Lecture Notes on Elliptic Partial Differential Equations

Luigi Ambrosio *

Contents

*PhD course given in 2009-2010 and then in 2012-2013, 2014-2015, lectures typed by A.Carlotto and A.Massaccesi

Preface

Prerequisites: basic knowledge of Functional Analysis and Measure Theory, preferably also a basic knowledge of Sobolev spaces of functions of one independent variable.

$B_r(x)$	Ball with center x and radius r (also $B_r = B_r(0)$, $B = B_1$)
$A \subset B$	Inclusion in the weak sense
$A \Subset B$	$\overline{A} \subset B$ (typically used for pairs of open sets)
$\mathscr{L}^n$	Lebesgue measure in $\mathbb{R}^n$
$C^k(\Omega)$	Functions continuously k-differentiable in Ω
$L^p(\Omega)$	Lebesgue L^p space
$\partial_i u,\ \partial_{x_i} u,\ \nabla_i u,\ \frac{\partial u}{\partial x_i}$	i-th partial derivative (weak or classical)
∇u	Gradient of u
$\fint_\Omega f\, d\mu$	Mean integral value, namely $\int_\Omega f\, d\mu / \mu(\Omega)$

1 Some basic facts concerning Sobolev spaces

In this book, we will make constant use of Sobolev spaces. Here, we will just summarize the basic facts needed in the sequel, referring for instance to [4] or [1] for a more detailed treatment of this topic.
Actually, it is possible to define Sobolev spaces in (at least) two different ways, whose (partial) equivalence is discussed below.

Definition 1.1. *Let $\Omega \subset \mathbb{R}^n$ be an open and bounded domain and fix an exponent p with $1 \le p < \infty$. We can consider the class of regular functions $C^1\big(\overline{\Omega}\big)$ (i.e. the subset of $C^1(\Omega)$ consisting of functions u such that both u and ∇u admit a continuous extension on $\partial\Omega$) endowed with the norm*

$$\|u\|_{W^{1,p}} =^p \sqrt{\|u\|_{L^p}^p + \|\nabla u\|_{L^p}^p}. \tag{1.1}$$

We define the space $H^{1,p}(\Omega)$ to be the completion with respect to the $W^{1,p}$ norm of $C^1(\overline{\Omega})$.

For unbounded domains, including the whole space $\mathbb{R}^n$, the definition is similar and based on the completion of

$$\left\{u \in C^1(\overline{\Omega}) :\ u \in L^p(\Omega),\ |\nabla u| \in L^p(\Omega)\right\}.$$

Notice that $H^{1,p}(\Omega) \subset L^p(\Omega)$.

On the other hand, we can adopt a different viewpoint, inspired by the theory of distributions.

Definition 1.2. *Let $\Omega \subset \mathbb{R}^n$ be an open domain and consider the space $C_c^\infty(\Omega)$ whose elements will be called test functions. For $i = 1, \ldots, n$, we say that $u \in L^1_{\mathrm{loc}}(\Omega)$ has i-th derivative in weak sense equal $g_i \in L^1_{\mathrm{loc}}(\Omega)$ if*

$$\int_\Omega u \partial_i \varphi \, dx = -\int_\Omega \varphi g_i \, dx \qquad \forall \varphi \in C_c^\infty(\Omega). \tag{1.2}$$

Whenever such $g_1, \ldots, g_n$ exist, we say that u is differentiable in weak sense and we write $g_i = \partial_i u$ and

$$\nabla u = \big(\partial_1 u, \ldots, \partial_n u\big).$$

For $1 \leq p \leq \infty$ we define the space $W^{1,p}(\Omega)$ as the subset of $L^p(\Omega)$ whose elements u are weakly differentiable with corresponding derivatives $\partial_i u$ also belonging to $L^p(\Omega)$.

It is clear that if g_i exists, it must be uniquely determined up to Lebesgue negligible sets, since $h \in L^1_{\mathrm{loc}}(\Omega)$ and

$$\int_\Omega h\varphi\, dx = 0 \qquad \forall \varphi \in C_c^\infty(\Omega)$$

implies $h = 0$. This implication can be easily proved by approximation, showing that the property above is stable under convolution, namely $h_\varepsilon = h * \rho_\varepsilon$ satisfies

$$\int_{\Omega_\varepsilon} h_\varepsilon \varphi\, dx = \int_\Omega h\varphi * \rho_\varepsilon = 0 \qquad \forall \varphi \in C_c^\infty(\Omega_\varepsilon),$$

where Ω_ε is the (slightly) smaller domain

$$\Omega_\varepsilon := \{x \in \Omega : \ \mathrm{dist}(x, \partial\Omega) > \varepsilon\}, \tag{1.3}$$

$\rho_\varepsilon(x) = \varepsilon^{-n}\rho(x/\varepsilon)$ with ρ smooth, even and compactly supported in the unit ball and we used the simmetry property (a consequence of Fubini's theorem).

$$\int (a * \rho_\varepsilon) b\, dx = \int a(b * \rho_\varepsilon)\, dx. \tag{1.4}$$

Obviously, classical derivatives are weak derivatives and the notation $\partial_i u$ (or, equivalently, $\partial_{x_i} u$, $\nabla_i u$ or even $\frac{\partial u}{\partial x_i}$) is justified.

Another classical way to relate weak and strong derivatives is via convolution: namely if u has weak i-th derivative equal to g, then

$$\partial_i(u * \rho_\varepsilon) = g * \rho_\varepsilon \qquad \text{in } \Omega_\varepsilon, \text{ in the classical sense.} \tag{1.5}$$

Knowing the identity (1.5) for smooth functions, its validity can be easily extended considering both sides as weak derivatives and using (1.4):

$$\int_\Omega (u * \rho_\varepsilon)\partial_i \varphi\, dx = \int_\Omega u(\partial_i \varphi) * \rho_\varepsilon\, dx = \int_\Omega u \partial_i(\varphi * \rho_\varepsilon)\, dx = -\int_\Omega g\varphi * \rho_\varepsilon\, dx = -\int_\Omega g * \rho_\varepsilon \varphi\, dx$$

for all $\varphi \in C_c^\infty(\Omega_\varepsilon)$. Now, the smoothness of $u * \rho_\varepsilon$ tells us that the derivative in the left hand side of (1.5) is (equivalent to) a classical one.

Another consequence of (1.5) is:

Theorem 1.3 (Constancy theorem). *If $u \in L^1_{\rm loc}(\Omega)$ 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 $u = c$ $\mathscr{L}^n$-a.e. in B. In particular, if Ω is connected, $u = c$ $\mathscr{L}^n$-a.e. in Ω for some $c \in \mathbb{R}$.*

Proof. Again we argue by approximation, using the fact that (1.5) ensures that the function $u * \rho_\varepsilon$ are locally constant in Ω_ε and taking the $L^1_{\rm loc}$ limit as $\varepsilon \to 0$. $\square$

Notice also that Definition 1.2 covers the case $p = \infty$, while it is not immediately clear how to adapt Definition 1.1 to cover this case: usually H Sobolev spaces are defined for $p < \infty$ only.

In the next proposition we consider the relation of $W^{1,\infty}$ with Lipschitz functions. We omit, for brevity, the simple proof, based once more on convolutions.

Proposition 1.4 (Lipschitz versus $W^{1,\infty}$ functions). *If $\Omega \subset \mathbb{R}^n$ is open, then* $\mathrm{Lip}(\Omega) \subset W^{1,\infty}(\Omega)$ *and*

$$\|Du\|_{L^\infty(\Omega)} \le \mathrm{Lip}(u, \Omega). \tag{1.6}$$

In addition, if Ω is convex then $\mathrm{Lip}(\Omega) = W^{1,\infty}(\Omega)$ *and equality holds in* (1.6).

Since $H^{1,p}(\Omega)$ is defined by means of approximation by regular functions, for which (1.2) is just the elementary "integration by parts formula", it is clear that $H^{1,p}(\Omega) \subset W^{1,p}(\Omega)$; in addition, the same argument shows that the weak derivative of $u \in H^{1,p}(\Omega)$, in the sense of W Sobolev spaces, is precisely the strong $L^p(\Omega, \mathbb{R}^n)$ limit of ∇u_h, where $u_h \in C^1(\overline{\Omega})$ are strongly convergent to u. This allows to show by approximation some basic calculus rules in H Sobolev spaces for weak derivatives, as the chain rule

$$\nabla(\phi \circ u) = \phi'(u)\nabla u \qquad \phi \in C^1(\mathbb{R}) \text{ Lipschitz with } \phi(0) = 0,\ u \in H^{1,p}(\Omega)\,, \tag{1.7}$$

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 equibounded $C^1(\overline{\Omega})$ functions) the Leibniz rule

$$\nabla(uv) = u\nabla v + v\nabla u \qquad u,\ v \in H^{1,p}(\Omega) \cap L^\infty(\Omega)\,. \tag{1.8}$$

On the other hand, we don't have to prove the same formulas 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 (1.5), since in this case $\Omega_\varepsilon = \mathbb{R}^n$), one can prove the following result:

Theorem 1.5 ($H = W$). *If either $\Omega = \mathbb{R}^n$ or Ω is a bounded regular domain, then*

$$H^{1,p}(\Omega) = W^{1,p}(\Omega) \qquad 1 \le p < \infty. \tag{1.9}$$

With the word *regular* we mean that Ω is the epigrah of a Lipschitz function of $(n-1)$-variables, written in a suitable system of coordinates, near to any boundary point. However the equality $H = W$ is not true in general, as the following example shows.

Example 1.6. In the Euclidean plane $\mathbb{R}^2$, consider the open unit ball $x^2+y^2 < 1$ deprived of one of its radii, say for instance the segment Σ 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 $1 \leq p < 2$. It is immediate to see that $v \in C^\infty(\Omega)$ and that its gradient is p-integrable, hence $v \in W^{1,p}(\Omega)$. On the other hand, $v \notin H^{1,p}(\Omega)$ because the definition we have given would require the existence of regular approximations for v up to the boundary: more precisely, one can easily show, using Fubini's theorem and polar coordinates, that any $u \in H^{1,p}(\Omega)$ satisfies

$$\omega \mapsto u(re^{i\omega}) \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}) \qquad \text{for } \mathscr{L}^1\text{-a.e. } r \in (0,1). \tag{1.10}$$

Indeed, if $u_n \in C^0(\overline{\Omega}) \cap C^1(\Omega)$ converge to u strongly in $H^{1,p}(\Omega)$ and (possibly extracting a subsequence) $\sum_n \|\nabla u_{n+1} - \nabla u_n\|_p < \infty$, for all $\delta \in (0,1)$ the inequality $|\partial_\theta v| \leq |\nabla v|/r$ gives

$$\int_\delta^1 \sum_n \left(\int_{-\pi}^{\pi} |\frac{\partial u_{n+1}}{\partial\theta} - \frac{\partial u_n}{\partial\theta}|^p \, d\theta \right)^{1/p} dr \leq \delta^{-1-1/p} \sum_n \|\nabla u_{n+1} - \nabla u_n\|_p < \infty.$$

Since $\delta > 0$ is arbitrary, it follows that for $\mathscr{L}^1$-a.e. $r \in (0,1)$ the 2π-periodic continuous functions $\theta \mapsto u_n(re^{i\theta})$ have derivatives strongly convergent in $L^p_{\mathrm{loc}}(\mathbb{R})$, and therefore (by the fundamental theorem of calculus) are equicontinuous. Any limit point of these functions in $L^p_{\mathrm{loc}}(\mathbb{R})$ must then be 2π-periodic, continuous and $W^{1,p}$. If, by contradiction, we take $u = v$, a similar Fubini argument shows that $u_n(re^{i\theta})$ converge 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 v can't be in $H^{1,p}(\Omega)$.

Remark 1.7. Taking into account the example above, we mention the Meyers-Serrin theorem [24], ensuring that, for *any* open set $\Omega \subset \mathbb{R}^n$ and $1 \leq p < \infty$, the identity

$$\overline{C^\infty(\Omega) \cap W^{1,p}(\Omega)}^{W^{1,p}} = W^{1,p}(\Omega) \tag{1.11}$$

holds. The proof can be achieved by (1.5) and a partition of unity.
The previous example underlines the crucial role played by the boundary behaviour, when we try to approximate a function in $W^{1,p}$ by $C^1(\overline{\Omega})$ (or even $C^0(\overline{\Omega}) \cap C^1(\Omega)$) 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 $C^1(\Omega) \cap L^p(\Omega)$ functions with gradient in $L^p(\Omega)$ instead of $C^1(\overline{\Omega})$ functions, the identity $H = W$ would

be true unconditionally. In the case $p = \infty$, the construction in the Meyers-Serrin theorem provides for all $u \in W^{1,\infty}(\Omega)$ a sequence $(u_n) \subset C^\infty(\Omega)$ converging to u uniformly in Ω, with $\sup_\Omega |\nabla u_n|$ convergent to $\|\nabla u\|_\infty$. Again, this might lead to 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 $H^{1,p}(\Omega)$ in order to work with functions vanishing at the boundary.

Definition 1.8. *Given $\Omega \subset \mathbb{R}^n$ open, we define $H_0^{1,p}(\Omega)$ to be the completion of $C_c^1(\Omega)$ with respect to the $W^{1,p}$ norm.*

It is clear that $H_0^{1,p}(\Omega)$, being *complete*, is a *closed* subspace of $H^{1,p}(\Omega)$. Notice also that $H^{1,p}(\mathbb{R}^n)$ coincides with $H_0^{1,p}(\mathbb{R}^n)$. To see this, suffices to show that any function $u \in C^1(\mathbb{R}^n)$ with both $|u|$ and $|\nabla u|$ in $L^p(\mathbb{R}^n)$ belongs to $H_0^{1,p}(\mathbb{R}^n)$. We can indeed approximate any such function u, strongly in $H^{1,p}$ norm, by the compactly supported functions $\chi_R u$, where $\chi_R : \mathbb{R}^n \to [0,1]$ are smooth, 2-Lipschitz, identically equal to 1 on $\overline{B}_R$ and identically equal to 0 on $\mathbb{R}^n \setminus \overline{B}_{R+1}$.

We now turn to some classical inequalities.

Theorem 1.9 (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 $C(\Omega, p)$, depending only on Ω and p, such that*

$$\|u\|_{L^p} \le C(\Omega, p) \|\nabla u\|_{L^p} \qquad \forall u \in H_0^{1,p}(\Omega). \tag{1.12}$$

In addition $C(\Omega) \le C(n,p)\mathrm{diam}(\Omega)$.

The proof of this result can be simplified by means of these properties:

- $H_0^{1,p}(\Omega) \subset H_0^{1,p}(\Omega')$ if $\Omega \subset \Omega'$ (*monotonicity*);
- If $C(\Omega, p)$ denotes the best constant, then $C(\lambda\Omega, p) = \lambda C(\Omega, p)$ (*scaling invariance*) and $C(\Omega + h, p) = C(\Omega, p)$ (*translation invariance*).

The first fact is a consequence of the definition of the spaces $H_0^{1,p}$ in terms of regular functions, while the second one (translation invariance is obvious) follows by:

$$u_\lambda(x) = u(\lambda x) \in H_0^{1,p}(\Omega) \qquad \forall u \in H_0^{1,p}(\lambda\Omega). \tag{1.13}$$

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 axis and length 2. We write $x = (x_1, x')$ with $x' = (x_2, \ldots, x_n)$. By density, we may also assume $u \in C_c^1(\Omega)$ and hence use the following representation formula:

$$u(x_1, x') = \int_{-1}^{x_1} \frac{\partial u}{\partial x_1}(t, x')\, dt. \tag{1.14}$$

Hölder's inequality gives

$$|u|^p(x_1, x') \le 2^{p-1} \int_{-1}^{1} \left|\frac{\partial u}{\partial x_1}\right|^p (t, x')dt \tag{1.15}$$

and hence we just need to integrate w.r.t. x_1 to get

$$\int_{-1}^{1} |u|^p(x_1, x')\, dx_1 \le 2^p \int_{-1}^{1} \left|\frac{\partial u}{\partial x_1}\right|^p (t, x')dt. \tag{1.16}$$

Now, integrating w.r.t. x', repeating the previous argument for all the variables x_j, $j = 1, \dots, n$ and summing all such inequalities we obtain the thesis with $C(Q, p) \le 2/n^{1/p}$. □

Theorem 1.10 (Rellich). *Let Ω be an open bounded subset with regular boundary and let $p \in [1, \infty)$. Then the immersion $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ 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*

$$T : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n) \tag{1.17}$$

such that

$$\begin{cases} Tu = u & \text{in } \Omega; \\ \\ \operatorname{supp}(Tu) \subset \Omega' & , \end{cases}$$

being Ω' a fixed bounded domain in $\mathbb{R}^n$ containing $\overline{\Omega}$. When Ω is an halfspace the operator can be achieved simply by a 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 Lipschitz maps (we will use these ideas later on, treating the boundary regularity of solutions to elliptic PDE's). The global construction is then obtained thanks to a partition of unity.

The operator T allows basically a reduction to the case $\Omega = \mathbb{R}^n$, considered in the next theorem.

Theorem 1.11. *The immersion $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^p_{\text{loc}}(\mathbb{R}^n)$ is compact, namely if $(u_k) \subset W^{1,p}(\mathbb{R}^n)$ is bounded, then (u_n) has limit points in the $L^p_{\text{loc}}(\mathbb{R}^n)$ topology, and any limit point belongs to $W^{1,p}(\mathbb{R}^n)$.*

Remark 1.12. It should be noted that the immersion $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ is obviously continuous, but certainly not compact: just take a fixed element in $W^{1,p}(\mathbb{R}^n)$ and supported 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 $W^{1,p}(\mathbb{R}^n)$ but no subsequence converges in $L^p(\mathbb{R}^n)$ (indeed, all functions have the same 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 will be often considered in the sequel.

Proof. Basically, it is enough to prove that a bounded family $\mathcal{F} \subset W^{1,p}(\mathbb{R}^n)$ is totally bounded in $L^p_{loc}(\mathbb{R}^n)$. To obtain this, observe firstly that given any Borel domain $A \subset \mathbb{R}^n$ and any $\varphi \in W^{1,p}(A_{|h|})$ we have

$$\|\tau_h\varphi - \varphi\|_{L^p(A)} \leq |h| \, \|\nabla\varphi\|_{L^p(A_{|h|})} \tag{1.18}$$

where $A_{|h|}$ is the $|h|$ −neighbourhood of the set A and $\tau_h\varphi(x) = \varphi(x+h)$. By approximation, we can assume with no loss of generality that $\varphi \in C^1(A_{|h|})$. The inequality (1.18) follows by the elementary representation

$$(\tau_h\varphi - \varphi)(x) = \int_0^1 \langle \nabla\varphi(x+sh), h\rangle \; ds \tag{1.19}$$

since

$$\begin{aligned} \|\tau_h\varphi - \varphi\|^p_{L^p(A)} &\leq \int_A \int_0^1 |\langle \nabla\varphi(x+sh), h\rangle|^p \; ds \, dx & (1.20)\\ &\leq |h|^p \int_0^1 \int_{A_{|h|}} |\nabla\varphi(y)|^p \; dy \, ds = |h|^p \, \|\nabla\varphi\|^p_{L^p(A_{|h|})} & (1.21) \end{aligned}$$

where we used the Cauchy-Schwarz inequality and Fubini's theorem. Hence, denoting by $(\rho_\varepsilon)_{\varepsilon>0}$ any rescaled family of smooth mollifiers such that $\mathrm{supp}(\rho_\varepsilon) \subset B(0,\varepsilon)$, we have that for any $R > 0$

$$\sup_{\varphi\in\mathcal{F}} \|\varphi * \rho_\varepsilon - \varphi\|_{L^p(B_R)} \to 0 \tag{1.22}$$

for $\varepsilon \to 0$. In fact, since $\varphi * \rho_\varepsilon$ is a mean, weighted by ρ_ε, of translates of φ

$$\varphi * \rho_\varepsilon = \int \tau_{-y}\varphi\rho_\varepsilon(y) \, dy \, ,$$

by the previous result we deduce

$$\sup_{\varphi\in\mathcal{F}} \|\varphi * \rho_\varepsilon - \varphi\|_{L^p(B_R)} \leq \varepsilon \sup_{\varphi\in\mathcal{F}} \left(\int_{B_{R+\varepsilon}} |\nabla\varphi|^p \; dx \right)^{1/p} . \tag{1.23}$$

To conclude we just need to observe that the regularized family $\{\varphi * \rho_\varepsilon, \varphi \in \mathcal{F}\}$ is relatively compact in $L^p_{loc}(\mathbb{R}^n)$ for any fixed $\varepsilon > 0$. But this is easy since the Young inequality implies

$$\sup_{B_R} |\varphi * \rho_\varepsilon| \leq \|\varphi\|_{L^1(B_{R+\varepsilon})} \, \|\rho_\varepsilon\|_\infty \tag{1.24}$$

and similarly

$$\sup_{B_R} |\nabla(\varphi * \rho_\varepsilon)| \le \|\varphi\|_{L^1(B_{R+\varepsilon})} \|\nabla \rho_\varepsilon\|_\infty \tag{1.25}$$

so the claim follows by means of the Ascoli-Arzelá theorem. Notice that we used the gradient bounds on elements of $\mathcal{F}$ only in (1.23). □

We also need to mention another inequality due to Poincaré.

Theorem 1.13 (Poncaré inequality, second version). *Let us consider a bounded, regular and connected domain $\Omega \subset \mathbb{R}^n$ and an exponent $1 \le p < \infty$, so that by Rellich's theorem we have the compact immersion $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$. Then, there exists a constant $C(\Omega, p)$ such that*

$$\int_\Omega |u - u_\Omega|^p \, dx \le C \int_\Omega |\nabla u|^p \, dx \qquad \forall u \in W^{1,p}(\Omega) \tag{1.26}$$

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 $(u_n) \subset W^{1,p}(\Omega)$ such that

- $(u_n)_\Omega = 0$ for all $n \in \mathbb{N}$;
- $\int_\Omega |u_n|^p \, dx = 1$ for all $n \in \mathbb{N}$;
- $\int_\Omega |\nabla u_n|^p \, dx \to 0$ for $n \to \infty$.

By Rellich's theorem there exists (up to a subsequence) a limit point $u \in L^p$, that is $u_n \to u$ in $L^p(\Omega)$. It is now a general fact that if (∇u_n) has some weak limit point g then necessarily $g = \nabla u$. Therefore, in this case we have by comparison $\nabla u = 0$ in $L^p(\Omega)$ and hence, by connectivity of the domain and the constancy theorem, we deduce that u must be equivalent to a constant. By taking limits we see that u satisfies at the same time

$$\int_\Omega u \, dx = 0 \qquad \text{and} \qquad \int_\Omega |u|^p \, dx = 1, \tag{1.27}$$

which is clearly impossible. □

Note that the previous proof is not constructive and crucially relies on the general compactness result by Rellich.

Remark 1.14. It should be observed that the previous proof, even though very simple, is far from giving the sharp constant for the Poincaré inequality (1.26). 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). Many more results are instead available for the sharp constant in the Poincaré inequality (1.12).

2 Variational formulation of some PDEs

After the introductory section, whose main purpose was to fix the notation and recall some basic tools of the theory of Sobolev spaces, we are now ready to discuss some basic elliptic PDEs.

Let us consider the generalised Poisson equation

$$\begin{cases} -\Delta u = f - \sum_\alpha \partial_\alpha f_\alpha & \text{in } \Omega; \\ u \in H_0^{1,2}(\Omega) \end{cases} \tag{2.1}$$

with data $f,\ f_\alpha \in L^2(\Omega)$ for some fixed bounded and regular domain Ω. This equation has to be intended in a weak sense, that is, we look for $u \in H_0^{1,2}(\Omega)$ satisfying

$$\int_\Omega \langle \nabla u, \nabla \varphi \rangle \, dx = \int_\Omega (f\varphi + \sum_\alpha f_\alpha \partial_\alpha \varphi) \, dx \qquad \forall \varphi \in C_c^\infty(\Omega). \tag{2.2}$$

Equivalently, by continuity of the bilinear form and density of $C_c^\infty(\Omega)$, the previous condition could be requested for any $\varphi \in H_0^{1,2}(\Omega)$.

In order to obtain existence we just need to apply Riesz's theorem to the associated *linear* functional $F(v) = \int_\Omega (fv + \sum_\alpha f_\alpha \partial_\alpha v) \, dx$ on the Hilbert space $H_0^{1,2}(\Omega)$ endowed with the scalar product

$$(u, v) = \int_\Omega \langle \nabla u, \nabla v \rangle \, dx \tag{2.3}$$

which is equivalent to the usual one thanks to the Poincaré inequality (first version) proved in Theorem 1.9.

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

- more general differential operators instead of $-\Delta$;
- inhomogeneous or mixed boundary conditions;
- systems instead of single equations.

Our purpose now is to briefly discuss each of these situations.

2.1 Elliptic operators

The first variant is to consider scalar problems having the form

$$\begin{cases} -\sum_{\alpha,\beta} \partial_\alpha (A^{\alpha\beta} \partial_\beta u) = f - \sum_\alpha \partial_\alpha f_\alpha & \text{in } \Omega; \\ u \in H_0^{1,2}(\Omega) \end{cases}$$

where, as before $f,\ f_\alpha \in L^2(\Omega)$, 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 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, precisely $u(x) = v(A^{-1}x)$, transforms this problem into one of the form (2.1). For this reason it is convenient to deal immediately with the case of a non-constant matrix $A(x) \in \mathbb{R}^{n\times n}$ satisfying:

(i) A is a Borel and L^∞ function defined on Ω;

(ii) $A(x)$ is symmetric for a.e. $x \in \Omega$;

(iii) there exists a positive constant c such that

$$A(x) \geq cI \quad \text{for a.e. } x \in \Omega\,. \tag{2.4}$$

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

$$\int_\Omega \langle A\nabla u, \nabla\varphi\rangle \, dx = \int_\Omega (f\varphi + \sum_\alpha f_\alpha \partial_\alpha \varphi)\, dx \qquad \forall \varphi \in C_c^\infty(\Omega). \tag{2.5}$$

By continuity and density, also in this case it is equivalent to require the validity of the identity above for all $\varphi \in H_0^{1,2}(\Omega)$. In order to obtain existence we could easily modify the previous argument when $|A| \in L^\infty(\Omega)$, using the equivalent scalar product

$$\langle u, v\rangle := \int_\Omega \sum_{\alpha,\beta} A^{\alpha\beta}\partial_\alpha u \partial_\beta v\, dx\,.$$

However, in order to include also unbounded A's, thus dropping assumption (i), we prefer here to proceed differently and introduce some ideas that belong to the so-called *direct method* of the Calculus of Variations. Let us consider the functional $F : H_0^{1,2}(\Omega) \to \mathbb{R}$

$$F(v) = \int_\Omega \frac{1}{2}\langle A\nabla v, \nabla v\rangle \, dx - \int_\Omega fv\, dx - \sum_\alpha \int_\Omega f_\alpha \partial_\alpha v\, dx. \tag{2.6}$$

First we note that, thanks to the assumption (2.4) on A, for all $\varepsilon > 0$ it holds

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

Choosing $\varepsilon < c/2$ we get

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

and now, thanks to the Poincaré inequality, we can choose possibly ε even smaller to get

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

In particular F is *coercive*, that is

$$\lim_{\|v\|_{H_0^{1,2}(\Omega)}\to+\infty} F(v) = +\infty \tag{2.7}$$

and consequently, in order to look for its minimum points it is enough to consider a closed ball of $H_0^{1,2}(\Omega)$. Now, take any minimizing sequence (u_n) of F: since $H_0^{1,2}(\Omega)$ is (being Hilbert) a reflexive space we can assume, possibly extracting a subsequence, that $u_n \rightharpoonup u$ for some $u \in H_0^{1,2}(\Omega)$. Using Fatou's lemma and the fact that $u_h \to u$ in $H^{1,2}$ implies $\nabla u_{h(k)} \to \nabla u$ a.e. in Ω for a suitable subsequence $h(k)$, it is not difficult to prove that F is lower semicontinuous (we shall also prove this in Theorem 3.2, in a more general framework). In addition, F is convex, being the sum of a linear and a convex functional. It follows that F is also *weakly* lower semicontinuous, hence

$$F(u) \leq \liminf_{n\to\infty} F(u_n) = \inf_{H_0^{1,2}(\Omega)} F \tag{2.8}$$

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

If A is bounded, since F is a C^1 functional on $H_0^{1,2}(\Omega)$ we get $dF(u) = 0$, where dF is the differential in the Gateaux sense of F:

$$dF(u)\,[\varphi] := \lim_{\varepsilon\to0} \frac{F(u+\varepsilon\varphi) - F(u)}{\varepsilon} \qquad \forall\varphi \in H_0^{1,2}(\Omega)\,.$$

Here a simple computation gives

$$dF(u)\,[\varphi] = \int_\Omega \langle A\nabla u, \nabla\varphi\rangle\,dx - \int_\Omega f\varphi\,dx - \sum_\alpha \int_\Omega f_\alpha \partial_\alpha\varphi\,dx \tag{2.9}$$

and the desired result follows. Even in the case when $|A| \in L^1_{\mathrm{loc}}$ we can still differentiate the functional, but a priori only along directions in $\varphi \in C_c^\infty(\Omega)$, and recover the weak formulation of our PDE.

2.2 Inhomogeneous boundary conditions

We now turn to study the boundary value problem for $u \in H^{1,2}(\Omega)$

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

with $f,\ f_\alpha \in L^2(\Omega)$ and a suitable class of functions $g \in L^2(\partial\Omega)$. Since the immersion $H^{1,2}(\Omega) \hookrightarrow C(\overline{\Omega})$ does not hold if $n \geq 2$, the boundary condition has to be considered in the *weak sense* described below.

Here and in the sequel, unless otherwise stated, we indicate with Ω an open, bounded and regular subset of $\mathbb{R}^n$.

Theorem 2.1. *For any $p \in [1, \infty)$ the restriction operator*

$$T : C^1(\overline{\Omega}) \to C^0(\partial\Omega) \tag{2.10}$$

satisfies $\|Tu\|_{L^p(\partial\Omega)} \leq C(p, \Omega)\|u\|_{W^{1,p}(\Omega)}$. Therefore it can be uniquely extended to a linear and continuous operator from $W^{1,p}(\Omega)$ to $L^p(\partial\Omega)$.

Proof. We prove the result only in the case when Ω is the subgraph of a Lipschitz function f inside the rectangle $\Omega' \times (a, b)$, with $\Omega \subset \mathbb{R}^{n-1}$ open, with $a' = \inf f > a$, proving the estimate on the portion

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

of its boundary (here we use the notation $x = (x', t)$ with $x' \in \Omega'$ and $t \in (a, b)$). The general case can be easily achieved by a partition of unity argument.

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 w.r.t. 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 that inserting the area element $\sqrt{1 + |\nabla f(x')|^2}$ and using the inequality $|r + s|^p \leq 2^{p-1}(|r|^p + |s|^p)$ gives

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

where L is the Lipschitz constant of f.

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

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

$\square$

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

$$Tu = g. \tag{2.11}$$

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

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i} \, dx = -\int_\Omega \varphi \frac{\partial u}{\partial x_i} \, dx + \int_{\partial\Omega} \varphi \, Tu \, \nu_i \, d\sigma \qquad \forall \varphi \in C^1(\overline{\Omega}) \tag{2.12}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit normal vector, pointing out of Ω. Indeed, using the equality $H^{1,p}(\Omega) = W^{1,p}(\Omega)$ of Theorem 1.5 one can start from the classical divergence theorem with $u \in C^1(\overline{\Omega})$ and then argue by approximation.

Remark 2.2. It is possible to show that the previously defined restriction 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

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

see [1], with $s = 1 - 1/p$. The borderline case $p = 1$ is special, and in this case Gagliardo proved in [13] the surjectivity of T.

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

$$\begin{cases} -\Delta v = \widetilde{f} - \sum_\alpha \partial_\alpha \widetilde{f}_\alpha & \text{in } \Omega; \\ \\ v \in H^{1,2}_0(\Omega) \, . \end{cases}$$

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

Finally, let us discuss the so-called Neumann boundary conditions, involving the behaviour of the normal derivative of u on the boundary. We consider a problem of the form

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

with $A^{\alpha\beta}$ a real matrix and $\lambda > 0$ a fixed constant. For the sake of brevity, we just discuss the case $A^{\alpha\beta} = \delta_{\alpha\beta}$ so that the problem above becomes

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

In order to give it a clear meaning, note that if u, $v \in C^1(\overline{\Omega})$ then

$$\int_\Omega \langle \nabla u, \nabla v \rangle \, dx = -\int_\Omega v \Delta u \, dx + \int_{\partial\Omega} v \frac{\partial u}{\partial \nu} \, d\sigma \tag{2.13}$$

and so in this case it is natural to ask that for any $v \in C^1(\overline{\Omega})$ the desired solution u satisfies

$$\int_\Omega [\langle \nabla u, \nabla v \rangle + \lambda u v] \, dx = \int_\Omega v f \, dx + \int_{\partial\Omega} v g \, d\sigma. \tag{2.14}$$

In order to obtain existence (and uniqueness) for this problem when $g \in L^2(\partial\Omega)$, it is enough to apply Riesz's theorem to the bilinear form on $H^{1,2}(\Omega)$

$$a(u,v) = \int_\Omega [\langle \nabla u, \nabla v \rangle + \lambda u v] \, dx \tag{2.15}$$

which is clearly equivalent to the standard Hilbert product on the same space (since $\lambda > 0$) and the continuous linear functional $F(v) = \int_\Omega v f \, dx + \int_{\partial\Omega} v g \, d\sigma$.

2.3 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 $\alpha, \beta, \ldots$) in order to indicate 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, $j \in \{1, 2, \ldots, m\}$). In many cases, we will need to work with *four indices matrices* (i.e. rank four tensors) like $A^{\alpha\beta}_{ij}$, 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 vector-valued case. The first idea is to define the *Legendre condition*

$$\sum_{\alpha,\beta,i,j} A_{ij}^{\alpha\beta} \xi_\alpha^i \xi_\beta^j \geq c\,|\xi|^2 \quad \forall \xi \in \mathbb{R}^{m\times n} \tag{2.16}$$

where $\mathbb{R}^{m\times n}$ indicates, as above, the space of $m \times n$ real matrices. Let us apply it in order to obtain existence and uniqueness for the system

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

with data f_i, $f_i^\alpha \in L^2(\Omega)$.[1] The weak formulation of the problem is obviously

$$\int_\Omega \sum_{i,j,\alpha,\beta} A_{ij}^{\alpha\beta} \partial_\beta u^j \partial_\alpha \varphi^i \, dx = \int_\Omega \left[\sum_i f_i \varphi^i + \sum_{i,\alpha} f_i^\alpha \partial_\alpha \varphi^i \right] dx \tag{2.17}$$

for every $\varphi \in [C_c^1(\Omega)]^m$ and $i = 1,\ldots,m$. 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 (α,β) and (i,j)), then it defines a scalar product on $H_0^1(\Omega;\mathbb{R}^m)$ by the formula

$$(\varphi,\psi) = \int_\Omega \sum_{i,j,\alpha,\beta} A_{ij}^{\alpha\beta} \partial_\alpha \varphi^i \partial_\beta \psi^j \, dx. \tag{2.18}$$

If, moreover, A satisfies the Legendre condition (2.16) for some $c > 0$, it is immediate to see that this scalar product is equivalent to the standard one (with $A_{i,j}^{\alpha\beta} = \delta^{\alpha\beta}\delta_{ij}$) and so we are led to apply again Riesz's theorem to conclude the proof.

From now on, we will often adopt Einstein's summation convention on repeated indices, using it without explicit mention.

It should be noted that in the proof of some existence result (and, in particular, in the scalar case) the symmetry hypothesis w.r.t. the transformation $(\alpha,i) \to (\beta,j)$ is not necessary, since we can exploit the following:

Theorem 2.3 (Lax-Milgram)**.** *Let H be a (real) Hilbert space and let $a : H \times H \to \mathbb{R}$ a bilinear, continuous and coercive form so 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 $v \in H$.

[1] Note that we sometimes omit the Sobolev exponent when this is equal to two: for instance $H_0^1(\Omega)$ stands for $H_0^{1,2}(\Omega)$.

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

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

and 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 $\|T\| \leq C$. Now we introduce the auxiliary bilinear form

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

which 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 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 achieved setting $u_F = T^*u'_F$:

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

□

As indicated above, we now want to formulate a different notion of ellipticity for the vector case. To this aim, it is useful to analyse more in detail the scalar case. We have 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*).

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 ($m = 1$) and fails in the vectorial case ($m > 1$).

Proposition 2.4. *Let (C) and (E) as above. Then, (C) is equivalent to (E).*

Proof. Let is 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 $u,\, v \in H^1_0(\Omega;\mathbb{C})$

$$a_A(u,v) = \int_\Omega \left\langle A\nabla u, \overline{\nabla v}\right\rangle\, dx = \int_\Omega \sum_{\alpha,\,\beta=1}^n A^{\alpha\beta}\partial_\alpha u\overline{\partial_\beta u}\, dx\,.$$

A simple computation shows that (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 part of u)

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

Hence, (C) implies

$$\Re a_A(u,u) \geq \lambda \int_\Omega |\nabla u|^2\, dx\,. \tag{2.19}$$

Now consider a function $\varphi \in C^\infty_c(\Omega)$ and define $u_\tau(x) = \varphi(x)e^{i\tau x\cdot\xi}$. We have that

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

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

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

Hence, exploiting our coercivity assumption and letting $\tau \to +\infty$ in (2.19) we get

$$\left(A^{\alpha\beta}\xi_\alpha\xi_\beta - \lambda\,|\xi|^2\right)\int_\Omega \varphi^2\, dx \geq 0 \tag{2.20}$$

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 Borel and L^∞ function in Ω and we can conclude that (E) holds for a.e. $x \in \Omega$: we just need to choose, in the very last step, an appropriate sequence of rescaled and normalized mollifiers concentrating around x_0, for any Lebesgue point x_0 of A. The conclusion comes, in this situation, by Lebesgue differentiation theorem.

For the reader's convenience we recall here some basic facts concerning Lebesgue points (see also Section 13). Given $f \in L^1_{\rm loc}(\mathbb{R}^n)$ 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

$$\lim_{r\downarrow 0} \fint_{B_r(x_0)} |f(y) - \lambda|\, dy = 0. \tag{2.21}$$

In this case λ is unique and it is sometimes written

$$\lambda = \widetilde{f}(x_0) = \widetilde{\lim_{x \to x_0}} f(x). \tag{2.22}$$

Notice that both the notion of Lebesgue point and $\widetilde{f}$ *are invariant* in the Lebesgue equivalence class of f. The *Lebesgue differentiation theorem* says 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 $m > 1$: in fact, the coercivity condition

$$a_A(u,u) \geq \lambda \int_\Omega |\nabla u|^2 \, dx \qquad u \in H^1(\Omega; \mathbb{R}^m) \tag{2.23}$$

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*

$$A_{ij}^{\alpha\beta} \xi_\alpha^i \xi_\beta^j \geq \lambda \, |\xi|^2 \qquad \text{for all } \xi = a \otimes b \,, \tag{2.24}$$

that is the *Legendre condition* restricted to rank one matrices $\xi_\alpha^i = a_\alpha b^i$. Explicit examples show that the Legendre-Hadamard condition is in general *strictly weaker* than the Legendre condition.

