpha}\eta(x)\varphi(x))\,dx
\end{aligned}$$

and hence, we can write the equation solved by ηu in the form

$$\begin{aligned}
&\sum_{\alpha,\beta,j}\int_{B_R(x_0)} A_{ij}^{\alpha\beta}(x)\partial_{x_\beta}(\eta u^j)(x)\partial_{x_\alpha}\varphi(x)\,dx \\
&\qquad = \int_{B_R(x_0)} f_i(x)\eta(x)\varphi(x)+\sum_\alpha F_i^\alpha(x)\partial_{x_\alpha}(\eta\varphi)(x) \\
&\qquad\quad + \sum_{\alpha,\beta,j} A_{ij}^{\alpha\beta}(x)\left(u^j(x)\partial_{x_\beta}\eta(x)\partial_{x_\alpha}\varphi(x)-\partial_{x_\beta}u^j(x)\partial_{x_\alpha}\eta(x)\varphi(x)\right)\,dx \\
&\qquad = \int_{B_R(x_0)} \tilde{f}_i(x)\varphi(x)+\sum_\alpha \tilde{F}_i^\alpha(x)\partial_{x_\alpha}\varphi(x)\,dx
\end{aligned}$$

where we have defined

$$\tilde{f}_i(x) := f_i(x)\eta(x)+\sum_\alpha\left(F_i^\alpha(x)\partial_{x_\alpha}\eta(x)-\sum_{\beta,j}A_{ij}^{\alpha\beta}(x)\partial_{x_\beta}u^j(x)\partial_{x_\alpha}\eta(x)\right)$$

and

$$\tilde{F}_i^\alpha(x) := F_i^\alpha(x)\eta(x)+\sum_{\beta,j}A_{ij}^{\alpha\beta}(x)u^j(x)\partial_{x_\beta}\eta(x).$$

At that stage, if we freeze the coefficients at the point x_0 we can rewrite this equation in the form

$$-\sum_{\alpha,\beta,j}\partial_{x_\alpha}\left((A_{ij}^{\alpha\beta}(x_0)\partial_{x_\beta}(\eta u^j)\right) = \tilde{f}_i-\sum_\alpha\partial_{x_\alpha}[\tilde{F}_i^\alpha-\sum_{\beta,j}(A(x)-A(x_0))_{ij}^{\alpha\beta}\partial_{x_\beta}(\eta u)^j],$$

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

Step 2: in order to write the datum $\tilde{f}$ in divergence form, let us consider the problem

$$\begin{cases} -\Delta w = \tilde{f} \\ w \in H^1_0(\Omega; \mathbb{R}^m) \end{cases}$$

which we can always solve provided we require Ω to be bounded and regular, which can always be assumed as our problem is of purely local character. Thanks to the previous L^p regularity result for the case of constant coefficients, since $\tilde{f} \in L^{s\wedge q}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ (because we assumed that $|\nabla u| \in L^{s\wedge p}_{\mathrm{loc}}(\Omega; \mathbb{R})$), we have $|\nabla^2 w| \in L^{s\wedge q}_{\mathrm{loc}}(\Omega; \mathbb{R})$ (see also Remark 3.33). By virtue of the Sobolev embedding theorem, we then get $|\nabla w| \in L^{(s\wedge q)^*}_{\mathrm{loc}}(\Omega; \mathbb{R})$, hence

$$|\nabla w| \in L^{s^*\wedge q^*}_{\mathrm{loc}}(\Omega; \mathbb{R}) = L^{s^*\wedge p}_{\mathrm{loc}}(\Omega; \mathbb{R}),$$

since we are indeed assuming $q^* = p$. Now we define

$$F^*(x) := \tilde{F}(x) + \nabla w(x) \in L^{s^*\wedge p}_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n}).$$

Step 3: set $E = H^{1,s^*\wedge p}_0(B_R(x_0); \mathbb{R}^m)$, denote by $\|\cdot\|_E$ the corresponding Hilbertian norm and let us define the operator $\Theta : E \to E$ which associates to each $V \in E$ the function $v \in E$ that solves

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big((A^{\alpha\beta}_{ij}(x_0)\partial_{x_\beta}(v^j)\big) = -\sum_{\alpha} \partial_{x_\alpha}(F^*)^\alpha_i - \sum_{\alpha,\beta,j} \partial_{x_\alpha}[(A(x_0) - A(x))^{\alpha\beta}_{ij}\partial_{x_\beta}V^j]. \quad (3.46)$$

The operator Θ is well-defined because $|F^*| \in L^{s^*\wedge p}(B_R(x_0); \mathbb{R})$, as we saw in the previous step, and we can take advantage of regularity theory for constant coefficients operators, which provides a scaling-invariant constant $c > 0$, independent of R, such that

$$\|\nabla(v_1 - v_2)\|_{L^{s^*\wedge p}} \le c\|\,(A(x_0) - A)\,\nabla(V_1 - V_2)\|_{L^{s^*\wedge p}}.$$

Hence, it follows that the operator Θ is indeed a contraction, since one can write

$$\|v_1 - v_2\|_E = \|\nabla(v_1 - v_2)\|_{L^{s^*\wedge p}} \le c\|\,(A(x_0) - A)\,\nabla(V_1 - V_2)\|_{L^{s^*\wedge p}} \le \frac{1}{2}\|\nabla(V_1 - V_2)\|_E$$

if R is sufficiently small, only depending on the oscillation of A in $B_R(x_0)$. That being said, let us denote by $v_* \in E$ the unique fixed point of $\Theta : E \to E$. We already know that the function ηu already solves (3.46), but in the larger space $H^{1,s\wedge p}_0(B_R(x_0); \mathbb{R}^m)$. Thus we gain $\eta u \in H^{1,s^*\wedge p}_0(B_R(x_0); \mathbb{R}^m)$ if we are able to show that $v_* = \eta u$, and to see this

it suffices to show that uniqueness holds in the larger space as well. With that goal in mind, we consider the difference

$$v_{**} := v_* - \eta u \in H_0^{1,s\wedge p}(B_R(x_0);\mathbb{R}^m) \subset H_0^1(B_R(x_0);\mathbb{R}^m)$$

which clearly satisfies (in weak sense)

$$-\sum_{\alpha,\beta,j} \partial_{x_\alpha}\big((A_{ij}^{\alpha\beta}(x)\partial_{x_\beta}(v_{**}^j)\big) = 0.$$

At that stage we obtain $v_{**} \equiv 0$ as a direct consequence of the variational characterization of the solution. This concludes the proof. □

3.6 De Giorgi's solution of Hilbert's *XIX* problem

We start by briefly recalling here the context and setting of Hilbert's *XIX* problem [52]. One is concerned with local minimizers u of scalar functionals of the form

$$w \longmapsto \int_\Omega L(\nabla w)\,dx$$

where $L \in C^2$ satisfies the following ellipticity property: there exist two positive constants $\lambda \le \Lambda$ such that $\Lambda I \ge \nabla^2 L(p) \ge \lambda I$ for all $p \in \mathbb{R}^n$ (this implies in particular that $|\nabla^2 L|$ is uniformly bounded). We have already seen that under these assumptions it is possible to derive the Euler-Lagrange equation satisfied by u, which takes the form

$$\sum_\alpha \partial_{x_\alpha}\left(\partial_{p^\alpha} L(\nabla u)\right) = 0.$$

By differentiation, for any direction $\gamma \in \{1,\dots,n\}$, the equation for $v := \partial_{x_\gamma} u$ is

$$\sum_{\alpha,\beta} \partial_{x_\alpha}\left(\partial_{p^\alpha p^\beta} L\left(\nabla u\right)\partial_{x_\beta} v\right) = 0.$$

Recall that, in order to derive this equation, we needed to work with the approximation $\Delta_{h,\gamma} u$ (in lieu of $\partial_{x_\gamma} u$), with the interpolating operator

$$\widetilde{A}_h^{\alpha\beta}(x) := \int_0^1 \partial_{p^\alpha p^\beta} L(t\nabla u(x+he_\gamma) + (1-t)\nabla u(x))\,dt$$

and to exploit the Caccioppoli-Leray inequality in gaining those uniform integral estimates that are necessary in order to take the limit as $h \to 0$ in the approximation scheme in question.

One of the striking ideas of De Giorgi [26] was basically to split the problem, that is to deal with u and v separately, as ∇u is only involved in the coefficients of the equation for v. The key point of the regularization procedure is then to show that under no regularity assumption on ∇u (i.e. nothing more than measurability), if v is a solution of the equation above, then $v \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R})$, with α depending only on n and on the ellipticity constants λ, Λ. If this is true and we now assume $L \in C^{2,\beta}(\mathbb{R}^n;\mathbb{R})$, we can proceed as follows:

$$v \in C^{0,\alpha}(\Omega;\mathbb{R}) \Rightarrow u \in C^{1,\alpha}(\Omega;\mathbb{R}) \Rightarrow \partial_p\partial_p L(\nabla u) \in C^{0,\alpha\beta}(\Omega;\mathbb{R}) \Rightarrow v \in C^{1,\alpha\beta}(\Omega;\mathbb{R}).$$

Notice that the above implications rely upon the fact that $\partial_p\partial_p L$ is Hölder continuous by assumption and on the Schauder estimates contained in the statement of Theorem 3.17. Since v is any partial derivative of u, we eventually get $u \in C^{2,\alpha\beta}(\Omega;\mathbb{R})$. If L is more regular, by continuing this iteration (now using Schauder regularity for elliptic equations whose coefficients are $C^{1,\gamma}$, $C^{2,\gamma}$ and so on) we obtain the implication

$$L \in C^{\infty}(\Omega;\mathbb{R}) \Rightarrow v \in C^{\infty}(\Omega;\mathbb{R})$$

and also, by the tools developed in [53] by E. Hopf, one can conclude that in fact

$$L \in C^{\omega}(\Omega;\mathbb{R}) \Rightarrow v \in C^{\omega}(\Omega;\mathbb{R}),$$

which is the complete solution of the problem raised by Hilbert.

Actually, let us remark that the problem in question has already been solved in the simpler case when $n = 2$: by means of Widman's hole-filling technique, we could prove that $|\nabla v| \in L^{2,\alpha}(\Omega;\mathbb{R})$ and hence $v \in \mathcal{L}^{2,\alpha+2}(\Omega;\mathbb{R})$ for some $\alpha > 0$, which is enough, if $n = 2$, to conclude that $u \in C^{0,\frac{\alpha}{2}}(\Omega;\mathbb{R})$. This was presented in detail in Subsection 3.2.

In order to proceed, we start by accurately describing our setup. Let Ω be an open domain in $\mathbb{R}^n$, consider two constants $0 < \lambda \leq \Lambda < +\infty$ and let $A^{\alpha\beta}$ be a Borel symmetric matrix satisfying the inequalities $\lambda I \leq A(x) \leq \Lambda I$ for $\mathscr{L}^n$-a.e. $x \in \Omega$. We want to show that if $u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ solves the problem

$$-\sum_{\alpha,\beta} \partial_{x_\alpha}\left(A^{\alpha\beta}(x)\partial_{x_\beta}u(x)\right) = 0$$

then u is locally Hölder-continuous. Some notation is needed: for $B_\rho(x) \subset \Omega$ we define the superlevels

$$A(k,\rho) := \{u > k\} \cap B_\rho(x),$$

where the dependence on the center x shall be omitted. This should not generate confusion, since we work with a fixed center, unless otherwise stated. In this section we derive various functional inequalities, however we shall not be concerned with finding the

sharpest constants, but only with the functional dependence of these quantities. Therefore, in order to avoid unnecessary complications of the notation, we use the same symbol (generally c) to indicate *different* constants, possibly varying from one passage to the next one. However we indicate the functional dependence explicitly whenever this is appropriate and so we use expressions like $c(n)$ or $c(n, \lambda, \Lambda)$ many times.

Theorem 3.36 (Caccioppoli inequality on superlevel sets)**.** *For any $k \in \mathbb{R}$ and radii $0 < \rho < R$ such that $B_\rho(x) \subset B_R(x) \Subset \Omega$ we have*

$$\int_{A(k,\rho)} |\nabla u|^2 \, dy \leq \frac{c_{CL,L}}{(R-\rho)^2} \int_{A(k,R)} (u-k)^2 \, dy, \tag{3.47}$$

with $c_{CL,L} = 16\Lambda^2/\lambda^2$.

Remark 3.37. It should be noted that the previous theorem generalizes the Caccioppoli-Leray inequality, since we do not restrict to $\rho = R/2$ and we work with superlevel sets $A(k, \rho)$ and $A(k, R)$.

Theorem 3.38 (Chain rule)**.** *If $u \in W^{1,1}_{\mathrm{loc}}(\Omega; \mathbb{R})$, then for any $k \in \mathbb{R}$ the function $(u-k)^+$ belongs to $W^{1,1}_{\mathrm{loc}}(\Omega; \mathbb{R})$. Moreover we have that $\nabla(u-k)^+ = \nabla u$ a.e. on $\{u > k\}$, while $\nabla(u-k)^+ = 0$ a.e. on $\{u \leq k\}$.*

Proof. Since this theorem is rather classical, we just sketch its proof. By the arbitrariness of u, the problem is clearly translation-invariant and we can assume without loss of generality $k = 0$. Consider the family of functions defined by

$$\varphi_\varepsilon(t) := \begin{cases} \sqrt{t^2 + \varepsilon^2} - \varepsilon & \text{if } t \geq 0 \\ 0 & \text{else.} \end{cases}$$

Notice that the corresponding derivatives $\varphi'_\varepsilon(t)$ are uniformly bounded and converge to the characteristic function $\chi_{\{t>0\}}$ as one lets $\varepsilon \to 0$. Moreover, let (u_n) be a sequence of C^1 functions approximating u in $W^{1,1}_{\mathrm{loc}}(\Omega; \mathbb{R})$. We have that for any $n \in \mathbb{N}$ and $\varepsilon > 0$ the classical chain rule gives $\nabla [\varphi_\varepsilon(u_n)] = \varphi'_\varepsilon(u_n) \nabla u_n$. Passing to the limit as $n \to \infty$ one obtains the equality $\nabla [\varphi_\varepsilon(u)] = \varphi'_\varepsilon(u) \nabla u$, to be understood in $L^1_{\mathrm{loc}}(\Omega; \mathbb{R})$. At that stage, we can pass to the limit as $\varepsilon \downarrow 0$ and use the dominated convergence theorem to conclude that $\nabla u^+ = \chi_{\{u>0\}} \nabla u$. □

We can now proceed to the proof of the Caccioppoli inequality on level sets.

Proof. Let η be a cutoff function supported in $B_R(x)$, with $\eta \equiv 1$ on $\overline{B}_\rho(x)$ and $|\nabla \eta| \leq 2/(R-\rho)$. If we exploit the weak form of our equation with the test function $\varphi := \eta^2 (u-k)^+$ we get

$$\begin{aligned} \int_{A(k,R)} \eta^2 \langle A\nabla u, \nabla u \rangle \, dy &= -2 \int_{B_R(x)} \eta \langle A \nabla u, \nabla \eta (u-k)^+ \rangle \, dy \\ &\leq \frac{\Lambda}{\varepsilon} \int_{A(k,R)} \eta^2 \, |\nabla u|^2 \, dy + \frac{4\varepsilon\Lambda}{(R-\rho)^2} \int_{A(k,R)} (u-k)^2 \, dy \end{aligned}$$

for any $\varepsilon > 0$, by our upper bound on $\nabla\eta$ and by Young's inequality. Hence, we set $\varepsilon = 2\Lambda/\lambda$ so that we obtain

$$\int_{A(k,R)} \eta^2 \langle A\nabla u, \nabla u\rangle \, dy \leq \frac{\lambda}{2} \int_{A(k,R)} \eta^2 |\nabla u|^2 \, dy + \frac{8\Lambda^2}{\lambda(R-\rho)^2} \int_{A(k,R)} (u-k)^2 \, dy.$$

Thanks to the uniform ellipticity assumption, and the fact that on the smaller ball η is identically equal to 1, we eventually get

$$\int_{A(k,\rho)} |\nabla u|^2 \, dy \leq \frac{16\Lambda^2}{\lambda^2(R-\rho)^2} \int_{A(k,R)} (u-k)^2 \, dy,$$

which is what we had claimed. □

The second great idea of De Giorgi was that (one-sided) regularity could be achieved for all functions satisfying the previous functional inequality, regardless of the fact that these were (or were not) solutions to an elliptic equation. For this reason he introduced a special class of objects.

Definition 3.39 (De Giorgi's class)**.** *We define the De Giorgi class $DG_+(\Omega)$ as follows:*

$$DG_+(\Omega) := \{u : \exists\, c \in \mathbb{R} \text{ s.t. } \forall\, k \in \mathbb{R}, B_r(x) \Subset B_R(x) \Subset \Omega, u \text{ satisfies } (3.47)\}.$$

In this case, we also define $c^+_{DG}(u)$ to be the least constant, greater or equal than 1, for which the condition (3.47) is verified.

Remark 3.40. From the previous proof, it should be clear that we do not really require u to be a solution, but just a *sub-solution* of our problem. In fact, we have proved that

$$-\sum_{\alpha,\beta} \partial_{x_\alpha}\left(A^{\alpha\beta}(x)\partial_{x_\beta} u(x)\right) \leq 0 \text{ in weak sense} \quad \Longrightarrow \quad u \in DG(\Omega),\ c^+_{DG}(u) \leq \frac{16\Lambda^2}{\lambda^2}.$$

In a similar way, the class $DG_-(\Omega;\mathbb{R})$ (corresponding to supersolutions) and $c^-_{DG}(u)$ could be defined by

$$\int_{\{u<k\}\cap B_\rho(x)} |\nabla u|^2 \, dy \leq \frac{c}{(R-\rho)^2} \int_{\{u<k\}\cap B_R(x)} (u-k)^2 \, dy$$

and obviously $u \mapsto -u$ maps $DG_+(\Omega)$ in $DG_-(\Omega)$ bijectively, with $c^+_{DG}(u) = c^-_{DG}(-u)$.

The main part of the program by De Giorgi can be divided into two steps, whose central goals correspond to proving the following two assertions:

(i) If $u \in DG^+(\Omega)$, then it satisfies a strong *maximum principle* in a quantitative form (more precisely the L^2 to L^∞ estimate in Theorem 3.43);

(ii) If both u and $-u$ belong to $DG_+(\Omega)$, then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R})$.

Let us start by discussing the first point. We define these two quantities, of crucial importance:

$$U(h,\rho) := \int_{A(h,\rho)} (u-h)^2\,dy, \qquad V(h,\rho) := \mathscr{L}^n\left(A(h,\rho)\right).$$

Theorem 3.41. *The following properties hold:*

(i) both U and V are non-decreasing functions of ρ, and non-increasing functions of h;

(ii) for any $h > k$ and $0 < \rho < R$ the following inequalities hold:

$$\begin{aligned} V(h,\rho) &\le \frac{1}{(h-k)^2}\,U(k,\rho), \\ U(k,\rho) &\le \frac{c\cdot c^+_{DG}(u)}{(R-\rho)^2}\,U(k,R)V^{2/n}(k,\rho), \end{aligned}$$

with $c = c(n)$.

Proof. The first statement and the first inequality in the second statement are trivial, since for the latter one

$$\begin{aligned} (h-k)^2 V(h,\rho) &= \int_{A(h,\rho)} (h-k)^2\,dy \le \int_{A(h,\rho)} (u-k)^2\,dy \\ &\le \int_{A(k,\rho)} (u-k)^2\,dy = U(k,\rho)\ . \end{aligned}$$

For the second inequality, let us introduce a Lipschitz cutoff function η satisfying $0 \le \eta \le 1$ at all points and supported in $B_{(R+\rho)/2}(x)$ with $\eta \equiv 1$ on $\overline{B}_\rho(x)$ and $|\nabla\eta| \le 4/(R-\rho)$. We need to note that

$$\int_{B_{(R+\rho)/2}(x)} \eta^2 |\nabla(u-k)^+|^2\,dy \le \frac{4c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)} (u-k)^2\,dy$$

and

$$\int_{B_{(R+\rho)/2}(x)} ((u-k)^+)^2 |\nabla\eta|^2\,dy \le \frac{16}{(R-\rho)^2}\int_{A(k,R)} (u-k)^2\,dy.$$

Combining these two inequalities, since $c^+_{DG}(u) \ge 1$ by its very definition, we get

$$\int_{B_{(R+\rho)/2}(x)} |\nabla(\eta(u-k)^+)|^2\,dy \le \frac{40c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)} (u-k)^2\,dy$$

and by the Sobolev embedding inequality for the function $\eta(u-k)^+$ this implies

$$\left(\int_{A(k,\rho)} (u-k)^{2^*}\,dy\right)^{2/2^*} \le \frac{c\cdot c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)} (u-k)^2\,dy$$

for some constant $c=c(n)$ only depending on the dimension n. In order to conclude, we just need to apply Hölder's inequality, for indeed

$$U(k,\rho)=\int_{A(k,\rho)} (u-k)^2\,dy \le \left(\int_{A(k,\rho)} (u-k)^{2^*}\,dy\right)^{2/2^*} V^{2/n}(k,\rho)$$

with $p=2^*/2=n/(n-2)$, $p'=n/2$. □

In proving the quantitative maximum principle that has been alluded to above, we will rather exploit a weaker form of such inequalities, namely

$$\begin{aligned} V(h,\rho) &\le \frac{1}{(h-k)^2}\,U(k,R),\\ U(h,\rho) &\le \frac{c\cdot c^+_{DG}(u)}{(R-\rho)^2}\,U(k,R)V^{2/n}(k,R)\,, \end{aligned}$$

for a constant $c=c(n)$.

We can view these inequalities as joint decay properties of U and V; in order to get the decay of a single scalar quantity, it is convenient to define $\varphi := U^\xi V^\eta$ for some choice of the (positive) real parameters ξ, η to be determined. We obtain:

$$U^\xi(h,\rho)V^\eta(h,\rho) \le \frac{c^\xi_{UV}}{(h-k)^{2\eta}}\frac{1}{(R-\rho)^{2\xi}}U^{\xi+\eta}(k,R)V^{\frac{2\xi}{n}}(k,R).$$

where $c_{UV}=c\cdot c^+_{DG}(u)$, a convention that will be systematically adopted in the sequel of this section. To the scope of determining some decay inequality for φ, we look for solutions (θ,ξ,η) to the system

$$\xi+\eta=\theta\xi, \qquad \frac{2\xi}{n}=\theta\eta.$$

Setting $\eta=1$ (by homogeneity this choice is not restrictive), we get $\xi=n\theta/2$ and we can use the first equation to conclude that

$$\theta=\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{2}{n}}. \tag{3.48}$$

Note that $\theta>1$: this fact will play a crucial role in the following proof. In any case, we get the decay relation

$$\varphi(h,\rho) \le \frac{c^\xi_{UV}}{(h-k)^2}\frac{1}{(R-\rho)^{2\xi}}\varphi^\theta(k,R).$$

Theorem 3.42. *Let $u \in DG_+(\Omega)$, $B_{R_0}(x) \Subset \Omega$. For any $h_0 \in \mathbb{R}$ there exists $d = d(h_0, R_0, c^+_{DG}(u))$ such that*

$$u \leq h_0 + d$$

$\mathscr{L}^n$-a.e. on $B_{R_0/2}(x)$, where

$$d^2 = p(n)[c^+_{DG}(u)]^{n\theta/2} \frac{\varphi(h_0, R_0)^{\theta-1}}{R_0^{n\theta}},$$

with the constant $p(n)$ depending only on the dimension n.

Corollary 3.43 (L^2 to L^∞ estimate)**.** *If $u \in DG_+(\Omega)$, then for any ball $B_{R_0}(x) \Subset \Omega$ and for any $h_0 \in \mathbb{R}$*

$$\operatorname*{ess\,sup}_{B_{R_0/2}(x)} u \leq h_0 + (\omega_n p(n))^{\frac{1}{2}} [c^+_{DG}(u)]^{n\theta/4} \left(\frac{U(h_0, R_0)}{\omega_n R_0^n}\right)^{\frac{1}{2}} \left(\frac{V(h_0, R_0)}{R_0^n}\right)^{\frac{\theta-1}{2}}.$$

Proof. This corollary comes immediately from Theorem 3.42, once we express φ in terms of U and V and recall that $\xi + 1 = \theta\xi$ (that is $\xi(\theta - 1) = 1$), by means of simple algebraic computations. □

Remark 3.44. From Corollary 3.43 with $h_0 = 0$, we can get the *maximum principle* for u, as anticipated above. Indeed

$$\operatorname*{ess\,sup}_{B_{R_0/2}(x)} (u^+)^2 \leq q(n)[c^+_{DG}(u)]^{n\theta/2} \fint_{B_{R_0}(x)} u^2 \, dy$$

with $q(n) = p(n)\omega_n^\theta$.

We are now ready to prove Theorem 3.42, the first main result of this section.

Proof. Define $h_p := h_0 + d - d/2^p$ and $R_p := R_0/2 + R_0/2^{p+1}$, so that $h_p \uparrow (h_0 + d)$ while $R_p \downarrow R_0/2$. Here $d \in \mathbb{R}$ is a parameter to be fixed in the sequel of the proof. From the decay inequality for φ we get

$$\varphi(h_{p+1}, R_{p+1}) \leq \varphi(h_p, R_p) \left[\varphi(h_p, R_p)^{\theta-1} c^\xi_{UV} \left(\frac{2^{p+2}}{R_0}\right)^{2\xi} \left(\frac{2^{p+1}}{d}\right)^2\right]$$

and letting $\psi_p := 2^{\mu p}\varphi(h_p, R_p)$ this becomes

$$\psi_{p+1} \leq \psi_p \left[2^\mu c^\xi_{UV} 2^{4\xi+2} 2^{p(2\xi+2)} R_0^{-2\xi} d^{-2} 2^{-\mu p(\theta-1)} \psi_p^{\theta-1}\right].$$

This is true for any $\mu \in \mathbb{R}$ but we fix it so that $(2\xi+2) = \mu(\theta-1)$, leading to a cancellation of two factors in the previous inequality. Having chosen μ, if we require d to satisfy

$$2^{\mu} c_{UV}^{\xi} 2^{4\xi+2} \psi_0^{\theta-1} R_0^{-2\xi} d^{-2} = 1$$

then we get at once $\psi_1 \leq \psi_0$. Hence, $2^{\mu} c_{UV}^{\xi} 2^{4\xi+2} \psi_1^{\theta-1} R_0^{-2\xi} d^{-2} \leq 1$ and the decay inequality yields $\psi_2 \leq \psi_1$. By induction, it follows that $\psi_p \leq \psi_0$ for all $p \in \mathbb{N}$. Therefore, we get $\varphi(h_p, R_p) \leq 2^{-\mu p} \varphi(h_0, R_0) \to 0$ and, hence, by the monotonicity properties of φ we derive

$$\varphi(h_0 + d, R_0/2) \leq \varphi(h_p, R_0/2) \leq \varphi(h_p, R_p) \to 0,$$

as it was claimed. But notice that the previous condition on d is satisfied if we set

$$d^2 = p(n)[c_{DG}^{+}(u)]^{n\theta/2} R_0^{-2\xi} \psi_0^{\theta-1}$$

thus the conclusion follows. □

We can now discuss the notion of *oscillation*, which will be crucial for the conclusion of De Giorgi's argument.

Definition 3.45. *Let $\Omega \subset \mathbb{R}^n$ be an open set, $B_r(x) \subset \Omega$ relatively compact and $u : \Omega \to \mathbb{R}$ a measurable function, which we assume to be locally bounded. We define its* oscillation *on $B_r(x)$ as*

$$\omega(B_r(x))(u) := \operatorname*{ess\,sup}_{B_r(x)} u - \operatorname*{ess\,inf}_{B_r(x)} u$$

When no confusion arises, we will omit the explicit dependence on the center of the ball, thus writing $\omega(r)$ in lieu of $\omega(B_r(x))$.

It is a direct consequence of the previous results that if $u \in DG_+(\Omega) \cap DG_-(\Omega)$, then

$$\operatorname*{ess\,sup}_{B_{r/2}(x)} u \leq c \left(\fint_{B_r(x)} u^2 \, dy \right)^{\frac{1}{2}}, \qquad -\operatorname*{ess\,inf}_{B_{r/2}(x)} u \leq c \left(\fint_{B_r(x)} u^2 \, dy \right)^{\frac{1}{2}}$$

for a constant c which is a function of the dimension n and of $c_{DG}(u)$. Here and in the sequel of our discussion we shall denote by $c_{DG}(u)$ the largest number between $c_{DG}^{+}(u)$ and $c_{DG}^{-}(u)$ and by $DG(\Omega)$ the intersection of the spaces $DG_+(\Omega)$ and $DG_-(\Omega)$.

Summing up what we have achieved so far, under these assumptions we get

$$\omega(B_{r/2}(x))(u) \leq 2c \left(\fint_{B_r(x)} u^2 \, dy \right)^{1/2}.$$

Let us now discuss the relation between the decay of the oscillation of u and its Hölder regularity. We prove this result passing through the theory of Campanato spaces.

Theorem 3.46. *Let $\Omega \subset \mathbb{R}^n$ be open, $\kappa \geq 0$, $\alpha \in (0,1]$ and let $u : \Omega \to \mathbb{R}$ be a measurable function such that for any $B_r(x) \subset \Omega$ we have $\omega(B_r(x)) \leq \kappa r^\alpha$. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R})$, that is, there exists in the Lebesgue equivalence class of u a $C^{0,\alpha}_{\mathrm{loc}}$ representative.*

Proof. By definition of essential extrema, for $\mathscr{L}^n$-a.e. $y \in B_r(x)$ we have that $\operatorname{ess\,inf}_{B_r(x)} u \leq u(y) \leq \operatorname{ess\,sup}_{B_r(x)} u$. If we denote, as above, by $u_{x,r}$ the average value of u over the ball $B_r(x)$, these inequalities imply at once that $\operatorname{ess\,inf}_{B_r(x)} u \leq u_{x,r} \leq \operatorname{ess\,sup}_{B_r(x)}$ and hence the inequality $|u - u_{x,r}| \leq \kappa r^\alpha$ is satisfied $\mathscr{L}^n$-a.e. in $B_r(x)$. Thereby, we have proved that $u \in \mathcal{L}^{2,n+2\alpha}(\Omega;\mathbb{R})$, but this gives $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R})$ (regularity is local since no assumption is made on Ω), which is the claim. $\square$

This theorem motivates our interest in the study of the oscillation of u, that will be carried on by means of some tools recalled in Section A.3.

Broadly speaking, De Giorgi's proof of Hölder continuity is geometric in spirit and ultimately based on the isoperimetric inequality. Notice that, as discussed in Section A.3, the isoperimetric inequality is also underlying the Sobolev inequalities, which we used in the proof of the sup estimate for functions in $DG_+(\Omega)$. To proceed further, we first need the following lemma.

Lemma 3.47 (Decay of V)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in DG_+(\Omega)$. Suppose that $B_{2r}(x) \Subset \Omega$ and $k_0 \leq \operatorname{ess\,sup}_{B_{2r}}(u) \leq M$ is such that*

$$V(k_0, r) \leq \frac{1}{2}\mathscr{L}^n(B_r(x)), \tag{3.49}$$

then the sequence of levels $k_\nu = M - (M - k_0)/2^\nu$ for $\nu \geq 0$ satisfies

$$\left(\frac{V(k_\nu, r)}{r^n}\right)^{2/1^*} \leq c\,\frac{c^+_{DG}(u)}{\nu},$$

for some constant $c = c(n)$.

Proof. Take two levels h, k such that $M \geq h \geq k \geq k_0$ and define $\overline{u} := u \wedge h - u \wedge k = (u \wedge h - k)^+$. By construction $\overline{u} \geq 0$ and since $u \in W^{1,1}(\Omega;\mathbb{R})$ we also have $\overline{u} \in W^{1,1}(\Omega;\mathbb{R})$. It is also clear that $\nabla \overline{u} \neq 0$ only on $A(k,r) \setminus A(h,r)$. Notice that (3.49) gives

$$\mathscr{L}^n(\{\overline{u} = 0\} \cap B_r(x)) \geq \mathscr{L}^n(\{u \leq k\} \cap B_r(x)) \geq \mathscr{L}^n(\{u \leq k_0\} \cap B_r(x)) \geq \frac{1}{2}\mathscr{L}^n(B_r(x))$$

and so we can apply the relative version of the critical Sobolev embedding (Theorem

A.28) and Hölder's inequality to get

$$\begin{aligned}(h-k)^{1^*}\mathscr{L}^n(A(h,r)) &= \int_{A(h,r)} \overline{u}^{1^*}\,dy \le c\left(\int_{B_r(x)} |\nabla \overline{u}|\,dy\right)^{1^*} \\ &= c\left(\int_{A(k,r)\setminus A(h,r)} |\nabla u|\,dy\right)^{1^*} \\ &\le c\left(\int_{A(k,r)} |\nabla u|^2\,dy\right)^{1^*/2} \left(\mathscr{L}^n(A(k,r)\setminus A(h,r))\right)^{1^*/2}\end{aligned}$$

for some $c = c(n)$. We can now exploit the De Giorgi property of u, namely the fact that

$$\int_{A(k,r)} |\nabla u|^2\,dy \le \frac{c^+_{DG}(u)}{r^2}\int_{B_{2r}(x)} (u-k)^2\,dy \le (M-k)^2\omega_n c^+_{DG}(u) r^{n-2}$$

in order to obtain

$$(h-k)^2 V(h,r)^{2/1^*} \le c\,\omega_n c^+_{DG}(u)(M-k)^2 r^{n-2}(V(k,r)-V(h,r)). \tag{3.50}$$

At this stage we can conclude the proof by applying (3.50) for $h = k_{i+1}$ and $k = k_i$, so that

$$\begin{aligned}\nu V(k_\nu,r)^{2/1^*} &\le \sum_{i=1}^{\nu} V(k_i,r)^{2/1^*} \\ &\le 4c\omega_n c^+_{DG}(u) r^{n-2}\sum_{i=1}^{\nu}\left[V(k_i,r)-V(k_{i+1},r)\right] \\ &\le 4c\omega_n^2 c^+_{DG}(u) r^{2n-2}.\end{aligned}$$

□

Theorem 3.48 ($C^{0,\alpha}$ regularity)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in DG(\Omega)$. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega;\mathbb{R})$, with $2\alpha = -\log_2\left(1-2^{-(\nu+2)}\right)$, where*

$$\nu = c\left[c_{DG}(u)\right]^{\frac{n\theta-1}{\theta-1}} \tag{3.51}$$

for $c = c(n)$ and $\theta > 1$ given by (3.48), the only positive solution to the quadratic equation $n\theta(\theta-1) = 2$.

Proof. Pick an $R > 0$ such that $B_{2R}(x) \Subset \Omega$ and consider for any $r \le R$ the functions $m(r) := \operatorname{ess\,inf}_{B_r(x)}(u)$ and $M(r) := \operatorname{ess\,sup}_{B_r(x)}(u)$. Moreover, set $\omega(r) = M(r) - m(r)$ and $\mu(r) := (m(r)+M(r))/2$. We apply the previous lemma to the sequence defined by

$k_\nu := M(2r) - \frac{\omega(2r)}{2^{\nu+1}}$ (so that $k_0 = \mu(2r)$), but to do this we need to check the hypothesis (3.49), which means

$$\mathscr{L}^n(\{u > \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x)).$$

However, it is certainly the case that either $\mathscr{L}^n(\{u > \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x))$ or $\mathscr{L}^n(\{u < \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x))$. We will proceed assuming the first alternative occurs, however the second case is analogous, provided we work with $-u$ instead of u. Notice that it is precisely here that we need the assumption that both u and $-u$ belong to $DG_+(\Omega)$.
Using Lemma 3.47 it is easily seen that the choice of ν as in (3.51), with $c = c(n)$ large enough (and chosen so that ν is a positive integer), provides

$$(\omega_n p(n))^{1/2}\left[c^+_{DG}(u)\right]^{n\theta/4}\left(\frac{V(k_\nu, r)}{r^n}\right)^{(\theta-1)/2} \leq \frac{1}{2}.$$

Moreover, notice that this choice of ν has been made independently of r and R (this is crucial for the validity of the scheme below). Now apply the maximum principle, as encoded in Corollary 3.43, to u with radii $r/2$ and r and $h_0 = M(2r) - \frac{\omega(2r)}{2^{\nu+1}} = k_\nu$ (for the previous choice of ν) to obtain

$$M\left(\frac{r}{2}\right) \leq h_0 + (\omega_n p(n))^{1/2}\left[c^+_{DG}(u)\right]^{n\theta/4}(M(2r) - h_0)\left(\frac{V(h_0, r)}{r^n}\right)^{(\theta-1)/2}$$

and, by the appropriate choice of ν that has been described, we deduce

$$M\left(\frac{r}{2}\right) \leq h_0 + \frac{M(2r) - h_0}{2} = \frac{M(2r) + h_0}{2} = M(2r) - \frac{\omega(2r)}{2^{\nu+2}}.$$

If we subtract the essential minimum $m(2r)$ and use $m(r/2) \geq m(2r)$ we finally get

$$\omega\left(\frac{r}{2}\right) \leq \omega(2r)\left(1 - \frac{1}{2^{\nu+2}}\right)$$

which is the desired decay estimate. By the standard iteration argument, as per Lemma 3.12 with the obvious changes, we find

$$\omega(r) \leq 4^\alpha \omega(R)\left(\frac{r}{R}\right)^\alpha \qquad 0 < r \leq R$$

for $2\alpha = -\log_2\left(1 - 2^{-(\nu+2)}\right)$ and the conclusion follows from Theorem 3.46. $\square$

4 Regularity for systems

In the last section of the previous chapter we presented De Giorgi's regularity result for solutions $u \in H^1(\Omega;\mathbb{R})$ of the *scalar* elliptic problem

$$\sum_{\alpha,\beta} \partial_{x_\alpha}(A^{\alpha\beta}(x)\partial_{x_\beta}u(x)) = 0$$

with bounded Borel coefficients $A^{\alpha\beta}$ satisfying $\lambda I \leq A \leq \Lambda I$: we proved that in fact $u \in C^{0,\vartheta}_{\mathrm{loc}}(\Omega;\mathbb{R})$, with $\vartheta = \vartheta(n,\lambda,\Lambda)$.

It is natural to investigate similar regularity properties for systems, still under no regularity assumption on A (for otherwise Schauder theory would be applicable). In [29], dating back to 1968, Ennio De Giorgi provided a counterexample showing that the scalar case is somewhat special: he obtained a surprisingly elementary example of a solution to an elliptic problem, actually the unique minimizer of a convex variational problem, which fails to be Hölder continuous.

This is the object of the first section of this chapter, and we will then proceed from there to the discussion of various *partial* regularity results for local minimizers of suitable elliptic functionals.

4.1 De Giorgi's counterexample to regularity for systems

When $m = n$, consider

$$u(x) := x|x|^\gamma. \tag{4.1}$$

We will prove (cf. equations (4.7), (4.8) and (4.9) below) that, choosing

$$\gamma = -\frac{n}{2}\left(1 - \frac{1}{\sqrt{(2n-2)^2+1}}\right), \tag{4.2}$$

the function u is the unique solution of the Euler-Lagrange equation associated with the uniformly convex functional

$$\mathcal{L}(u) := \frac{1}{2}\int_{B_1}\left[\left((n-2)\sum_i \partial_{x_i}u^i + n\sum_{i,j}\frac{x_ix_j}{|x|^2}\partial_{x_j}u^i(x)\right)^2 + |\nabla u(x)|^2\right]dx. \tag{4.3}$$

If $n \geq 3$ then $|u| \notin L^\infty(B_1;\mathbb{R})$, because

$$-\gamma = \frac{n}{2}\left(1 - \frac{1}{\sqrt{(2n-2)^2+1}}\right) \geq \frac{3}{2}\left(1 - \frac{1}{\sqrt{17}}\right) > 1$$

and this provides a counterexample not only to Hölder regularity, but also to local boundedness of solutions. In the case $n = m = 2$ we already know from Widman's technique

(see Remark 2.4) that any solution u to a uniformly elliptic problem in divergence form, as per equation (2.1) with $f = 0, F = 0$, is locally Hölder continuous, nevertheless De Giorgi's example shows that this regularity cannot be improved to local Lipschitz continuity.

Let us remark that the matrix $A(x)$ (that is uniquely characterized by the equation $\mathcal{L}(u) = \int_{B_1} \sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta}(x)\partial_{x_\alpha} u^i \partial_{x_\beta} u^j \, dx$) is smooth away from the origin, where it has a discontinuity determined by the term $x \otimes x/|x|^2$.

Remark 4.1. In this sole section, for the sake of notational convenience, we employ Latin letters to denote *repeated indices* in the domain and in the target space as exemplified above by the divergence operator which we wrote in the form $\sum_i \partial_{x_i} u^i$ in lieu of $\sum_{\alpha,i} \delta_i^\alpha \partial_{x_\alpha} u^i$ as our general conventions would impose. This choice allows a substantial simplification of the formulae we are about to present.

The Euler-Lagrange equations associated to (4.3) are as follows: for every $h = 1, \dots, n$ it must be

$$
\begin{aligned}
0 &= (n-2)\partial_{x_h}\left((n-2)\sum_{t=1}^n \partial_{x_t} u^t + n\sum_{s,t=1}^n \frac{x_s x_t}{|x|^2}\partial_{x_s} u^t\right) && (4.4)\\
&+ n\sum_{k=1}^n \partial_{x_k}\left[\frac{x_h x_k}{|x|^2}\left((n-2)\sum_{t=1}^n \partial_{x_t} u^t + n\sum_{s,t=1}^n \frac{x_s x_t}{|x|^2}\partial_{x_s} u^t\right)\right] && (4.5)\\
&+ \sum_{k=1}^n \partial^2_{x_k x_k} u^h \,. && (4.6)
\end{aligned}
$$

We are now going to prove that u is the unique minimizer of $\mathcal{L}$, with respect to the boundary data coinciding with the values attained by u itself on ∂B_1, and therefore that u solves the Euler-Lagrange equations presented above. More precisely, we proceed according to the following two steps:

(i) u, as defined by (4.1), belongs to $C^\infty(B_1 \setminus \{0\}; \mathbb{R}^n)$ and solves in $B_1 \setminus \{0\}$ the Euler-Lagrange equations in the classical sense;

(ii) $u \in H^1(B_1; \mathbb{R}^n)$ and is also a weak solution in B_1 of the system in question.

Let us perform step (i). Fix $h \in \{1, \dots, n\}$: using the elementary identity

$$\partial_{x_h}|x|^\gamma = \gamma x_h |x|^{\gamma-2}$$

we get at once $\Delta|x|^\gamma = \gamma(n+\gamma-2)|x|^{\gamma-2}$ and hence

$$\Delta\left(x_h|x|^\gamma\right) = x_h\Delta|x|^\gamma + 2\partial_{x_h}|x|^\gamma = (\gamma n + \gamma^2)x_h|x|^{\gamma-2}\,. \tag{4.7}$$

This is what we need to compute the third line in the above expression, cf. (4.6), for u given by (4.1). Instead, concerning both (4.4) and (4.5) we have to calculate

$$\sum_{t=1}^{n} \partial_{x_t} u^t = \sum_{t=1}^{n} \partial_{x_t} \left(x_t |x|^\gamma \right) = (n+\gamma)|x|^\gamma \ ,$$

and

$$\sum_{s,t=1}^{n} \frac{x_s x_t}{|x|^2} \partial_{x_s} u^t = \sum_{s,t=1}^{n} \frac{x_s x_t}{|x|^2} \left(\gamma x_s x_t |x|^{\gamma-2} + \delta_{st} |x|^\gamma \right) = (\gamma+1)|x|^\gamma.$$

At that stage, it is readily checked that (4.4) is given by

$$(n-2)\partial_{x_h}\left((n-2)\sum_{t=1}^{n} \partial_{x_t} u^t + n \sum_{s,t=1}^{n} \frac{x_s x_t}{|x|^2} \partial_{x_s} u^t\right) = \gamma(n-2)[(n-2)(n+\gamma)+n(\gamma+1)]x_h|x|^{\gamma-2}. \tag{4.8}$$

In order to compute the term (4.5) we further need

$$\sum_{k=1}^{n} \partial_{x_k} \left(x_h x_k |x|^{\gamma-2} \right) = (n+\gamma-1) x_h |x|^{\gamma-2},$$

and then we can conclude

$$n \sum_{k=1}^{n} \partial_{x_k} \left[\frac{x_h x_k}{|x|^2} \left((n-2) \sum_{t=1}^{n} \partial_{x_t} u^t + n \sum_{s,t=1}^{n} \frac{x_s x_t}{|x|^2} \partial_{x_s} u^t \right) \right] = n(n+\gamma-1)[(n-2)(n+\gamma)+n(\gamma+1)]x_h|x|^{\gamma-2} \ . \tag{4.9}$$

Combining together (4.7), (4.8) and (4.9), we see that $u(x) = x|x|^\gamma$ solves the Euler-Lagrange equation if and only if

$$(2n-2)^2 \left(\gamma + \frac{n}{2}\right)^2 + \gamma n + \gamma^2 = 0 \ ,$$

which leads to the choice (4.2) of γ.