Example 2.5. When $m = n = 2$, consider the tensor $A_{ij}^{\alpha\beta}$ implicitly defined by

$$A_{ij}^{\alpha\beta} \xi_\alpha^i \xi_\beta^j = \det(\xi) + \varepsilon \, |\xi|^2 \tag{2.25}$$

with $\varepsilon \geq 0$. Since $t \mapsto \det(M + tN)$ is linear for any rank one matrix N, 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 the Legendre condition with $\lambda = 0$ fails when $2\varepsilon < 1$.

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

Theorem 2.6 (Gårding)**.** *Assume that $A_{ij}^{\alpha\beta}$ satisfies the Legendre-Hadamard condition for some positive constant λ and the symmetry condition $A_{ij}^{\alpha\beta} = A^{\beta\alpha} ji$. 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, we denote by $\mathcal{S}(\mathbb{R}^n)$ 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

$$\widehat{\varphi}(\xi) = (2\pi)^{-n/2} \int \varphi(x) e^{-ix\cdot\xi}\, dx \tag{2.26}$$

and

$$\widetilde{\varphi}(x) = (2\pi)^{-n/2} \int \varphi(\xi) e^{ix\cdot\xi}\, d\xi\,. \tag{2.27}$$

We will also make use of the Plancherel identity:

$$\int \widehat{\varphi}\overline{\widehat{\psi}}\, d\xi = \int \varphi\overline{\psi}\, dx \qquad \forall \varphi,\, \psi \in \mathcal{S}(\mathbb{R}^n)\,. \tag{2.28}$$

Proof. By density it is enough to prove the result when $u \in [C_c^\infty(\mathbb{R}^n)]^m$. 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 $\varphi \in [\mathcal{S}(\mathbb{R}^n)]^m$. Consequently,

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

hence

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

the last passage being due to the Plancherel identity (2.28). Now we can apply our hypothesis to get

$$A_{jl}^{\alpha\beta} a_\alpha b^j a_\beta \overline{b^l} \geq \lambda |a|^2 |b|^2$$

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

$$a_A(u,u) \geq \lambda \int_{\mathbb{R}^n} |x|^2 \left|\varphi(x)\right|^2\, dx\,. \tag{2.29}$$

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

$$\int_{\mathbb{R}^n} \left|\nabla u\right|^2(\xi)\, d\xi = \int_{\mathbb{R}^n} |x|^2 \left|\varphi(x)\right|^2\, dx\,. \tag{2.30}$$

Comparing (2.29) and (2.30) we conclude the proof. $\square$

Remark 2.7. 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*, namely, they do not arise from the integration of a pointwise inequality. An important example is *Korn's inequality*:

$$\int_{\mathbb{R}^n} |\nabla u|^p \, dx \leq c(n,p) \int_{\mathbb{R}^n} \left| \frac{\nabla u + (\nabla u)^t}{2} \right|^p dx \qquad \text{for all } u \in C_c^\infty(\mathbb{R}^n, \mathbb{R}^n)\,, \tag{2.31}$$

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

$$\inf_{c \in \mathbb{R}^m,\, B^t = -B} \int_\Omega |u(x) - Bx - c|^p \, dx \leq C(\Omega, p) \int_\Omega \left| \frac{\nabla u + (\nabla u)^t}{2} \right|^p dx\,. \tag{2.32}$$

2.4 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 $u \in W^{1,\infty}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$, fix a Lagrangian L and define a functional

$$F(u, \Omega) = \int_\Omega L(x, u, \nabla u)\, dx\,.$$

We say that u is a *local minimizer* for F if

$$F(u, A) \leq F(v, A) \qquad \text{for all } v \in W^{1,\infty}_{\mathrm{loc}}(\Omega; \mathbb{R}^m) \text{ such that } \{v \neq u\} \Subset A \Subset \Omega. \tag{2.33}$$

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

$$\sup_K \left(|L| + |L_s| + |L_p|\right) < \infty \tag{2.34}$$

for any domain $K = \Omega' \times \{(s,p) | \, |s| + |p| \leq R\}$ 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, $\varphi \in W^{1,\infty}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ and $\Omega' \Subset \Omega$, and its derivative equals

$$\int_{\Omega'} L_s(x, u + t\varphi, \nabla u + t\nabla\varphi) \cdot \varphi + L_p(x, u + t\varphi, \nabla u + t\nabla\varphi) \cdot \nabla\varphi \, dx$$

(the assumption (2.34) is needed to differentiate under the integral sign). As a consequence, if a locally Lipschitz function u is a local minimizer and $\varphi \in C_c^\infty(\Omega';\mathbb{R}^m)$, since $F(u,\Omega') \le F(u+t\varphi,\Omega')$ we can differentiate at $t=0$ to obtain

$$\int_{\Omega'} \left[\sum_i L_{s_i}(x,u,\nabla u)\varphi^i + \sum_{\alpha,i} L_{p_i^\alpha}(x,u,\nabla u)\frac{\partial \varphi^i}{\partial x_\alpha} \right] dx = 0\,. \tag{2.35}$$

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

$$\frac{\partial}{\partial x_\alpha} L_{p_i^\alpha}(x,u,\nabla u) = L_{s^i}(x,u,\nabla u) \qquad i=1,2,\ldots,m\,.$$

Exploiting this idea, we can associate to many classes of PDEs 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), equations having the form

$$-\Delta u = g(x,u) \tag{2.36}$$

derive from the functional

$$L(x,s,p) = \frac{1}{2}\,|p|^2 - \int_0^s g(x,r)\,dr\,. \tag{2.37}$$

Adding stronger hypotheses on the Lagrangian L, in analogy with what has been done above, i.e. requiring that

$$\sup_K \left(|L_{ss}| + |L_{sp}| + |L_{pp}|\right) < \infty$$

for any domain $K = \Omega' \times \{(s,p) |\, |s| + |p| \le R\}$ with $\Omega' \Subset \Omega$, we can find another *necessary* minimality condition corresponding to

$$\frac{d^2}{dt^2} F(u+t\varphi)\bigg|_{t=0} \ge 0\,,$$

namely

$$0 \le \Gamma(\varphi,\varphi) = \int_\Omega \left[A\nabla\varphi\nabla\varphi + B\nabla\varphi\cdot\varphi + C\varphi\cdot\varphi\right] dx \qquad \forall \varphi \in C_c^\infty(\Omega;\mathbb{R}^m)\,, \tag{2.38}$$

where the dependence on x and all indices are omitted for brevity and

$$\begin{cases} A(x) = L_{pp}(x,u(x),\nabla u(x))\,; \\ B(x) = L_{ps}(x,u(x),\nabla u(x))\,; \\ C(x) = L_{ss}(x,u(x),\nabla u(x))\,. \end{cases} \tag{2.39}$$

We can finally obtain *pointwise* conditions on the local minimizer u by means of the following theorem, whose proof can be obtained arguing as in the proof that coercivity implies ellipticity (Proposition 2.4).

Theorem 2.8. *Consider the bilinear form on $H_0^1(\Omega;\mathbb{R}^m)$ defined by*

$$\Theta(u,v) = \int_\Omega (A\nabla u\nabla v + B\nabla u\cdot v + Cu\cdot v)\,dx\,, \tag{2.40}$$

where $A = A_{ij}^{\alpha\beta}(x)$, $B = B_{ij}^{\alpha}(x)$ and $C = C_{ij}(x)$ are Borel and L^∞ functions. If $\Theta(u,u) \geq 0$ for all $u \in H_0^1(\Omega;\mathbb{R}^m)$ then $A(x)$ satisfies the Legendre-Hadamard condition with $\lambda = 0$ for a.e. $x \in \Omega$.

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

3 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

$$F(u) := \int_\Omega L(x, u(x), \nabla u(x))\,dx \tag{3.1}$$

in suitable function spaces (including for instance the boundary conditions).

Before stating Tonelli's theorem, we recall some useful facts about uniformly integrable maps. A comprehensive treatment of this subject can be found for instance in [28], see also [3, Theorem 1.38].

Theorem 3.1 (Dunford-Pettis)**.** *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and let $\mathcal{F} \subset L^1(X, \mathcal{A}, \mu)$. Then the following facts are equivalent:*

(i) the family $\mathcal{F}$ is sequentially relatively compact w.r.t. 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

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

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

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

Theorem 3.2 (Tonelli). *Let $L(x,s,p) : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$ be a Lagrangian satisfying the following properties:*

(1) L is non-negative;

(2) L is lower semicontinuous w.r.t. s and the partial derivatives $L_{p_i^\alpha}$ exist and are continuous w.r.t. s;

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

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

$$\liminf_{h\to\infty} F(u_h) \geq F(u) \ . \tag{3.2}$$

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

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

and, possibly extracting one more subsequence,

$$u_{h(k)} \longrightarrow u \quad \text{a.e. in } \Omega\,.$$

Thanks to the Dunford-Pettis Theorem we can also assume the weak-L^1 convergence

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

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 $\nabla u = g$.

Thanks to Egorov's Theorem, for all $\varepsilon > 0$ there exists a compact subset $K_\varepsilon \subset \Omega$ such that

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

Because of the convexity hypothesis (3) and the non-negativity of L, we can estimate

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

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

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

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

and as $\varepsilon \to 0$ we achieve the desired inequality (3.2). $\square$

Before stating the following corollary we recall Rellich's theorem (see Theorem 1.10) which provides the compactness of the inclusion $W^{1,1}(\Omega) \subset L^1(\Omega)$ whenever $\Omega \subset \mathbb{R}^n$ is an open, bounded and regular set.

Corollary 3.3. *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded and regular set and let L be a Borel Lagrangian satisfying hypotheses (2), (3) from Theorem 3.2 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

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

admits a solution.

Proof. It is a classical application of the direct method of Calculus of Variations, where hypothesis $(1')$ provides the sequential relative compactness of sublevels $\{F \leq t\}$ 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 3.2. $\square$

At this point one could ask whether the convexity assumption in Theorem 3.2 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 2.5 fits again. Let us define a weaker, although less transparent, convexity condition, introduced by Morrey.

Definition 3.4 (Quasiconvexity)**.** *A continuous function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ 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*

$$\fint_\Omega F(A + \nabla\varphi)\, dx \geq F(A) \qquad \forall\, \varphi \in C_c^\infty(\Omega; \mathbb{R}^m)\ . \tag{3.3}$$

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

Remark 3.5. Obviously we can replace the left-hand side in (3.3) with the quantity $\fint_{\{\nabla\varphi\neq 0\}} F(A + \nabla\varphi)\, dx$: this follows from the equality

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

This proves that the dependence from Ω of this notion is only seeming. Another way to see this relies on the observation that whenever (3.3) is valid for Ω, then:

- it is valid for every $\Omega' \subset \Omega$, thanks to the previous observation;
- it is valid for $x_0 + \lambda\Omega$, for $x_0 \in \mathbb{R}^n$ and $\lambda > 0$, considering the transformation $\varphi(x) \mapsto \varphi(x_0 + \lambda x)/\lambda$.

Finally, a simple approximation argument gives

$$\fint_\Omega F(A + \nabla\varphi)\, dx \geq F(A) \qquad \forall\, \varphi \in C^1_c(\Omega; \mathbb{R}^m)\ . \tag{3.4}$$

The definition of quasiconvexity is related to Jensen's inequality, which we briefly recall here.

Theorem 3.6 (Jensen)**.** *Let us consider a probability measure μ on a convex domain $X \subset \mathbb{R}^p$, with $\int_X |y|\, d\mu(y) < \infty$, and a convex, lower semicontinuous function $F : X \to \mathbb{R} \cup \{+\infty\}$. Then*

$$\int_X F(y)\, d\mu(y) \geq F\left(\int_X y\, d\mu(y)\right)\ . \tag{3.5}$$

Notice that the inequality above makes sense: either $F \equiv +\infty$ or it is finite at least one point. In the second case the negative part of F has at most linear growth and the integral in the left hand side makes sense, finite or infinite.

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 $\mathscr{L}^n/\mathscr{L}^n(\Omega)$. If $F : \mathbb{R}^{m\times n} \to \mathbb{R}\cup\{+\infty\}$ is lower semicontinuous and convex, thanks to Jensen's inequality one has

$$\fint_\Omega F(f(x))\, dx = \int_{\mathbb{R}^{m\times n}} F(y)\, d\mu(y) \geq F\left(\int_{\mathbb{R}^{m\times n}} y\, d\mu(y)\right) = F\left(\fint_\Omega f\, dx\right). \tag{3.6}$$

Quasiconvexity should be considered as a weak version of convexity: indeed, if F is convex then the inequality (3.6) holds for all maps f, thanks to Jensen's inequality; on the other hand the condition (3.3) concerns only gradient maps (more precisely gradients of maps coinciding with an affine function on the boundary of the domain). If we go back to the formulation (3.5), we should say that quasiconvexity should be understood as (3.5) for measures μ in $\mathbb{R}^{mn}$ generated by gradient maps.

Proposition 3.7. *Any convex lower semicontinuous function $F : \mathbb{R}^{m\times n} \to \mathbb{R} \cup \{+\infty\}$ is quasiconvex.*

Proof. Fix $\varphi \in C^\infty_c(\Omega; \mathbb{R}^m)$ and consider the law μ of the map $x \mapsto f(x) = A + \nabla\varphi(x)$ with respect to the rescaled Lebesgue measure $\mathscr{L}^n/\mathscr{L}^n(\Omega)$. Since $\nabla\varphi$ is bounded the measure μ has compact support and one has

$$\int_{\mathbb{R}^n} y\, d\mu(y) = \fint_\Omega A + \nabla\varphi(x)\, dx = A\,.$$

From (3.6) we conclude. □

Remark 3.8. The following chain of implications holds:

convexity $\implies$ quasiconvexity $\implies$ $F_{pp}(A)$ satisfies Legendre-Hadamard with $\lambda = 0$.

All these notions are equivalent when either $n = 1$ or $m = 1$; more generally:

- An integration by parts easily yields

$$\int_\Omega (A + \det\nabla\varphi)\, dx = A|\Omega| \qquad \forall\varphi \in C_c^\infty(\Omega; \mathbb{R}^n)\,.$$

  Hence, Example 2.5 provides a quasiconvex function that is not convex when $n = m = 2$, and considering the determinant of a 2×2 minor the example fits also the case $\min\{m, n\} \geq 2$;

- when $\max\{n, m\} \geq 3$ and $\min\{n, m\} \geq 2$, there exist highly nontrivial examples showing that the Legendre-Hadamard condition does not imply quasiconvexity;

- 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 Morrey-Tonelli's theorem. The following Theorem 3.12 confirms this fact.

Definition 3.9 (w^*-convergence in $W^{1,\infty}$)**.** *Let us consider an open set $\Omega \subset \mathbb{R}^n$ and $f_k \in W^{1,\infty}(\Omega)$. We write $f_k \to f$ in w^*-$W^{1,\infty}(\Omega)$ if*

- *$f_k \to f$ uniformly in Ω;*
- *$\|\nabla f_k\|_{L^\infty}$ is uniformly bounded.*

Proposition 3.10. *If $f_k \to f$ in w^*-$W^{1,\infty}(\Omega)$, then $f \in W^{1,\infty}(\Omega)$ 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). Obviously an analogous statement holds for $\mathbb{R}^m$-valued maps.

Before stating Morrrey's lower semicontinuity theorem we give a quick proof of Rademacher's differentiability theorem.

Theorem 3.11 (Rademacher)**.** *Any locally Lipschitz function $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable $\mathscr{L}^n$-a.e. and its differential coincides $\mathscr{L}^n$-a.e. with the weak gradient.*

Proof. Fix a Lebesgue point x_0 of the weak gradient ∇u, namely $\fint_{B_r(x_0)} |\nabla u - L|\, dx \to 0$ as $r \downarrow 0$ for some linear map $L : \mathbb{R}^n \to \mathbb{R}^m$. We shall prove that f is differentiable at x_0 and that the (classical) gradient ∇f at x_0 is equal to L.

First of all, it is easy to see that this property 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 f_r are equi-Lipschitz in $\overline{B}_1$ and equi-bounded (because $f_r(0) = 0$), hence f_r is relatively compact in $C^0(\overline{B}_1)$ as $r \downarrow 0$. Hence, 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 (understanding here gradients as weak gradients!)

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

It follows that $\nabla f_r \to L$ in $L^1(B_1; \mathbb{R}^{m\times n})$, hence $\nabla f_0 = L$ $\mathscr{L}^n$-a.e. in B_1. By the constancy theorem we get $f_0(y) = L(y) + c$ for some constant c, which obviously should be 0 because $f_0(0) = \lim_i f_{r_i}(0) = 0$. $\square$

Theorem 3.12 (Morrey)**.** *Assume that $L : \Omega \times \mathbb{R}^m \times \mathbb{R}^{mn} \to [0,\infty)$ is continuous, and that the functional F in* (3.1) *is lower semicontinuous w.r.t. the w^*-$W^{1,\infty}(\Omega;\mathbb{R}^m)$ convergence at some function u. Then $L(x, u(x), \cdot)$ is quasiconvex at $\nabla u(x)$ for almost every $x \in \Omega$.*
Conversely, under the same assumptions on L, if $L(x, s, \cdot)$ is quasiconvex for all $(x, s) \in \Omega \times \mathbb{R}^m$, then F is lower semicontinuous w.r.t. w^*-$W^{1,\infty}(Q;\mathbb{R}^m)$ convergence.

Proof. (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

$$F_r(v) := \int_Q L(x_0 + ry, u(x_0 + ry) + rv(y), \nabla u(x_0 + ry) + \nabla v(y))\, dy\,.$$

The formal limit as $r \downarrow 0$ of F_r, namely

$$F_0(v) := \int_Q L(x_0, u(x_0), \nabla u(x_0) + \nabla v(y))\, dy$$

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

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

- 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

$$\lim_{r\to 0^+}\sup_h |F_r(v_h)-F_0(v_h)| = 0\,.$$

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 1-periodic function ψ such that $\psi|_Q=\varphi$ and the sequence of highly oscillating (because h^{-1}-periodic) functions

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

which obviously converge uniformly to 0. Since $\nabla v_h(x)=\nabla\psi(hx)$ we have also $v_h \stackrel{*}{\rightharpoonup} 0$ in $W^{1,\infty}(Q;\mathbb{R}^m)$, so that thanks to the lower semicontinuity of F_0 at 0 one has

$$\begin{aligned} H(0) &= F_0(0) \le \liminf_{h\to\infty}\int_Q H(\nabla v_h(x))\,dx = \liminf_{h\to\infty} h^{-n}\int_{Q_h} H(\nabla\psi(y))\,dy \\ &= \int_Q H(\nabla\psi(y))\,dy = \int_Q H(\nabla\varphi(y))\,dy\;, \end{aligned}$$

which is exactly the quasiconvexity property for $L(x,u(x),\cdot)$ at $\nabla u(x_0)$.

(Sufficiency of quasiconvexity) We split the proof in several steps, reducing ourselves to progressively 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 $L\in C^0(\overline{\Omega}\times\mathbb{R}^m\times\mathbb{R}^{mn})$. Since Ω can be written as the disjoint union of half-open disjoint cubes, by the superadditivity of the lim inf we can also assume that $\Omega=Q$ is a n-cube with side length ℓ. We also set

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

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 we need only to consider Lagrangians $L_1(x,p)$ independent of s (just take $L_1(x,p)=L(x,u(x),p)$).

The next step is to reduce ourselves to Lagrangians independent of x. To this aim, consider a modulus of continuity for L_1 in the ball $\overline{B}_M$ and a decomposition of Q in 2^{kn} cubes Q_i with side length $\ell/2^k$ and centers x_i. Then, adding and subtracting $L(x_i,\nabla u_h(x_i))$ and using once more the superadditivity of $\liminf$, yields

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

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

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

we obtain

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

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

Hence, we are led to show the lower semicontinuity property for Lagrangians $L_2(p)=L_1(x_i,p)$ independent of x. In this proof 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

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

and quasiconvex functions g satisfy the Legendre-Hadamard condition, hence $g(p)$ as function of p_i^α is convex.

Now, let us consider a quasiconvex Lagrangian $L_2(p)$. 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. We can apply the quasiconvexity inequality to $\varphi = (u_h - u)\psi$ and the local Lipschitz property with $R(p) = L_2(p) - L_2(0)$ to get

$$\begin{aligned} R(A) &\leq \fint_\Omega R((1-\psi)A + \psi \nabla u_h + (u_h - u) \otimes \nabla \psi)\, dx \\ &\leq C(|A| + \|\nabla u_h\|_\infty) \fint_\Omega (1-\psi)\, dx + C\|\nabla \psi\|_\infty \fint_\Omega |u_h - u|\, dx + \fint_\Omega R(\nabla u_h)\,, \end{aligned}$$

so that passing to the limit first as $h \to \infty$ and then as $\psi \uparrow 1$ gives the result.

Finally, we consider the general case, using Rademacher's theorem and a blow-up argument. Assume that the $\liminf \int_\Omega L_2(\nabla u_h)\, dx$ is a limit, that we call L, 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 $C_c(\Omega)$, to some measure μ. Recall that the evaluations on compact sets K and open sets A are respectively upper and lower semicontinuous w.r.t. weak convergence, i.e.

$$\mu(K) \geq \limsup_{h\to\infty} \mu_h(K), \qquad \mu(A) \leq \liminf_{h\to\infty} \mu_h(A)\,. \tag{3.7}$$

In particular $\mu(\Omega) \leq L$, 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, suffices to show that

$$\liminf_{r\downarrow 0} \frac{\mu(\overline{B}_r(x_0))}{\omega_n r^n} \geq L_2(\nabla u(x_0)) \qquad \text{for a.e. } x_0 \in \Omega\,. \tag{3.8}$$

We shall prove this property at any differentiability point x_0 of u. To this aim, let $r_i \to 0$ be a sequence on which the liminf is achieved, and $\varepsilon > 0$. For any i we can choose $h_i \geq i$ so large that

$$\int_{B_{r_i}(x_0)} L_2(\nabla u_{h_i})\, dx \leq \mu(\overline{B}_{r_i}(x_0)) + \frac{r_i^n}{i}\,, \qquad \fint_{B_{r_i}(x_0)} |u_{h_i} - u|\, dx \leq \frac{r_i}{i}\,. \tag{3.9}$$

Now, rescale as follows

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

to obtain functions v_i satisfying

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

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

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

□

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 3.13 (Acerbi-Fusco)**.** *Suppose that a Borel Lagrangian $L(x, s, p)$ is continuous in (s, p) and satisfies*

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

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 w.r.t. the weak $W^{1,\alpha}(\Omega; \mathbb{R}^m)$-topology.

4 Regularity Theory

We begin studying the local behaviour of (weak) solutions of the system of equations

$$\begin{cases} -\partial_\alpha\big(A_{ij}^{\alpha\beta}\partial_\beta u^j\big) = f_i - \partial_\alpha F_i^\alpha & \qquad i = 1, \ldots, m \\ u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m) \end{cases} \tag{4.1}$$

with $A_{ij}^{\alpha\beta} \in L^\infty(\Omega)$, $f_i \in L^2_{\mathrm{loc}}(\Omega)$ and $F_i^\alpha \in L^2_{\mathrm{loc}}(\Omega)$. From now on we shall use $|\cdot|$ for the Hilbert-Schmidt norm of matrices and tensors, even though some estimates would still be valid with the (smaller) operator norm.

Theorem 4.1 (Caccioppoli-Leray inequality)**.** *If the Borel coefficients $A_{ij}^{\alpha\beta}$ satisfy the Legendre condition $(L)_\lambda$ with $\lambda > 0$ and*

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

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

$$c\int_{B_{R/2}(x_0)} |\nabla u|^2\,dx \le R^{-2}\int_{B_R(x_0)} |u(x) - k|^2\,dx + R^2\int_{B_R(x_0)} |f(x)|^2\,dx + \int_{B_R(x_0)} |F(x)|^2\,dx \ . \tag{4.2}$$

Before proceeding to the proof, some remarks are in order.

Remark 4.2. (1) The validity of (4.2) for all $k \in \mathbb{R}^m$ depends on the translation invariance of the PDE. Moreover, the inequality (and the PDE as well) has a natural scaling invariance: if we think of u as an adimensional quantity, then all sides have dimension length^{n-2}, because $f \sim$ length^{-2} and $F \sim$ length^{-1}.

(2) The Caccioppoli-Leray inequality is meaningful because for a general function u the gradient ∇u can not be controlled by the variance of u! Precisely because of this fact we can expect that several useful (regularity) informations can be drawn from it. We will see indeed that CL inequalities are very "natural" and useful in the context of regularity theory.

Remark 4.3 (Absorbtion scheme)**.** In the regularity theory it often happens that one can estimate, for some $\alpha < 1$,

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

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

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

for $p = 1/\alpha$ one obtains

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

Now, if we choose $\varepsilon = \varepsilon(p)$ sufficiently small, so that $\dfrac{\varepsilon^p}{p} \leq \dfrac{1}{2}$, we get

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

Let us prove Theorem 4.1.

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

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

where $\eta \in C_c^{\infty}(B_R)$, $\eta \equiv 1$ in $B_{R/2}$, $0 \leq \eta \leq 1$ and $|\nabla \eta| \leq 4/R$.
Since u solves (4.1), we have that

$$\int A\nabla u \nabla \Phi - \int f\Phi - \int F \cdot \nabla \Phi = 0 \tag{4.3}$$

where integrations are understood to be on B_R. Moreover

$$\nabla \Phi = \eta^2 \nabla u + 2\eta u \otimes \nabla \eta \, , \tag{4.4}$$

so completing (4.3) with (4.4) we obtain

$$\int \eta^2 A \nabla u \, \nabla u + 2 \int \eta A \nabla u \, (u \otimes \nabla \eta) - \int f\Phi - \int \eta^2 F \nabla u - 2 \int \eta F \, (u \otimes \nabla \eta) = 0 \ . \tag{4.5}$$

Let us deal with each addendum separately.

- By the Legendre condition

$$\int_{B_R} \eta^2 A_{ij}^{\alpha\beta} \partial_\alpha u^i \partial_\beta u^j \geq \lambda \int_{B_R} \eta^2 |\nabla u|^2 \ .$$

- We have

$$\begin{aligned} 2\int \eta A \nabla u \, (u \otimes \nabla \eta) &\leq 2\int \eta |A||\nabla u||u||\nabla \eta| \leq \frac{8\Lambda}{R} \int (\eta |\nabla u|)\, |u| \\ &\leq \frac{4\Lambda\varepsilon}{R} \int \eta^2 |\nabla u|^2 + \frac{4\Lambda}{R\varepsilon} \int |u|^2 \,, \end{aligned}$$

  where the first estimate is due to 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.

- By the Young inequality

$$\int_{B_R} \eta^2 |f_i u^i| \leq \int_{B_R} |f||u| \leq \frac{1}{2R^2} \int_{B_R} |u|^2 + \frac{R^2}{2} \int_{B_R} |f|^2 \ .$$

- Similarly, since $\eta^4 \leq \eta^2$, one has

$$\int \eta^2 |F_i^\alpha \partial_\alpha u^i| \leq \frac{\lambda}{4} \int \eta^2 |\nabla u|^2 + \frac{1}{\lambda} \int |F|^2 \ .$$

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

$$2\int_{B_R} \eta |F_i^\alpha u^i \partial_\alpha \eta| \leq \frac{8}{R} \int_{B_R} |F||u| \leq 4 \int_{B_R} |F|^2 + \frac{4}{R^2} \int_{B_R} |u|^2 \ .$$

From (4.5) it follows that

$$\begin{aligned} \lambda \int_{B_R} \eta^2 |\nabla u|^2 &\leq \int_{B_R} \eta^2 A \nabla u \, \nabla u \\ &= -2\int_{B_R} \eta A \nabla u \, (u \otimes \nabla \eta) + \int_{B_R} f \Phi + \int_{B_R} \eta^2 F \nabla u + 2\int_{B_R} \eta F \, (u \otimes \nabla \eta) \\ &\leq \frac{4\Lambda\varepsilon}{R} \int_{B_R} \eta^2 |\nabla u|^2 + \frac{\lambda}{4} \int_{B_R} \eta^2 |\nabla u|^2 \qquad (4.6) \\ &+ \left(\frac{4\Lambda}{R\varepsilon} + \frac{1}{2R^2} + \frac{4}{R^2} \right) \int_{B_R} |u|^2 + \frac{R^2}{2} \int_{B_R} |f|^2 + \left(\frac{1}{\lambda} + 4 \right) \int_{B_R} |F|^2 \,. \end{aligned}$$

By choosing ε sufficiently small, in such a way that $4\Lambda\varepsilon/R = \lambda/4$, one can absorb line (4.6), and the thesis follows noticing that

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

□

Remark 4.4 (Widman's hole-filling technique)**.** There exists a sharper version of the Caccioppoli-Leray inequality, let us illustrate it in the simpler case $f = 0$, $F = 0$. Indeed, since

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

following the proof of Theorem 4.1 one obtains

$$\int_{B_{R/2}} |\nabla u(x)|^2 \, dx \leq \frac{c}{R^2} \int_{B_R \setminus B_{R/2}} |u(x) - k|^2 \, dx \ . \tag{4.7}$$

Setting $k := \fint_{B_{R/2}} u$, the Poincaré inequality in the domain $B_1 \setminus \overline{B}_{1/2}$ and a scaling argument give

$$\int_{B_{R/2}} |\nabla u(x)|^2 \, dx \leq c \int_{B_R \setminus B_{R/2}} |\nabla u(x)|^2 \, dx \ . \tag{4.8}$$

Adding to (4.8) the term $c \int_{B_{R/2}} |\nabla u(x)|^2 \, dx$, we get

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

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

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

Iterating (4.7) and interpolating (i.e. considering the integer k such that $2^{-k-1}R < r \leq 2^{-k}R$), it is not difficult to get

$$\int_{B_r} |\nabla u(x)|^2 \, dx \leq 2^\alpha \left(\frac{r}{R}\right)^\alpha \int_{B_R} |\nabla u(x)|^2 \, dx \qquad 0 < r \leq R \tag{4.9}$$

with $(1/2)^\alpha = \theta$, i.e. $\alpha = \ln_2(1/\theta)$. When $n = 2$, this implies that $u \in C^{0,\alpha/2}$, as we will see.

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

Definition 4.5 (Reverse Hölder's inequality)**.** *Let $\alpha \in (1, \infty)$. A non-negative function $f \in L^{\alpha}_{\mathrm{loc}}(\Omega)$ satisfies a* reverse Hölder's inequality *with exponent α if there exists a constant $c > 0$ such that*

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

For the sake of completeness, we now recall the Sobolev inequalities. 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. 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.

Theorem 4.6 (Sobolev inequalities)**.** *Let Ω be either the whole space $\mathbb{R}^n$ or a bounded regular domain.*

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

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

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

- *If $p > n$,*

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

Remark 4.7. Combining the Poincaré inequality with the inequality

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

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

$$\left(\int_{B_1} |v - \overline{v}|^{p^*} \right)^{1/p^*} \le c \left(\int B_1 |\nabla v|^p \right)^{1/p} ,$$

for some constant c depending on c_I and c_P. By a rescaling argument this gives

$$\left(\fint_{B_R} |u - \overline{u}|^{p^*} \right)^{1/p^*} \le cR \left(\fint_{B_R} |\nabla u|^p \right)^{1/p} \ . \tag{4.10}$$

[3]The result is basically sharp, as the example of $(-\ln|x|)^{\alpha} \in W^{1,n}(B_1)$ for $n > 1$ and $\alpha \in (0, 1 - 1/n)$ shows.

If u solves (4.1) with $f = F = 0$, combining (4.10) with the CL inequality when $p^* = 2$ (that is, $p = 2n/(n+2) < 2$), we write

$$c_{CL} R \left(\fint_{B_{R/2}} |\nabla u|^2 \right)^{1/2} \leq \left(\fint_{B_R} |u - \overline{u}|^2 \right)^{1/2} \leq cR \left(\fint_{B_R} |\nabla u|^p \right)^{1/p} .$$

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

$$\left(\fint_{B_{R/2}} |\nabla u|^2 \right)^{1/2} \leq C \left(\fint_{B_R} |\nabla u|^p \right)^{1/p} .$$

Remark 4.8 (Embedding for higher order Sobolev spaces)**.** Recall first that higher order Sobolev spaces $W^{k,p}(\Omega)$ are recursively defined ($k \geq 1$ integer, $1 \leq p \leq \infty$)

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

Together with the Sobolev embedding in Theorem 4.6, with $p > n$, another way to gain continuity is using the Sobolev spaces $W^{k,p}$, with k sufficiently large. In fact, we can arbitrarily expand the chain

$$W^{2,p} \hookrightarrow W^{1,p^*} \hookrightarrow L^{(p^*)^*} .$$

Iterating the $*$ operation k-times we get

$$\frac{1}{p^{*\cdots*}} = \frac{1}{p} - \frac{k}{n} ,$$

therefore if $k > [\frac{n}{p}]$ (where $[\cdot]$ denotes the integer part) we obtain $W^{k,p} \subset C^{0,\alpha}$ with any positive α with $\alpha < 1 - n/p + [n/p]$.

4.1 Nirenberg method

For the moment let us consider a (local) solution u to the Poisson equation

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

Our aim is to prove that u belongs to $H^2_{\text{loc}}(\Omega)$.

When we talk about an *a priori* estimate, we mean this argument: suppose that we already know that $\frac{\partial u}{\partial x_i} \in H^1_{\text{loc}}(\Omega)$, then it is not difficult to check (using the fact that higher order weak derivative, as well as classical ones, commute) that this function solves

$$-\Delta \left(\frac{\partial u}{\partial x_i} \right) = \frac{\partial f}{\partial x_i}$$

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

$$\int_{B_{R/2}(x_0)} \left|\nabla\left(\frac{\partial u}{\partial x_i}\right)\right|^2 \le \frac{c}{R^2}\int_{B_R(x_0)} \left|\frac{\partial u}{\partial x_i}\right|^2 + \int_{B_R(x_0)} |f|^2 \,. \tag{4.11}$$

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, estimate (4.11) applies to $u * \rho_\varepsilon$ with $f * \rho_\varepsilon$ in place of f, since $u * \rho_\varepsilon$ satisfies

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

Passing to the limit as $\varepsilon \to 0$ we obtain that $u \in H^2_{\mathrm{loc}}(\Omega)$ and that the same inequality holds for u, starting from the assumption $u \in H^1_{\mathrm{loc}}(\Omega)$.

The situation is much more complex when the coefficients $A^{\alpha\beta}_{ij}$ are not constant and therefore differentiation provides a worse right hand side in the PDE. Nirenberg's idea is to introduce partial discrete derivatives

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

Remark 4.9. Some basic properties of differentiation are still true and easy to prove:

- (sort of) Leibniz property

$$\Delta_{h,i}(ab) = (\tau_{h,i}a)\Delta_{h,i}b + (\Delta_{h,i}a)b \ ;$$

- integration by parts (relying ultimately on the translation invariance of Lebesgue measure)

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

In the next lemma we show that membership to $W^{1,p}$ with $p > 1$ can be characterized in terms of uniform L^p bounds on $\Delta_{h,i}u$; notice that one implication was already established in (1.18).

Lemma 4.10. *Consider $u \in L^p_{\mathrm{loc}}(\Omega)$, with $1 < p \le \infty$ and fix $i \in \{1,\dots,n\}$. The partial derivative $\frac{\partial u}{\partial x_i}$ belongs to $L^p_{\mathrm{loc}}(\Omega)$ if and only if*

$$\forall\,\Omega' \Subset \Omega \quad \exists\, c(\Omega') \quad \text{s.t.} \quad \left|\int_{\Omega'} (\Delta_{h,i}u)\,\varphi\right| \le c(\Omega')\|\varphi\|_{L^{p'}(\Omega')} \quad \forall\,\varphi \in C^1_c(\Omega')\ ,$$

with $1/p + 1/p' = 1$.

Proof. The first implication has been proved in (1.18), because we know that $\Delta_{h,i}u$ is bounded in $L^p_{\mathrm{loc}}(\Omega)$ when $h \to 0$, so we can conclude with Hölder's inequality.

Now fix $\Omega' \Subset \Omega$,

$$\left|\int_{\Omega'} u\frac{\partial\varphi}{\partial x_i}\,dx\right| = \left|\lim_{h\to 0}\int_{\Omega'} u\Delta_{-h,i}\varphi\,dx\right| = \left|-\lim_{h\to 0}\int_{\Omega'} (\Delta_{h,i}u)\,\varphi\,dx\right| \le c(\Omega')\|\varphi\|_{L^{p'}(\Omega')}\ ;$$

because of the duality relation between $L^p(\Omega')$ and $L^{p'}(\Omega')$, the weak derivative $\partial_{x_i}u$ exists and belongs to $L^p_{\mathrm{loc}}(\Omega)$. $\square$

Let us see how Lemma 4.10 contributes to regularity theory, still in the simplified case of the Poisson equation. Suppose $f \in H^1_{\mathrm{loc}}(\Omega)$ in the Poisson equation, then translation invariance and linearity allow us to write

$$-\Delta\tau_{h,i}u = \tau_{h,i}f \quad\Longrightarrow\quad -\Delta(\Delta_{h,i}u) = \Delta_{h,i}f\ .$$

Thanks to Lemma 4.10, $\Delta_{h,i}f$ is bounded in $L^2_{\mathrm{loc}}(\Omega)$, then by the Caccioppoli-Leray inequality $|\nabla\Delta_{h,i}u|$ is bounded in $L^2_{\mathrm{loc}}(\Omega)$.
As $\Delta_{h,i}(\nabla u) = \nabla\Delta_{h,i}u$ is bounded in $L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^n)$, thanks to Lemma 4.10 again (applied componentwise) we get

$$\frac{\partial}{\partial x_i}(\nabla u) \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^n)\ .$$

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

Theorem 4.11. *Let Ω be an open regular domain in $\mathbb{R}^n$. Consider a function $A \in C^{0,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^{m^2\times n^2})$ such that $A(x) := A^{\alpha\beta}_{ij}(x)$ satisfies the Legendre-Hadamard condition for a given constant $\lambda > 0$. Then, for every choice of subsets $\Omega' \Subset \Omega'' \Subset \Omega$ there exists a constant $c := c(\Omega',\Omega'',A)$ such that*

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

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

$$-\mathrm{div}(A\nabla u) = f - \mathrm{div}(F)$$

with data $f \in L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ and $F \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^{m\times n})$.

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

Remark 4.12. Although the thesis 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

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

for any $x_0 \in \Omega'$. The general result can be easily obtained by a compactness and covering argument.

Notice also that the statement as given is redundant, since the term $\mathrm{div}(F)$ can always be absorbed into f. We will see however that the optimal estimate 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 changing f). In addition, we prove the result under the stronger assumption that the Legendre condition with constant λ holds uniformly in Ω.

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

$$\int_\Omega A\nabla u \nabla\varphi\, dx = \int_\Omega f\varphi\, dx \tag{4.12}$$

for all $\varphi \in C_c^\infty(\Omega; \mathbb{R}^m)$. If we apply it to the test function $\tau_{-h}\varphi$ with $|h| \ll 1$ and we do a change of variable, we find

$$\int_\Omega \tau_h(A\nabla u)\nabla\varphi\, dx = \int_\Omega \tau_h f\varphi\, dx\,. \tag{4.13}$$

Subtracting (4.12) to equation (4.13) and dividing by h, we get (thanks to the discrete Leibniz property)

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

which is nothing but the weak form of the equation

$$-\mathrm{div}((\tau_h A)\nabla v) = f' - \mathrm{div}(G) \tag{4.14}$$

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 will be to use the Caccioppoli-Leray inequality. However, a direct application of the CL inequality would lead to an estimate having the L^2 norm of f' on the right hand side, and we know from Lemma 4.10 that this norm can be uniformly bounded in h only if $f \in H^1_{\mathrm{loc}}$. Hence, rather than applying CL directly, we will 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 cut-off 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 insert in (4.14) the test function $\Phi := \eta^2\Delta_h u = \eta^2 v$ with $|h| < R/2$.
Using Young inequality as in Theorem 4.1 (see (4.6)), we get

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

with Λ depending only on A. As in the proof of Theorem 4.1, we absorb the term with $\|\eta\nabla v\|^2_{L^2(B_R)}$ in the left side of the inequality, so that, up to some constant $c > 0$ depending on (λ, Λ, R), we get

$$
c\int_{B_{2R}}\eta^2|\nabla v|^2\,dx \le \int_{B_R}|v|^2\,dx + \int_{B_R}|G|^2\,dx + \int\eta^2 v\Delta_h f\,dx\ . \tag{4.15}
$$

We now study each term of (4.15) separately. Firstly

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

by means of (1.18). 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}: 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 Young inequality, we can write

$$
\left|\int_{B_R}\eta^2 v\Delta_h f\,dx\right| \le \tilde{\varepsilon}\int_{B_R}|\Delta_{-h}(\eta^2 v)|^2\,dx + \frac{1}{\tilde{\varepsilon}}\int_{B_R}|f|^2\,dx\,. \tag{4.16}
$$

The first term in the right hand side of (4.16) can be estimated with (since $|\nabla\eta|^2 \le 64/R^2$)

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

so that choosing ε sufficiently small and using the inequality $\eta^4 \le \eta^2$ we can absorb the first term and use once more the CL inequality to estimate $\int_{B_{R+h}}|v|^2\,dx$.
The term involving the integral $\|G\|^2_{L^2(B_R)}$ 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 4.10 and then letting $h \to 0$ in the estimate involving $v = \Delta_h u$). $\square$

Remark 4.13. 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 behaviour 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 later on.

5 Decay estimates for systems with constant coefficients

Our next target towards the development of a regularity theory is now to derive some decay estimates for constant coefficients differential operators. Let $A = A_{ij}^{\alpha\beta}$ be a matrix satisfying the Legendre-Hadamard condition for some $\lambda > 0$, let $\Lambda = |A|$ and consider the problem

$$\begin{cases} -\mathrm{div}(A\nabla u) = 0 \\ \\ u \in H^1_{\mathrm{loc}}\left(\Omega; \mathbb{R}^m\right). \end{cases}$$

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

$$\int_{B_r(x_0)} |u|^2 \, dx \le c(n,\lambda,\Lambda) \left(\frac{r}{R}\right)^n \int_{B_R(x_0)} |u|^2 \, dx \tag{5.1}$$

$$\int_{B_r(x_0)} |u - u_{r,x_0}|^2 \, dx \le c(n,\lambda,\Lambda) \left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u - u_{R,x_0}|^2 \, dx \tag{5.2}$$

with $c(n,\lambda,\Lambda)$ depending only on n, λ and Λ.
Here u_{s,x_0} denotes as usual the mean value of u on $B_s(x_0)$.

Proof of (5.1)**.** By a standard rescaling argument, it is enough to study the case $R = 1$. For the sequel, let k be the smallest integer such that $k > \left[\frac{n}{2}\right]$ (and consequently $H^k \hookrightarrow C^{0,\alpha}$ with $\alpha = k - \left[\frac{n}{2}\right]$). First of all, by the Caccioppoli-Leray inequality, we have that

$$\int_{B_{1/2}(x_0)} |\nabla u|^2 \, dx \le c_1 \int_{B_1(x_0)} |u|^2 \, dx\,.$$

Now, for any $\alpha \in \{1, 2, \ldots, n\}$, we know that $\partial_\alpha u \in H^{1,2}_{\mathrm{loc}}(\Omega)$ by Theorem 4.11, and since the matrix A has constant coefficients it will solve the same equation. Hence, we can iterate the argument in order to get an estimate having the form

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

for some constant $c_k > 0$. Consequently, thanks to our choice of the integer k, we can find a constant κ 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 better to consider 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 $c(\lambda, \Lambda) \geq \kappa \omega_n$. If $r \in (2^{-k}, 1)$, then it is clear that $\int_{B_r(x_0)} |u|^2 \, dx \leq \int_{B_1(x_0)} |u|^2 \, dx$ and so, since we have a lower bound for r, we just need to choose $c(\lambda, \Lambda)$ such that $c(\lambda, \Lambda) \geq 2^{kn}$.

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

Proof of (5.2)**.** Again, it is useful to study two cases separately. If $r \leq R/2$, then by the Poincaré inequality there exists a constant $c(n)$ such that

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

and so

$$\begin{aligned} \int_{B_r(x_0)} |u - u_{x_0,r}|^2 \ dx \ &\leq \ c(n) r^2 \left(\frac{r}{R/2} \right)^n \int_{B_{R/2}(x_0)} |\nabla u|^2 \, dx \\ &\leq \ c(n, \lambda, \Lambda) \left(\frac{r}{R/2} \right)^{n+2} \int_{B_R(x_0)} |u - u_{R,x_0}|^2 \ dx \end{aligned}$$

respectively by the previous result applied to the gradient ∇u and finally by the Cacciop-poli-Leray inequality. For the case $R/2 < r \leq R$ we need to use the following fact, that will be discussed below: the mean value $u_{x_0,r}$ is a minimizer for the function

$$m \longmapsto \int_{B_r(x_0)} |u - m|^2 \ dx \, . \tag{5.3}$$

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

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

Let us go back to the study of

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

for $1 \le p < \infty$ and $u \in L^p(\Omega)$ 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, for $p \neq 2$. Of course

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

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

$$\int_\Omega |u-u_\Omega|^p \, dx \le 2^p \int_\Omega |u-m|^p \, dx \, . \tag{5.4}$$

Since the problem is clearly translation invariant, it is sufficient to prove inequality (5.4) 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.

6 Regularity up to the boundary

Let us first consider a simple special case. Suppose we have to deal with the problem

$$\begin{cases} -\Delta u = f \\ \\ u \in H_0^1(R) \, , \end{cases} \tag{6.1}$$

where $R := (-a,a)^{n-1} \times (0,a)$ is a rectangle in $\mathbb{R}^n$ with sides parallel to the coordinate axes. Let us use coordinates $x = (x', x_n)$ with $x' \in \mathbb{R}^{n-1}$ and assume $f \in L^2(R)$. The rectangle $R' = (-a/2, a/2)^{n-1} \times (0, a/2)$ is not relatively compact in R, nevertheless via Nirenberg's method we may find estimates of the form

$$\int_{R'} |\partial_{x_s} \nabla u|^2 \, dx \leq \frac{c}{a^2} \int_R |\nabla u|^2 \, dx + c \int_R |f|^2$$

for $s = 1, 2, \ldots, n-1$, provided $u = 0$ on $R \cap \{x_n = 0\}$. Indeed, we are allowed to use test functions $\varphi = \eta^2 \Delta_{h,s} u$, where the support of η can touch the hyperplane $\{x_n = 0\}$ (because of the homogeneous Dirichlet boundary condition on u). The equation (6.1) may be rewritten as

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

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

$$\frac{\partial u}{\partial x_i} \qquad i = 1, \ldots, n-1$$

have 0 trace on $\{x_n = 0\}$, and this is crucial for the iteration of this argument with higher order derivatives (see also Theorem 6.2 below).

Now we want to use this idea in order to study the regularity up to the boundary for problems like

$$\begin{cases} -\mathrm{div}(A \nabla u) = f + \mathrm{div} F \\ \\ u \in H_0^1(\Omega; \mathbb{R}^m) \end{cases}$$

under the following hypotheses:

- $f \in L^2(\Omega; \mathbb{R}^m)$;
- $F \in H^1(\Omega; \mathbb{R}^{m \times n})$;
- $A \in C^{0,1}(\Omega; \mathbb{R}^{m^2 \times n^2})$;
- $A(x)$ satisfies the Legendre-Hadamard condition uniformly in Ω;
- Ω has a C^2 boundary, in the sense that it is, up to a rigid motion, locally the subgraph of a C^2 function.

Theorem 6.1. *Under the previous assumptions, the function u belongs to $H^2(\Omega; \mathbb{R}^m)$ and*

$$\|u\|_{H^2} \leq c(\Omega, A, n) \big[\|f\|_2 + \|F\|_{H^1} \big]$$

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 neighbourhood U of x_0 in Ω such that $u \in H^2(U)$. Without loss of generality we assume $x_0 = 0$. We consider first the case of a flat boundary.

Step 1. (Flat boundary) By applying Nirenberg's method as described above for the case of the constant coefficient operator $-\Delta$ we get $\partial_{x_\alpha} u^i \in H^1(R')$ for $\alpha = 1, 2, \dots, n-1$ and $i = 1, 2, \dots, m$, and

$$-\mathrm{div}\left(A\nabla\left(\frac{\partial u}{\partial x_\alpha}\right)\right) = \frac{\partial f}{\partial x_\alpha} + \mathrm{div}\left(\frac{\partial F}{\partial x_\alpha}\right) + \mathrm{div}\left(\frac{\partial A}{\partial x_\alpha}\nabla u\right) . \tag{6.2}$$

Anyway, we cannot include in the previous conclusion the second derivatives $\partial^2_{x_n x_n} u^i$ and here we really need to refine a bit the strategy seen above for the Poisson equation. Actually, this is not complicated because the equation readily implies that $\partial_{x_n}(A^{nn}_{ij}\partial_{x_n}u^j) \in L^2(R')$ for any $i \in \{1, 2, \dots, m\}$. Formally this implies, by the Leibniz rule, that $A^{nn}_{ij}\partial^2_{x_n x_n}u^j$ belong to $L^2(R')$; this is formal because one of the factors is only a distribution (not yet a function). To make this rigorous, we use the difference quotients in the x_n direction and the discrete Lebniz rule: since by Lemma 4.10 the difference quotients $\Delta_h(A^{nn}_{ij}\partial_{x_n}u^j)$ have uniformly bounded L^2 norm in $R'_h = \{x \in R:\ \mathrm{dist}(x, \partial R') > h\}$, we obtain that the same is true for $A^{nn}_{ij}\Delta_h\partial_{x_n}u^j$. Since the matrix A^{nn}_{ij} is invertible with $\mathrm{det}A^{nn}_{ij} \geq \lambda^m$ (as a consequence of the Legendre-Hadamard condition) we get

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

which gives $\partial^2_{x_n x_n}u^j \in L^2(R')$.

Step 2. (Straightnening of the boundary) There exist $h \in C^2(\mathbb{R}^{n-1})$ and $V = (-b, b)^n$ such that (up to a rigid motion, choosing the hyperplane $\{x_n = 0\}$ as the tangent one to $\partial\Omega$ at 0)

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

Consequently, we can define the change of variables $x'_n = x_n - h(x')$ and the function $H(x', x_n) = (x', x_n - h(x')))$ that maps $\Omega \cap V$ onto $H(\Omega \cap V)$, which contains a rectangle $R = (-a, a)^{n-1} \times (0, a)$. We set $\Omega' := H^{-1}(R) \subset V \cap \Omega$ and $U := H^{-1}(R')$, with $R' = (-a/2, a/2)^{n-1} \times (0, a/2)$.

It is clear that H is invertible and, called G its inverse, both H and G are C^2 functions. Moreover ∇H is a triangular matrix with $\det(\nabla H) = 1$. Besides, the maps G and H induce isomorphisms between both H^1 and H^2 spaces (via change of variables in the definition of weak derivative, as we will see in a moment). To conclude, it suffices to show that $v = u \circ G$ belongs to $H^2(R'; \mathbb{R}^m)$. To this aim, we check that v solves in R the PDE

$$\begin{cases} -\mathrm{div}(\widetilde{A}\nabla v) = \widetilde{f} + \mathrm{div}\widetilde{F} \\ \\ v = 0 \quad \text{on } \{x'_n = 0\} \cap R \end{cases} \tag{6.3}$$

where of course the boundary condition has to be interpreted in the weak sense and

$$\widetilde{f} = f \circ G, \quad \widetilde{F} = (F \cdot DH) \circ G, \quad \widetilde{A} = \left[DH \cdot A \cdot (DH)^t\right] \circ G$$

(here contractions are understood with respect to the greek indices, the only ones involved in the change of variables, see (6.4) below). These formulas can be easily derived by 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

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

just letting $x = G(y)$, but then $\det(\nabla G) = 1$ and we can set $\varphi = \psi \circ H$ so that equivalently $\psi = \varphi \circ G$ and

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

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

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

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 ask whether the conditions on A (for instance, the Legendre-Hadamard condition) still hold true for $\widetilde{A}$. This is the case and we can verify it directly by means of the expression of $\widetilde{A}$ above. In fact,

$$\widetilde{A}_{ij}^{\alpha'\beta'} = \left(\frac{\partial H^{\alpha'}}{\partial x_\alpha}A_{ij}^{\alpha\beta}\frac{\partial H^{\beta'}}{\partial x_\beta}\right) \circ G \tag{6.4}$$

and so, for any $\widetilde{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$

$$\begin{aligned}
\widetilde{A}_{ij}^{\alpha'\beta'}(y)\widetilde{a}_{\alpha'}\widetilde{a}_{\beta'}b^ib^j &= A_{ij}^{\alpha\beta}(G(y))\left(\frac{\partial H^{\alpha'}}{\partial x_\alpha}(G(y))\widetilde{a}_{\alpha'}\right)\left(\frac{\partial H^{\beta'}}{\partial x_\beta}(G(y))\widetilde{a}_{\beta'}\right)b^ib^j \\
&\geq \lambda\,|\nabla H(G(y))\widetilde{a}|^2\,|b|^2 \geq \lambda\left|(\nabla H(G(y)))^{-1}\right|^{-2}|\widetilde{a}|^2\,|b|^2
\end{aligned}$$

since clearly

$$|\widetilde{a}|^2 \leq \left|(\nabla H(G(y)))^{-1}\right|^2|\nabla H(G(y))\widetilde{a}|^2\,.$$

Hence, $\widetilde{A}$ satisfies the Legendre-Hadamard condition for an appropriate constant $\lambda' > 0$ depending on λ and H, and of course $\widetilde{A} \in C^{0,1}(R)$.
Through this transformation of the domain, we can finally apply Step 1 and find that $v \in H^2(R')$. Coming back to the original variables we obtain the H^2 regularity of u.
$\square$

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

Theorem 6.2. *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 idea consists in differentiating the starting equation with respect to each fixed direction to get an equation having the form of (6.3), as in (6.2), provided we set $\widetilde{F} = \frac{\partial F}{\partial x_\alpha} + \frac{\partial A}{\partial x_\alpha} \nabla u$.

7 Interior regularity for nonlinear problems

So far, we have just dealt with linear problems and the richness 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 see now that Nirenberg's technique is particularly appropriate to deal also with nonlinear PDE's, as those arising from Euler-Lagrange equations of non-quadratic functionals.

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

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

(ii) F satisfies a uniform Legendre condition, i.e. $\partial_{p_i^\alpha} \partial_{p_j^\beta} F(p) \xi_i^\alpha \xi_j^\beta \geq \lambda |\xi|^2$ for all $\xi \in \mathbb{R}^{m \times n}$, for some $\lambda > 0$ independent of $p \in \mathbb{R}^{m \times n}$.

Let $B_i^\alpha := \frac{\partial F}{\partial p_i^\alpha}$ and $A_{ij}^{\alpha\beta} := \frac{\partial^2 F}{\partial p_i^\alpha \partial p_j^\beta}$ 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 $u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ be a local minimizer of the functional

$$w \longmapsto I(w) := \int_\Omega F(\nabla w)\, dx \,.$$

The implication

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

is strongly related to Hilbert's XIX problem (initially posed in 2-dimensions space and in the category of analytic functions). In the sequel we will first treat the case $n = 2$ and much later the case $n \geq 3$, which is significantly harder.

Recall that u is a local minimizer for I if, for any

$$u \in H^1_{\mathrm{loc}}(\Omega;\mathbb{R}^m),\ \mathrm{spt}(u-u') \subset \Omega' \Subset \Omega \quad \Longrightarrow \quad \int_{\Omega'} F(\nabla u')\,dx \geq \int_{\Omega'} F(\nabla u)\,dx\,.$$

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

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

Now, suppose s is a fixed coordinate direction (and let e_s 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 $\tau_{-h}\varphi$, we get

$$\int_\Omega \tau_h(B^\alpha_i(\nabla u))\frac{\partial\varphi^i}{\partial x_\alpha}\,dx = 0$$

and consequently, subtracting this to the previous one

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

However, as a consequence of the regularity of F, we can write

$$B^\alpha_i(\nabla u(x+he_s)) - B^\alpha_i(\nabla u(x)) = \int_0^1 \frac{d}{dt}B^\alpha_i(t\nabla u(x+he_s) + (1-t)\nabla u(x))\,dt$$

$$= \left[\int_0^1 A^{\alpha\beta}_{ij}(t\nabla u(x+he_s) + (1-t)\nabla u(x))\,dt\right]\left[\frac{\partial u^j}{\partial x_\beta}(x+he_s) - \frac{\partial u^j}{\partial x_\beta}(x)\right]$$

and setting

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

we rewrite the previous condition as

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

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

$$-\mathrm{div}(\widetilde{A}_h\nabla w) = 0\,. \tag{7.1}$$

It is obvious by the definition that $\widetilde{A}^{\alpha\beta}_{ij,h}(x)$ 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 (7.1) to obtain constants C_1 and C_2, not depending on h, such that

$$\int_{B_R(x_0)} |\nabla(\Delta_{h,s}u)|^2 \, dx \leq \frac{C_1}{R^2} \int_{B_{2R}(x_0)} |\Delta_{h,s}u|^2 \, dx \leq C_2$$

for any $B_R(x_0) \subset B_{2R}(x_0) \Subset \Omega$. Consequently, by Lemma 4.10 we deduce that

$$u \in H^2_{\text{loc}}(\Omega; \mathbb{R}^m). \tag{7.2}$$

Moreover, we have that

- $\Delta_{h,s}u \to \partial u/\partial x_s$ in L^2_{loc} (this is clearly true if u is regular and then we can exploit the fact that the operators $\Delta_{h,s}$ are equibounded, still by Lemma 4.10);
- $\partial u/\partial x_s$ satisfies, in a weak sense, the equation

$$-\text{div}\left(A(\nabla u)\nabla \frac{\partial u}{\partial x_s}\right) = 0 \ . \tag{7.3}$$

  In fact

$$A^{\alpha\beta}_{ij}(t\nabla u(x + he_s) + (1-t)\nabla u(x)) \overset{h\to 0}{\to} A^{\alpha\beta}_{ij}(\nabla u(x))$$

  in L^p for any $1 \leq p < \infty$, as an easy consequence of the continuity of translations in L^p and the continuity of A.

In order to solve Hilbert's XIX problem, we would like to apply a classical result by Schauder saying that if w is a weak solution of the problem $-\text{div}(B\nabla w) = 0$, then $B \in C^{0,\alpha} \Rightarrow w \in C^{1,\alpha}$, and so $u \in C^{2,\alpha}$. But we first need to improve the regularity of $B(x) = A(\nabla u(x))$. As a matter of fact, at this point we just know that $A(\nabla u) \in H^1_{\text{loc}}$, while we need $A(\nabla u) \in C^{0,\alpha}$. When $n = 2$ we can apply Widman's technique (see (4.9)) to the PDE (7.3) to obtain Hölder regularity of ∇u, both in the scalar and in the vectorial case. The situation is much harder in the case $n > 2$, since this requires deep new ideas: the celebrated theory by De Giorgi-Nash-Moser which solves the problem in the scalar case. We will see that in the vectorial case new difficulties arise.

8 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 are relevant, besides the standard Lebesgue spaces, in the regularity theory, as we will see.

Definition 8.1 (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*

$$\|u\|_{\alpha,A} := \sup_{x \neq y \in A} \frac{|u(x) - u(y)|}{|x-y|^\alpha}.$$

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

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 neighbourhood $U_x \Subset \Omega$ such that $\|u\|_{\alpha,U_x} < \infty$. The corresponding vector space is denoted by $C^{0,\alpha}_{\text{loc}}(\Omega;\mathbb{R}^m)$.

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

$$\|u\|_{C^{k,\alpha}} = \sum_{|i| \leq k} \left\|\nabla^i u\right\|_{C^{0,\alpha}} .$$

Definition 8.3 (Morrey spaces)**.** *Assume $\Omega \subset \mathbb{R}^n$ open, $\lambda \geq 0$ and $1 \leq p < \infty$. We say that $f \in L^p(\Omega)$ belongs to $L^{p,\lambda}(\Omega)$ 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

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

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

Remark 8.4. 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, for any $1 \leq p < \infty$ and $\lambda \geq 0$;

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

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

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

(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 asking $\lambda \leq \lambda_c$ 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 immediate applications of Lebesgue Differentiation Theorem. Finally the last one relies on Hölder inequality:

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

Definition 8.5 (Campanato spaces)**.** *Assume $\Omega \subset \mathbb{R}^n$ open, $\lambda > 0$, $1 \leq p < \infty$. A function $f \in L^p(\Omega)$ belongs to the Campanato space $\mathcal{L}^{p,\lambda}$ if*

$$\|f\|^p_{\mathcal{L}^{p,\lambda}} := \sup_{x_0 \in \Omega,\ 0<r<d_\Omega} r^{-\lambda} \int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx < \infty \ , \tag{8.1}$$

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

$$f_{x_0,r} := \fint_{\Omega(x_0,r)} f(x) \, dx \ . \tag{8.2}$$

The mean $f_{x_0,r}$ defined in (8.2) might not be optimal in the calculation of the sort of p-variance in (8.1), anyway it gives equivalent results, thanks to (5.4).

Remark 8.6. As in Remark 8.4, we briefly highlight the main properties of Campanato spaces.

(i) As defined in (8.1), $\|\cdot\|_{\mathcal{L}^{p,\lambda}}$ 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, $p \leq q$ and $(n-\lambda)/p \geq (n-\mu)/q$.

(iii) $C^{0,\alpha} \subset \mathcal{L}^{p,n+\alpha p}$, because

$$\int_{\Omega(x_0,r)} |f(x) - f_{x_0,r}|^p \, dx \leq \|f\|^p_{C^{0,\alpha}} r^{\alpha p} \mathscr{L}^n (B(x_0,r)) = \|f\|^p_{C^{0,\alpha}} \omega_n r^{n+\alpha p} \ .$$

We will see that a converse statement holds (namely functions in these Campanato spaces have Hölder continuous representatives in their Lebesgue equivalence class), and this is very useful: we can replace the pointwise definition of Hölder spaces with an integral one.

Actually, Campanato spaces are interesting only when $\lambda \geq n$, exactly 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 $c_* > 0$ satisfying

$$\mathscr{L}^n \left(\Omega \cap B_r(x_0)\right) \geq c_* r^n \quad \forall x_0 \in \overline{\Omega},\ \forall r \in (0, d_\Omega)\,. \tag{8.3}$$

For instance this assumption includes domains which are locally subgraphs of Lipschitz functions, while it rules out domains with outer cusps.

Theorem 8.7. *Let $\Omega \subset \mathbb{R}^n$ be an open bounded set satisfying* (8.3) *and let* $0 \leq \lambda < n$. *Then the spaces $L^{p,\lambda}$ and $\mathcal{L}^{p,\lambda}$ are equivalent, i.e.*

$$\|\cdot\|_{L^{p,\lambda}} \simeq \|\cdot\|_{\mathcal{L}^{p,\lambda}} + \|\cdot\|_{L^p}\ .$$

Proof. All through the proof we denote by c a generic constant depending from the constant c_* in (8.3) and from n, p, λ. We allow it to vary, even within the same line.

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

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

thus we can estimate

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

Conversely, we would like to estimate $r^{-\lambda}\int_{\Omega(x_0,r)} |f(x)|^p\, dx$ with $\|f\|_{\mathcal{L}^{p,\lambda}} + \|f\|_p$ for every $0 < r < d_\Omega$ and every $x_0 \in \Omega$. As a first step, by triangular inequality we separate

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

so we took out the problematic summand $|f_{x_0,r}|^p$.

In order to estimate $|f_{x_0,r}|^p$, let us bring in an inequality involving means on concentric balls: when $x_0 \in \Omega$ is fixed and $0 < r < \rho < d_\Omega$, it holds

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

thus we obtained that

$$|f_{x_0,r} - f_{x_0,\rho}| \le c\|f\|_{\mathcal{L}^{p,\lambda}} r^{-\frac{n}{p}} \rho^{\frac{\lambda}{p}} = c\|f\|_{\mathcal{L}^{p,\lambda}} \left(\frac{\rho}{r}\right)^{\frac{n}{p}} \rho^{\frac{\lambda-n}{p}} \ . \tag{8.4}$$

Now fix a radius $R > 0$: if $r = 2^{-(k+1)}R$ and $\rho = 2^{-k}R$, inequality (8.4) means that

$$|f_{x_0,R/2^{k+1}} - f_{x_0,R/2^k}| \le c\|f\|_{\mathcal{L}^{p,\lambda}} \left(\frac{R}{2^k}\right)^{\frac{\lambda-n}{p}} \ , \tag{8.5}$$

and, adding up when $k = 0, \dots, N-1$, it means that

$$|f_{x_0,R/2^N} - f_{x_0,R}| \le c\|f\|_{\mathcal{L}^{p,\lambda}} R^{\frac{\lambda-n}{p}} \frac{2^{N\frac{n-\lambda}{p}} - 1}{2^{\frac{n-\lambda}{p}} - 1} \le c\|f\|_{\mathcal{L}^{p,\lambda}} \left(\frac{R}{2^N}\right)^{\frac{\lambda-n}{p}} \ . \tag{8.6}$$

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 $r = R/2^N$. By triangular inequality

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

since

$$|f_{x_0,R}| \le c(d_\Omega)\|f\|_{L^p} \ ,$$

the only thing left to conclude is to apply inequality (8.6) in this case:

$$|f_{x_0,r} - f_{x_0,R}|^p \le c\|f\|^p_{\mathcal{L}^{p,\lambda}} r^{\lambda-n} \ ,$$

that is all we needed to conclude that

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

□

Remark 8.8. When the dimension of the domain space is n, the Campanato space $\mathcal{L}^{1,n}$ is very important in harmonic analysis and elliptic regularity theory: after John-Nirenberg seminal paper, this space is called BMO (bounded mean oscillation). It consists of the space of all functions $f : \Omega \to \mathbb{R}$ such that there exists a constant C satisfying the inequality

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

Notice that $L^\infty(\Omega) \subsetneq BMO(\Omega)$: for example, consider $\Omega = (0,1)$ and $f(x) = \ln x$. For any a, $r > 0$ it is easy to check that

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

hence $\ln x \in BMO(\Omega)$. For simplicity, we replaced the mean $⨍_a^{a+r} \ln s\, ds$ with $\ln(a+r)$, but, up to a multiplicative factor 2, this does not make a difference. On the contrary $\ln x \notin L^\infty(\Omega)$.

Theorem 8.9 (Campanato)**.** *With the previous notation, when $n < \lambda \leq n+p$ Campanato spaces $\mathcal{L}^{p,\lambda}$ are equivalent to 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 8.7, the letter c denotes a generic constant depending on the exponents, the space dimension n and the constant c_* in (8.3).

Let $\lambda = n + \alpha p$. We already observed in Remark 8.6 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 in the Lebesgue equivalence class of f which belongs to $C^{0,\alpha}$.

Recalling inequality (8.5) with fixed radius $R > 0$ and $x \in \Omega$, we obtain that the sequence $(f_{x,R/2^k})$ has the Cauchy property. Hence we define

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

Clearly

$$⨍_{\Omega(x,R/2^k)} |f(y) - f_{x,R/2^k}|^p\, dy \longrightarrow 0 \quad \Longrightarrow \quad ⨍_{\Omega(x,R/2^k)} |f(y) - \tilde{f}(x)|^p\, dy \longrightarrow 0 \ , \tag{8.7}$$

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

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

so that (8.7) implies that

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

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

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

We employ again an inequality from the proof of Theorem 8.7: letting $N \to \infty$ in (8.6), we get that

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

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

$$|\tilde{f}(x) - \tilde{f}(y)| \leq |\tilde{f}(x) - f_{x,R}| + |f_{x,R} - f_{y,R}| + |f_{y,R} - \tilde{f}(y)| \leq c|x-y|^\alpha + |f_{x,R} - f_{y,R}| \ .$$

The theorem will be proved if we can estimate $|f_{x,R} - f_{y,R}|$. To this aim, we use the inclusion $\Omega(y, R/2) \subset \Omega(x, R)$ to get

$$\begin{aligned} c_* 2^{-n} R^n |f_{x,R} - f_{y,R}|^p &\leq \int_{\Omega(y,R/2)} |f_{x,R} - f_{y,R}|^p \, ds \\ &\leq 2^{p-1} \left(\int_{\Omega(x,R)} |f(s) - f_{x,R}|^p \, ds + \int_{\Omega(y,R)} |f(s) - f_{y,R}|^p \right) \\ &\leq 2^p \|f\|^p_{\mathcal{L}^{p,\lambda}} R^\lambda , \end{aligned}$$

and finally

$$|f_{x,R} - f_{y,R}| \leq c \|f\|_{\mathcal{L}^{p,\lambda}} R^{\frac{\lambda - n}{p}} \leq c|x - y|^\alpha .$$

□

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

Proposition 8.10 (Inclusions between Lebesgue and Morrey spaces, Morrey and Campanato spaces)**.** *For all* $p \in (1, \infty)$, $L^p_{\rm loc}(\Omega) \subset L^{1,n/p'}_{\rm loc}(\Omega)$. *In addition*

$$|\nabla u| \in L^{p,\lambda}_{\rm loc}(\Omega) \quad \Longrightarrow \quad u \in \mathcal{L}^{p,\lambda+p}_{\rm loc}(\Omega) . \tag{8.8}$$

Corollary 8.11 (Sobolev embedding for $p > n$)**.** *If* $p > n$*, then* $W^{1,p}(\Omega) \subset C^{0,\alpha}_{\rm loc}(\Omega)$*, with* $\alpha = 1 - n/p$*. If* Ω *is bounded and regular, then* $W^{1,p}(\Omega) \subset C^{0,\alpha}(\Omega)$*.*

Proof. By the previous proposition we get

$$u \in W^{1,p}_{\rm loc} \quad \Longrightarrow \quad |\nabla u| \in L^{1,n/p'}_{\rm loc}(\Omega) = L^{1,n-n/p}(\Omega) = L^{1,n-1+\alpha}(\Omega) . \tag{8.9}$$

Applying (8.8) and (8.9), we get $u \in \mathcal{L}^{1,n+\alpha}_{\rm loc}(\Omega)$, so that $u \in C^{0,\alpha}_{\rm loc}(\Omega)$. 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. □

9 XIX Hilbert problem and its solution in the two-dimensional case

Let $\Omega \subset \mathbb{R}^n$ open, let $F \in C^3(\mathbb{R}^{m\times n})$ and let us consider a local minimizer u of the functional

$$v \mapsto \int_\Omega F(\nabla v) \, dx \tag{9.1}$$

as in Section 2.4. We assume that $\nabla^2 F(p)$ satisfies the Legendre condition (2.16) with $\lambda > 0$ independent of p and is uniformly bounded.

We have seen that u satisfies the Euler-Lagrange equations, for (9.1) they are

$$\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha}(\nabla u)\right) = 0 \qquad i = 1, \dots, m\,. \tag{9.2}$$

We have also seen in Section 7 how, differentiating (9.2) along the direction x_s, one can obtain

$$\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha p_j^\beta}(\nabla u)\frac{\partial^2 u^j}{\partial x_\beta \partial x_s}\right) = 0 \qquad i = 1, \dots, m\,. \tag{9.3}$$

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

$$\begin{aligned} w(x) &:= \frac{\partial u}{\partial x_s}(x) \in L^2(\Omega, \mathbb{R}^m)\ , \\ A(x) &:= \nabla^2 F(\nabla u(x))\ , \end{aligned}$$

thus (9.3) can be written as

$$\operatorname{div}(A\nabla w) = \frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha p_j^\beta}(\nabla u)\frac{\partial^2 u^j}{\partial x_\beta \partial x_s}\right) = 0\ . \tag{9.4}$$

Since $w \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ by (7.2), we can use the Caccioppoli-Leray inequality for w, in the sharp version of Remark 4.4. 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

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

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

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

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

$$\int_{B_{R/2}(x_0)} |\nabla w|^2\, dx \leq \left(\frac{1}{2}\right)^\alpha \int_{B_R(x_0)} |\nabla w|^2\, dx\ . \tag{9.5}$$

In order to get a power decay inequality from (9.5), we state this basic iteration lemma.

Lemma 9.1. *Consider a non-decreasing function $f : (0, R_0] \to \mathbb{R}$ satisfying*

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

Then

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

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

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

It is clear from the iteration of the hypothesis that

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

thus, by monotonicity,

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

□

Thanks to Lemma 9.1, we are ready to transform (9.5) in

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

therefore $|\nabla w| \in L^{2,\alpha}_{\mathrm{loc}}(\Omega)$. So, as we remarked in the proof of Corollary 8.11, this gives $w \in \mathcal{L}^{2,\alpha+2}_{\mathrm{loc}}(\Omega)$. All these facts are true in any number n of space dimensions, but when $n = 2$ we can apply Campanato Theorem to get

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

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

The Schauder theory that we will consider in the next section (just apply Theorem 10.4 to $\partial_{x_s} u$, solving the PDE (9.4)) will allow us to conclude that

$$u \in C^{2,\alpha/2}_{\mathrm{loc}}(\Omega) \, .$$

As long as F is sufficiently regular, the iteration of this argument solves XIX Hilbert's regularity problem in the C^∞ category.

We close this section with a more technical but useful iteration lemma in the same spirit of Lemma 9.1.

Lemma 9.2 (Iteration Lemma). *Consider a non-decreasing real function $f : (0, R_0] \to \mathbb{R}$ which satisfies for some coefficients $A > 0$, $B \geq 0$ and exponents $\alpha > \beta$ the following inequality*

$$f(\rho) \leq A\left[\left(\frac{\rho}{R}\right)^{\alpha} + \varepsilon\right] f(R) + BR^{\beta} \qquad \forall\, 0 < \rho \leq R \leq R_0 \, . \tag{9.6}$$

If

$$\varepsilon \leq \left(\frac{1}{2A}\right)^{\frac{\alpha}{\alpha-\gamma}} \tag{9.7}$$

for some $\gamma \in (\beta, \alpha)$, *then*

$$f(\rho) \leq c(\alpha, \beta, \gamma, A)\left[\left(\frac{\rho}{R}\right)^{\gamma} f(R) + B\rho^{\beta}\right] \qquad \forall\, 0 < \rho \leq R \leq R_0\,. \tag{9.8}$$

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

$$2A\tau^{\alpha} = \tau^{\gamma}\ , \tag{9.9}$$

thus (9.7) gives the inequality

$$\varepsilon \leq \tau^{\alpha}\ . \tag{9.10}$$

The following basic estimate uses the hypothesis (9.6) jointly with (9.9) and (9.10):

$$\begin{aligned} f(\tau R) &\leq A(\tau^{\alpha} + \varepsilon) f(R) + BR^{\beta} \\ &\leq 2A\tau^{\alpha} f(R) + BR^{\beta} = \tau^{\gamma} f(R) + BR^{\beta}\ . \end{aligned} \tag{9.11}$$

The iteration of (9.11) easily gives

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

It now can be easily proven by induction that

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

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

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

we conclude choosing the constant $c(\alpha, \beta, \gamma, A)$ in such a way that the last line in the following chain of inequalities holds:

$$\begin{aligned} f(\rho) &\leq f(\tau^N R) \leq \tau^{N\gamma} f(R) + \frac{BR^{\beta}\tau^{(N-1)\beta}}{1 - \tau^{(\gamma-\beta)}} \\ &\leq \tau^{-\gamma}\left(\tau^{(N+1)\gamma} f(R)\right) + \frac{\tau^{-2\beta}}{1 - \tau^{(\gamma-\beta)}}\left(BR^{\beta}\tau^{(N+1)\beta}\right) \\ &< \tau^{-\gamma}\left(\left(\frac{\rho}{R}\right)^{\gamma} f(R)\right) + \frac{\tau^{-2\beta}}{1 - \tau^{(\gamma-\beta)}}\left(B\rho^{\beta}\right) \\ &\leq c(\alpha, \beta, \gamma, A)\left(\left(\frac{\rho}{R}\right)^{\gamma} f(R) + B\rho^{\beta}\right)\ . \end{aligned}$$

□

Remark 9.3. The fundamental gain in Lemma 9.2 is the passage from R^β to ρ^β and the removal of ε, provided that ε is small enough. These improvements can be obtained at the price of passing from the power α to the worse power $\gamma < \alpha$.

10 Schauder theory

We are treating Schauder theory in a local form in $\Omega \subset \mathbb{R}^n$, just because it would be too long and technical to deal also with boundary regularity (some ideas are analogous to those used in Section 6). We shall describe first a model result for constant coefficient operators, and then we will consider the case of Hölder continuous coefficients.
We recall the usual PDE we are studying, in a divergence form:

$$\begin{cases} \operatorname{div}(A\nabla u) = \operatorname{div} F & \text{in } \Omega\,; \\ \\ u \in H^1_{\text{loc}}(\Omega;\mathbb{R}^m)\,. \end{cases} \tag{10.1}$$

Theorem 10.1. *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*

$$F \in \mathcal{L}^{2,\mu}_{\text{loc}}(\Omega) \qquad \Longrightarrow \qquad \nabla u \in \mathcal{L}^{2,\mu}_{\text{loc}}(\Omega)\ .$$

Proof. In this proof, $c = c(n,\lambda,|A|)$ and its value can change from line to line. Since the estimates we make are local, we assume with no loss of generality that $F \in \mathcal{L}^{2,\mu}(\Omega)$. Let us fix a ball $B_R \Subset \Omega$ with center $x_0 \in \Omega$ and compare with u the solution v of the homogeneous problem

$$\begin{cases} -\operatorname{div}(A\nabla v) = 0 & \text{in } B_R; \\ \\ v = u & \text{in } \partial B_R. \end{cases} \tag{10.2}$$

Since ∇v belongs to H^1 for previous results concerning H^2 regularity and its components $\frac{\partial v}{\partial x_\alpha}$ solve the same problem (because we supposed to have constant coefficients), we can use the decay estimates (5.1) and (5.2).
So, if $0 < \rho < R$, (5.2) provides us with the following inequality:

$$\int_{B_\rho} |\nabla v(x) - (\nabla v)_\rho|^2\,dx \le c\left(\frac{\rho}{R}\right)^{n+2} \int_{B_R} |\nabla v(x) - (\nabla v)_R|^2\,dx\ . \tag{10.3}$$

Now we try to employ (10.3) to get some estimate for u, the original "non-homogeneous", solution of (10.1). Obviously, we can write

$$u = w + v\ ,$$

where $w \in H_0^1(B_R;\mathbb{R}^m)$. Thus (first using $\nabla u = \nabla v + \nabla w$, then the minimality of the mean and (10.3), eventually $\nabla v = \nabla u - \nabla w$ and $(\nabla w)_R = 0$)

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

The auxiliary function

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

is non decreasing because of the minimality property of the mean $(\nabla u)_\rho$, when one minimizes $m \mapsto \int_{B_\rho} |\nabla u(x) - m|^2\,dx$. In order to get that f satisfies the hypothesis of Lemma 9.2, we have to estimate $\int_{B_R} |\nabla w|^2\,dx$. We can consider w as a function in $H^1(\mathbb{R}^n)$ (null out of Ω) so, by Gårding inequality (choosing the test function $\varphi = w$),

$$\begin{aligned}
\int_{B_R} |\nabla w(x)|^2\,dx\ &\leq\ c\int_{B_R} A\nabla w(x)\nabla w(x)\,dx \\
&=\ c\int_{B_R} F(x)\nabla w(x)\,dx = c\int (F(x) - F_R)\nabla w(x)\,dx
\end{aligned} \tag{10.4}$$

because $\mathrm{div}(A\nabla w) = \mathrm{div}F$ by linearity. Applying Young inequality to (10.4) and then absorbing $\int_{B_R} |\nabla w|^2\,dx$ in the left side of (10.4), we get

$$\int_{B_R} |\nabla w(x)|^2\,dx \leq c\int_{B_R} |F(x) - F_R|^2\,dx \leq c\|F\|^2_{\mathcal{L}^{2,\mu}} R^\mu\ ,$$

because $F \in \mathcal{L}^{2,\mu}$.