Let us now discuss step (ii), checking first that $u \in H^1(B_1;\mathbb{R}^n)$. As $|\nabla u(x)| \sim |x|^\gamma$ and $2\gamma > -n$, it is easy to show that $|\nabla u| \in L^2(B_1;\mathbb{R})$. Moreover, for every $\varphi \in C_c^\infty(B_1 \setminus \{0\};\mathbb{R})$, by the classical integration by parts formula, we have that

$$\int_{B_1} \partial_{x_\alpha} u^i(x)\varphi(x)\,dx = -\int_{B_1} u^i(x)\partial_{x_\alpha}\varphi(x)\,dx \tag{4.10}$$

for any choice of the indices α and i. Thanks to Lemma 4.2 below, we are allowed to approximate in H^1-norm every $\varphi \in C_c^\infty(B_1;\mathbb{R})$ with a sequence $(\varphi_k) \subset C_c^\infty(B_1 \setminus \{0\};\mathbb{R})$.

Then we can pass to the limit in (4.10) (because $|\nabla u| \in L^2(B_1;\mathbb{R})$) to conclude that indeed $u \in H^1(B_1;\mathbb{R}^n)$. Now, using the fact that the Euler-Lagrange equation holds in the weak sense in $B_1 \setminus \{0\}$ (because it holds in the classical sense), we have

$$\int_{B_1} \sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta}(x)\partial_{x_\alpha} u^i(x)\partial_{x_\beta}\varphi^j(x)\,dx = 0 \tag{4.11}$$

for every $\varphi \in C_c^\infty(B_1 \setminus \{0\};\mathbb{R}^n)$. Using Lemma 4.2 again, we can extend (4.11) to every $\varphi \in C_c^\infty(B_1;\mathbb{R}^n)$, thus obtaining that u satisfies the equation in the weak sense over the whole ball.

Finally, since the functional $\mathcal{L}$ in (4.3) is convex, the Euler-Lagrange equation is satisfied by u if and only if u is a minimizer of $\mathcal{L}(u)$ with boundary condition

$$u(x) = x \quad \text{in } \partial B_1.$$

Thus the function u is not only a solution of the Euler-Lagrange equation associated to the functional $\mathcal{L}$ defined by (4.3), but also a minimizer of the same functional for fixed boundary data, as claimed.

Lemma 4.2. *Let $m \geq 1$ and $n > 2$. For every $\varphi \in C_c^\infty(B_1;\mathbb{R}^m)$ there exists a sequence (φ_k) of functions belonging to $C_c^\infty(B_1 \setminus \{0\};\mathbb{R}^m)$ and converging to φ strongly in $H^1(B_1;\mathbb{R}^m)$.*

Proof. Consider $\psi \in C_c^\infty(\mathbb{R}^n;\mathbb{R})$ with $\psi \equiv 1$ on $\overline{B}_1$, then rescale ψ setting $\psi_k(x) := \psi(kx)$. Set $\varphi_k := \varphi(1-\psi_k)$; in the L^2-topology we have $\varphi - \varphi_k = \varphi\psi_k \to 0$ and $(\nabla\varphi)\psi_k \to 0$. Since

$$\nabla(\varphi - \varphi_k) = (\nabla\varphi)\psi_k + \varphi\nabla\psi_k,$$

the claim follows from verifying that

$$\int_{B_1} \varphi^2(x)|\nabla\psi_k(x)|^2\,dx \to 0.$$

Indeed

$$\begin{aligned}\int_{B_1} \varphi^2(x)|\nabla\psi_k(x)|^2\,dx &\leq (\sup\varphi^2)k^2\int_{B_1}|\nabla\psi(kx)|^2\,dx \\ &\leq (\sup\varphi^2)k^{2-n}\int_{\mathbb{R}^n}|\nabla\psi(x)|^2\,dx \longrightarrow 0\ ,\end{aligned}$$

thanks to the fact that $n > 2$. This completes the proof. □

We conclude noticing that the restriction $n \geq 3$ in the proof of Lemma 4.2 is not really needed. Indeed, when $n = 2$ we have

$$\inf \left\{ \int_{B_1} |\nabla \psi(x)|^2 \, dx : \psi \in C_c^\infty(B_1; \mathbb{R}), \, \psi = 1 \text{ in a neighborhood of } 0 \right\} = 0. \tag{4.12}$$

Let us prove (4.12): we first observe that

$$\inf \left\{ \int_0^1 r|a'(r)|^2 \, dr \ : \ a(0) = 1, \, a(1) = 0 \right\} = 0,$$

because, for any $\gamma > 0$, one can take $a_\gamma(r) := 1 - r^\gamma$, so that

$$\int_0^1 r|a_\gamma'(r)|^2 \, dr = \frac{\gamma}{2} \xrightarrow{\gamma \to 0} 0 \ .$$

Then, equation (4.12) is justified considering suitable approximations of a_γ, for instance $\min\{1 - r^\gamma, 1 - \gamma\}/(1 - \gamma)$ and their mollifications (which are equal to 1 in a neighborhood of 0).

In a more general perspective, let us recall that the *p-capacity* of a compact set $K \subset \mathbb{R}^n$ is defined by

$$\inf \left\{ \int_{\mathbb{R}^n} |\nabla \phi|^p \, dx : \phi \in C_c^\infty(\mathbb{R}^n; \mathbb{R}), \ \phi \equiv 1 \text{ in a neighborhood of } K \right\}$$

and thus we proved that singletons have null 2-capacity in $\mathbb{R}^n$ for $n \geq 2$.

Using (4.12) to remove the point singularity also in the case $n = 2$, it follows that the functional $\mathcal{L}$ defined by (4.3) and its minimizer u defined by (4.1) with $\gamma = -(1 - 1/\sqrt{5})$ are a counterexample to Lipschitz regularity of minimizers of uniformly convex variational problems.

4.2 Partial regularity for systems: basic definitions and ancillary results

As we have seen with De Giorgi's counterexample, it is impossible to expect an "everywhere" regularity result for weak solutions to elliptic systems: however we can pursue a different goal, a "partial" regularity result, namely we can aim at proving a suitable degree of regularity of the solution away from a small singular set. This strategy dates back to De Giorgi himself, and it was implemented for the first time in the study of minimal surfaces, see [28].

Definition 4.3 (Regular and singular sets)**.** *For a function $u : \Omega \to \mathbb{R}^m$ we call* regular set *of u the set*

$$\Omega_{\text{reg}}(u) := \left\{ x \in \Omega : \exists\, r > 0 \text{ s.t. } B_r(x) \subset \Omega \text{ and } u \in C^1\left(B_r(x); \mathbb{R}^m\right) \right\}.$$

Correspondingly, the singular set *is*

$$\Sigma(u) := \Omega \setminus \Omega_{\text{reg}}(u).$$

Notice that the in previous definition we actually mean that $x \in \Omega_{\text{reg}}(u)$ if the function u has a C^1 representative in $B_r(x)$, consistently with the fact that we are always tacitly working with equivalence classes. The set $\Omega_{\text{reg}}(u)$ is then obviously the largest open subset A of Ω such that u coincides $\mathscr{L}^n$-a.e. in A with a C^1 function v.

Let us summarize here, specified for elliptic systems, some results that we have already presented in the introductory discussion of Chapter 3:

(a) If we are looking at the problem from the variational point of view, studying local minimizers $u \in H^1_{\text{loc}}(\Omega; \mathbb{R}^m)$ of $v \mapsto \int_\Omega L(\nabla v)\, dx$, with $L \in C^2(\mathbb{R}^{m\times n}; \mathbb{R})$ and $|\nabla^2 L(p)| \le \Lambda$, we have shown that any such u satisfies the Euler-Lagrange equations associated with the functional in question. More precisely, if

$$\int_{\Omega'} L(\nabla u(x))\, dx \le \int_{\Omega'} L(\nabla v(x))\, dx \qquad \forall\, v \text{ s.t. } \{u \neq v\} \Subset \Omega' \Subset \Omega,$$

then

$$\partial_{x_\alpha}\left(\partial_{p_i^\alpha} L(\nabla u)\right) = 0 \qquad \forall\, i = 1, \dots, m.$$

(b) If $\nabla^2 L$ satisfies a uniform Legendre condition for some $\lambda > 0$, following Nirenberg's argument (cf. Section 2.3) we have $\nabla u \in H^1_{\text{loc}}(\Omega; \mathbb{R}^{m\times n})$ and (by differentiation of the Euler-Lagrange equations with respect to x_γ)

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(\partial_{p_i^\alpha p_j^\beta} L(\nabla u) \partial^2_{x_\beta x_\gamma} u^j\right) = 0 \quad \forall\, i = 1, \dots, m, \gamma = 1, \dots, n. \tag{4.13}$$

To proceed further with our discussion, we first need to introduce an important concept.

Definition 4.4 (Uniform quasiconvexity)**.** *A continuous function $L : \mathbb{R}^{m\times n} \to \mathbb{R}$ is said to be λ-uniformly quasiconvex at $A \in \mathbb{R}^{m\times n}$ if, for all $\Omega \subset \mathbb{R}^n$ open and bounded, it holds*

$$\fint_\Omega \left(L(A + \nabla\varphi(x)) - L(A)\right) dx \ge \lambda \fint_\Omega |\nabla\varphi|^2\, dx \qquad \forall\, \varphi \in C_c^\infty(\Omega; \mathbb{R}^m).$$

We say that L is λ-uniformly quasiconvex if it is λ-uniformly quasiconvex at every point $A \in \mathbb{R}^{m\times n}$.

For $\lambda > 0$ this is obviously a strengthening of the condition given in Definition 1.12.

Remark 4.5. Notice that L is λ-uniformly quasiconvex if and only if $L(p) - \lambda|p|^2$ is quasiconvex. In this case, $(\nabla^2 L - 2\lambda I)$ satisfies the Legendre-Hadamard condition with parameter 0 or, equivalently, $\nabla^2 L$ satisfies the Legendre-Hadamard condition with parameter λ.

In this section we shall provide a fairly complete proof of the following result, following with minor variants the original argument in [34].

Theorem 4.6 (Evans, [34])**.** *If $L \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$ is λ-uniformly quasiconvex with $\lambda > 0$ and satisfies*

$$|\nabla^2 L(p)| \le \Lambda \qquad \forall p \in \mathbb{R}^{m\times n}, \tag{4.14}$$

for some $\Lambda > 0$, then any local minimizer u belongs to $C^{1,\gamma}(\Omega_{\rm reg};\mathbb{R}^m)$ for some $\gamma = \gamma(n, m, \lambda, \Lambda)$ and

$$\mathscr{L}^n(\Omega \setminus \Omega_{\rm reg}) = 0.$$

At this stage we should point out that the growth condition (4.14) is a bit restrictive if we want to allow certain standard examples of quasiconvex functions, e.g. convex functions of determinants of minors of ∇u. This issue arises, for instance, if one considers the high-dimensional counterparts of the functional (considered for $m = n = 2$, and which *does satisfy* the assumptions of Theorem 4.6)

$$L(\nabla u) := |\nabla u|^2 + \sqrt{1 + \det(\nabla u)^2},$$

namely generalizations of the form

$$L(\nabla u) := |\nabla u|^2 + \sqrt{1 + \sum_M (M\nabla u)^2},$$

where $M\nabla u$ denotes a 2×2 minor of ∇u.

A more general growth condition considered in [34], and motivated by such examples, is

$$|\nabla^2 L(p)| \le c\left(1 + |p|^{q-2}\right) \qquad \text{with } q \ge 2, \tag{4.15}$$

which leads to the estimates $|\nabla L(p)| \le c(1 + |p|^{q-1})$ and $|L(p)| \le c(1 + |p|^q)$.

Before presenting the proof of Theorem 4.6, let us give a short list of significant results concerning the regularity and the size of the singular set for local minimizers (in $H^1_{\rm loc}$) of variational problems under the general assumptions that $L \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$ and $|\nabla^2 L(p)| \le \Lambda$ uniformly in the domain under consideration:

(i) If $\nabla^2 L \geq \lambda I$ for some $\lambda > 0$, then Giaquinta and Giusti (see [42] and [43]) proved a much stronger estimate on the size of the singular set, namely (here $\mathscr{H}^k$ denotes the Hausdorff measure, see Appendix C)

$$\mathscr{H}^{n-2+\varepsilon}\left(\Sigma(u)\right) = 0 \qquad \forall\, \varepsilon > 0.$$

(ii) If $\nabla^2 L \geq \lambda I$ for some $\lambda > 0$ and it is globally uniformly continuous, then we have even $\mathscr{H}^{n-2}\left(\Sigma(u)\right) = 0$.

(iii) If, in the setting of Evans' theorem, u is assumed to be locally Lipschitz (but with no extra hypotheses on L besides the general ones), then Kristensen and Mingione proved in [64] that there exists $\delta > 0$ such that

$$\mathscr{H}^{n-\delta}\left(\Sigma(u)\right) = 0.$$

(iv) On the contrary, when $n = 2$ and $m = 3$, there exists a Lipschitz solution u for the system $\sum_\alpha \partial_{x_\alpha}\left(\partial_{p_i^\alpha} L(\nabla u)\right) = 0,\ i = 1, 2, 3$ (with L is smooth and satisfies the Legendre-Hadamard condition), provided in [77], such that

$$\Omega_{\text{reg}}(u) = \emptyset.$$

This last result clarifies once and for all that partial regularity can be expected for (local) minimizers only. We will see how local minimality (and not only the validity of the Euler-Lagrange equations) plays a role in the proof of Evans' result.

We will start with a decay lemma relative to constant coefficients operators.

Lemma 4.7. *There exists a positive constant $c_E = c_E(n, m, \lambda, \Lambda)$ such that, for every constant matrix A satisfying the Legendre-Hadamard condition with constant λ as well as the inequality $|A| \leq \Lambda$, any solution $u \in H^1(B_r(x); \mathbb{R}^m)$ of*

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(A_{ij}^{\alpha\beta} \partial_{x_\beta} u^j\right) = 0 \qquad i = 1, \ldots, m \quad \text{in } B_r(x) \tag{4.16}$$

satisfies

$$\fint_{B_{\alpha r}(x)} |\nabla u(y) - (\nabla u)_{x,\alpha r}|^2\, dy \leq c_E \alpha^2 \fint_{B_r(x)} |\nabla u(y) - (\nabla u)_{x,r}|^2\, dy \qquad \forall \alpha \in (0, 1).$$

Proof. First of all, let us observe that $A = A_{ij}^{\alpha\beta}$ being constant one can prove that in fact $u \in H^2(B_r(x); \mathbb{R}^m)$ and each of its partial derivatives satisfies equation (4.16).

That being said, as a consequence of what we showed in Section 2.4 about decay estimates for systems with constant coefficients, considering (2.17) with $\rho = \alpha r$ and $\alpha < 1$, we have that

$$\int_{B_{\alpha r}(x)} |\nabla u(y) - (\nabla u)_{x,\alpha r}|^2 \, dy \leq c_E \left(\frac{\alpha r}{r}\right)^{n+2} \int_{B_r(x)} |\nabla u(y) - (\nabla u)_{x,r}|^2 \, dy. \tag{4.17}$$

It is enough to consider the mean of (4.17), so that

$$\fint_{B_{\alpha r}(x)} |\nabla u(y) - (\nabla u)_{x,\alpha r}|^2 \, dy \leq c_E \alpha^2 \fint_{B_r(x)} |\nabla u(y) - (\nabla u)_{x,r}|^2 \, dy.$$

□

Definition 4.8 (Excess)**.** *For any function $u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ and any ball $B_\rho(x) \Subset \Omega$ the* excess *of u in $B_\rho(x)$ is defined by*

$$\mathrm{Exc}\,(u, B_\rho(x)) := \left(\fint_{B_\rho(x)} |\nabla u(y) - (\nabla u)_{x,\rho}|^2 \, dy \right)^{1/2}.$$

When we consider functions L satisfying the more general growth condition (4.15), then we should modify the definition of excess as follows, see [34]:

$$\mathrm{Exc}\,(u, B_\rho(x))^2 = \fint_{B_\rho(x)} \left(1 + |\nabla u(y) - (\nabla u)_{x,\rho}|^{q-2}\right) |\nabla u(y) - (\nabla u)_{x,\rho}|^2 \; dy.$$

However, in our presentation we will only cover the case $q = 2$.

Remark 4.9 (Properties of the excess)**.** We list here the basic properties of the excess, they are trivial to check.

(i) Any additive perturbation by an affine function $p(x)$ does not change the excess, that is

$$\mathrm{Exc}\,(u + p, B_\rho(x)) = \mathrm{Exc}\,(u, B_\rho(x)).$$

(ii) The excess is positively 1-homogeneous, that is for any number $\lambda \geq 0$

$$\mathrm{Exc}\,(\lambda u, B_\rho(x)) = \lambda \mathrm{Exc}\,(u, B_\rho(x)).$$

(iii) We have the following scaling property: said $u^{(\rho)}(x) = u(\rho x)$ then

$$\mathrm{Exc}\left(\frac{u^{(\rho)}}{\rho}, B_1(0)\right) = \mathrm{Exc}\,(u, B_\rho(0)).$$

Remark 4.10. The name "excess" is inspired by De Giorgi's theory of minimal surfaces, as proposed in [27] and [28], see also [46] for a modern presentation. The excess of a set E at a point is defined (for sufficiently regular sets) by

$$\mathrm{Exc}\,(E, B_\rho(x)) := \fint_{B_\rho(x)\cap\partial E} |\nu_E(y) - \nu_E(x)|^2 \, d\mathscr{H}^{n-1}(y),$$

where ν_E is the inner normal of the set E. The correspondence between $\mathrm{Exc}\,(u, B_\rho(x))$ and $\mathrm{Exc}\,(E, B_\rho(x))$ is easily described assuming the set ∂E to be, in a neighborhood of x, the graph of a function u in a coordinate system where $\nabla u(x') = 0$ for $x = (x', x_n)$. Indeed, the identity $\nu_E = (-\nabla u, 1)/\sqrt{1 + |\nabla u|^2}$ and the area formula for graphs give

$$\int_{B_\rho(x)\cap\partial E} |\nu_E(y) - \nu_E(x)|^2 \, d\mathscr{H}^{n-1}(y) = 2\int_{\pi(B_\rho(x)\cap\partial E)} \left(\sqrt{1 + |\nabla u(z)|^2} - 1\right) dz$$
$$\sim \int_{\pi(B_\rho(x)\cap\partial E)} |\nabla u(z)|^2 \, dz,$$

where $\pi(B_\rho(x)\cap\partial E)$ denotes the projection of the $B_\rho(x)\cap\partial E$ on the tangent hyperplane to E at x, which is a horizontal plane in the coordinate system that we have chosen.

The main ingredient in the proof of Evans' theorem is the decay property of the excess: there exists a critical threshold such that, if the excess in a given ball is below the threshold, then decay occurs at all smaller scales.

Theorem 4.11 (Excess decay)**.** *Let L be as in Theorem 4.6. For every $M \geq 0$ and all $\alpha \in (0, 1/4)$ there exists $\varepsilon_0 = \varepsilon_0(n, m, \lambda, \Lambda, M, \alpha) > 0$ satisfying the following implication: if*

(a) $u \in H^1(B_r(x); \mathbb{R}^m)$ is a local minimizer in $B_r(x)$ of $v \mapsto \int L(\nabla v)\, dx$,

(b) $|(\nabla u)_{x,r}| \leq M$,

(c) $\mathrm{Exc}\,(u, B_r(x)) < \varepsilon_0$,

then

$$\mathrm{Exc}\,(u, B_{\alpha r}(x)) \leq c_{\mathrm{Exc}}\alpha\, \mathrm{Exc}\,(u, B_r(x))$$

with c_{Exc} depending only on n, m, λ and Λ. When $\nabla^2 L$ is uniformly continuous, condition (b) is not needed for the validity of the implication and ε_0 can be taken independent of M (even though it will depend on the modulus of continuity of $\nabla^2 L$).

Proof. We choose c_{Exc} in such a way that $16 c_E c_{P,II} c_{CL,N} < c_{\mathrm{Exc}}^2$, where c_E is the constant appearing in the statement of Lemma 4.7, $c_{P,II}$ is the constant in the Poincaré

inequality for functions having null mean value (Theorem A.16) and $c_{CL,N}$ is the constant of Proposition 4.12 below.

The proof we are about to present proceeds by contradiction, assuming that the statement fails for some α and M (for simplicity we keep L fixed in the contradiction argument, but a slightly more complex proof would give the stronger result): in step 2 we will normalize the excesses, obtaining functions w_k with $\mathrm{Exc}\,(w_k, B_\alpha(0)) \geq c_{\mathrm{Exc}}\alpha$ while $\mathrm{Exc}\,(w_k, B_1(0)) = 1$. Each w_k is a solution of

$$\partial_{x_\alpha}\left(\partial_{p_i^\alpha} L(\nabla w_k)\right) = 0.$$

We will then see in step 3 that, taking the limit as $k \to \infty$, any limit point w_∞ with respect to the weak H^1 topology solves a limit equation of the form

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha}\left(\partial_{p_i^\alpha p_j^\beta} L(p_\infty)\partial_{x_\beta} w_\infty^j\right) = 0 \qquad i = 1, \ldots, m.$$

Using Lemma 4.7 in combination with Proposition 4.12 we will then reach the contradiction.

Step 1. By contradiction, we have $M \geq 0$, $\alpha \in (0, 1/4)$ and local minimizers $u_k : \Omega \to \mathbb{R}^m$ in $B_{r_k}(x_k)$ with

$$\varepsilon_k := \mathrm{Exc}\,(u_k, B_{r_k}(x_k)) \longrightarrow 0$$

satisfying

$$|(\nabla u_k)_{x_k,r_k}| \leq M \tag{4.18}$$

but

$$\mathrm{Exc}\,(u_k, B_{\alpha r_k}(x_k)) > c_{\mathrm{Exc}}\alpha\,\mathrm{Exc}\,(u_k, B_{r_k}(x_k)) \qquad \forall\, k \in \mathbb{N}.$$

Step 2. Suitably rescaling and translating the functions u_k, we can assume that $x_k = 0$, $r_k = 1$ and $(u_k)_{0,1} = 0$ for all k. Setting $p_k := (\nabla u_k)_{0,1}$, the hypothesis (4.18) gives, up to subsequences,

$$p_k \longrightarrow p_\infty \in \mathbb{R}^{m\times n}. \tag{4.19}$$

In the case when $\nabla^2 L$ is uniformly continuous no uniform bound on p_k is assumed and we shall replace (4.19) with

$$\nabla^2 L(p_k) \to A_\infty \in \mathbb{R}^{m^2\times n^2}. \tag{4.20}$$

Notice that (4.20) holds under (4.19), simply with $A_\infty = \nabla^2 L(p_\infty)$. Furthermore, observe that in either case A_∞ satisfies a Legendre-Hadamard condition with constant λ and $|A_\infty| \leq \Lambda$.

We do a second translation in order to annihilate the mean of the gradients of minimizers: let us define

$$v_k(x) := u_k(x) - p_k x,$$

so that $(v_k)_{0,1} = 0$ and $(\nabla v_k)_{0,1} = 0$. According to property (i) of Remark 4.9 the excess does not change, so still

$$\mathrm{Exc}\,(v_k, B_1) = \varepsilon_k \longrightarrow 0$$

and

$$\mathrm{Exc}\,(v_k, B_\alpha) > c_{\mathrm{Exc}}\alpha\,\varepsilon_k.$$

Notice that the rescaled function v_k minimizes the integral functional associated to

$$p \mapsto L(p + p_k) - L(p_k) - \nabla L(p_k)p.$$

In order to get some contradiction, our aim is to find a "limit problem" whose solutions satisfy a suitable decay property. To that scope let us define

$$w_k := \frac{v_k}{\varepsilon_k} \qquad k \in \mathbb{N}.$$

It is trivial to check that $(w_k)_{B_1} = (\nabla w_k)_{B_1} = 0$, moreover

$$\mathrm{Exc}\,(w_k, B_1) = 1 \qquad \text{and} \qquad \mathrm{Exc}\,(w_k, B_\alpha) > c_{\mathrm{Exc}}\alpha. \tag{4.21}$$

A key point for the sequel of the proof is to notice that w_k is a local minimizer of the map $v \mapsto \int L_k(\nabla v)\,dx$, where

$$L_k(p) := \frac{1}{\varepsilon_k^2}\left[L(\varepsilon_k p + p_k) - L(p_k) - \nabla L(p_k)\varepsilon_k p\right].$$

Step 3. We now study both the limit of L_k and the limit of w_k, as $k \to \infty$. Since $L_k \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$, by Taylor expansion we are able to identify a limit Lagrangian, given by

$$L_\infty(p) = \frac{1}{2}\langle A_\infty p, p\rangle,$$

to which $L_k(p)$ converges uniformly on compact subsets of $\mathbb{R}^{m\times n}$. Indeed, this is clear with $A_\infty = \nabla^2 L(p_\infty)$ in the case when $p_k \to p_\infty$; however it is still true with A_∞ given by (4.20) when $\nabla^2 L$ is uniformly continuous, writing $L_k(p) = \frac{1}{2}\langle \nabla^2 L(p_k + \theta\varepsilon_k p)p, p\rangle$ with $\theta = \theta(k,p) \in (0,1)$.
Once we have the limit problem defined by L_∞, we drive our attention to w_k: it is a bounded sequence in $H^1(B_1;\mathbb{R}^m)$ because the excesses are constant, so by Rellich's compactness theorem (cf. Theorem A.13) we have that (possibly extracting one more subsequence)

$$w_k \longrightarrow w_\infty \qquad \text{in } L^2(B_1;\mathbb{R}^m), \quad \nabla w_k \rightharpoonup \nabla w_\infty \qquad \text{in } L^2(B_1;\mathbb{R}^{m\times n}).$$

The analysis of the limit problem now requires the verification that w_∞ solves the Euler-Lagrange equation associated to L_∞: to that scope we just need to pass to the limit in the Euler-Lagrange equation satisfied by w_k, namely

$$\sum_{\alpha,i} \int_{B_1} \frac{1}{\varepsilon_k} \left(\partial_{p_i^\alpha} L(p_k + \varepsilon_k \nabla w_k(x)) - \partial_{p_i^\alpha} L(p_k) \right) \partial_{x_\alpha} \varphi^i(x)\, dx = 0 \qquad \forall\, \varphi \in C_c^\infty(B_1; \mathbb{R}^m).$$

Writing the difference quotient of ∇L by means of the mean value theorem and exploiting the fact that $\nabla^2 L(p_k) \to A_\infty$ we obtain

$$\int_{B_1} \langle A_\infty \nabla w_\infty(x), \nabla \varphi(x) \rangle\, dx = 0 \qquad \forall\, \varphi \in C_c^\infty(B_1; \mathbb{R}^m), \tag{4.22}$$

provided we show that (here $\theta = \theta(x, \alpha, i, k) \in (0,1)$)

$$\lim_{k\to\infty} \sum_{\alpha,\beta,i,j} \int_{B_1} \left| \partial^2_{p_i^\alpha p_j^\beta} L(p_k + \theta \varepsilon_k \nabla w_k) - (A_\infty)_{ij}^{\alpha\beta} \right| dx = 0.$$

This can be obtained splitting the integral into regions given by $\{|\nabla w_k| \le c\}$ and $\{|\nabla w_k| > c\}$, with c fixed. The first contribution goes to zero, thanks to the convergence of p_k to p_∞ or, when p_k is possibly unbounded, thanks to the uniform continuity of $\nabla^2 L$. The second contribution tends to 0 as $L \uparrow \infty$ uniformly in k, since $|\nabla^2 L| \le \Lambda$ and the L^2-norm of ∇w_k is uniformly bounded by virtue of the fact that $\mathrm{Exc}\,(w_k, B_1) = 1$ for any k.

Step 4. Equality (4.22) means that

$$\sum_{\alpha,\beta,j} \partial_{x_\alpha} \left((A_\infty)_{ij}^{\alpha\beta} \partial_{x_\beta} w_\infty^j \right) = 0, \quad i = 1, \dots, m$$

in a weak sense: since the equation has constant coefficients we can then apply Lemma 4.7 to get

$$\fint_{B_{2\alpha}} |\nabla w_\infty(x) - (\nabla w_\infty)_{0,2\alpha}|^2\, dx \le 4 c_E \alpha^2 \fint_{B_1} |\nabla w_\infty(x)|^2\, dx \le 4 c_E \alpha^2. \tag{4.23}$$

On the other hand, using Proposition 4.12 below (with $L = L_k$, where we shall notice that still $|\nabla^2 L_k| \le \Lambda$ and that the functions L_k are uniformly λ-quasiconvex) we get

$$c_{\mathrm{Exc}}^2 \alpha^2 < \left(\mathrm{Exc}\,(w_k, B_\alpha) \right)^2 \le \frac{c_{CL,N}}{\alpha^2} \fint_{B_{2\alpha}} |w_k(x) - (w_k)_{0,2\alpha} - (\nabla w_k)_{0,2\alpha} x|^2\, dx,$$

hence passing to the limit as $k \to \infty$ gives

$$\frac{c_{\mathrm{Exc}}^2}{c_{CL,N}} \alpha^4 \le \fint_{B_{2\alpha}} |w_\infty(x) - (w_\infty)_{0,2\alpha} - (\nabla w_\infty)_{0,2\alpha} x|^2\, dx.$$

However, the Poincaré inequality and (4.23) give

$$\begin{aligned}\fint_{B_{2\alpha}} |w_\infty(x) - (w_\infty)_{0,2\alpha} - (\nabla w_\infty)_{0,2\alpha} x|^2 \, dx &\le 4c_{P,II}\alpha^2 \fint_{B_{2\alpha}} |\nabla w_\infty(x) - (\nabla w_\infty)_{0,2\alpha}|^2 \, dx \\ &\le 16 c_{P,II} c_E \alpha^4.\end{aligned}$$

Taking into account our definition of c_{Exc} we have reached a contradiction. □

The following proposition can be considered as a nonlinear Caccioppoli inequality. It can be derived without using the Euler-Lagrange equation (which would not help) and exploiting the minimality instead.

Proposition 4.12 (Caccioppoli inequality for minimizers)**.** *There exists $c_{CL,N} = c_{CL,N}(\lambda, \Lambda)$ such that, if L is λ-uniformly quasiconvex with $|\nabla^2 L| \le \Lambda$ and if u is a local minimizer in Ω, then*

$$\fint_{B_r(x_0)} |\nabla u - A|^2 \, dx \le \frac{c_{CL,N}}{r^2} \fint_{B_{2r}(x_0)} |u - a - A(x - x_0)|^2 \, dx$$

for all balls $B_{2r}(x_0) \Subset \Omega$, all $A \in \mathbb{R}^{m\times n}$ and $a \in \mathbb{R}^m$.

Proof. By translation invariance, both with respect to the domain and the target, we can assume $a = 0$, $x_0 = 0$. Let $r \le t < s \le 2r$ and let $\zeta \in C_c^\infty(B_s; \mathbb{R})$ with $\zeta \equiv 1$ on B_t, $0 \le \zeta \le 1$ and $|\nabla \zeta| \le 2(s - t)$. We exploit the functions $\zeta, 1 - \zeta$ to define suitable cutoff versions of $u - Ax$, specifically we set $\phi = \zeta(u - Ax)$, $\psi = (1 - \zeta)(u - Ax)$. Hence, the equation $\phi + \psi = u - Ax$ gives

$$\nabla\phi + \nabla\psi = \nabla u - A.$$

From the λ-uniform quasiconvexity we get

$$\begin{aligned}\int_{B_s} L(A) + \lambda|\nabla\phi|^2 \, dx &\le \int_{B_s} L(A + \nabla\phi) \, dx \\ &= \int_{B_s} L(\nabla u - \nabla\psi) \, dx \\ &\le \int_{B_s} L(\nabla u) - \nabla L(\nabla u)\nabla\psi + c|\nabla\psi|^2 \, dx,\end{aligned} \tag{4.24}$$

with $c = c(\Lambda)$. On the other hand, since u is a local minimum, we have

$$\begin{aligned}\int_{B_s} L(\nabla u) \, dx &\le \int_{B_s} L(\nabla u - \nabla\phi) \, dx \\ &= \int_{B_s} L(A + \nabla\psi) \, dx \\ &\le \int_{B_s} L(A) + \nabla L(A)\nabla\psi + c|\nabla\psi|^2 \, dx.\end{aligned} \tag{4.25}$$

Combining (4.24) with (4.25) we get

$$\lambda \int_{B_s} |\nabla \phi|^2 \, dx \le \int_{B_s} |\nabla L(A) - \nabla L(\nabla u)||\nabla \psi| + c|\nabla \psi|^2 \, dx,$$

so that (recalling that $\zeta \equiv 1$ and $\psi \equiv 0$ on B_t)

$$\int_{B_t} |\nabla u - A|^2 \, dx \le c \int_{B_s \setminus B_t} |\nabla u - A||\nabla \psi| + |\nabla \psi|^2 \, dx,$$

with $c = c(\lambda, \Lambda)$.

Now, since $|\nabla \psi| \le |\nabla u - A| + 2|u - Ax|/(s-t)$, we get

$$\int_{B_t} |\nabla u - A|^2 \, dx \le c \int_{B_s \setminus B_t} |\nabla u - A|^2 \, dx + \frac{c}{(s-t)^2} \int_{B_s \setminus B_t} |u - Ax|^2 \, dx$$

for some new constant $c = c(\lambda, \Lambda)$ and we apply the hole-filling technique to derive

$$\int_{B_t} |\nabla u - A|^2 \, dx \le \theta \int_{B_s} |\nabla u - A|^2 \, dx + \frac{c}{(s-t)^2} \int_{B_{2r}} |u - Ax|^2 \, dx$$

with $\theta = c/(c+1) < 1$. At this point, since the inequality is true for all $r \le t \le s \le 2r$, a standard iteration scheme gives the result. Indeed, let $\tau \in (0,1)$ with $\theta < \tau^2$ and define $t_i = (1 - \tau^i/2)2r$, so that $t_0 = r$, $t_i \uparrow 2r$ and $t_{i+1} - t_i = r(1-\tau)\tau^i$. It is clear that one can choose the constant τ depending on λ, Λ only. By iteration of the inequality

$$\int_{B_{t_i}} |\nabla u - A|^2 \, dx \le \theta \int_{B_{t_{i+1}}} |\nabla u - A|^2 \, dx + \frac{c}{r^2(1-\tau)^2} \tau^{-2i} \int_{B_{2r}} |u - Ax|^2 \, dx$$

we get

$$\begin{aligned} \int_{B_{t_0}} |\nabla u - A|^2 \, dx &\le \theta^N \int_{B_{t_N}} |\nabla u - A|^2 \, dx + \frac{c}{r^2(1-\tau)^2} \sum_{i=0}^{N-1} (\theta/\tau^2)^i \int_{B_{2r}} |u - Ax|^2 \, dx \\ &\le \theta^N \int_{B_{2r}} |\nabla u - A|^2 \, dx + \frac{c\tau^2}{r^2(1-\tau)^2(\tau^2 - \theta)} \int_{B_{2r}} |u - Ax|^2 \, dx \end{aligned}$$

for any integer $N \ge 1$. Letting $N \to \infty$ we obtain the result. □

4.3 Partial regularity for systems: $\mathscr{L}^n(\Sigma(u)) = 0$

After having proven Theorem 4.11 about the decay of the excess, we will see how it can be used to prove partial regularity results for systems, starting with Theorem 4.6.

We briefly recall that $\Omega_{\text{reg}}(u)$ denotes the largest open set contained in Ω where $u : \Omega \to \mathbb{R}^m$ admits a C^1 representative, while $\Sigma(u) := \Omega \setminus \Omega_{\text{reg}}(u)$. Our aim, in the next two sections, is to show that for a locally minimizing solution of an elliptic system the following facts hold:

- $\mathscr{L}^n\big(\Sigma(u)\big) = 0$;
- $\mathscr{H}^{n-2+\varepsilon}\big(\Sigma(u)\big) = 0$ for all $\varepsilon > 0$ in the uniformly convex case and $\mathscr{H}^{n-2}\left(\Sigma(u)\right) = 0$ if $\nabla^2 L$ is also uniformly continuous.

In order to exploit Theorem 4.11 and prove that $\mathscr{L}^n\big(\Sigma(u)\big) = 0$, we fix once for all the constant $\alpha \in (0, 1/4)$ in such a way that $c_{\mathrm{Exc}}\alpha < 1/2$ (recall that c_{Exc} depends only on the dimensions and on the ellipticity constants). Then, we fix $M \geq 0$, so that there is an associated $\varepsilon_0 = \varepsilon_0(n, m, \lambda, \Lambda, M)$ for which the decay property of the excess applies and the excess itself is halved when passing from the scale r to the scale αr.

Definition 4.13. *We shall set*

$$\Omega_M(u) := \big\{x \in \Omega : \exists\, B_r(x) \Subset \Omega \text{ with } \big|(\nabla u)_{x,r}\big| < M/2 \text{ and } \mathrm{Exc}\,(u, B_r(x)) < \varepsilon_1\big\}$$

where ε_1 *satisfies*

$$2^{n/2}\varepsilon_1 \leq \varepsilon_0 \tag{4.26}$$

and

$$(2^{n/2} + 2(2/\alpha)^{n/2})\varepsilon_1 \leq M/2. \tag{4.27}$$

Remark 4.14. Let us remark that the set $\Omega_M(u) \subset \Omega$ is open by definition. Moreover, by the Lebesgue approximate continuity theorem (see, for instance, Section 1.3) it is easy to see that

$$\mathscr{L}^n\left(\{|\nabla u| < M/2\} \setminus \Omega_M(u)\right) = 0. \tag{4.28}$$

Moreover, using (4.28) together with the fact that (by definition of local minimizer) $u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$, we obtain that

$$\mathscr{L}^n\left(\Omega \setminus \bigcup_{M\in\mathbb{N}} \Omega_M(u)\right) = \mathscr{L}^n\left(\Omega \setminus \bigcup_{M\in\mathbb{N}} \{|\nabla u| < M/2\}\right) = 0. \tag{4.29}$$

By the previous remark, if we are able to prove that

$$\Omega_M(u) \subset \Omega_{\mathrm{reg}}(u) \qquad \forall M > 0, \tag{4.30}$$

we obtain $\mathscr{L}^n\left(\Sigma(u)\right) = 0$. So, the rest of this section is devoted to the proof of the inclusion above, with M fixed.

Proof. Fix $x \in \Omega_M(u)$: according to Definition 4.13 there exists $r > 0$ such that $B_r(x) \Subset \Omega$, $|(\nabla u)_{B_r(x)}| < M/2$ and $\mathrm{Exc}\,(u, B_r(x)) < \varepsilon_1$. We prove that

$$B_{r/2}(x) \subset \Omega_{\mathrm{reg}}(u).$$

To this scope, let us fix $y \in B_{r/2}(x)$.

Step 1. Thanks to our choice of ε_1 (see equation (4.26) in Definition 4.13) we have

$$\begin{aligned}
\mathrm{Exc}\left(u, B_{r/2}(y)\right) &= \left(\fint_{B_{r/2}(y)} |\nabla u(z) - (\nabla u)_{y,r/2}|^2 \, dz\right)^{1/2} \\
&\leq \left(\fint_{B_{r/2}(y)} |\nabla u(z) - (\nabla u)_{x,r}|^2 \, dz\right)^{1/2} \\
&\leq 2^{n/2} \left(\fint_{B_r(x)} |\nabla u(z) - (\nabla u)_{x,r}|^2 \, dz\right)^{1/2} = 2^{n/2} \mathrm{Exc}\left(u, B_r(x)\right) < \varepsilon_0
\end{aligned}$$

so, momentarily ignoring the hypothesis that $|(\nabla u)_{B_{r/2}(y)}|$ should be bounded by M (we are postponing this to Step 2 of this proof), Theorem 4.11 gives *tout court*

$$\mathrm{Exc}\left(u, B_{\alpha r/2}(y)\right) \leq \frac{1}{2} \mathrm{Exc}\left(u, B_{r/2}(y)\right) < \frac{1}{2} \varepsilon_0,$$

thus, just iterating Theorem 4.11, we get

$$\mathrm{Exc}\left(u, B_{\alpha^k r/2}(y)\right) \leq 2^{-k} \mathrm{Exc}\left(u, B_{r/2}(y)\right). \tag{4.31}$$

As we have often seen through these notes, we can then apply an interpolation argument to a sequence of radii with ratio α to obtain, for every $\rho \in (0, r/2]$ and every $y \in B_{r/2}(x)$, that

$$\mathrm{Exc}\left(u, B_\rho(y)\right) \leq \alpha^{-\mu - n/2} \left(\frac{\rho}{r/2}\right)^\mu \mathrm{Exc}\left(u, B_{r/2}(y)\right) \leq \alpha^{-\mu - n/2} \left(\frac{\rho}{r/2}\right)^\mu \varepsilon_0$$

where we have set $\mu = \left(\log_2(1/\alpha)\right)^{-1}$. We conclude that the components of ∇u belong to the Campanato space $\mathcal{L}^{2,n+2\mu}(B_{r/2}(x); \mathbb{R}^{m \times n})$ and hence, invoking Theorem 3.9, u belongs to $C^{1,\mu}(B_{r/2}(x); \mathbb{R}^m)$ which proves the inclusion $B_{r/2}(x) \subset \Omega_{\mathrm{reg}}(u)$.

Step 2. Now that we have explained how the proof runs through the iterative application of Theorem 4.11, we deal with the initially neglected hypothesis, that is $|(\nabla u)_{y,r/2}| < M$ and, at each subsequent step, $|(\nabla u)_{y,\alpha^k r/2}| < M$. Observe that in Step 1 of this proof we never used (4.27).
Since $x \in \Omega_M(u)$ and r fulfills Definition 4.13, for the first step it is sufficient to use the

triangle inequality in (4.32) and Hölder's inequality in (4.33): in fact we can estimate

$$\begin{aligned}
\left|(\nabla u)_{y,r/2}\right| &= \left|\fint_{B_{r/2}(y)} (\nabla u(z) - (\nabla u)_{x,r})\, dz + (\nabla u)_{x,r}\right| \\
&\leq \fint_{B_{r/2}(y)} \left|\nabla u(z) - (\nabla u)_{x,r}\right| dz + \left|(\nabla u)_{x,r}\right| \qquad (4.32)\\
&= \frac{2^n}{\omega_n r^n} \int_{B_{r/2}(y)} \left|\nabla u(z) - (\nabla u)_{x,r}\right| dz + \left|(\nabla u)_{x,r}\right| \\
&\leq 2^{n/2} \left(\fint_{B_r(x)} \left|\nabla u(z) - (\nabla u)_{x,r}\right|^2 dz\right)^{1/2} + \left|(\nabla u)_{x,r}\right| \qquad (4.33)
\end{aligned}$$

so that we can conclude

$$\left|(\nabla u)_{y,r/2}\right| \leq 2^{n/2}\mathrm{Exc}\,(u, B_r(x)) + \left|(\nabla u)_{x,r}\right| < 2^{n/2}\varepsilon_1 + M/2 < M. \qquad (4.34)$$

We now show inductively that, for every integer $k \geq 1$,

$$\left|(\nabla u)_{y,\alpha^k r/2}\right| \leq M/2 + \varepsilon_1 \left(2^{n/2} + (2/\alpha)^{n/2} \sum_{j=0}^{k-1} 2^{-j}\right). \qquad (4.35)$$

If we recall (4.27), it is clear that (4.35) implies

$$\left|(\nabla u)_{y,\alpha^k r/2}\right| \leq M \qquad \forall\, k \geq 1.$$

The case $k = 1$ follows directly from (4.34), because, estimating as in (4.32) and (4.33), we immediately get

$$\begin{aligned}
\left|(\nabla u)_{y,\alpha r/2}\right| &\leq \fint_{B_{\alpha r/2}(y)} \left|\nabla u(z) - (\nabla u)_{y,r/2}\right| dz + \left|(\nabla u)_{y,r/2}\right| \\
&\leq \alpha^{-n/2}\mathrm{Exc}\left(u, B_{r/2}(y)\right) + \left|(\nabla u)_{y,r/2}\right| \\
&\leq \alpha^{-n/2}2^{n/2}\varepsilon_1 + 2^{n/2}\varepsilon_1 + M/2.
\end{aligned}$$

We shall now prove equation (4.35) for $(k+1)$, given the earlier inductive steps. With the same procedure, we estimate again

$$\begin{aligned}
\left|(\nabla u)_{y,\alpha^{k+1}r/2}\right| &\leq \fint_{B_{\alpha^{k+1}r/2}(y)} \left|\nabla u(z) - (\nabla u)_{y,\alpha^k r/2}\right| dz + \left|(\nabla u)_{y,\alpha^k r/2}\right| \\
&\leq \alpha^{-n/2}\mathrm{Exc}\left(u, B_{\alpha^k r/2}(y)\right) + \left|(\nabla u)_{y,\alpha^k r/2}\right| \\
&\leq \alpha^{-n/2}2^{n/2-k}\varepsilon_1 + M/2 + \varepsilon_1 \left(2^{n/2} + (2/\alpha)^{n/2} \sum_{j=0}^{k-1} 2^{-j}\right) \qquad (4.36)
\end{aligned}$$

where (4.36) is obtained combining the estimate on the excess (4.31) with the inductive hypothesis (4.35). A straightforward rearrangement of the terms leads to the conclusion. □

4.4 Partial regularity for systems: $\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0$

We are now ready to show that, if $L \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$ satisfies the Legendre condition for some $\lambda > 0$ as well as

$$|\nabla^2 L(p)| \leq \Lambda < +\infty \qquad \forall\, p \in \mathbb{R}^{m\times n},$$

then we have a stronger upper bound on the size of the singular set, namely

$$\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0 \qquad \forall\, \varepsilon > 0, \tag{4.37}$$

where we have set, as usual, $\Sigma(u) := \Omega \setminus \Omega_{\mathrm{reg}}(u)$. To this scope, we will make use of Proposition C.7.