Therefore we obtained the decay inequality of Lemma 9.2 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 $\nabla u \in \mathcal{L}^{2,\mu}$. □

Corollary 10.2. *With the previous notation, when $\mu = n + 2\alpha$, Theorem 10.1 and Campanato Theorem 8.9 yield that*

$$F \in C^{0,\alpha} \quad \Longrightarrow \quad \nabla u \in C^{0,\alpha}\ .$$

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

Theorem 10.3. *Considering again* (10.1)*, suppose that* $A_{ij}^{\alpha\beta} \in C(\Omega)$ *and* A *satisfies a (locally) uniform Legendre-Hadamard condition for some* $\lambda > 0$*. If* $F \in L^{2,\mu}_{\rm loc}$ *with* $\mu < n$*, then* $|\nabla u| \in L^{2,\mu}_{\rm loc}$*.*

Naturally, since $\mu < n$, Campanato spaces and Morrey spaces coincide, so that we used Morrey spaces for simplicity.

Proof. Here is an example of Korn's technique of freezing of coefficients. We use the same convention on c of the previous proof, namely $c = c(n, \lambda, \sup |A|)$.
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 (10.1) solves

$$\mathrm{div}(A(x_0)\nabla u(x)) = \mathrm{div}\tilde{F}(x) \qquad \text{with} \qquad \tilde{F}(x) := F(x) + \big((A(x_0) - A(x))\nabla u(x)\big) \ .$$

Write $u = v + w$, where v solves the homogeneous PDE (10.2) with frozen coefficients $A(x_0)$. Using (5.1) for v we obtain

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

Thanks to the continuity property of A, there exists a (local) modulus of continuity ω of A which allows us to estimate

$$\int_{B_R} |\tilde{F}(x)|^2\, dx \leq 2\int_{B_R} |F(x) - F_R|^2\, dx + 2\omega^2(R)\int_{B_R} |\nabla u(x)|^2\, dx \ . \tag{10.5}$$

Consequently, as $F \in L^{2,\mu}_{\rm loc}$,

$$\int_{B_R} |\tilde{F}(x) - F_R|^2\, dx \leq \tilde{c}R^\mu + 2\omega^2(R)\int_{B_R} |\nabla u(x)|^2\, dx$$

with $\tilde{c}$ depending only on $\|F\|_{L^{2,\mu}_{\rm loc}}$. We are ready to use Lemma 9.2 with $f(\rho) := \int_{B_\rho} |\nabla u(x)|^2\, dx$, $\alpha = n$, $\beta = \mu < n$ and $\varepsilon = \omega^2(R)$: it tells us that if R is under a threshold depending only on c, α, β, ω and $\|F\|_{L^{2,\mu}_{\rm loc}}$ we have

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

so that $|\nabla u| \in L^{2,\mu}_{\rm loc}$. $\square$

We can now prove Schauder theorem for elliptic PDE's in divergence form. In the non-divergence form the result is (in the scalar case)

$$\sum_{\alpha,\beta} A_{\alpha\beta}\frac{\partial^2 u}{\partial x_\alpha \partial x_\beta} \in C^{0,\alpha} \qquad \Longrightarrow \qquad u \in C^{2,\alpha}\,, \tag{10.6}$$

if A is of class $C^{0,\alpha}$. The proof follows similar lines, i.e. starting for second derivative decay estimates for constant coefficient operators, and then freezing the coefficients. Notice also that both (10.6) and Theorem 10.4 below are easily seen to be optimal, considering 1-dimensional ODE's $au'' = f$ or $(au')' = f'$.

Theorem 10.4 (Schauder)**.** *Suppose that the coefficients $A_{ij}^{\alpha\beta}(x)$ of the PDE* (10.1) *belong to $C^{0,\alpha}(\Omega)$ and A satisfies a (locally) uniform Legendre-Hadamard in Ω for some $\lambda > 0$. Then the following implication holds*

$$F \in C^{0,\alpha}_{\rm loc} \quad \Longrightarrow \quad \nabla u \in C^{0,\alpha}_{\rm loc}\,,$$

that is to say

$$F \in \mathcal{L}^{2,n+2\alpha}_{\rm loc} \quad \Longrightarrow \quad \nabla u \in \mathcal{L}^{2,n+2\alpha}_{\rm loc}\,.$$

Proof. With the same idea of freezing coefficients (and the same notation, too), we estimate by (5.1)

$$\int_{B_\rho} |\nabla u(x) - (\nabla u)_\rho|^2\,dx \le c\left(\frac{\rho}{R}\right)^{n+2}\int_{B_R} |\nabla u(x) - (\nabla u)_R|^2\,dx + c\int_{B_R} |\tilde F(x) - F_R|^2\,dx\,. \tag{10.7}$$

Additionally, the Hölder property of A makes us rewrite (10.5) as

$$\int_{B_R} |\tilde F(x) - F_R|^2\,dx \le 2\int_{B_R} |F(x) - F_R|^2\,dx + cR^{2\alpha}\int_{B_R} |\nabla u(x)|^2\,dx\,. \tag{10.8}$$

Since $F \in C^{0,\alpha}_{\rm loc}$, we obtain

$$\int_{B_R} |\tilde F(x) - F_R|^2\,dx \le cR^{n+2\alpha} + cR^{2\alpha}\int_{B_R} |\nabla u(x)|^2\,dx\,.$$

Theorem 10.3 with $\mu = n - \alpha < n$ tells us that $|\nabla u| \in L^{2,\mu}$, thus

$$\int_{B_R} |\tilde F(x) - F_R|^2\,dx \le cR^{n+2\alpha} + cR^{n+\alpha}\,. \tag{10.9}$$

Adding (10.9) to (10.7) and applying Lemma 9.2 with exponents $n+2$ and $n+\alpha$, we get $\nabla u \in \mathcal{L}^{2,n+\alpha}$, so that $\nabla u \in C^{0,\alpha/2}$, in particular $|\nabla u|$ is locally bounded. Using this information we can improve (10.9) as follows:

$$\int_{B_R} |\tilde F(x) - F_R|^2\,dx \le cR^{n+2\alpha}\,.$$

Now we reach the conclusion, again by Lemma 9.2 with exponents $n+2$ and $n+2\alpha$. □

11 Regularity in L^p spaces

In this section we deal with elliptic regularity in the category of L^p spaces, obviously a natural class of spaces besides Morrey, Hölder and Campanato spaces.

Lemma 11.1. *In a measure space $(\Omega, \mathcal{F}, \mu)$, consider a $\mathcal{F}$-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 $1 \le p < \infty$:

$$\int_\Omega f^p(x)\, d\mu(x) \ = \ p\int_0^\infty t^{p-1}F(t)\, dt \tag{11.1}$$

$$\int_{\{f>s\}} f^p(x)\, d\mu(x) \ = \ p\int_s^\infty t^{p-1}F(t)\, dt + s^pF(s) \qquad 0 \le s < \infty \ . \tag{11.2}$$

Proof. It is a simple consequence of Fubini's Theorem that

$$\begin{aligned}\int_\Omega f^p(x)\, d\mu(x) &= \int_\Omega p\left(\int_0^{f(x)} t^{p-1}\, dt\right) d\mu(x) = p\int_0^\infty t^{p-1}\left(\int \chi_{\{t<f(x)\}}\, d\mu(x)\right) dt \\ &= p\int_0^\infty t^{p-1}F(t)\, dt \ .\end{aligned}$$

Equation (11.2) follows from (11.1) applied to the function $f\chi_{\{f>s\}}$. □

Theorem 11.2 (Markov inequality)**.** *In a measure space $(\Omega, \mathcal{F}, \mu)$, a function $f \in L^p(\Omega, \mathcal{F}, \mu)$ satisfies (with the convention $0 \times \infty = 0$)*

$$t^p\mu\left(\{|f| \ge t\}\right) \le \int_\Omega |f|^p\, d\mu \qquad \forall t \ge 0 \ . \tag{11.3}$$

Proof. We begin with the trivial pointwise inequality

$$s\chi_{\{g\ge s\}}(x) \le g(x) \qquad \forall\, x \in \Omega \tag{11.4}$$

for g nonnegative. Thus, integrating (11.4) in Ω we obtain

$$s\mu\left(\{g \ge s\}\right) \le \int_\Omega g\, d\mu \ .$$

The thesis follows choosing $s = t^p$ and $g = |f|^p$. □

The Markov inequality inspires the definition of a space which is weaker than L^p, but still keeps (11.3).

Definition 11.3 (Marcinkiewicz space)**.** *Given a measure space $(\Omega, \mathcal{F}, \mu)$ and an exponent $1 \le p < \infty$, the Marcinkiewicz space $L^p_w(\Omega,\mu)$ is defined by*

$$L^p_w(\Omega,\mu) := \left\{ f : \Omega \to \mathbb{R}\ \mathcal{F}\text{-measurable} \,\Big|\, \sup_{t>0} t^p \mu\left(\{|f| > t\}\right) < \infty \right\} .$$

We denote[4] *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 11.4. If μ is a finite measure, then

$$q < p \quad \Longrightarrow \quad L^p \subset L^p_w \subset L^q \ .$$

The first inclusion is due to Markov inequality (11.2), on the other hand, if $f \in L^p_w$, then

$$\begin{aligned} \int_\Omega |f|^q \, d\mu(x) &= q \int_0^\infty t^{q-1} F(t)\, dt \le q \left(\int_0^1 t^{q-1} F(t)\, dt + \int_1^\infty t^{q-1} F(t)\, dt \right) \\ &\le q\mu(\Omega) + q \int_1^\infty t^{q-1} \|f\|^p_{L^p_w} t^{-p}\, dt = q\mu(\Omega) + \frac{q}{p-q} \|f\|^p_{L^p_w} \ . \end{aligned}$$

Definition 11.5 (Maximal operator)**.** *When $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ we define the maximal function $\mathcal{M}f$ by*

$$\mathcal{M}f(x) := \sup_{Q_r(x)} \fint_{Q_r(x)} |f(y)|\, dy \ , \tag{11.5}$$

where $Q_r(x)$ is the n-dimensional cube with center x and side length r.

It is easy to check that at $\mathcal{M}f(x) \ge \tilde{f}(x)$ at Lebesgue points, so that $\mathcal{M}f \ge f$ $\mathscr{L}^n$-a.e. in $\mathbb{R}^n$. On the other hand, it is important to remark that the maximal operator $\mathcal{M}$ does not map L^1 into L^1.

Example 11.6. In dimension $n = 1$, consider $f = \chi_{[0,1]} \in L^1$. Then

$$\mathcal{M}f(x) = \frac{1}{2|x|} \qquad \text{when } |x| \ge 1 \ ,$$

so $\mathcal{M}f \notin L^1$. In fact, it is easy to prove that $\mathcal{M}f \in L^1$ implies $|f| = 0$ $\mathscr{L}^n$-a.e. in $\mathbb{R}^n$.

[4] Pay attention to the lack of subadditivity of $\|\cdot\|_{L^p_w}$: the notation is misleading, 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 greater. On the other hand, it is easily seen that $\|f+g\|_{L^p_w} \le 2\|f\|_{L^p_w} + 2\|g\|_{L^p_w}$

However, if $f \in L^1$, the maximal operator $\mathcal{M}f$ belongs to the weaker Marcinkiewicz space L^1_w, as we are going to see in Theorem 11.8. We first recall the Vitali covering theorem, in a version valid in *any* metric space.

Lemma 11.7 (Vitali)**.** *Let $\mathcal{F}$ be a finite family of balls in a metric space (X, d). Then, there exists $\mathcal{G} \subset \mathcal{F}$, made of disjoint balls, satisfying*

$$\bigcup_{B\in\mathcal{F}} B \subset \bigcup_{B\in\mathcal{G}} \hat{B} \ .$$

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 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 $\mathcal{G}$ of chosen balls is, by construction, disjoint. If $B \in \mathcal{F} \setminus \mathcal{G}$, then B has not been chosen because it intersects one of the balls in $\mathcal{G}$; the first of these balls B_f has radius larger than the radius of B (otherwise B would have been chosen before B_f), hence $B \subset \widehat{B_f}$. □

Theorem 11.8 (Hardy-Littlewood maximal theorem)**.** *The maximal operator $\mathcal{M}f$ defined in* (11.5) *satifies*

$$\|\mathcal{M}f\|_{L^1_w} \leq 3^n \|f\|_{L^1} \qquad \forall f \in L^1(\mathbb{R}^n) \ .$$

Proof. Fix $t > 0$ and a compact set $K \subset \{\mathcal{M}f > t\}$: by inner regularity of the Lebesgue measure we will reach the conclusion showing 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

$$\int_{Q_{r(x)}(x)} |f(y)|\, dy \geq t(r(x))^n \ .$$

Compactness allows us to cover K with a finite number of cubes

$$K \subset \bigcup_{i\in I} Q_{r(x_i)}(x_i) \ ,$$

then Vitali's lemma stated for the distance induced by the sup norm in $\mathbb{R}^n$ allows us to find $J \subset I$ 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

$$\mathscr{L}^n(K) \leq \sum_{j \in J} 3^n (r(x_j))^n \leq \frac{3^n}{t} \sum_{i \in I} \int_{Q_{r(x_i)}(x_i)} f(y)\, dy \leq \frac{3^n}{t} \|f\|_{L^1} \ .$$

□

12 Some classical interpolation theorems

In the sequel, we will make extensive use of some classical interpolation theorems, that are basic tools in Functional and Harmonic Analysis.

Assume $(X, \mathcal{F}, \mu)$ is a measure space. For the sake of brevity, we will say that a linear operator T mapping a vector space $D \subset L^p(X, \mu)$ into $L^q(X, \mu)$ is of type (p, q) if it is continuous with respect to the $L^p - L^q$ topologies. If this happens, obviously T can be extended (by Hahn-Banach) to a linear continuous operator from $L^p(X, \mu)$ to $L^q(X, \mu)$ and the extension is unique if D is dense.

The inclusion $L^p \cap L^q \subset L^r$ for $p \leq q$ and $r \in [p, q]$ can be better specified with the following result.

Theorem 12.1 (Riesz-Thorin interpolation)**.** *Let p, $q \in [1, \infty]$ with $p \leq q$ and $T : L^p(X, \mu) \cap L^q(X, \mu) \to L^p(X, \mu) \cap L^q(X, \mu)$ 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 give the proof of this theorem, whose proof follows the lines of the more general Marcinkiewicz theorem below (a standard reference is [26]). In the sequel we shall consider operators T that are not necessarily linear, but Q-subadditive for some $Q \geq 0$, namely

$$|T(f + g)| \leq Q(|T(f)| + |T(g)|) \qquad \forall f,\, g \in D \ .$$

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 $k > 0$ (all L^p spaces are stable under truncations).

Definition 12.2 (Strong and weak (p, p) operators)**.** *Let $s \in [1, \infty]$, $D \subset L^s(X, \mu)$ a linear subspace and let $T : D \subset L^s(X, \mu) \to L^s(X, \mu)$, not necessarily linear. We say that T is of strong type (s, s) if $\|T(u)\|_s \leq C\|u\|_s$ 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

$$\mu\left(\{x : \; |Tu(x)| > \alpha\}\right) \le C \frac{\|u\|_s^s}{\alpha^s} \quad \forall \alpha > 0,\, u \in D$$

for some constant C independent of u and α. 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 12.3 (Marcinkiewicz Interpolation Theorem)**.** *Assume that p, $q \in [1, \infty]$ with $p < q$, $D \subset L^p(X, \mu) \cap L^q(X, \mu)$ is a linear space stable under truncations and $T : D \to L^p(X, \mu) \cap L^q(X, \mu)$ 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 12.4. 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 we limit ourselves to prove the theorem under this additional hypothesis.

Proof. We can truncate $f \in D$ as follows:

$$f = g + h, \quad g(x) = f(x)\chi_{\{|f| \le \gamma s\}}(x), \quad h(x) = f(x)\chi_{\{|f| > \gamma s\}}(x)\,,$$

where γ is an auxiliary parameter to be fixed later. By assumption $g \in D \cap L^\infty(X, \mu)$ while $h \in D \cap L^1(X, \mu)$ 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 $D \cap L^\infty(X, \mu)$ into $L^\infty(X, \mu)$. Choose γ so that $QA_\infty\gamma = 1/2$, therefore

$$\{|T(f)| > s\} \subset \{|T(h)| > \frac{s}{2Q}\}$$

and so

$$\mu\big(\{|T(f)| > s\}\big) \le \mu\big(\{|T(h)| > \frac{s}{2Q}\}\big) \le \left(\frac{2A_1 Q}{s}\right) \int_X |h| \, d\mu \le \left(\frac{2A_1 Q}{s}\right) \int_{\{|f| > \gamma s\}} |f| \, d\mu\,,$$

where A_1 is the constant appearing in the weak $(1, 1)$ estimate. By integration of the previous inequality, we get

$$p \int_0^\infty s^{p-1} \mu\left(\{|T(f)| > s\}\right) \, ds \le 2A_1 Qp \int_0^\infty \int_{\{|f| \ge \gamma s\}} s^{p-2} |f| \, d\mu \, ds$$

and by means of the Fubini-Tonelli Theorem we finally get

$$\|T(f)\|_p^p \le 2A_1 Qp \int_X \left(\int_0^{|f(x)|/\gamma} s^{p-2} \, ds\right) |f(x)| \, d\mu(x) = \frac{2A_1 Qp}{(p-1)\gamma^{p-1}} \|f\|_p^p$$

and the conclusion follows. □

Remark 12.5 (The limit case $p = 1$)**.** In the limit case $p = 1$ we can argue similarly to find

$$\begin{aligned}&\int_1^\infty \mu\left(\{|T(f)| > s\}\right) ds \\ \le\ & 2A_1Q \int_{\{|f|\ge\gamma\}} \left(\int_1^{|f(x)|/\gamma} s^{-1}\, ds\right) |f(x)|\, d\mu(x) = 2A_1Q \int_{\{f\ge\gamma\}} |f| \log|f|\, d\mu.\end{aligned}$$

Therefore, a slightly better integrability of $|f|$ provides at least integrability of $|T(f)|$ on bounded sets.

Remark 12.6. 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, can be easily generalized, for instance to *pseudo-metric spaces* (i.e. when the distance fulfils only the triangle and symmetry assumptions) endowed with a *doubling measure*, that is a measure μ such that $\mu(B_{2r}(x)) \le \beta\mu(B_r(x))$ for some constant β not depending on the radius and the center of the ball. Notice that in this case the constant in the weak $(1, 1)$ bound of the maximal operator does not exceed β^2, since $\mu(B_{3r}(x)) \le \beta^2\mu(B_r(x))$.

13 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.

Theorem 13.1. *Let (X, d, μ) be a metric space with a finite doubling measure on its Borel σ-algebra and $p \in [1, \infty)$. If $f \in L^p(\mu)$ then for μ-a.e. $x \in X$ we have that*

$$\lim_{r\downarrow 0} ⨍_{B_r(x)} |f(y) - f(x)|^p\, d\mu(y) = 0.$$

Proof. Let

$$\Lambda_t := \left\{x \in X \,|\, \limsup_{r\downarrow 0} ⨍_{B_r(x)} |f(y) - f(x)|^p\, d\mu(y) > t\right\}.$$

The thesis can be achieved showing that for any $t > 0$ we have $\mu(\Lambda_t) = 0$, since the stated property holds out of $\cup_n \Lambda_{1/n}$. Now, we can exploit the metric structure of X in order to approximate f in $L^1(\mu)$ norm by means of continuous and bounded functions: for any $\varepsilon > 0$ we can write $f = g + h$ with $g \in C_b(X)$ and $\|h\|^p_{L^p} \le t\varepsilon$. Hence, it is enough to prove that for any $t > 0$ we have $\mu(A_t) = 0$ where

$$A_t := \left\{x \in X \,|\, \limsup_{r\downarrow 0} ⨍_{B_r(x)} |h(y) - h(x)|^p\, d\mu(y) > t\right\}.$$

This is easy, because by definition

$$A_t \subset \left\{ |h|^p > \frac{t}{2^{p+1}} \right\} \cup \left\{ \mathcal{M}(|h|^p) > \frac{t}{2^{p+1}} \right\}$$

and, if we consider the corresponding measures, we have (taking Remark 12.6 into account)

$$\mu(A_t) \le \frac{2^{p+1}}{t} \|h\|_{L^p}^p + \frac{2^{p+1}}{t} M \|h\|_{L^p}^p \le 2^{p+1}(1+M)\varepsilon$$

where M is the constant in the weak $(1,1)$ bound. Since $\varepsilon > 0$ is arbitrary we get the thesis. □

Remark 13.2. All the previous results have been derived for the maximal operator defined in terms of *centered balls*, that is

$$Mf(x) = \sup_{r>0} \fint_{B_r(x)} f(y)\, dy$$

and the Lebesgue differentiation theorem has been stated according to this setting. However, it is clear that we can generalize everything to any metric space (X, d, μ) with a finite doubling measure and a suitable family of sets $\mathcal{F} := \cup_{x \in X} \mathcal{F}_x$ with

$$M_{\mathcal{F}} f(x) = \sup_{A \in \mathcal{F}_x} \fint_A f(y)\, dy$$

provided there exists a universal constant $C > 0$ such that

$$\text{for all } A \in \mathcal{F}_x \text{ there exists } r > 0 \text{ such that } A \subset B_r(x) \text{ and } \mu(A) \ge C\mu(B_r(x)) . \tag{13.1}$$

Indeed, even though one might define the maximal operator with this larger family of mean values, suffices just to notice that

$$\fint_A |f(y) - f(x)|\, d\mu(y) \le \frac{1}{C} \fint_{B_r(x)} |f(y) - f(x)|\, d\mu(y) ,$$

provided $B_r(x)$ is chosen according to (13.1).

In Euclidean spaces, an important example to which the previous remark applies, in connection with Calderón-Zygmund theory, is given by

$$\mathcal{F}_x := \{Q \text{ cube}, \, x \in Q\} ,$$

consequently Lebesgue theorem gives

$$\lim_{x \in Q, \, |Q| \to 0} \fint_Q |f(y) - f(x)|^p \, dy = 0$$

for a.e. $x \in \mathbb{R}^n$, Notice that requiring $|Q| \to 0$ (i.e. $\mathrm{diam}(Q) \to 0$) is essential to "factor" continuous functions as in the proof of Theorem 13.1.

14 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 14.1. *Let $f \in L^1(Q)$, $f \geq 0$ and consider a positive real number α such that $⨍_Q f\,dx \leq \alpha$. 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) $\alpha < ⨍_{Q_i} f\,dx \leq 2^n\alpha \quad \forall i$;

(iii) $f \leq \alpha$ *a.e. on* $Q \setminus \cup_i Q_i$.

Remark 14.2. 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

$$\alpha \sum_i \mathscr{L}^n(Q_i) < \sum_i \int_{Q_i} f\,dx \leq \|f\|_1 .$$

The proof is based on a so-called *stopping-time* argument.

Proof. Divide the cube Q in 2^n subcubes by means of n bisections of Q with hyperplanes parallel to the sides of the cube itself. We will call this process *dyadic decomposition*. Then

- if $⨍_{Q_i} f > \alpha$ we do not divide Q_i anymore;
- else 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 are obvious by construction: indeed, if Q_i is a chosen cube then its parent cube $\widetilde{Q}_i$ satisfies $⨍_{\widetilde{Q}_i} f \leq \alpha$, which gives easily $⨍_{Q_i} f \leq 2^n\alpha$. For the third one, note that if $x \in Q \setminus \cup_i Q_i$, then there exists a sequence of subcubes $(\widetilde{Q}_j)$ with $x \in \cap_j \widetilde{Q}_j$ and $\mathscr{L}^n\big(\widetilde{Q}_j\big) \to 0$, $⨍_{\widetilde{Q}_j} f\,dx \leq \alpha$. Thanks to the Lebesgue differentiation theorem we get $f(x) \leq \alpha$ for a.e. $x \in Q \setminus \cup_i Q_i$. □

Remark 14.3 (Again in the limit case $p = 1$)**.** Using the Calderon-Zygmund decomposition, for $\alpha > \|f\|_1$ we can reverse somehow the weak $(1,1)$ estimate:

$$\int_{\{|f|>\alpha\}} |f|\,dx \le \sum_i \int_{Q_i} |f|\,dx \le 2^n \alpha \mathscr{L}^n(Q_i) \le 2^n \alpha \mathscr{L}^n\big(\{M|f| > \alpha/2^n\}\big)\,,$$

because the cubes Q_i are contained in $\{M|f| > \alpha/2^n\}$. Using this inequality we can also reverse the implication of Remark 12.5, namely assuming with no loss of generality that $f \ge 0$ and $\int f\,dx = 1$:

$$\begin{aligned}\int_{\{f>1\}} f\log f\,dx &= \int_0^\infty \int_{\{\log f>t\}} f\,dxdt = \int_1^\infty \frac{1}{s}\int_{\{f>s\}} f\,dx \\ &\le 2\int_1^\infty \mathscr{L}^n\big(\{Mf > \frac{s}{2}\}\big)\,ds = 2\int (Mf - \frac{1}{2})^+\,dx\,.\end{aligned}$$

15 The BMO space

Given a cube $Q \subset \mathbb{R}^n$, we define

$$BMO(Q) := \left\{ u \in L^1(Q) |\ \sup_{Q' \subset Q} \fint_{Q'} |u - u_{Q'}|\ dx < \infty \right\}\,,$$

where $u_{Q'}$ denotes the mean value of u on Q'. We also define the seminorm $\|u\|_{BMO}$ as the supremum in the right hand side. An elementary argument replacing balls with concentric cubes shows that $BMO(Q) \sim \mathcal{L}^{1,n}$, that is the two spaces consist of the same elements and the corresponding semi-norms are equivalent. Here we recall the inclusion already discussed in Remark 8.8.

Theorem 15.1. *For any cube $Q \subset \mathbb{R}^n$ the following inclusion holds:*

$$W^{1,n}(Q) \hookrightarrow BMO(Q).$$

Proof. First, notice that $W^{1,n}(Q) \hookrightarrow \{u|\,|\nabla u| \in L^{1,n-1}(Q)\}$, as an immediate consequence of the Hölder inequality. Then, by Poincaré inequality, there exists a dimensional constant $C > 0$ such that for any $Q' \subset Q$ with sides of length h

$$\int_{Q'} |u - u_{Q'}|\ dx \le Ch \int_{Q'} |\nabla u|\ dx \le C\,|\nabla u|_{L^{1,n-1}}\,h^n\ .$$

$\square$

However, it should be clear that the previous inclusion is far from being an equality as elementary examples show, see Remark 8.8. We shall extend now to n-dimensional spaces the example in Remark 8.8, stating first a simple sufficient (and necessary, as we will see) condition for BMO.

Proposition 15.2. *Let $u : Q \to \mathbb{R}$ be a measurable function such that, for some $b > 0$, $B \geq 0$, the following property holds:*

$$\forall\, C \subset Q \text{ cube}, \quad \exists\, a_C \in \mathbb{R} \text{ s.t. } \mathscr{L}^n\big(C \cap \{|u - a_C| > \sigma\}\big) \leq B e^{-b\sigma}\, |C| \;\; \forall \sigma \geq 0\ . \tag{15.1}$$

Then $u \in BMO(Q)$.

The proof of the proposition above is simple, since

$$\frac{1}{2}\int_C |u - u_C|\, dx \leq \int_C |u - a_C|\, dx = \int_0^\infty \mathscr{L}^n\big(C \cap \{|u - a_C| \geq \sigma\}\big)\, d\sigma \leq \frac{B}{b}|C|\ .$$

Example 15.3. Thanks to Proposition 15.2 we can check that $\ln|x| \in BMO\big((0,1)^n\big)$. Indeed, $\ln|x|$ satisfies (15.1) (the parameters b and B will be made precise later). To see this, fix a cube C, with h the length of the side of C. We define, respectively,

$$\xi := \max_{x \in C} |x|\ , \qquad \eta := \min_{x \in C} |x|\ , \qquad a_C := \ln \xi\ ,$$

so that

$$a_C - u = \ln\left(\frac{\xi}{|x|}\right) \geq 0\ .$$

We estimate the Lebesgue measure of $C \cap \{\xi \geq |x| e^\sigma\}$: naturally we can assume that $\xi \geq \eta e^\sigma$, otherwise there is nothing to prove, so

$$\xi e^{-\sigma} \geq \eta \geq \xi - \mathrm{diam}(C) \geq \xi - \sqrt{n} h\ ,$$

then

$$\xi \leq \frac{\sqrt{n} h}{1 - e^{-\sigma}}\ .$$

Finally

$$\frac{1}{h^n}\mathscr{L}^n\left(C \cap \{|u - a_C| \geq \sigma\}\right) \leq \frac{1}{h^n}\mathscr{L}^n\big(B_{\xi e^{-\sigma}}\big) \leq \frac{(\sqrt{n})^n \omega_n}{(1 - e^{-\sigma})^n} e^{-n\sigma}\ ,$$

so that distinguishing the cases $\sigma \leq 1$ and $\sigma > 1$ we see that (15.1) holds with $b = n$ and $B = \max\{e^n, (\sqrt{n})^n \omega_n\, (1 - e^{-1})^{-n}\}$.

The following theorem by John and Nirenberg was first proved in [21].

Theorem 15.4 (John-Nirenberg, first version)**.** *There exist constants* c_1, c_2 *depending only on the dimension* n *such that*

$$\mathscr{L}^n\left(\{|u-u_Q|>t\}\right) \le c_1 e^{-c_2 t/\|u\|_{BMO}} \mathscr{L}^n(Q) \qquad \forall u \in BMO(Q)\setminus\{0\}\,. \tag{15.2}$$

Remark 15.5. In the proof we present here, we will find explicitly $c_1 = e$ and $c_2 = 1/(2^n e)$. However, these constants are not sharp.

Before presenting the proof, we discuss here two very important consequences of this result.

Corollary 15.6 (Exponential integrability of BMO functions)**.** *For any* $c < c_2$ *there exists* $K(c, c_1, c_2)$ *such that*

$$⨍_Q e^{c|u-u_Q|/\|u\|_{BMO}}\,dx \le K(c,c_1,c_2) \qquad \forall u \in BMO(Q)\setminus\{0\}\,.$$

Proof. It is a simple 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 $\|u\|_{BMO(Q)} = 1$, $\mathscr{L}^n(Q) = 1$ and we used the John-Nirenberg inequality. $\square$

Remark 15.7 (Better integrability of $W^{1,n}$ functions)**.** The previous theorem tells that the class BMO (and hence also $W^{1,n}$) has exponential integrability properties. This result can be in part refined by the celebrated *Moser-Trudinger inequality*, that we quote here without proof.

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

$$C(\Omega) := \sup\left\{\int_\Omega \exp\bigl(\alpha_n|u|^{n/(n-1)}\bigr)\,dx \,:\, u \in W_0^{1,n}(\Omega),\ \int_\Omega |\nabla u|^n\,dx \le 1\right\} < \infty\,.$$

This inequality has been first proved in [27].

Theorem 15.8. *If* $p \in [1,\infty)$ *we have*

$$\left(⨍_Q |u-u_Q|^p\,dx\right)^{1/p} \le c(n,p)\|u\|_{BMO} \qquad \forall u \in BMO(Q)\ .$$

Consequently the following isomorphisms hold:

$$\mathcal{L}^{p,n}(Q) \sim BMO(Q) \sim \mathcal{L}^{1,n}(Q)\ . \tag{15.3}$$

The proof of Theorem 15.8 relies on a simple and standard computation, similar to the one presented before in order to get exponential integrability. Indeed, assuming $\|u\|_{BMO} = 1$, (15.2) gives

$$\fint_Q |u - u_Q|^p \, dx = p \int_0^\infty \mathscr{L}^n\big(\{|u - u_Q| > s\}\big) s^{p-1} \, ds \leq c_1 p \int_0^\infty e^{-c_2 s} s^{p-1} \, ds \, .$$

We can now conclude this section, by proving the John-Nirenberg inequality (15.2).

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 $k \geq 1$ a countable family of subcubes $\{Q_i^k\}_{i \in I_k}$ contained in Q such that

(i) $|u(x) - u_Q| \leq 2^n k\alpha$ a.e. on $Q \setminus \cup_{i \in I_k} Q_i^k$;

(ii) $\sum_{i \in I_k} \mathscr{L}^n(Q_i^k) \leq \alpha^{-k} \mathscr{L}^n(Q)$.

The combination of linear growth in (i) and geometric decay in (ii) leads to the exponential decay of the repartition function: indeed, choose k such that $2^n\alpha k \leq t < 2^n\alpha(k+1)$, then

$$\mathscr{L}^n\left(\{|u - u_Q| > t\}\right) \leq \mathscr{L}^n\left(\{|u - u_Q| > 2^n\alpha k\}\right) \leq \alpha^{-k} \mathscr{L}^n(Q)$$

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

$$\alpha^{-k} = c_1 e^{-c_2 2^n \alpha(k+1)}$$

and consequently we determine the constants c_1, c_2, requiring

$$e^{c_2 2^n \alpha} = \alpha, \qquad c_1 e^{-c_2 2^n \alpha} = 1 \ .$$

By the first relation $c_2 = \log \alpha / (2^n \alpha)$ and we maximize 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 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 14.2, while (i) is obvious since $|u(x) - u_Q| \leq \alpha$ 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 cubes $\{Q_{i,l}^2\}$, each contained in one of the previous ones. Moreover Remark 14.2 and the induction assumption give

$$\sum_{i,l} \mathscr{L}^n(Q_{i,l}^2) \leq \sum_i \frac{1}{\alpha} \int_{Q_i^1} |u - u_{Q_i^1}|\, dx \leq \sum_i \frac{1}{\alpha} \mathscr{L}^n(Q_i^1) \leq \frac{1}{\alpha^2} \mathscr{L}^n(Q)\,,$$

which is (ii). In order to get (i), notice that

$$Q \setminus \bigcup Q_{i,l}^2 \subset \left(Q \setminus \bigcup_i Q_i^1\right) \cup \left(\bigcup_i (Q_i^1 \setminus \bigcup_l Q_{i,l}^2)\right)$$

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 $k>1$ and this is what we need to conclude the argument and the proof. □

The John-Nirenberg theorem stated in Theorem 15.4 can be extended considering the L^p norms, so that the case of BMO maps corresponds to the limit as $p \to \infty$.

Theorem 15.9 (John-Nirenberg, second version)**.** *For any $p \in [1,\infty)$ and $u \in L^p(Q)$ define*

$$K_p^p(u) := \sup\left\{\sum_i \mathscr{L}^n(Q_i)\left(\fint_{Q_i} |u(x) - u_{Q_i}|\, dx\right)^p \mid \{Q_i\} \text{ partition of } Q\right\}\,.$$

Then there exists a constant $c = c(p,n)$ such that

$$\|u - u_Q\|_{L^p_w} \leq c(p,n) K_p(u)\,.$$

The proof of Theorem 15.9 is basically the same as Theorem 15.4, the goal being to prove the polynomial decay

$$|\{|u-u_Q| > t\}| \leq \frac{c(p,n)}{t^p} K_p(u) \qquad t > 0$$

instead of an exponential decay.

The following important result improves the classical interpolation theorems in L^p spaces, replacing L^∞ with BMO. This is crucial for the application to elliptic PDE's, as we will see.

Theorem 15.10 (Stampacchia's interpolation)**.** *Let $D \subset L^\infty(Q;\mathbb{R}^s)$ be a linear space and $p \in [1,\infty)$. Consider a linear operator $T : D \to BMO(Q_0)$, continuous with respect to the norms $(L^\infty(Q;\mathbb{R}^s), BMO(Q_0))$ and $(L^p(Q;\mathbb{R}^s), L^p(Q_0))$. Then for every $r \in [p,\infty)$ the operator T is continuous with respect to the $(L^r(Q;\mathbb{R}^s), L^r(Q_0))$ topologies.*

Proof. For simplicity we assume $s = 1$ (the proof is the same in the general case). We fix a partition $\{Q_i\}$ of Q and we regularize the operator T with respect to $\{Q_i\}$ (even if we do not write the dependence of $\tilde{T}$ from $\{Q_i\}$ for brevity):

$$\tilde{T}(u)(x) := \fint_{Q_i} |Tu(y) - (Tu)_{Q_i}|\, dy \qquad \forall\, x \in Q_i\ .$$

We claim that $\tilde{T}$ satisfies the assumptions of Marcinkiewicz theorem. Indeed

(1) $\tilde{T}$ is obviously 1-subadditive;

(2) $L^\infty \to L^\infty$ continuity holds by the inequality

$$\|\tilde{T}u\|_{L^\infty} = \sup_i \fint_{Q_i} |Tu(y) - (Tu)_{Q_i}|\, dy \le \|Tu\|_{BMO} \le c\|u\|_{L^\infty}\ ;$$

(3) $L^p \to L^p$ continuity holds too, in fact, by Jensen's inequality,

$$\begin{aligned}
\|\tilde{T}u\|^p_{L^p} &= \sum_i \mathscr{L}^n(Q_i) \left(\fint_{Q_i} |Tu(y) - (Tu)_{Q_i}|\, dy \right)^p \\
&\le \sum_i \int_{Q_i} |Tu(y) - (Tu)_{Q_i}|^p\, dy \\
&\le 2^{p-1} \sum_i \int_{Q_i} \left(|Tu(y)|^p + |(Tu)_{Q_i}|^p \right) dy \le 2^p \|Tu\|^p_{L^p} \le c2^p \|u\|^p_{L^p}\ .
\end{aligned}$$

Thanks to Marcinkiewicz theorem the operator

$$\tilde{T} : D \subset L^r(Q) \longrightarrow L^r(Q_0) \tag{15.4}$$

is continuous for every $r \in [p,\infty]$, and its continuity constant c can be bounded independently of the chosen partition $\{Q_i\}$.

In order to get information from Theorem 15.9, for $r \in [p, \infty)$, we estimate

$$K_r^r(Tu) = \sup_{\{Q_i\}} \sum_i \mathscr{L}^n(Q_i) \left(\fint_{Q_i} |Tu(y) - (Tu)_{Q_i}|\, dy \right)^r = \sup_{\{Q_i\}} \|\tilde{T}_{\{Q_i\}} u\|_{L^r}^r \leq c\|u\|_{L^r} ,$$

where we used the continuity property of $\tilde{T} : L^r(Q) \to L^r(Q_0)$ stated in (15.4). Therefore, by Theorem 15.9, we get

$$\|Tu - (Tu)_Q\|_{L_w^r} \leq c(r, n, T)\|u\|_{L^r} \qquad \forall u \in D .$$

Since $u \mapsto (Tu)_Q$ obviously satisfies a similar L_w^r estimate, we conclude that $\|Tu\|_{L_w^r} \leq c(r, n, T)\|u\|_{L^r}$ 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. □

We are now ready to employ these harmonic analysis tools to the study of regularity in L^p spaces for elliptic PDEs, considering first the case of constant coefficients. Suppose that $\Omega \subset \mathbb{R}^n$ is an open, bounded set with Lipschitz boundary $\partial\Omega$, suppose that the coefficients $A_{ij}^{\alpha\beta}$ satisfy the Legendre-Hadamard condition with $\lambda > 0$ and consider the divergence form of the PDE

$$\begin{cases} -\mathrm{div}(A\nabla u) = \mathrm{div} F \\ u \in H_0^1(\Omega; \mathbb{R}^m) . \end{cases} \tag{15.5}$$

In the spirit of Theorem 15.10, we define

$$TF := \nabla u.$$

Thanks to Campanato regularity theory, we already got the continuity of $T : \mathcal{L}^{2,\lambda} \to \mathcal{L}^{2,\lambda}$ when $0 \leq \lambda < n+2$, thus choosing $\lambda = n$ and using the isomorphism (15.3) we see that T is continuous as an operator

$$T : L^\infty(\Omega; \mathbb{R}^{m\times n}) \longrightarrow BMO(\Omega; \mathbb{R}^{m\times n}) . \tag{15.6}$$

Remark 15.11. Let us remark the importance of weakening the norm in the target space in (15.6): we passed from L^∞ (for which, as we will see, no estimate is possible) to BMO. For BMO the regularity result for PDEs is true and Theorem 15.10 allows us to interpolate between 2 and ∞.

We are going to apply Theorem 15.10 with $D = L^\infty(\Omega; \mathbb{R}^s)$ and $s = m \times n$. By the global Caccioppoli-Leray inequality (see Theorem 6.1) we obtain the second hypothesis of Theorem 15.10: $T : L^2(\Omega; \mathbb{R}^{m\times n}) \to L^2(\Omega; \mathbb{R}^{m\times n})$ is continuous. Therefore

$$T : D \to L^p(\Omega; \mathbb{R}^{m\times n}) \tag{15.7}$$

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 (15.5), we have proved the following result:

Theorem 15.12. *For all $p \in [2, \infty)$ the operator $F \mapsto \nabla u$ in (15.5) maps $L^p(\Omega; \mathbb{R}^{m\times n})$ into $L^p(\Omega; \mathbb{R}^{m\times n})$ continuously.*

Our intention is now to extend the previous result for $p \in (1, 2)$, by a duality argument.

Lemma 15.13 (Helmholtz decomposition)**.** *If $p \geq 2$ and B is a matrix satisfying the Legendre-Hadamard inequality, a map $G \in L^p(\Omega; \mathbb{R}^{m\times n})$ can always be written as a sum*

$$G = B\nabla\phi + \tilde{G} \ , \tag{15.8}$$

where (understanding the divergence w.r.t. the spatial components)

$$\mathrm{div}(\tilde{G}) = 0 \qquad \text{in } \Omega$$

and, for some constant $c_* > 0$, the following inequality holds:

$$\|\nabla\phi\|_{L^p} \leq c_* \|G\|_{L^p} \ . \tag{15.9}$$

Proof. It is sufficient to solve in $H^1_0(\Omega; \mathbb{R}^m)$ the PDE

$$-\mathrm{div}(B\nabla\phi) = -\mathrm{div}G \ .$$

and set $\tilde{G} := G - B\nabla\phi$. The estimate (15.9) is just a consequence of Theorem 15.12. □

Fix $q \in (1, 2)$, so that its conjugated exponent p is larger than 2, and set $D := L^2(\Omega; \mathbb{R}^{m\times n})$. Our aim is to prove that $T : L^2 \to L^q$ is (L^q, L^q)-continuous. We are going to show that, for every $F \in D$, TF belongs to $(L^p)' \sim L^q$. In the chain of inequalities that follows we are using A^*, that is the adjoint matrix of A, which certainly keeps the Legendre-Hadamard property. Lemma 15.13 is used in order to decompose the generic function $G \in L^p$ as in (15.8), so

$$\begin{aligned}
\sup_{\|G\|_{L^p}\leq 1} \langle TF, G\rangle &= \sup_{\|G\|_{L^p}\leq 1} \int TF(x)G(x)\,dx \\
&= \sup_{\|G\|_{L^p}\leq 1} \int \nabla u(x) \left(A^*\nabla\phi(x) + \tilde{G}(x)\right)\,dx \\
&\leq \sup_{\|\nabla\phi\|_{L^p}\leq c_*} \int (A\nabla u(x))\,\nabla\phi(x)\,dx \\
&= \sup_{\|\nabla\phi\|_{L^p}\leq c_*} \int F(x)\nabla\phi(x)\,dx \leq c_*\|F\|_{L^q} \ .
\end{aligned}$$

If we approximate now $F \in L^q$ in the L^q topology by functions $F_n \in L^2$ we can use the (L^q, L^q)-continuity to prove existence of weak solutions to the PDE in $H^{1,q}_0$, when the right hand side is L^q only. Notice that the solutions obtained in this way have no variational character anymore, since their energy $\int A\nabla u\nabla u\,dx$ is infinite (for this reason they are sometimes called *very weak* solutions). Since the variational characterization is lacking, the uniqueness of these solutions needs a new argument, based on Helmholtz decomposition.

Theorem 15.14. *For all $q\in(1,2)$ there exists a continuous operator $T: L^q(\Omega;\mathbb{R}^{m\times n})\to H^{1,q}_0(\Omega;\mathbb{R}^m)$ mapping F to the unique weak solution u to* (15.5).

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 $u\in H^{1,q}_0$ and that $-\mathrm{div}(A\nabla u)=0$ implies $u=0$. To this aim, we define $G=|\nabla u|^{q-2}\nabla u\in L^p$ and we apply Helmholtz decomposition $G=A^*\nabla\phi+\tilde G$ with $\phi\in H^{1,p}_0$ and $\tilde G\in L^p$ divergence-free. By a density argument w.r.t. u and w.r.t. ϕ (notice that the exponents are dual) we have $\int\tilde G\nabla u\,dx=0$ and $\int A\nabla u\nabla\phi\,dx=0$, hence

$$\int_\Omega|\nabla u|^q\,dx=\int_\Omega G\nabla u\,dx=\int_\Omega A^*\nabla\phi\nabla u\,dx=\int_\Omega A\nabla u\nabla\phi\,dx=0.$$

□

Remark 15.15 (General Helmholtz decomposition)**.** Thanks to Theorem 15.14, the Helmholtz decomposition showed above is possible for every $p\in(1,\infty)$.

Remark 15.16 ($W^{2,p}$ estimates)**.** By differentiating the equation and multiplying by cut-off functions, we easily see that Theorem 15.12 and Theorem 15.14 yield

$$-\mathrm{div}(A\nabla u)=f,\quad |\nabla u|\in L^p_{\mathrm{loc}},\ f\in L^p_{\mathrm{loc}}\qquad\Longrightarrow\qquad u\in W^{2,p}_{\mathrm{loc}}.$$

Remark 15.17 (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_i are discontinuous) functions $F_k\in L^\infty(\Omega_i;\mathbb{R}^{m\times n})$ with $\|F_k\|_\infty=1$ and solutions $u_k\in H^1_0(\Omega_k;\mathbb{R}^m)$ to the equation (15.5) 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

$$\|\nabla u\|_{L^1} \le c\|F\|_{L^1}$$

whenever $u \in H^1_0(\Omega;\mathbb{R}^m)$ solves equation (15.5) 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 weak gradient $Du = (D_1 u, \ldots, D_n u)$ is a vector-valued measure satisfying

$$\sum_{\alpha,\beta}\int_\Omega A^{\alpha\beta}\partial_{x_\alpha}\phi\, dD_\beta u = -\sum_\alpha \int_\Omega \partial_{x_\alpha}\phi\, d\mu^\alpha \qquad \forall \phi \in C^\infty_c(\Omega;\mathbb{R}). \tag{15.10}$$

and

$$|Du|(\Omega) \le c|\mu|(\Omega), \tag{15.11}$$

where $|\mu|$ (resp. $|Du|$) denote the total variation of the measure μ (resp. Du). On the other hand, we claim that the inequality (15.11) can't be true. In fact, when $n = 2$ and $m = 1$, consider the identity matrix $A^{\alpha\beta} := \delta^{\alpha\beta}$ and the corresponding Laplace equation

$$-\Delta v = \delta_0\ , \tag{15.12}$$

where δ_0 is the Dirac measure supported in 0. The well-known fundamental solution of (15.12) is

$$v(x) = -\frac{\log|x|}{2\pi} \in C^\infty(\mathbb{R}^2 \setminus \{0\})\ ,$$

so that $v \in W^{1,p}_{\rm loc}(\mathbb{R}^2)$ for any $p < 2$, with $\nabla v(x) = -(2\pi)^{-1}x/|x|^2$, and (understanding the second derivative in the pointwise sense) $|\nabla^2 v| \notin L^1(\Omega)$, since

$$\nabla^2 v(x) = -\frac{1}{2\pi|x|^2}\left(I - 2\frac{x\otimes x}{|x|^2}\right)\ .$$

For any $\eta \in C^\infty_c(\Omega)$ 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 $w = \partial_{x_1}(\eta v) \in L^1(\mathbb{R}^2)$ is a distributional solution in $\mathbb{R}^2$ to the equation

$$-\Delta w = \sum_\alpha \partial_{x_\alpha} \mu^\alpha .$$

It follows that $\tilde{u} = w - u$, with u as in (15.10) is a distributional solution to 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 for $\partial_{x_1} v$ in $B_{1/2}$, and a similar argument gives the same property 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 constant to continuous coefficients, using Korn's technique.

Theorem 15.18. *In an open set $\Omega \subset \mathbb{R}^n$ let $u \in H^1_{\rm loc}(\Omega;\mathbb{R}^m)$ be a solution to the PDE*

$$-\mathrm{div}(A\nabla u) = f + \mathrm{div} F$$

with coefficients $A \in C(\overline{\Omega};\mathbb{R}^{n^2m^2})$ which satisfy a uniform Legendre-Hadamard condition for some $\lambda > 0$. Moreover, if $p \in (1,\infty)$, let us suppose that $F \in L^p_{\rm loc}$ and $f \in L^q_{\rm loc}$, where the Sobolev conjugate exponent $q^* = qn/(n-q)$ coincides with p. Then $|\nabla u| \in L^p_{\rm loc}(\Omega)$.

Proof. We give the proof for $p \geq 2$ (the other cases come again by duality). Let us fix $s \geq 2$ and let us show that

$$|\nabla u| \in L^{s\wedge p}_{\rm loc}(\Omega) \quad \Longrightarrow \quad |\nabla u| \in L^{s^*\wedge p}_{\rm loc}(\Omega)\ . \tag{15.13}$$

Proving (15.13) ends the proof because $|\nabla u| \in L^2_{\rm loc}(\Omega)$ (case $s = 2$) and in finitely many steps s^* becomes larger than p.

Fix a point $x_0 \in \Omega$ and a radius $R > 0$ such that $B_R(x_0) \Subset \Omega$: we choose a cut-off function $\eta \in C^\infty_c(B_R(x_0))$, with $0 \leq \eta \leq 1$ and $\eta \equiv 1$ in $B_{R/2}(x_0)$.
We claim that ηu belongs to $H^{1,s^*\wedge p}_0(B_R(x_0))$ if $R \ll 1$, as it is the unique fixed point of a contraction in $H^{1,s^*\wedge p}_0(B_R(x_0))$, that we are going to define and study in some steps. This implies, in particular, that $|\nabla u| \in L^{s^*\wedge p}(B_{R/2}(x_0))$.
(1) We start localizing the equation. Replacing φ with $\eta\varphi$ in the PDE, by algebraic

computations we obtain

$$\begin{aligned}
&\int_{B_R(x_0)} A(x)\nabla(\eta u)(x)\nabla\varphi(x)\,dx \\
= &\int_{B_R(x_0)} A(x)\left(\eta(x)\nabla u(x) + u(x)\otimes\nabla\eta(x)\right)\nabla\varphi(x)\,dx \\
= &\int_{B_R(x_0)} A(x)\left(\nabla u(x)\nabla(\eta\varphi)(x) + u(x)\otimes\nabla\eta(x)\nabla\varphi(x) - \nabla u(x)\left(\nabla\eta(x)\varphi(x)\right)\right)\,dx \\
= &\int_{B_R(x_0)} f(x)\eta(x)\varphi(x) + F(x)\nabla(\eta\varphi)(x) + A(x)\left(u(x)\otimes\nabla\eta(x)\nabla\varphi(x) - \nabla u(x)\nabla\eta(x)\varphi(x)\right)dx \\
= &\int_{B_R(x_0)} \tilde{f}(x)\varphi(x) + \tilde{F}(x)\nabla\varphi(x)\,dx\ ,
\end{aligned}$$

defining

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

and

$$\tilde{F}(x) := F(x)\eta(x) + A(x)u(x)\otimes\nabla\eta(x)\ .$$

Thus ηu satisfies

$$-\mathrm{div}(A(x_0)\nabla(\eta u)) = \tilde{f} + \mathrm{div}[\tilde{F} + (A - A(x_0))\nabla(\eta u)]. \tag{15.14}$$

(2) In order to write $\tilde{f}$ in divergence form, let us consider the problem

$$\begin{cases} -\Delta w = \tilde{f} \\ w \in H^1_0(\Omega;\mathbb{R}^m) \end{cases}.$$

Thanks to the previous L^p regularity result for constant coefficients PDEs, since $\tilde{f} \in L^{s\wedge q}_{\mathrm{loc}}$ (because we assumed that $|\nabla u| \in L^{s\wedge p}_{\mathrm{loc}}$), we have $|\nabla^2 w| \in L^{s\wedge q}_{\mathrm{loc}}$ (see also Remark 15.16). By Sobolev immersion we get $|\nabla w| \in L^{(s\wedge q)^*}_{\mathrm{loc}}$, hence

$$|\nabla w| \in L^{s^*\wedge q^*}_{\mathrm{loc}} = L^{s^*\wedge p}_{\mathrm{loc}}\ .$$

Now we define

$$F^*(x) := \tilde{F}(x) + \nabla w(x) \in L^{s^*\wedge p}_{\mathrm{loc}}\ .$$

(3) Let $E = H^{1,s^*\wedge p}_0(B_R(x_0);\mathbb{R}^m)$ and let us define the operator $\Theta : E \to E$ which associates to each $V \in E$ the function $v \in E$ that solves

$$-\mathrm{div}\left(A(x_0)\nabla v\right) = \mathrm{div}F^* - \mathrm{div}\left((A(x_0) - A)\nabla V\right)\ . \tag{15.15}$$

The operator Θ is well-defined because $|F^*| \in L^{s^*\wedge p}(B_R(x_0))$ (we saw this in step (2)) and we can take advantage of regularity theory for constant coefficients operators. The operator Θ is a contraction, in fact

$$\|\nabla(v_1 - v_2)\|_E \leq c\| \left(A(x_0) - A\right) \nabla(V_1 - V_2)\|_E \leq \frac{1}{2}\|\nabla(V_1 - V_2)\|_{L^{s^*\wedge p}(B_R(x_0))}$$

if R is sufficiently small, according to the continuity of A. Here we use the fact that the constant c in the first inequality is scale invariant, so it can be "beaten" by the oscillation of A in $B_R(x_0)$, if R is small enough.

Let us call $v_* \in E$ the unique fixed point of (15.15). According to (15.14), ηu already solves (15.15), but in the larger space $H_0^{1,s\wedge p}$. Thus $\eta u \in H_0^{1,s^*\wedge p}$ 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.

Consider the difference $v' := v_* - \eta u \in H_0^{1,s\wedge p}(B_R(x_0); \mathbb{R}^m) \subset H_0^1(B_R(x_0); \mathbb{R}^m)$: v' is a weak solution of

$$-\mathrm{div}\left(A(x)\nabla v'\right) = 0 \ ,$$

hence $v' \equiv 0$ (we can indeed use the variational characterization of the solution). This concludes the proof. □

16 De Giorgi's solution of Hilbert's XIX problem

16.1 The basic estimates

We briefly recall here the setting of Hilbert's XIX problem, that has already been described and solved in dimension 2.

We deal with local minimizers v of scalar functionals

$$v \longmapsto \int_\Omega F(\nabla v)\, dx$$

where $F \in C^{2,\beta}(\mathbb{R}^n)$ (at least, for some $\beta > 0$) satisfies the following ellipticity property: there exist two positive constants $\lambda \leq \Lambda$ such that $\Lambda I \geq \nabla^2 F(p) \geq \lambda I$ for all $p \in \mathbb{R}^n$ (this implies in particular that $|\nabla^2 F|$ is uniformly bounded). We have already seen that under these assumptions it is possible to derive the Euler-Lagrange equations $\mathrm{div} F_p(\nabla v) = 0$. By differentiation, for any direction $s \in \{1, \ldots, n\}$, the equation for $u := \partial v / \partial x_s$ is

$$\frac{\partial}{\partial x_\alpha}\left(F_{p^\alpha p^\beta}\left(\nabla v\right) \frac{\partial}{\partial x_\beta} u\right) = 0 \ .$$

Recall also the fact that, in order to obtain this equation, we needed to work with the approximation $\Delta_{h,s}v$ and with the interpolating operator

$$\widetilde{A}_h(x) := \int_0^1 F_{pp}(t\nabla v(x + he_s) + (1 - t)\nabla v(x))\, dt$$

and to exploit the Caccioppoli-Leray inequality.

One of the striking ideas of De Giorgi was basically to split the problem, that is to deal with u and v separately, as ∇v is only involved in the coefficients of the equation for u. The key point of the regularization procedure is then to show that under no regularity assumption on ∇v (i.e. not more than measurability), if u is a solution of this equation, then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$, with α depending only on n and on the ellipticity constants λ, Λ. If this is true, we can proceed as follows:

$$u \in C^{0,\alpha} \Rightarrow v \in C^{1,\alpha} \Rightarrow F_{pp}(\nabla v) \in C^{0,\alpha\beta} \Rightarrow u \in C^{1,\alpha\beta}\,,$$

where the implications rely upon the fact that F_{pp} is Hölder continuous and on the Schauder estimates of Theorem 10.4. Since u is any partial derivative of v, we eventually get $v \in C^{2,\alpha\beta}$. If F is more regular, by continuing this iteration (now using Shauder regularity for PDE's whose coefficients are $C^{1,\gamma}$, $C^{2,\gamma}$ and so on) we obtain

$$F \in C^\infty \Rightarrow v \in C^\infty$$

and also, by the tools developed in [20], that $F \in C^\omega \Rightarrow v \in C^\omega$, which is the complete solution of the problem raised by Hilbert.

Actually, we have solved this problem in the special case $n = 2$, since, by means of Widman's technique, we could prove that $|\nabla u| \in L^{2,\alpha}$ and hence $u \in \mathcal{L}^{2,\alpha+2}$ for some $\alpha > 0$. This is enough, if $n = 2$, to conclude that $u \in C^{0,\alpha/2}$.

First of all, let us fix our setting. Let Ω be an open domain in $\mathbb{R}^n$, $0 < \lambda \leq \Lambda < \infty$ and let $A^{\alpha\beta}$ be a Borel symmetric matrix satisfying a.e. the condition $\lambda I \leq A(x) \leq \Lambda I$. We want to show that if $u \in H^1_{\mathrm{loc}}$ solves the problem

$$-\mathrm{div}\,(A(x)\nabla u(x)) = 0$$

then $u \in C^{0,\alpha}_{\mathrm{loc}}$. Some notation is needed: for $B_\rho(x) \subset \Omega$ we define

$$A(k, \rho) := \{u > k\} \cap B_\rho(x)\ ,$$

where the dependence on the center x can be omitted. This should not create confusion, since we will often work with a fixed center. In this section, we will derive many functional inequalities, but typically we are not interested in finding the sharpest constants, but only on the functional dependence of these quantities. Therefore, in order to avoid complications of the notation we will use the same symbol (generally c) to indicate *different* constants, possibly varying from one passage to the next one. However we will try to indicate the functional dependence explicitly whenever this is appropriate and so we will use expressions like $c(n)$ or $c(n, \lambda, \Lambda)$ many times.

Theorem 16.1 (Caccioppoli inequality on level sets). *For any $k \in \mathbb{R}$ and $B_\rho(x) \subset B_R(x) \Subset \Omega$ we have*

$$\int_{A(k,\rho)} |\nabla u|^2 \, dy \leq \frac{c}{(R-\rho)^2} \int_{A(k,R)} (u-k)^2 \, dy \tag{16.1}$$

with $c = 16\Lambda^2/\lambda^2$.

Remark 16.2. It should be noted that the previous theorem generalizes the Caccioppoli-Leray inequality, since we do not ask $\rho = R/2$ and we introduce the sublevels.

Theorem 16.3 (Chain rule). *If $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$, then for any $k \in \mathbb{R}$ the function $(u-k)^+$ belongs to $W^{1,1}_{\mathrm{loc}}(\Omega)$. 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 the 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 $\varphi_\varepsilon(t) := \sqrt{t^2+\varepsilon^2} - \varepsilon$ for $t \geq 0$ and identically zero elsewhere, whose derivatives are uniformly bounded and converge to $\chi_{\{t>0\}}$. Moreover, let (u_n) be a sequence of C^1_{loc} functions approximating u in $W^{1,1}_{\mathrm{loc}}$. We have that for any $n \in \mathbb{N}$ and $\varepsilon > 0$ the classical chain rule gives $\nabla\left[\varphi_\varepsilon(u_n)\right] = \varphi'_\varepsilon(u_n)\nabla u_n$. Passing to the limit as $n \to \infty$ gives $\nabla\left[\varphi_\varepsilon(u)\right] = \varphi'_\varepsilon(u)\nabla u$. Now, 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 come to the proof of the Caccioppoli inequality on level sets.

Proof. Let η be a cut-off function supported in $B_R(x)$, with $\eta \equiv 1$ on $\overline{B}_\rho(x)$ and $|\nabla\eta| \leq 2/(R-\rho)$. If we apply the weak form of our equation to the test function $\varphi := \eta^2(u-k)^+$ we get

$$\begin{aligned}\int_{A(k,R)} \eta^2 A\nabla u\nabla u \, dy &= -2\int_{B_R(x)} \eta A\nabla u\nabla\eta(u-k)^+ \, dy \\ &\leq \frac{\Lambda}{\varepsilon}\int_{A(k,R)} \eta^2 \left|Du\right|^2 \, dy + \frac{4\varepsilon\Lambda}{(R-\rho)^2}\int_{A(k,R)} (u-k)^2 \, dy\end{aligned}$$

for any $\varepsilon > 0$, by our upper bound and by Young inequality. Here we set $\varepsilon = 2\Lambda/\lambda$ so that, thanks to the uniform ellipticity assumption, we obtain

$$\int_{A(k,R)} \eta^2 A\nabla u\nabla u \, dy \leq \frac{\lambda}{2}\int_{A(k,R)} \eta^2|\nabla u|^2 \, dy + \frac{8\Lambda^2}{\lambda(R-\rho)^2}\int_{A(k,R)} (u-k)^2 \, dy \, .$$

Since 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 our thesis. □

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 solutions to an elliptic equation. For this reason he introduced a special class of objects.

Definition 16.4 (De Giorgi's class)**.** *We define the De Giorgi class $DG_+(\Omega)$ as follows:*

$$DG_+(\Omega) := \{u \,|\, \exists\, c \in \mathbb{R} \text{ s.t. } \forall\, k \in \mathbb{R}, B_r(x) \Subset B_R(x) \Subset \Omega, u \text{ satisfies } (16.1)\} \ .$$

In this case, we also define $c^+_{DG}(u)$ to be the minimal constant larger than 1 *for which the condition* (16.1) *is verified.*

Remark 16.5. 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

$$-\mathrm{div}\,(A\nabla u) \leq 0 \ \text{ in } \mathcal{D}'(\Omega) \quad \Longrightarrow \quad u \in DG(\Omega),\ c^+_{DG}(u) \leq \frac{16\Lambda^2}{\lambda^2}.$$

In a similar way, the class $DG_-(\Omega)$ (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:

(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 16.8);

(ii) If both u and $-u$ belong to $DG_+(\Omega)$, then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$.

Let us start by discussing the first point. We define these two crucial quantities:

$$U(h,\rho) := \int_{A(h,\rho)} (u-h)^2\, dy, \qquad V(h,\rho) := \mathscr{L}^n\left(A(h,\rho)\right).$$

Theorem 16.6. *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:

$$\begin{aligned} V(h,\rho) &\leq \frac{1}{(h-k)^2}\, U(k,\rho), \\ U(k,\rho) &\leq \frac{c(n)\cdot c^+_{DG}(u)}{(R-\rho)^2}\, U(k,R) V^{2/n}(k,\rho)\,. \end{aligned}$$

Proof. The first statement the first inequality in the second statement are trivial, since

$$\begin{aligned}(h-k)^2V(h,\rho) &= \int_{A(h,\rho)}(h-k)^2\,dy \le \int_{A(h,\rho)}(u-k)^2\,dy\\ &\le \int_{A(k,\rho)}(u-k)^2\,dy = U(k,\rho)\ .\end{aligned}$$

For the second inequality, let us introduce a Lipschitz cut-off function η 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

$$\int_{B_{(R+\rho)/2}}\eta^2|\nabla(u-k)^+|^2\,dy \le \frac{4c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)}(u-k)^2\,dy$$

and

$$\int_{B_{(R+\rho)/2}}((u-k)^+)^2|\nabla\eta|^2\,dy \le \frac{16}{(R-\rho)^2}\int_{A(k,R)}(u-k)^2\,dy\ .$$

Combining these two inequalities, since $c^+_{DG}(u)\ge 1$, we get

$$\int_{B_{(R+\rho)/2}}|\nabla(\eta(u-k)^+)|^2\,dy \le \frac{40c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)}(u-k)^2\,dy$$

and by the Sobolev embedding inequality with the function $\eta(u-k)^+$ this implies

$$\left(\int_{A(k,\rho)}(u-k)^{2^*}\,dy\right)^{2/2^*} \le \frac{c(n)\cdot c^+_{DG}(u)}{(R-\rho)^2}\int_{A(k,R)}(u-k)^2\,dy$$

for some constant $c(n)$ depending on the dimension n. In order to conclude, we just need to apply Hölder's inequality, in fact

$$U(k,\rho)=\int_{A(k,\rho)}(u-k)^2\,dy \le \left(\int_{A(k,\rho)}(u-k)^{2^*}\,dy\right)^{2/2^*}V(k,\rho)^{2/n}$$

with $p=2^*/2=n/(n-2)$, $p'=n/2$. □

The previous inequalities can be slightly weakened, writing

$$\begin{aligned}V(h,\rho) &\le \frac{1}{(h-k)^2}U(k,R),\\ U(h,\rho) &\le \frac{c(n)\cdot c^+_{DG}(u)}{(R-\rho)^2}U(k,R)V^{2/n}(k,R)\end{aligned}$$

and we shall use these to obtain the quantitative maximum principle.

We can view these inequalities as joint decay properties of U and V; in order to get the decay of a single 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:

$$U^\xi(h,\rho)V^\eta(h,\rho) \le \frac{C^\xi}{(h-k)^{2\eta}}\frac{1}{(R-\rho)^{2\xi}}U^{\xi+\eta}(k,R)V^{2\xi/n}(k,R).$$

where $C := c(n)\cdot c^+_{DG}(u)$, a convention that will be systematically adopted in the sequel. Since we are looking for 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 get

$$\theta = \frac{1}{2}+\sqrt{\frac{1}{4}+\frac{2}{n}} \ . \tag{16.2}$$

Note that $\theta > 1$: this fact will play a crucial role in the following proof. In any case, we get the decay relation

$$\varphi(h,\rho) \le \frac{C^\xi}{(h-k)^{2\eta}}\frac{1}{(R-\rho)^{2\xi}}\varphi^\theta(k,R) \ .$$

Theorem 16.7. *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 $\varphi(h_0 + d, R_0/2) = 0$. Moreover, we can take*

$$d^2 = c'(n)[c^+_{DG}(u)]^{n\theta/2}\frac{\varphi(h_0,R_0)^{\theta-1}}{R_0^{n\theta}} \ ,$$

with the constant $c'(n)$ depending only on the dimension n. In particular $u \le h_0 + d$ $\mathscr{L}^n$-a.e. on $B_{R_0/2}(x)$.

Corollary 16.8 (L^2 to L^∞ estimate)**.** *If $u \in DG_+(\Omega)$, then for any $B_{R_0}(x) \subset \Omega$ and for any $h_0 \in \mathbb{R}$*

$$\operatorname*{ess\,sup}_{B_{R_0/2}(x)} u \le h_0 + c''(n)[c^+_{DG}(u)]^{n\theta/4}\left(\frac{1}{\omega_n R_0^n}\int_{A(h_0,R_0)}(u-h_0)^2\,dy\right)^{1/2}\left(\frac{V(h_0,R_0)}{R_0^n}\right)^{(\theta-1)/2} .$$

Proof. This corollary comes immediately from Theorem 16.7, 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 16.9. From Corollary 16.8 with $h_0 = 0$, we can get the *maximum principle* for u, as anticipated above. In fact

$$\operatorname*{ess\,sup}_{B_{R_0/2}(x)} (u^+)^2 \leq q(n)[c^+_{DG}(u)]^{n\theta/2} \fint_{B_{R_0}(x)} u^2 \, dy$$

with $q(n)$ easily estimated in terms of $c''(n)$ and ω_n.

We are now ready to prove Theorem 16.7, the main result of this section.

Proof. Define $k_p := h_0 + d - d/2^p$ and $R_p := R_0/2 + R_0/2^{p+1}$, so that $k_p \uparrow (h_0 + d)$ while $R_p \downarrow R_0/2$. Here $d \in \mathbb{R}$ is a parameter to be fixed in the sequel. From the decay inequality for φ we get

$$\varphi(k_{p+1}, R_{p+1}) \leq \varphi(k_p, R_p) \left[\varphi(k_p, R_p)^{\theta-1} C^\xi \left(\frac{2^{p+2}}{R_0} \right)^{2\xi} \left(\frac{2^{p+1}}{d} \right)^2 \right]$$

and letting $\psi_p := 2^{\mu p} \varphi(k_p, R_p)$ this becomes

$$\psi_{p+1} \leq \psi_p \left[2^\mu C^\xi 2^{4\xi+2} 2^{p(2\xi+2)} R_0^{-2\xi} d^{-2} 2^{-\mu p(\theta-1)} \psi_p^{\theta-1} \right] .$$

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 choose d as follows

$$2^\mu C^\xi 2^{4\xi+2} \psi_0^{\theta-1} R_0^{-2\xi} d^{-2} = 1$$

then $\psi_1 \leq \psi_0$. Hence, $2^\mu C^\xi 2^{4\xi+2} \psi_1^{\theta-1} R_0^{-2\xi} d^{-2} \leq 1$ and the decay inequality yields $\psi_2 \leq \psi_1$. By induction, it follows that $\psi_p \leq \psi_0$, $\forall p \in \mathbb{N}$. In that case, $\varphi(k_p, R_p) \leq 2^{-\mu p} \varphi(h_0, R_0) \to 0$ and, since by monotonicity

$$\varphi(h_0 + d, R_0/2) \leq \varphi(k_p, R_0/2) \leq \varphi(k_p, R_p) ,$$

we get the thesis. But the previous condition on d is satisfied if

$$d^2 \geq c'(n)[c^+_{DG}(u)]^{n\theta/2} R_0^{-2\xi} \psi_0^{\theta-1}$$

and the desired claim follows. □

We are now in position to discuss the notion of *oscillation*, which will be crucial for the conclusion of the argument by De Giorgi.

Definition 16.10. *Let $\Omega \subset \mathbb{R}^n$ be an open set, $B_r(x) \subset \Omega$ and $u : \Omega \to \mathbb{R}$ a measurable function. 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 identifying $\omega(r) = \omega(B_r(x))$.

It is an immediate consequence of the previous results that if $u \in DG_+(\Omega) \cap DG_-(\Omega)$, then

$$\operatorname*{ess\,sup}_{B_{r/2}(x)} u \leq \zeta \left(\fint_{B_r(x)} u^2 \, dy \right)^{1/2}, \qquad -\operatorname*{ess\,inf}_{B_{r/2}(x)} u \leq \zeta \left(\fint_{B_r(x)} u^2 \, dy \right)^{1/2}$$

for a constant ζ, which is a function of the dimension n and of $c_{DG}(u)$. Here and in the sequel we shall denote by $c_{DG}(u)$ the maximum of $c^+_{DG}(u)$ and $c^-_{DG}(u)$ and by $DG(u)$ the intersection of the spaces $DG_+(\Omega)$ and $DG_-(\Omega)$.

Consequently, under the same assumptions,

$$\omega(B_{r/2}(x))(u) \leq 2\zeta \left(\fint_{B_r(x)} u^2 \, dy \right)^{1/2} .$$

Let us see the relation between the decay of the oscillation of u and the Hölder regularity of u. We prove this result passing through the theory of Campanato spaces (a more elementary proof is based on the observation that the Lebesgue representative defined at approximate continuity points is Hölder continuous).

Theorem 16.11. *Let $\Omega \subset \mathbb{R}^n$ be open, $c \geq 0$, $\alpha \in (0, 1]$ and let $u : \Omega \to \mathbb{R}$ be a measurable function such that for any $B_r(x) \subset \Omega$ we have $\omega(B_r(x)) \leq cr^\alpha$. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$, 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 $\operatorname{ess\,inf}_{B_r(x)} u \leq u(y) \leq \operatorname{ess\,sup}_{B_r(x)} u$. These inequalities imply $\operatorname{ess\,inf}_{B_r(x)} u \leq u_{B_r(x)} \leq \operatorname{ess\,sup}_{B_r(x)}$ and hence that $\mathscr{L}^n$-a.e. in $B_r(x)$ the inequality $|u - u_{B_r(x)}| \leq cr^\alpha$ holds. We have proved that $u \in \mathcal{L}^{2,n+2\alpha}(\Omega)$, but this gives $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$ (regularity is local since no assumption is made on Ω), which is the thesis. □

This theorem motivates our interest in the study of oscillation of u, that will be carried on by means of some tools we have not introduced so far.

16.2 Some useful tools

De Giorgi's proof of Hölder continuity is geometric in spirit and ultimately based on the isoperimetric inequality. Notice that, as we will see, the isoperimetric inequality is also underlying the Sobolev inequalities, which we used in the proof of the sup estimate for functions in $DG_+(\Omega)$.

We will 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.

Of course, regular sets are a very unnatural (somehow too restrictive) setting for isoperimetric inequalities, but it is sufficient for our purposes. We state without proof two isoperimetric inequalities:

Theorem 16.12 (Isoperimetric inequality)**.** *Let $E \subset \mathbb{R}^n$ be a regular set such that $\sigma_{n-1}(\partial E) < \infty$. Then*

$$\min\{\mathscr{L}^n(E), \mathscr{L}^n(\mathbb{R}^n \setminus E)\} \leq c(n)\,[\sigma_{n-1}(\partial E)]^{1^*}$$

with $c(n)$ a dimensional constant.

It is also well-known that the best constant $c(n)$ in the previous inequality is

$$\mathscr{L}^n(B_1)/[\sigma_{n-1}(\partial B_1)]^{1^*} = \omega_n/[n\omega_n]^{1^*}\ ,$$

that is, balls have the best isoperimetric ratio.

Theorem 16.13 (Relative isoperimetric inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open and bounded set, with $\partial\Omega$ Lipschitz. Let $E \subset \Omega$ with $\Omega \cap \partial E \in C^1$. Then*

$$\min\{\mathscr{L}^n(E), \mathscr{L}^n(\Omega \setminus E)\} \leq c(\Omega)\,[\sigma_{n-1}(\Omega \cap \partial E)]^{1^*}\ .$$

Let us introduce another classical tool in Geometric Measure Theory.

Theorem 16.14 (Coarea formula)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in C^\infty(\Omega)$ be non-negative, then*

$$\int_\Omega |\nabla u|\,dx = \int_0^\infty \sigma_{n-1}\left(\Omega \cap \{u = t\}\right)\,dt\ .$$

Remark 16.15. It should be observed that the right-hand side of the previous formula is well-defined, since by the classical *Sard's theorem*

$$u \in C^\infty(\Omega) \quad \Longrightarrow \quad \mathscr{L}^1\big(\{u(x):\ x \in \Omega,\ \nabla u(x) = 0\}\big) = 0\ .$$

By the implicit function theorem this implies that almost every sublevel set $\{u < t\}$ is regular.

Proof. A complete proof will not be described here since it is far from the main purpose of these lectures, however we sketch the main points. The interested reader may consult, for instance, [12].

We first prove $\int_\Omega |\nabla u|\,dx \leq \int_0^\infty \sigma_{n-1}\left(\Omega \cap \{u = t\}\right)\,dt$. Consider the pointwise identity

$$u(x) = \int_0^\infty \chi_{\{u>t\}}(x)\,dt$$

that implies

$$\begin{aligned}
\int_\Omega |\nabla u|\,dx &= \sup_{\varphi \in C^1_c,\,|\varphi|\leq 1} \int_\Omega \langle \nabla u, \varphi\rangle\,dx = \sup_{\varphi \in C^1_c,\,|\varphi|\leq 1} \int_\Omega u\,\mathrm{div}\varphi\,dx \\
&= \sup_{\varphi \in C^1_c,\,|\varphi|\leq 1} \int_0^\infty \left(\int_\Omega (\mathrm{div}\varphi)\chi_{\{u>t\}}\,dx\right)\,dt \\
&\leq \int_0^\infty \left(\sup_{\varphi \in C^1_c,\,|\varphi|\leq 1} \int_{\{u>t\}} \mathrm{div}\varphi\,dx\right)\,dt.
\end{aligned}$$

Hence, by the Gauss-Green theorem (with ν_t outer normal to $\{u > t\}$) we obtain

$$\int_\Omega |\nabla u|\, dx \le \int_0^\infty \left(\sup_{\varphi \in C^1_c,\, |\varphi| \le 1} \int_{\Omega \cap \{u=t\}} \langle \varphi, \nu_t \rangle \, d\sigma_{n-1} \right) dt \le \int_0^\infty \sigma_{n-1} \left(\Omega \cap \{u = t\} \right) \, dt \,,$$

again exploiting the fact that for a.e. t the set $\{u = t\}$ is the (regular) boundary of $\{u > t\}$.

Let us consider the converse inequality, namely

$$\int_\Omega |\nabla u|\, dx \ge \int_0^\infty \sigma_{n-1} \left(\Omega \cap \{u = t\} \right) \, dt \ .$$

It is not restrictive to assume that Ω is a cube. This is trivial (with equality) if u is continuous and piecewise linear, since on each part of a triangulation of Ω the coarea formula is just Fubini's Theorem. The general case is obtained by approximation, choosing piecewise affine functions which converge to u in $W^{1,1}(\Omega)$ and using Fatou's lemma and the lower semicontinuity of $E \mapsto \sigma_{n-1}(\Omega \cap \partial E)$ (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 16.16. *Let $G : [0, \infty) \to [0, \infty)$ a non-increasing measurable function. Then for any $\alpha \ge 1$ we have*

$$\alpha \int_0^\infty t^{\alpha - 1} G(t)\, dt \le \left(\int_0^\infty G^{1/\alpha}(t)\, dt \right)^\alpha \ .$$

Proof. It is sufficient to prove that for any $T > 0$ we have the finite time inequality

$$\alpha \int_0^T t^{\alpha - 1} G(t)\, dt \le \left(\int_0^T G^{1/\alpha}(t)\, dt \right)^\alpha \ . \tag{16.3}$$

Since G is non-increasing, we can observe that

$$G^{1/\alpha}(t) \le \fint_0^t G^{1/\alpha}(s)\, ds \ ,$$

which is equivalent to

$$t^{\alpha - 1} G(t) \le \left(\int_0^t G^{1/\alpha}(s)\, ds \right)^{\alpha - 1} G^{1/\alpha}(t) \ .$$

Then, multiplying both sides by α, (16.3) follows by integration. □

We are now ready to derive the Sobolev inequalities stated in Theorem 4.6.