Let us remark that, in comparison with Theorem 4.6, we have slightly but significantly changed the properties of the system, replacing the weaker hypothesis of uniform quasiconvexity with the Legendre condition for some positive λ (i.e. uniform convexity). In fact, thanks to the Legendre condition the sequence of difference quotients $\Delta_{h,\gamma}(\nabla u)$ satisfies an equielliptic family of systems. Then, via Caccioppoli-Leray inequality 2.1 the sequence $\Delta_{h,\gamma}(\nabla u)$ is uniformly bounded in L^2_{loc}. The existence of second derivatives in L^2_{loc} is useful to estimate the size of the singular set. We will also obtain a stronger version of (4.37) for systems in which $\nabla^2 L$ is uniformly continuous, see Corollary 4.17.

As for the strategy: in Proposition 4.15 we are going to split the singular set $\Sigma(u)$ in two other sets, $\Sigma_1(u)$ and $\Sigma_2(u)$, and then we are going to estimate separately the Hausdorff measure of each of them with the aid of Proposition 4.16 and Theorem 4.19, respectively.

Proposition 4.15. *Consider, as above, a variational problem defined by $L \in C^2(\mathbb{R}^{m\times n};\mathbb{R})$ with $|\nabla^2 L| \leq \Lambda$, satisfying the Legendre condition for some $\lambda > 0$. If u is a local minimizer of such a problem, define the sets*

$$\Sigma_1(u) := \left\{ x \in \Omega : \limsup_{r\to 0} r^{2-n} \int_{B_r(x)} |\nabla^2 u(y)|^2\, dy > 0 \right\}$$

and

$$\Sigma_2(u) := \left\{ x \in \Omega : \limsup_{r\to 0} \left|(\nabla u)_{x,r}\right| = +\infty \right\}.$$

Then $\Sigma(u) \subset \Sigma_1(u) \cup \Sigma_2(u)$. If, in addition, $\nabla^2 L$ is uniformly continuous, we have $\Sigma(u) \subset \Sigma_1(u)$.

Proof. Fix $x \in \Omega$ such that $x \notin \Sigma_1(u) \cup \Sigma_2(u)$, then

- there exists $M_1 < +\infty$ such that $\left|(\nabla u)_{x,r}\right| < M_1$ for arbitrarily small radii $r > 0$;
- thanks to the Poincaré inequality, Theorem A.16,

$$\mathrm{Exc}\,(u, B_r(x))^2 \leq c_{P,II} r^{2-n} \int_{B_r(x)} |\nabla^2 u(y)|^2 \, dy \longrightarrow 0.$$

Thus, for some $M = M(M_1, n, m, \lambda, \Lambda) > 0$, we have that $x \in \Omega_M(u)$, where $\Omega_M(u)$ has been specified in Definition 4.13, and hence $\Omega_M(u) \subset \Omega_{\mathrm{reg}}$ due to the argument we presented in the proof of Theorem 4.6.

The second part of the statement can be proven noticing that, in the case when $\nabla^2 L$ is uniformly continuous, no bound on $|(\nabla u)_{x,r}|$ is needed in the theorem concerning the decay of the excess, and hence also in the characterization of the regular set. □

Proposition 4.16. *If $u \in H^2_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$, we have that*

$$\mathscr{H}^{n-2}\left(\Sigma_1(u)\right) = 0.$$

Proof. We shall employ Proposition C.7 for the absolutely continuous measure defined by $\mu := |\nabla^2 u|^2 \mathscr{L}^n$. Obviously, we choose $k = (n-2)$ and we have that μ vanishes on sets with finite $\mathscr{H}^{n-2}$-measure. The conclusion follows once we observe that

$$\Sigma_1(u) = \bigcup_{\nu=1}^{\infty} \left\{ x \in \Omega : \limsup_{r \to 0} \frac{\mu(\overline{B}_r(x))}{\omega_{n-2} r^{n-2}} > \frac{1}{\nu} \right\}.$$

□

Thanks to the second part of the statement of Proposition 4.15 we get:

Corollary 4.17. *If we add the uniform continuity of $\nabla^2 L$ to the hypotheses of Proposition 4.16, we can conclude that*

$$\mathscr{H}^{n-2}\left(\Sigma(u)\right) = 0. \tag{4.38}$$

The estimate on the Hausdorff measure of $\Sigma_2(u)$ is a bit more delicate and goes through the estimate of the Hausdorff measure of the so-called *approximate discontinuity set* S_v of a function v.

Definition 4.18. *Given a function $v \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R}^d)$, we set*

$$\Omega \setminus S_v := \left\{ x \in \Omega : \exists\, z \in \mathbb{R}^d \text{ s.t. } \lim_{r \downarrow 0} \fint_{B_r(x)} |v(y) - z| \, dy = 0 \right\}.$$

When such a z exists, it is unique and we will call it approximate limit *of v at the point x.*

Here is the general result in question:

Theorem 4.19. *If $v \in W^{1,p}(\Omega;\mathbb{R}^d)$, $1 \le p \le n$, then*

$$\mathscr{H}^{n-p+\varepsilon}(S_v) = 0 \qquad \forall\, \varepsilon > 0.$$

Notice that the statement is trivial (i.e. $S_v = \emptyset$) in the case $p > n$, by the Sobolev embedding Theorem 2.6. On the other hand, as p increases from 1 to n, the Hausdorff dimension of the approximate discontinuity set moves from $n-1$ to 0.

Applying this theorem to $v = \nabla u \in H^1(\Omega;\mathbb{R}^{m\times n})$, for $p = 2$, we come to the conclusion that $\mathscr{H}^{n-2+\varepsilon}\left(\Sigma_2(u)\right) = 0$ for all $\varepsilon > 0$.

Let us now proceed with the proof of Theorem 4.19.

Proof. For the sake of simplicity, we present the proof of this result in the scalar case, namely when $d = 1$.

Step 1. Fix $0 < \eta < \rho$, we observe that

$$\left|v_{x,\rho} - v_{x,\eta}\right| \le \int_\eta^\rho t^{-n} \int_{B_t(x)} |\nabla v(y)|\, dy\, dt. \tag{4.39}$$

The proof of this inequality is straightforward when v is C^1:

$$\begin{aligned}\left|v_{x,\rho} - v_{x,\eta}\right| &= \Big|\fint_{B_1} \left(v(x+\rho z) - v(x+\eta z)\right) dz\Big| = \Big|\fint_{B_1}\int_\eta^\rho \langle \nabla v(x+tz), z\rangle\, dt\, dz\Big| \\ &\le \fint_{B_1}\int_\eta^\rho |\nabla v(x+tz)|\, dt\, dz = \int_\eta^\rho t^{-n}\int_{B_t(x)} |\nabla v(y)|\, dy\, dt.\end{aligned}$$

and follows, for $v \in W^{1,p}(\Omega;\mathbb{R}^d)$, by a direct approximation argument in (4.39).

Now, suppose that x is a point for which $\int_{B_t(x)} |\nabla v(y)|\, dy = o(t^{n-1+\varepsilon})$ for some $\varepsilon > 0$, then it is also true that $\rho^{-(n-1)}\int_{B_\rho(x)} |\nabla v(y)|\, dy \to 0$ and the map $r \mapsto v_{x,r}$ admits a limit z as $r \to 0$ because it is a Cauchy sequence. This key claim relies on (4.39). At that stage, thanks to the Poincaré inequality (Theorem A.16)

$$\fint_{B_r(x)} |v(y) - v_{x,r}|\, dy \le c_{P,II}\; r^{-(n-1)} \int_{B_r(x)} |\nabla v(y)|\, dy \xrightarrow{r\to 0} 0,$$

therefore

$$\fint_{B_r(x)} |v(y) - z|\, dy \xrightarrow{r\to 0} 0,$$

that is to say, $x \notin S_v$. Hence, this discussion implies that, for all $\varepsilon > 0$,

$$\Omega \setminus S_v \supset \left\{ x \in \Omega : \int_{B_t(x)} |\nabla v(y)|\, dy = o(t^{n-1+\varepsilon}) \right\}. \tag{4.40}$$

Step 2. In order to refine (4.40) suppose, that

$$\int_{B_t(x)} |\nabla v(y)|^p \, dy = o(t^{n-p+\varepsilon})$$

for some $\varepsilon > 0$. Then, by Hölder's inequality,

$$\int_{B_t(x)} |\nabla v(y)| \, dy \leq o(t^{n/p-1+\varepsilon/p}) t^{n/p'} = o(t^{n-1+\varepsilon/p}).$$

For this reason we can deduce from (4.40) the inclusion

$$\Omega \setminus S_v \supset \left\{ x \in \Omega : \int_{B_t(x)} |\nabla v(y)|^p \, dy = o(t^{n-p+\varepsilon}) \right\} \qquad \forall \varepsilon > 0. \tag{4.41}$$

In view of Proposition C.7, arguing as we have done above for the set Σ_1, the complement of the set $\{x \in \Omega : \int_{B_t(x)} |\nabla v(y)|^p \, dy = o(t^{n-p+\varepsilon})\}$ is $\mathscr{H}^{n-p+\varepsilon}$-negligible, hence the approximate discontinuity set S_v is $\mathscr{H}^{n-p+\varepsilon}$-negligible, too.

□

Remark 4.20. In the case $p = 1$ it is even possible to prove that S_v is σ-finite with respect to $\mathscr{H}^{n-1}$, so the quantitative description of the approximate discontinuity set with the scale of Hausdorff measures is sharp. On the contrary, in the case $p > 1$ the right scale for the quantitative description of the approximate discontinuity set are provided by the so-called capacities (see [98]).

5 Viscosity solutions

In this section we want to present the notion of *viscosity solution* for equations having the general form

$$E(x, u(x), \nabla u(x), \nabla^2 u(x)) = 0. \tag{5.1}$$

The idea behind this approach is a second-order comparison principle, which makes it suitable for dealing with both elliptic and parabolic problems. Consistently with this goal, we shall assume u to be defined on some locally compact domain $A \subset \mathbb{R}^n$, so that we require every point in the domain A to have a compact neighborhood. This topological assumption is actually very useful, as it allows to deal at the same time with open and closed domains, as well as with domains of the form $\mathbb{R}^{n-1} \times [0, \infty)$, which typically occur in the study of parabolic problems.

For a survey on viscosity solutions see [22], while we refer to reader to [16] for a thorough treatment of fully nonlinear elliptic equations.

5.1 Basic definitions

We first need to recall two classical ways to regularize a function.

Definition 5.1 (u.s.c. and l.s.c. regularizations)**.** *Let $A' \subset A$ be a dense subset and $u : A' \to \overline{\mathbb{R}}$. We define its* upper regularization u^* *on A by means of one of the following three equivalent formulae:*

(i) $u^*(x) := \sup \{\limsup_h u(x_h) : (x_h) \subset A',\ x_h \to x\}$;

(ii) $u^*(x) := \inf_{r>0} \left\{\sup_{B_r(x)\cap A'} u\right\}$;

(iii) $u^*(x) := \inf \{v(x) : v$ *is upper semi-continuous and* $v \geq u\}$.

Similarly we can define the lower regularization u_* *by:*

(i') $u_*(x) := \inf \{\liminf_h u(x_h) : (x_h) \subset A',\ x_h \to x\}$;

(ii') $u_*(x) := \sup_{r>0} \left\{\inf_{B_r(x)\cap A'} u\right\}$;

(iii') $u_*(x) := \sup \{v(x) : v$ *is lower semi-continuous and* $v \leq u\}$.

The lower regularization is also characterized by the identity $u_* = -(-u)^*$.

Remark 5.2. One clearly has the pointwise inequalities $u_* \leq u \leq u^*$ on the subset where u is defined. Moreover, u is continuous at a point $x \in A$ (or, more precisely, it has a continuous extension in case $x \in A \setminus A'$) if and only if $u_*(x) = u(x) = u^*(x)$.

We now assume that $E : L \subset A \times \mathbb{R} \times \mathbb{R}^n \times \mathrm{Sym}^{n\times n} \to \mathbb{R}$, with L dense. Here and in the sequel we denote by $\mathrm{Sym}^{n\times n}$ the space of symmetric $n \times n$ matrices over $\mathbb{R}$.

Definition 5.3 (Subsolution)**.** *A function $u : A \to \mathbb{R}$ is a* subsolution *for the equation* (5.1) *(and we shall write $E \leq 0$) if the two following conditions hold:*

(a) the upper semi-continuous regularization u^* is a real-valued function;

(b) for any $x \in A$, if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is C^∞ and $u^* - \varphi|_A$ has a local maximum at x, then

$$E_*(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \leq 0 \ . \tag{5.2}$$

It is obvious from the definition that the property of being a subsolution is invariant under u.s.c. regularization, i.e. u is a subsolution if and only if u^* is a subsolution, too.

The geometric idea behind this definition is to use a local *comparison principle*, since assuming that $u^* - \varphi$ has a maximum at x implies, if u is smooth, that $\nabla u^*(x) = \nabla\varphi(x)$ and $\nabla^2 u^*(x) \leq \nabla^2\varphi(x)$ (as bilinear forms). This should be compared to the theory of distributions (hence to the classical study of weak solutions to partial differential equations), where integration by parts is exploited to the scope of transferring derivatives from u to the test function φ.

Similarly, we give the following:

Definition 5.4 (Supersolution)**.** *A function $u : A \to \mathbb{R}$ is a supersolution for the equation* (5.1) *(and we shall write $E \geq 0$) if the two following conditions hold:*

(a) the lower semi-continuous u_* is a real-valued function;

(b) for any $x \in A$, if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is C^∞ and $u_* - \varphi|_A$ has a local minimum at x, then

$$E^*(x, u_*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \geq 0. \tag{5.3}$$

Combining the two definitions above, we introduce the following one:

Definition 5.5. *A function $u : A \to \mathbb{R}$ is a* viscosity solution *of the equation* (5.1) *(and we shall write $E = 0$) if it is both a subsolution and a supersolution.*

The notion of viscosity solutions was introduced by Evans in [35], but the phrase "viscosity solution" is due to Crandall and Lions [23] (see also [21] for first-order problems and the following [22] and [57] for second-order problems, as in our case).

Remark 5.6. Without loss of generality, we can always assume in the definition of subsolution that the value of the local maximum is zero, that is $u^*(x) - \varphi(x) = 0$. This is true because the test function φ is arbitrary and the value of φ at x does not appear in (5.2). Also, possibly subtracting $|y - x|^4$ to φ (so that first and second derivatives of φ at x remain unchanged), we can assume with no loss of generality that the local maximum is strict. Analogous remarks hold for supersolutions.

Remark 5.7. A trivial example of viscosity solution is given by the Dirichlet function $\chi_{\mathbb{Q}}$ on $\mathbb{R}$, which is easily seen to be a solution to the equation $u' = 0$ in the sense above. This example shows that some continuity assumption is needed, in order to hope for reasonable existence and uniqueness results.

Remark 5.8. Rather surprisingly, a solution of $E = 0$ in the viscosity sense does not necessarily solve $-E = 0$ in the viscosity sense. To display this phenomenon, consider the equation defined by

$$E(x, u(x), u'(x), u''(x)) = |u'(x)| - 1,$$

as well as

$$-E(x, u(x), u'(x), u''(x)) = 1 - |u'(x)|.$$

We will prove later (as an application of Theorem 5.12) that the Lipschitz function $f(t) = \min\{1 - t, 1 + t\}$, that is easily seen to be a supersolution of the first problem, is actually a solution thereof, while it is not even a subsolution for the second one, since choosing $\varphi = 1$ leads to violate the condition $1 - |\varphi'(0)| \leq 0$, corresponding to (5.2). We have instead the following parity properties:

(a) Let E be *odd* in the triple (u, p, S). If u satisfies $E \leq 0$, then $-u$ satisfies $E \geq 0$.

(b) Let E be *even* in the triple (u, p, S). If u satisfies $E \leq 0$, then $-u$ satisfies $-E \geq 0$.

We now spend some words on two ways of simplifying the conditions that need to be checked in order prove that a certain function is indeed a subsolution or a supersolution of a given equation of the type considered above. For the sake of definiteness, we will explicitly refer to the case of subsolutions, leaving the obvious variations needed when dealing with supersolutions to the reader.

We have already seen in Remark 5.6 that one can assume without loss of generality that $u^* - \varphi$ has a strict local maximum, equal to 0, at x. We further claim that one can also work, equivalently, with the larger class of C^2 functions φ, in a neighborhood of x. One implication is trivial, so let us discuss the converse one. Let then $\varphi \in C^2$ and assume $u^*(y) - \varphi(y) \leq 0$ for $y \in \overline{B}_r(x)$, with equality only when $y = x$. By employing suitable mollifiers, we can build a sequence (φ_k) in $C^\infty(\overline{B}_r(x); \mathbb{R})$ with $\varphi_k \to \varphi$ in $C^2(\overline{B}_r(x); \mathbb{R})$. Let then x_k be a maximum in $\overline{B}_r(x)$ of the function $u^* - \varphi_k$. Since $\varphi_k \to \varphi$ uniformly, it is easy to see that any limit point of (x_k) has to be a maximum for $u^* - \varphi$, hence it must be x; in addition the convergence of the maximal values yields $u^*(x_k) \to u^*(x)$. The subsolution property, applied with φ_k at x_k, gives

$$E_*(x_k, u^*(x_k), \nabla\varphi_k(x_k), \nabla^2\varphi_k(x_k)) \leq 0$$

and we can now let $k \to \infty$ and use the lower semicontinuity of E_* to prove our claim.

Actually, it is rather easy now to see that the subsolution property at $x \in A$ is also equivalent to the conditions

$$E_*(x, u^*(x), p, S) \leq 0 \quad \forall\ (p, S) \in J_2^+ u^*(x)$$

where $J_2^+ u^*$ is the second-order superjet of u, namely the set

$$J_2^+ u^*(x) := \left\{(p, S) :\ u^*(y) \leq u^*(x) + \langle p, y - x\rangle + \tfrac{1}{2}\langle S(y - x), y - x\rangle + o(|y - x|^2)\right\}.$$

Indeed, let $P(y) := u^*(x) + \langle p, y - x\rangle + \frac{1}{2}\langle S(y - x), y - x\rangle$, so that $u^*(y) \leq P(y) + o(|y - x|^2)$, with equality when $y = x$. For any $\varepsilon > 0$ we have $u^*(y) \leq P(y) + \varepsilon|y - x|^2$ on a sufficiently small neighborhood of x with equality at $y = x$, thus we can apply (5.2) to this smooth function to get

$$E_*(x, u^*(x), p, S + 2\varepsilon I) = E_*(x, u^*(x), \nabla P(x), \nabla^2 P(x) + 2\varepsilon I) \leq 0$$

and by lower semicontinuity of E_* we can let $\varepsilon \to 0$ and prove the claim. Of course, if we are dealing with first-order equations, only the first-order superjet is needed.

Remark 5.9. After these preliminary remarks, it should be clear that this theory, despite its elegance, has two main restrictions: on the one hand it is only suited to first or second-order equations (since no information on third derivatives can be obtained via local comparison), on the other hand this approach cannot be generalized to the case of systems.

5.2 Viscosity versus classical solutions

We first observe that a classical solution is *not* always a viscosity solution. To see this, consider on $\mathbb{R}$ the problem $u'' - 2 = 0$. The function $f(t) = t^2$ is clearly a classical solution, but it is not a viscosity solution, because it is not even a viscosity supersolution (choose $\varphi \equiv 0$ and observe that equation (5.2) is patently violated at the origin).

Since we can always take $u = \varphi$ if u is at least C^2, the following theorem is trivial:

Theorem 5.10 (C^2 viscosity solutions are classical solutions)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, $u \in C^2(\Omega; \mathbb{R})$ and E be continuous. If u is a viscosity solution of* (5.1) *on Ω, then it is also a classical solution of the same problem.*

The converse holds if $S \mapsto E_*(x, u, p, S)$ and $S \mapsto E^*(x, u, p, S)$ are non-increasing in $\mathrm{Sym}^{n\times n}$:

Theorem 5.11 (Classical solutions are viscosity solutions)**.** *If u is a classical subsolution (resp. supersolution) of* (5.1)*, then it is also a viscosity subsolution (resp. supersolution) of the same problem whenever $E_*(x, u, p, \cdot)$ (resp. $E^*(x, u, p, \cdot)$) is non-increasing in* $\mathrm{Sym}^{n\times n}$.

Proof. We just study the case of subsolutions. For a test function φ, if $u-\varphi$ has a local maximum at a point x then we know by elementary calculus that $\nabla u(x)=\nabla\varphi(x)$ and $\nabla^2 u(x)\le\nabla^2\varphi(x)$ and, by definition, $E_*(x,u(x),\nabla u(x),\nabla^2 u(x))\le 0$. Consequently, exploiting our monotonicity assumption we obtain $E_*(x,u(x),\nabla\varphi(x),\nabla^2\varphi(x))\le 0$ and the conclusion follows. □

Before going further, we need to spend some words on conventions. First of all, it should be clear that this theory also applies to parabolic equations such as $(\partial_t-\Delta)u-g=0$, simply letting $x:=(y,t)\in\mathbb{R}^n\times[0,\infty)$, with $A=\mathbb{R}^n\times[0,\infty)$. Secondly, it is worth remarking that some authors adopt a different sign convention, which we might call *elliptic convention*, which is "opposite" to the one we gave before. Indeed, according to such a convention, when (for instance) dealing with a problem of the form $E(\nabla^2 u)=0$, it is required for a subsolution that if $u^*-\varphi$ has a maximum at x then $E(\nabla^2\varphi(x))\ge 0$ (i.e. a subsolution of $-E(\nabla^2 u)=0$ in our terminology). As a consequence, in the previous theorem, one should replace "non-increasing" with "non-decreasing."

Now, we are ready to introduce the first important tool for our discussion.

Theorem 5.12. *Let $\mathcal{F}$ be a nonempty family of viscosity subsolutions of* (5.1) *in A and let $u:A\to\overline{\mathbb{R}}$ be defined by*

$$u(x):=\sup\{v(x):\ v\in\mathcal{F}\}.$$

Then u is a viscosity subsolution of the same problem on the domain $A\cap\{u^*<+\infty\}$ (since the set $\{u^*<+\infty\}$ is open, the domain is still locally compact).

Proof. Assume as usual that $u^*-\varphi$ has a strict local maximum at x, equal to 0, and denote by K the compact set $\overline{B}_r(x)\cap A$ for some $r>0$ to be chosen sufficiently small, so that x is the unique maximum of $u^*-\varphi$ on K.

By means of a standard argument one can find a sequence (x_h) in K, convergent to x, and a sequence of functions $(v_h)\subset\mathcal{F}$ such that $u^*(x)=\lim_h u(x_h)=\lim_h v_h(x_h)$. Hence, if we denote by y_h a maximum point of $v_h^*-\varphi$ on K, then

$$u^*(y_h)-\varphi(y_h)\ge v_h^*(y_h)-\varphi(y_h)\ge v_h^*(x_h)-\varphi(x_h)\ge v_h(x_h)-\varphi(x_h). \tag{5.4}$$

Since by our construction we have $v_h(x_h)-\varphi(x_h)\to 0$ for $h\to\infty$, we get that every limit point $y\in K$ of (y_h) satisfies

$$u^*(y)-\varphi(y)\ge 0.$$

Hence y is a maximum in K of $u^*-\varphi$, $u^*(y)-\varphi(y)=0$ and y must coincide with x. As a result $y_h\to x$, $\limsup_h(u^*(y_h)-\varphi(y_h))\le u^*(x)-\varphi(x)=0$ and, combining this information with (5.4), we obtain $v_h^*(y_h)\to u^*(x)$. In order to conclude, we just need to consider the viscosity condition at the points y_h, which reads

$$E_*(y_h,v_h^*(y_h),\nabla\varphi(y_h),\nabla^2\varphi(y_h))\le 0,$$

and let $h \to \infty$ to get

$$E_*(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \leq 0.$$

□

We can now state a first existence result. This result builds on Perron's method [80] for harmonic functions and thus is named after Perron. The application of this technique to viscosity solutions is due to Ishii [54] (for the Hamilton-Jacobi equation) and was then extended to fully nonlinear second-order elliptic partial differential equations in [55] and [56]. See also [45] for a survey on the topic.

Theorem 5.13 (Perron, [80])**.** *Let f and g be respectively a viscosity subsolution and a viscosity supersolution of* (5.1)*, such that $f_* > -\infty$ and $g^* < +\infty$ on A. If $f \leq g$ on A and the functions $E_*(x, u, p, \cdot)$ and $E^*(x, u, p, \cdot)$ are non-increasing (in* $\mathrm{Sym}^{n\times n}$*), then there exists a viscosity solution u of* (5.1) *satisfying $f \leq u \leq g$.*

Proof. Let

$$\mathcal{F} := \{v : v \text{ is a viscosity subsolution of (5.1) and } v \leq g\}.$$

We know that $f \in \mathcal{F}$, so that this set is not empty. Hence, we can define pointwise a function $u : A \to \overline{\mathbb{R}}$ as

$$u(x) := \sup\{v(x) : v \in \mathcal{F}\}.$$

A fortiori, by our definition of $\mathcal{F}$, we have that $u \leq g$ and therefore $u^* \leq g^* < +\infty$, so that by Theorem 5.12 u is a subsolution on A. Since $u_* \geq f_* > -\infty$ in A, we just need to prove that it is also a supersolution on the same domain.

Pick a test function φ such that $u_* - \varphi$ has a relative minimum, equal to 0, at x_0. Without loss of generality, we can assume that

$$u_*(x) - \varphi(x) \geq |x - x_0|^4 \qquad \text{on } A \cap \overline{B}_r(x_0) \tag{5.5}$$

for some sufficiently small $r > 0$. Assume by contradiction that

$$E^*(x_0, u_*(x_0), \nabla\varphi(x_0), \nabla^2\varphi(x_0)) < 0 \tag{5.6}$$

and define a function $w := \max\{\varphi + \delta^4, u\}$ for some parameter $\delta > 0$ to be fixed later in the proof. In fact, we claim that for any sufficiently small $\delta > 0$ each of the following assertions hold:

- w is a viscosity subsolution of (5.1);
- $\{w > u\} \neq \emptyset$;
- $w \leq g$;

and hence w belongs to $\mathcal{F}$ and is larger than u, contradicting the very definition of u. Let us now justify such claims, which will then complete the proof.

It is easily proved, again by contradiction and exploiting the fact that E^* is upper semicontinuous, that for $\delta > 0$ sufficiently small we have

$$E^*(x, \varphi(x) + \delta^4, \nabla\varphi(x), \nabla^2\varphi(x)) \leq 0 \qquad \text{on } \overline{B}_{2\delta}(x_0) \cap A.$$

This means that $\varphi + \delta^4$ is a classical subsolution of (5.1) on this domain and hence, by our monotonicity hypothesis, it has to be also a viscosity subsolution. Consequently, by a very special case of the previous theorem, we get that the function w is a viscosity subsolution of (5.1) on $\overline{B}_{2\delta}(x_0) \cap A$. Moreover, by (5.5), we know that $w = u$ on $A \cap B_r(x) \setminus \overline{B}_\delta(x_0)$. Since the notions of viscosity subsolution and supersolution are clearly local, w is a global viscosity subsolution on A.[6]

To prove that $\{w > u\} \neq \emptyset$ we just need to observe that, for any $\delta > 0$, $u_*(x_0) = \varphi(x_0) < \varphi(x_0) + \delta^4$, and hence, by the very definition of $u_*(x_0)$ there ought to exist a sequence (x_h) converging to x_0 and such that $u(x_h) \to u_*(x_0)$, thereby implying for h sufficiently large the inequality $u(x_h) < \varphi(x_h) + \delta^4 \leq w(x_h)$.

Finally, we have to show that $w \leq g$: this completes the proof of the claim and gives the desired contradiction. To this aim, it is enough to prove that there exists $\delta > 0$ such that $\varphi + \delta^4 \leq g$ on $A \cap \overline{B}_\delta(x_0)$. But this readily follows, by an elementary argument, showing that $\varphi(x_0) = u_*(x_0) < g_*(x_0)$. Again, assume by contradiction that $u_*(x_0) = g_*(x_0)$: if this were the case, the function $g_* - \varphi$ would have a local minimum at x_0 and so, since g_* is a viscosity supersolution, we would get

$$E^*(x_0, g_*(x_0), \nabla\varphi(x_0), \nabla^2\varphi(x_0)) \geq 0,$$

which is in contrast with (5.6). □

5.3 The distance function

Our next goal is now to study the uniqueness problem, which is actually very delicate as anticipated in Remark 5.7. We shall start here with the discussion of a special case.

Let $C \subset \mathbb{R}^n$ be a closed set, $C \neq \emptyset$ and let $u(x) := \text{dist}(x, C)$. We claim that this distance function u is a viscosity solution of the equation $|p|^2 - 1 = 0$ on $A := \mathbb{R}^n \setminus C$.

First of all, it is a viscosity supersolution in A. This follows by Theorem 5.12 (in the obvious symmetric version for supersolutions), once we observe that $u(x) = \inf_{y \in C} |x - y|$ and that, for any $y \in C$, the function $x \mapsto |x - y|$ is a classical supersolution in A (because $y \notin A$) and hence a viscosity supersolution of our problem.

The fact that u is also a subsolution is proven as a direct consequence of the following general implication.

[6] We mean that, if $A = A_1 \cup A_2$ and we know that u is a subsolution both on A_1 and A_2, relatively open in A, then u is also a subsolution on A.

Remark 5.14. Let $f : A \to \overline{\mathbb{R}}$ be a Lipschitz function, then

$$\mathrm{Lip}(f, A) \le 1 \;\Rightarrow\; |\nabla f|^2 - 1 \le 0 \qquad \text{on } A \text{ in the sense of viscosity solutions.}$$

Indeed, let x be a local maximum for $f - \varphi$, so that $f(y) - \varphi(y) \le f(x) - \varphi(x)$ for any $y \in B_r(x)$ (and r small enough). When $\mathrm{Lip}(f, A) \le 1$ this implies, on the same domain, the inequality $\varphi(y) - \varphi(x) \ge f(y) - f(x) \ge -|y - x|$ and, by Taylor expanding the smooth function φ, we finally get

$$\langle \nabla\varphi(x), y - x\rangle + o(|y - x|) \ge -|y - x|.$$

This readily implies the claim.

The converse implication is less trivial, but still true. Namely

$$|\nabla f|^2 - 1 \le 0 \text{ on } A \text{ in the sense of viscosity solutions} \quad \Rightarrow \quad \mathrm{Lip}(f, A) \le 1$$

for f continuous, or at least upper semicontinuous. This is proven exploiting the regularizations $f^\varepsilon(x) := \sup_y\big(f(y) - |x - y|^2/\varepsilon\big)$ that we will study more in detail later on in this chapter. We just sketch here the structure of the argument:

(i) still $|\nabla f^\varepsilon|^2 - 1 \le 0$ on A, in the sense of viscosity solutions;

(ii) $|\nabla f^\varepsilon|^2 - 1 \le 0$ pointwise $\mathscr{L}^n$-a.e., because f^ε is semiconvex, hence locally Lipschitz, and therefore the inequality holds at any differentiability point by the superjet characterization of viscosity subsolutions;

(iii) by Proposition A.6 one obtains $\mathrm{Lip}(f^\varepsilon, A) \le 1$;

(iv) $f^\varepsilon \downarrow f$ as one lets $\varepsilon \downarrow 0$ and hence $\mathrm{Lip}(f, A) \le 1$.

We now come to the uniqueness result.

Theorem 5.15. *Let $C \subset \mathbb{R}^n$ be a closed set as above, $A = \mathbb{R}^n \setminus C$ and let $u \in C(\overline{A}; \mathbb{R})$ be a non-negative viscosity solution of $|\nabla u|^2 - 1 = 0$ on A with $u = 0$ on ∂A. Then $u(x) = \mathrm{dist}(x, C)$.*

Proof. By our assumptions we can clearly extend u continuously to $\mathbb{R}^n$, so that $u = 0$ identically on C. It is straightforward to verify that $|\nabla u|^2 - 1 \le 0$ in the sense of viscosity solutions on $\mathbb{R}^n$ (actually, one only need to check this fact at the boundary points, namely on ∂A, where the implication (b) in Definition 5.3 is easy to check: if the graph of a smooth φ touches the graph of u at a point $x \in \partial A$ from above, then necessarily $\varphi \ge 0$ in a neighborhood of x and $\varphi(x) = u(x) = 0$ hence $\nabla\varphi(x) = 0$, which implies the claim). Consequently, thanks to the previous regularization argument, $\mathrm{Lip}(u, \mathbb{R}^n) \le 1$ and thus,

for any $y \in C$, we have that $u(x) \leq |x-y|$, which means $u(x) \leq \mathrm{dist}(x, C)$. In the sequel, in order to simplify the notation, we write $w(x)$ for the distance function $\mathrm{dist}(x, C)$.

It remains to show that $w \leq u$. Assume first that A is bounded: we will show later on that this is not restrictive. By contradiction, assume that $w(x_0) > u(x_0)$ for some x_0; in this case there exist $\lambda_0 > 0$ and $\gamma_0 > 0$ such that

$$\sup_{x,y\in\mathbb{R}^n} \left\{ w(x) - (1+\lambda)u(y) - \frac{1}{2\varepsilon}|x-y|^2 \right\} \geq \gamma_0$$

for all $\varepsilon > 0$ and $\lambda \in (0, \lambda_0)$. Indeed, it suffices to bound from below the supremum with $w(x_0) - (1+\lambda)u(x_0)$, which is larger than $\gamma_0 := (w(x_0) - u(x_0))/2$ for $\lambda > 0$ small enough.

Moreover, for $\varepsilon > 0$ and $\lambda \in (0, \lambda_0)$, the supremum is actually a maximum because it is clear that we can localize x in A (otherwise the whole sum above is non-positive) and y in a bounded set of $\mathbb{R}^n$ (because w is bounded on A, and again for $|y-x|$ large the whole sum is non-positive). So, let $(\overline{x}, \overline{y})$ be a maximizing couple, omitting for notational simplicity the dependence on the parameters ε, λ. The function $x \mapsto w(x) - \frac{1}{2\varepsilon}|x-\overline{y}|^2$ has a maximum at $x = \overline{x}$ and so we can exploit the fact that $w(\cdot)$ is a viscosity solution of our equation (with respect to the test function $\varphi(x) = |x-\overline{y}|^2/(2\varepsilon)$) to derive $|\nabla\varphi|^2(\overline{x}) \leq 1$, that is

$$\frac{|\overline{x}-\overline{y}|}{\varepsilon} \leq 1.$$

We also claim that necessarily $\overline{y} \in A$, if ε is sufficiently small, precisely if $\varepsilon < \gamma_0$. Indeed, assume by contradiction that $\overline{y} \notin A$, so that $u(\overline{y}) = 0$, then by the triangle inequality

$$\gamma_0 \leq w(\overline{x}) - \frac{1}{2\varepsilon}|\overline{x}-\overline{y}|^2 \leq |\overline{x}-\overline{y}| - \frac{1}{2\varepsilon}|\overline{x}-\overline{y}|^2 \leq |\overline{x}-\overline{y}|.$$

As a consequence, we get $\gamma_0 \leq |\overline{x}-\overline{y}| \leq \varepsilon$, which gives a contradiction.

Now, choosing $\varepsilon > 0$ so that $\overline{y} \in A$, the function $y \mapsto (1+\lambda)u(y) + \frac{1}{2\varepsilon}|\overline{x}-y|^2$ has a minimum at $y = \overline{y}$ and arguing as above we obtain

$$\left| \frac{\overline{x}-\overline{y}}{\varepsilon} \right| \geq (1+\lambda),$$

which is not compatible with $|\overline{x}-\overline{y}| \leq \varepsilon$. Hence, at least when A is bounded, we have proven that $w = u$.

In the general case, fix a constant $R > 0$ and define

$$u_R(x) := \min\left\{u(x), \mathrm{dist}(x, \mathbb{R}^n \setminus \overline{B}_R)\right\}.$$

This is a supersolution of our problem on $A \cap B_R$, since $u(x)$ is a supersolution on A and $\mathrm{dist}(x, \mathbb{R}^n \setminus \overline{B}_R)$ is a supersolution on B_R (again by Theorem 5.12). Moreover, $\mathrm{Lip}(u_R, \mathbb{R}^n) \leq 1$ implies that u_R is a global subsolution and hence the discussion

above can be applied to the function u_R (which satifies the null boundary conditions on $\partial(A \cap B_R) = (\partial A \cap B_R) \cup (A \cap \partial B_R))$ to give

$$u_R(x) = \mathrm{dist}(x, \mathbb{R}^n \setminus (A \cap B_R)).$$

At this stage, the conclusion comes at once by simply letting $R \to \infty$. $\square$

Remark 5.16. We can also apply a similar argument to the study of the case when $A = \mathbb{R}^n$ (thus $\partial A = \emptyset$) and we replace u by $u - \inf u$. In the spirit of the classical Liouville's theorems we can say that "the equation $|\nabla u|^2 - 1 = 0$ does not have entire viscosity solutions on $\mathbb{R}^n$ that are bounded from below". Nevertheless, there exist trivial examples of functions that solve this equation in the viscosity sense and are unbounded from below (e.g. take $u(x) = x_i$ for some $i \in \{1, \dots, n\}$).

5.4 Maximum principle for semiconvex functions

We now turn to the case of second-order problems having the form $F(\nabla u, \nabla^2 u) = 0$ on an open domain $A \subset \mathbb{R}^n$. We always assume that $F(p, S)$ is non-increasing in its second variable $S \in \mathrm{Sym}^{n\times n}$, so that classical solutions are viscosity solutions.

Let us begin with some heuristics. Let f, $g \in C^2(A; \mathbb{R}) \cap C(\overline{A}; \mathbb{R})$, with A bounded, and assume that f is a subsolution on A, g is a supersolution on A, $f \le g$ on ∂A and that, at all points, at least one of the two inequalities $F(\nabla f, \nabla^2 f) \le 0$, $F(\nabla g, \nabla^2 g) \ge 0$ is strict. Then $f \le g$ in A. Indeed, assume by contradiction $\sup_A (f - g) > 0$, then there exists $x_0 \in A$ which is an interior maximum point for the function $f - g$. Consequently $\nabla f(x_0) = \nabla g(x_0)$ as well as $\nabla^2 f(x_0) \le \nabla^2 g(x_0)$. These two facts imply, by the monotonicity of F, that

$$F(\nabla f(x_0), \nabla^2 f(x_0)) \ge F(\nabla g(x_0), \nabla^2 g(x_0)) \ . \tag{5.7}$$

On the other hand, f (resp. g) is also a regular subsolution (resp. supersolution) so that

$$F(\nabla f(x_0), \nabla^2 f(x_0)) \le 0, \qquad F(\nabla g(x_0), \nabla^2 g(x_0)) \ge 0 \ . \tag{5.8}$$

Hence, if we compare (5.7) with (5.8), we find a contradiction as soon as one of the two inequalities in (5.8) is strict.

In order to hope for a comparison principle, this argument shows the necessity to approximate subsolutions (supersolutions) with strict subsolutions (respectively, strict supersolutions), and this is always linked to some form of strict monotonicity of the equation, variable from case to case.

When aiming at a general uniqueness result for viscosity solutions, we cannot just argue as in the case of the distance function. Some partial uniqueness results are available in [21] and [23] (for first-order problems) and in [71] (under convexity assumptions). A breakthrough in the problem was a strategy introduced by Jensen in [58] (see also [60],

[59], [57] and [93] for extensions of the result). The first step in that direction is to obtain a refined version of the maximum principle *à la* Bony [7]: with that goal in mind, we start with an elementary observation.

Remark 5.17. If $(p, S) \in J_2^- u(x)$ and u has a relative maximum at x, then necessarily $p = 0$ and $S \leq 0$. To see this, it is enough to apply the definitions: by our two hypotheses

$$0 \geq u(y) - u(x) \geq \langle p, y - x\rangle + \frac{1}{2}\langle S(y-x), y-x\rangle + o(|y-x|^2)$$

and hence

$$\langle p, \frac{y-x}{|y-x|}\rangle \leq O(|y-x|) \ \Rightarrow p = 0,$$
$$\frac{\langle S(y-x), y-x\rangle}{|y-x|^2} \leq o(1) \ \Rightarrow S \leq 0.$$

We are now ready to state and prove Jensen's maximum principle for semiconvex functions.

Theorem 5.18 (Jensen's maximum principle, [58])**.** *Let $\Omega \subset \mathbb{R}^n$ be open, let $u : \Omega \to \mathbb{R}$ be semiconvex and let $x_0 \in \Omega$ be a local maximum for u. Then, there exist a sequence (x_k) in Ω converging to x_0 and a sequence of real numbers $\varepsilon_k \downarrow 0$ such that u is pointwise second-order differentiable at x_k and*

$$\nabla u(x_k) \to 0\,, \qquad \nabla^2 u(x_k) \leq \varepsilon_k I.$$

The proof is based on the following lemma. In the sequel we shall denote by $\mathrm{sc}(u, \Omega)$ the least non-negative constant c such that u is $(-c)$-convex, i.e. $u + c|x|^2/2$ is convex (recall Definition D.8).

Theorem 5.19. *Let $B_R \subset \mathbb{R}^n$ denote the open ball of radius R centered at the origin and let $u \in C(\overline{B}_R; \mathbb{R})$ be semiconvex, with*

$$\max_{\overline{B}_R} u > \max_{\partial B_R} u$$

(notice that this implies $\mathrm{sc}(u, B) > 0$, since $\max_{\overline{B}_R} w = \max_{\partial B_R} w$ for any convex function w). Then, if we let

$$G^\delta = \left\{x \in B_R : \exists\, p \in \overline{B}_\delta \ s.t.\ u(y) \leq u(x) - \langle p, x-y\rangle,\ \forall y \in B_R\right\},$$

it must be

$$\mathscr{L}^n(G^\delta) \geq \frac{\omega_n \delta^n}{[\mathrm{sc}(u, B)]^n} \tag{5.9}$$

for all $0 < \delta < \left(\max_{\overline{B}} u - \min_{\overline{B}} u\right)/(2R)$.