Theorem 16.17 (Sobolev embedding, $p = 1$)**.** *For any $u \in W^{1,1}(\mathbb{R}^n)$ we have that*

$$\left(\int_{\mathbb{R}^n} |u|^{1^*}\, dx\right)^{1/1^*} \leq c(n) \int_{\mathbb{R}^n} |\nabla u|\, dx \ .$$

Consequently, we have the following continuous embeddings:

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

(2) for any $\Omega \subset \mathbb{R}^n$ open, regular and bounded $W^{1,1}(\Omega) \hookrightarrow L^{1^*}(\Omega)$.

Proof. By Theorem 16.3 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

$$\int_{\mathbb{R}^n} u^{1^*}\, dx = 1^* \int_0^\infty t^{1/(n-1)} \mathscr{L}^n(\{u > t\})\, dt \leq \left(\int_0^\infty \mathscr{L}^n(\{u > t\})^{1/1^*}\, dt\right)^{1^*}$$

thanks to Lemma 16.16. Consequently, the isoperimetric inequality and the coarea formula give

$$\int_{\mathbb{R}^n} u^{1^*}\, dx \leq c(n) \left(\int_0^\infty \sigma_{n-1}\big(\{u = t\}\big)\, dt\right)^{1^*} = c(n) \left(\int_{\mathbb{R}^n} |\nabla u|\ dx\right)^{1^*} .$$

The continuous embedding in (2) follows by the global one in (1) applied to an extension of u (recall that regularity of Ω yields the existence of a continuous extension operator from $W^{1,1}(\Omega)$ to $W^{1,1}(\mathbb{R}^n)$). □

Theorem 16.18 (Sobolev embeddings, $1 < p < n$)**.** *For any $u \in W^{1,p}(\mathbb{R}^n)$ we have that*

$$\left(\int_{\mathbb{R}^n} |u|^{p^*}\, dx\right)^{1/p^*} \leq c(n,p) \left(\int_{\mathbb{R}^n} |\nabla u|^p\, dx\right)^{1/p} .$$

Consequently, the have the following continuous embeddings:

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

(2) for any $\Omega \subset \mathbb{R}^n$ open, regular and bounded $W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$.

Proof. Again, it is enough to study the case $u \geq 0$. We can exploit the case $p = 1$ to get

$$\left(\int_{\mathbb{R}^n} u^{\alpha 1^*}\, dx\right)^{1/1^*} \leq c(n) \int_{\mathbb{R}^n} \alpha u^{\alpha-1} |\nabla u|\, dx \qquad \forall\, \alpha > 1$$

and, by Hölder's inequality, the right hand side can be estimated from above with

$$c(n)\alpha \left[\int_{\mathbb{R}^n} u^{(\alpha-1)p'}\, dx\right]^{1/p'} \left[\int_{\mathbb{R}^n} |\nabla u|^p\, dx\right]^{1/p} .$$

Now, choose α such that $\alpha 1^* = (\alpha-1)p'$. Consequently

$$\left(\int_{\mathbb{R}^n} u^{\alpha 1^*}\, dx\right)^{1/1^*-1/p'} \le c(n,p) \left(\int_{\mathbb{R}^n} |\nabla u|^p\, dx\right)^{1/p} ,$$

but $1/1^* - 1/p' = 1/p^*$, $\alpha 1^* = p^*$ and the claim follows. The second part of the statement can be obtained as in Theorem 16.17. □

We will also make use of the following refinement of the Poincaré inequality in $W_0^{1,1}$: 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 16.19. *Let $u \in W^{1,1}(B_R)$ with $u \ge 0$ and suppose that $\mathscr{L}^n(\{u=0\}) \ge \mathscr{L}^n(B_R)/2$. Then*

$$\left(\int_{B_R} u^{1^*}\, dx\right)^{1/1^*} \le c(n) \int_{B_R} |\nabla u|\, dx .$$

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 just needed to mimic the previous argument substituting the isoperimetric inequality with the relative isoperimetric inequality, that is, here

$$\mathscr{L}^n(B_R \cap \{u>t\}) \le c(n)\sigma_{n-1}\left[\mathscr{L}^n(B_R \cap \{u=t\})\right]^{1^*} .$$

We leave the details to the reader. □

16.3 Proof of Hölder continuity

We divide the final part of the proof in two parts.

Lemma 16.20 (Decay of V)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in DG_+(\Omega)$. Suppose that $B_{2r} \Subset \Omega$ and $k_0 < \operatorname{ess\,sup}_{B_{2r}}(u) \le M$ satisfies*

$$V(k_0, r) \le \frac{1}{2}\mathscr{L}^n(B_r) , \tag{16.4}$$

then the sequence of levels $k_\nu = M - (M-k_0)/2^\nu$ for $\nu \ge 0$ satisfies

$$\left(\frac{V(k_\nu, r)}{r^n}\right)^{2(n-1)/n} \le \frac{c(n)\, c^+_{DG}(u)}{\nu} .$$

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 $u \in W^{1,1}(\Omega)$ we also have $\overline{u} \in W^{1,1}(\Omega)$. It is also clear that $\nabla \overline{u} \neq 0$ only on $A(k,r) \setminus A(h,r)$. Notice that

$$\mathscr{L}^n(\{\overline{u} = 0\} \cap B_r) \geq \mathscr{L}^n(\{u \leq k\} \cap B_r) \geq \mathscr{L}^n(\{u \leq k_0\} \cap B_r) \geq \frac{1}{2}\mathscr{L}^n(B_r)$$

and so we can apply the relative version of the critical Sobolev embedding and Hölder's inequality to get

$$\begin{aligned}
(h-k)^{1^*}\mathscr{L}^n(A(h,r)) &= \int_{A(h,r)} \overline{u}^{1^*}\, dy \leq c(n)\left(\int_{B_r} |\nabla \overline{u}|\, dy\right)^{1^*} \\
&= c(n)\int_{A(k,r)\setminus A(h,r)} |\nabla u|\, dy \\
&\leq c(n)\left(\int_{A(k,r)} |\nabla u|^2\, dy\right)^{1^*/2} \mathscr{L}^n(A(k,r)\setminus A(h,r))^{1^*/2}.
\end{aligned}$$

We can now exploit the De Giorgi property of u that is

$$\int_{A(k,r)} |\nabla u|^2\, dy \leq \frac{c^+_{DG}(u)}{r^2}\int_{B_{2r}} (u-k)^2\, dy \leq (M-k)^2\omega_n c^+_{DG}(u) r^{n-2}$$

in order to obtain

$$(h-k)^2\mathscr{L}^n\left(A(h,r)\right)^{2/1^*} \leq c(n)c^+_{DG}(u)(M-k)^2 r^{n-2}(V(k,r)-V(h,r))\ . \tag{16.5}$$

Here we can conclude the proof by applying (16.5) for $h = k_{i+1}$ and $k = k_i$, so that

$$\begin{aligned}
\nu V(k_\nu, r)^{2/1^*} &\leq \sum_{i=1}^{\nu} V(k_i, r)^{2/1^*} \\
&\leq 4c(n)c^+_{DG}(u) r^{n-2}\sum_{i=1}^{\nu}\left[V(k_i,r) - V(k_{i+1},r)\right] \\
&\leq 4c(n)c^+_{DG}(u)\omega_n r^{2n-2}\ .
\end{aligned}$$

□

Theorem 16.21 ($C^{0,\alpha}$ regularity)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and let $u \in DG(\Omega)$. Then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$, with $2\alpha = -\log_2\left(1 - 2^{-(\nu+2)}\right)$,*

$$\nu = 2c(n)\left[c_{DG}(u)\right]^{(n\theta-1)/(\theta-1)} \tag{16.6}$$

and $\theta > 1$ given by (16.2)*, solution to the equation $n\theta(\theta-1) = 2$.*

Proof. Pick an $R > 0$ such that $B_{2R}(x) \Subset \Omega$ and consider for any $r \leq R$ 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 $k_\nu := M(2r) - \frac{\omega(2r)}{2^{\nu+1}}$, but to do this we should check the hypothesis (16.4), which means

$$\mathscr{L}^n(\{u > \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x)).$$

Anyway, either $\mathscr{L}^n(\{u > \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x))$ or $\mathscr{L}^n(\{u < \mu(2r)\} \cap B_r(x)) \leq \frac{1}{2}\mathscr{L}^n(B_r(x))$. The second case is analogous, provided we work with $-u$ instead of u, and it is precisely here that we need the assumption that both u and $-u$ belong to $DG_+(\Omega)$. Using Lemma 16.20 it is easily seen that the choice of ν as in (16.6), with $c(n)$ large enough, provides

$$c''(n)\left[c^+_{DG}(u)\right]^{n\theta/4}\left(\frac{V(k_\nu, r)}{r^n}\right)^{(\theta-1)/2} \leq \frac{1}{2}\,,$$

where $c''(n)$ is the dimensional constant in Theorem 16.8. Moreover, this choice of ν has been made independently of of r and R (this is crucial for the validity of the scheme below).

Now apply the maximum principle in Theorem 16.8 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

$$M\left(\frac{r}{2}\right) \leq h_0 + c''(n)\left[c^+_{DG}(u)\right]^{n\theta/4}(M(2r) - h_0)\left(\frac{V(h_0, r)}{r^n}\right)^{(\theta-1)/2}$$

and, by the appropriate choice of ν that has been described, we deduce

$$M\left(\frac{r}{2}\right) \leq h_0 + \frac{M(2r) - h_0}{2} = \frac{M(2r) + h_0}{2} = M(2r) - \frac{1}{2^{\nu+2}}\omega(2r).$$

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[5], 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 16.11. $\square$

[5] We refer to Lemma 9.1, with the obvious changes.

17 Regularity for systems

17.1 De Giorgi's counterexample to regularity for systems

In the previous section we saw De Giorgi's regularity result for solutions $u \in H^1(\Omega)$ of the elliptic equation

$$\mathrm{div}\,(A(x)\nabla u(x)) = 0$$

with bounded Borel coefficients A satisfying $\lambda I \le A \le \Lambda I$. It turned out that $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$, with $\alpha = \alpha(n, \lambda, \Lambda)$.

It is natural to investigate about similar regularity properties for systems, still under no regularity assumption on A (otherwise, Schauder theory is applicable). In 1968, in [8], Ennio De Giorgi provided a counterexample showing that the scalar case is special. De Giorgi's example not only solves an elliptic PDE, but it is also the minimum of a convex variational problem.

When $m = n$, consider

$$u(x) := x|x|^\alpha \ . \tag{17.1}$$

We will show in (17.7), (17.8) and (17.9) that, choosing

$$\alpha = -\frac{n}{2}\left(1 - \frac{1}{\sqrt{(2n-2)^2+1}}\right) \ , \tag{17.2}$$

the function u is the solution of the Euler-Lagrange equation associated with the uniformly convex functional (here $\nabla \cdot u$ stands for the divergence $\sum_i \partial_{x_i} u^i$)

$$L(u) := \int_{B_1} \left((n-2)\nabla \cdot u(x) + n\frac{x \otimes x}{|x|^2}\nabla u(x)\right)^2 + |\nabla u(x)|^2 \, dx \ . \tag{17.3}$$

If $n \ge 3$ then $|u| \notin L^\infty(B_1)$, because

$$-\alpha = \frac{n}{2}\left(1 - \frac{1}{\sqrt{(2n-2)^2+1}}\right) \ge \frac{3}{2}\left(1 - \frac{1}{\sqrt{17}}\right) > 1$$

and this provides a counterexample not only to Hölder regularity, but also to local boundedness of solutions. In the case $n = 2$ we already know from Widman's technique (see Remark 4.4) that u is locally Hölder continuous, nevertheless De Giorgi's example will show that this regularity cannot be improved to local Lipschitz.

Calling $A(x)$ the matrix such that $L(u) = \int_{B_1} \langle A(x)\nabla u, \nabla u\rangle \, dx$, we remark that A has a discontinuity at the origin (determined by the term $x \otimes x/|x|^2$).

The Euler-Lagrange equation associated to (17.3) is the following (in the weak distributional sense): for every $h = 1, \ldots, n$ it must be

$$
\begin{aligned}
0 &= (n-2)\frac{\partial}{\partial x_h}\left((n-2)\sum_{t=1}^{n}\frac{\partial u^t}{\partial x_t} + n\sum_{s,t=1}^{n}\frac{x_s x_t}{|x|^2}\frac{\partial u^t}{\partial x_s}\right) &(17.4)\\
&+ n\sum_{k=1}^{n}\frac{\partial}{\partial x_k}\left[\frac{x_h x_k}{|x|^2}\left((n-2)\sum_{t=1}^{n}\frac{\partial u^t}{\partial x_t} + n\sum_{s,t=1}^{n}\frac{x_s x_t}{|x|^2}\frac{\partial u^t}{\partial x_s}\right)\right] &(17.5)\\
&+ \sum_{k=1}^{n}\frac{\partial^2 u^h}{\partial x_k^2}\,. &(17.6)
\end{aligned}
$$

We are going to prove in a few steps that u is the unique minimizer of L, with boundary data given by u itself, and that u solves the Euler-Lagrange equations. The steps are:

(i) u, as defined in (17.1), belongs to $C^\infty(B_1 \setminus \{0\}; \mathbb{R}^n)$ and solves in $B_1 \setminus \{0\}$ the Euler-Lagrange equations;

(ii) $u \in H^1(B_1; \mathbb{R}^n)$ and is also a weak solution in B_1 of the system.

Let us perform step (i). Fix $h \in \{1, \ldots, n\}$, and use extensively the identity

$$\frac{\partial}{\partial x_h}|x|^\alpha = \alpha x_h |x|^{\alpha-2}\,.$$

Then $\Delta|x|^\alpha = (n\alpha + \alpha^2 - 2\alpha)|x|^{\alpha-2}$ and

$$\Delta\left(x_h|x|^\alpha\right) = x_h\Delta|x|^\alpha + \frac{\partial}{\partial x_h}|x|^\alpha = (\alpha n + \alpha^2)x_h|x|^{\alpha-2} \tag{17.7}$$

and this is what we need to put in (17.6) when u is given by (17.1). For both (17.4) and (17.5) we have to calculate

$$\sum_{t=1}^{n}\frac{\partial}{\partial x_t}\left(x_t|x|^\alpha\right) = (n+\alpha)|x|^\alpha\,,$$

and

$$\sum_{s,t=1}^{n}\frac{x_s x_t}{|x|^2}\frac{\partial u^t}{\partial x_s} = \sum_{s,t=1}^{n}\frac{x_s x_t}{|x|^2}\left(\alpha x_s x_t|x|^{\alpha-2} + \delta_{st}|x|^\alpha\right) = (\alpha+1)|x|^\alpha\,.$$

Therefore (17.4) is given by

$$(n-2)\frac{\partial}{\partial x_h}\left((n-2)\sum_{t=1}^{n}\frac{\partial u^t}{\partial x_t} + n\sum_{s,t=1}^{n}\frac{x_s x_t}{|x|^2}\frac{\partial u^t}{\partial x_s}\right) = \alpha(n-2)[(n-2)(n+\alpha)+n(\alpha+1)]x_h|x|^{\alpha-2}\,. \tag{17.8}$$

In order to compute the term (17.5) we first get

$$\sum_{k=1}^{n} \frac{\partial}{\partial x_k} \left(x_h x_k |x|^{\alpha-2} \right) = (n+\alpha-1) x_h |x|^{\alpha-2} \ ,$$

and therefore we obtain

$$n \sum_{k=1}^{n} \frac{\partial}{\partial x_k} \left[\frac{x_h x_k}{|x|^2} \left((n-2) \sum_{t=1}^{n} \frac{\partial u^t}{\partial x_t} + n \sum_{s,t=1}^{n} \frac{x_s x_t}{|x|^2} \frac{\partial u^t}{\partial x_s} \right) \right]$$
$$= n(n+\alpha-1)[(n-2)(n+\alpha) + n(\alpha+1)] x_h |x|^{\alpha-2} \ . \quad (17.9)$$

Putting together (17.7), (17.8) and (17.9), $u(x) = x|x|^{\alpha}$ solves the Euler-Lagrange equation if and only if

$$(2n-2)^2 \left(\alpha + \frac{n}{2} \right)^2 + \alpha n + \alpha^2 = 0 \ ,$$

which leads to the choice (17.2) of α.

Let us now perform step (ii), checking first that $u \in H^1$. As $|\nabla u(x)| \sim |x|^{\alpha}$ and $2\alpha > -n$, it is easy to show that $|\nabla u| \in L^2(B_1)$. Moreover, for every $\varphi \in C_c^{\infty}(B_1 \setminus \{0\})$ we have classical integration by parts formula

$$\int \nabla u(x) \varphi(x) \, dx = - \int u(x) \nabla \varphi(x) \, dx \ . \quad (17.10)$$

Thanks to Lemma 17.1 below, we are allowed to approximate in $H^1(B_1; \mathbb{R}^n)$ norm every $\varphi \in C_c^{\infty}(B_1)$ with a sequence $(\varphi_k) \subset C_c^{\infty}(B_1 \setminus \{0\})$. Then we can pass to the limit in (17.10) because $|\nabla u| \in L^2(B_1)$ to obtain $u \in H^1(B_1; \mathbb{R}^m)$. Now, using the fact that the Euler-Lagrange PDE holds in the weak sense in $B_1 \setminus \{0\}$ (because it holds in the classical sense), we have

$$\int_{B_1} A(x) \nabla u(x) \nabla \varphi(x) \, dx = 0 \quad (17.11)$$

for every $\varphi \in C_c^{\infty}(B_1 \setminus \{0\}; \mathbb{R}^n)$. Using Lemma 17.1 again, we can extend (17.11) to every $\varphi \in C_c^{\infty}(B_1; \mathbb{R}^n)$, thus obtaining the validity of the Euler-Lagrange PDE in the weak sense in the whole ball.

Finally, since the functional L in (17.3) is convex, the Euler-Lagrange equation is satisfied by u if and only if u is a minimizer of $L(u)$ with boundary condition

$$u(x) = x \quad \text{in } \partial B_1 \ .$$

This means that De Giorgi's counterexample holds not only for solution of the Euler-Lagrange equation, but also for minimizers.

Lemma 17.1. *Assume that $n > 2$. For every $\varphi \in C_c^\infty(B_1)$ there exists $\varphi_k \in C_c^\infty(B_1 \setminus \{0\})$ such that φ_k tends to φ strongly in $W^{1,2}(B_1)$.*

Proof. Consider $\psi \in C_c^\infty(\mathbb{R}^n)$ 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 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 thesis is equivalent to verify that

$$\int_{B_1} \varphi(x)^2 |\nabla\psi_k(x)|^2 \, dx \to 0 \ ,$$

but

$$\begin{aligned} \int_{B_1} \varphi(x)^2 |\nabla\psi_k(x)|^2 \, dx &\le (\sup \varphi^2) k^2 \int_{B_1} |\nabla\psi(kx)|^2 \, dx \\ &\le (\sup \varphi^2) k^{2-n} \int_{\mathbb{R}^n} |\nabla\psi(x)|^2 \, dx \longrightarrow 0 \ , \end{aligned}$$

where we used the fact that $n > 2$. $\square$

We conclude noticing that the restriction $n \ge 3$ in the proof of Lemma 17.1 is not really needed. Indeed, when $n = 2$ we have

$$\inf \left\{ \int |\nabla\psi(x)|^2 \, dx \mid \psi \in C_c^\infty(B_1), \ \psi = 1 \text{ in a neighbourhood of } 0 \right\} = 0 \ . \qquad (17.12)$$

Let us prove (17.12): we first prove that

$$\inf \left\{ \int_0^1 r|a'(r)|^2 \, dr \mid a(0) = 1, \ a(1) = 0 \right\} = 0 \ ,$$

considering radial functions $\psi(x) = a(|x|)$. We can take $a_\gamma(r) := 1 - r^\gamma$, so

$$\int_0^1 r|a'_\gamma(r)|^2 \, dr = \frac{\gamma}{2} \overset{\gamma \to 0}{\longrightarrow} 0 \ .$$

Then, 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 neighbourhood of 0) we prove (17.12).

Using (17.12) to remove the point singularity also in the case $n = 2$, it follows that the functional $L(u)$ and its minimizer are a counterexample to Lipschitz regularity.

In a more general perspective, we recall that the p-capacity of a compact set $K \subset \mathbb{R}^n$ is defined by

$$\inf \left\{ \int_{\mathbb{R}^n} |\nabla\phi|^p \, dx \mid \phi \in C_c^\infty(\mathbb{R}^n), \ \phi \equiv 1 \text{ in a neighbourhood of } K \right\} .$$

We proved that singletons have null 2-capacity in $\mathbb{R}^n$ for $n \ge 2$.

18 Partial regularity for systems

As we have seen with De Giorgi's counterexample, it is impossible to expect an "everywhere" regularity result for elliptic systems: the main idea is to pursue a different goal, a "partial" regularity result, away from a singular set. This strategy goes back to De Giorgi himself, and it was implemented for the first time in the study of minimal surfaces.

Definition 18.1 (Regular and singular sets)**.** *For a generic function $u : \Omega \to \mathbb{R}$ we call* regular set *of u the set*

$$\Omega_{\mathrm{reg}}(u) := \left\{ x \in \Omega \,\middle|\, \exists\, r > 0 \text{ s.t. } B_r(x) \subset \Omega \text{ and } u \in C^1\left(B_r(x)\right) \right\} .$$

Analogously, the singular set *is*

$$\Sigma(u) := \Omega \setminus \Omega_{\mathrm{reg}}(u) .$$

The set $\Omega_{\mathrm{reg}}(u)$ is obviously the largest open subset A of Ω such that $u \in C^1(A)$. Briefly, let us recall here the main results we have already obtained for elliptic systems.

(a) If we are looking at the problem from the variational point of view, studying local minimizers $u \in H^1_{\mathrm{loc}}$ of $v \mapsto \int_\Omega F(Dv)\,dx$, with $F \in C^2(\mathbb{R}^{m\times n})$, $|D^2F(p)| \le \Lambda$, we already have the validity of the Euler-Lagrange equations. More precisely, if

$$\int_{\Omega'} F(\nabla u(x))\,dx \le \int_{\Omega'} F(\nabla v(x))\,dx \qquad \forall\, v \text{ s.t. } \{u \ne v\} \Subset \Omega' \Subset \Omega ,$$

then

$$\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha}(\nabla u)\right) = 0 \quad \forall\, i = 1, \ldots, m .$$

(b) If F satisfies a uniform Legendre condition for some $\lambda > 0$, by Nirenberg method we have $\nabla u \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^{m\times n})$ and (by differentiation of the (EL) equations with respect to x_γ)

$$\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha p_j^\beta}(\nabla u)\frac{\partial^2 u^j}{\partial x_\beta \partial x_\gamma}\right) = 0 \quad \forall\, i = 1, \ldots, m, \gamma = 1, \ldots, n . \tag{18.1}$$

Definition 18.2 (Uniform quasiconvexity)**.** *We say that F is λ-uniformly quasiconvex if*

$$\int_\Omega F(A + \nabla\varphi(x)) - F(A)\,dx \ge \lambda \int_\Omega |\nabla\varphi|^2\,dx \qquad \forall\, \varphi \in C^\infty_c(\Omega; \mathbb{R}^m) .$$

In this section we shall provide a fairly complete proof of the following result, following with minor variants the original proof in [10].

Theorem 18.3 (Evans). *If $F \in C^2(\mathbb{R}^{m\times n})$ is λ-uniformly quasiconvex with $\lambda > 0$ and satisfies*

$$|\nabla^2 F(p)| \leq \Lambda \qquad \forall p \in \mathbb{R}^{m\times n}\,, \tag{18.2}$$

for some $\Lambda > 0$. Then any local minimizer u belongs to $C^{1,\gamma}(\Omega_{\mathrm{reg}})$ for some $\gamma = \gamma(n, m, \lambda, \Lambda)$ and

$$\mathscr{L}^n(\Omega \setminus \Omega_{\mathrm{reg}}) = 0\ .$$

The following list summarizes some results in the spirit of Theorem 18.3. At this stage we should point out that the growth condition (18.2) is a bit restrictive if we want to allow the standard examples of quasiconvex functions, i.e. convex functions of determinants of minors of ∇u; it includes for instance functions of the form

$$F(\nabla u) := |\nabla u|^2 + \sqrt{1 + \sum_M (M\nabla u)^2}$$

where $M\nabla u$ is a 2×2 minor of ∇u.

A more general growth condition considered in [10] is

$$|\nabla^2 F(p)| \leq C_0\left(1 + |p|^{q-2}\right) \qquad \text{with } q \geq 2\ , \tag{18.3}$$

which leads to the estimates $|\nabla F(p)| \leq C_1(1 + |p|^{q-1})$ and $|F(p)| \leq C_2(1 + |p|^q)$.

(i) If $\nabla^2 F \geq \lambda I$ for some $\lambda > 0$, then Giaquinta and Giusti (see [16] and [18]) proved a much stronger estimate on the size of the singular set, namely (here $\mathscr{H}^k$ denotes the Hausdorff measure, that we will introduce later on)

$$\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0 \qquad \forall\, \varepsilon > 0\ .$$

(ii) If $\nabla^2 F \geq \lambda I$ for some $\lambda > 0$ and it is globally uniformly continuous, then we have even $\mathscr{H}^{n-2}(\Sigma(u)) = 0$.

(iii) If u is locally Lipschitz, then Kristensen and Mingione proved in [23] that there exists $\delta > 0$ such that

$$\mathscr{H}^{n-\delta}(\Sigma(u)) = 0\ .$$

(iv) On the contrary, when $n = 2$ and $m = 3$, there exists a Lipschitz solution u for the system $\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha}(\nabla u)\right)$ (with F smooth and satisfying the Legendre-Hadamard condition), provided in [25], such that

$$\Omega_{\mathrm{reg}}(u) = \emptyset\ .$$

This last result clarifies once 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 18.4. *There exists a constant $C_* = C_*(n, m, \lambda, \Lambda) \in (0, 1)$ such that, for every constant matrix A satisfying the Legendre-Hadamard condition with λ and the inequality $|A| \le \Lambda$, any solution $u \in H^1(B_r; \mathbb{R}^m)$ of*

$$\operatorname{div}(A\nabla u) = 0 \quad \text{in } B_r$$

satisfies

$$⨍_{B_{\alpha r}} |\nabla u(x) - (\nabla u)_{B_{\alpha r}}|^2 \, dx \le C_* \alpha^2 ⨍_{B_r} |\nabla u(x) - (\nabla u)_{B_r}|^2 \, dx \qquad \forall \alpha \in (0, 1) \ .$$

Proof. As a consequence of what we proved in the section about decay estimates for systems with constant coefficients, considering (5.2) with $\rho = \alpha r$ and $\alpha < 1$, we have that

$$\int_{B_{\alpha r}} |\nabla u(x) - (\nabla u)_{B_{\alpha r}}|^2 \, dx \le c(n, m, \lambda, \Lambda) \left(\frac{\alpha r}{r}\right)^{n+2} \int_{B_r} |\nabla u(x) - (\nabla u)_{B_r}|^2 \, dx \ . \tag{18.4}$$

It is enough to consider the mean of (18.4), so that

$$⨍_{B_{\alpha r}} |\nabla u(x) - (\nabla u)_{B_{\alpha r}}|^2 \, dx \le c(n, m, \lambda, \Lambda) \alpha^2 ⨍_{B_r} |\nabla u(x) - (\nabla u)_{B_r}|^2 \, dx \ .$$

□

Definition 18.5 (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*

$$\operatorname{Exc}(u, B_\rho(x)) := \left(⨍_{B_\rho(x)} |\nabla u(y) - (\nabla u)_{B_\rho(x)}|^2 \, dy \right)^{1/2} \ .$$

When we consider functions F satisfying the more general growth condition (18.3), then we should modify the definition of excess as follows, see [10]:

$$\operatorname{Exc}(u, B_\rho(x))^2 = ⨍_{B_\rho(x)} \left(1 + |\nabla u(y) - (\nabla u)_{B_\rho(x)}|^{q-2}\right) |\nabla u(y) - (\nabla u)_{B_\rho(x)}|^2 \ dy.$$

However, in our presentation we will cover only the case $q = 2$.

Remark 18.6 (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

$$\operatorname{Exc}(u+p, B_\rho(x)) = \operatorname{Exc}(u, B_\rho(x)) \ .$$

(ii) The excess is positively 1-homogeneous, that is for any number $\lambda \geq 0$

$$\operatorname{Exc}(\lambda u, B_\rho(x)) = \lambda \operatorname{Exc}(u, B_\rho(x)) \ .$$

(iii) We have the following scaling property:

$$\operatorname{Exc}\left(\frac{u(\rho \cdot +x_0)}{\rho}, B_1(0)\right) = \operatorname{Exc}(u, B_\rho(x_0)) \ .$$

Remark 18.7. The name "excess" is inspired by De Giorgi's theory of minimal surfaces, presented in [6] and [7], see also [15] for a modern presentation. The excess of a set E at a point is defined (for regular sets) by

$$\operatorname{Exc}(E, B_\rho(x)) := \fint_{B_\rho(x)\cap\partial E} |\nu_E(y) - \nu_E(x)|^2 \, d\sigma_{n-1}(y) \ ,$$

where ν_E is the inner normal of the set E. The correspondence between $\operatorname{Exc}(u, B_\rho(x))$ and $\operatorname{Exc}(E, B_\rho(x))$ can be made more evident seeing near x the set ∂E as the graph associated to a function u, in a coordinate system where $\nabla u(x) = 0$. Indeed, the identity $\nu_E = (-\nabla u, 1)/\sqrt{1+|\nabla u|^2}$ and the area formula for graphs give

$$\int_{B_\rho(x)\cap\partial E} |\nu_E(y) - \nu_E(x)|^2 \, d\sigma_{n-1}(y) = 2\int_{\pi(B_\rho(x)\cap\partial E)} \left(\sqrt{1+|\nabla u(z)|^2} - 1\right) dz \sim \int_{B_\rho(z)} |\nabla u(z)|^2 \, dz \, ,$$

where $\pi(B_\rho(x)\cap\partial E)$ denotes the projection of the $B_\rho(x)\cap\partial E$ on the hyperplane.

The main ingredient in the proof of Evans' theorem will be the decay property of the excess: there exists a critical threshold such that, if the decay in the ball is below the threshold, then decay occurs in the smaller balls.

Theorem 18.8 (Excess decay)**.** *Let F be as in Theorem 18.3. 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* $v \mapsto \int F(\nabla v)\, dx$*,*

(b) $|(\nabla u)_{B_r(x)}| \leq M$*,*

(c) $\operatorname{Exc}(u, B_r(x)) < \varepsilon_0$*,*

then

$$\operatorname{Exc}(u, B_{\alpha r}(x)) \le C_e \alpha \operatorname{Exc}(u, B_r(x))$$

with C_e depending only on (n, m, λ, Λ). When $\nabla^2 F$ is uniformly continuous, condition (b) is not needed for the validity of the implication and ε_0 is independent of M.

Proof. We choose C_e in such a way that $16C_*^2 C_P C^* < C_e^2$, where C_* is the constant of Lemma 18.4, C_P is the constant in the Poincaré inequality and C^* is the constant of Proposition 18.9 below.

The proof is by contradiction, assuming that the statement fails for some α and M (for simplicity we keep F fixed in the contradiction argument, but a slightly more complex proof would give the stronger result): in step (ii) we will normalize the excesses, obtaining functions w_k with $\operatorname{Exc}(w_k, B_\alpha(0)) \ge C_e \alpha$ while $\operatorname{Exc}(w_k, B_1(0)) = 1$. Each w_k is a solution of

$$\frac{\partial}{\partial x_\alpha}\left(F_{p_i^\alpha}(\nabla w_k)\right) = 0 \ .$$

We will see in step (iii) that, passing through the limit as $k \to \infty$, any limit point w_∞ w.r.t. the weak H^1 topology solves

$$\operatorname{div}\left(F_{p_i^\alpha p_j^\beta}(p_\infty)\nabla w_\infty\right) = 0.$$

Using Lemma 18.4 in combination with Proposition 18.9 we will reach the contradiction.
(i) By contradiction, we have $M \ge 0$, $\alpha \in (0, 1/4)$ and local minimizers $u_k : \Omega \to \mathbb{R}^m$ in $B_{r_k}(x_k)$ with

$$\varepsilon_k := \operatorname{Exc}(u_k, B_{r_k}(x_k)) \longrightarrow 0$$

satisfying

$$\left|(\nabla u_k)_{B_{r_k}(x_k)}\right| \le M \tag{18.5}$$

but

$$\operatorname{Exc}(u_k, B_{\alpha r_k}(x_k)) > C_e \alpha \operatorname{Exc}(u_k, B_{r_k}(x_k)) \qquad \forall\, k \in \mathbb{N} \ .$$

(ii) Suitably rescaling and translating the functions u_k, we can assume that $x_k = 0$, $r_k = 1$ and $(u_k)_{B_1} = 0$ for all k. Setting $p_k := (\nabla u_k)_{B_1}$, the hypothesis (18.5) gives, up to subsequences,

$$p_k \longrightarrow p_\infty \in \mathbb{R}^{m \times n} \ . \tag{18.6}$$

We start here a parallel and simpler path through this proof, in the case when $\nabla^2 F$ is uniformly continuous: in this case no uniform bound on p_k is needed and we can replace (18.6) with

$$\nabla^2 F(p_k) \to A_\infty \in \mathbb{R}^{m^2 \times n^2} \ . \tag{18.7}$$

Notice that (18.7) holds under (18.6), simply with $A_\infty = \nabla^2 F(p_\infty)$. Notice also that, in any case, A_∞ satisfies a (LH) condition with constant λ (this can be achieved using oscillating test functions, as we did to show that quasi-convexity implies the Legendre-Hadamard condition) and $|A_\infty| \le \Lambda$

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 $(v_k)_{B_1} = 0$ and $(\nabla v_k)_{B_1} = 0$. According to property (i) of Remark 18.6 the excess does not change, so still
$$\mathrm{Exc}\,(v_k, B_1) = \varepsilon_k \longrightarrow 0$$
and
$$\mathrm{Exc}\,(v_k, B_\alpha) > C_e \alpha\, \varepsilon_k \ .$$
During these operations, we need not lose sight of the variational problem we are solving, for example every function v_k minimizes the integral functional associated to
$$p \mapsto F(p + p_k) - F(p_k) - \nabla F(p_k) p \ .$$
In order to get some contradiction, our aim is to find a "limit problem" with some decaying property. 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
$$\mathrm{Exc}\,(w_k, B_1) = 1 \qquad \text{and} \qquad \mathrm{Exc}\,(w_k, B_\alpha) > C_e \alpha \ . \tag{18.8}$$
The key point of the proof is that w_k is a local minimizer of $v \mapsto \int F_k(\nabla v)\, dx$, where
$$F_k(p) := \frac{1}{\varepsilon_k^2} \left[F(\varepsilon_k p + p_k) - F(p_k) - \nabla F(p_k) \varepsilon_k p \right] \ .$$
Here we used the fact local minimality w.r.t. to an integrand F is preserved if we multiply F by a positive constant or add to F an affine function.
(iii) We now study both the limit of F_k and the limit of w_k, as $k \to \infty$. Since $F_k \in C^2(\mathbb{R}^{m \times n})$, by Taylor expansion we are able to identify a limit Lagrangian, given by
$$F_\infty(p) = \frac{1}{2} \langle A_\infty p, p \rangle \ ,$$
to which $F_k(p)$ converge uniformly on compact subsets of $\mathbb{R}^{m\times n}$. Indeed, this is clear with $A_\infty = \nabla^2 F(p_\infty)$ in the case when $p_k \to p_\infty$; it is still true with A_∞ given by (18.7) when $\nabla^2 F$ is uniformly continuous, writing $F_k(p) = \frac{1}{2}\langle \nabla^2 F(p_k + \theta \varepsilon_k p) p, p \rangle$ with $\theta = \theta(k,p) \in (0,1)$.

Once we have the limit problem defined by F_∞, we drive our attention to w_k: it is a bounded sequence in $H^{1,2}(B_1;\mathbb{R}^m)$ because the excesses are constant, so by Rellich theorem we have that (possibly extracting one more subsequence)

$$w_k \longrightarrow w_\infty \qquad \text{in } L^2(B_1;\mathbb{R}^m)$$

and, as a consequence,

$$\nabla w_k \rightharpoonup \nabla w_\infty \qquad \text{in } L^2(B_1;\mathbb{R}^m)\ . \tag{18.9}$$

The analysis of the limit problem now requires the verification that w_∞ solves the Euler equation associated to F_∞. We need just to pass to the limit in the (EL) equation satisfied by w_k, namely

$$\sum_{\alpha,i}\int_{B_1}\frac{1}{\varepsilon_k}\left(\frac{\partial F}{\partial p_i^\alpha}(p_k+\varepsilon_k\nabla w_k(x))-\frac{\partial F}{\partial p_i^\alpha}(p_k)\right)\frac{\partial \phi^i}{\partial x_\alpha}(x)\,dx=0 \qquad \forall\,\varphi\in C^\infty_c(B_1;\mathbb{R}^m)\ .$$

Writing the difference quotient of ∇F with the mean value theorem and using $\nabla^2 F(p_k)\to A_\infty$ we obtain

$$\int_{B_1}\langle A_\infty\nabla w_\infty(x),\nabla\varphi(x)\rangle\,dx=0 \qquad \forall\,\varphi\in C^\infty_c(B_1;\mathbb{R}^m)\ , \tag{18.10}$$

provided we show that (here $\theta=\theta(x,\alpha,\beta)\in(0,1)$)

$$\lim_{k\to\infty}\sum_{\alpha,\beta}\sum_{i,j}\int_{B_1}|\frac{\partial^2 F}{\partial p_i^\alpha p_j^\beta}(p_k+\theta\varepsilon_k\nabla w_k)-(A_\infty)_{ij}^{\alpha\beta}|\,dx=0\,.$$

This can be obtained splitting the integral into the regions $\{|\nabla w_k|\le L\}$ and $\{|\nabla w_k|>L\}$, with L 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 $\nabla^2 F$. The second contribution tends to 0 as $L\uparrow\infty$ uniformly in k, since $|\nabla^2 F|\le\Lambda$ and $\|\nabla w_k\|_2\le 1$.
(iv) Equality (18.10) means that

$$\mathrm{div}\,(A_\infty\nabla w_\infty)=0$$

in a weak sense: since the equation has constant coefficients we can apply Lemma 18.4 to get

$$\fint_{B_{2\alpha}}|\nabla w_\infty(x)-(\nabla w_\infty)_{B_{2\alpha}}|^2\,dx\le 4C_*^2\alpha^2\fint_{B_1}|\nabla w_\infty(x)|^2\,dx\le 4C_*^2\alpha^2. \tag{18.11}$$

On the other hand, using Proposition 18.9 below we get

$$C_e^2\alpha^2<\big(\mathrm{Exc}\,(w_k,B_\alpha)\big)^2\le\frac{C^*}{\alpha^2}\fint_{B_{2\alpha}}|w_k-(w_k)_{2\alpha}-(\nabla w_k)_{2\alpha}(x)|^2\,dx\ ,$$

hence passing to the limit as $k \to \infty$ gives

$$\frac{C_e^2}{C^*}\alpha^4 \le \fint_{B_{2\alpha}} |w_\infty - (w_\infty)_{2\alpha} - (\nabla w_\infty)_{2\alpha}(x)|^2\,dx\ .$$

On the other hand, the Poincaré inequality and (18.11) gives

$$\fint_{B_{2\alpha}} |w_\infty - (w_\infty)_{2\alpha} - (\nabla w_\infty)_{2\alpha}(x)|^2\,dx \le 4C_P\alpha^2 \fint_{B_{2\alpha}} |\nabla w_\infty - (\nabla w_\infty)_{2\alpha}|^2\,dx \le 16 C_P C_*^2\alpha^4\ .$$

Taking into account our definition of C_e 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 using the minimality instead.