Proof. We assume first that u also belongs to $C^1(B_R;\mathbb{R})$. Let us explicitly observe that, for any given $\delta > 0$, one has the inclusion $\nabla u(G^\delta) \subset \overline{B}_\delta$: indeed, straight from the definition of the set G^δ, we get that the function $y \mapsto u(y) - \langle p, y\rangle$ has a local maximum at x. Hence necessarily $\nabla u(x) = p \in \overline{B}_\delta$. In order to proceed and gain the opposite inclusion, pick a $\delta > 0$, so small that $2R\delta < \max_{\overline{B}_R} u - \max_{\partial B_R} u$, and consider a perturbation $u(y) - \langle p, y\rangle$ with $|p| \le \delta$. We claim that such a function necessarily attains its maximum in B_R. Indeed, this comes from the two inequalities

$$\max_{\partial B_R} (u - \langle p, y\rangle) \le \max_{\partial B_R} u + \delta R$$

and

$$\max_{\overline{B}_R}(u - \langle p, y\rangle) \ge \max_{\overline{B}_R} u - \delta R.$$

Consequently, there exists $x \in B_R$ such that $\nabla u(x) = p$. This shows that $\nabla u(G^\delta) = \overline{B}_\delta$. To proceed in the proof, we employ the area formula (see Theorem D.14). In this case, it gives

$$\int_{G^\delta} |\det \nabla^2 u|\, dx = \int_{\overline{B}_\delta} \operatorname{card}\big(\{x \in G^\delta :\ \nabla u(x) = p\}\big)\ dp \ge \omega_n \delta^n$$

by the previous statement. On the other hand

$$\int_{G^\delta} \big|\det \nabla^2 u\big|\ dx \le [\mathrm{sc}(u, B)]^n\, \mathscr{L}^n(G^\delta),$$

because the points in G^δ are maxima for the function $u(y) - \langle p, y\rangle$: this implies $\nabla^2 u(x) \le 0$ for any $x \in G^\delta$ and, by semiconvexity, $\nabla^2 u(x) \ge -\mathrm{sc}(u, B_R)I$, where I is the identity matrix in $\mathrm{Sym}^{n\times n}$. If we combine these two inequalities, we get (5.9).
In the general case we argue by approximation, finding radii $r_h \uparrow R$ and smooth functions u_h in $\overline{B}_{r_h}$ such that $u_h \to u$ locally uniformly in B_R and $\limsup_h \mathrm{sc}(u_h, B_{r_h}) \le \mathrm{sc}(u, B_R)$; to conclude, it suffices to notice that any limit of points in $G^\delta(u_h) \cap B_{r_h}$ belongs to $G^\delta(u)$, hence $\mathscr{L}^n(G^\delta(u)) \ge \limsup_h \mathscr{L}^n(G^\delta(u_h) \cap B_{r_h})$. □

We can now get back to justifying Jensen's maximum principle.

Proof. Recall that, by assumption, x_0 is a local maximum of u. We can choose $R > 0$ sufficiently small so that $u \le u(x_0)$ in $\overline{B}_R(x_0)$ and, without loss of generality, we can assume $u(x_0) = 0$. This becomes a strict local maximum for the function $\widetilde{u}(x) = u(x) - |x - x_0|^4$. It is also easy to verify that $\widetilde{u}$ is semiconvex in $\overline{B}_R(x_0)$. We now apply Theorem 5.19 to $\widetilde{u}$: for any $\delta = 1/k$ with k large enough we obtain that $\mathscr{L}^n(G^{1/k}) > 0$ and (thanks to Theorem D.15) this means that there exists a sequence of points (x_k) such that $\widetilde{u}$ is pointwise second-order differentiable at x_k and, for appropriate vectors p_k with $|p_k| \le 1/k$, the function $\widetilde{u}(y) - \langle p_k, y\rangle$ has a local maximum at x_k. Since $|p_k| \to 0$, any limit point of (x_k) for $k \to \infty$ has to be a local maximum for $\widetilde{u}$, but in $\overline{B}_R(x_0)$ this necessarily

implies $x_k \to x_0$. Moreover $p_k = \nabla\widetilde{u}(x_k) \to 0$ and $\nabla^2\widetilde{u}(x_k) \le 0$ (cf. Remark 5.17). As a consequence

$$\nabla u(x_k) = \nabla\widetilde{u}(x_k) + 4|x_k - x_0|^2(x_k - x_0) \to 0$$

and the identity

$$\nabla^2|z|^4 = 4|z|^2 I + 8z \otimes z \tag{5.10}$$

gives

$$\begin{aligned} \nabla^2 u(x_k) &= \nabla^2\widetilde{u}(x_k) + 8(x_k - x_0) \otimes (x_k - x_0) + 4|x_k - x_0|^2 I \\ &\le \nabla^2\widetilde{u}(x_k) + 12|x_k - x_0|^2 I. \end{aligned}$$

Setting $\varepsilon_k = 12|x_k - x_0|^2$ we get the thesis. □

We now introduce another important tool in the theory of viscosity solutions.

Definition 5.20 (Inf and sup-convolutions)**.** *Given $u : A \to \mathbb{R}$ and a parameter $\varepsilon > 0$, for every $x \in \mathbb{R}^n$ we define the regularized functions*

$$u^\varepsilon(x) := \sup_{y \in A} \left\{ u(y) - \frac{1}{\varepsilon}|x - y|^2 \right\} \tag{5.11}$$

which are called sup-convolutions of u and satisfy $u^\varepsilon \ge u$ in A, and, for every $x \in \mathbb{R}^n$

$$u_\varepsilon(x) := \inf_{y \in A} \left\{ u(y) + \frac{1}{\varepsilon}|x - y|^2 \right\} \tag{5.12}$$

which are called inf-convolutions of u and satisfy $u_\varepsilon \le u$ in A.

In the next proposition we summarize the main properties of sup-convolutions; analogous properties hold for inf-convolutions.

Proposition 5.21 (Properties of sup-convolutions)**.** *Assume that u is u.s.c. on A and that $u(x) \le c(1 + |x|)$ for some constant $c \ge 0$, then*

(i) *u^ε is semiconvex and $\mathrm{sc}(u^\varepsilon, \mathbb{R}^n) \le 2/\varepsilon$;*

(ii) *$u^\varepsilon \ge u$ and $u^\varepsilon \downarrow u$ pointwise in A. If u is continuous, then $u^\varepsilon \downarrow u$ locally uniformly;*

(iii) *if $F(\nabla u, \nabla^2 u) \le 0$ in the sense of viscosity solutions on A, then $F(\nabla u^\varepsilon, \nabla^2 u^\varepsilon) \le 0$ on A^ε, where*

$$A^\varepsilon := \{x \in \mathbb{R}^n : \textit{the supremum in } (5.11) \textit{ is attained}\}.$$

Proof. To prove *(i)*, notice that, by the linear growth assumption, the function u^ε is real-valued for any $\varepsilon > 0$. Moreover by its very definition

$$u^\varepsilon(x) + \frac{1}{\varepsilon}|x|^2 = \sup_{y\in A}\left(u(y) - \frac{1}{\varepsilon}|y|^2 + \frac{2}{\varepsilon}\langle x, y\rangle\right)$$

and the functions in the right-hand side are affine with respect to x. It follows that the left-hand side is convex, which means $\mathrm{sc}(u^\varepsilon, \mathbb{R}^n) \le 2/\varepsilon$.

Concerning *(ii)*, the inequality $u^\varepsilon \ge u$ and the monotonicity in ε are trivial. In addition, we can take quasi-maxima (y_ε) satisfying

$$u^\varepsilon(x) \le u(y_\varepsilon) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon \le c(1 + |y_\varepsilon|) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon \le c(1 + |x| + |\delta_\varepsilon|) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon$$

with $\delta_\varepsilon = |y_\varepsilon - x|$. Via these inequalities, one first sees that $y_\varepsilon \to x$ as $\varepsilon \to 0$, and then, exploiting the upper semicontinuity of u and neglecting the quadratic term in the first inequality, one gets

$$u(x) \ge \limsup_{\varepsilon\to 0} u(y_\varepsilon) \ge \limsup_{\varepsilon\to 0} u^\varepsilon(x),$$

which gives pointwise convergence of (u_ε) to u. If u is continuous, the claim comes from Dini's monotone convergence theorem and the local compactness of A.

Proceeding to *(iii)*, let now $x_0 \in A^\varepsilon$ and let $y_0 \in A$ be a corresponding maximum point, so that $u^\varepsilon(x_0) = u(y_0) - |x_0 - y_0|^2/\varepsilon$. Let then φ be a smooth function such that $u^\varepsilon - \varphi$ has a local maximum in x_0 and, without loss of generality, $u^\varepsilon(x_0) = \varphi(x_0)$; correspondingly let $r > 0$ be such that $u^\varepsilon \le \varphi$ on $B_r(x_0)$. Define $\psi(x) := \varphi(x - y_0 + x_0)$: we claim that $u - \psi$ has a local maximum at y_0 with value $|x_0 - y_0|^2/\varepsilon$. If we prove this claim, then it must be

$$F(\nabla\psi(y_0), \nabla^2\psi(y_0)) \le 0$$

and, by the definition of ψ, this is indeed equivalent to *(iii)*, namely

$$F(\nabla\varphi(x_0), \nabla^2\varphi(x_0)) \le 0.$$

Thus it is enough to prove the claim. On the one hand

$$u(y_0) - \psi(y_0) = u(y_0) - \varphi(x_0) = u(y_0) - u^\varepsilon(x_0) = \frac{1}{\varepsilon}|x_0 - y_0|^2,$$

while on the other hand $u^\varepsilon(x) \le \varphi(x)$ in $B_r(x_0)$ gives

$$u(y) - \frac{1}{\varepsilon}|x - y|^2 \le \varphi(x) \quad \forall x \in B_r(x_0),\ \forall y \in A.$$

Letting $y = x - x_0 + y_0 \in A$ with $x \in B_r(x_0)$, this implies

$$u(y) - \psi(y) \le \frac{1}{\varepsilon}|x_0 - y_0|^2 \quad \forall y \in A \cap B_r(y_0)$$

as was to be shown. □

Remark 5.22. *We will also need an x-dependent version of the previous result, that reads as follows: if $F(x, \nabla u, \nabla^2 u) \leq 0$ in the sense of viscosity solutions on A, then for all $\delta > 0$ there holds $F^\delta(x, \nabla u^\varepsilon, \nabla^2 u^\varepsilon) \leq 0$ on $A^{\varepsilon,\delta}$, where*

$$A^{\varepsilon,\delta} := \{x \in \mathbb{R}^n : \textit{the supremum in } (5.11) \textit{ is attained at some } y \in B_\delta(x) \cap A\},$$

and

$$F^\delta(x, p, S) := \inf \{F(y, p, s) : y \in B_\delta(x) \cap A\}.$$

An analogous result holds for supersolutions.

5.5 Existence and uniqueness results

In this section we collect some existence and uniqueness results for certain classes of second-order equations. The main tool we will employ to that aim is the comparison principle, that is stated below (see also [58], [60] and [93], for instance). Throughout this section we shall always assume that A is a bounded open set in $\mathbb{R}^n$.

Proposition 5.23 (Comparison principle)**.** *Let $F : A \times \mathrm{Sym}^{n\times n} \to \mathbb{R}$ be continuous and satisfy, for some $\lambda > 0$, the strict monotonicity condition*

$$F(x, S + tI) \geq F(x, S) + \lambda t, \qquad \forall t \geq 0$$

as well as the uniform continuity assumption

$$\textit{the family of functions } \left\{F(\cdot, S) \ : S \in \mathrm{Sym}^{n\times n}\right\} \textit{ is equicontinuous in } A.$$

Let $\underline{u}$, $\overline{u} : A \to \mathbb{R}$ be respectively a bounded u.s.c. subsolution and a bounded l.s.c. supersolution to $-F(x, \nabla^2 u) = 0$ in A, with $(\underline{u})^* \leq (\overline{u})_*$ on ∂A. Then $\underline{u} \leq \overline{u}$ on A.

Notice that the uniform continuity assumption, though restrictive, covers all equations of the form $G(\nabla^2 u) + f = 0$ with f continuous in A.

A direct consequence of the comparison principle (that is obtained considering the special case when $\underline{u} = \overline{u} = u$) is the following uniqueness result for fully nonlinear boundary value problems:

Theorem 5.24 (Uniqueness of continuous solutions)**.** *Let F be as in Proposition 5.23 and $h \in C(\partial A; \mathbb{R})$. Then the problem*

$$\begin{cases} -F(x, \nabla^2 u(x)) = 0 & \textit{in } A \\ \\ u = h & \textit{on } \partial A \end{cases} \tag{5.13}$$

admits at most one viscosity solution $u \in C(\overline{A}; \mathbb{R})$.

At the level of existence, we can instead exploit Theorem 5.13 to obtain the following result.

Theorem 5.25 (Existence of continuous solutions)**.** *Let F be as in Proposition 5.23 and let f, $g : A \to \mathbb{R}$ be respectively a subsolution and a supersolution of $-F(x, \nabla^2 u) = 0$ in A, such that $f_* > -\infty$, $g^* < +\infty$ and $f \leq g$ on A. If $g^* \leq f_*$ on ∂A, then there exists a solution u to* (5.13)*, with $u = g^* = f_*$.*

In order to prove this last result, it suffices to take any solution u given by Perron's method (see Theorem 5.13), so that $f \leq u \leq g$ in A. It follows that $u^* \leq g^* \leq f_* \leq u_*$ on ∂A and the comparison principle (with $\underline{u} = u^*$, $\overline{u} = u_*$) gives $u^* \leq u_*$ on A, i.e. u is continuous.

The rest of the section is devoted to the proof of the comparison principle, which exploits a number of different tools: doubling of variables, inf and sup-convolutions (see Definition 5.20) and Jensen's maximum principle (see Theorem 5.18). The next lemma is useful to reduce ourselves, in the proof of the comparison principle, to the case of a strict subsolution $\underline{u}$.

Lemma 5.26. *Let F, $\underline{u}$ and $\overline{u}$ be as in Proposition 5.23 and set*

$$F_\gamma(x, S) := F(x, S - \gamma I) \leq F(x, S) - \gamma\lambda,$$

with $\gamma > 0$. For any $\delta > 0$, consider the function

$$v_{\delta,\gamma} := \underline{u} - \delta + \frac{\gamma}{2}|x|^2.$$

Then:

(i) $v_{\delta,\gamma}$ satisfies $-F_\gamma(x, \nabla^2 v_{\delta,\gamma}) \leq 0$ in the viscosity sense;

(ii) if $\delta \geq \delta(\gamma, A)$, then $v_{\delta,\gamma} \leq \overline{u}$ on ∂A and $\delta(\gamma, A) \to 0$ as $\gamma \downarrow 0$.

(iii) if the comparison principle holds for $v_{\delta,\gamma}$ for any $\delta > \delta(\gamma, A)$, that is to say

$$v_{\delta,\gamma} \leq \overline{u} \quad \text{on } A, \ \forall\, \delta > \delta(\gamma, A), \tag{5.14}$$

then $\underline{u} \leq \overline{u}$ on A.

Proof. Statement *(i)* is a direct consequence of the definition of the functional F_γ together with the identity $\nabla^2 \left(\varphi + \gamma \frac{|x|^2}{2}\right) = \nabla^2 \varphi + \gamma I$ for every test function $\varphi \in C^\infty(\mathbb{R}^n; \mathbb{R})$. Statement *(ii)* follows by the fact that $(\underline{u})^* < (\overline{u})_*$ on ∂A. Finally, if (5.14) holds, then

$$\underline{u} - \delta \leq v_{\delta,\gamma} \leq \overline{u} \qquad \text{on } A,$$

and the comparison principle for $\underline{u}$ is obtained by letting $\gamma \downarrow 0$, which allows to choose arbitrarily small δ in view of *(ii)*. □

We shall now proceed to the proof of Proposition 5.23.

Proof. Thanks to Lemma 5.26, without loss of generality (possibly by replacing $\underline{u}$ by $v_{\delta,\gamma}$) we can assume that $\underline{u}$ satisfies the stronger property

$$-F_\gamma(x, \nabla^2 \underline{u}) \leq 0$$

in the viscosity sense, for some $\gamma > 0$.

Assume by contradiction that $d_0 := \underline{u}(x_0) - \overline{u}(x_0) > 0$ for some $x_0 \in A$, and with a slight abuse of notation let us consider the sup convolution

$$u^\varepsilon(x) := \sup_{x' \in A} \left(\underline{u}(x') - \frac{1}{\varepsilon}|x - x'|^2 \right) = \max_{x' \in \overline{A}} \left((\underline{u})^*(x') - \frac{1}{\varepsilon}|x - x'|^2 \right) \tag{5.15}$$

of $\underline{u}$ and the inf convolution

$$u_\varepsilon(y) := \inf_{y' \in A} \left(\overline{u}(y') + \frac{1}{\varepsilon}|y - y'|^2 \right) = \min_{y' \in \overline{A}} \left((\overline{u})_*(y') + \frac{1}{\varepsilon}|y - y'|^2 \right) \tag{5.16}$$

of $\overline{u}$; since $u^\varepsilon \geq \underline{u}$ and $u_\varepsilon \leq \overline{u}$ on A we have

$$\max_{\overline{A} \times \overline{A}} \left(u^\varepsilon(x) - u_\varepsilon(y) - \frac{1}{4\varepsilon}|x - y|^4 \right) \geq u^\varepsilon(x_0) - u_\varepsilon(x_0) \geq \underline{u}(x_0) - \overline{u}(x_0) = d_0.$$

We shall denote by $(x_\varepsilon, y_\varepsilon) \in \overline{A} \times \overline{A}$ a maximizing pair, so that

$$d_0 + \frac{1}{4\varepsilon}|x_\varepsilon - y_\varepsilon|^4 \leq u^\varepsilon(x_\varepsilon) - u_\varepsilon(y_\varepsilon) \leq \sup \underline{u} - \inf \overline{u}. \tag{5.17}$$

Furthermore, let us denote by $x'_\varepsilon \in \overline{A}$ and $y'_\varepsilon \in \overline{A}$ a maximizing (respectively: minimizing) point for the variational problem defined by (5.15) with $x = x_\varepsilon$ (respectively: (5.16) with $y = y_\varepsilon$). We claim that:

(a) $\liminf_{\varepsilon \downarrow 0} \operatorname{dist}(x_\varepsilon, \partial A) > 0$ and $\liminf_{\varepsilon \downarrow 0} \operatorname{dist}(y_\varepsilon, \partial A) > 0$;

(b) set $M := \max\{\operatorname{osc}_A(\underline{u}), \operatorname{osc}_A(\overline{u})\}$, for ε small enough, the supremum in (5.15) with any $x \in A$ satisfying $|x - x_\varepsilon| < \varepsilon$ is attained at a point $x' \in A$ with $|x' - x|^2 \leq M\varepsilon$ and the infimum in (5.16) with any $y \in A$ satisfying $|y - y_\varepsilon| < \varepsilon$ is attained at a point $y' \in A$ with $|y' - y|^2 \leq M\varepsilon$.

To prove (a), notice that, if $(\bar{x}, \bar{y})$ is any limit point of $(x_\varepsilon, y_\varepsilon)$ as $\varepsilon \downarrow 0$, then (5.17) gives $\bar{x} = \bar{y}$ and

$$d_0 \leq \limsup_{\varepsilon \downarrow 0} \left((\underline{u})^*(x'_\varepsilon) - (\overline{u})_*(y'_\varepsilon) - \frac{|x_\varepsilon - x'_\varepsilon|^2 + |y_\varepsilon - y'_\varepsilon|^2}{\varepsilon} \right).$$

Since $\overline{u}$ and $\underline{u}$ are assumed to be bounded, such inequality implies that $|x_\varepsilon - x'_\varepsilon| \to 0$, $|y_\varepsilon - y'_\varepsilon| \to 0$, hence $(x'_\varepsilon, y'_\varepsilon) \to (\bar{x}, \bar{x})$ as well and by semicontinuity $d_0 \le (\underline{u})^*(\bar{x}) - (\overline{u})_*(\bar{x})$. By assumption $(\underline{u})^* \le (\overline{u})_*$ on ∂A, therefore $\bar{x} \in A$ and this gives (a). To prove (b), it suffices to choose (thanks to (a)) constants $\varepsilon_0 > 0$ and $\delta_0 > 0$ small enough, so that $\mathrm{dist}(x_\varepsilon, \partial A) \ge \delta_0$ for $\varepsilon \in (0, \varepsilon_0)$. In general, for $x \in A$ we have

$$\underline{u}(x') - \frac{1}{\varepsilon}|x' - x|^2 \le \underline{u}(x) + M - \frac{1}{\varepsilon}|x' - x|^2$$

which implies that the supremum in the definition of $u^\varepsilon(x)$ is unchanged if we maximize in the closed ball $\overline{B}_{\sqrt{M\varepsilon}}(x)$ centered at x with radius $\sqrt{M\varepsilon}$. If $|x - x_\varepsilon| < \varepsilon$ and $\varepsilon < \varepsilon_0$, since $\mathrm{dist}(x_\varepsilon, \partial A) \ge \delta_0$, this implies that the ball $\overline{B}_{\sqrt{M\varepsilon}}(x)$ is contained in A for ε small enough, hence the supremum is attained. The argument for y_ε is similar.

Let us fix ε small enough so that (b) holds and both x'_ε and y'_ε belong to A, and let us apply Jensen's maximum principle to the locally semiconvex[7] function

$$w(x, y) := u^\varepsilon(x) - u_\varepsilon(y) - \frac{1}{4\varepsilon}|x - y|^4$$

to find sequences $z_l := (x_{\varepsilon,l}, y_{\varepsilon,l}) \to (x_\varepsilon, y_\varepsilon)$ and $\delta_l \downarrow 0$ such that w is pointwise second-order differentiable at z_l, $\nabla w(z_l) \to 0$ and $\nabla^2 w(z_l) \le \delta_l I$. By statement (b) and Remark 5.22, for l large enough we have

$$\inf_{|x - x_{\varepsilon,l}|^2 \le M\varepsilon} -F_\gamma(x, \nabla^2 u^\varepsilon(x_{\varepsilon,l})) \le 0, \qquad \sup_{|y - y_{\varepsilon,l}|^2 \le M\varepsilon} -F(y, \nabla^2 u_\varepsilon(y_{\varepsilon,l})) \ge 0. \tag{5.18}$$

On the other hand, the upper bound on $\nabla^2 w(z_l)$ together with (5.10) gives

$$\begin{cases} \nabla^2 u^\varepsilon(x_{\varepsilon,l}) - 2\varepsilon^{-1}(x_{\varepsilon,l} - y_{\varepsilon,l}) \otimes (x_{\varepsilon,l} - y_{\varepsilon,l}) - \varepsilon^{-1}|x_{\varepsilon,l} - y_{\varepsilon,l}|^2 I \le \delta_l I \\ -\nabla^2 u_\varepsilon(y_{\varepsilon,l}) - 2\varepsilon^{-1}(x_{\varepsilon,l} - y_{\varepsilon,l}) \otimes (x_{\varepsilon,l} - y_{\varepsilon,l}) - \varepsilon^{-1}|x_{\varepsilon,l} - y_{\varepsilon,l}|^2 I \le \delta_l I. \end{cases} \tag{5.19}$$

By (5.19) we obtain that $\nabla^2 u^\varepsilon(x_{\varepsilon,l})$ are uniformly bounded from above, and they are also uniformly bounded from below because u^ε is locally semiconvex. Since similar remarks apply to $\nabla^2 u_\varepsilon(y_{\varepsilon,l})$, we can assume with no loss of generality that $\nabla^2 u^\varepsilon(x_{\varepsilon,l}) \to X_\varepsilon$ and $\nabla^2 u_\varepsilon(y_{\varepsilon,l}) \to Y_\varepsilon$ as one lets $l \to \infty$. If we now differentiate w along a direction (ξ, ξ) with $\xi \in \mathbb{R}^n$, we may use the fact that along these directions the fourth-order term is constant to get

$$\langle \nabla^2 u^\varepsilon(x_{\varepsilon,l})\xi, \xi\rangle - \langle \nabla^2 u_\varepsilon(y_{\varepsilon,l})\xi, \xi\rangle \le 2\delta_l|\xi|^2.$$

Taking limits, this proves that $X_\varepsilon \le Y_\varepsilon$. On the other hand, from (5.18) we get

$$-\sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, X_\varepsilon) \le 0 \quad \text{and} \quad -\inf_{y \in \overline{B}_{\sqrt{M\varepsilon}}(y_\varepsilon)} F(y, Y_\varepsilon) \ge 0.$$

[7]The local semiconvexity of w follows from Proposition 5.21.

Now, the inequality $F \geq F_\gamma + \lambda\gamma$ yields

$$\sup_{x\in\overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F(x, Y_\varepsilon) \geq \sup_{x\in\overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, Y_\varepsilon) + \lambda\gamma \geq \sup_{x\in\overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, X_\varepsilon) + \lambda\gamma \geq \lambda\gamma.$$

Hence

$$\sup_{x\in\overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F(x, Y_\varepsilon) - \inf_{y\in\overline{B}_{\sqrt{M\varepsilon}}(y_\varepsilon)} F(y, Y_\varepsilon) \geq \lambda\gamma.$$

Since γ and λ are fixed positive constants independent of ε, and since $|x_\varepsilon - y_\varepsilon| \to 0$, this contradicts the uniform continuity of $F(\cdot, S)$ for a sufficiently small ε. $\square$

5.6 The ABP maximum principle

Let us start this section, devoted to the proof of the ABP maximum principle (see the statement of Theorem 5.40), with some preparatory results and introducing some useful notation.

Consider a paraboloid with opening $\Theta \in \mathbb{R}$, namely the graph of a quadratic polynomial of the form

$$P(x) = c + \langle p, x\rangle + \frac{\Theta}{2}|x|^2,$$

for some $c \in \mathbb{R}, p \in \mathbb{R}^n$ and $\Theta \in \mathbb{R}$.

Definition 5.27 (Tangent paraboloids)**.** *Given a function $u : \Omega \to \mathbb{R}$ and a subset $A \subset \Omega \subset \mathbb{R}^n$, we define*

$$\overline{\theta}(x_0, A, u) := \inf\left\{\Theta : \exists\, P \text{ with opening } \Theta,\ u(x_0) = P(x_0) \text{ and } u \leq P \text{ on } A\right\}.$$

Analogously we set

$$\underline{\theta}(x_0, A, u) := \sup\left\{\Theta : \exists\, P \text{ with opening } \Theta,\ u(x_0) = P(x_0) \text{ and } u \geq P \text{ on } A\right\},$$

so that $\underline{\theta}(x_0, A, u) = -\overline{\theta}(x_0, A, -u)$. Finally, denoting by $\pm$ the positive and negative parts, we set

$$\theta(x_0, A, u) := \max\left\{\underline{\theta}^-(x_0, A, u), \overline{\theta}^+(x_0, A, u)\right\} \geq 0.$$

Given $\Omega \subset \mathbb{R}^n$ open, $x_0 \in \Omega$, a function $u : \Omega \to \mathbb{R}$ and $h > 0$, let us consider the symmetric (second) difference quotient in the direction $\xi \in \mathbb{R}^n$ that is defined by setting

$$\Delta^2_{h,\xi} u(x_0) := \Delta_{h,\xi}(\Delta_{-h,\xi} u)(x_0) = \Delta_{-h,\xi}(\Delta_{h,\xi} u)(x_0)$$

namely

$$\Delta^2_{h,\xi} u(x_0) = \frac{u(x_0 + h\xi) + u(x_0 - h\xi) - 2u(x_0)}{h^2} \sim \frac{\partial^2 u}{\partial \xi^2}(x_0).$$

This quantity is well-defined if $h|\xi| < \operatorname{dist}(x_0, \partial\Omega)$ and is identically equal to Θ on paraboloids with opening Θ. Notice that the symmetric difference quotient satisfies, by applying twice the discrete integration by parts formula for $\Delta_{h,\xi}$ (see Remark 2.11) and exploiting the fact that $\Delta_{h,\xi} \cdot \Delta_{-h,\xi} = \Delta_{-h,\xi} \cdot \Delta_{h,\xi} u$, the identity

$$\int_\Omega u \Delta^2_{h,\xi} \phi \, dx = \int_\Omega \phi \Delta^2_{h,\xi} u \, dx, \tag{5.20}$$

whenever $u \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R})$, $\phi \in L^\infty(\Omega; \mathbb{R})$ has compact support, $|\xi| = 1$ and the h-neighborhood of $\operatorname{supp}(\phi)$ is contained in Ω.

Remark 5.28 (Maximum principle for Δ^2_ξ)**.** If a paraboloid P with opening Θ "touches" u from above (i.e. $P(x_0) = u(x_0)$ and $P(x) \geq u(x)$ in some ball $B_r(x_0) \subset \Omega$), then

$$\Delta^2_{h,\xi} u(x_0) \leq \Delta^2_{h,\xi} P(x_0) = \Theta \qquad \text{whenever } \ |\xi| = 1 \text{ and } |h| \leq r,$$

and a similar property holds for paraboloids touching from below. Thus, taking the infimum when approximating from above and the supremum when approximating from below, we deduce the inequalities

$$\underline{\theta}(x_0, B_r(x_0), u) \leq \Delta^2_{h,\xi} u(x_0) \leq \overline{\theta}(x_0, B_r(x_0), u) \ \text{ whenever } \ |\xi| = 1 \text{ and } |h| \leq r, \tag{5.21}$$

and hence

$$|\Delta^2_{h,\xi} u(x_0)| \leq \theta(x_0, B_r(x_0), u) \qquad \text{whenever } \ |\xi| = 1 \text{ and } |h| \leq r. \tag{5.22}$$

Proposition 5.29. *If $u : \Omega \to \mathbb{R}$ satisfies*

$$\theta_\varepsilon := \theta(\,\cdot\,, B_\varepsilon(\cdot) \cap \Omega, u) \in L^p(\Omega; \mathbb{R})$$

for some $\varepsilon > 0$ and $1 < p \leq \infty$, then u belongs to $W^{2,p}(\Omega; \mathbb{R})$ and, more precisely,

$$\|\nabla^2 u\|_{L^p} \leq c \|\theta_\varepsilon\|_{L^p}, \tag{5.23}$$

for a geometric constant $c = c(n)$.

Proof. Let us observe first that, in order to gain (5.23), it suffices to prove

$$\|\nabla^2_{\xi\xi} u\|_{L^p} \leq \|\theta_\varepsilon\|_{L^p} \qquad \forall \xi \in \mathbb{R}^n, \ |\xi| = 1.$$

Indeed, by polarization (cf. equation (D.3)), it is possible to obtain from the above inequality that

$$\|\nabla^2_{\xi\eta} u\|_{L^p} \leq \frac{1}{2} \left(|\xi|^2 + |\eta|^2 \right) \|\theta_\varepsilon\|_{L^p} \qquad \forall \xi, \eta \in \mathbb{R}^n.$$

Thus, replacing ξ by $s\xi$ and η by η/s and minimizing with respect to the parameter $s \in \mathbb{R}$, one obtains the general estimate on mixed second derivatives:

$$\|\nabla^2_{\xi\eta} u\|_{L^p} \leq |\xi||\eta|\|\theta_\varepsilon\|_{L^p} \qquad \forall \xi, \eta \in \mathbb{R}^n.$$

For any $\varphi \in C_c^\infty(\Omega; \mathbb{R})$ one has

$$\begin{aligned} &\left| \int_\Omega u(x) \frac{\partial^2 \varphi}{\partial \xi^2}(x)\, dx \right| = \left| \lim_{h \to 0} \int_\Omega u(x) \Delta^2_{h,\xi} \varphi(x)\, dx \right| \\ = &\left| \lim_{h \to 0} \int_\Omega (\Delta^2_{h,\xi} u(x)) \varphi(x)\, dx \right| \leq \|\theta_\varepsilon\|_{L^p} \|\varphi\|_{L^{p'}} \end{aligned}$$

where we go from the first to the second line with (5.20) and the inequality follows from our assumption (5.22). Thanks to the Riesz representation theorem (see e.g. Theorem 6.16 in [84]), we know that the map $\varphi \mapsto \int_\Omega u(x) \frac{\partial^2 \varphi}{\partial \xi^2}(x)\, dx$ can be represented by integration against an element of $L^p(\Omega; \mathbb{R})$, which then represents (by definition) the second derivative $\nabla^2_{\xi\xi} u$ in the sense of distributions and which satisfies (5.23).

□

Remark 5.30. Here and below we denote by $\|N\|$ the operator norm of a matrix N, that is defined by

$$\|N\| := \sup_{x \in \mathbb{R}^n,\, |x|=1} |Nx|.$$

When N is assumed to be symmetric, as it is always the case in this section, it is easily checked that the value of $\|N\|$ can be computed as

$$\sup_{x \in \mathbb{R}^n,\, |x|=1} \langle Nx, x \rangle$$

hence also as

$$\max_{i=1,\dots,n} |\nu_i|$$

where $\nu_1, \dots, \nu_n$ are the eigenvalues of N. In other words, the norm $\|\cdot\|$ is the largest modulus of the eigenvalues in the spectrum $\sigma(N)$. From (5.22) we get

$$\|\nabla^2 u(x_0)\| \leq \theta(x_0, B_\varepsilon(x_0), u) \qquad \text{for all } \varepsilon > 0 \text{ with } B_\varepsilon(x_0) \subset \Omega, \tag{5.24}$$

at any point x_0 where u has a second-order Taylor expansion.

Corollary 5.31. *If $\Omega \subset \mathbb{R}^n$ is convex and $\theta_\varepsilon = \theta(\,\cdot\,, B_\varepsilon(\cdot) \cap \Omega, u) \in L^\infty(\Omega; \mathbb{R})$ for some $\varepsilon > 0$, then*

$$\mathrm{Lip}(\nabla u, \Omega) \leq \|\theta_\varepsilon\|_{L^\infty}.$$

Proof. The previous proposition shows that $u \in W^{2,\infty}(\Omega;\mathbb{R})$ and, on the other hand, equation (5.24) provides a pointwise control on $\nabla^2 u$ (recall that semiconvex/semiconcave functions have a second-order Taylor expansion $\mathscr{L}^n$-a.e. for one can reduce to applying Theorem D.15). We further recall that whenever Ω is convex and $v : \Omega \to \mathbb{R}$ is a scalar function we have $\|\nabla v\|_{L^\infty} = \mathrm{Lip}(v,\Omega)$ (while, in general, one only has the one-sided estimate $\|\nabla v\|_{L^\infty} \le \mathrm{Lip}(v,\Omega)$). If instead v takes values in $\mathbb{R}^n$ (as it is in our case, where we aim to take $v := \nabla u : \Omega \to \mathbb{R}^n$), then, by the same smoothing argument used in the scalar case, we can always show that

$$\|\|\nabla v\|\|_{L^\infty} = \mathrm{Lip}(v,\Omega) \tag{5.25}$$

because, when v is continuously differentiable, there holds

$$\Big|v(x) - v(y)\Big| = \left|\int_0^1 \nabla v((1-t)x + ty)(x-y)\,dt\right| \le |x-y| \int_0^1 \|\nabla v\|((1-t)x+ty)\,dt.$$

Therefore from (5.24) and (5.25) we conclude. □

At this point our aim is the study of a nonlinear elliptic equation of the form

$$-F(\nabla^2 u(x)) + f(x) = 0 \tag{5.26}$$

with F non-decreasing on $\mathrm{Sym}^{n\times n}$. A fundamental (though linear) example is given by the case when F is the trace operator on symmetric matrices, which corresponds to the Laplace equation $-\Delta u = f$. Another example of great interest is given by the genuinely nonlinear Monge-Ampère equation $-\det(\nabla^2 u) + f = 0$ (see [12], [13] as well as [31] and [38]).

Definition 5.32 (Ellipticity)**.** *We shall say that the problem* (5.26) *is elliptic with constants* $\Lambda \ge \lambda > 0$ *if for all matrices* $M \in \mathrm{Sym}^{n\times n}$

$$\lambda\|N\| \le F(M+N) - F(M) \le \Lambda\|N\| \qquad \forall\, N \in \mathrm{Sym}^{n\times n},\ N \ge 0. \tag{5.27}$$

Remark 5.33. Every symmetric matrix N admits a unique decomposition as a sum

$$N = N^+ - N^- \ ,$$

with $N^+, N^- \ge 0$ and $N^+N^- = 0$. It can be obtained simply diagonalizing $N = \sum_{i=1}^n \rho_i e_i \otimes e_i$ and then choosing $N^+ := \sum_{\rho_i>0} \rho_i e_i \otimes e_i$ and $N^- = -\sum_{\rho_i\le 0} \rho_i e_i \otimes e_i$. Observing this, we are able to rewrite the ellipticity condition replacing (5.27) with the equivalent requirement that for every $M \in \mathrm{Sym}^{n\times n}$

$$F(M+N) - F(M) \le \Lambda\|N^+\| - \lambda\|N^-\| \qquad \forall\, N \in \mathrm{Sym}^{n\times n}. \tag{5.28}$$

Indeed, it suffices to write

$$F(M+N)-F(M)=\big(F(M-N^-+N^+)-F(M-N^-)\big)+\big(F(M-N^-)-F(M)\big)$$

and to apply to the first term the estimate from above and to the second one the estimate from below.

Example 5.34. Consider the special case

$$F(M)=\mathrm{tr}(BM)$$

where $B=(b_{ij})_{i,j=1,\dots,n}$ belongs to the set

$$\mathcal{A}_{\lambda,\Lambda}:=\left\{B\in\mathrm{Sym}^{n\times n}:\ \lambda I\le B\le\Lambda I\right\}.$$

To verify (5.27), let the symmetric matrix $N\ge 0$ be assigned and choose the coordinate system in which $N=\mathrm{diag}(\rho_1,\dots,\rho_n)$, thus (since $b_{ii}\ge\lambda$ and $\rho_i\ge 0$ for all $i=1,\dots,n$)

$$F(M+N)-F(M)=\mathrm{tr}(BN)=\sum_{i=1}^n b_{ii}\rho_i\ge\lambda\sum_{i=1}^n\rho_i\ge\lambda\rho_{\max}=\lambda\|N\|.$$

Analogously, since $b_{ii}\le\Lambda$ one has

$$F(M+N)-F(M)=\mathrm{tr}(BN)=\sum_{i=1}^n b_{ii}\rho_i\le\Lambda\sum_{i=1}^n\rho_i\le n\Lambda\rho_{\max}=n\Lambda\|N\|.$$

After this introductory part about definitions and notation, we enter in the core of the matter of the Hölder regularity for viscosity solutions: as in De Giorgi's work on Hilbert's *XIX* problem, the regularity results we shall present will be directly obtained from inequalities derived from ellipticity, without specific reference to the original equation.

Definition 5.35 (Pucci's extremal operators)**.** *Given ellipticity constants $\Lambda\ge\lambda>0$ and a symmetric matrix M with spectrum $\sigma(M)$, Pucci's extremal operators are defined by setting $\mathcal{M}^{\pm}_{\lambda,\Lambda}(0)=0$ and*

$$\begin{aligned}\mathcal{M}^-_{\lambda,\Lambda}(M)&:=\lambda\sum_{\rho\in\sigma(M)\cap(0,\infty)}\rho+\Lambda\sum_{\rho\in\sigma(M)\cap(-\infty,0)}\rho,\\ \mathcal{M}^+_{\lambda,\Lambda}(M)&:=\Lambda\sum_{\rho\in\sigma(M)\cap(0,\infty)}\rho+\lambda\sum_{\rho\in\sigma(M)\cap(-\infty,0)}\rho.\end{aligned}$$

We will omit the dependence on λ and Λ, when clear from the context.

Remark 5.36. Resuming the discussion of Example 5.34, it is easy to show that

$$\mathcal{M}^-_{\lambda,\Lambda}(M) = \inf_{B\in\mathcal{A}_{\lambda,\Lambda}} \operatorname{tr}(BM) \tag{5.29}$$

$$\mathcal{M}^+_{\lambda,\Lambda}(M) = \sup_{B\in\mathcal{A}_{\lambda,\Lambda}} \operatorname{tr}(BM). \tag{5.30}$$

Indeed, denoting with (b_{ij}) the coefficients of the matrix $B \in \mathcal{A}_{\lambda,\Lambda}$ in the system of coordinates where M is diagonal, with $M = \operatorname{diag}(\rho_1, \dots, \rho_n)$ we get

$$\lambda \sum_{\rho_i>0} \rho_i + \Lambda \sum_{\rho_i<0} \rho_i \le \operatorname{tr}(BM) = \sum_{i=1}^n b_{ii}\rho_i \le \Lambda \sum_{\rho_i>0} \rho_i + \lambda \sum_{\rho_i<0} \rho_i \tag{5.31}$$

and the equality in (5.31) holds, for instance, if

$$B = \sum_{\rho_i>0} \lambda e_i \otimes e_i + \sum_{\rho_i<0} \Lambda e_i \otimes e_i,$$

as far as the first inequality is concerned, and if

$$B = \sum_{\rho_i>0} \Lambda e_i \otimes e_i + \sum_{\rho_i<0} \lambda e_i \otimes e_i,$$

for the second one. In both cases, we have denoted by $\{e_1, \dots, e_n\}$ the aforementioned diagonalizing basis for M.

Remark 5.37. Pucci's extremal operators satisfy the following properties:

(a) trivially $\mathcal{M}^- \le \mathcal{M}^+$ and $\mathcal{M}^-(-M) = -\mathcal{M}^+(M)$ for every symmetric matrix M, moreover $\mathcal{M}^\pm$ are positively 1-homogeneous;

(b) $\mathcal{M}^\pm$ are elliptic (i.e. they satisfy (5.27)) with constants $\lambda, n\Lambda$, because of Example 5.34 and equations (5.29), (5.30) which represent $\mathcal{M}^\pm$ as an envelope of a family of functionals with ellipticity constants $\lambda, n\Lambda$;

(c) for every couple of symmetric matrices M, N it is simple to obtain from (5.29) and (5.30) that

$$\mathcal{M}^+(M) + \mathcal{M}^-(N) \le \mathcal{M}^+(M+N) \le \mathcal{M}^+(M) + \mathcal{M}^+(N)$$

and, similarly,

$$\mathcal{M}^-(M) + \mathcal{M}^-(N) \le \mathcal{M}^-(M+N) \le \mathcal{M}^-(M) + \mathcal{M}^+(N);$$

(d) thanks to (5.28), one has for any symmetric matrix M the inequalities

$$\mathcal{M}^-_{\lambda/n,\Lambda}(M) \le F(M) \le \mathcal{M}^+_{\lambda/n,\Lambda}(M) \tag{5.32}$$

whenever F is elliptic with constants λ, Λ and $F(0) = 0$.

Definition 5.38. *In the setting above, we shall set*

$$\begin{aligned} \mathrm{Sub}_{\lambda,\Lambda}(f) &:= \left\{u : \Omega \to \mathbb{R} : \, -\mathcal{M}^+_{\lambda,\Lambda}(\nabla^2 u) + f \le 0 \ \textit{in}\ \Omega\right\} \\ \mathrm{Sup}_{\lambda,\Lambda}(f) &:= \left\{u : \Omega \to \mathbb{R} : \, -\mathcal{M}^-_{\lambda,\Lambda}(\nabla^2 u) + f \ge 0 \ \textit{in}\ \Omega\right\}. \end{aligned}$$

Furthermore, we define

$$\mathrm{Sol}_{\lambda,\Lambda}(f) := \mathrm{Sub}_{\lambda/n,\Lambda}(-|f|) \cap \mathrm{Sup}_{\lambda/n,\Lambda}(|f|). \tag{5.33}$$

Remark 5.39. Roughly speaking, the classes defined above correspond to De Giorgi's classes $DG_\pm(\Omega)$, since u being a solution to (5.26) (with F having ellipticity constants λ and Λ and $F(0) = 0$) implies $u \in \mathrm{Sol}_{\lambda,\Lambda}(f)$ by virtue of Remark 5.37(d); thus, if we are able to infer regularity of functions in $\mathrm{Sol}_{\lambda,\Lambda}(f)$ then we can "forget" the specific form of the equation.

From now onward, we shall not explicitly indicate the dependence of the maximal operators $\mathcal{M}^+, \mathcal{M}^-$ on the ellipticity coefficients λ and Λ. Notice that, since $\mathcal{M}^+ \ge \mathcal{M}^-$, the intersection of the two sets above can be nonempty.

The key estimate we want to prove is named after Aleksandrov, Bakelman and Pucci in [14], [10] and [16] (building on results in [81] and [4]) and is therefore called *ABP weak maximum principle*. It plays in this regularity theory essentially the same role played by the Caccioppoli-Leray inequality in the standard linear elliptic theory. See also [94] for another interesting insight into the problem.

In the sequel of this chapter we call *universal* a constant which depends only on the space dimension n and on the ellipticity constants λ, Λ.