Proposition 18.9 (Caccioppoli inequality for minimizers)**.** *There exists $C^* = C(n,m,\lambda,\Lambda)$ such that if F is λ-quasiconvex with $|\nabla^2 F| \le \Lambda$ and if u is a local minimizer in Ω, then*

$$\fint_{B_{r/2}(x_0)} |\nabla u - A|^2\,dx \le \frac{C^*}{r^2}\fint_{B_r(x_0)} |u - a - A(x-x_0)|^2\,dx$$

for all balls $B_r(x_0) \Subset \Omega$, all $A \in \mathbb{R}^{m\times n}$ and $a \in \mathbb{R}^m$.

Proof. By translation invariance we can assume $a = 0$, $x_0 = 0$. Let $r/2 \le t < s \le r$ and let $\zeta \in C_c^\infty(B_s)$ with $\zeta \equiv 1$ on B_t, $0 \le \zeta \le 1$ and $|\nabla\zeta| \le 2(s-t)$. Define $\phi = \zeta(u - Ax)$, $\psi = (1-\zeta)(u - Ax)$, so that $\phi + \psi = u - Ax$ gives

$$\nabla\phi + \nabla\psi = \nabla u - A\ .$$

From the λ-uniform quasiconvexity we get

$$\begin{aligned}\int_{B_s} F(A) + \lambda|\nabla\phi|^2\,dx &\le \int_{B_s} F(A+\nabla\phi)\,dx\\ &= \int_{B_s} F(\nabla u - \nabla\psi)\,dx \\ &\le \int_{B_s} F(\nabla u) - \nabla F(\nabla u)\nabla\psi + C|\nabla\psi|^2\,dx\ ,\end{aligned} \tag{18.12}$$

with $C = C(\Lambda)$. On the other hand, since u is a local minimum, we have

$$\begin{aligned}\int_{B_s} F(\nabla u)\,dx &\le \int_{B_s} F(\nabla u - \nabla\phi)\,dx\\ &= \int_{B_s} F(A+\nabla\psi)\,dx \\ &\le \int_{B_s} F(A) + \nabla F(A)\nabla\psi + C|\nabla\psi|^2\,dx\ .\end{aligned} \tag{18.13}$$

Combining (18.12) with (18.13) we get

$$\lambda \int_{B_s} |\nabla \phi|^2 \, dx \leq \int_{B_s} |\nabla F(A) - \nabla F(\nabla u)||\nabla \psi| + C|\nabla \psi|^2 \, dx \ ,$$

so that (using also that $\psi \equiv 0$ on B_t)

$$\int_{B_t} |\nabla u - A|^2 \, dx \leq C \int_{B_s \setminus B_t} |\nabla u - A||\nabla \psi| + |\nabla \psi|^2 \, dx \ ,$$

with $C = C(\lambda, \Lambda)$.

Now, since $|\nabla \psi| \leq |\nabla u - A| + 2|u - Ax|/(s-t)$, we get

$$\int_{B_t} |\nabla u - A|^2 \, dx \leq D \int_{B_s \setminus B_t} |\nabla u - A|^2 \, dx + \frac{D}{(s-t)^2} \int_{B_r} |u - Ax|^2 \, dx$$

for some new constant $D = D(\lambda, \Lambda)$. Now we apply the hole-filling technique to get

$$\int_{B_t} |\nabla u - A|^2 \, dx \leq \theta \int_{B_s} |\nabla u - A|^2 \, dx + \frac{D}{(s-t)^2} \int_{B_r} |u - Ax|^2 \, dx \ .$$

with $\theta = D/(D+1) < 1$. At this point, since the inequality is true for all $r/2 \leq t \leq s \leq r$, a standard iteration scheme gives the result. Indeed, let $\tau \in (0,1)$ with $\theta < \tau^2$ and define $t_i = (1 - \tau^i/2)r$, so that $t_0 = r/2$, $t_i \uparrow r$ and $t_{i+1} - t_i = r(1-\tau)\tau^i/2$. By iteration of the inequality

$$\int_{B_{t_i}} |\nabla u - A|^2 \, dx \leq \theta \int_{B_{t_{i+1}}} |\nabla u - A|^2 \, dx + \frac{4D}{r^2(1-\tau)^2} \tau^{-2i}$$

we get

$$\begin{aligned} \int_{B_{t_0}} |\nabla u - A|^2 \, dx &\leq \theta^N \int_{B_{t_N}} |\nabla u - A|^2 \, dx + \frac{4D}{r^2(1-\tau)^2} \sum_{i=0}^{N-1} (\theta/\tau^2)^i \\ &\leq \theta^N \int_{B_r} |\nabla u - A|^2 \, dx + \frac{4D\tau^2}{r^2(1-\tau)^2(\tau^2 - \theta)} \end{aligned}$$

for any integer $N \geq 1$. As $N \to \infty$ we get the result. □

18.1 Partial regularity for systems: $\mathscr{L}^n(\Sigma(u)) = 0$

After proving Theorem 18.8 about the decay of the excess, we will see how it can be used to prove partial regularity for systems.

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 is to show that for a solution of an elliptic system the following facts:

- $\mathscr{L}^n\big(\Sigma(u)\big) = 0$;
- $\mathscr{H}^{n-2+\varepsilon}\big(\Sigma(u)\big) = 0$ for all $\varepsilon > 0$ in the uniformly convex case and $\mathscr{H}^{n-2}(\Sigma(u)) = 0$ if $\nabla^2 F$ is also uniformly continuous.

In order to exploit Theorem 18.8 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 $C_e\alpha < 1/2$ (recall that C_e 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 with halving of the excess from the scale r to the scale αr.

Definition 18.10. *We will call*

$$\Omega_M(u) := \left\{x \in \Omega \,\middle|\, \exists B_r(x) \Subset \Omega \text{ with } \left|(\nabla u)_{B_r(x)}\right| < M_1 \text{ and } \mathrm{Exc}\,(u, B_r(x)) < \varepsilon_1 \right\}$$

where

$$M_1 := M/2 \tag{18.14}$$

and ε_1 verifies

$$2^{n/2}\varepsilon_1 \leq \varepsilon_0 \tag{18.15}$$

and for $\alpha \in (0, 1/4)$ fixed, chosen in such a way that $C_e\alpha < 1/2$,

$$(2^{n+1} + \alpha^{-n}2^{1+n/2})\varepsilon_1 \leq M\ . \tag{18.16}$$

Remark 18.11. The set $\Omega_M(u) \subset \Omega$ of Definition 18.10 is open, since the inequalities are strict. Moreover, by Lebesgue approximate continuity theorem (that is, if $f \in L^p(\Omega)$, then for $\mathscr{L}^n$-almost every x one has $⨍_{B_r(x)} |f(y) - f(x)|^p\, dy \to 0$ as $r \downarrow 0$), it is easy to see that

$$\mathscr{L}^n\left(\{|\nabla u| < M_1\} \setminus \Omega_M(u)\right) = 0\ . \tag{18.17}$$

Finally, using (18.17), we realize that

$$\mathscr{L}^n\left(\Omega \setminus \bigcup_{M\in\mathbb{N}} \Omega_M(u)\right) = \mathscr{L}^n\left(\Omega \setminus \bigcup_{M\in\mathbb{N}} \{|\nabla u| < M_1\}\right) = 0\ . \tag{18.18}$$

By the previous remark, if we are able to prove that

$$\Omega_M(u) \subset \Omega_{\mathrm{reg}} \qquad \forall M > 0, \tag{18.19}$$

we obtain $\mathscr{L}^n(\Sigma(u)) = 0$. So, the rest of this section will be devoted to the proof of the inclusion above, with M fixed.

Fix $x \in \Omega_M(u)$, according to Definition 18.10 there exists $r > 0$ such that $B_r(x) \Subset \Omega$, $|(\nabla u)_{B_r(x)}| < M_1$ and $\mathrm{Exc}\,(u, B_r(x)) < \varepsilon_1$. We will prove that

$$B_{r/2}(x) \subset \Omega_{\mathrm{reg}}(u)\ ,$$

so let us fix $y \in B_{r/2}(x)$.
(1) Thanks to our choice of ε_1 (see property (18.15) of Definition 18.10) we have

$$
\begin{aligned}
\operatorname{Exc}\left(u, B_{r/2}(y)\right) &= \left(\fint_{B_{r/2}(y)} |\nabla u(z) - (\nabla u)_{B_{r/2}(y)}|^2 \, dz\right)^{1/2} \\
&\leq \left(\fint_{B_{r/2}(y)} |\nabla u(z) - (\nabla u)_{B_r(x)}|^2 \, dz\right)^{1/2} \\
&\leq 2^{n/2} \left(\fint_{B_r(x)} |\nabla u(z) - (\nabla u)_{B_r(x)}|^2 \, dz\right)^{1/2} = 2^{n/2} \operatorname{Exc}\left(u, B_r(x)\right) < \varepsilon_0
\end{aligned}
$$

so, momentarily ignoring the hypothesis that $|(\nabla u)_{B_{r/2}(y)}|$ should be bounded by M (we are postponing this to point (2) of this proof), Theorem 18.8 gives *tout court*

$$
\operatorname{Exc}\left(u, B_{\alpha r/2}(y)\right) \leq \frac{1}{2} \operatorname{Exc}\left(u, B_{r/2}(y)\right) < \frac{1}{2} \varepsilon_0 \ ,
$$

thus, just iterating Theorem 18.8, we get

$$
\operatorname{Exc}\left(u, B_{\alpha^k r/2}(y)\right) \leq 2^{-k} \operatorname{Exc}\left(u, B_{r/2}(y)\right) \ . \tag{18.20}
$$

As we have often seen through these notes, we can apply an interpolation argument to a sequence of radii with ratio α to obtain

$$
\operatorname{Exc}\left(u, B_\rho(y)\right) \leq \alpha^\mu \left(\frac{\rho}{r/2}\right)^\mu \operatorname{Exc}\left(u, B_{r/2}(y)\right) \leq \alpha^\mu \left(\frac{\rho}{r/2}\right)^\mu \varepsilon_0 \qquad \forall \rho \in (0, r/2], \ y \in B_{r/2}(x)
$$

with $\mu = (\log_2(1/\alpha))^{-1}$. We conclude that the components of ∇u belong to the Campanato space $\mathcal{L}^{2,n+2\mu}(B_{r/2}(x))$ and then u belongs to $C^{1,\mu}(B_{r/2}(x))$.
(2) Now that we have explained how the proof runs through the iterative application of Theorem 18.8, we deal with the initially neglected hypothesis, that is $|(\nabla u)_{B_{r/2}(y)}| < M$ and, at each subsequent step, $|(\nabla u)_{B_{\alpha^k r/2}(y)}| < M$. Remember that in part (1) of this proof we never used (18.14) and (18.16).
Since $x \in \Omega_M(u)$ and r fulfills Definition 18.10, for the first step it is sufficient to use the

triangular inequality in (18.21) and Hölder's inequality in (18.22): in fact we can estimate

$$\begin{aligned}
\left|(\nabla u)_{B_{r/2}(y)}\right| &= \left|⨍_{B_{r/2}(y)} \left(\nabla u(z) - (\nabla u)_{B_r(x)}\right)\, dz + (\nabla u)_{B_r(x)}\right| \\
&\leq ⨍_{B_{r/2}(y)} \left|\nabla u(z) - (\nabla u)_{B_r(x)}\right| dz + \left|(\nabla u)_{B_r(x)}\right| \qquad (18.21) \\
&\leq \left(\frac{2^n}{\omega_n r^n} \int_{B_r(x)} \left|\nabla u(z) - (\nabla u)_{B_r(x)}\right| dz\right) + \left|(\nabla u)_{B_r(x)}\right| \\
&\leq 2^n \left(⨍_{B_r(x)} \left|\nabla u(z) - (\nabla u)_{B_r(x)}\right|^2 dz\right)^{1/2} + \left|(\nabla u)_{B_r(x)}\right| \qquad (18.22) \\
&\leq 2^n \mathrm{Exc}\,(u, B_r(x)) + \left|(\nabla u)_{B_r(x)}\right| < 2^n\varepsilon_1 + M_1 < M\ . \qquad (18.23)
\end{aligned}$$

We now show inductively that for every integer $k \geq 1$

$$\left|(\nabla u)_{B_{\alpha^k r/2}(y)}\right| \leq M_1 + 2^n\varepsilon_1 + \alpha^{-n}\varepsilon_1 2^{n/2} \sum_{j=0}^{k-1} 2^{-j}\ . \qquad (18.24)$$

If we recall (18.14) and (18.16), it is clear that (18.24) implies

$$\left|(\nabla u)_{B_{\alpha^k r/2}(y)}\right| \leq M$$

for every $k \geq 1$.
The first step ($k = 1$) follows from (18.23), because, estimating as in (18.21) and (18.22), we immediately get

$$\begin{aligned}
\left|(\nabla u)_{B_{\alpha r/2}(y)}\right| &\leq ⨍_{B_{\alpha r/2}(y)} \left|\nabla u(z) - (\nabla u)_{B_{r/2}(y)}\right| dz + \left|(\nabla u)_{B_{r/2}(y)}\right| \\
&\leq \alpha^{-n}\mathrm{Exc}\left(u, B_{r/2}(y)\right) + \left|(\nabla u)_{B_{r/2}(y)}\right| \\
&\leq \alpha^{-n}2^{n/2}\varepsilon_1 + 2^n\varepsilon_1 + M_1\ .
\end{aligned}$$

Being the first step already proved, we fix our attention on the $(k+1)^{\text{th}}$ step. With the same procedure, we estimate again

$$\begin{aligned}
\left|(\nabla u)_{B_{\alpha^{k+1} r/2}(y)}\right| &\leq ⨍_{B_{\alpha^{k+1} r/2}(y)} \left|\nabla u(z) - (\nabla u)_{B_{\alpha^k r/2}(y)}\right| dz + \left|(\nabla u)_{B_{\alpha^k r/2}(y)}\right| \\
&\leq \alpha^{-n}\mathrm{Exc}\left(u, B_{\alpha^k r/2}(y)\right) + \left|(\nabla u)_{B_{\alpha^k r/2}(y)}\right| \\
&\leq \alpha^{-n}2^{n/2-k}\varepsilon_1 + M_1 + 2^n\varepsilon_1 + \alpha^{-n}\varepsilon_1 2^{n/2} \sum_{j=0}^{k-1} 2^{-j} \qquad (18.25)
\end{aligned}$$

where (18.25) is obtained joining the estimate on the excess (18.20) with the inductive hypothesis (18.24).

In order to carry out our second goal, namely to prove that

$$\mathscr{H}^{n-2+\varepsilon}\big(\Sigma(u)\big) = 0 \qquad \forall\, \varepsilon > 0 \ ,$$

we need some basic results concerning Hausdorff measures.

18.2 Hausdorff measures

Definition 18.12. *Consider a subset $B \subset \mathbb{R}^n$, $k \geq 0$ and fix $\delta \in (0, \infty]$. The so-called pre-Hausdorff measures $\mathscr{H}^k_\delta$ are defined by*

$$\mathscr{H}^k_\delta(B) := c_k \inf \left\{ \sum_{i=1}^{\infty} [\mathrm{diam}(B_i)]^k \middle| B \subset \bigcup_{i=1}^{\infty} B_i,\ \mathrm{diam}(B_i) < \delta \right\} ,$$

while $\mathscr{H}^k$ is defined by

$$\mathscr{H}^k(B) := \lim_{\delta \to 0} \mathscr{H}^k_\delta(B) \ , \tag{18.26}$$

the limit in (18.26) *being well defined because $\delta \mapsto \mathscr{H}^k_\delta(B)$ is non-increasing. The constant $c_k \in (0, \infty)$ will be conveniently fixed in Remark 18.14.*

It is easy to check that $\mathscr{H}^k$ is the counting measure when $k = 0$ (provided $c_0 = 1$) and $\mathscr{H}^k$ is identically 0 when $k > n$.

The *spherical* Hausdorff measure $\mathscr{S}^k$ has a definition analogous to Definition 18.12, but only covers made with balls are allowed, so that

$$\mathscr{H}^k_\delta \leq \mathscr{S}^k_\delta \leq 2^k \mathscr{H}^k_\delta, \qquad \mathscr{H}^k \leq \mathscr{S}^k \leq 2^k \mathscr{H}^k \ . \tag{18.27}$$

Remark 18.13. Simple but useful properties of Hausdorff measures are:

(i) The Hausdorff measures are translation invariant, that is

$$\mathscr{H}^k(B + h) = \mathscr{H}^k(B) \quad \forall\, B \subset \mathbb{R}^n,\ \forall\, h \in \mathbb{R}^n \ ,$$

and (positively) k-homogeneous, that is

$$\mathscr{H}^k(\lambda B) = \lambda^k \mathscr{H}^k(B) \quad \forall\, B \subset \mathbb{R}^n,\ \forall \lambda > 0 \ .$$

(ii) The Hausdorff measures are countably subadditive, which means that whenever we have a countable cover of a subset B, namely $B \subset \cup_{i \in I} B_i$, then

$$\mathscr{H}^k(B) \leq \sum_{i \in I} \mathscr{H}^k(B_i) \ .$$

(iii) For every set $A \subset \mathbb{R}^n$ the map $B \mapsto \mathscr{H}^k(A \cap B)$ is σ-additive on Borel sets, which means that whenever we have a countable pairwise disjoint cover of a Borel set B by Borel sets B_i, we have

$$\mathscr{H}^k(A \cap B) = \sum_{i \in I} \mathscr{H}^k(A \cap B_i) \ .$$

(iv) Having fixed the subset $B \subset \mathbb{R}^n$ and $\delta > 0$, we have that

$$k > k' \quad \Longrightarrow \quad \mathscr{H}^k_\delta(B) \leq \delta^{k-k'} \mathscr{H}^{k'}_\delta(B) \ . \tag{18.28}$$

In particular, looking at (18.28) when $\delta \to 0$, we deduce that

$$\mathscr{H}^{k'}(B) < +\infty \quad \Longrightarrow \quad \mathscr{H}^k(B) = 0$$

or, equivalently,

$$\mathscr{H}^k(B) > 0 \quad \Longrightarrow \quad \mathscr{H}^{k'}(B) = +\infty \ .$$

Remark 18.14. When k is an integer, the choice of c_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 that

$$\mathscr{L}^k_\pi(B) = \mathscr{H}^k(B) \ , \tag{18.29}$$

where $\mathscr{L}^k_\pi$ is the k-dimensional Lebesgue measure on $\pi \sim \mathbb{R}^k$. It is useful to remember the isodiametric inequality among all sets with prescribed diameter, balls have the largest volume: more precisely, if $\omega_k := \mathscr{L}^k(B_1(0))$, for every Borel subset $B \subset \mathbb{R}^k$ there holds

$$\mathscr{L}^k(B) \leq \omega_k \left(\frac{\operatorname{diam}(B)}{2} \right)^k \ . \tag{18.30}$$

Thanks to (18.30), it can be easily proved that equality (18.29) holds if we choose

$$c_k = \frac{\omega_k}{2^k} \ .$$

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 local parametrizations and partitions of unity.

Proposition 18.15. *Consider a locally finite measure $\mu \geq 0$ on the family of Borel sets $\mathcal{B}(\mathbb{R}^n)$ and, fixing $t > 0$, set*

$$B := \left\{ x \mid \limsup_{r \to 0} \frac{\mu(\overline{B}_r(x))}{\omega_k r^k} > t \right\} , \tag{18.31}$$

then B is a Borel set and

$$\mu(B) \geq t\mathscr{S}^k(B) .$$

Moreover, if μ vanishes on $\mathscr{H}^k$-finite sets, then $\mathscr{H}^k(B) = 0$.

A traditional proof of Proposition 18.15 is based on Besicovitch covering theorem, whose statement is included below for the sake of completeness. We present instead a proof based on a more general and robust covering theorem, valid in general metric spaces.

Theorem 18.16 (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 $A_1, \ldots, A_{\xi(n)} \subset A$ such that*

(a) for all $j = 1, \ldots, \xi$, the balls in $\{B_{\rho(x)}(x)\}_{x \in A_j}$ are pairwise disjoint;

(b) the ξ families still cover the set A, that is

$$A \subset \bigcup_{j=1}^{\xi} \left(\bigcup_{x \in A_j} B_{\rho(x)}(x) \right) .$$

Let us introduce now the general covering theorem.

Definition 18.17 (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 18.18. *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 $\mathcal{F}' = \{\overline{B}_i\}_{i \geq 1} \subset \mathcal{F}$ such that at least one of the following conditions holds:*

(i) $\sum_{i=1}^{\infty} [\mathrm{r}(B_i)]^k = \infty$,

(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 $\mathcal{F}_0 := \mathcal{F}$. Surely, there exists a closed ball, let us call it $\overline{B}_1$, such that

$$\mathrm{r}\left(\overline{B}_1\right) > \frac{1}{2}\sup\left\{\mathrm{r}(\overline{B})|\,\overline{B} \in \mathcal{F}_0\right\} \ .$$

Now put

$$\mathcal{F}_1 := \{\overline{B} \in \mathcal{F}_0|\,\overline{B} \cap \overline{B}_1 = \emptyset\} \ ,$$

and choose among them a ball $\overline{B}_2 \in \mathcal{F}_1$ such that

$$\mathrm{r}(\overline{B}_2) > \frac{1}{2}\sup\left\{\mathrm{r}(\overline{B})|\,B \in \mathcal{F}_1\right\} \ .$$

If we try to go on analogously, the only chance by which the construction has to stop is that for some $l \in \mathbb{N}$ the family $\mathcal{F}_l = \emptyset$, so we are 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 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, and in particular $\mathrm{diam}(\overline{B}_i) \to 0$, then we have to find (ii) again.

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

$$\overline{B}_{r(x)}(x) \cap \overline{B}_{i(x)} \neq \emptyset \ . \tag{18.32}$$

In fact if

$$\overline{B}_{r(x)}(x) \cap \overline{B}_i = \emptyset \qquad \forall\, i > i_0 \ , \tag{18.33}$$

then

$$r_i \geq \frac{r(x)}{2} \qquad \forall\forall\, i > i_0 \tag{18.34}$$

but $r_i \to 0$, so (18.34) leads to a contradiction. Without loss of generality, we can think that $i(x)$ is the first index larger than i_0 for which (18.32) holds, too. Since, by construction, $r_{i(x)} > \frac{1}{2}\sup\{\mathrm{r}(\overline{B})|\overline{B} \in \mathcal{F}_{i(x)-1}\}$ (and $\overline{B}_{r(x)}(x) \in \mathcal{F}_{i(x)-1}$ by the minimality of $i(x)$), then $r(x) \leq 2r_{i(x)}$.

Since the balls intersect, 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

$$A \setminus \bigcup_{i=1}^{i_0} \overline{B}_i \subset \bigcup_{i=i_0+1}^{\infty} \overline{B}_{5r_i}(y_i)\ . \tag{18.35}$$

Choosing i_0 such that $10r_i < \delta$ for every $i > i_0$, (18.35) says that

$$\mathscr{H}^k_\delta\left(A \setminus \bigcup_{i=1}^{\infty} \overline{B}_i\right) \le \mathscr{H}^k_\delta\left(A \setminus \bigcup_{i=1}^{i_0} \overline{B}_i\right) \le \sum_{i=i_0+1}^{\infty} \omega_k (10r_i)^k\ .$$

We conclude remarking that when $\delta \to 0$, $i_0 \to +\infty$ and

$$\mathscr{H}^k\left(A \setminus \bigcup_{i=1}^{\infty} \overline{B}_i\right) \le \lim_{i_0\to\infty} \omega_k \sum_{i=i_0+1}^{\infty} (10r_i)^k = 0\ .$$

□

Now we are able to prove Proposition 18.15.

Proof. Intersecting B with balls, one easily reduces to the case of a bounded set B. Hence, we can assume B bounded and μ finite measure. Fix $\delta > 0$, an open set $A \supset B$ and consider the family

$$\mathcal{F} := \left\{\overline{B}_r(x) \,\middle|\, r < \delta/2,\, \overline{B}_r(x) \subset A,\, \mu\left(B_r(x)\right) > t\omega_k r^k\right\}\ , \tag{18.36}$$

that is a fine cover of B. Applying Theorem 18.18, we get a subfamily $\mathcal{F}' \subset \mathcal{F}$ whose elements we will denote by

$$\overline{B}_i = \overline{B}_{r_i}(x_i)\ .$$

First we exclude possibility (i) of Theorem 18.18: as a matter of fact

$$\sum_{i=1}^{\infty} r_i^k < \frac{1}{t\omega_k}\sum_{i=1}^{\infty} \mu(\overline{B}_i) \le \frac{\mu(A)}{t\omega_k} < \infty\ .$$

Since (ii) holds and we can compare $\mathscr{H}^k_\delta$ with $\mathscr{S}^k_\delta$ via (18.27), to get

$$\mathscr{S}^k_\delta(B) \le \mathscr{S}^k_\delta\left(\bigcup_{i=1}^{\infty} \overline{B}_i\right) \le \sum_{i=1}^{\infty} \omega_k r_i^k < \frac{1}{t}\sum_{i=1}^{\infty} \mu(\overline{B}_i) \le \frac{\mu(A)}{t}\ , \tag{18.37}$$

As $\delta \downarrow 0$ we get $t\mathscr{S}^k(B) \le \mu(A)$ and the outer regularity of μ gives $t\mathscr{S}^k(B) \le \mu(B)$.

Finally, the last statement of the proposition can be achieved noticing that the inequality (18.37) gives that $\mathscr{S}^k(B)$ is finite; if we assume that μ vanishes on sets with finite k-dimensional measure we obtain that $\mu(B) = 0$; applying once more the inequality we get $\mathscr{S}^k(B) = 0$. □

18.3 Partial regularity for systems: $\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0$

Aware of the usefulness of Proposition 18.15 for our purposes, we are now ready to obtain that if $F \in C^2(\mathbb{R}^{m\times n})$ satisfies the Legendre condition for some $\lambda > 0$ and satisfies also

$$|\nabla^2 F(p)| \leq \Lambda < \infty \qquad \forall\, p \in \mathbb{R}^{m\times n}$$

then we have a stronger upper bound on the size of the singular set, namely

$$\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0 \qquad \forall\, \varepsilon > 0\ , \tag{18.38}$$

where, as usual, $\Sigma(u) := \Omega \setminus \Omega_{\mathrm{reg}}(u)$.

Let us remark that, with respect to the first partial regularity result and with respect to Evans Theorem 18.3, we 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 $\Delta_{h,s}(\nabla u)$ satisfies an equielliptic family of systems, then, via Caccioppoli inequality the sequence $\Delta_{h,s}(\nabla u)$ 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 (18.38) for systems in which $\nabla^2 F$ is uniformly continuous, we will see it in Corollary 18.21.

As for the strategy: in Proposition 18.19 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 18.20 and Theorem 18.23, respectively.

Proposition 18.19. *Consider, as previously, a variational problem defined by $F \in C^2(\mathbb{R}^{m\times n})$ with $|\nabla^2 F| \leq \Lambda$, satisfying the Legendre condition for some $\lambda > 0$. If u is a local minimizer of such a problem, define the sets*

$$\Sigma_1(u) := \left\{ x \in \Omega \,\middle|\, \limsup_{r\to 0} r^{2-n} \int_{B_r(x)} |\nabla^2 u(y)|^2\, dy > 0 \right\}$$

and

$$\Sigma_2(u) := \left\{ x \in \Omega \,\middle|\, \limsup_{r\to 0} \left|(\nabla u)_{B_r(x)}\right| = +\infty \right\}\ .$$

Then $\Sigma(u) \subset \Sigma_1(u) \cup \Sigma_2(u)$. If in addition $\nabla^2 F$ 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 $\left|(\nabla u)_{B_r(x)}\right| < M_1$ for arbitrarily small radii $r > 0$;

- thanks to Poincaré inequality

$$\operatorname{Exc}(u, B_r(x))^2 \le C(n) r^{2-n} \int_{B_r(x)} |\nabla^2 u(y)|^2 \, dy \longrightarrow 0 \ ;$$

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 18.10, and $\Omega_M(u) \subset \Omega_{\text{reg}}$ due to (18.19).

The second part of the statement can be achieved noticing that, in the case when $\nabla^2 F$ is uniformly continuous, no bound on $|(\nabla u)_{B_r(x)}|$ is needed in the decay theorem and in the characterization of the regular set. $\square$

Proposition 18.20. *If $u \in W^{2,2}_{\text{loc}}(\Omega)$, we have that*

$$\mathscr{H}^{n-2}(\Sigma_1(u)) = 0 \ .$$

Proof. Let us employ Proposition 18.15 with the absolutely continuous measure $\mu := |\nabla^2 u|^2 \mathscr{L}^n$. Obviously we choose $k = (n-2)$ and we have that μ vanishes on sets with finite $\mathscr{H}^{n-2}$-measure. The thesis follows when we observe that

$$\Sigma_1(u) = \bigcup_{\nu=1}^{\infty} \left\{ x \in \Omega \,\middle|\, \limsup_{r \to 0} \frac{\mu(\overline{B}_r(x))}{\omega_{n-2} r^{n-2}} > \frac{1}{\nu} \right\} \ .$$

$\square$

By the second part of the statement of Proposition 18.19 we get:

Corollary 18.21. *If we add the uniform continuity of $D^2 F$ to the hypotheses of Proposition 18.20, we can conclude that*

$$\mathscr{H}^{n-2}(\Sigma(u)) = 0 \ . \tag{18.39}$$

The estimate on the Hausdorff measure of $\Sigma_2(u)$ is a bit more complex and passes through the estimate of the Hausdorff measure of the so-called *approximate discontinuity set* S_v of a function v.

Definition 18.22. *Given a function $v \in L^1_{\text{loc}}(\Omega)$, we put*

$$\Omega \setminus S_v := \left\{ x \in \Omega \,\middle|\, \exists\, z \in \mathbb{R} \ s.t. \ \lim_{r \downarrow 0} \fint_{B_r(x)} |v(y) - z| \, dy = 0 \right\} \ .$$

When such a z exists, it is unique and we will call it approximate limit *of v at the point x.*

Theorem 18.23. *If $v \in W^{1,p}(\Omega)$, $1 \le p \le n$, then*

$$\mathscr{H}^{n-p+\varepsilon}(S_v) = 0 \qquad \forall \varepsilon > 0 \ .$$

Notice that the statement is trivial in the case $p > n$, by the Sobolev Embedding Theorem (i.e. $S_v = \emptyset$): as p increases the Hausdorff dimension of the approximate discontinuity set moves from $n-1$ to 0.

Applying this theorem to $v = \nabla u \in H^{1,2}(\Omega; \mathbb{R}^{m\times n})$, $p = 2$, we get that $\mathscr{H}^{n-2+\varepsilon}(\Sigma_2(u)) = 0$.

Proof. (1) Fix $0 < \eta < \rho$, we claim that

$$n\omega_n \big|(v)_{B_\eta(x)} - (v)_{B_\rho(x)}\big| \le (n-1)\int_0^\rho t^{-n}\int_{B_t(x)} |\nabla v(y)|\, dy\, dt + \rho^{-(n-1)}\int_{B_\rho(x)} |\nabla v(y)|\, dy \ ; \tag{18.40}$$

we will show this in the part (3) of this proof.

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 sequence $(v)_{B_r(x)}$ admits a limit z as $r \to 0$ because it is a Cauchy sequence. Thanks to the Poincaré inequality

$$\fint_{B_r(x)} |v(y) - (v)_{B_r(x)}|\, dy \le C(n) r^{-(n-1)} \int_{B_r(x)} |\nabla v(y)|\, dy \stackrel{r\to 0}{\longrightarrow} 0 \ ,$$

therefore

$$\fint_{B_r(x)} |v(y) - z|\, dy \stackrel{r\to 0}{\longrightarrow} 0 \ ,$$

that is to say, $x \notin S_v$. This chain of implications means that, for all $\varepsilon > 0$,

$$\Omega \setminus S_v \supset \left\{ x \in \Omega \,\Big|\, \int_{B_t(x)} |\nabla v(y)|\, dy = o(t^{n-1+\varepsilon}) \right\} \ . \tag{18.41}$$

(2) In order to refine (18.41) 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,

$$\int_{B_t(x)} |\nabla v(y)|\, dy \le o(t^{n/p-1+\varepsilon/p}) t^{n/p'} = o(t^{n-1+\varepsilon/p}) \ .$$

For this reason we can deduce from (18.41) the inclusion

$$\Omega \setminus S_v \supset \left\{ x \in \Omega \,\Big|\, \int_{B_t(x)} |\nabla v(y)|^p\, dy = o(t^{n-p+\varepsilon}) \right\} \qquad \forall \varepsilon > 0 \ . \tag{18.42}$$

In view of Proposition 18.15, the complement of the set $\{x \in \Omega \mid \int_{B_t(x)} |\nabla v(y)|^p\, dy = o(t^{n-p+\varepsilon})\}$ is $\mathscr{H}^{n-p+\varepsilon}$-negligible, hence the jump set S_v is $\mathscr{H}^{n-p+\varepsilon}$-negligible, too.

(3) This third part is devoted to the proof of (18.40); for the sake of simplicity we put $x = 0$. Let us consider the characteristic function χ_{B_1}; since we would like to differentiate the map

$$\rho \mapsto \rho^{-n} \int \chi\left(\frac{y}{\rho}\right) v(y)\, dy\ ,$$

a possible proof of (18.40) is based on a regularization of χ, differentiation and passage to the limit.

We produce instead a direct proof based on a *ad hoc* calibration: we need a vector field ϕ with $\operatorname{supp}\phi \subset \overline{B}_\rho$ whose divergence almost coincides with the operator acting on v in left member of (18.40), that is

$$\mathrm{div}\phi = n\left(\eta^{-n}\chi_{B_\eta} - \rho^{-n}\chi_{B_\rho}\right)\ . \tag{18.43}$$

Therefore,

$$\phi(x) := x\left(\left(\eta^{-n} \wedge |x|^{-n}\right) - \left(\rho^{-n} \wedge |x|^{-n}\right)\right)$$

verifies (18.43) and, with the notation $\mu = |\nabla v|\chi_{B_\rho}\mathscr{L}^n$, there holds

$$\frac{n}{\eta^n}\int_{B_\eta} v(y)\, dy - \frac{n}{\rho^n}\int_{B_\rho} v(y)\, dy = \int v(y)\mathrm{div}\phi(y)\, dy \tag{18.44}$$

$$= -\int \phi(y)\cdot\nabla v(y)\, dy \le \int_{B_\rho} |\phi(y)||\nabla v(y)|\, dy \le \int_{\mathbb{R}^n} |y|^{-(n-1)}\, d\mu(y) \tag{18.45}$$

$$= \int_0^\infty \mu\left(|y|^{-(n-1)} > t\right)\, dt = (n-1)\int_0^\infty s^{-n}\mu(B_s)\, ds \tag{18.46}$$

$$= (n-1)\int_0^\rho s^{-n}\int_{B_s} |\nabla v(y)|\, dy\, ds + (n-1)\int_\rho^\infty s^{-n}\int_{B_\rho} |\nabla v(y)|\, dy\, ds$$

$$= (n-1)\int_0^\rho s^{-n}\int_{B_s} |\nabla v(y)|\, dy\, ds + \rho^{-(n-1)}\int_{B_\rho} |\nabla v(y)|\, dy\ ,$$

where we pass from (18.44) to (18.45) by the divergence theorem, from (18.45) to (18.46) by Cavalieri's principle and then it is all change of variables and Fubini's theorem. □

Remark 18.24. 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 measurement of the discontinuity set with the scale of Hausdorff measures is sharp. On the contrary, in the case $p > 1$ the right scale for the measurement of the approximate discontinuity set are the so-called capacities.

19 Some tools from convex and nonsmooth analysis

19.1 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.

In the sequel we 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 $u \in C^2(\Omega)$ this is equivalent to say that $\nabla^2 u(x) \ge 0$, in the sense of symmetric operators, for all $x \in \Omega$.

Definition 19.1 (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 19.2. According to Definition 19.1, it is easy to show that

$$\partial u(x) = \{p \in \mathbb{R}^n | \liminf_{t \to 0^+} \frac{u(x+tv) - u(x)}{t} \ge \langle p, v\rangle \quad \forall v \in \mathbb{R}^n\} \ . \tag{19.1}$$

Indeed, when $p \in \partial u(x)$ the relation

$$\frac{u(x+tv) - u(x)}{t} \ge \langle p, v\rangle$$

passes through the limit. Conversely, let us recall the monotonicity property of difference quotients of a convex function, i.e.

$$\frac{u(x+t'v) - u(x)}{t'} \le \frac{(1 - t'/t)\, u(x) + (t'/t)u(x+tv) - u(x)}{t'} = \frac{u(x+tv) - u(x)}{t} \ , \tag{19.2}$$

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 (19.2) yields that the $\liminf$ in (19.1) is a limit.

Remark 19.3. The following properties are easy to check:

(i) The graph of the subdifferential, i.e. $\{(x,p)|p \in \partial u(x)\} \subset \Omega \times \mathbb{R}^n$, is closed, in fact convex functions are continuous (suffices, by (ii) below, to show that they are locally bounded to obtain even local Lipschitz continuity).

(ii) 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 in (19.2), 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 starting 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 $B_R(x_0)$, or even $B_r(x_0)$ by $\overline{B}_r(x_0)$ in the inequality above. Equivalently

$$\operatorname*{ess\,sup}_{B_r(x_0)} |\nabla u| \leq \frac{\operatorname{osc}(u, B_R(x_0))}{R-r} ,$$

because of (1.6).

(iii) As a consequence of (ii) and Rademacher's Theorem, $\partial u(x) \neq \emptyset$ for all $x \in \Omega$. In addition, a convex function u belongs to C^1 if and only if $\partial u(x)$ is a singleton for every $x \in \Omega$. Indeed, if $\{x_h\}$ are differentiability points of u such that $x_h \to x$ and $\nabla u(x_h)$ has at least two distinct limit points, then $\partial u(x)$ is not a singleton. Hence ∇u has a continuous extension to the whole of Ω and $u \in C^1$.