Theorem 5.40 (Aleksandrov-Bakelman-Pucci weak maximum principle, [14, 10])**.** *Let u be in $\mathrm{Sup}(f) \cap C(\overline{B}_r;\mathbb{R})$ with $u \ge 0$ on ∂B_r and $f \in C(\overline{B}_r;\mathbb{R})$. Then*

$$\max_{\overline{B}_r} u^- \le c_{ABP} r \left(\int_{\{u=\Gamma_u\}} \left(f^+\right)^n \, dx \right)^{1/n}$$

where $c_{ABP} = c_{ABP}(\lambda, \Lambda, n)$ is universal and Γ_u is defined below.

In particular, u supersolution with $f = 0$ provides the minimum principle: $u \ge 0$ on ∂B_r implies $u \ge 0$ on B_r. Since f^+ measures, in some sense, how far u is from being concave, the estimate above can be seen as a quantitative formulation of the fact that a concave function in a ball attains its minimum on the boundary of the ball.

Definition 5.41. *Assume the function u^- is extended to all $\overline{B}_{2r} \setminus \overline{B}_r$ as the null function (this extension is continuous, since u^- is null on ∂B_r). We then define for $x \in \overline{B}_{2r}$*

$$\Gamma_u(x) := \sup \left\{ L(x) : \ L \text{ affine}, \ L \leq -u^- \text{ on } \overline{B}_{2r} \right\}.$$

We will call contact region *the set $\{u = \Gamma_u\} \subset \overline{B}_r$. Moreover notice that*

$$\{u = \Gamma_u\} = \{-u^- = \Gamma_u\} \cap \overline{B}_r \subset \overline{B}_{2r}.$$

In order to prove the ABP estimate we set $m := \max_{\overline{B}_r} u^-$ and assume with no loss of generality that $m > 0$.

The following facts are either trivial consequences of the definitions or easy applications of the tools introduced in Appendix D, devoted to convex analysis: first of all $-m \leq \Gamma_u \leq 0$, hence $\Gamma_u \in W^{1,\infty}_{\text{loc}}(B_{2r}; \mathbb{R})$ and finally, since Γ_u is differentiable almost everywhere by Rademacher's theorem and the graph of the subdifferential is closed, we get $\partial \Gamma_u(x) \neq \emptyset$ for all $x \in B_{2r}$. We will use this last property to provide a supporting hyperplane to Γ_u at any point in the closed ball $\overline{B}_r$.

We need some preliminary results, here is the first one.

Theorem 5.42. *Assume $u \in C(\overline{B}_r; \mathbb{R})$, $u \geq 0$ on ∂B_r and $\Gamma_u \in C^{1,1}(B_r; \mathbb{R})$. Then*

$$\max_{\overline{B}_r} u^- \leq cr \left(\int_{B_r} \det \nabla^2 \Gamma_u \, dx \right)^{1/n}$$

with $c = c(n)$.

Proof. Let $x_1 \in \overline{B}_r$ be such that $u^-(x_1) = m$ (where, as we set above, $m = \max_{\overline{B}_r} u^-$). Fix ξ with $|\xi| < m/(3r)$ and denote by L_α the affine function $L_\alpha(x) = -\alpha + \langle x, \xi \rangle$. It is obvious that, if α is large enough, then the corresponding hyperplane lies below the graph of $-u^-$ and by continuity of u there is a minimum value of α such that this happens, that is to say $-u^- \geq L_\alpha$ on $\overline{B}_{2r}$. The graph of $-u^-$ will then meet the hyperplane L_α at some point, say $x_0 \in \overline{B}_{2r}$. If it were $|x_0| > r$, then $L_\alpha(x_0) = 0$, but on the other hand $|L_\alpha(x_1)| \geq m$ and, since $|x_0 - x_1| \leq 3r$, L_α would have slope $|\xi| \geq m/3r$, which is a contradiction. Hence any contact point x_1 must lie inside the ball B_r; from the pointwise inequality $-u^- \geq \Gamma_u \geq L_\alpha$ we get $\nabla \Gamma_u(x_1) = \xi$ and therefore $B_{m/(3r)} \subset \nabla \Gamma_u(B_r)$. If we measure the corresponding volumes and use the area formula, we get

$$\omega_n \left(\frac{m}{3r} \right)^n \leq \int_{B_r} \det \nabla^2 \Gamma_u \, dx$$

or, equivalently,

$$m \leq 3\omega_n^{-1/n} r \left(\int_{B_r} \det \nabla^2 \Gamma_u \, dx \right)^{1/n}.$$

This proves the claim with $c = 3\omega_n^{-1/n}$. $\square$

Remark 5.43. The previous theorem implies the ABP estimate, provided we show that

- $\Gamma_u \in C^{1,1}(B_r;\mathbb{R})$, as a consequence of $u \in \mathrm{Sup}(f)$. Even though $\nabla^2\Gamma_u$ exists $\mathscr{L}^n$-a.e. thanks to the Aleksandrov Theorem, its existence is not sufficient to provide the validity of the area formula for $\nabla\Gamma_u$ (for instance in one space dimension a counterexample is provided by an antiderivative of the Cantor-Vitali function);
- $\mathscr{L}^n$-a.e. on $\{u > \Gamma_u\}$ (the so-called *non-contact region*) one has $\det(\nabla^2\Gamma_u) = 0$;
- $\mathscr{L}^n$-a.e. on $\{u = \Gamma_u\}$ one has $\det(\nabla^2\Gamma_u) \le (cf^+)^n$, with c universal constant.

Let us now discuss each of the items above. The next theorem shows that regularity, as measured in terms of opening of paraboloids touching Γ_u from above, propagates from the contact set to the non-contact set. It turns out that the regularity in the contact set is a direct consequence of the supersolution property.

Theorem 5.44 (Propagation of regularity)**.** *Let $u \in C(\overline{B}_r;\mathbb{R}), \Gamma_u$ as in Definition 5.41 and suppose there exist $\varepsilon \in (0,r]$ and $M \ge 0$ such that, for all $x_0 \in \{u = \Gamma_u\}$, there exists a paraboloid with opening less than M which has a contact point from above with the graph of Γ_u in $B_\varepsilon(x_0)$. Then $\Gamma_u \in C^{1,1}(B_r;\mathbb{R})$ and $\det \nabla^2\Gamma_u = 0$ $\mathscr{L}^n$-a.e. on $\{u > \Gamma_u\}$.*

With the notation introduced before, the assumption of Theorem 5.44 means

$$\overline{\theta}(x_0, B_\varepsilon(x_0), \Gamma_u) \le M \qquad \forall x_0 \in \overline{B}_r \cap \{u = \Gamma_u\} \ .$$

Since Γ_u is convex, the corresponding quantity $\underline{\theta}^-$ is null, so that $\theta = \overline{\theta}^+ \le M$. Recall also that we have already proved that $\overline{\theta}, \underline{\theta} \in L^\infty$ implies $u \in C^{1,1}$ in Corollary 5.31.

Theorem 5.45 (Regularity at contact points)**.** *Consider $v \in \mathrm{Sup}(f)$ in B_δ, φ convex in B_δ with $0 \le \varphi \le v$ and $v(0) = \varphi(0) = 0$. Then $\varphi(x) \le c\left(\sup_{B_\delta} f^+\right)|x|^2$ in $B_{\nu\delta}$, where ν and c are universal constants.*

We can get a naïve interpretation of this result (or, rather, of its infinitesimal version as $\delta \downarrow 0$) by means of this formal argument: by virtue of the assumption $v \in \mathrm{Sup}(f)$, we have the inequality $\mathcal{M}^-(\nabla^2 v(0)) \le f(0)$, while the fact that $v - \varphi$ has a local minimum at 0 gives $\mathcal{M}^-(\nabla^2\varphi(0)) \le \mathcal{M}^-(\nabla^2 v(0)) \le f(0)$.

That being said, let us discuss how these tools allow to prove the ABP estimate.

Proof of Theorem 5.40. Pick a point $x_0 \in \{u = \Gamma_u\}$ and let L be a supporting hyperplane for Γ_u at x_0, so that $\Gamma_u \ge L$ and $\Gamma_u(x_0) = L(x_0)$. Recalling Theorem 5.45, define $\varphi := \Gamma_u - L$, $v := -u^- - L$ (and notice that v is a supersolution in B_{2r} with datum $f^+\chi_{\overline{B}_r}$ namely $v \in \mathrm{Sup}(f^+\chi_{\overline{B}_r})$). Now, $\varphi(x_0) = v(x_0)$ implies, by means of Theorem 5.45,

$$\overline{\theta}(x_0, B_{\nu\delta}(x_0), \varphi) \le c \sup_{B_\delta(x_0)\cap\overline{B}_r} f^+ \qquad \forall x_0 \in \{u = \Gamma_u\} \tag{5.34}$$

with ν and c universal, for all $\delta \in (0, r)$. Hence

$$\overline{\theta}(x_0, B_{\nu\delta}(x_0), \Gamma_u) \leq c \sup_{B_\delta(x_0)\cap\overline{B}_r} f^+ \qquad \forall x_0 \in \{u = \Gamma_u\}. \tag{5.35}$$

By Theorem 5.44 we get $\Gamma_u \in C^{1,1}(B_r; \mathbb{R})$ and $\det(\nabla^2\Gamma_u) = 0$ $\mathscr{L}^n$-almost everywhere in the non-contact region. Finally, in order to get the desired estimate, we have to show that $\mathscr{L}^n$-almost everywhere in the contact region one has $\det(\nabla^2\Gamma_u) \leq (cf^+)^n$. But this comes at once by passing to the limit as $\delta \to 0$ in (5.35) at any differentiability point x_0 of Γ_u. In fact, by virtue of the remarks presented before the statement of Theorem 5.45, hence ultimately relying on Corollary 5.31, all the eigenvalues of $\nabla^2\Gamma_u(x_0)$ do not exceed $cf^+(x_0)$ and the conclusion follows. □

Now we prove Theorem 5.45.

Proof. Let $r \in (0, \delta/4)$ and set $\overline{s} := \left(\sup_{\overline{B}_r} \varphi\right)/r^2$. Let then $\bar{x} \in \partial B_r$ be a maximum point of φ on $\overline{B}_r$ (by convexity the maximum is attained at the boundary). Possibly acting with a rotation, we can write $x = (x', x_n)$, $x' \in \mathbb{R}^{n-1}$, $x_n \in \mathbb{R}$, and assume $\bar{x} = (0, r)$. Consider the intersection A of the closed strip defined by the hyperplanes $x_n = r$ and $x_n = -r$ with the ball $\overline{B}_{\delta/2}$. We clearly have that $\partial A = A_1 \cup A_2 \cup A_3$, where $A_1 = \overline{B}_{\delta/2} \cap \{x_n = r\}$, $A_2 = \overline{B}_{\delta/2} \cap \{x_n = -r\}$ and $A_3 = \partial B_{\delta/2} \cap \{|x_n| < r\}$.

We claim that $\varphi \geq \varphi(\bar{x})$ on A_1. To this aim, we first prove that $\varphi(y) \leq \varphi(\bar{x}) + o(|y - \bar{x}|)$ for $y \to \bar{x}$, $y \in H := \{x_n = r\}$. In fact, this comes from $\varphi(ry/|y|) \leq \varphi(\bar{x})$ and $\varphi(y) - \varphi(ry/|y|) = o(|y - \bar{x}|)$, because φ is Lipschitz continuous. On the other hand, we have that $\xi \in \partial\varphi_{|H}(\bar{x})$ implies $\varphi(y) \geq \varphi(\bar{x}) + \langle \xi, y - \bar{x}\rangle$ for all $y \in H$. Hence, by comparison, it must be $\xi = 0$ and so $\varphi(y) \geq \varphi(\bar{x})$ on A_1 (this can be seen as a nonsmooth version of the Lagrange multipliers theorem).

As a second step, set

$$p(x) := \frac{\overline{s}}{8}(x_n + r)^2 - 4\frac{\overline{s}}{\delta^2} r^2 |x'|^2$$

and notice that the following properties hold:

(a) on A_1, $p(x) \leq \overline{s}r^2/2 = \varphi(\bar{x})/2 \leq \varphi(x)/2$;

(b) on A_2, $p(x) \leq 0 \leq \varphi(x)$ (and in particular $p(x) \leq v(x)$);

(c) on A_3, $\delta^2/4 = |x'|^2 + x_n^2 \leq |x'|^2 + r^2 \leq |x'|^2 + \delta^2/16$, which implies $|x'|^2 \geq (3/16)\delta^2$. By means of such estimate we get $p(x) \leq (\overline{s}/2)r^2 - (3/4)\overline{s}r^2 \leq 0 \leq \varphi$.

Combining (a), (b), (c) above we get $p \leq v$ on ∂A. Since $p(0) = \overline{s}r^2/8 > 0 = \varphi(0)$ we can rigidly move down this paraboloid until we get a limit paraboloid $p' = p - \alpha$ (for some translation parameter $\alpha > 0$) lying below the graph of v and touching it at some point, say y. Since $p \leq v$ on ∂A, the point y is in the interior of A.

By the supersolution property $\mathcal{M}^-(\nabla^2 p) \le f(y) \le \sup_{B_\delta} f$ we get (since we have an explicit expression for p)

$$\lambda \frac{\overline{s}}{4} - 8(n-1)\Lambda \overline{s} \frac{r^2}{\delta^2} \le \sup_{B_\delta} f.$$

But now we can fix r such that $8(n-1)\Lambda r^2/\delta^2 \le \lambda/8$ (it is done by taking r so that the inequality $8r \le \delta\sqrt{\lambda/((n-1)\Lambda)}$ is satisfied): therefore we have $\overline{s} \le \frac{8}{\lambda} \sup_{B_\delta} f$. The statement follows, with $c = 8/\lambda$ and $\nu := \frac{1}{8}\sqrt{\lambda/((n-1)\Lambda)}$. □

It remains to prove Theorem 5.44.

Proof. Recall first that we are assuming the existence of a uniform estimate

$$\overline{\theta}(x, B_\varepsilon(x), \Gamma_u) \le M \qquad \forall x \in \overline{B}_r \cap \{u = \Gamma_u\}.$$

Thanks to Proposition 5.29, we are able to obtain $C^{1,1}$ regularity of Γ_u as soon as we are able to propagate this estimate also to non-contact points.

Consider now any point $x_0 \in \{u > \Gamma_u\}$ and call L a supporting hyperplane for Γ_u at x_0, so that $L(x_0) = \Gamma_u(x_0)$. Notice that $\{-u^- = L\} \cap \overline{B}_r \subset \{u = \Gamma_u\}$. We claim that:

(a) There exist $n+1$ points $x_1, \dots, x_{n+1} \in \{-u^- = L\}$ such that $x_0 \in S := \mathrm{co}(x_1, \dots, x_{n+1})$ (here and in the sequel co stands for convex hull) and, moreover, all such points belong to $\overline{B}_r$ with at most one exception lying on ∂B_{2r}. In addition $\Gamma_u \equiv L$ on S;

(b) $x_0 = \sum_{i=1}^{n+1} t_i x_i$ with at least one index i verifying both $x_i \in \overline{B}_r \cap \{-u^- = L\}$ and $t_i \ge 1/(3n)$.

To show the utility of this claim, just consider how these two facts imply the thesis: on the one hand, if $\nabla\Gamma_u$ is differentiable at x_0, we get $\det \nabla^2\Gamma_u(x_0) = 0$ because $\Gamma_u = L$ on S and $\dim(S) \ge 1$. On the other hand we may assume, without loss of generality, that $x_1 \in \{-u^- = L\} \cap \overline{B}_r$ and $t_1 \ge 1/(3n)$ so that, since

$$x_0 + h = t_1\left(x_1 + \frac{h}{t_1}\right) + t_2 x_2 + \cdots + t_{n+1} x_{n+1},$$

one has

$$\begin{aligned}
\Gamma_u(x_0 + h) &\le t_1 \Gamma_u\left(x_1 + h/t_1\right) + t_2 \Gamma_u(x_2) + \cdots + t_{n+1}\Gamma_u(x_{n+1}) \\
&\le t_1 \left[L(x_1) + M\left|\frac{h}{t_1}\right|^2\right] + t_2 L(x_2) + \cdots + t_{n+1} L(x_{n+1}) \\
&= L(x_0) + M|h|^2/t_1 \le \Gamma_u(x_0) + 3nM|h|^2
\end{aligned}$$

and this estimate is clearly uniform since we only require $|h/t_1| \le \varepsilon$, which is implied by $|h| \le \varepsilon/(3n)$.

Hence, the problem is reduced to prove the two claims above. To proceed, we need a standard result in convex analysis (first proved by Carathéodory [18] for closed sets and later extended by Steinitz [89]), which is recalled here for completeness.

Theorem 5.46 (Carathéodory, [18])**.** *Let V be a n-dimensional real vector space. If $C \subset V$, then for every $x \in \mathrm{co}(C)$ (the convex hull of C) there exist $x_1, \dots, x_{n+1} \in C$, $t_1, \dots, t_{n+1} \in [0,1]$ such that*

$$x = \sum_{i=1}^{n+1} t_i x_i \qquad \text{and} \qquad \sum_{i=1}^{n+1} t_i = 1.$$

Set then $C' := \left\{x \in \overline{B}_{2r} : L(x) = -u^-(x)\right\}$ and $C = \mathrm{co}(C')$. We notice that $C' \neq \emptyset$: indeed if, by contradiction, $C = \emptyset$, then there exists $\delta > 0$ such that $L + \delta \leq -u^-$. Therefore $\Gamma_u \geq L + \delta$, contradicting the assumption that x_0 is a contact point. We claim that $x_0 \in C$: in fact, if this were not the case, there would exist $\eta > 0$ and a hyperplane L' such that $L'(x_0) > 0$ and $L'(y) < 0$ if $y \in C_\eta := \left\{y \in \overline{B}_{2r} \,|\, \mathrm{dist}(y, C) < \eta\right\}$, therefore $L + \delta L' \leq -u^-$ on C_η for all $\delta > 0$. Let us notice that, on $\overline{B}_{2r} \setminus C_\eta \subset \overline{B}_{2r} \setminus C'$, the function $-u^- - L$ is strictly positive and, thanks to the compactness of $\overline{B}_{2r} \setminus C_\eta$, there exists $\delta > 0$ such that

$$L(x) + \delta L'(x) \leq -u^-(x), \qquad \forall\, x \in \overline{B}_{2r} \setminus C_\eta.$$

Hence, we would have $(L + \delta L')(x_0) > L(x_0) = \Gamma_u(x_0)$ and, at the same time,

$$L + \delta L' \leq -u^- \quad \text{on } \overline{B}_{2r},$$

which contradicts the maximality of Γ_u.

Thanks to Carathéodory's theorem, we can write $x_0 = \sum_{i=1}^{n+1} t_i x_i$ with $x_i \in \{-u^- = L\}$. In case there were distinct points x_i, x_j with $|x_i| > r$ and $|x_j| > r$ (and so $L(x_i) = 0$, $L(x_j) = 0$) then (considering a point z on the open segment between x_i and x_j) the function Γ_u would achieve its maximum, equal to 0, in the interior of B_{2r} and so, by the convexity of Γ_u, it would be $\Gamma_u \equiv 0$ on B_{2r}, in contrast with the assumption $m = \max u^- > 0$. The same argument also proves that exceptional points out of $\overline{B}_r$, if any, must lie on ∂B_{2r}.

Let us now prove that $\Gamma_u(x) = L(x)$ on $S := \mathrm{co}(x_1, \dots, x_{n+1})$. The inequality $\geq$ is trivial, the converse one is clear for each $x = x_i$, since $L \leq \Gamma_u \leq -u^-$, and it is obtained by means of the convexity of Γ_u at all points in S.

Now we prove part (b) of the claim. If all points x_j verify $|x_j| \leq r$, then $\max t_i \geq \frac{1}{n+1} > \frac{1}{3n}$. Otherwise, if one point, say x_{n+1}, satisfies $|x_{n+1}| = 2r$, then $t_i < 1/(3n)$ for all $i = 1, \dots, n$ implies $t_{n+1} > 2/3$ and therefore

$$r \geq |x_0| \geq 2t_{n+1} r - \sum_{i=1}^{n} t_i |x_i| > \frac{4}{3} r - \frac{n}{3n} r = r.$$

□

5.7 A regularity result for viscosity solutions

The scope of this section is to prove the Harnack inequality (see [65] and [66]) for functions in the class $\mathrm{Sol}(f) := \mathrm{Sub}(-|f|) \cap \mathrm{Sup}(|f|)$, following a strategy devised in [14] (see also [16]). Here, according to Definition 5.38, the sets $\mathrm{Sup}(|f|)$ and $\mathrm{Sub}(-|f|)$ are defined through Pucci's extremal operators (with fixed ellipticity constants $0 < \lambda \leq \Lambda$)[8], namely by declaring

$$u \in \mathrm{Sub}(-|f|) \quad \Longleftrightarrow \quad -\mathcal{M}^+(\nabla^2 u) - |f| \leq 0 \tag{5.36}$$

$$u \in \mathrm{Sup}(|f|) \quad \Longleftrightarrow \quad -\mathcal{M}^-(\nabla^2 u) + |f| \geq 0 \tag{5.37}$$

in the sense of viscosity solutions.

We shall use the standard notation $Q_r(x)$ for the closed n-cube in $\mathbb{R}^n$ with side length r, $Q_r = Q_r(0)$ and always assume that f is continuous. In the proof of Lemma 5.52 below, however, we shall apply the ABP estimate to a function $w \in \mathrm{Sup}(g)$ with g upper semicontinuous. Since there exists g_n continuous with $g_n \downarrow g$ and $w \in \mathrm{Sup}(g_n)$, the ABP estimate holds, by approximation, even in this case. Here is the statement we want to present:

Theorem 5.47. *Consider a function $u : Q_1 \to \mathbb{R}$ with $u \geq 0$ and $u \in \mathrm{Sol}(f) \cap C(Q_1; \mathbb{R})$. There exists a universal constant c_H such that*

$$\sup_{Q_{1/2}} u \leq c_H \left(\inf_{Q_{1/2}} u + \|f\|_{L^n(Q_1;\mathbb{R})} \right). \tag{5.38}$$

Let us show how (5.38) leads to the Hölder regularity result for viscosity solutions of the fully nonlinear elliptic equation

$$-F(\nabla^2 u(x)) + f(x) = 0. \tag{5.39}$$

Step 1. As usual, we need to control the oscillation (now on cubes), defined by

$$\omega_r := M_r - m_r$$

with $M_r := \sup_{Q_r} u$ and $m_r := \inf_{Q_r} u$.
Using the notation of Theorem 5.47, we claim that there exists a universal constant $\mu \in (0, 1)$ such that

$$\omega_{1/2} \leq \mu\, \omega_1 + 2\|f\|_{L^n(Q_1;\mathbb{R})}. \tag{5.40}$$

Indeed, we apply the Harnack inequality (5.38)

[8] Notice that $\mathrm{Sup}(f) \subset \mathrm{Sup}(|f|)$ and $\mathrm{Sub}(f) \subset \mathrm{Sub}(-|f|)$.

- to the function $u - m_1$, so that

$$M_{1/2} - m_1 \le c_H \left(m_{1/2} - m_1 + \|f\|_{L^n(Q_1;\mathbb{R})} \right), \tag{5.41}$$

- to the function $M_1 - u$, so that

$$M_1 - m_{1/2} \le c_H \left(M_1 - M_{1/2} + \|f\|_{L^n(Q_1;\mathbb{R})} \right), \tag{5.42}$$

and adding (5.41) and (5.42) we get

$$\omega_1 + \omega_{1/2} \le c_H \left(\omega_1 - \omega_{1/2} + 2\|f\|_{L^n(Q_1;\mathbb{R})} \right)$$

which proves (5.40) because

$$\omega_{1/2} \le \frac{c_H - 1}{c_H + 1}\,\omega_1 + 2\frac{c_H}{c_H + 1}\,\|f\|_{L^n(Q_1;\mathbb{R})} \le \frac{c_H - 1}{c_H + 1}\,\omega_1 + 2\,\|f\|_{L^n(Q_1;\mathbb{R})}.$$

Notice that $\mu = (c_H - 1)/(c_H + 1)$, with c_H being the universal constant in (5.38): it is crucial for the decay of the oscillation that $\mu < 1$.

Step 2. Thanks to a rescaling argument (which we will be also used in the proof of the Harnack inequality), we can generalize (5.40) as follows. Fix a radius $0 < r \le 1$ and set

$$u_r(y) := \frac{u(ry)}{r^2}, \qquad f_r(y) = f(ry) \qquad \text{with } y \in Q_1.$$

Notice that (5.40) holds also for u_r (with the corresponding source f_r). Moreover, passing to a smaller scale, the L^n-norm improves in the sense that $\|f_r\|_{L^n(Q_1;\mathbb{R})} = r^{-1}\|f\|_{L^n(Q_r;\mathbb{R})}$. For simplicity we keep the notation ω_r for the oscillation of the function u, we use $\mathrm{osc}(\cdot, Q_r)$ otherwise. Based on the above remarks, we can estimate

$$\begin{aligned} \omega_{r/2} &= r^2 \mathrm{osc}(u_r, Q_{1/2}) \le \mu r^2 \mathrm{osc}(u_r, Q_1) + 2r^2 \|f_r\|_{L^n(Q_1;\mathbb{R})} \\ &= \mu\omega_r + 2r\|f\|_{L^n(Q_r;\mathbb{R})} \le \mu\omega_r + 2r\|f\|_{L^n(Q_1;\mathbb{R})}. \end{aligned}$$

Step 3. By the iteration lemma we used so frequently in the elliptic regularity chapters[9], we are immediately able to conclude that

$$\omega_r \le c\omega_1 r^{\min\{1,\alpha\}} \quad \forall r \in (0,1] \qquad \text{with } \left(\frac{1}{2}\right)^{\alpha} = \mu,$$

and with $c = c(\mu, \|f\|_{L^n(Q_1;\mathbb{R})})$ (i.e. a constant depending only on μ and $\|f\|_{L^n(Q_1;\mathbb{R})}$), thus we have Hölder regularity.

In order to prove the Harnack inequality, we go through the following reformulation of Theorem 5.47.

[9] See, for instance, Lemma 3.13.

Theorem 5.48. *There exist universal positive constants ε_0 and c_{HV} such that, if $u : Q_{4\sqrt{n}} \to [0,\infty)$ belongs to $\mathrm{Sol}(f) \cap C(Q_{4\sqrt{n}};\mathbb{R})$ on $Q_{4\sqrt{n}}$, then*

$$\inf_{Q_{1/4}} u \le 1 \qquad \Longrightarrow \qquad \sup_{Q_{1/4}} u \le c_{HV} \tag{5.43}$$

provided

$$\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \le \varepsilon_0.$$

Remark 5.49. Theorem 5.47 and Theorem 5.48 are easily seen to be equivalent: since we will prove the second one, it is more important for us to check that Theorem 5.47 follows from Theorem 5.48.
For some positive $\delta > 0$ (needed to avoid a potential division by 0) consider the function

$$v := \frac{u}{\delta + \inf_{Q_{1/4}} u + \|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})}/\varepsilon_0}.$$

Denoting by f_v the source term associated with v, namely

$$f_v := \frac{f}{\delta + \inf_{Q_{1/4}} u + \|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})}/\varepsilon_0}$$

the homogeneity of Pucci's operators gives $\|f_v\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \le \varepsilon_0$. Since $\inf_{Q_{1/4}} v \le 1$ we have (by Theorem 5.48) $\sup_{Q_{1/4}} v \le c_{HV}$, hence

$$\sup_{Q_{1/4}} u \le c_{HV} \left(\inf_{Q_{1/4}} u + \delta + \|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})}/\varepsilon_0 \right).$$

We let $\delta \to 0$ and we obtain Harnack inequality with the cubes $Q_{1/4}$, $Q_{4\sqrt{n}}$; hence by the same scaling argument we have already used, this means that for all $r \in (0, 1/4)$

$$\sup_{Q_r} u \le c_{HV} \left(\inf_{Q_r} u + r\|f\|_{L^n(Q_{16r\sqrt{n}};\mathbb{R})}/\varepsilon_0 \right). \tag{5.44}$$

Now, we pass to the cubes $Q_{1/2}$, Q_1 with a simple covering argument: there exists an integer $N = N(n)$ such that for all $x \in Q_{1/2}$, $y \in Q_{1/2}$ we can find points x_i, $1 \le i \le N$, with $x_i = x$, $x_N = y$ and $x_{i+1} \in Q_r(x_i)$ for $1 \le i < N$, with $r = r(n)$ so small that all cubes $Q_{16r\sqrt{n}}(x_i)$ are contained in Q_1. Applying repeatedly (5.44) we get (5.38) with $c_H \sim c_{HV}^N$.

We shall now describe the strategy of the proof of Theorem 5.48, even if the full proof will be completed at the end of this section. We study the map

$$t \mapsto \mathscr{L}^n \left(\{u > t\} \cap Q_1 \right)$$

in order to prove:

- a decay estimate of the form $\mathscr{L}^n(\{u > t\} \cap Q_1) \leq dt^{-\varepsilon}$, thanks to the fact that $u \in \mathrm{Sup}(|f|)$ (see Lemma 5.52);
- the full thesis of Theorem 5.48 using the fact that $u \in \mathrm{Sol}(f) \subset \mathrm{Sub}(-|f|)$, too.

The first goal will be achieved using the Aleksandrov-Bakelman-Pucci inequality of the previous section. The structure of the proof is reminescent of De Giorgi's argument and, like it was said, we complete it through the following lemmata and remarks.

This preliminary lemma is a self-improvement principle, whose proof exploits the Calderón-Zygmund construction. Heuristically, the assumption (5.45) means that B contains all regions (i.e. dyadic cubes) where the measure of A is sufficiently large. This allows to prove the improved implication of (5.46).

Lemma 5.50 (Dyadic Lemma, Krylov-Safonov [66])**.** *Consider two Borel sets $A, B \subset Q_1$. If the implication*

$$\mathscr{L}^n(A \cap Q) > \delta \mathscr{L}^n(Q) \implies \tilde{Q} \subset B, \tag{5.45}$$

holds for any dyadic cube $Q \subset Q_1$, with $\tilde{Q}$ being the predecessor of Q, then

$$\mathscr{L}^n(A) \leq \delta \mathscr{L}^n(Q_1) \implies \mathscr{L}^n(A) \leq \delta \mathscr{L}^n(B). \tag{5.46}$$

Proof. We apply the construction of Calderón-Zygmund (seen in the proof of Theorem B.18, for instance) to $f = \chi_A$ with $\alpha = \delta$: there exists a countable family of cubes $\{Q_i\}_{i \in I}$, pairwise disjoint, such that

$$\chi_A \leq \delta \qquad \mathscr{L}^n\text{-a.e. on } Q_1 \setminus \bigcup_{i \in I} Q_i \tag{5.47}$$

and $\mathscr{L}^n(A \cap Q_i) > \delta \mathscr{L}^n(Q_i)$ for all $i \in I$. Since $\delta < 1$ and χ_A is a characteristic function, (5.47) means that $A \subset \bigcup_{i \in I} Q_i$ up to Lebesgue negligible sets. Moreover, if $\tilde{Q}_i$ are the predecessors of Q_i, from (5.45) we get $\tilde{Q}_i \subset B$ for all i and

$$\mathscr{L}^n(A \cap \tilde{Q}_i) \leq \delta \mathscr{L}^n(\tilde{Q}_i) \qquad \forall\, i \in I. \tag{5.48}$$

This is due to the fact that a cube Q, in the Calderón-Zygmund construction, is divided in subcubes as long as $\mathscr{L}^n(A \cap Q) \leq \delta \mathscr{L}^n(Q)$. Thus (note that we sum on $\tilde{Q}_i$ rather than on i, because different cubes might have the same predecessor)

$$\mathscr{L}^n(A) \leq \sum_{\tilde{Q}_i} \mathscr{L}^n(A \cap \tilde{Q}_i) \leq \sum_{\tilde{Q}_i} \delta \mathscr{L}^n(\tilde{Q}_i) \leq \delta \mathscr{L}^n(B).$$

□

It is bothering, but necessary to go on with the proof, to deal at the same time with balls and cubes: balls emerge from the radial construction in the next lemma and cubes are needed in Calderón-Zygmund-type constructions.

Lemma 5.51 (Truncation Lemma)**.** *There exists a function $\varphi \in C^\infty(\mathbb{R}^n;\mathbb{R})$ such that*

(i) $\varphi \geq 0$ *on* $\mathbb{R}^n \setminus B_{2\sqrt{n}}$ *;*

(ii) $\varphi \leq -2$ *on the cube* Q_3 *;*

(iii) $\mathcal{M}^+(\nabla^2\varphi) \leq c\chi_{Q_1}$ *on* $\mathbb{R}^n$ *;*

for some universal constant $c > 0$*.*

Proof. We recall some useful inclusions:

$$B_{1/2} \subset Q_1 \subset Q_3 \subset B_{3\sqrt{n}/2} \subset B_{2\sqrt{n}}.$$

For M_1, $M_2 > 0$ and $\alpha > 0$ we define

$$\varphi(x) = M_1 - M_2|x|^{-\alpha} \quad \text{when } |x| \geq 1/2.$$

Since φ is an increasing function of $|x|$, we can find $M_1 = M_1(\alpha) > 0$ and $M_2 = M_2(\alpha) > 0$ such that

(i) $\varphi\big|_{\partial B_{2\sqrt{n}}} \equiv 0$, so that $\varphi \geq 0$ on $\mathbb{R}^n \setminus B_{2\sqrt{n}}$;

(ii) $\varphi\big|_{\partial B_{3\sqrt{n}/2}} \equiv -2$, so that $\varphi \leq -2$ on $Q_3 \setminus B_{1/2}$, since this set is contained in $B_{3\sqrt{n}/2}$.

After choosing a smooth extension for φ on $B_{1/2}$, still bounded from above by -2, we conclude checking that there exists an exponent α that is suitable to verify the third property of the statement, that needs to be checked only on $\mathbb{R}^n \setminus Q_1$. We compute

$$\nabla^2\left(|x|^{-\alpha}\right) = -\frac{\alpha}{|x|^{\alpha+2}}I + \frac{\alpha(\alpha+2)}{|x|^{\alpha+4}}\, x \otimes x,$$

thus the eigenvalues of $\nabla^2\varphi$ when $|x| \geq 1/2$ are $M_2\alpha|x|^{-(\alpha+2)}$ with multiplicity $n-1$ and $-M_2\alpha(\alpha+1)|x|^{-(\alpha+2)}$ with multiplicity 1 (this is the eigenvalue due to the radial direction). Hence, when $|x| \geq 1/2$, we have

$$\mathcal{M}^+(\nabla^2\varphi) = \frac{M_2}{|x|^{\alpha+2}}(\Lambda(n-1)\alpha - \lambda\alpha(\alpha+1))$$

so that $\mathcal{M}^+(\nabla^2\varphi) \leq 0$ on $\mathbb{R}^n \setminus B_{1/2}$ if we choose $\alpha = \alpha(n,\lambda,\Lambda)$ large enough. Since $B_{1/2} \subset Q_1$ and φ is smooth, we conclude that (iii) holds for a suitable constant $c = c(\varphi)$. $\square$

Lemma 5.52 (Decay Lemma)**.** *There exist universal constants $\varepsilon_0 > 0$, $M > 1$ and $\mu \in (0,1)$ such that, if $u \in \mathrm{Sup}(|f|)$, $u \geq 0$ on $Q_{4\sqrt{n}}$, $\inf_{Q_3} u \leq 1$ and $\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0$, then for every integer $k \geq 1$*

$$\mathscr{L}^n\left(\{u > M^k\} \cap Q_1\right) \leq (1-\mu)^k. \tag{5.49}$$

Proof. We prove the first step, that is

$$\mathscr{L}^n(\{u > M\} \cap Q_1) \leq (1-\mu), \tag{5.50}$$

with φ given by Lemma 5.51, $M := \max \varphi^- \geq 2$ and μ and ε_0 are respectively given by

$$\mu := (2c_{ABP}c)^{-n}, \qquad \varepsilon_0 = \frac{1}{2c_{ABP}}, \tag{5.51}$$

where c_{ABP} is the universal constant of the Aleksandrov-Bakelman-Pucci estimate of Theorem 5.40 and c as in the statement of Lemma 5.51. Since u is non-negative, in order to obtain a meaningful result from the ABP estimate, we apply the estimate in the ball $B_{2\sqrt{n}}$ for the function w, defined as the function u additively perturbed with the truncation function φ. If $w := u + \varphi$, then

$$w \geq 0 \quad \text{on } \partial B_{2\sqrt{n}} \tag{5.52}$$

because $u \geq 0$ on $Q_{4\sqrt{n}} \supset B_{2\sqrt{n}}$ and $\varphi \geq 0$ on $\mathbb{R}^n \setminus B_{2\sqrt{n}}$. Moreover

$$\inf_{B_{2\sqrt{n}}} w \leq \inf_{Q_3} w \leq -1 \tag{5.53}$$

because $Q_3 \subset B_{2\sqrt{n}}$ and $\varphi \leq -2$ on $B_{2\sqrt{n}}$, and, at the same time, we are assuming that $\inf_{Q_3} u \leq 1$. Directly from the definition of $\mathrm{Sup}(|f|)$ we get $-\mathcal{M}^-(\nabla^2 u) + |f| \geq 0$, moreover $\mathcal{M}^+(\nabla^2\varphi) \leq c\chi_{Q_1}$. Since in general $\mathcal{M}^-(A+B) \leq \mathcal{M}^-(A) + \mathcal{M}^+(B)$ (see Remark 5.37), then

$$-\mathcal{M}^-(\nabla^2 w) + (|f| + c\chi_{Q_1}) \geq (-\mathcal{M}^-(\nabla^2 u) + |f|) + (-\mathcal{M}^+(\nabla^2\varphi) + c\chi_{Q_1}) \geq 0. \tag{5.54}$$

The inequality (5.54) means that $w \in \mathrm{Sup}(|f| + c\chi_{Q_1})$. Thanks to the ABP estimate (which we can apply to w thanks to (5.52) and (5.54)) we get

$$\max_{\overline{B}_{2\sqrt{n}}} w^- \leq c_{ABP}\left(\int_{\{w=\Gamma_w\}} (|f| + c\chi_{Q_1})^n \, dy\right)^{1/n}. \tag{5.55}$$

Now, remembering that (5.53) holds and that, by definition, $\{w = \Gamma_w\} \subset \{w \leq 0\}$, we can expand (5.55) with

$$\begin{aligned} 1 &\leq \max_{x \in \overline{B}_{2\sqrt{n}}} w^-(x) \leq c_{ABP}\left(\int_{\{w\leq 0\}} (|f| + c\chi_{Q_1})^n \, dy\right)^{1/n} && (5.56)\\ &\leq c_{ABP}\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} + c_{ABP}c\mathscr{L}^n\left(Q_1 \cap \{w \leq 0\}\right)^{1/n} && (5.57)\\ &\leq c_{ABP}\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} + c_{ABP}c\mathscr{L}^n\left(Q_1 \cap \{u \leq M\}\right)^{1/n} && (5.58) \end{aligned}$$

where we go from line (5.56) to line (5.57) by Minkowski inequality and from line (5.57) to line (5.58) because, if $w(x) \leq 0$, then $u(x) \leq -\varphi(x)$ and then $u(x) \leq M$. Exploiting our choice of ε_0 we obtain from (5.58) the lower bound

$$\mathscr{L}^n \left(Q_1 \cap \{u \leq M\}\right)^{1/n} \geq \frac{1}{2c_{ABP}c}. \tag{5.59}$$

Thus, if μ is given by (5.51), we obtain (5.50).
We prove the inductive step: suppose that (5.49) holds for every $j \leq k-1$. We exploit the dyadic Lemma 5.50 with $A = \{u > M^k\} \cap Q_1$, $B = \{u > M^{k-1}\} \cap Q_1$ and $\delta = 1 - \mu$. Naturally $A \subset B \subset Q_1$ and $\mathscr{L}^n(A) \leq \delta$; if we are able to check that (5.45) holds, then

$$\mathscr{L}^n \left(Q_1 \cap \{u > M^k\}\right) \leq (1-\mu)\mathscr{L}^n \left(Q_1 \cap \{u > M^{k-1}\}\right) \leq (1-\mu)^k.$$

Concerning (5.45), suppose by contradiction that for some dyadic cube $Q \subset Q_1$ we have that

$$\mathscr{L}^n(A \cap Q) > \delta \mathscr{L}^n(Q) \tag{5.60}$$

but $\tilde{Q} \not\subset B$, $\tilde{Q}$ being the predecessor of Q, as usual: there exists $z \in \tilde{Q}$ such that $u(z) \leq M^{k-1}$. Let us rescale and translate the problem, putting $\tilde{u}(y) := u(x)M^{-(k-1)}$ with $x = x_0 + 2^{-i}y$ if Q has edge length 2^{-i} and centre x_0 (so that, in this transformation Q becomes the unit cube Q_1 and $\tilde{Q}$ is contained in Q_3). Because of the rescaling technique, we need to adapt f, that is define a new datum

$$\tilde{f}(y) := \frac{f(x)}{2^{2i}M^{k-1}}.$$

The point of this definition of $\tilde{f}$ is to ensure that $\tilde{u} \in \mathrm{Sup}(|\tilde{f}|)$, in fact formally[10]

$$(-\mathcal{M}^-(\nabla^2\tilde{u}) + |\tilde{f}|)(y) = \frac{1}{2^{2i}M^{k-1}} \left(-\mathcal{M}^-(\nabla^2 u) + |f|\right)(x) \geq 0,$$

because Pucci's operators are positively 1-homogeneous. Since the point corresponding to z belongs to Q_3, we get

$$\inf_{y \in Q_3} \tilde{u}(y) \leq \frac{u(z)}{M^{k-1}} \leq 1.$$

If $\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0$, then, applying what we already saw in (5.59) to $\tilde{u}$ instead of u,

$$\mu \leq \mathscr{L}^n \left(\{\tilde{u} \leq M\} \cap Q_1\right) = 2^{ni}\mathscr{L}^n \left(\{u \leq M^k\} \cap Q\right).$$

This means that $\mu\mathscr{L}^n(Q) \leq \mathscr{L}^n \left(\{u \leq M^k\} \cap Q\right)$ and, passing to the complement,

$$\mathscr{L}^n \left(\{u > M^k\} \cap Q\right) \leq (1-\mu)\mathscr{L}^n(Q),$$

[10] Thereby we mean that the identity in question has to be interpreted, as always with viscosity solutions, in terms of comparisons involving test functions.

which contradicts (5.60). In order to complete our proof, we show that $\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0$. To this aim, let us remark that in general the rescaling technique does not cause any problem at the level of the source term f. Indeed

$$\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} = \frac{1}{M^{k-1}2^i}\|f\|_{L^n(Q_{4\sqrt{n}/2^i}(x_0);\mathbb{R})} \leq \|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0.$$

□

Corollary 5.53. *There exist universal constants $\varepsilon > 0$ and $d \geq 0$ such that, if $u \in \mathrm{Sup}(|f|)$, $u \geq 0$ on $Q_{4\sqrt{n}}$, $\inf_{Q_3} u \leq 1$ and $\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0$, then*

$$\mathscr{L}^n\left(\{u > t\} \cap Q_1\right) \leq dt^{-\varepsilon} \qquad \forall\, t > 0. \tag{5.61}$$

Proof. This corollary is obtained by Lemma 5.52 choosing ε such that $(1-\mu) = M^{-\varepsilon}$ and $d' = M^{\varepsilon} = (1-\mu)^{-1}$: interpolating, for every $t \geq M$ there exists $k \in \mathbb{N}$ such that $M^{k-1} \leq t < M^k$, so

$$\mathscr{L}^n\left(\{u > t\} \cap Q_1\right) \leq \mathscr{L}^n\left(\{u > M^{k-1}\} \cap Q_1\right) \leq M^{-\varepsilon(k-1)} \leq d'(M^k)^{-\varepsilon} \leq d't^{-\varepsilon}.$$

Choosing $d \geq d'$ such that $1 \leq dt^{-\varepsilon}$ for all $t \in (0, M)$, we conclude. □

In the next lemma we use both the subsolution and the supersolution property to improve the decay estimate on $\mathscr{L}^n(\{u > t\})$ (the supersolution property is used to secure the validity of (5.61)). The statement is a bit technical and the reader might wonder about the choice of the scale ℓ_j as given in the statement of the lemma; it turns out, see (5.65), that this is (somehow) the smallest scale r on which we are able to say that $\mathscr{L}^n\left(\{u \geq \nu^j\} \cap Q_r\right) \ll r^n$, knowing that the global volume $\mathscr{L}^n\left(\{u \geq \nu^j\} \cap Q_1\right)$ is bounded by $d(\nu^j)^{-\varepsilon}$.

Lemma 5.54. *Suppose that $u \in \mathrm{Sub}(-|f|)$ is non-negative on $Q_{4\sqrt{n}}$ and $\|f\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} \leq \varepsilon_0$, with ε_0 given by the decay Lemma 5.52. Assume that (5.61) holds. Then there exist universal constants $M_0 > 1$ and $\sigma > 0$ such that, if*

$$x_0 \in Q_{1/2} \text{ and } u(x_0) \geq M_0\nu^{j-1} \text{ for some } j \geq 1,$$

then there exists

$$x_1 \in \overline{Q}_{l_j}(x_0) \text{ such that } u(x_1) \geq M_0\nu^j,$$

where $\nu := 2M_0/(2M_0 - 1) > 1$ and $\ell_j := \sigma M_0^{-\varepsilon/n}\nu^{-\varepsilon j/n}$.

Proof. First of all, we fix a large universal constant $\sigma > 0$ such that

$$\frac{1}{2}\left(\frac{\sigma}{4\sqrt{n}}\right)^n > d2^{\varepsilon} \tag{5.62}$$

and then we choose another universal constant M_0 so large that

$$dM_0^{-\varepsilon} < \frac{1}{2} \tag{5.63}$$

and

$$\sigma M_0^{-\varepsilon/n} < 2\sqrt{n}. \tag{5.64}$$

In the sequel of the proof we employ the following simple estimate on the superlevels of the function u:

$$\begin{aligned} \mathscr{L}^n\left(\{u \geq \nu^j M_0/2\} \cap Q_{\ell_j/(4\sqrt{n})}(x_0)\right) &\leq \mathscr{L}^n\left(\{u \geq \nu^j M_0/2\} \cap Q_1\right) \\ &\leq d\left(\nu^j M_0/2\right)^{-\varepsilon} \\ &< \frac{1}{2}\left(\frac{\sigma}{4\sqrt{n}}\right)^n \nu^{-j\varepsilon} M_0^{-\varepsilon} \\ &= \frac{1}{2}\left(\frac{\ell_j}{4\sqrt{n}}\right)^n. \end{aligned} \tag{5.65}$$

We used condition (5.62) on σ and the definition of ℓ_j, as given in the statement of the lemma.

By contradiction, assume that for some $j \geq 1$ we have

$$\max_{\overline{Q}_{\ell_j}(x_0)} u < M_0 \nu^j. \tag{5.66}$$

Under this assumption, we claim that the superlevel can be estimated as follows:

$$\mathscr{L}^n\left(\{u < \nu^j M_0/2\} \cap Q_{\ell_j/(4\sqrt{n})}(x_0)\right) < \frac{1}{2}\mathscr{L}^n\left(Q_{\ell_j/(4\sqrt{n})}(x_0)\right). \tag{5.67}$$

Obviously the validity of (5.65) and (5.67) at the same time is the contradiction that will conclude the proof, so we need only to show (5.67).

Therefore, define the auxiliary function

$$v(y) := \frac{\nu M_0 - u(x)\nu^{-(j-1)}}{(\nu-1)M_0} = 2\left(M_0 - \frac{u(x)}{\nu^j}\right),$$

where $x = x_0 + \frac{\ell_j}{4\sqrt{n}} y$ and the second identity is a consequence of the relation $M_0 = \nu/[2(\nu-1)]$. Since $y \in Q_{4\sqrt{n}} \iff x \in Q_{\ell_j}(x_0)$, by (5.66) the function v is defined and positive on $Q_{4\sqrt{n}}$. In addition, using the first equality in the definition of v, we immediately see that $u(x_0) \geq M_0 \nu^{j-1}$ implies $\inf_{Q_{4\sqrt{n}}} v \leq 1$. Using the second equality we see that (modulo the change of variables)

$$\{v > M_0\} = \{u < \nu^j M_0/2\}.$$

Moreover, if we compute the datum f_v which corresponds to v, since the rescaling radius is $\ell_j/(4\sqrt{n})$, we get

$$f_v(y) = -\frac{2\ell_j^2}{16n\nu^j} f(x)$$

so that

$$\|f_v\|_{L^n(Q_{4\sqrt{n}};\mathbb{R})} = \frac{2\ell_j}{4\sqrt{n}\nu^j}\|f\|_{L^n(Q_{\ell_j}(x_0);\mathbb{R})} \leq \varepsilon_0, \tag{5.68}$$

because

$$\frac{2\ell_j}{4\sqrt{n}\nu^j} = \frac{\sigma M_0^{-\varepsilon/n}}{2\sqrt{n}}\nu^{-(\varepsilon j/n+j)} < 1$$

thanks to (5.64). The estimate in (5.68) allows us to use Corollary 5.53 for v, that is

$$\mathscr{L}^n\left(\{v > M_0\} \cap Q_1\right) \leq dM_0^{-\varepsilon},$$

and we can use this, together with (5.63), to obtain that (5.67) holds:

$$\mathscr{L}^n\left(\{u < \nu^j M_0/2\} \cap Q_{\ell_j/(4\sqrt{n})}(x_0)\right) \leq dM_0^{-\varepsilon}\mathscr{L}^n\left(Q_{\ell_j/(4\sqrt{n})}(x_0)\right) < \frac{1}{2}\mathscr{L}^n\left(Q_{\ell_j/(4\sqrt{n})(x_0)}\right).$$

□

We can now complete the proof of Theorem 5.48, using Lemma 5.54. Notice that in Theorem 5.48 we made all assumptions needed to apply Lemma 5.54, taking also Corollary 5.53 into account, which ensures the validity of (5.61). Roughly speaking, if we assume, by (a sort of) contradiction, that u is not bounded from above by $M\nu^{k_0}$ on $Q_{1/4}$ for k_0 sufficiently large, then, thanks to Lemma 5.54, we should be able to find recursively a sequence (x_j) with the property that

$$u(x_j) \geq M_0\nu^j \quad \text{and} \quad x_{j+1} \in Q_{\ell_j}(x_j).$$

Since $\sum_j \ell_j < +\infty$, the sequence (x_j) converges, and we eventually find a contradiction at the limit point. However, in order to iterate Lemma 5.54 we have to confine the whole sequence in the cube $Q_{1/2}$ (for this purpose it is convenient to use the distance induced by the L^∞ norm in $\mathbb{R}^n$, whose balls are actually Euclidean cubes). To achieve this, we fix a universal positive integer j_0 such that $\sum_{j\geq j_0}\ell_j < 1/4$ and we assume, by contradiction, that there exists a point $x_0 \in Q_{1/4}$ with $u(x_0) \geq M_0\nu^{j_0-1}$. This time, the sequence (x_k) we generate iterating Lemma 5.54 is contained in $Q_{1/2}$ and

$$u(x_k) \geq M_0\nu^{j_0+k-1}. \tag{5.69}$$

When $k \to \infty$ in (5.69) we obtain the contradiction. This way, we obtained also an "explicit" expression of the universal constant in (5.43), in fact we proved that

$$\sup_{x\in Q_{1/4}} u(x) \leq M_0\nu^{j_0-1}.$$

A Some basic facts concerning Sobolev spaces

A.1 Two definitions and their comparison

We summarize here some basic facts on Sobolev spaces, which are needed throughout this monograph. For a more detailed treatment of these topics, the reader may consult [2], see also [8], [35], [36], [69] or [82].

It is possible to define Sobolev spaces in (at least) two different ways, whose partial equivalence is discussed below.

Definition A.1. *Let $\Omega \subset \mathbb{R}^n$ be an open domain and fix $p \in [1, \infty)$. Consider the subspace of regular functions $C^1(\overline{\Omega}; \mathbb{R})$ (i.e. the subset of $C^1(\Omega; \mathbb{R})$ consisting of functions u such that both u and ∇u admit a continuous extension to $\overline{\Omega}$) such that the norm*

$$\|u\|_{W^{1,p}(\Omega;\mathbb{R})} := \Big(\|u\|^p_{L^p(\Omega;\mathbb{R})} + \|\nabla u\|^p_{L^p(\Omega;\mathbb{R})} \Big)^{1/p}. \tag{A.1}$$

is finite. We define the Sobolev space $H^{1,p}(\Omega; \mathbb{R})$ to be the completion of such a subspace of $C^1(\overline{\Omega}; \mathbb{R})$ with respect to the $W^{1,p}$ norm.

This definition applies equally well to both bounded and unbounded domains, including the whole space $\mathbb{R}^n$. However, when $\Omega \subset \mathbb{R}^n$ is bounded, finiteness of $\|u\|_{W^{1,p}}$ is trivially satisfied for any $u \in C^1(\overline{\Omega}; \mathbb{R})$.

Notice that, even though *a priori* $H^{1,p}(\Omega; \mathbb{R})$ is an abstract completion, since $C^1_c(\Omega; \mathbb{R}) \subset L^p(\Omega; \mathbb{R})$ and the norm used for the completion is stronger, we can and will realize $H^{1,p}(\Omega; \mathbb{R})$ as a subset of $L^p(\Omega; \mathbb{R})$.

Alternatively, one can adopt a different viewpoint, inspired by the theory of distributions (see [86]).

Definition A.2. *Let $\Omega \subset \mathbb{R}^n$ be an open domain and consider the space $C^\infty_c(\Omega; \mathbb{R})$, whose elements will be called* test functions. *For $\alpha = 1, \dots, n$, we say that $u \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R})$ has α-th derivative in* weak sense *equal to $g_\alpha \in L^1_{\mathrm{loc}}(\Omega; \mathbb{R})$ if*

$$\int_\Omega u \partial_{x_\alpha} \varphi \, dx = - \int_\Omega \varphi g_\alpha \, dx \qquad \forall \, \varphi \in C^\infty_c(\Omega; \mathbb{R}). \tag{A.2}$$

Whenever such $g_1, \dots, g_n$ exist, we say that u is differentiable in weak sense and we write $g_\alpha = \partial_{x_\alpha} u$ and

$$\nabla u = \big(\partial_{x_1} u, \dots, \partial_{x_n} u \big).$$

For $p \in [1, \infty]$ we define the space $W^{1,p}(\Omega; \mathbb{R})$ as the subset of $L^p(\Omega; \mathbb{R})$ whose elements u are weakly differentiable with corresponding derivatives $\partial_{x_\alpha} u$ also belonging to $L^p(\Omega; \mathbb{R})$.

Remark A.3. Given $\Omega \subset \mathbb{R}^n$ an open domain, we shall say that $u \in H^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R})$ (respectively $u \in W^{1,p}_{\mathrm{loc}}(\Omega;\mathbb{R})$) if for every $x_0 \in \Omega$ there exists $r_0 > 0$ such that $B_{r_0}(x_0) \subset \Omega$ and $u \in H^{1,p}(B_{r_0}(x_0);\mathbb{R})$ (respectively $u \in W^{1,p}(B_{r_0}(x_0);\mathbb{R})$). Equivalently, one has $u \in H^{1,p}(\Omega';\mathbb{R})$ or $u \in W^{1,p}(\Omega';\mathbb{R})$ for any relatively compact open subdomain $\Omega' \Subset \Omega$. Analogous definitions are adopted, with straightforward modifications, for Lebesgue spaces L^p_{loc} and for higher-order Sobolev spaces, $H^{k,p}_{\mathrm{loc}}$ and $W^{k,p}_{\mathrm{loc}}$.

It is clear that if g_α exists, it must be uniquely determined modulo negligible sets (for the standard Lebesgue measure in $\mathbb{R}^n$), since $h \in L^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ and

$$\int_\Omega h\varphi\, dx = 0 \qquad \forall \varphi \in C^\infty_c(\Omega;\mathbb{R}) \tag{A.3}$$

implies $h = 0$. This implication can be easily proved by approximation, showing that the property above is stable under convolution. To that scope, let us consider a smooth, even and compactly supported function $\rho \in C^\infty_c(B_1;\mathbb{R})$ normalized to have unit L^1-norm and define $\rho_\varepsilon(x) := \varepsilon^{-n}\rho(x/\varepsilon)$. Then $h_\varepsilon = h * \rho_\varepsilon$ satisfies

$$\int_{\Omega_\varepsilon} h_\varepsilon\varphi\, dx = \int_\Omega h\varphi * \rho_\varepsilon\, dx = 0 \qquad \forall \varphi \in C^\infty_c(\Omega_\varepsilon;\mathbb{R}), \tag{A.4}$$

where Ω_ε is the (slightly) smaller domain

$$\Omega_\varepsilon := \{x \in \Omega : \ \mathrm{dist}(x, \partial\Omega) > \varepsilon\}\,, \tag{A.5}$$

and we exploited Fubini's theorem to prove the identity

$$\int_{\mathbb{R}^n} (a * \rho_\varepsilon) b\, dx = \int_{\mathbb{R}^n} a(b * \rho_\varepsilon)\, dx, \tag{A.6}$$

which holds true, for instance, for all $a \in L^p(\mathbb{R}^n;\mathbb{R}), b \in L^q(\mathbb{R}^n;\mathbb{R})$ with $p, q \geq 1$ satisfying $1 \leq 1/p + 1/q$. Hence if $h \in L^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ is assumed to satisfy (A.3) then from (A.4) we can trivially deduce that $h_\varepsilon = 0$ identically on Ω_ε for any $\varepsilon > 0$ and thus the conclusion follows from the fact that $\Omega_\varepsilon \uparrow \Omega$ and $h_\varepsilon \to h$ in $L^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ as we let $\varepsilon \to 0$.

Getting back to the discussion of weak derivatives of Sobolev functions, we shall recall a basic criterion.

Proposition A.4 (Stability of weak derivatives)**.** *Assume $u_k \in W^{1,p}(\Omega;\mathbb{R})$, $1 < p < \infty$ for $k \in \mathbb{N}$. If the sequence (u_k) converges strongly in $L^p(\Omega;\mathbb{R})$ to u and (∇u_k) is bounded in $L^p(\Omega;\mathbb{R}^n)$, then $u \in W^{1,p}(\Omega;\mathbb{R})$ and (∇u_k) weakly converges to ∇u in $L^p(\Omega;\mathbb{R}^n)$.*

Proof. By the reflexivity of $L^p(\Omega;\mathbb{R}^n)$, the sequence (∇u_k) has subsequential limits in the weak L^p topology as $k \to \infty$. By taking limits in the definition of weak gradient, we obtain that any limit point $g \in L^p(\Omega;\mathbb{R}^n)$ of (∇u_k) is the weak derivative of u. It follows that $u \in W^{1,p}(\Omega;\mathbb{R}^n)$, and that $g = \nabla u$ is uniquely determined. Again the reflexivity of $L^p(\Omega;\mathbb{R}^n)$ gives us that the whole family (∇u_k) weakly converges to ∇u in $L^p(\Omega;\mathbb{R}^n)$ as $k \to \infty$, which completes the argument. $\square$

Obviously, classical derivatives are weak derivatives and thus the notation $\partial_{x_\alpha} u$ is justified. We shall also recall a useful fact concerning the relation between weak and strong derivatives: if u has weak α-th derivative equal to g, then

$$\partial_{x_\alpha}(u * \rho_\varepsilon) = g * \rho_\varepsilon \qquad \text{in } \Omega_\varepsilon, \text{ in the classical sense.} \tag{A.7}$$

Knowing the identity (A.7) for smooth functions, its validity can be easily extended considering both sides as weak derivatives and using (A.6):

$$\begin{aligned}\int_\Omega (u * \rho_\varepsilon)\partial_{x_\alpha}\varphi\, dx &= \int_\Omega u(\partial_{x_\alpha}\varphi * \rho_\varepsilon)\, dx = \int_\Omega u\partial_{x_\alpha}(\varphi * \rho_\varepsilon)\, dx \\ &= -\int_\Omega g(\varphi * \rho_\varepsilon)\, dx = -\int_\Omega (g * \rho_\varepsilon)\varphi\, dx\end{aligned}$$

for all $\varphi \in C_c^\infty(\Omega;\mathbb{R})$ and for every $\varepsilon < \operatorname{dist}(\partial\Omega, \operatorname{supp}(\varphi))$. Now, the smoothness of $u * \rho_\varepsilon$ tells us that the derivative in the left-hand side of (A.7) is (equivalent to) a classical one.

An important consequence of (A.7) is the following assertion:

Theorem A.5 (Constancy theorem)**.** *If $u \in L^1_{\mathrm{loc}}(\Omega;\mathbb{R})$ satisfies $\nabla u = 0$ in the weak sense, then for any ball $B \subset \Omega$ there exists a constant $c \in \mathbb{R}$ such that $u(x) = c$ for $\mathscr{L}^n$-a.e. $x \in B$. In particular, if Ω is connected, the function u coincides $\mathscr{L}^n$-a.e. with a constant.*

Proof. Again we argue by approximation, using the fact that (A.7) ensures that the function $u * \rho_\varepsilon$ are locally constant in Ω_ε and taking the L^1_{loc} limit as $\varepsilon \to 0$. □

Notice also that Definition A.2 covers the case $p = \infty$, while it is not immediately clear how to adapt Definition A.1 to cover this case: usually H Sobolev spaces are defined for $p < \infty$ only. In fact, the formal extension of Definition A.1 to the borderline case $p = \infty$ would determine the space of equivalence classes (modulo coincidence $\mathscr{L}^n$-a.e.) of those functions $u \in C^1(\overline{\Omega};\mathbb{R})$ such that both u and ∇u are uniformly bounded in Ω. This should be compared with the characterization of the corresponding (much larger) W space, namely $W^{1,\infty}$, that we are about to discuss.

In the next proposition we indeed consider the relation of $W^{1,\infty}$ with the class of uniformly Lipschitz functions. For the sake of brevity, we omit the simple proof which also relies on the use of convolutions.

Proposition A.6 (Lipschitz versus $W^{1,\infty}$ functions)**.** *If $\Omega \subset \mathbb{R}^n$ is open, then one has the inclusion* $\operatorname{Lip}(\Omega;\mathbb{R}) \subset W^{1,\infty}(\Omega;\mathbb{R})$ *and*

$$\|\nabla u\|_{L^\infty(\Omega;\mathbb{R})} \le \operatorname{Lip}(u,\Omega). \tag{A.8}$$

In addition, if Ω is convex then $\operatorname{Lip}(\Omega;\mathbb{R}) = W^{1,\infty}(\Omega;\mathbb{R})$ *and equality holds in* (A.8).

In order to avoid ambiguities, we remind the reader of the definition of Lipschitz seminorm:

$$\mathrm{Lip}(u,\Omega) := \sup_{x\neq y\in\Omega} \frac{|u(x)-u(y)|}{|x-y|}$$

and recall that the space $\mathrm{Lip}(\Omega;\mathbb{R})$ consists of those functions such that this seminorm is finite.

Since $H^{1,p}(\Omega;\mathbb{R})$ is defined by means of approximation by regular functions, for which (A.2) is just the elementary "integration by parts formula", it is clear that $H^{1,p}(\Omega;\mathbb{R}) \subset W^{1,p}(\Omega;\mathbb{R})$; in addition, the same argument shows that the weak derivative of $u \in H^{1,p}(\Omega;\mathbb{R})$, in the sense of W Sobolev spaces, is precisely the strong $L^p(\Omega;\mathbb{R}^n)$ limit of ∇u_h, where $u_h \in C^1(\overline{\Omega};\mathbb{R})$ are strongly convergent to u. This allows to show by approximation some basic calculus rules for weak derivatives in H Sobolev spaces, such as the chain rule

$$\nabla(\phi\circ u) = \phi'(u)\nabla u \qquad \phi\in C^1(\mathbb{R};\mathbb{R}) \text{ and Lipschitz with } \phi(0)=0,\ u\in H^{1,p}(\Omega;\mathbb{R}) \tag{A.9}$$

and, with a little more effort (because one has first to show using the chain rule that bounded $H^{1,p}$ functions can be strongly approximated in $H^{1,p}$ by functions $C^1(\overline{\Omega};\mathbb{R})$ that are uniformly bounded together with their gradients) the Leibniz rule

$$\nabla(uv) = u\nabla v + v\nabla u \qquad u,\, v\in H^{1,p}(\Omega;\mathbb{R})\cap L^\infty(\Omega;\mathbb{R})\,. \tag{A.10}$$

On the other hand, we do not have to prove the same formulae for the W Sobolev spaces. Indeed, using convolutions and a suitable extension operator described below (in the case $\Omega=\mathbb{R}^n$ the proof is a direct application of (A.7), since in this case $\Omega_\varepsilon=\mathbb{R}^n$), one can prove the following result.

Theorem A.7 ($H=W$, [73])**.** *If either $\Omega=\mathbb{R}^n$ or Ω is a bounded regular domain, then*

$$H^{1,p}(\Omega;\mathbb{R}) = W^{1,p}(\Omega;\mathbb{R}) \qquad 1\le p<\infty. \tag{A.11}$$

Recall that the word *regular* is used in this monograph to describe any domain Ω that is locally the epigraph of a C^1 function of $(n-1)$-variables, written in a suitable system of coordinates, near any boundary point.

However the equality $H=W$ is not true in general, as the following example shows.

Example A.8. In the Euclidean plane $\mathbb{R}^2$, consider the open unit ball $\{x^2+y^2<1\}$ with one of its radii removed, say for instance the segment given by $(-1,0]\times\{0\}$. We can define on this domain Ω a function v having values in $(-\pi,\pi)$ and representing the angle in polar coordinates. Fix an exponent $1\le p<2$. It is immediate to see that $v\in C^\infty(\Omega;\mathbb{R})$ and that its gradient is p-integrable, hence $v\in W^{1,p}(\Omega;\mathbb{R})$. On the other

hand, we claim that $v \notin H^{1,p}(\Omega;\mathbb{R})$. Indeed, one can easily show, using Fubini's theorem and working in polar coordinates, that any $u \in H^{1,p}(\Omega;\mathbb{R})$ satisfies

$$\theta \mapsto u(re^{i\theta}) \in W^{1,p}_{\mathrm{loc}}(\mathbb{R};\mathbb{R}) \qquad \text{for } \mathscr{L}^1\text{-a.e. } r \in (0,1). \tag{A.12}$$

Indeed, if the sequence $u_n \in C^0(\overline{\Omega};\mathbb{R}) \cap C^1(\Omega;\mathbb{R})$ converges to u strongly in $H^{1,p}(\Omega;\mathbb{R})$ and (possibly extracting a subsequence) $\sum_n \|\nabla u_{n+1} - \nabla u_n\|_{L^p(\Omega;\mathbb{R})} < +\infty$, the pointwise inequality $|\nabla w| \geq r^{-1}|\partial_\theta w|$ gives

$$\int_0^1 \sum_n \left(\int_{-\pi}^{\pi} r^{1-p} \left|\partial_\theta u_{n+1} - \partial_\theta u_n\right|^p \, d\theta \right)^{1/p} dr \leq \sum_n \|\nabla u_{n+1} - \nabla u_n\|_{L^p(\Omega;\mathbb{R})} < +\infty.$$

It follows that

$$\sum_n \left(\int_{-\pi}^{\pi} \left|\partial_\theta u_{n+1} - \partial_\theta u_n\right|^p \, d\theta \right)^{1/p} < +\infty$$

for $\mathscr{L}^1$-a.e. $r \in (0,1)$, so that the 2π-periodic continuous functions $\theta \mapsto u_n(re^{i\theta})$ have derivatives strongly convergent in $L^p_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$, and therefore (by the fundamental theorem of calculus) are equicontinuous. Any limit point of these functions in $L^p_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ must then be equivalent to a 2π-periodic and continuous function. If, by contradiction, we take $u = v$, a similar Fubini argument shows that, whenever $\sum_n \|u_{n+1} - u_n\|_{L^p(\Omega;\mathbb{R})} < +\infty$, the sequence $u_n(re^{i\theta})$ converges in $L^p(-\pi,\pi)$ to the function v for $\mathscr{L}^1$-a.e. $r \in (0,1)$. But, the function $v(r,\theta) = \theta \in (-\pi,\pi)$ has no continuous 2π-periodic extension. Therefore we get a contradiction and we conclude that v cannot be in $H^{1,p}(\Omega;\mathbb{R})$.

Remark A.9. Taking into account the example above, we mention the more general result by Meyers-Serrin [73], asserting that, for *any* open set $\Omega \subset \mathbb{R}^n$ and $1 \leq p < \infty$, the identity

$$\overline{C^\infty(\Omega;\mathbb{R}) \cap W^{1,p}(\Omega;\mathbb{R})}^{W^{1,p}} = W^{1,p}(\Omega;\mathbb{R}) \tag{A.13}$$

holds. The proof relies on equation (A.7) and a partition of unity argument.

The previous example underlines the crucial role played by the boundary behavior, when we try to approximate a function in $W^{1,p}$ by $C^1(\overline{\Omega};\mathbb{R})$ (or even $C^0(\overline{\Omega};\mathbb{R}) \cap C^1(\Omega;\mathbb{R})$) functions. In the Meyers-Serrin theorem, instead, no smoothness up to the boundary is required for the approximating sequence. So, if we had defined the H spaces using $C^1(\Omega;\mathbb{R}) \cap L^p(\Omega;\mathbb{R})$ functions with gradient in $L^p(\Omega;\mathbb{R})$ instead of $C^1(\overline{\Omega};\mathbb{R})$ functions, the identity $H = W$ would have been true unconditionally. In the case $p = \infty$, the construction in the Meyers-Serrin theorem provides for all $u \in W^{1,\infty}(\Omega;\mathbb{R})$ a sequence $(u_n) \subset C^\infty(\Omega;\mathbb{R})$ converging to u locally uniformly in Ω, with $\sup_\Omega |\nabla u_n|$ convergent to $\|\nabla u\|_\infty$. Again, this fact can be exploited to give a definition of $H^{1,\infty}$ for which $H^{1,\infty} = W^{1,\infty}$ unconditionally.

As it will be clear soon, we also need to define an appropriate subspace of $H^{1,p}(\Omega;\mathbb{R})$ in order to work with functions vanishing at the boundary.

Definition A.10. *Given an open domain $\Omega \subset \mathbb{R}^n$, we define $H^{1,p}_0(\Omega;\mathbb{R})$ to be the completion of $C^1_c(\Omega;\mathbb{R})$ with respect to the $W^{1,p}$ norm.*

It is clear that $H^{1,p}_0(\Omega;\mathbb{R})$, being *complete*, is a *closed* subspace of $H^{1,p}(\Omega;\mathbb{R})$. Notice also that $H^{1,p}(\mathbb{R}^n;\mathbb{R})$ coincides with $H^{1,p}_0(\mathbb{R}^n;\mathbb{R})$. To see this, it suffices to show that any function $u \in C^1(\mathbb{R}^n;\mathbb{R})$ with both $|u|$ and $|\nabla u|$ in $L^p(\mathbb{R}^n;\mathbb{R})$ belongs to $H^{1,p}_0(\mathbb{R}^n;\mathbb{R})$. We can indeed approximate any such function u, strongly in $H^{1,p}$ norm, by the compactly supported functions $\chi_R u$, where the truncations $\chi_R : \mathbb{R}^n \to [0,1]$ are smooth, 2-Lipschitz, identically equal to 1 on $\overline{B}_R$ and identically equal to 0 on $\mathbb{R}^n \setminus \overline{B}_{R+1}$.

Remark A.11. Notice that one could equivalently define $H^{1,p}_0(\Omega;\mathbb{R})$ to be the completion of $C^\infty_c(\Omega;\mathbb{R})$ with respect to the $W^{1,p}$ norm. The proof relies on the use of convolutions, and we leave the easy details to the reader.

A.2 Poincaré inequalities

Theorem A.12 (Poincaré inequality, first version)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open bounded set and $p \in [1,\infty)$. Then there exists a constant $c_{P,I} = c_{P,I}(\Omega,p)$, depending only on Ω and p, such that*

$$\|u\|_{L^p(\Omega;\mathbb{R})} \le c_{P,I}\, \|\nabla u\|_{L^p(\Omega;\mathbb{R})} \qquad \forall u \in H^{1,p}_0(\Omega;\mathbb{R}). \tag{A.14}$$

In addition we can take $c_{P,I} \le c_P(n,p) \cdot \mathrm{diam}(\Omega)$.

The proof of this result can be simplified by observing that:

- $H^{1,p}_0(\Omega;\mathbb{R}) \subset H^{1,p}_0(\Omega',\mathbb{R})$ if $\Omega \subset \Omega'$ (*monotonicity*);
- if $c_{P,I}(\Omega,p)$ denotes the best Poincaré constant, then $c_{P,I}(\lambda\Omega,p) = \lambda c_{P,I}(\Omega,p)$ (*scaling invariance*) and $c_{P,I}(\Omega+h,p) = c_{P,I}(\Omega,p)$ (*translation invariance*).

The first fact is a consequence of the definition of the spaces $H^{1,p}_0$ in terms of regular functions, while the second one (translation invariance is obvious) follows by observing that

$$u_\lambda(x) = u(\lambda x) \in H^{1,p}_0(\Omega;\mathbb{R}) \text{ and } |\nabla u_\lambda(x)|^p = \lambda^p |\nabla u(\lambda x)|^p \quad \forall u \in H^{1,p}_0(\lambda\Omega;\mathbb{R}). \tag{A.15}$$

Proof. By the monotonicity and scaling properties, it is enough to prove the inequality for $\Omega = Q \subset \mathbb{R}^n$ where Q is the cube centered at the origin, with sides parallel to the coordinate axes and length 2. We write $x = (x_1, x')$ with $x' = (x_2,\dots,x_n)$. By density, we may also assume $u \in C^1_c(\Omega;\mathbb{R})$ and hence use the following representation formula:

$$u(x_1,x') = \int_{-1}^{x_1} \partial_{x_1} u(t,x')\, dt. \tag{A.16}$$

Hölder's inequality gives

$$|u|^p (x_1, x') \leq 2^{p-1} \int_{-1}^{1} |\partial_{x_1} u|^p (t, x') \, dt \tag{A.17}$$

and hence we just need to integrate with respect to x_1 to get

$$\int_{-1}^{1} |u|^p (x_1, x') \, dx_1 \leq 2^p \int_{-1}^{1} |\partial_{x_1} u|^p (t, x') \, dt. \tag{A.18}$$

Now, integrating with respect to x', we obtain the desired conclusion. □

Theorem A.13 (Rellich)**.** *Let Ω be an open bounded subset with regular boundary and let $p \in [1, \infty)$. Then the immersion $W^{1,p}(\Omega; \mathbb{R}) \hookrightarrow L^p(\Omega; \mathbb{R})$ is compact.*

We do not give a complete proof of this result. Instead, we observe that it can be obtained using an appropriate, linear and continuous *extension operator*

$$T : W^{1,p}(\Omega; \mathbb{R}) \to W^{1,p}(\mathbb{R}^n; \mathbb{R}) \tag{A.19}$$

such that

$$\operatorname{supp}(Tu) \subset \Omega', \quad \text{and} \quad Tu = u \quad \text{in } \Omega$$

where Ω' is a fixed open and bounded domain in $\mathbb{R}^n$ containing $\overline{\Omega}$. When Ω is a halfspace the operator can be defined by means of a standard reflection argument; in the general case the construction relies on the fact that the boundary of $\partial\Omega$ is regular and so can be locally *straightened* by means of C^1 maps (we use these ideas also in treating the boundary regularity of solutions to elliptic problems). The global construction is then obtained by means of a partition of unity.

The operator T allows a reduction to the case $\Omega = \mathbb{R}^n$, which is considered in the next theorem.

Theorem A.14. *For any $p \in [1, \infty)$ the immersion $W^{1,p}(\mathbb{R}^n; \mathbb{R}) \hookrightarrow L^p_{\text{loc}}(\mathbb{R}^n; \mathbb{R})$ is compact, namely if a sequence $(u_k) \subset W^{1,p}(\mathbb{R}^n; \mathbb{R})$ is bounded, then (u_k) has limit points in the $L^p_{\text{loc}}(\mathbb{R}^n; \mathbb{R})$ topology. Moreover, if $p > 1$ any limit point belongs to $W^{1,p}(\mathbb{R}^n; \mathbb{R})$.*

Remark A.15. It should be noted that the immersion $W^{1,p}(\mathbb{R}^n; \mathbb{R}) \hookrightarrow L^p(\mathbb{R}^n; \mathbb{R})$ is obviously continuous, but certainly not compact: just take a fixed non-zero element in $W^{1,p}(\mathbb{R}^n; \mathbb{R})$ with support in the unit square and consider the sequence of its translates along vectors τ_h with $|\tau_h| \to \infty$. Of course this is a bounded sequence in $W^{1,p}(\mathbb{R}^n; \mathbb{R})$ but no subsequence converges in $L^p(\mathbb{R}^n; \mathbb{R})$ (indeed, all functions have the same non-zero L^p norm, while it is easily seen that their L^p_{loc} limit is 0).

Let us now briefly sketch the main points of the proof of this theorem, since some of the ideas we use here are often considered in this monograph (see in part. Lemma 2.12).

Proof. Let us first observe that given any bounded, regular domain $A \subset \mathbb{R}^n$, an $|h|$-neighborhood $A_{|h|}$ of the set A and any $\varphi \in W^{1,p}(A_{|h|};\mathbb{R})$ we have

$$\|\tau_h\varphi - \varphi\|_{L^p(A;\mathbb{R})} \leq |h| \, \|\nabla\varphi\|_{L^p(A_{|h|};\mathbb{R})} \tag{A.20}$$

for $\tau_h\varphi(x) := \varphi(x+h)$. By approximation, we can assume with no loss of generality that $\varphi \in C^1(A_{|h|};\mathbb{R})$. The inequality (A.20) follows by the elementary representation

$$(\tau_h\varphi - \varphi)(x) = \int_0^1 \langle \nabla\varphi(x+sh), h\rangle \, ds, \tag{A.21}$$

because

$$\begin{aligned} \|\tau_h\varphi - \varphi\|^p_{L^p(A;\mathbb{R})} &\leq \int_A \int_0^1 |\langle \nabla\varphi(x+sh), h\rangle|^p \, ds\, dx && \text{(A.22)} \\ &\leq |h|^p \int_0^1 \int_{A_{|h|}} |\nabla\varphi(y)|^p \, dy\, ds = |h|^p \, \|\nabla\varphi\|^p_{L^p(A_{|h|};\mathbb{R})} && \text{(A.23)} \end{aligned}$$

where we used the (pointwise) Cauchy-Schwarz inequality followed by Hölder's inequality and Fubini's theorem. Hence, denoting by $(\rho_\varepsilon)_{\varepsilon>0}$ any rescaled family of smooth mollifiers such that $\mathrm{supp}(\rho_\varepsilon) \subset B_\varepsilon(0)$, we claim that for any $R > 0$

$$\sup_{k\in\mathbb{N}} \|\varphi_k * \rho_\varepsilon - \varphi_k\|_{L^p(B_R;\mathbb{R})} \to 0 \tag{A.24}$$

as one lets $\varepsilon \to 0$. Indeed, since $\varphi_k * \rho_\varepsilon$ is a mean, weighted by ρ_ε, of translates of φ_k, that is to say

$$\varphi_k * \rho_\varepsilon = \int_{\mathbb{R}^n} \tau_{-y}\varphi_k \rho_\varepsilon(y) \, dy,$$

then, by the previous result, we deduce

$$\sup_{k\in\mathbb{N}} \|\varphi_k * \rho_\varepsilon - \varphi_k\|_{L^p(B_R;\mathbb{R})} \leq \varepsilon \sup_{k\in\mathbb{N}} \left(\int_{B_{R+\varepsilon}} |\nabla\varphi_k|^p \, dx \right)^{1/p}. \tag{A.25}$$

To conclude we need to observe that the regularized sequence $(\varphi_k * \rho_\varepsilon)$ has a subsequence converging in $L^p_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ for any fixed $\varepsilon > 0$. But, in turn, that is easy since the Hölder inequality implies (with $1/p + 1/p' = 1$)

$$\sup_{B_R} |\varphi_k * \rho_\varepsilon| \leq \|\varphi_k\|_{L^p(B_{R+\varepsilon};\mathbb{R})} \, \|\rho_\varepsilon\|_{L^{p'}(\mathbb{R}^n;\mathbb{R})} \tag{A.26}$$

and similarly

$$\sup_{B_R} |\nabla(\varphi_k * \rho_\varepsilon)| \leq \|\varphi_k\|_{L^p(B_{R+\varepsilon};\mathbb{R})} \|\nabla \rho_\varepsilon\|_{L^{p'}(\mathbb{R}^n;\mathbb{R})} \tag{A.27}$$

so the claim follows by means of the Ascoli-Arzelà theorem and a standard diagonal argument. Notice that we used the gradient bounds on elements of our sequence only in (A.25).

If $p > 1$ the conclusion that any limit point belongs to $W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ follows from a direct application of Proposition A.4. To go from there to the assertion that in fact any limit point belongs to $W^{1,p}(\mathbb{R}^n;\mathbb{R})$ we just need to use the lower semicontinuity of the $W^{1,p}$ norm with respect to the corresponding weak convergence: indeed if $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ is a limit point then for any $R > 0$ we have

$$\|u\|_{W^{1,p}(B_R;\mathbb{R})} \leq \liminf_{k\to\infty} \|u_k\|_{W^{1,p}(B_R;\mathbb{R})} \leq \liminf_{k\to\infty} \|u_k\|_{W^{1,p}(\mathbb{R}^n;\mathbb{R})}$$

which is uniformly bounded by assumption. Thereby, a standard application of the monotone convergence theorem allows to complete the proof. □

We also need to mention another inequality due to Poincaré. The difference with respect to Theorem A.12 is that we do not impose any boundary behavior to the functions.

Theorem A.16 (Poincaré inequality, second version)**.** *Let us consider a bounded, regular and connected domain $\Omega \subset \mathbb{R}^n$ and an exponent $1 \leq p < \infty$, so that by Rellich's theorem we have the compact immersion $W^{1,p}(\Omega;\mathbb{R}) \hookrightarrow L^p(\Omega;\mathbb{R})$. Then, there exists a constant $c_{P,II} = c_{P,II}(\Omega, p)$ such that*

$$\int_\Omega |u - u_\Omega|^p \, dx \leq c_{P,II} \int_\Omega |\nabla u|^p \, dx \qquad \forall u \in W^{1,p}(\Omega;\mathbb{R}), \tag{A.28}$$

where $u_\Omega = \fint_\Omega u \, dx$.

Proof. By contradiction, if the desired inequality were not true, exploiting its homogeneity and translation invariance we could find a sequence $(u_h) \subset W^{1,p}(\Omega;\mathbb{R})$ such that

- $(u_h)_\Omega = 0$ for all $h \in \mathbb{N}$;
- $\int_\Omega |u_h|^p \, dx = 1$ for all $h \in \mathbb{N}$;
- $\int_\Omega |\nabla u_h|^p \, dx \to 0$ for $h \to \infty$.

By Rellich's theorem there exists a limit point $u \in L^p(\Omega;\mathbb{R})$. Invoking Proposition A.4 we get $u \in W^{1,p}(\Omega;\mathbb{R})$ and by virtue of the third condition above $\nabla u = 0$ in $L^p(\Omega;\mathbb{R})$. Hence, by connectivity of the domain and the constancy theorem A.5, we deduce that u must be equivalent to a constant. By taking limits we see that u satisfies at the same time

$$\int_\Omega u \, dx = 0 \qquad \text{and} \qquad \int_\Omega |u|^p \, dx = 1, \tag{A.29}$$

which is clearly impossible. □

Note that the previous proof is not constructive and crucially relies on the general compactness result by Rellich.

Remark A.17. It should be observed that the previous proof, even though very simple, is far from giving the sharp constant for the Poincaré inequality (A.28). The determination of the sharp constant is a difficult problem, solved only in very special cases (for instance on intervals of the real line and $p = 2$, by Fourier analysis). More results are instead available for the sharp constant in the Poincaré inequality (A.14) (see, for instance, [9]).

Remark A.18. All definitions and results in this chapter, including most remarkably Theorem A.12 and Theorem A.16, can easily be extended to vector-valued maps with changes of purely notational character.

Remark A.19. As it is customary in the literature, we will use the notation $H^k(\Omega;\mathbb{R})$ in lieu of $H^{k,2}(\Omega;\mathbb{R})$ for *Hilbertian* Sobolev spaces, namely in the special case when $p = 2$. Unless otherwise stated, this convention will always be tacitly adopted throughout this monograph.

A.3 Sobolev inequalities

In this section, we present a proof of the basic Sobolev inequalities ensuring improved summability of functions whose gradient is itself integrable. To that scope, we first need to recall the statement of two isoperimetric inequalities.

We say that a set $E \subset \mathbb{R}^n$ is regular if it is locally the epigraph of a C^1 function. In this case, it is well-known that by local parametrizations and a partition of unity, we can define $\sigma_{n-1}(\partial E)$, the $(n-1)$-dimensional surface measure of ∂E. In fact, this coincides with the $(n-1)$ dimensional Hausdorff measure of the same set (see Appendix C), so we shall simply write $\mathscr{H}^{n-1}(\partial E)$ in lieu of $\sigma_{n-1}(\partial E)$ for notational consistency.

Of course, regular sets provide a very unnatural (somehow too restrictive) setting for isoperimetric inequalities, but they suffice for our purposes. We shall now state, without proof, two isoperimetric inequalities:

Theorem A.20 (Isoperimetric inequality)**.** *Let $E \subset \mathbb{R}^n$ be a regular set such that $\mathscr{H}^{n-1}(\partial E) < +\infty$. Then*

$$\min\{\mathscr{L}^n(E), \mathscr{L}^n(\mathbb{R}^n \setminus E)\} \leq c_I \left[\mathscr{H}^{n-1}(\partial E)\right]^{1^*}$$

with $c_I = c_I(n)$ a dimensional constant.

In the previous statement and throughout this monograph, we let p^* denote the Sobolev dual exponent of $p \geq 1$, defined by $p^* = np/(n-p)$. As a special case, $1^* = n/(n-1)$.

It is well-known that the best constant $c_I(n)$ in the previous inequality is given by

$$\mathscr{L}^n(B_1)/[\mathscr{H}^{n-1}(\partial B_1)]^{1^*} = \omega_n/[n\omega_n]^{1^*},$$

that is to say: balls have the best isoperimetric ratio.

Theorem A.21 (Relative isoperimetric inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded and regular set, which we further assume to be connected. Then there exists a constant $c_{I,R} = c_{I,R}(\Omega)$ such that, for every $E \subset \Omega$ with $\Omega \cap \partial E \in C^1$, one has*

$$\min\{\mathscr{L}^n(E), \mathscr{L}^n(\Omega \setminus E)\} \leq c_{I,R} \left[\mathscr{H}^{n-1}(\Omega \cap \partial E)\right]^{1^*}.$$

Let us introduce another classical tool in Geometric Measure Theory.

Theorem A.22 (Coarea formula)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in C^\infty(\Omega;\mathbb{R})$ be non-negative, then*

$$\int_\Omega |\nabla u|\, dx = \int_0^\infty \mathscr{H}^{n-1}\left(\Omega \cap \{u = t\}\right)\, dt.$$

Remark A.23. It should be observed that the right-hand side of the previous formula is well-defined, since by the classical *Sard's theorem*

$$u \in C^\infty(\Omega;\mathbb{R}) \quad \Longrightarrow \quad \mathscr{L}^1\big(\{u(x) :\ x \in \Omega,\ \nabla u(x) = 0\}\big) = 0.$$

By the implicit function theorem this implies that almost every sublevel set $\{u < t\}$ is regular and that its boundary actually coincides with the level set $\{u = t\}$.

Proof. A complete proof will not be described here since it is far from the main purpose of this textbook, however we sketch the main points. The interested reader may consult, for instance, [37].

We first prove the inequality $\int_\Omega |\nabla u|\, dx \leq \int_0^\infty \mathscr{H}^{n-1}\left(\Omega \cap \{u = t\}\right)\, dt$. The pointwise identity

$$u(x) = \int_0^\infty \chi_{\{u>t\}}(x)\, dt$$

implies, by applying Fubini's theorem, that

$$\begin{aligned}
\int_\Omega |\nabla u|\, dx &= \sup_{\varphi \in C^1_c,\, |\varphi| \leq 1} \int_\Omega \langle \nabla u, \varphi\rangle\, dx = \sup_{\varphi \in C^1_c,\, |\varphi| \leq 1} \int_\Omega u\, \mathrm{div}\varphi\, dx \\
&= \sup_{\varphi \in C^1_c,\, |\varphi| \leq 1} \int_0^\infty \left(\int_\Omega (\mathrm{div}\varphi)\chi_{\{u>t\}}\, dx\right) dt \\
&\leq \int_0^\infty \left(\sup_{\varphi \in C^1_c,\, |\varphi| \leq 1} \int_{\{u>t\}} \mathrm{div}\varphi\, dx\right) dt.
\end{aligned}$$

Hence, by the Gauss-Green theorem (with ν_t outer normal to $\{u > t\}$) we obtain

$$\int_\Omega |\nabla u|\, dx \leq \int_0^\infty \left(\sup_{\varphi \in C^1_c,\, |\varphi| \leq 1} \int_{\Omega \cap \{u=t\}} \langle \varphi, \nu_t\rangle\ d\mathscr{H}^{n-1}\right) dt \leq \int_0^\infty \mathscr{H}^{n-1}\left(\Omega \cap \{u = t\}\right)\, dt,$$

exploiting the fact that for $\mathscr{L}^1$-a.e. $t \in \mathbb{R}$ the set $\{u = t\}$ is the (regular) boundary of $\{u > t\}$.

Let us then consider the converse inequality, namely

$$\int_\Omega |\nabla u|\, dx \geq \int_0^\infty \mathscr{H}^{n-1}\left(\Omega \cap \{u = t\}\right)\, dt.$$

It is not restrictive to assume that Ω is a cube. At that stage, one can prove the claim above (with equality, in fact) if u is continuous and piecewise linear, since on each subdomain of a triangulation of Ω the coarea formula just reduces to Fubini's Theorem. The general case is obtained by an approximation argument, choosing piecewise affine functions which converge to u in $W^{1,1}(\Omega;\mathbb{R})$ and using Fatou's lemma and the lower semicontinuity of $E \mapsto \mathscr{H}^{n-1}(\Omega \cap \partial E)$ (this, in turn, follows by the sup formula we already used in the proof of the first inequality). We omit the details. □

In order to deduce the desired Sobolev embeddings, we need a technical lemma.

Lemma A.24. *Let $G : [0,\infty) \to [0,\infty)$ a non-increasing measurable function. Then for any $\alpha \geq 1$ we have*

$$\alpha \int_0^\infty t^{\alpha-1} G(t)\, dt \leq \left(\int_0^\infty G^{1/\alpha}(t)\, dt\right)^\alpha.$$

Proof. It is sufficient to prove that for any $T > 0$ we have the inequality

$$\alpha \int_0^T t^{\alpha-1} G(t)\, dt \leq \left(\int_0^T G^{1/\alpha}(t)\, dt\right)^\alpha. \tag{A.30}$$

Since G is non-increasing, we can observe that

$$G^{1/\alpha}(t) \leq \fint_0^t G^{1/\alpha}(s)\, ds,$$

which, raising both sides to the power $\alpha - 1$, is equivalent to

$$t^{\alpha-1} G(t) \leq \left(\int_0^t G^{1/\alpha}(s)\, ds\right)^{\alpha-1} G^{1/\alpha}(t).$$

Then, multiplying both sides by α, (A.30) follows by integration on $[0,T]$. □

We are now ready to derive the Sobolev inequalities.

Theorem A.25 (Sobolev embedding, $p = 1$)**.** *For any $u \in W^{1,1}(\mathbb{R}^n;\mathbb{R})$ we have that*

$$\left(\int_{\mathbb{R}^n} |u|^{1^*}\, dx\right)^{1/1^*} \leq c_{S,1} \int_{\mathbb{R}^n} |\nabla u|\, dx,$$

where $c_{S,1} = c_{S,1}(n)$. As a result, we have the following continuous embeddings:

(1) $W^{1,1}(\mathbb{R}^n;\mathbb{R}) \hookrightarrow L^{1^*}(\mathbb{R}^n;\mathbb{R})$;

(2) for any $\Omega \subset \mathbb{R}^n$ *open, bounded and regular* $W^{1,1}(\Omega;\mathbb{R}) \hookrightarrow L^{1^*}(\Omega;\mathbb{R})$.

Proof. By a truncation argument it is possible to reduce the thesis to the case $u \geq 0$, and smoothing reduces the proof to the case $u \in C^\infty$. Under these assumptions we have

$$\int_{\mathbb{R}^n} u^{1^*}\,dx = 1^* \int_0^\infty t^{1^*-1}\mathscr{L}^n(\{u>t\})\,dt \leq \left(\int_0^\infty \mathscr{L}^n(\{u>t\})^{1/1^*}\,dt\right)^{1^*}$$

thanks to Lemma A.24. Consequently, since the fact that $u \in L^1$ ensures that all sets $\{u>t\}$ have finite measure, the isoperimetric inequality and the coarea formula give

$$\int_{\mathbb{R}^n} u^{1^*}\,dx \leq c_I \left(\int_0^\infty \mathscr{H}^{n-1}\big(\{u=t\}\big)\,dt\right)^{1^*} = c_I \left(\int_{\mathbb{R}^n} |\nabla u|\,dx\right)^{1^*}.$$

The continuous embedding in (2) follows by the global one in (1) applied to an extension of u (recall that regularity of $\partial\Omega$ yields the existence of a continuous extension operator from $W^{1,1}(\Omega;\mathbb{R})$ to $W^{1,1}(\mathbb{R}^n;\mathbb{R})$). □

Theorem A.26 (Sobolev embeddings, $1<p<n$)**.** *For any* $u \in W^{1,p}(\mathbb{R}^n;\mathbb{R})$ *we have that*

$$\left(\int_{\mathbb{R}^n} |u|^{p^*}\,dx\right)^{1/p^*} \leq c_{S,p}\left(\int_{\mathbb{R}^n} |\nabla u|^p\,dx\right)^{1/p},$$

where $c_{S,p} = c_{S,p}(n,p)$ *is a constant depending on* p *and the dimension* n*. As a result, the have the following continuous embeddings:*

(1) $W^{1,p}(\mathbb{R}^n;\mathbb{R}) \hookrightarrow L^{p^*}(\mathbb{R}^n;\mathbb{R})$;

(2) for any $\Omega \subset \mathbb{R}^n$ *open, bounded and regular* $W^{1,p}(\Omega;\mathbb{R}) \hookrightarrow L^{p^*}(\Omega;\mathbb{R})$.

Proof. Again, it is enough to study the case $u \geq 0$. Furthermore, given the identities $W^{1,p}(\mathbb{R}^n;\mathbb{R}) = H^{1,p}(\mathbb{R}^n;\mathbb{R}) = H^{1,p}_0(\mathbb{R}^n;\mathbb{R})$ (see Theorem A.7 and the remarks presented after Definition A.10) we can assume, without loss of generality, that u is smooth and compactly supported. That being said, we can exploit the case $p=1$ to get

$$\left(\int_{\mathbb{R}^n} u^{\alpha 1^*}\,dx\right)^{1/1^*} \leq c_{S,1}\int_{\mathbb{R}^n} \alpha u^{\alpha-1}|\nabla u|\,dx \qquad \forall\,\alpha>1$$

and, by Hölder's inequality, the right-hand side can be estimated from above with

$$c_{S,1}\alpha\left[\int_{\mathbb{R}^n} u^{(\alpha-1)p'}\,dx\right]^{1/p'}\left[\int_{\mathbb{R}^n} |\nabla u|^p\,dx\right]^{1/p}.$$

Now, choose α such that $\alpha 1^* = (\alpha - 1)p'$, namely we set $\alpha = p^*/1^* > 1$. Consequently, by simply combining the two inequalities above

$$\left(\int_{\mathbb{R}^n} u^{\alpha 1^*}\,dx\right)^{1/1^*-1/p'} \leq c\left(\int_{\mathbb{R}^n} |\nabla u|^p\,dx\right)^{1/p},$$

but $1/1^* - 1/p' = 1/p^*$ and thus the claim follows. The second part of the statement can be obtained as described in the proof of Theorem A.25. □

Remark A.27. *While natural from the perspective of Functional Analysis, the above assumptions are not really sharp. In fact, one can prove the following: if $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ and $|\nabla u| \in L^p(\mathbb{R}^n;\mathbb{R})$ then there exists a (unique) constant $c \in \mathbb{R}$ such that $u - c \in L^{p^*}(\mathbb{R}^n;\mathbb{R})$, furthermore*

$$\left(\int_{\mathbb{R}^n} |u-c|^{p^*}\,dx\right)^{1/p^*} \leq c_{S,p}\left(\int_{\mathbb{R}^n} |\nabla u|^p\,dx\right)^{1/p}.$$

The necessary modifications that are needed to prove this assertion are left as an exercise for the reader.

We will also make use of the following refinement of the Poincaré inequality: even though no assumption is made on the behaviour of u at the boundary of the domain, it is still possible to control the L^{1^*} norm with the gradient.

Theorem A.28. *Let $u \in W^{1,1}(B_R;\mathbb{R})$ with $u \geq 0$ and suppose that one has $\mathscr{L}^n(\{u=0\}) \geq \mathscr{L}^n(B_R)/2$. Then*

$$\left(\int_{B_R} u^{1^*}\,dx\right)^{1/1^*} \leq c\int_{B_R} |\nabla u|\,dx,$$

where $c = c(n)$ is a constant depending only on the dimension n.

Proof. This result is the local version of the embedding $W^{1,1} \hookrightarrow L^{1^*}$. Hence, in order to give the proof, it is enough to follow the argument presented for Theorem A.25 replacing the use of the isoperimetric inequality with that of the relative isoperimetric inequality, that is

$$\mathscr{L}^n(B_R \cap \{u > t\}) \leq c\mathscr{H}^{n-1}(B_R \cap \{u = t\})^{1^*}.$$

To that scope we proceed as follows. First of all observe that the assumption $\mathscr{L}^n(\{u=0\}) \geq \omega_n R^n/2$ is equivalent to having $\mathscr{L}^n(\{u<\delta\}) \geq \omega_n R^n/2$ for every $\delta > 0$. Moreover, by a perturbation argument we can reduce to the case $\mathscr{L}^n(\{u \leq 0\}) > \omega_n R^n/2$, so that $\mathscr{L}^n\{u \leq \delta\} > \omega_n R^n/2$ for all $\delta > 0$. Given u as in the statement, let (u_h) be an approximating sequence of smooth functions, obtained invoking Theorem A.7. Since the constructions by Meyers-Serrin preserves the sign, we can assume $\mathscr{L}^n(\{u_h < \delta\}) \geq \omega_n R^n/2$

for h large enough (depending on δ). Hence one can apply the relative isoperimetric inequality to the function $\varphi(u_h)$ where φ is gotten by smoothing the function $z \mapsto (z-\delta)^+$. At this stage, the conclusion comes by first letting $h \to \infty$ and then $\delta \to 0$. We leave the details to the reader. $\square$

B Miscellaneous results in real and harmonic analysis

The following two lemmata are useful in the study of weak L^p spaces and of the Hardy-Littlewood maximal operator, which is the one of the key points we wish to recall in this appendix.

Lemma B.1. *In a measure space $(\Omega, \mathcal{A}, \mu)$, consider an $\mathcal{A}$-measurable function $f : \Omega \to [0, +\infty]$ and set*

$$F(t) := \mu\left(\{x \in \Omega : \ f(x) > t\}\right).$$

The following equalities hold for $1 \leq p < \infty$:

$$\int_\Omega f^p(x)\, d\mu(x) = p \int_0^\infty t^{p-1} F(t)\, dt \tag{B.1}$$

$$\int_{\{f>s\}} f^p(x)\, d\mu(x) = p \int_s^\infty t^{p-1} F(t)\, dt + s^p F(s) \qquad 0 \leq s < +\infty. \tag{B.2}$$

Proof. It is a simple consequence of Fubini's Theorem that

$$\begin{aligned}\int_\Omega f^p(x)\, d\mu(x) &= \int_\Omega p \left(\int_0^{f(x)} t^{p-1}\, dt \right) d\mu(x) = p \int_0^\infty t^{p-1} \left(\int_\Omega \chi_{\{f(x)>t\}}\, d\mu(x) \right) dt \\ &= p \int_0^\infty t^{p-1} F(t)\, dt.\end{aligned}$$

Equation (B.2) follows from (B.1) applied to the function $f\chi_{\{f>s\}}$. □

Theorem B.2 (Markov inequality)**.** *In a measure space $(\Omega, \mathcal{A}, \mu)$, an $\mathcal{A}$-measurable function $f : \Omega \to [0, +\infty]$ satisfies (with the standard convention $0 \times \infty = 0$)*

$$t^p \mu\left(\{f \geq t\}\right) \leq \int_\Omega f^p\, d\mu \qquad \forall t \geq 0. \tag{B.3}$$

Proof. We begin with the trivial pointwise inequality

$$s\chi_{\{g \geq s\}}(x) \leq g(x) \qquad \forall\, x \in \Omega \tag{B.4}$$

for g non-negative. Thus, integrating (B.4) in Ω we obtain

$$s\mu\left(\{g \geq s\}\right) \leq \int_\Omega g\, d\mu.$$

The thesis follows choosing $s = t^p$ and $g = f^p$. □

B.1 Weak L^p spaces and the maximal operator

The Markov inequality inspires the definition of a space which is weaker than L^p, but whose elements still satisfy (B.3).

Definition B.3 (Marcinkiewicz space)**.** *Given a measure space $(\Omega, \mathcal{A}, \mu)$ and an exponent $1 \le p < \infty$, the Marcinkiewicz space $L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R})$ is defined by*

$$L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R}) := \big\{ f : \Omega \to \mathbb{R} \quad \mathcal{A}\text{-measurable} \ : \ \sup_{t>0} t^p \mu\left(\{|f| > t\}\right) < +\infty \big\}.$$

For $f \in L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R})$ we shall denote with $\|f\|^p_{L^p_w}$ the smallest constant c satisfying

$$t^p \mu\left(\{|f| > t\}\right) \le c \quad \forall\, t > 0\,.$$

Remark B.4. Pay attention to the lack of subadditivity of $\|\cdot\|_{L^p_w}$: the notation is misleading, as this is not a norm. For instance both $1/x$ and $1/(1-x)$ have weak L^1 norm equal to 1 on $\Omega = (0,1)$, but their sum has weak L^1 norm strictly larger than two. On the other hand, it is easily seen that $\|f+g\|_{L^p_w} \le 2\|f\|_{L^p_w} + 2\|g\|_{L^p_w}$.

Remark B.5. Observe that, due to Markov inequality (B.2),

$$L^p((\Omega, \mathcal{A}, \mu); \mathbb{R}) \subset L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R}).$$

Moreover, if μ is a finite measure, then

$$q < p \quad \Longrightarrow \quad L^p((\Omega, \mathcal{A}, \mu); \mathbb{R}) \subset L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R}) \subset L^q((\Omega, \mathcal{A}, \mu); \mathbb{R}).$$

Indeed, if $f \in L^p_w((\Omega, \mathcal{A}, \mu); \mathbb{R})$ and $F(t) = \mu(\{|f| > t\})$, then

$$\begin{aligned} \int_\Omega |f|^q \, d\mu(x) &= q \int_0^\infty t^{q-1} F(t)\, dt \le q \left(\int_0^1 t^{q-1} F(t)\, dt + \int_1^\infty t^{q-1} F(t)\, dt \right) \\ &\le \mu(\Omega) + q \int_1^\infty t^{q-1} \|f\|^p_{L^p_w} t^{-p}\, dt = \mu(\Omega) + \frac{q}{p-q} \|f\|^p_{L^p_w}. \end{aligned}$$

Definition B.6 (Maximal operator)**.** *When $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R})$ we define the maximal function $\mathcal{M}f$ by*

$$\mathcal{M}f(x) := \sup_{r>0} \fint_{Q_r(x)} |f(y)|\, dy, \tag{B.5}$$

where $Q_r(x)$ is the n-dimensional open cube with center x and side length r.

It is easy to check that $\mathcal{M}f(x) \ge |f(x)|$ at Lebesgue points, so that $\mathcal{M}f \ge |f|$ $\mathscr{L}^n$-a.e. in $\mathbb{R}^n$ (see also Theorem B.16). On the other hand, it is important to remark that the maximal operator $\mathcal{M}$ does not map L^1 into L^1.

Example B.7. In dimension $n = 1$, consider $f = \chi_{[0,1]} \in L^1(\mathbb{R};\mathbb{R})$. Then

$$\mathcal{M}f(x) = \frac{1}{2|x|} \qquad \text{when } |x| \geq 1,$$

so $\mathcal{M}f \notin L^1(\mathbb{R};\mathbb{R})$. In fact, it is easy to prove that $\mathcal{M}f \in L^1(\mathbb{R};\mathbb{R})$ implies $|f| = 0$ $\mathscr{L}^n$-a.e. in $\mathbb{R}^n$.

However, if $f \in L^1(\mathbb{R}^n;\mathbb{R})$, the maximal function $\mathcal{M}f$ belongs to the weaker Marcinkiewicz space $L^1_w(\mathbb{R}^n;\mathbb{R})$, as we are going to see in Theorem B.9. We first recall the Vitali covering theorem, in a version valid in *any* metric space.

Lemma B.8 (Vitali)**.** *Let $\mathcal{E}$ be a finite family of open balls*[11] *in a metric space (X, d). Then, there exists a subfamily $\mathcal{E}' \subset \mathcal{E}$, consisting of disjoint balls, satisfying*

$$\bigcup_{B\in\mathcal{E}} B \subset \bigcup_{B\in\mathcal{E}'} \hat{B}.$$

Here, for B ball, $\hat{B}$ denotes the ball with the same center and triple radius.

Proof. The initial remark is that if B_1 and B_2 are intersecting balls then $B_1 \subset \widehat{B_2}$, provided the radius of B_2 is larger than or equal to the radius of B_1. Assume that the family of balls is ordered in such a way that their radii are non-increasing. Pick the first ball B_1, then pick the first ball among those that do not intersect B_1 and continue in this way, until either there is no ball left or all the balls left intersect one of the chosen balls. The family $\mathcal{E}'$ of chosen balls is, by construction, disjoint. If $B \in \mathcal{E} \setminus \mathcal{E}'$, then B has not been chosen because it intersects one of the balls in $\mathcal{E}'$; the first of these balls (say B_e) has radius larger than or equal to the radius of B (otherwise B would have been chosen before B_e), hence $B \subset \widehat{B_e}$. □

Theorem B.9 (Hardy-Littlewood maximal theorem)**.** *The maximal operator $\mathcal{M}f$ defined in* (B.5) *satifies*

$$\|\mathcal{M}f\|_{L^1_w} \leq 3^n \|f\|_{L^1} \qquad \forall f \in L^1(\mathbb{R}^n;\mathbb{R}).$$

Proof. Fix $t > 0$ and a compact set $K \subset \{\mathcal{M}f > t\}$: by inner regularity of the Lebesgue measure we reach the conclusion if we show that

$$\mathscr{L}^n(K) \leq \frac{3^n}{t}\|f\|_{L^1}.$$

Since $K \subset \{\mathcal{M}f > t\}$, for any $x \in K$ there exists a radius $r(x)$ such that

$$\int_{Q_{r(x)}(x)} |f(y)|\, dy \geq t(r(x))^n.$$

[11] The same statement holds for a finite family of closed balls, too.

Compactness allows us to cover K with a finite number of open cubes, i.e.

$$K \subset \bigcup_{i \in I} Q_{r(x_i)}(x_i).$$

Now Vitali's lemma (B.8), applied with the distance induced by the sup norm in $\mathbb{R}^n$, allows us to find a subfamily $J \subset I$ such that the cubes $Q_{r(x_j)}(x_j)$, $j \in J$, are pairwise disjoint and

$$\bigcup_{j \in J} Q_{3r(x_j)}(x_j) \supset \bigcup_{i \in I} Q_{r_i}(x_i) \supset K.$$

We conclude that

$$\mathscr{L}^n(K) \leq \sum_{j \in J} 3^n (r(x_j))^n \leq \frac{3^n}{t} \sum_{j \in J} \int_{Q_{r(x_j)}(x_j)} |f(y)|\, dy \leq \frac{3^n}{t} \|f\|_{L^1}.$$

□

B.2 Some classical interpolation theorems

In the sequel of this chapter, we will make extensive use of some classical interpolation theorems, that are basic tools in functional and harmonic analysis.

Assume $(X, \mathcal{A}, \mu)$ is a measure space. For the sake of brevity, we will say that a linear operator T mapping a vector space $D \subset L^p((X, \mathcal{A}, \mu); \mathbb{R})$ into $L^q((X, \mathcal{A}, \mu); \mathbb{R})$ is *of type* (p, q) if it is continuous with respect to the $L^p - L^q$ topologies. If this happens and D is dense, T can be uniquely extended to a linear continuous operator from $L^p((X, \mathcal{A}, \mu); \mathbb{R})$ to $L^q((X, \mathcal{A}, \mu); \mathbb{R})$.

The inclusion $L^p \cap L^q \subset L^r$ for $p \leq q$ and $r \in [p, q]$ can be better understood by means of the following general result.

Theorem B.10 (Riesz-Thorin interpolation)**.** *Let $p, q \in [1, \infty]$ with $p \leq q$ and let $T : L^p((X, \mathcal{A}, \mu); \mathbb{R}) \cap L^q((X, \mathcal{A}, \mu); \mathbb{R}) \to L^p((X, \mathcal{A}, \mu); \mathbb{R}) \cap L^q((X, \mathcal{A}, \mu); \mathbb{R})$ be a linear operator which is both of type (p, p) and (q, q). Then T is of type (r, r) for all $r \in [p, q]$.*

We do not present the proof of this theorem, since it follows the lines of the more general Marcinkiewicz theorem below (a standard reference is [88]). In the sequel we shall consider operators T that are not necessarily linear, but rather Q-subadditive for some $Q \geq 0$, meaning that

$$|T(f + g)| \leq Q(|T(f)| + |T(g)|) \qquad \forall f,\, g \in D.$$

For instance, the maximal operator is 1-subadditive. We also say that a space D of real-valued functions is stable under truncations if $f \in D$ implies $f\chi_{\{|f|<k\}} \in D$ for all $k > 0$. We remark that all L^p spaces are stable under truncations.

Definition B.11 (Strong and weak (p,p) operators)**.** *Let $s \in [1,\infty]$, $D \subset L^s((X,\mathcal{A},\mu);\mathbb{R})$ a linear subspace and let $T : D \subset L^s((X,\mathcal{A},\mu);\mathbb{R}) \to L^s((X,\mathcal{A},\mu);\mathbb{R})$, not necessarily linear. We say that T is of strong type (s,s) if $\|T(u)\|_{L^s} \le C\|u\|_{L^s}$ for all $u \in D$, for some constant C independent of u.*
If $s < \infty$, we say that T is of weak type (s,s) if $\|Tu\|_{L^s_w} \le c\|u\|_{L^s}$ for some constant c, namely if

$$\mu\left(\{x : \ |Tu(x)| > \alpha\}\right) \le c^s \frac{\|u\|^s_{L^s}}{\alpha^s} \quad \forall \alpha > 0,\, u \in D.$$

Finally, by convention, T is called of weak type (∞,∞) if it is of strong type (∞,∞).

We can derive an appropriate interpolation theorem even in the case of weak continuity.

Theorem B.12 (Marcinkiewicz Interpolation Theorem)**.** *Assume that p, $q \in [1,\infty]$ with $p < q$, $D \subset L^p((X,\mathcal{A},\mu);\mathbb{R}) \cap L^q((X,\mathcal{A},\mu);\mathbb{R})$ is a linear space stable under truncations and $T : D \to L^p((X,\mathcal{A},\mu);\mathbb{R}) \cap L^q((X,\mathcal{A},\mu);\mathbb{R})$ is Q-subadditive, of weak type (p,p) and of weak type (q,q). Then T is of strong type (r,r) for all $r \in (p,q)$.*

Remark B.13. The most important application of the previous result is perhaps the study of the boundedness of maximal operators (see the next remark). In that case, one typically works with $p = 1$ and $q = \infty$ and thus we shall limit ourselves to prove the theorem with this choice of the exponents.

Proof. We can truncate $f \in D$ as follows:

$$f = g + h, \quad g(x) = f(x)\chi_{\{|f|<\gamma s\}}(x), \quad h(x) = f(x)\chi_{\{|f|\ge\gamma s\}}(x)\,,$$

where γ is an auxiliary parameter to be fixed later. By assumption $g \in D \cap L^\infty((X,\mathcal{A},\mu);\mathbb{R})$ while $h \in D \cap L^1((X,\mathcal{A},\mu);\mathbb{R})$ by linearity of D. Hence

$$|T(f)| \le Q|T(g)| + Q|T(h)| \le QA_\infty \gamma s + Q|T(h)|$$

with A_∞ as the operator norm of T acting from $D \cap L^\infty((X,\mathcal{A},\mu);\mathbb{R})$ into $L^\infty((X,\mathcal{A},\mu);\mathbb{R})$, in the setting of the previous Definition B.11. Choose γ so that $QA_\infty\gamma = 1/2$, therefore

$$\{|T(f)| > s\} \subset \{Q|T(h)| > s/2\}$$

and so

$$\mu\big(\{|T(f)| > s\}\big) \le \mu\left(\left\{|T(h)| > \tfrac{s}{2Q}\right\}\right) \le \left(\tfrac{2A_1Q}{s}\right)\int_X |h|\,d\mu \le \left(\tfrac{2A_1Q}{s}\right)\int_{\{|f|\ge\gamma s\}} |f|\,d\mu,$$

where A_1 is the constant appearing in the weak $(1,1)$ estimate. By integration of the previous inequality, we get for any given $1 < r < \infty$

$$r\int_0^\infty s^{r-1}\mu\left(\{|T(f)| > s\}\right)\,ds \le 2A_1Qr\int_0^\infty \int_{\{|f|\ge\gamma s\}} s^{r-2}|f|\,d\mu\,ds$$

and by means of Lemma B.1 and the Fubini-Tonelli Theorem we finally get

$$\|T(f)\|_{L^r}^r \leq 2A_1 Qr \int_X \left(\int_0^{|f(x)|/\gamma} s^{r-2}\, ds \right) |f(x)|\, d\mu(x) = \frac{2A_1 Qr}{(r-1)\gamma^{r-1}} \|f\|_{L^r}^r$$

so that the conclusion follows. □

Remark B.14 (The limit case $r = 1$)**.** In the limit case $r = 1$ we can argue similarly to find

$$\begin{aligned}\int_1^\infty \mu\left(\{|T(f)| > s\}\right) ds &\leq 2A_1 Q \int_{\{|f|\geq\gamma\}} \left(\int_1^{|f(x)|/\gamma} s^{-1}\, ds \right) |f(x)|\, d\mu(x) \\ &= 2A_1 Q \left(\int_{\{f\geq\gamma\}} |f| \log |f|\, d\mu - \log\gamma \int_{\{f\geq\gamma\}} |f|\, d\mu \right).\end{aligned}$$

Therefore, a slightly better integrability of $|f|$ provides non-trivial information about the integrability of $T(f)$: more precisely, one obtains that $(|T(f)| - 1)^+$ is integrable. Roughly speaking, in $\{f \geq \gamma\}$ the summability of $T(f)$ and the behavior of the entropy are equivalent.

Remark B.15. As a byproduct of the previous result, we have that the maximal operator $\mathcal{M}$ defined in the previous section is of strong type (p, p) for any $p \in (1, \infty]$ (and only of weak type $(1, 1)$). These facts, which have been derived for simplicity in the standard Euclidean setting and dealing with cubic neighborhoods, can be easily generalized, for instance to *pseudo-metric spaces* (i.e. when the distance fulfills only the triangle and symmetry assumptions) endowed with a *doubling measure*, that is a measure μ such that $\mu(B_{2r}(x)) \leq \beta\mu(B_r(x))$ for some constant β not depending on the radius and the center of the ball. Of course, in this case one has to consider the maximal operator $\mathcal{M}_{\mathcal{B}}$ whose definition involves metric balls, namely

$$\mathcal{M}_{\mathcal{B}} f(x) := \sup_{r>0} \fint_{B_r(x)} |f(y)|\, dy.$$

Notice that the constant in the weak $(1, 1)$ bound of the maximal operator does not exceed β^2, since $\mu(B_{3r}(x)) \leq \beta^2\mu(B_r(x))$.

B.3 Lebesgue differentiation theorem

In this section, we want to give a direct proof, based on the $(1, 1)$-weak continuity of the maximal operator $\mathcal{M}$, of the classical Lebesgue differentiation theorem. In fact, to that scope it is convenient to rather employ the modified maximal operator $\mathcal{M}_{\mathcal{B}}$ whose definition involves Euclidean balls rather than cubes.

Theorem B.16. *Let (X, d) be a metric space endowed with a finite doubling measure μ on its Borel σ-algebra $\mathcal{A}$ and let $p \in [1, \infty)$. If $f \in L^p((X, \mathcal{A}, \mu); \mathbb{R})$ then for μ-a.e. $x \in X$ we have that*

$$\lim_{r \downarrow 0} \fint_{B_r(x)} |f(y) - f(x)|^p \, d\mu(y) = 0.$$

Proof. For $t > 0$, let

$$\Lambda_t := \left\{ x \in X : \limsup_{r \downarrow 0} \fint_{B_r(x)} |f(y) - f(x)|^p \, d\mu(y) > t \right\}.$$

The claim is proven by showing that for any $t > 0$ we have $\mu(\Lambda_t) = 0$, since the stated property holds out of the set $\cup_n \Lambda_{1/n}$. Now, we can exploit the metric structure of X in order to approximate f in $L^p((X, \mathcal{A}, \mu); \mathbb{R})$ norm by means of continuous and bounded functions: for any $\varepsilon > 0$ we can write $f = g + h$ with g bounded and continuous and $\|h\|_{L^p}^p \leq t\varepsilon$. Hence, it is enough to prove that for any $t > 0$ we have $\mu(\Lambda'_t) = 0$ where

$$\Lambda'_t := \left\{ x \in X : \limsup_{r \downarrow 0} \fint_{B_r(x)} |h(y) - h(x)|^p \, d\mu(y) > t \right\}.$$

This is easy, because by definition

$$\Lambda'_t \subset \left\{ |h|^p > \frac{t}{2^p} \right\} \cup \left\{ \mathcal{M}(|h|^p) > \frac{t}{2^p} \right\}$$

and, if we consider the corresponding measures, we have (taking Remark B.15 into account)

$$\mu(\Lambda'_t) \leq \frac{2^p}{t} \|h\|_{L^p}^p + \frac{2^p}{t} M \|h\|_{L^p}^p \leq 2^p (1 + M)\varepsilon$$

where M is the constant in the weak $(1, 1)$ bound. Since $\varepsilon > 0$ is arbitrary we conclude $\mu(\Lambda'_t) = 0$, as needed. □

Remark B.17. The Lebesgue differentiation theorem has been stated, as it is customary in most literature, for centered balls. However, one can generalize everything to any metric measure space (X, d, μ) with μ finite and doubling, and a suitable family of sets $\mathcal{E} := \cup_{x \in X} \mathcal{E}_x$, provided there exists a constant $c > 0$ such that

$$\forall\, A \in \mathcal{E}_x \; \exists\, r > 0 \text{ s.t. } A \subset B_r(x) \text{ and } \mu(A) \geq c\mu(B_r(x)) \,. \tag{B.6}$$

Indeed, while such a family may in principle be much larger than the one considered above, it suffices to notice that

$$\fint_A |f(y) - f(x)| \, d\mu(y) \leq \frac{1}{c} \fint_{B_r(x)} |f(y) - f(x)| \, d\mu(y),$$

provided $B_r(x)$ is chosen according to (B.6).

In Euclidean spaces, an important example to which the previous remark applies, in connection with the Calderón-Zygmund theory, is given by

$$\mathcal{E}_x := \{Q \text{ cube}, x \in Q\}.$$

In that case, the corresponding version of the Lebesgue theorem asserts that

$$\lim_{x \in Q,\, |Q| \to 0} \fint_Q |f(y) - f(x)|^p \, dy = 0$$

for $\mathscr{L}^n$-a.e. $x \in \mathbb{R}^n$. Notice that requiring $|Q| \to 0$ (i.e. $\operatorname{diam}(Q) \to 0$) is essential to "factor" continuous functions as in the proof of Theorem B.16.

B.4 Calderón-Zygmund decomposition

We need to introduce another powerful tool, that will be applied to the study of the BMO spaces. Here and below Q will indicate an open cube in $\mathbb{R}^n$ and similarly Q' or Q''.

Theorem B.18. *Let $f \in L^1(Q;\mathbb{R})$, $f \geq 0$ and consider a positive real number α such that $\fint_Q f\,dx \leq \alpha$. Then, there exists a finite or countable family of open cubes $\{Q_i\}_{i \in I}$ with $Q_i \subset Q$ and sides parallel to the ones of Q, such that*

(i) $Q_i \cap Q_j = \emptyset$ *if* $i \neq j$;

(ii) $\alpha < \fint_{Q_i} f\,dx \leq 2^n \alpha$ *for every* $i \in I$;

(iii) $f \leq \alpha$ $\mathscr{L}^n$*-a.e. on* $Q \setminus \cup_i Q_i$.

Remark B.19. The remarkable (and useful) aspect of this decomposition is that the "bad" set $\{f > \alpha\}$ is almost all packed inside a family of cubes, carefully chosen in such a way that still the mean values inside the cubes is of order α. As a consequence of the existence of this decomposition, we have that

$$\alpha \sum_{i \in I} \mathscr{L}^n(Q_i) < \sum_{i \in I} \int_{Q_i} f\,dx \leq \|f\|_1.$$

The proof is based on a so-called *stopping-time* argument.

Proof. Divide $\mathscr{L}^n$-almost all of the cube Q in 2^n open subcubes Q_i by means of n bisections of Q with hyperplanes parallel to the sides of the cube itself. We will call this process a *dyadic decomposition*. Then

(a) if $\fint_{Q_i} f > \alpha$ we stop and do not divide Q_i anymore;

(b) if $\fint_{Q_i} f \leq \alpha$ we iterate the process on Q_i.

At each step we collect the cubes that verify the first condition and put together all such cubes, thus forming a countable family. The first two properties we need to check are obvious by construction: indeed, if Q_i is a chosen cube then its parent cube Q_i^* satisfies $\fint_{Q_i^*} f \leq \alpha$, which gives easily $\fint_{Q_i} f \leq 2^n\alpha$. For the third one note that, modulo a set of points of $\mathscr{L}^n$-negligible measure, if $x \in Q \setminus \cup_i Q_i$, then there exists a sequence of subcubes (Q_j^*) with $x \in \cap_j Q_j^*$ and $\mathscr{L}^n\big(Q_j^*\big) \to 0$, $\fint_{Q_j^*} f\,dx \leq \alpha$. Thanks to the Lebesgue differentiation theorem we get $f(x) \leq \alpha$ for $\mathscr{L}^n$-a.e. $x \in Q \setminus \cup_i Q_i$. $\square$

Remark B.20 (Again in the limit case $p = 1$)**.** Using the Calderón-Zygmund decomposition, for $\alpha \geq \|f\|_{L^1}$ (and assuming, for the sole notational simplicity, to work on the open unit cube $Q = (0,1)^n \subset \mathbb{R}^n$) we can somehow reverse the weak $(1,1)$ estimate:

$$\int_{\{|f|>\alpha\}} |f|\,dx \leq \sum_i \int_{Q_i} |f|\,dx \leq \sum_i 2^n\alpha\mathscr{L}^n(Q_i) \leq 2^n\alpha\mathscr{L}^n\big(\{\mathcal{M}|f| > 2^{-n}\alpha\}\big).$$

The last inequality relies on the fact that the cubes Q_i are contained in $\{\mathcal{M}|f| > 2^{-n}\alpha\}$ by virtue of the definition of the maximal operator.

Using this inequality we can also reverse the implication of Remark B.14, namely assuming with no loss of generality that $f \geq 0$ and $\int_Q f\,dx = 1$:

$$\begin{aligned}\int_{\{f>1\}} f\log f\,dx &= \int_0^\infty \left(\int_{\{\log f>t\}} f\,dx\right) dt = \int_1^\infty \left(\frac{1}{s}\int_{\{f>s\}} f\,dx\right) ds \\ &\leq 4^n \int_1^\infty \mathscr{L}^n\big(\{\mathcal{M}f > s\}\big)\,ds = 4^n \int_Q (\mathcal{M}f - 1)^+\,dx.\end{aligned}$$

C Hausdorff measures

C.1 Basic definitions

Definition C.1. *Consider a subset $S \subset \mathbb{R}^n$, $k \geq 0$ and fix $\delta \in (0, \infty]$. The so-called pre-Hausdorff measures $\mathscr{H}^k_\delta$ are defined by*

$$\mathscr{H}^k_\delta(S) := c_{\mathscr{H}^k} \inf \left\{ \sum_{i=1}^{\infty} [\mathrm{diam}(S_i)]^k : \, S \subset \bigcup_{i=1}^{\infty} S_i, \, \mathrm{diam}(S_i) < \delta \right\} ,$$

while $\mathscr{H}^k$ is defined by

$$\mathscr{H}^k(S) := \lim_{\delta \to 0} \mathscr{H}^k_\delta(S) \ , \tag{C.1}$$

the limit in (C.1) *being well-defined because the map $\delta \mapsto \mathscr{H}^k_\delta(S)$ is non-increasing. The dimensional constant $c_{\mathscr{H}^k} \in (0, \infty)$ will be conveniently fixed in Remark C.3.*

It is easy to check that $\mathscr{H}^k$ is the counting measure when $k = 0$ (provided $c_{\mathscr{H}^0} = 1$) and $\mathscr{H}^k$ is identically equal to 0 when $k > n$.

The *spherical* Hausdorff measure $\mathscr{S}^k$ has a definition analogous to Definition C.1, but only covers made with balls are allowed, so that for all $\delta \in (0, \infty]$ one has

$$\mathscr{H}^k_\delta \leq \mathscr{S}^k_\delta \leq 2^k \mathscr{H}^k_\delta \tag{C.2}$$

and hence $\mathscr{H}^k \leq \mathscr{S}^k \leq 2^k \mathscr{H}^k$.

Remark C.2. Simple but useful properties of Hausdorff measures are:

(i) The Hausdorff measures are translation invariant, that is to say

$$\mathscr{H}^k(S + h) = \mathscr{H}^k(S) \quad \forall \, S \subset \mathbb{R}^n, \ \forall \, h \in \mathbb{R}^n,$$

and (positively) k-homogeneous, that is to say

$$\mathscr{H}^k(\lambda S) = \lambda^k \mathscr{H}^k(S) \quad \forall \, S \subset \mathbb{R}^n, \ \forall \lambda > 0.$$

(ii) The Hausdorff measures are countably subadditive, which means that whenever we have a countable cover of a subset S, namely $S \subset \cup_{i \in I} S_i$, then

$$\mathscr{H}^k(S) \leq \sum_{i \in I} \mathscr{H}^k(S_i).$$

(iii) For every set $A \subset \mathbb{R}^n$ the map $S \mapsto \mathscr{H}^k(A \cap S)$ is σ-additive on Borel sets, which means that whenever we have a countable, pairwise disjoint, cover of a Borel set S by Borel sets S_i, we have

$$\mathscr{H}^k(A \cap S) = \sum_{i \in I} \mathscr{H}^k(A \cap S_i).$$

(iv) Having fixed the subset $S \subset \mathbb{R}^n$ and $\delta > 0$, we have that

$$k > k' \quad \Longrightarrow \quad \mathscr{H}^k_\delta(S) \leq \left(\frac{c_{\mathscr{H}^k}}{c_{\mathscr{H}^{k'}}}\right) \delta^{k-k'} \mathscr{H}^{k'}_\delta(S). \tag{C.3}$$

In particular, looking at (C.3) when $\delta \to 0$, we deduce that

$$\mathscr{H}^{k'}(S) < +\infty \quad \Longrightarrow \quad \mathscr{H}^k(S) = 0$$

or, equivalently,

$$\mathscr{H}^k(S) > 0 \quad \Longrightarrow \quad \mathscr{H}^{k'}(S) = +\infty.$$

Remark C.3. When k is an integer, the choice of $c_{\mathscr{H}^k}$ is meant to be consistent with the usual notion of k-dimensional area: if B is a Borel subset of a k-dimensional plane $\pi \subset \mathbb{R}^n$, $1 \leq k \leq n$, then we would like to have

$$\mathscr{L}^k_\pi(B) = \mathscr{H}^k(B), \tag{C.4}$$

where $\mathscr{L}^k_\pi$ is the k-dimensional Lebesgue measure on $\pi \equiv \mathbb{R}^k$. In that respect, it is useful to remember the isodiametric inequality asserting that, among all sets with prescribed diameter, balls have the largest volume: more precisely, if $\omega_k := \mathscr{L}^k(B_1)$, for every Borel subset $B \subset \mathbb{R}^k$ there holds

$$\mathscr{L}^k(B) \leq \omega_k \left(\frac{\operatorname{diam}(B)}{2}\right)^k. \tag{C.5}$$

Thanks to (C.5), it can be easily proved that equality (C.4) holds if we choose

$$c_{\mathscr{H}^k} = \frac{\omega_k}{2^k}.$$

Recall also that ω_k can be computed by the formula $\omega_k = \pi^{k/2}/\Gamma(1 + k/2)$, where Γ is Euler's function

$$\Gamma(t) := \int_0^\infty s^{t-1} e^{-s}\, ds.$$

More generally, with this choice of the normalization constant, if B is contained in an embedded C^1-manifold M of dimension k in $\mathbb{R}^n$, then

$$\mathscr{H}^k(B) = \sigma_k(B)$$

where σ_k is the classical k-dimensional surface measure defined on Borel subsets of M by decomposition into sufficiently small pieces and local parametrizations.

C.2 An *ad hoc* covering theorem

To the scope of proving the main result in this appendix, Proposition C.7, one would normally appeal to the Besicovitch covering theorem, whose statement is included below for the sake of completeness (see also, for instance, [36]). We present instead a proof based on a more robust covering theorem, valid in general metric spaces.

Theorem C.4 (Besicovitch)**.** *There exists an integer $\xi = \xi(n)$ with the following property: if $A \subset \mathbb{R}^n$ is bounded and $\rho : A \to (0, \infty)$, there exist sets $A_1, \dots, A_{\xi(n)} \subset A$ such that*

(a) for all $j = 1, \dots, \xi$, the balls in the family $\{B_{\rho(x)}(x)\}_{x \in A_j}$ are pairwise disjoint;

(b) the ξ families still cover the set A, that is to say

$$A \subset \bigcup_{j=1}^{\xi} \left(\bigcup_{x \in A_j} B_{\rho(x)}(x) \right).$$

Let us move then to the aforementioned general covering theorem. We first need a definition.

Definition C.5 (Fine cover)**.** *A family $\mathcal{F}$ of closed balls in a metric space (X, d) is a* fine cover *of a set $A \subset X$ if*

$$\inf \left\{ r > 0 : \overline{B}_r(x) \in \mathcal{F} \right\} = 0 \qquad \textit{for all } x \in A.$$

Theorem C.6. *Fix $k \geq 0$, consider a fine cover $\mathcal{F}$ of $A \subset X$, with (X, d) metric space. Then there exists a countable and pairwise disjoint subfamily $\mathcal{F}' = \{\overline{B}_i\}_{i \geq 1} \subset \mathcal{F}$, with $\overline{B}_i := \overline{B}_{r_i}(x_i)$ for all $i \geq 1$, such that at least one of the following conditions holds:*

(i) $\sum_{i=1}^{\infty} (r_i)^k = \infty$,

(ii) $\mathscr{H}^k \left(A \setminus \bigcup_{i=1}^{\infty} \overline{B}_i \right) = 0$.

Proof. The subfamily $\mathcal{F}'$ is chosen inductively, beginning with $\mathcal{F}_0 := \mathcal{F}$. First of all, notice that there exists a closed ball, let us call it $\overline{B}_1$, such that

$$r_1 > \frac{1}{2} \sup \left\{ r : \overline{B}_r(x) \in \mathcal{F}_0 \right\}.$$

Now set

$$\mathcal{F}_1 := \{\overline{B} \in \mathcal{F}_0 : \overline{B} \cap \overline{B}_1 = \emptyset\},$$

and, if $\mathcal{F}_1 \neq \emptyset$, choose a ball $\overline{B}_2 \in \mathcal{F}_1$ such that

$$r_2 > \frac{1}{2} \sup \left\{ r : \overline{B}_r(x) \in \mathcal{F}_1 \right\}.$$

If we try to proceed analogously, removing all balls intersecting previously chosen balls, the only reason why this construction may have to stop is that for some $\ell \in \mathbb{N}$ the family $\mathcal{F}_\ell$ is empty, so we would be getting (because the cover is fine) that the union of the chosen balls covers the whole of A and therefore option (ii) in the statement.
Otherwise, assuming that the construction does not stop, we get a countable family $\mathcal{F}' = \{\overline{B}_i\}_{i\geq 1} = \{\overline{B}_{r_i}(y_i)\}_{i\geq 1}$. We prove that if (i) does not hold, so that in particular $\operatorname{diam}(\overline{B}_i) \to 0$, then we can prove (ii) again. Notice that in the following argument we can also assume, without loss of generality, that for every $i_0 \in \mathbb{N}_*$ the set $A \setminus \bigcup_1^{i_0} \overline{B}_i$ is not empty (which, a priori, is not implied by the condition that $\mathcal{F}_i \neq \emptyset$ for all i), for otherwise we would trivially get (ii) anyway.

Then, fix an index $i_0 \in \mathbb{N}$: for every $x \in A \setminus \bigcup_1^{i_0} \overline{B}_i$ there exists a ball $\overline{B}_{r(x)}(x) \in \mathcal{F}$ such that

$$\overline{B}_{r(x)}(x) \cap \bigcup_{i=1}^{i_0} \overline{B}_i = \emptyset,$$

because $\mathcal{F}$ is a fine cover of A and the complement of $\cup_1^{i_0} \overline{B}_i$ is open in X. On the other hand, we claim that there exists an integer $i(x) > i_0$ such that

$$\overline{B}_{r(x)}(x) \cap \overline{B}_{i(x)} \neq \emptyset. \tag{C.6}$$

Indeed, if

$$\overline{B}_{r(x)}(x) \cap \overline{B}_i = \emptyset \qquad \forall\, i > i_0, \tag{C.7}$$

then

$$r_i \geq \frac{r(x)}{2} \qquad \forall\, i > i_0 \tag{C.8}$$

but $r_i \to 0$, so (C.8) leads to a contradiction. Without loss of generality, we can assume that $i(x)$ is the first index larger than i_0 for which (C.6) holds. Since, by construction, $r_{i(x)} > \frac{1}{2} \sup\{r : \overline{B}_r(y) \in \mathcal{F}_{i(x)-1}\}$ and $\overline{B}_{r(x)}(x) \in \mathcal{F}_{i(x)-1}$ by the minimality of $i(x)$, we also have $r(x) \leq 2r_{i(x)}$.

At that stage, since $\overline{B}_{r(x)}(x) \cap \overline{B}_{i(x)} \neq \emptyset$, the inequality $d(x, y_{i(x)}) \leq r(x) + r_{i(x)} \leq 3r_{i(x)}$ gives

$$\overline{B}_{r(x)}(x) \subset \overline{B}_{5r_{i(x)}}(y_{i(x)})$$

and therefore

$$A \setminus \bigcup_{i=1}^{i_0} \overline{B}_i \subset \bigcup_{i=i_0+1}^{\infty} \overline{B}_{5r_i}(y_i). \tag{C.9}$$

Then, given $\delta > 0$ and chosen i_0 such that $10r_i < \delta$ for every $i > i_0$, (C.9) implies that

$$\mathscr{H}^k_\delta\left(A \setminus \bigcup_{i=1}^{\infty} \overline{B}_i\right) \leq \mathscr{H}^k_\delta\left(A \setminus \bigcup_{i=1}^{i_0} \overline{B}_i\right) \leq \sum_{i=i_0+1}^{\infty} c_{\mathscr{H}^k}(10r_i)^k.$$

We conclude remarking that when $\delta \to 0$, $i_0 \to +\infty$ and thus, by virtue of the summability assumption

$$\mathscr{H}^k\left(A \setminus \bigcup_{i=1}^{\infty} \overline{B}_i\right) \leq \lim_{i_0 \to \infty} c_{\mathscr{H}^k} \sum_{i=i_0+1}^{\infty} (10r_i)^k = 0. \quad \square$$

C.3 A comparison theorem for Hausdorff measures

Proposition C.7. *Consider a locally finite measure $\mu \geq 0$ on the family of Borel sets $\mathscr{B}(\mathbb{R}^n)$ and, fixing $t > 0$, set*

$$E := \left\{x : \limsup_{r \to 0} \frac{\mu(\overline{B}_r(x))}{\omega_k r^k} > t\right\}, \tag{C.10}$$

then E is a Borel set and

$$\mu(E) \geq t\mathscr{S}^k(E).$$

Moreover, if μ vanishes on $\mathscr{H}^k$-finite sets, then $\mathscr{H}^k(E) = 0$.

Using the covering theorem presented in the previous section, we are now able to prove Proposition C.7.

Proof. Given $R > 0$ set $E_R := E \cap B_R(0)$. The argument below will show that the conclusions stated above hold true for the set E_R, whatever value of $R > 0$, and then conclude the proof by simply letting $R \to +\infty$.

Hence, can assume without loss of generality that the set E is bounded. Given $\delta > 0$, we can then fix an open set $A \supset E$ with $\mu(A) < +\infty$ and consider the family

$$\mathcal{F} := \left\{\overline{B}_r(x) : r < \delta/2, \overline{B}_r(x) \subset A, \mu\left(\overline{B}_r(x)\right) > t\omega_k r^k\right\} \tag{C.11}$$

that is a fine cover of E. Applying Theorem C.6, we get a disjoint subfamily $\mathcal{F}' \subset \mathcal{F}$ whose elements are denoted by

$$\overline{B}_i = \overline{B}_{r_i}(x_i).$$

First we exclude possibility (i) of Theorem C.6: by the very definition of the class $\mathcal{F}'$

$$\sum_{i=1}^{\infty} r_i^k < \frac{1}{t\omega_k} \sum_{i=1}^{\infty} \mu(\overline{B}_i) \leq \frac{\mu(A)}{t\omega_k} < +\infty.$$

Since (ii) holds and we can compare $\mathscr{H}^k_\delta$ with $\mathscr{S}^k_\delta$ via (C.2), one has the following inequalities

$$\mathscr{S}^k_\delta(E) \leq \mathscr{S}^k_\delta\left(\bigcup_{i=1}^{\infty} \overline{B}_i\right) \leq \sum_{i=1}^{\infty} \omega_k r_i^k < \frac{1}{t}\sum_{i=1}^{\infty} \mu(\overline{B}_i) \leq \frac{\mu(A)}{t}. \tag{C.12}$$

As $\delta \downarrow 0$ we get $t\mathscr{S}^k(E) \leq \mu(A)$ and the outer regularity of μ gives $t\mathscr{S}^k(E) \leq \mu(E)$.

Finally, the last statement of the proposition can be proved noticing that the inequality (C.12) gives that $\mathscr{S}^k(E)$ is finite; if we assume that μ vanishes on sets with finite k-dimensional measure we obtain that $\mu(E) = 0$; applying once more the same inequality we get $\mathscr{S}^k(E) = 0$. $\square$

D Some tools from convex and nonsmooth analysis

D.1 Subdifferential of a convex function

In this section we briefly recall some classical notions and results from convex and nonsmooth analysis, which will be useful in dealing with uniqueness and regularity results for viscosity solutions to partial differential equations.

Let us consider a convex open subset Ω of $\mathbb{R}^n$ and a convex function $u : \Omega \to \mathbb{R}$. Recall that u is convex if

$$u\big((1-t)x + ty\big) \le (1-t)u(x) + tu(y) \qquad \forall x, y \in \Omega,\ t \in [0,1].$$

If $u \in C^2(\Omega;\mathbb{R})$ this is equivalent to saying that $\nabla^2 u(x) \ge 0$, in the sense of symmetric operators, for all $x \in \Omega$.

Definition D.1 (Subdifferential)**.** *For each $x \in \Omega$, the subdifferential $\partial u(x)$ is the set*

$$\partial u(x) := \{p \in \mathbb{R}^n : u(y) \ge u(x) + \langle p, y - x\rangle \ \ \forall\, y \in \Omega\}.$$

Obviously, $\partial u(x) = \{\nabla u(x)\}$ at any differentiability point.

Remark D.2. According to Definition D.1, it is easy to show that

$$\partial u(x) = \left\{ p \in \mathbb{R}^n : \liminf_{t \to 0^+} \frac{u(x+tv) - u(x)}{t} \ge \langle p, v\rangle \quad \forall v \in \mathbb{R}^n \right\}. \tag{D.1}$$

Indeed, when $p \in \partial u(x)$ we can just take the limit (in fact: the liminf) as one lets $t \to 0^+$ in the inequality

$$\frac{u(x+tv) - u(x)}{t} \ge \langle p, v\rangle.$$

Conversely, let us recall the monotonicity property of difference quotients of a convex function, i.e.

$$\frac{u(x+t'v) - u(x)}{t'} \le \frac{(1 - t'/t)\, u(x) + (t'/t) u(x+tv) - u(x)}{t'} = \frac{u(x+tv) - u(x)}{t}, \tag{D.2}$$

for any $0 < t' < t$. Hence, for every $y \in \Omega$, we have (choosing $t = 1$, $v = y - x$)

$$u(y) - u(x) \ge \frac{u(x+t'v) - u(x)}{t'} \ge \langle p, y - x\rangle + \frac{o(t')}{t'}.$$

The same monotonicity property (D.2) yields that the $\liminf$ in (D.1) is actually a limit.

Remark D.3. The following properties are easy to check:

(i) Convex functions are locally Lipschitz in Ω: to see this, fix a point $x_0 \in \Omega$ and x, $y \in B_r(x_0) \Subset B_R(x_0) \Subset \Omega$. Thanks to the monotonicity of difference quotients seen above (as in equation (D.2)), we can estimate

$$\frac{u(y)-u(x)}{|y-x|} \leq \frac{u(y_R)-u(x)}{|y_R-x|} \leq \frac{\operatorname{osc}(u, \overline{B}_R(x_0))}{R-r},$$

where $y_R \in \partial B_R(x_0)$ is on the halfline emanating from x and containing y. Reversing the roles of x and y we get

$$\operatorname{Lip}(u, B_r(x_0)) \leq \frac{\operatorname{osc}(u, \overline{B}_R(x_0))}{R-r}.$$

This proves the local Lipschitz continuity and we can use this information to replace $\overline{B}_R(x_0)$ by $B_R(x_0)$ in the inequality above. Therefore, invoking Rademacher's Theorem 1.19, we get

$$\operatorname*{ess\,sup}_{B_r(x_0)} |\nabla u| \leq \frac{\operatorname{osc}(u, B_R(x_0))}{R-r},$$

because of (A.8).

(ii) One has that $\partial u(x) \neq \emptyset$ for all $x \in \Omega$ and the graph of the subdifferential, namely the set $\{(x,p) : p \in \partial u(x)\}$ is a closed subset of $\Omega \times \mathbb{R}^n$. These facts follow at once from from the equality $\partial u = \{\nabla u\}$ at differentiability points, and from the definition of subdifferential.

(iii) A convex function u belongs to $C^1(\Omega;\mathbb{R})$ if and only if $\partial u(x)$ is a singleton for every $x \in \Omega$. One implication is straightforward, so let us discuss the other one. To that aim, suppose by contradiction that x_h are differentiability points of u such that $x_h \to x$ and the sequence $(\nabla u(x_h))$ has at least two distinct limit points, say $p_1, p_2 \in \mathbb{R}^n$. But then, passing to the limit in equation (D.1) one obtains that $p_1, p_2 \in \partial u(x)$ and thus $\partial u(x)$ is not a singleton, a contradiction. Hence ∇u has a continuous extension to the whole of Ω and $u \in C^1(\Omega;\mathbb{R})$.

(iv) More generally, given convex functions $f_k : \Omega \to \mathbb{R}$, locally uniformly converging in Ω to f, and $x_k \to x \in \Omega$, any sequence (p_k) with $p_k \in \partial f_k(x_k)$ is bounded (by the local Lipschitz condition) and any limit point p of (p_k) satisfies

$$p \in \partial f(x).$$

Indeed, it suffices to pass to the limit as $k \to \infty$ in the inequalities

$$f_k(y) \geq f_k(x_k) + \langle p_k, y - x_k \rangle \quad \forall y \in \Omega.$$

As a first result of nonsmooth analysis, we state the following theorem.

Theorem D.4 (Nonsmooth mean value theorem)**.** *Consider a convex function $f : \Omega \to \mathbb{R}$ and a couple of points x, $y \in \Omega$. There exist z in the closed segment between x and y and $p \in \partial f(z)$ such that*

$$f(x) - f(y) = \langle p, x - y \rangle.$$

Proof. Choose a positive convolution kernel ρ with support contained in the closed unit ball $\overline{B}_1(0) \subset \mathbb{R}^n$ and consider the smooth functions $f_\varepsilon := f * \rho_\varepsilon$, which are easily seen to be convex in the set Ω_ε defined as per equation (A.5), because

$$\begin{aligned} f_\varepsilon((1-t)x + ty) &= \int_\Omega f((1-t)x + ty - \varepsilon\xi)\rho(\xi)\, d\xi \\ &\leq \int_\Omega \left((1-t)f(x - \varepsilon\xi) + tf(y - \varepsilon\xi)\right)\rho(\xi)\, d\xi \\ &= (1-t)f_\varepsilon(x) + tf_\varepsilon(y). \end{aligned}$$

Thanks to the classical mean value theorem for regular functions, for every $\varepsilon > 0$ there exists $z_\varepsilon = (1 - \theta_\varepsilon)x + \theta_\varepsilon y$, with $\theta_\varepsilon \in (0, 1)$, such that

$$f_\varepsilon(x) - f_\varepsilon(y) = \langle p_\varepsilon, x - y \rangle$$

with $p_\varepsilon = \nabla f_\varepsilon(z_\varepsilon) \in \partial f_\varepsilon(z_\varepsilon)$. Since $(z_\varepsilon, p_\varepsilon)$ are uniformly bounded as $\varepsilon \to 0$ (the claim for (p_ε) following from the fact that the oscillation of f_ε is bounded by the oscillation of f), we can find $\varepsilon_k \to 0$ with $\theta_{\varepsilon_k} \to \theta \in [0, 1]$ and $p_{\varepsilon_k} \to p$. Now, $f_\varepsilon \to f$ locally uniformly on compact subdomains of Ω, hence

$$f(x) - f(y) = \langle p, x - y \rangle$$

and Remark D.3, part (iv) allows to conclude that $p \in \partial f((1-\theta)x + \theta y)$, which completes the proof. □

As an application of the nonsmooth mean value theorem, we can derive a pointwise version of Remark D.3, part (iii). Notice that we will follow a similar idea to achieve second-order differentiability.

Proposition D.5. *If $f : \Omega \to \mathbb{R}$ is convex, then f is differentiable at $x \in \Omega$ if and only if $\partial f(x)$ is a singleton. If this is the case, $\partial f(x) = \{\nabla f(x)\}$.*

Proof. One implication is trivial. For the other one, assume that $\partial f(x) = \{p\}$ and notice that closedness of the graph of ∂f and the local Lipschitz property of f give that $x_h \to x$ and $p_h \in \partial f(x_h)$ imply $p_h \to p$. Then, given any $y \in \Omega$, the nonsmooth mean value theorem ensures the existence of $p_y \in \mathbb{R}^n$ such that

$$f(y) - f(x) = \langle p_y, y - x \rangle.$$

The fact that $p_y \to p$ as we let $y \to x$ implies at once the existence of the limit of the ratio $(f(y) - f(x))/|x - y|$ with

$$\lim_{t \to 0} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|x - y|} = 0$$

This means that $p = \nabla f(x)$, the classical gradient of f at x, as we had to prove.

□

Remark D.6. Recall that a continuous function $f : \Omega \to \mathbb{R}$ is convex if and only if its distributional Hessian $\nabla^2 f$ is non-negative, namely if for every non-negative $\varphi \in C_c^\infty(\Omega; \mathbb{R})$ and every $\xi \in \mathbb{R}^n$ there holds

$$\int_\Omega f(x) \frac{\partial^2 \varphi}{\partial \xi^2}(x)\, dx \geq 0.$$

This result is easily checked by approximation via convolution kernels, because, still in the weak sense (as discussed in Appendix A)

$$\nabla^2 (f * \rho_\varepsilon) = \left(\nabla^2 f\right) * \rho_\varepsilon \ \text{ in } \Omega_\varepsilon.$$

Although we shall not need this fact in the sequel, except in Remark D.16, let us mention, for the sake of completeness, that the positivity condition on every second weak derivative implies that $\nabla^2 f$ derivative is representable by a symmetric matrix-valued measure. To see this, thanks to the classical polarization identity

$$\partial_{\xi+\eta} \partial_{\xi+\eta} f - \partial_{\xi-\eta} \partial_{\xi-\eta} f = 4 \partial_\xi \partial_\eta f \tag{D.3}$$

it suffices to apply the following more general result to the second derivatives $\partial_\xi \partial_\xi f$:

Lemma D.7. *Consider a positive distribution $T \in \mathscr{D}'(\Omega; \mathbb{R})$, i.e. assume*

$$\forall\, \varphi \in C_c^\infty(\Omega; \mathbb{R}), \varphi \geq 0 \quad \Longrightarrow \quad \langle T, \varphi \rangle \geq 0.$$

Then there exists a locally finite non-negative measure μ in Ω such that

$$\langle T, \psi \rangle = \int_\Omega \psi \, d\mu \quad \forall \psi \in C_c^\infty(\Omega; \mathbb{R}).$$

Proof. Fix an open set $\Omega' \Subset \Omega$, define $K := \overline{\Omega'}$ and choose a non-negative cutoff function $\varphi \in C_c^\infty(\Omega; \mathbb{R})$ with $\varphi|_K \equiv 1$. For every test function $\psi \in C_c^\infty(\Omega'; \mathbb{R})$, since $(\|\psi\|_{L^\infty} \varphi - \psi) \geq 0$ and T is a positive distribution, we have

$$\langle T, \psi \rangle \leq \langle T, \|\psi\|_{L^\infty} \varphi \rangle = c \|\psi\|_{L^\infty},$$

where $c = c(\Omega') := \langle T, \varphi \rangle$. Replacing ψ by $-\psi$, the same estimate holds with $|\langle T, \psi \rangle|$ in the left-hand side. By the Riesz-Markov-Kakutani representation theorem (see e.g. Theorem 2.14 in [84]) we obtain the existence of μ. □

Definition D.8 (λ-convexity, uniform convexity, semiconvexity)**.** *Given $\lambda \in \mathbb{R}$, we say that a function $f : \Omega \to \mathbb{R}$ is λ-convex if*

$$\int_\Omega f(x) \frac{\partial^2 \varphi}{\partial \xi^2}(x)\, dx \geq \lambda \int_\Omega \varphi(x)\, dx$$

for every non-negative $\varphi \in C_c^\infty(\Omega; \mathbb{R})$ and for every unit vector $\xi \in \mathbb{R}^n$ (in short $\nabla^2 f \geq \lambda I$). We shall also say that

- *f is* uniformly convex *if $\lambda > 0$;*
- *f is* semiconvex *if $\lambda \leq 0$.*

Notice that, with the notation of Definition D.8, a function f is λ-convex if and only if $f(x) - \lambda |x|^2/2$ is convex.

Analogous concepts can be given in the concave case, namely λ-concavity (i.e. $\nabla^2 f \leq \lambda I$), uniform concavity, semiconcavity. Thus we say that a function $f : \Omega \to \mathbb{R}$ is λ-concave if

$$\int_\Omega f(x) \frac{\partial^2 \varphi}{\partial \xi^2}(x)\, dx \leq \lambda \int_\Omega \varphi(x)\, dx$$

for every non-negative $\varphi \in C_c^\infty(\Omega; \mathbb{R})$ and for every unit vector $\xi \in \mathbb{R}^n$ and

- f is uniformly concave if $\lambda < 0$;
- f is semiconvex if $\lambda \geq 0$.

An important class of semiconcave functions is given by squared distance functions:

Example D.9. Given a closed set $E \subset \mathbb{R}^n$, the square of the distance from E is 2-concave. Indeed,

$$\mathrm{dist}^2(x, E) - |x|^2 = \inf_{y \in E} (x - y)^2 - |x|^2 = \inf_{y \in E} |y|^2 - 2\langle x, y \rangle \tag{D.4}$$

and since the functions $x \mapsto |y|^2 - 2\langle x, y \rangle$ are affine, their infimum over $y \in E$ is concave.

Particularly in the duality theory of convex functions, it is useful to extend the concept of convexity to functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. The concept of subdifferential at points x, where $f(x) < +\infty$, extends immediately and, in the interior of the convex set $\{f < +\infty\}$, we recover all the properties stated before (in particular: local Lipschitz continuity, mean value theorem). Conversely, given $f : \Omega \to \mathbb{R}$ convex with Ω convex, a canonical extension $\tilde{f}$ of f to the whole of $\mathbb{R}^n$ is given by

$$\tilde{f}(x) := \begin{cases} \inf \{\liminf_{h \to \infty} f(x_h) : \ x_h \in \Omega, \ x_h \to x\} & \text{if } x \in \overline{\Omega} \\ +\infty & \text{if } x \in \mathbb{R}^n \setminus \overline{\Omega}. \end{cases}$$

Thereby, one obtains a convex and lower semicontinuous extension of f. For these reasons, we will consider convex and lower-semicontinuous functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Notice that also the notion of λ-convexity extends, by simply requiring that $f(x) - \lambda|x|^2/2$ is convex.

Proposition D.10. *Given a convex lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, its subdifferential ∂f satisfies for all x, $y \in \{f < +\infty\}$ the monotonicity property*

$$\langle p - q, x - y\rangle \geq 0 \qquad \forall p \in \partial f(x),\ \forall q \in \partial f(y).$$

Proof. It is sufficient to sum the inequalities satisfied, respectively, by p and q, i.e.

$$\begin{aligned} f(y) - f(x) &\geq \langle p, y - x\rangle \\ f(x) - f(y) &\geq \langle q, x - y\rangle. \end{aligned}$$

□

Remark D.11 (Inverse of the subdifferential)**.** We observe the following:

(i) If $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is λ-convex, the argument presented for Proposition D.10 proves that for every $p \in \partial f(x)$ and every $q \in \partial f(y)$, we have

$$\langle p - q, x - y\rangle \geq \lambda|x - y|^2. \tag{D.5}$$

(ii) If $\lambda > 0$, for every $p \in \mathbb{R}^n$ no more than one $x \in \{f < +\infty\}$ can satisfy $p \in \partial f(x)$, because, through (D.5), we get

$$p \in \partial f(x) \cap \partial f(y) \implies 0 = \langle p - p, x - y\rangle \geq \lambda|x - y|^2 \implies x = y.$$

In particular, setting

$$L := \bigcup_{f(x)<+\infty} \partial f(x),$$

there exists a single-valued and onto map

$$(\partial f)^{-1} : L \to \{x :\ f(x) < +\infty,\ \partial f(x) \neq \emptyset\}$$

such that $p \in \partial f((\partial f)^{-1}(p))$. In addition, $L = \mathbb{R}^n$: given p, to find x such that $p \in \partial f(x)$ it suffices to minimize $y \mapsto f(y) - \langle p, y\rangle$ and to take x as the (unique) minimum point.

(iii) Moreover, $(\partial f)^{-1}$ is a Lipschitz map: rewriting equation (D.5) for $(\partial f)^{-1}$ we get

$$\begin{aligned} \lambda|(\partial f)^{-1}(p) - (\partial f)^{-1}(q)|^2 &\leq \langle p - q, (\partial f)^{-1}(p) - (\partial f)^{-1}(q)\rangle \\ &\leq |p - q||(\partial f)^{-1}(p) - (\partial f)^{-1}(q)|, \end{aligned}$$

thus $\mathrm{Lip}((\partial f)^{-1}) \leq 1/\lambda$.

The conjugate of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, not identically equal to $+\infty$, is defined as

$$f^*(x^*) := \sup_{x \in \mathbb{R}^n} \langle x^*, x \rangle - f(x).$$

We immediately point out that f^* is convex and lower semicontinuous, because it is the supremum of a family of affine functions. The assumption that $f(x) < +\infty$ for at least one x ensures that in fact $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Moreover, notice that the following equivalent characterization holds: f^* is the smallest function satisfying

$$\langle x, y \rangle \leq f(x) + f^*(y) \qquad \forall x,\, y \in \mathbb{R}^n. \tag{D.6}$$

A similar "variational" characterization of the subdifferential is that $x^* \in \partial f(x)$ if and only if $z \mapsto \langle x^*, z \rangle - f(z)$ attains its maximum at $z = x$:

$$x^* \in \partial f(x) \qquad \Longleftrightarrow \qquad f^*(x^*) = \langle x^*, x \rangle - f(x). \tag{D.7}$$

We proceed with the following general result concerning the representation of convex lower semicontinuous functions:

Theorem D.12. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex lower semicontinuous function, not identically equal to $+\infty$. Then f^* is not identically equal to $+\infty$ and $(f^*)^* = f$. In particular, any convex lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ not identically equal to $+\infty$ is representable as g^* for some $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ not identically equal to $+\infty$.*

Proof. If $f(x_0) < +\infty$ we can use Hahn-Banach theorem in $\mathbb{R}^{n+1}$ (with a small open ball centered at $\{(x_0, f(x_0) - 1)\}$ and the epigraph of f, which is a convex set) to find an affine function $\ell(x) = \langle p, x \rangle + c$ such that $\ell \leq f$. This yields $f^*(p) < +\infty$, so that $(f^*)^*$ makes sense. Now, the variational characterization of the conjugate function based on (D.6) gives that $(f^*)^* \leq f$. On the other hand, the operator $g \mapsto (g^*)^*$ is order-preserving and coincides, as it is easily seen, with the identity on affine functions. Since convex lower semicontinuous functions are suprema of affine functions (again as an application of the Hahn-Banach theorem), these two facts combine together to give $(f^*)^* \geq f$ on convex lower semicontinuous functions, thereby completing the proof. $\square$

Based on this fact, it is easily seen that (D.7) gives the equivalence

$$x \in \partial f^*(x^*) \qquad \Longleftrightarrow \qquad x^* \in \partial f(x). \tag{D.8}$$

In particular, in the case when f is λ-convex for some $\lambda > 0$, from the quadratic growth of f we obtain that f^* is finite and that $\partial f^* = (\partial f)^{-1}$ is single-valued and Lipschitz, therefore $f^* \in C^{1,1}(\mathbb{R}^n; \mathbb{R})$.

D.2 Convex functions and Measure Theory

We shall now recall some classical results in Measure Theory, in order to have the necessary tools to prove Aleksandrov's theorem D.15, on the second differentiability of convex functions as encoded in the following definition.

Definition D.13 (Pointwise second-order differentiability)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $x \in \Omega$. A function $f : \Omega \to \mathbb{R}$ is pointwise second-order differentiable at x if there exist $p \in \mathbb{R}^n$ and $S \in \mathrm{Sym}^{n\times n}$ such that*

$$f(y) = f(x) + \langle p, y - x\rangle + \frac{1}{2}\langle S(y - x), y - x\rangle + o(|y - x|^2).$$

Notice that pointwise second-order differentiability implies first-order differentiability, and that $p = \nabla f(x)$ (here understood in the pointwise sense). Furthermore, the requirement that $S \in \mathrm{Sym}^{n\times n}$ in Definition D.13 is actually unnecessary, since one can replace, in the expansion above, any matrix S with its symmetric part (due to the fact that the anti-symmetric part determines a null bilinear form).

Let us first state the following classical result, whose proof can be found, for instance, in [36] and [37].

Theorem D.14 (Area formula)**.** *Consider a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$ and a Borel set $A \subset \mathbb{R}^n$. Then the function*

$$N(y, A) := \mathrm{card}\left(f^{-1}(y) \cap A\right)$$

is $\mathscr{L}^n$-measurable[12] *and*

$$\int_A |\det \nabla f(x)|\, dx = \int_{\mathbb{R}^n} N(y, A)\, dy \geq \mathscr{L}^n(f(A)).$$

Also, let us remind the reader that, thanks to Rademacher's Theorem 1.19, we can, with a slight abuse of notation, denote by ∇f both the weak gradient and its pointwise representative, at least for locally Lipschitz functions. We are now ready to prove the main result of this section.

Theorem D.15 (Aleksandrov)**.** *Any convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is $\mathscr{L}^n$-a.e. pointwise second-order differentiable in the interior of $\{f < +\infty\}$.*

Proof. The proof is based on the inverse function $\Psi = (\partial f)^{-1}$, introduced in Remark D.11. Obviously, there is no loss of generality in supposing that f is λ-convex for some $\lambda > 0$ since we can always replace f with $f + g$ where g is smooth, real-valued, and λ-convex.

[12] In particular, notice that $f(A) = \{N > 0\}$.

We briefly recall, from Remark D.11, that ∂f associates to each $x \in \mathbb{R}^n$ the subdifferential set, on the contrary Ψ is a single-valued map which associates to each $p \in \mathbb{R}^n$ the point x such that $p \in \partial f(x)$. Let us define the set of "bad" points

$$\Sigma := \{p : \nexists \nabla\Psi(p) \ \text{ or } \exists \ \nabla\Psi(p) \text{ and } \det \nabla\Psi(p) = 0\}.$$

Since Ψ is Lipschitz, Rademacher's Theorem 1.19 and the area formula D.14 give

$$\mathscr{L}^n\left(\Psi(\Sigma)\right) \leq \int_\Sigma |\det \nabla\Psi| \, dp = 0.$$

We shall prove that the stated differentiability property holds at all points $x \notin \Psi(\Sigma)$. Let us write $x = \Psi(p)$ with $p \notin \Sigma$, so that $\nabla f(x) = p$, there exists the derivative $\nabla\Psi(p)$ and, since it is invertible, we can set

$$S(x) := \left(\nabla\Psi(p)\right)^{-1} \ .$$

If $y = \Psi(q)$, we get

$$\begin{aligned} S(x)^{-1}\left(q - p - S(x)(y-x)\right) &= -\left(y - x - \nabla\Psi(p)(q-p)\right) \\ &= -\left(\Psi(q) - \Psi(p) - \nabla\Psi(p)(q-p)\right) \\ &= o(|q-p|) = o(|y-x|) \end{aligned}$$

where the last equality relies on the differentiability of the map Ψ at p and on the fact that $\det(\nabla\Psi(p)) \neq 0$. Therefore

$$\lim_{\substack{y \to x \\ q \in \partial f(y)}} \frac{|q - \nabla f(x) - S(x)(y-x)|}{|y-x|} = 0. \tag{D.9}$$

We claim that such equation (D.9), together with Theorem D.4 (the nonsmooth mean value theorem), implies the second-order differentiability of f at the point x. Indeed, let

$$\tilde{f}(y) := f(y) - f(x) - \langle \nabla f(x), (y-x)\rangle - \frac{1}{2}\langle S(x)(y-x), (y-x)\rangle.$$

Since

$$\partial\tilde{f}(y) = \partial f(y) - \nabla f(x) - S(x)(y-x)$$

we can rephrase (D.9) as $\lim_{q \in \partial\tilde{f}(x),\, y \to x} |q|/|y-x| = 0$. Now, let $\theta \in [0,1]$ and $q \in \partial\tilde{f}((1-\theta)y + \theta x)$ be such that $\tilde{f}(y) = \langle q, y-x\rangle$ (since $\tilde{f}(x) = 0$): we immediately find

$$\tilde{f}(y) = \langle q, y-x\rangle = o(|y-x|^2).$$

Hence, going back to the very definition of $\tilde{f}$, the statement follows. □

Remark D.16 (Characterization of S)**.** A blow-up analysis, analogous to the one performed in the proof of Rademacher's theorem, shows that the matrix S in Aleksandrov's theorem is the density of the measure $\nabla^2 f$ with respect to $\mathscr{L}^n$, see [3] for details.

References

[1] E. Acerbi and N. Fusco, *Semicontinuity problems in the calculus of variations*, Arch. Rational Mech. Anal., 86 (1984), pp. 125–145.

[2] R. A. Adams, *Sobolev spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[3] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbf{R}^n$, Math. Z., 230 (1999), pp. 259–316.

[4] A. D. Aleksandrov, *Uniqueness conditions and bounds for the solution of the Dirichlet problem*, Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom., 18 (1963), pp. 5–29.

[5] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[6] L. Beck, *Elliptic regularity theory*, vol. 19 of Lecture Notes of the Unione Matematica Italiana, Springer, Cham; Unione Matematica Italiana, Bologna, 2016. A first course.

[7] J.-M. Bony, *Principe du maximum dans les espaces de Sobolev*, C. R. Acad. Sci. Paris Sér. A-B, 265 (1967), pp. A333–A336.

[8] H. Brezis, *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].

[9] Y. D. Burago and V. A. Zalgaller, *Geometric inequalities*, vol. 285 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1988. Translated from the Russian by A. B. Sosinskiĭ, Springer Series in Soviet Mathematics.

[10] X. Cabré, *On the Alexandroff-Bakelman-Pucci estimate and the reversed Hölder inequality for solutions of elliptic and parabolic equations*, Comm. Pure Appl. Math., 48 (1995), pp. 539–570.

[11] R. Caccioppoli, *Limitazioni integrali per le soluzioni di un'equazione lineare ellitica a derivate parziali*, Giorn. Mat. Battaglini (4), 4(80) (1951), pp. 186–212.

[12] L. Caffarelli, L. Nirenberg, and J. Spruck, *The Dirichlet problem for nonlinear second-order elliptic equations. I. Monge-Ampère equation*, Comm. Pure Appl. Math., 37 (1984), pp. 369–402.

[13] L. Caffarelli, L. Nirenberg, and J. Spruck, *Nonlinear second-order elliptic equations. V. The Dirichlet problem for Weingarten hypersurfaces*, Comm. Pure Appl. Math., 41 (1988), pp. 47–70.

[14] L. A. Caffarelli, *Interior a priori estimates for solutions of fully nonlinear equations*, Ann. of Math. (2), 130 (1989), pp. 189–213.

[15] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, vol. 43 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 1995.

[16] ——, *Fully nonlinear elliptic equations*, vol. 43 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 1995.

[17] S. Campanato, *Proprietà di hölderianità di alcune classi di funzioni*, Ann. Scuola Norm. Sup. Pisa (3), 17 (1963), pp. 175–188.

[18] C. Carathéodory, *über den variabilitätsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen*, Mathematische Annalen, 64 (1907), pp. 95–115.

[19] L. Carbone and R. De Arcangelis, *Unbounded functionals in the calculus of variations*, vol. 125 of Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, Chapman & Hall/CRC, Boca Raton, FL, 2002. Representation, relaxation, and homogenization.

[20] F. Clarke, *Functional analysis, calculus of variations and optimal control*, vol. 264 of Graduate Texts in Mathematics, Springer, London, 2013.

[21] M. G. Crandall, L. C. Evans, and P.-L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 282 (1984), pp. 487–502.

[22] M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.), 27 (1992), pp. 1–67.

[23] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), pp. 1–42.

[24] B. Dacorogna, *Direct methods in the calculus of variations*, vol. 78 of Applied Mathematical Sciences, Springer, New York, second ed., 2008.

[25] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 3*, Springer-Verlag, Berlin, 1990. Spectral theory and applications, With the collaboration of Michel Artola and Michel Cessenat, Translated from the French by John C. Amson.

[26] E. De Giorgi, *Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari*, Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3), 3 (1957), pp. 25–43.

[27] ——, *Complementi alla teoria della misura $(n-1)$-dimensionale in uno spazio n-dimensionale*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61, Editrice Tecnico Scientifica, Pisa, 1961.

[28] ——, *Frontiere orientate di misura minima*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61, Editrice Tecnico Scientifica, Pisa, 1961.

[29] ——, *Un esempio di estremali discontinue per un problema variazionale di tipo ellittico*, Boll. Un. Mat. Ital. (4), 1 (1968), pp. 135–137.

[30] C. De Lellis, *Rectifiable sets, densities and tangent measures*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008.

[31] G. De Philippis, *Regularity of Optimal Transport Maps and Applications*, vol. 17 of Theses (Scuola Normale Superiore), Edizioni della Normale, 1 ed., 2013.

[32] G. Dolzmann, B. Kirchheim, and J. Kristensen, *Conditions for equality of hulls in the calculus of variations*, Arch. Ration. Mech. Anal., 154 (2000), pp. 93–100.

[33] N. Dunford and B. J. Pettis, *Linear operations on summable functions*, Trans. Amer. Math. Soc., 47 (1940), pp. 323–392.

[34] L. C. Evans, *Quasiconvexity and partial regularity in the calculus of variations*, Arch. Rational Mech. Anal., 95 (1986), pp. 227–252.

[35] ——, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.

[36] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[37] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[38] A. Figalli, *The Monge-Ampère Equation and Its Applications*, Zurich Lectures in Advanced Mathematics, European Mathematical Society Publishing House, 2017.

[39] E. Gagliardo, *Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili*, Rend. Sem. Mat. Univ. Padova, 27 (1957), pp. 284–305.

[40] F. W. Gehring, *The L^p-integrability of the partial derivatives of a quasiconformal mapping*, Acta Math., 130 (1973), pp. 265–277.

[41] M. Giaquinta, *Introduction to Regularity Theory for Nonlinear Elliptic Systems*, Birkhauser Verlag, 1993.

[42] M. Giaquinta and E. Giusti, *On the regularity of the minima of variational integrals*, Acta Math., 148 (1982), pp. 31–46.

[43] ——, *The singular set of the minima of certain quadratic functionals*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 11 (1984), pp. 45–55.

[44] M. Giaquinta and L. Martinazzi, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, vol. 11 of Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], Edizioni della Normale, Pisa, second ed., 2012.

[45] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, vol. 224 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, second ed., 1983.

[46] E. Giusti, *Minimal surfaces and functions of bounded variation*, vol. 80 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 1984.

[47] ——, *Direct methods in the calculus of variations*, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[48] L. Gårding, *Dirichlet's problem for linear elliptic partial differential equations*, Math. Scand., 1 (1953), pp. 55–72.

[49] Q. Han and F. Lin, *Elliptic partial differential equations*, vol. 1 of Courant Lecture Notes in Mathematics, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, second ed., 2011.

[50] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, at the University Press, 1952. 2d ed.

[51] H. Helmholtz, *Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen*, J. Reine Angew. Math., 55 (1858), pp. 25–55.

[52] D. Hilbert, *Mathematical problems*, Bull. Amer. Math. Soc., 8 (1902), pp. 437–479.

[53] E. Hopf, *Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung*, Math. Z., 34 (1932), pp. 194–233.

[54] H. Ishii, *Perron's method for Hamilton-Jacobi equations*, Duke Math. J., 55 (1987), pp. 369–384.

[55] ——, *On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs*, Comm. Pure Appl. Math., 42 (1989), pp. 15–45.

[56] ——, *Perron's method for monotone systems of second-order elliptic partial differential equations*, Differential Integral Equations, 5 (1992), pp. 1–24.

[57] H. Ishii and P.-L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations, 83 (1990), pp. 26–78.

[58] R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rational Mech. Anal., 101 (1988), pp. 1–27.

[59] ——, *Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations*, Indiana Univ. Math. J., 38 (1989), pp. 629–667.

[60] R. Jensen, P.-L. Lions, and P. E. Souganidis, *A uniqueness result for viscosity solutions of second order fully nonlinear partial differential equations*, Proc. Amer. Math. Soc., 102 (1988), pp. 975–978.

[61] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math., 14 (1961), pp. 415–426.

[62] D. Kinderlehrer and P. Pedregal, *Gradient Young measures generated by sequences in Sobolev spaces*, J. Geom. Anal., 4 (1994), pp. 59–90.

[63] V. A. Kondratiev and O. A. Oleinik, *On Korn's inequalities*, C. R. Acad. Sci. Paris Sér. I Math., 308 (1989), pp. 483–487.

[64] J. Kristensen and G. Mingione, *The singular set of Lipschitzian minima of multiple integrals*, Arch. Ration. Mech. Anal., 184 (2007), pp. 341–369.

[65] N. V. Krylov and M. Safonov, *An estimate for the probability of a diffusion process hitting a set of positive measure*, Doklady Akademii Nauk SSSR, 245 (1979), pp. 18–20.

[66] ——, *A property of the solutions of parabolic equations with measurable coefficients*, Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya, 44 (1980), pp. 161–175.

[67] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations*, Translated from the Russian by Scripta Technica, Inc. Translation editor: Leon Ehrenpreis, Academic Press, New York-London, 1968.

[68] P. D. Lax and A. N. Milgram, *Parabolic equations*, in Contributions to the theory of partial differential equations, Annals of Mathematics Studies, no. 33, Princeton University Press, Princeton, N. J., 1954, pp. 167–190.

[69] G. Leoni, *A first course in Sobolev spaces*, vol. 181 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2017.

[70] J. Leray, *Majoration des dérivées secondes des solutions d'un problème de Dirichlet*, J. Math. Pures Appl., 17 (1938), pp. 89–104.

[71] P.-L. Lions, *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. II. Viscosity solutions and uniqueness*, Comm. Partial Differential Equations, 8 (1983), pp. 1229–1276.

[72] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, vol. 44 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[73] N. G. Meyers and J. Serrin, $H = W$, Proc. Nat. Acad. Sci. U.S.A., 51 (1964), pp. 1055–1056.

[74] G. Mingione, *Nonlinear Calderón-Zygmund theory*, Zürich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich. In preparation.

[75] C. B. Morrey, Jr., *Quasi-convexity and the lower semicontinuity of multiple integrals*, Pacific J. Math., 2 (1952), pp. 25–53.

[76] J. Moser, *A sharp form of an inequality by N. Trudinger*, Indiana Univ. Math. J., 20 (1970/71), pp. 1077–1092.

[77] S. Müller and V. Šverák, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. of Math. (2), 157 (2003), pp. 715–742.

[78] L. Nirenberg, *Estimates and existence of solutions of elliptic equations*, Comm. Pure Appl. Math., 9 (1956), pp. 509–529.

[79] P. Pedregal, *Parametrized measures and variational principles*, vol. 30 of Progress in Nonlinear Differential Equations and their Applications, Birkhäuser Verlag, Basel, 1997.

[80] O. Perron, *Eine neue Behandlung der ersten Randwertaufgabe für* $\Delta u = 0$, Math. Z., 18 (1923), pp. 42–54.

[81] C. Pucci, *Limitazioni per soluzioni di equazioni ellittiche*, Ann. Mat. Pura Appl. (4), 74 (1966), pp. 15–30.

[82] J. Rauch, *Partial differential equations*, vol. 128 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1991.

[83] M. Renardy and R. C. Rogers, *An introduction to partial differential equations*, vol. 13 of Texts in Applied Mathematics, Springer-Verlag, New York, second ed., 2004.

[84] W. Rudin, *Real and complex analysis*, McGraw-Hill Book Co., New York, third ed., 1987.

[85] J. Schauder, *Über lineare elliptische Differentialgleichungen zweiter Ordnung*, Math. Z., 38 (1934), pp. 257–282.

[86] L. Schwartz, *Théorie des distributions*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Nouvelle édition, entiérement corrigée, refondue et augmentée, Hermann, Paris, 1966.

[87] G. Stampacchia, $\mathcal{L}^{(p,\lambda)}$*-spaces and interpolation*, Comm. Pure Appl. Math., 17 (1964), pp. 293–306.

[88] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

[89] E. Steinitz, *Bedingt konvergente reihen und konvexe systeme*, Journal für die reine und angewandte Mathematik (Crelle's Journal), 1913 (1913), pp. 128–176.

[90] F. J. Terpstra, *Die Darstellung biquadratischer Formen als Summen von Quadraten mit Anwendung auf die Variationsrechnung*, Math. Ann., 116 (1939), pp. 166–180.

[91] L. Tonelli, *La semicontinuità nel calcolo delle variazioni*, Rend. Circ. Mat. Palermo, 44 (1920), pp. 167–249.

[92] N. S. Trudinger, *On imbeddings into Orlicz spaces and some applications*, J. Math. Mech., 17 (1967), pp. 473–483.

[93] ——, *Comparison principles and pointwise estimates for viscosity solutions of nonlinear elliptic equations*, Rev. Mat. Iberoamericana, 4 (1988), pp. 453–468.

[94] ——, *On regularity and existence of viscosity solutions of nonlinear second order, elliptic equations*, in Partial differential equations and the calculus of variations, Vol. II, vol. 2 of Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, Boston, MA, 1989, pp. 939–957.

[95] V. r. Šverák, *Rank-one convexity does not imply quasiconvexity*, Proc. Roy. Soc. Edinburgh Sect. A, 120 (1992), pp. 185–189.

[96] K.-O. Widman, *Hölder continuity of solutions of elliptic systems*, Manuscripta Math., 5 (1971), pp. 299–308.

[97] K. Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

[98] W. P. Ziemer, *Weakly differentiable functions*, vol. 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

Index