(iv) 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

$$p \in \partial f(x) .$$

In fact, 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 19.4 (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 $\overline{B}_1$ and define the sequence of functions $f_\varepsilon := f * \rho_\varepsilon$, which are easily seen to be convex in the set Ω_ε in (1.3), because

$$\begin{aligned} f_\varepsilon((1-t)x + ty) &= \int_\Omega f((1-t)x + ty - \varepsilon\xi)\rho(\xi)\, d\xi \\ &\le \int_\Omega ((1-t)f(x - \varepsilon\xi) + tf(y - \varepsilon\xi))\,\rho(\xi)\, d\xi \\ &= (1-t)f_\varepsilon(x) + tf_\varepsilon(y) \ ; \end{aligned}$$

moreover $f_\varepsilon \to f$ locally uniformly. 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 $\varepsilon \to 0$, we can find $\varepsilon_k \to 0$ with $\theta_{\varepsilon_k} \to \theta \in [0, 1]$ and $p_{\varepsilon_k} \to p$. Remark 19.3(iv) allows us to conclude that $p \in \partial f((1 - \theta)x + \theta y)$ and

$$f(x) - f(y) = \langle p, x - y\rangle \ .$$

□

As an application of the nonsmooth mean value theorem, we can derive a pointwise version of Remark 19.3(iii). Notice that we will follow a similar idea to achieve second order differentiability.

Proposition 19.5. *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 closure 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, the nonsmooth mean value theorem gives

$$f(y) - f(x) = \langle p_{xy}, y - x\rangle = \langle p, y - x\rangle + \langle p_{xy}, x - y\rangle = \langle p, y - x\rangle + o(|y - x|) \ .$$

□

Remark 19.6. Recall that a continuous function $f : \Omega \to \mathbb{R}$ is convex if and only if its Hessian $\nabla^2 f$ is non-negative, i.e. for every non-negative $\varphi \in C_c^\infty(\Omega)$ 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 obtained by approximation by convolution, because, still in the weak sense,

$$\nabla^2 (f * \rho_\varepsilon) = \big(\nabla^2 f\big) * \rho_\varepsilon.$$

Although we shall not need this fact in the sequel, except in Remark 19.17, let us mention, for completeness, that the positivity condition on the weak derivative $\nabla^2 f$ implies that this derivative is representable by a symmetric matrix-valued measure. To see this, it suffices to apply the following result to the second derivatives $\nabla^2_{\xi\xi} f$:

Lemma 19.7. *Consider a positive distribution $T \in \mathscr{D}'(\Omega)$, i.e.*

$$\forall\, \varphi \in C_c^\infty(\Omega), \varphi \geq 0 \quad \Longrightarrow \quad \langle T, \varphi \rangle \geq 0 \ .$$

Then there exists a locally finite non-negative measure μ in Ω such that

$$\langle T, \psi \rangle = \int_\Omega \psi \, d\mu \quad \forall \psi \in C_c^\infty(\Omega) \ .$$

Proof. Fix an open set $\Omega' \Subset \Omega$, define $K := \overline{\Omega'}$ and choose a non-negative cut-off function $\varphi \in C_c^\infty(\Omega)$ with $\varphi|_K \equiv 1$. For every test function $\psi \in C_c^\infty(\Omega')$, since $(\|\psi\|_{L^\infty}\varphi - \psi) \geq 0$ and T is a positive distribution, we have

$$\langle T, \psi \rangle \leq \langle T, \|\psi\|_{L^\infty}\varphi \rangle = C(\Omega')\|\psi\|_{L^\infty} \ ,$$

where $C(\Omega') := \langle T, \varphi \rangle$. Replacing ψ by $-\psi$, the same estimate holds with $|\langle T, \psi \rangle|$ in the left hand side. By Riesz representation theorem we obtain the existence of μ. □

Definition 19.8 (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 $\varphi \in C_c^\infty(\Omega)$ and for every $\xi \in \mathbb{R}^n$ (in short $\nabla^2 f \geq \lambda I$). We say also that

- *f is* uniformly convex *if $\lambda > 0$;*
- *f is* semiconvex *if $\lambda \leq 0$.*

Notice that, with the notation of Definition 19.8, 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. An important class of semiconcave functions is given by squared distance functions:

Example 19.9. Given a closed set $E \subset \mathbb{R}^n$, the square of the distance from E is 2-concave. In fact,

$$\mathrm{dist}^2(x, E) - |x|^2 = \inf_{y \in E} (x-y)^2 - |x|^2 = \inf_{y \in E} |y|^2 - 2\langle x, y\rangle \ ; \tag{19.3}$$

since the functions $x \mapsto |y|^2 - 2\langle x, y\rangle$ are affine, their infimum over $y \in E$, that is (19.3), is concave.

Particularly in the duality theory of convex functions, it is useful to extend the concept and 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 (mean value theorem, local Lipschitz continuity). 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

$$\tilde{f}(x) := \inf \left\{ \liminf_{h\to\infty} f(x_h) : \ x_h \in \Omega, \ x_h \to x \right\} \ .$$

It provides a convex and lower semicontinuous extension of f, equal to $+\infty$ on $\mathbb{R}^n \setminus \overline{\Omega}$. For these reasons, in the sequel 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, just requiring that $f(x) - \lambda|x|^2/2$ is convex.

Proposition 19.10. *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 19.11 (Inverse of the subdifferential)**.** (i) If $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ is λ-convex, Proposition 19.10 proves that for every $p \in \partial f(x)$ and every $q \in \partial f(y)$, we have

$$\langle p - q, x - y\rangle \geq \lambda|x - y|^2 \ . \tag{19.4}$$

(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 (19.4), 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 $(\partial f)^{-1} : L \to \{x : \ \partial f(x) \neq \emptyset\}$ 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 (19.4) 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 $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Equivalently, f^* is the smallest function satisfying

$$\langle x, y \rangle \leq f(x) + f^*(y) \qquad \forall x,\, y \in \mathbb{R}^n . \tag{19.5}$$

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 $z = x$, so that:

$$x^* \in \partial f(x) \qquad \Longleftrightarrow \qquad f^*(x^*) = \langle x^*, x \rangle - f(x) . \tag{19.6}$$

Theorem 19.12. *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 hypograph 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 immediately $f^*(p) < \infty$, so that $(f^*)^*$ makes sense. Now, the variational characterization of the conjugate function based on (19.5) 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 $\ell(x) = \langle p, x\rangle + c$ (notice that ℓ^* is finite only at $x^* = p$ and $\ell^*(p) = -c$). Since convex lower semicontinuous functions are supremum of affine functions (again as an application of the Hahn-Banach theorem), these two facts yield $(f^*)^* \geq f$ on convex lower semicontinuous functions, completing the proof. □

A byproduct of the previous proof is that $(f^*)^* = f$ in the class of convex and lower semicontinuous functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, not identically equal to $+\infty$. This way (19.5) becomes completely symmetric and it is easily seen that (19.6) gives

$$x \in \partial f^*(x^*) \qquad \Longleftrightarrow \qquad x^* \in \partial f(x)\,. \tag{19.7}$$

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 $f^* \in C^{1,1}(\mathbb{R}^n)$.

19.2 Convex functions and Measure Theory

Now we recall some classical results in Measure Theory, in order to have the necessary tools to prove Alexandrov theorem 19.16 on differentiability of convex functions.

Thanks to the next classical result we can, with a slight abuse of notation, keep the same notation ∇f for the pointwise gradient and the weak derivative, at least for locally Lipschitz functions.

Theorem 19.13 (Rademacher)**.** *Any Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathscr{L}^n$-almost every point and the pointwise gradient ∇f coincides $\mathscr{L}^n$-a.e. with the distributional derivative ∇f.*

Proof. Fix a point x_0 which is a Lebesgue point of ∇f, i.e.

$$⨍_{B_r(x_0)} |\nabla f(y) - \nabla f(x_0)|\, dy \stackrel{r\to 0}{\longrightarrow} 0\ . \tag{19.8}$$

Defining

$$f_r(y) := \frac{1}{r}\left(f(x_0 + ry) - f(x_0)\right)$$

and noticing that $\nabla f_r(y) = \nabla f(x_0 + ry)$ (still in the distributional sense), we are able to rewrite (19.8) as

$$⨍_{B_1(0)} |\nabla f_r(y) - \nabla f(x_0)|\, dy \stackrel{r\to 0}{\longrightarrow} 0\ ,$$

where (f_r) is a sequence of functions with equibounded Lipschitz constant and $f_r(0) = 0$ for every $r > 0$. Thanks to the Ascoli-Arzelà theorem, as $r \downarrow 0$, this family of functions has limit points in the uniform topology. Any limit point g obviously satisfies $g(0) = 0$, and since ∇g is a limit point of ∇f_r in the weak* topology, the strong convergence of ∇f_r to $\nabla f(x_0)$ gives $\nabla g \equiv \nabla f(x_0)$, still in the weak sense. We conclude that $g(x) = \nabla f(x_0)x$, so that g is uniquely determined and

$$f_r(y) = \frac{1}{r}\left(f(x_0 + ry) - f(x_0)\right) \stackrel{r\to 0}{\longrightarrow} \nabla f(x_0)y$$

uniformly in $\overline{B}_1(0)$. This convergence property is immediately seen to be equivalent to the classical differentiability of f at x_0, with gradient equal to $\nabla f(x_0)$. $\square$

The proof of the following classical result can be found, for instance, in [11] and [12].

Theorem 19.14 (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[6] *and*

$$\int_A |\det \nabla f(x)|\, dx = \int_{\mathbb{R}^n} N(y, A)\, dy \geq \mathscr{L}^n(f(A)) \ .$$

Definition 19.15 (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). Also, the symmetry assumption on S is not restrictive, since in the formula S can also be replaced by its symmetric part.

We are now ready to prove the main result of this section, Alexandrov theorem.

Theorem 19.16 (Alexandrov)**.** *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 19.11. Obviously, there is no loss of generality supposing that f is λ-convex for some $\lambda > 0$.

[6] In particular, notice that $f(A) = \{N > 0\}$.

We briefly recall, from Remark 19.11, 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 Theorem 19.13 and the area formula 19.14 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 name

$$S(x) := \left(\nabla\Psi(p)\right)^{-1} \ .$$

If $y = \Psi(q)$, we get

$$\begin{aligned} S(x)^{-1}\left(q - p - S(x)(y - x)\right) &= -\left(y - x - \nabla\Psi(p)(q - p)\right) \\ &= -\left(\Psi(q) - \Psi(p) - \nabla\Psi(p)(q - p)\right) \\ &= o(|p - q|) = o(|x - y|) \ . \end{aligned}$$

Therefore

$$\lim_{\substack{y \to x \\ q \in \partial f(y)}} \frac{|q - \nabla f(x) - S(x)(y - x)|}{|y - x|} = 0 \ . \tag{19.9}$$

The result got in (19.9), together with the nonsmooth mean value Theorem 19.4, give us the second order expansion. In fact, 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 read (19.9) as $\displaystyle\lim_{q \in \partial \tilde{f}(x),\, y \to x} |q|/|y - x| = 0$. Now, choose $\theta \in [0, 1]$ and a vector $q \in \partial \tilde{f}((1 - \theta)y + \theta x)$ such that $\tilde{f}(y) = \langle q, y - x \rangle$ (since $\tilde{f}(x) = 0$) to find

$$\tilde{f}(y) = \langle q, y - x \rangle = o(|y - x|^2) \ .$$

By the very definition of $\tilde{f}$, the statement follows. $\square$

Remark 19.17 (Characterization of S)**.** A blow-up analysis, analogous to the one performed in the proof of Rademacher's theorem, shows that the matrix $S(x)$ in Alexandrov's theorem is the density of the measure $\nabla^2 f$ with respect to $\mathscr{L}^n$, see [2] for details.

20 Viscosity solutions

20.1 Basic definitions

In this section we want to give the notion of *viscosity solution* for general equations having the form

$$E(x, u(x), \nabla u(x), \nabla^2 u(x)) = 0 \tag{20.1}$$

where u is defined on some *locally compact* domain $A \subset \mathbb{R}^n$. This topological assumptions is actually very useful, because we can deal at the same time with open and closed domains, and also domain of the form $\mathbb{R}^{n-1} \times [0, \infty)$, which typically occur in parabolic problems.

We first need to recall two classical ways to regularize a function.

Definition 20.1 (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 one of the following equivalent formulas:*

$$\begin{aligned} u^*(x) &:= \sup\left\{\limsup_h u(x_h) \,|\, (x_h) \subset A',\ x_h \to x\right\} \\ &= \inf_{r>0} \sup_{B_r(x)\cap A'} u \\ &= \min\left\{v \,|\, v \text{ is u.s.c. and } v \geq u\right\}. \end{aligned}$$

Similarly we can define the lower regularization u_*

$$\begin{aligned} u_*(x) &:= \inf\left\{\liminf_h u(x_h) \,|\, (x_h) \subset A',\ x_h \to x\right\} \\ &= \sup_{r>0} \inf_{B_r(x)\cap A'} u \\ &= \max\left\{v \,|\, v \text{ is l.s.c. and } v \leq u\right\} \end{aligned}$$

which is also characterized by the identity $u_* = -(-u)^*$.

Remark 20.2. It is clear that pointwise $u_* \leq u \leq u^*$. In fact, 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 $u_*(x) = u^*(x)$.

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.

Definition 20.3 (Subsolution)**.** *A function $u : A \to \mathbb{R}$ is a subsolution for the equation* (20.1) *(and we write $E \leq 0$) if the two following conditions hold:*

(i) u^* is a real-valued function;

(ii) for any $x \in A$, if φ is C^∞ in a neighbourhood of x and $u^* - \varphi$ has a local maximum at x, then

$$E_*(x, u^*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \leq 0 \ . \tag{20.2}$$

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.

The geometric idea in 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 $\nabla^2 u^*(x) \leq \nabla^2\varphi(x)$. So, while in the classical theory of PDEs an integration by parts formula allows to transfer derivatives from u to the test function φ, here the comparison principle allows to transfer (to some extent, since only an inequality holds for second order derivatives) the derivatives from u to the test function φ.

Similarly, we give the following:

Definition 20.4 (Supersolution). *A function $u : A \to \mathbb{R}$ is a supersolution for the equation* (20.1) *(and we write $E \geq 0$) if the two following conditions hold:*

(i) u_* is a real-valued function;

(ii) for any $x \in A$, if φ is C^∞ in a neighbourhood of x and $u_* - \varphi$ has a local minimum at x, then

$$E^*(x, u_*(x), \nabla\varphi(x), \nabla^2\varphi(x)) \geq 0. \tag{20.3}$$

We finally say that u is a solution of our problem if it is both a subsolution and a supersolution.

Remark 20.5. 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 (20.2). 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 20.6. 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 20.7. Rather surprisingly, a solution of $E = 0$ in the viscosity sense does not necessarily solve $-E = 0$ in the viscosity sense. To show this, consider the equations $|f'| - 1 = 0$ and $1 - |f'| = 0$ and the function $f(t) = \min\{1 - t, 1 + t\}$. In this case, it is immediate to see that f is a subsolution of the first problem (and actually a solution, as we will see), but it is not a subsolution of the second problem, since we can choose identically $\varphi = 1$ to find that the condition $1 - |\varphi'(0)| \le 0$, corresponding to (20.2), is violated.

We have instead the following parity properties:

(a) Let E be *odd* in (u, p, S). If u verifies $E \le 0$, then $-u$ verifies $E \ge 0$.

(b) Let E be *even* in (u, p, S). If u verifies $E \le 0$, then $-u$ verifies $-E \ge 0$.

We now spend some words on the ways of simplifying the conditions that have to be checked in order prove the subsolution or supersolution property. We just examine the case of subsolutions, the case of supersolutions being the same (with obvious variants).

We have already seen in Remark 20.5 that we can assume without loss of generality that $u^* - \varphi$ has a strict local maximum, equal to 0, at x. We can also work equivalently with the larger class of C^2 functions φ, in a neighbourhood of x. One implication is trivial, let us see the converse one. Let $\varphi \in C^2$ and assume $u^*(y) - \varphi(y) \le 0$ for $y \in \overline{B}_r(x)$, with equality only when $y = x$. By appropriate mollifiers, we can build a sequence $(\varphi_k) \subset C^\infty(\overline{B}_r(x))$ with $\varphi_k \to \varphi$ in $C^2(\overline{B}_r(x))$. 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

$$E_*(x_k, u^*(x_k), \nabla\varphi_k(x_k), \nabla^2\varphi_k(x_k)) \le 0$$

and we can now let $k \to \infty$ and use the lower semicontinuity of E_* to get the thesis.

Actually, it is rather easy now to see that the subsolution property is even equivalent to

$$E_*(x, u^*(x), p, S) \le 0 \quad \forall\ (p, S) \in J_2^+ u^*(x)$$

where $J_2^+ u^*$ is the second-order super jet of u, namely

$$J_2^+ u^*(x) := \left\{ (p, S) \,\middle|\, u^*(y) \le u^*(x) + \langle p, y - x\rangle + \tfrac{1}{2}\langle S(y - x), y - x\rangle + o(|y - x|^2) \right\}.$$

Indeed, let $P(y) := u^*(x) + \langle p, y - x\rangle + \frac{1}{2}\langle S(y - x), y - x\rangle$, so that $u^*(y) \le P(y) + o(|y - x|^2)$, with equality when $y = x$. Hence, for any $\varepsilon > 0$ we have $u^*(y) \le P(y) + \varepsilon|y - x|^2$ on a sufficiently small neighbourhood of x with equality at $y = x$ and we can apply (20.2) to this smooth function to get

$$E_*(x, u^*(x), p, S + 2\varepsilon I) = E_*(x, u^*(x), \nabla P(x), \nabla^2 P(x) + 2\varepsilon I) \le 0$$

and by lower semicontinuity we can let $\varepsilon \to 0$ and prove the claim. Of course, if we are dealing with first order equations, only the first order super jet is needed.

Remark 20.8. After these preliminary facts, 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 comes from local comparison), on the other hand it cannot be generalized to vector-valued functions.

20.2 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 a viscosity supersolution (take $\varphi \equiv 0$ and study the situation at the origin).

Since we can always take $u=\varphi$ if u is at least C^2, the following theorem is trivial:

Theorem 20.9 (C^2 viscosity solutions are classical solutions)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, $u \in C^2(\Omega)$ and E continuous. If u is a viscosity solution of* (20.1) *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 20.10 (Classical solutions are viscosity solutions)**.** *If u is a classical subsolution (resp. supersolution) of* (20.1)*, 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 $\nabla^2 u(x) \leq \nabla^2\varphi(x)$ and by definition $E_*(x,u(x),\nabla u(x),\nabla^2 u(x)) \leq 0$. Consequently, exploiting our monotonicity assumption we obtain $E_*(x,u(x),\nabla\varphi(x),\nabla^2\varphi(x)) \leq 0$ 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 $(\partial_t-\Delta)u-g=0$ if we let $x:=(y,t)\in\mathbb{R}^n\times(0,\infty)$ with $A=\mathbb{R}^n\times(0,\infty)$. Secondly, it is worth remarking that some authors adopt a different convention, which we might call *elliptic convention*, which is "opposite" to the one we gave before. Indeed, according to this convention, if (for instance) we deal with a problem of the form $F(\nabla^2 u)=0$, we require for a subsolution that $u^*-\varphi$ has a maximum at x implies $F(\nabla^2\varphi(x)) \geq 0$ (i.e. a subsolution of $-F(\nabla^2 u)=0$ in our terminology). As a consequence, in the previous theorem, we should replace "non-increasing" with "non-decreasing."

Now, we are ready to introduce the first important tool for the following theorems.

Theorem 20.11. *Let $\mathcal{F}$ be a family of subsolutions of* (20.1) *in A and let $u : A \to \overline{\mathbb{R}}$ be defined by*

$$u(x) := \sup\{v(x) \mid v \in \mathcal{F}\}\,.$$

Then u is a subsolution of the same problem on the domain $A \cap \{u^* < \infty\}$ (since $\{u^* < \infty\}$ 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 to be chosen sufficiently small, so that x is the unique maximum of $u * -\varphi$ on K.

By a diagonal argument can find a sequence (x_h) inside 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 call y_h the maximum of $v_h^* - \varphi$ on K, then

$$u^*(y_h) - \varphi(y_h) \geq v_h^*(y_h) - \varphi(y_h) \geq v_h^*(x_h) - \varphi(x_h) \geq v_h(x_h) - \varphi(x_h).$$

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 of (y_h) satisfies

$$u^*(y) - \varphi(y) \geq 0.$$

Hence y is a maximum in K of $u^* - \varphi$, $u^*(y) - \varphi(y) = 0$ and y must coincide with x. Consequently $y_h \to x$, $\limsup_h (u^*(y_h) - \varphi(y_h) \leq u^*(x) - \varphi(x)$ and, by comparison, the same is true for the intermediate terms, so that $v_h^*(y_h) \to u^*(x)$. In order to conclude, we just need to consider the viscosity condition at the points y_h, which reads

$$E_*(y_h, v_h^*(y_h), \nabla\varphi(y_h), \nabla^2\varphi(y_h)) \leq 0 \ ,$$

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.

Theorem 20.12 (Perron). *Let f and g be respectively a subsolution and a supersolution of* (20.1)*, 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, then there exists a solution u of* (20.1) *satisfying $f \leq u \leq g$.*

Proof. Call

$$\mathcal{F} := \{v \,|\, v \text{ is a subsolution of (20.1) and } v \leq g\} \ .$$

We know that $f \in \mathcal{F}$, so that this set is not empty. Hence, we can define $u := \sup\{v| \ v \in \mathcal{F}\}$. By our definition of $\mathcal{F}$, we have that $u \leq g$ and therefore $u^* \leq g^* < +\infty$. Since $u^* \geq u_* \geq f_* > -\infty$, in A, by Theorem 20.11 u is a subsolution on A. Consequently, 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

$$u_*(x) - \varphi(x) \geq |x - x_0|^4 \qquad \text{on } A \cap \overline{B}_r(x) \tag{20.4}$$

for some sufficiently small $r > 0$. Assume by contradiction that

$$E^*(x_0, u_*(x_0), \nabla\varphi(x_0), \nabla^2\varphi(x_0)) < 0 \tag{20.5}$$

and define a function $w := \max\{\varphi + \delta^4, u\}$ for some parameter $\delta > 0$. We claim that:

(a) w is a subsolution of (20.1);

(b) $\{w > u\} \neq \emptyset$;

(c) $w \leq g$ (and hence $w \in \mathcal{F}$),

provided we choose δ sufficiently small in (a) and (c).

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 (20.1) 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 (20.1) on $\overline{B}_{2\delta}(x_0) \cap A$. Moreover, by (20.4), 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 subsolution on A.[7]

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 on any sequence (x_h) such that $u(x_h) \to u_*(x_0)$, we must have for h sufficiently large the inequality $u(x_h) < \varphi(x_h) + \delta^4$.

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 (20.5). □

20.3 The distance function

Our next goal is now to study the uniqueness problem, which is actually very delicate as the previous examples show. We begin here with a special case.

Let $C \subset \mathbb{R}^n$ be a closed set, $C \neq \emptyset$ and let $u(x) := \mathrm{dist}(x, C)$. We claim that the distance function is a viscosity solution of the equation $|p|^2 - 1 = 0$ on $A := \mathbb{R}^n \setminus C$.

First of all, it is clearly a viscosity supersolution in A. This follows by Theorem 20.11 (in the obvious version for supersolutions), once we observe that $u(x) = \inf_{y \in C} |x - y|$ and

[7] 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 it is also a subsolution on A.

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 follows by the general implication:

$$\mathrm{Lip}(f) \leq 1 \;\Rightarrow\; |\nabla f|^2 - 1 \leq 0 \qquad \text{in the sense of viscosity solutions.}$$

Indeed, let x be a local maximum for $f - \varphi$, so that $f(y) - \varphi(y) \leq f(x) - \varphi(x)$ for any $y \in B_r(x)$ (and r small enough). This is equivalent, on the same domain, to $\varphi(y) - \varphi(x) \geq f(y) - f(x) \geq -|y-x|$ and, by the Taylor expansion, we finally get

$$\langle \nabla\varphi(x), y - x\rangle + o(|y-x|) \geq -|y-x| \;.$$

This readily implies the claim.

The converse implication is less trivial, but still true! Namely

$$|\nabla f|^2 - 1 \leq 0 \text{ in the sense of viscosity solutions} \quad \Rightarrow \quad \mathrm{Lip}(f) \leq 1$$

for f continuous (or at least upper semicontinuous), which is proved by means of the regularizations $f^\varepsilon(x) := \sup_y \big(f(y) - |x-y|^2/\varepsilon\big)$ that we will study more in detail later on. We just sketch here the structure of the argument:

(1) still $|\nabla f^\varepsilon|^2 - 1 \leq 0$ in the sense of viscosity solutions;

(2) $|\nabla f^\varepsilon|^2 - 1 \leq 0$ pointwise $\mathscr{L}^n$-a.e., because f^ε is semiconcave, hence locally Lipschitz, and therefore the inequality holds at any differentiaiblity point by the super-jet characterization of viscosity subsolutions;

(3) by Proposition 1.4 one obtains $\mathrm{Lip}(f^\varepsilon) \leq 1$;

(4) $f^\varepsilon \downarrow f$ and hence $\mathrm{Lip}(f) \leq 1$.

We now come to our uniqueness result.

Theorem 20.13. *Let $C \subset \mathbb{R}^n$ be a closed set as above, $A = \mathbb{R}^n \setminus C$ and let $u \in C(\overline{A})$ be a non-negative viscosity solution of $|p|^2 - 1 = 0$ on A with $u = 0$ on ∂A. Then $C \neq \emptyset$ and* $u(x) = \mathrm{dist}(x, C)$.

Proof. By our assumptions we can clearly extend u continuously to $\mathbb{R}^n$, so that $u = 0$ identically on C. It is immediate to verify that $|\nabla u|^2 - 1 \leq 0$ in the sense of viscosity solutions on $\mathbb{R}^n$. Consequently, thanks to the previous regularization argument, $\mathrm{Lip}(u) \leq 1$ and hence, 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 will write $w(x)$ for the distance function $\mathrm{dist}(x, C)$, It remains to show that $w \le u$. 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

$$\sup_{x,y} \left\{ w(x) - (1+\lambda)u(y) - \frac{1}{2\varepsilon}|x-y|^2 \right\} \ge \gamma_0$$

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 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, call $(\overline{x}, \overline{y})$ 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 $|\nabla\varphi|^2(\overline{x}) \le 1$, that is

$$\frac{|\overline{x}-\overline{y}|}{\varepsilon} \le 1 \ .$$

We also claim that necessarily $\overline{y} \in A$, if ε is sufficiently small, precisely $\varepsilon < \gamma_0$. Indeed, assume by contradiction that $\overline{y} \notin A$, so that $w(\overline{y}) = 0$, then by the triangle inequality

$$\gamma_0 \le w(\overline{x}) - \frac{1}{2\varepsilon}|\overline{x}-\overline{y}|^2 \le |\overline{x}-\overline{y}| - \frac{1}{2\varepsilon}|\overline{x}-\overline{y}|^2 \le |\overline{x}-\overline{y}| \, .$$

As a consequence, we get $\gamma_0 \le |\overline{x}-\overline{y}| \le \varepsilon$, 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

$$\left| \frac{\overline{x}-\overline{y}}{\varepsilon} \right| \ge (1+\lambda) \ ,$$

which is not compatible with $|\overline{x}-\overline{y}| \le \varepsilon$. Hence, at least when A is bounded, we have proved that $w = u$.

In the general case, fix a constant $R > 0$ and define $u_R(x) := u(x) \wedge \mathrm{dist}(x, \mathbb{R}^n \setminus \overline{B}_R)$: 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 (by the infimum property). Moreover, $\mathrm{Lip}(u_R) \le 1$ implies that u_R is a global subsolution and we can apply the previous result (special case) to the function u_R to get

$$u_R(x) = d(x, \mathbb{R}^n \setminus (A \cap B_R)).$$

Letting $R \to \infty$ we first exclude $C = \emptyset$ since in that case $u_R \uparrow \infty$ which is not admissible since $u_R \le u$ and then (by $C \ne \emptyset$) we obtain $u(x) = \mathrm{dist}(x, C)$. $\square$

Remark 20.14. We can also give a different interpretation of the result above. 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 $i \in \{1, \ldots, n\}$).

20.4 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 will always assume that $F(p, S)$ is non-increasing in its second variable S, so that classical solutions are viscosity solutions.

Let us begin with some heuristics. Let f, $g \in C^2(A) \cap C(\overline{A})$, with A bounded, and assume that f is a subsolution on A, g is a supersolution on A, $f \leq g$ on ∂A and that one of the inequalities $F(\nabla f, \nabla^2 f) \leq 0$, $F(\nabla g, \nabla^2 g) \geq 0$ is always strict. Then $f \leq g$ in A. Indeed, assume by contradiction $\sup_A (f - g) > 0$, then there exists a $x_0 \in A$ which is a maximum for $f - g$. Consequently $\nabla f(x_0) = \nabla g(x_0)$ and also $\nabla^2 f(x_0) \leq \nabla^2 g(x_0)$. These two facts imply, by the monotonicity of F, that

$$F(\nabla f(x_0), \nabla^2 f(x_0)) \geq F(\nabla g(x_0), \nabla^2 g(x_0)) \ . \tag{20.6}$$

On the other hand, f (resp. g) is also a regular subsolution (resp. supersolution) so that

$$F(\nabla f(x_0), \nabla^2 f(x_0)) \leq 0, \qquad F(\nabla g(x_0), \nabla^2 g(x_0)) \geq 0 \ . \tag{20.7}$$

Hence, if we compare (20.6) with (20.7), we find a contradiction as soon as one of the two inequalities in (20.7) is strict.

In order to hope for a comparison principle, this argument shows the necessity to approximate subsolutions (or supersolutions) with strict subsolutions, and this is always linked to some form of strict monotonicity of the equation, variable from case to case (of course in the trivial case $F \equiv 0$ no comparison principle is possible). To clarify this point, let us consider the following example. Consider the space-time coordinates $x = (y, t)$ and a parabolic problem

$$F(\nabla_{y,t} u, \nabla^2_{y,t} u) = \partial_t u - G(\nabla^2_y u)$$

with G non-decreasing, in the appropriate sense. In this case, we can reduce ourselves to strict inequalities by performing the transformation $u \rightsquigarrow e^{\lambda t} u$.

In order to get a general uniqueness result for viscosity solution, we cannot just argue as in the case of the distance function and we need to follow a strategy introduced by Jensen. The first step is to obtain a refined versions of the maximum principle. We start with an elementary observation.

Remark 20.15. If $(p, S) \in J_2^- u(x)$ and u has a relative maximum at x, then necessarily $p = 0$ and $S \le 0$. To see this, it is enough to apply the definitions: by our two hypotheses

$$0 \ge u(y) - u(x) \ge \langle p, y - x\rangle + \frac{1}{2}\langle S(y-x), y-x\rangle + o(|y-x|^2)$$

and hence

$$\langle p, \frac{y-x}{|y-x|}\rangle \le o(|y-x|) \ \Rightarrow p = 0\,,$$
$$\frac{\langle S(y-x), y-x\rangle}{|y-x|^2} \le o(1) \ \Rightarrow S \le 0\,.$$

We are now ready to state and prove Jensen's maximum principle for semiconvex functions.

Theorem 20.16 (Jensen's maximum principle)**.** *Let $u : \Omega \to \mathbb{R}$ be semiconvex and let $x_0 \in \Omega$ a local maximum for u. Then, there exist a sequence (x_k) converging to x_0 and $\varepsilon_k \downarrow 0$ such that u is pointwise second order differentiable at x_k and*

$$\nabla u(x_k) \to 0 \qquad\qquad \nabla^2 u(x_k) \le \varepsilon_k I\ .$$

The proof is based on the following lemma. In the sequel we shall denote by $\mathrm{sc}(u, \Omega)$ the least nonnegative constant C such that u is $(-C)$-convex, i.e. $u + C|x|^2/2$ is convex (recall Definition 19.8).

Theorem 20.17. *Let $B \subset \mathbb{R}^n$ be a ball of radius R centered at the origin and $u \in C(\overline{B})$ semiconvex, with*[8]

$$\max_{\overline{B}} u > \max_{\partial B} u\ .$$

Then, if we let

$$G^\delta = \left\{x \in B \,\middle|\, \exists\, p \in \overline{B}_\delta \ s.t.\ u(y) \le u(x) + \langle p, x - y\rangle,\ \forall y \in B\right\}$$

it must be

$$\mathscr{L}^n(G^\delta) \ge \frac{\omega_n \delta^n}{[\mathrm{sc}(u,B)]^n} \tag{20.8}$$

for $0 < \delta < \left(\max_{\overline{B}} u - \min_{\overline{B}} u\right)/(2R)$.

Proof. We assume first that u is also in $C^1(B)$. Pick a $\delta > 0$, so small that $2R\delta < \max_{\overline{B}} u - \max_{\partial B} u$, and consider a perturbation $u(y) + \langle p, y\rangle$ with $|p| \le \delta$. We claim that such function necessarily attains its maximum in B. Indeed, this immediately comes from the two inequalities

$$\max_{\partial B}\left(u + \langle p, y\rangle\right) \le \max_{\partial B} u + \delta R$$

[8]Notice that this implies $\mathrm{sc}(u, B) > 0$, since $\max_{\overline{B}} = \max_{\partial B}$ for convex functions.

and

$$\max_{\overline{B}}(u + \langle p, y\rangle) \geq \max_{\overline{B}} u - \delta R \ .$$

Consequently, there exists $x \in B$ such that $\nabla u(x) = -p$. This shows that $\nabla u(G^\delta) = \overline{B}_\delta$. To go further, we need the area formula. In this case, it gives

$$\int_{G^\delta} |\det \nabla^2 u| \, dx = \int_{\overline{B}_\delta} \mathrm{card}\left(\{x \mid \nabla u(x) = p\}\right) \, dp \geq \omega_n \delta^n$$

by the previous statement. On the other hand

$$\int_{G^\delta} \left|\det \nabla^2 u\right| \, dx \leq [\mathrm{sc}(u, B)]^n \, \mathscr{L}^n(G^\delta) \ ,$$

because the points in G^δ are maxima for the function $u(y)+\langle p, y\rangle$: this implies $\nabla^2 u(x) \leq 0$ for any $x \in G_\delta$ and, by semiconvexity, $\nabla^2 u(x) \geq -\mathrm{sc}(u, B)I$. If we combine these two inequalities, we get (20.8).
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 and $\limsup_h \mathrm{sc}(u_h, B_{r_h}) \leq \mathrm{sc}(u, B)$; 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 $\mathscr{L}^n(G^\delta(u)) \geq \limsup_h \mathscr{L}^n(G^\delta(u_h) \cap B_{r_h})$. □

We can now prove Jensen's maximum principle. As a preliminary remark, observe that, in Definition 19.8 one has (for our u) $\lambda = 0$ then the claim is trivial, so that we can without loss of generality assume that $\lambda < 0$ and Theorem 20.17 applies.

Proof. Let x_0 be a local maximum of u. We can choose $R > 0$ sufficiently small so that $u \leq u(x_0)$ 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 20.17 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 the Alexandrov theorem) 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 $|p_k| \leq 1/k$, 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 $\nabla^2\widetilde{u}(x_k) \leq 0$. 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

$$\nabla^2|z|^4 = 4|z|^2 I + 8z \otimes z \tag{20.9}$$

gives

$$\begin{aligned} \nabla^2 u(x_k) &= \nabla^2\widetilde{u}(x_k) + 8(x_k - x_0) \otimes (x_k - x_0) + 4|x_k - x_0|^2 I \\ &\leq \nabla^2\widetilde{u}(x_k) + 12|x_k - x_0|^2 I \ . \end{aligned}$$

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 20.18 (Inf and sup-convolutions)**.** *Given $u : A \to \mathbb{R}$ and a parameter $\varepsilon > 0$, we can build the regularized functions*

$$u^\varepsilon(x) := \sup_{y\in A}\left\{u(y) - \frac{1}{\varepsilon}|x-y|^2\right\} \tag{20.10}$$

which are called sup-convolutions of u and satisfy $u^\varepsilon \geq u$, and

$$u_\varepsilon(x) := \inf_{y\in A}\left\{u(y) + \frac{1}{\varepsilon}|x-y|^2\right\}. \tag{20.11}$$

which are called inf-convolutions of u and satisfy $u_\varepsilon \leq u$.

In the next proposition we summarize the main properties of sup-convolutions; analogous properties hold for inf-convolutions.

Proposition 20.19 (Properties of sup-convolutions)**.** *Assume that u is u.s.c. on A and that $u(x) \leq K(1+|x|)$ for some constant $K \geq 0$, then*

(i) *u^ε is semiconvex and* $\mathrm{sc}(u^\varepsilon, \mathbb{R}^n) \leq 2/\varepsilon$;

(ii) *$u^\varepsilon \geq u$ and $u^\varepsilon \downarrow u$ pointwise in A. If u is continuous, then $u^\varepsilon \downarrow u$ locally uniformly;*

(iii) *if $F(\nabla u, \nabla^2 u) \leq 0$ in the sense of viscosity solutions on A, then $F(\nabla u^\varepsilon, \nabla^2 u^\varepsilon) \leq 0$ on A^ε, where*

$$A^\varepsilon := \{x \in \mathbb{R}^n |\ \textit{the supremum in } (20.10) \textit{ is attained}\}.$$

Proof. (i) First of all, 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 $\mathrm{sc}(u^\varepsilon, \mathbb{R}^n) \leq 2/\varepsilon$.
(ii) The inequality $u^\varepsilon \geq u$ and the monotonicity in ε are trivial. In addition, we can take quasi-maxima (y_ε) satisfying

$$u^\varepsilon(x) \leq u(y_\varepsilon) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon \leq K(1+|y_\varepsilon|) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon \leq K(1+|x|+|\delta_\varepsilon|) - \frac{\delta_\varepsilon^2}{\varepsilon} + \varepsilon\ .$$

with $\delta_\varepsilon = |y_\varepsilon - x|$. Via these two inequalities, one first sees that $y_\varepsilon \to x$ so that, exploiting the upper semicontinuity of u and neglecting the quadratic term in the first inequality we get

$$u(x) \geq \limsup_{\varepsilon \to 0} u(y_\varepsilon) \geq \limsup_{\varepsilon \to 0} u^\varepsilon(x) \ .$$

If u is continuous, the claim comes from Dini's monotone convergence theorem and the local compactness of A.
(iii) Let $x_0 \in A^\varepsilon$ and let $y_0 \in A$ be a corresponding maximum, 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, we can take $u^\varepsilon(x_0) = \varphi(x_0)$. Let us call r the radius such that $u^\varepsilon \leq \varphi$ 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

$$F(\nabla\psi(y_0), \nabla^2\psi(y_0)) \leq 0$$

and, by the definition of ψ, this is equivalent to

$$F(\nabla\varphi(x_0), \nabla^2\varphi(x_0)) \leq 0.$$

This 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 $u^\varepsilon(x) \leq \varphi(x)$ in $B_r(x_0)$ gives

$$u(y) - \frac{1}{\varepsilon}|x - y|^2 \leq \varphi(x) \quad \forall x \in B_r(x_0), \forall y \in A$$

and, letting $y = x - x_0 + y_0 \in A$ with $x \in B_r(x_0)$, this implies

$$u(y) - \psi(y) \leq \frac{1}{\varepsilon}|x_0 - y_0|^2 \quad \forall y \in A \cap B_r(y_0) \ .$$

□

Remark 20.20. *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^ε, where*

$$A^{\varepsilon,\delta} := \{x \in \mathbb{R}^n | \text{ the supremum in } (20.10) \text{ is attained at some } y \in B_\delta(x) \cap A\} \ ,$$

$$F^\delta(x, p, S) := \inf\{F(y, p, s) : \ y \in B_\delta(x) \cap A\} \ . \tag{20.12}$$

An analogous result holds for supersolutions

20.5 Existence and uniqueness results

In this section we will collect some existence and uniqueness results for second order equations. The main tool is the comparison principle, stated below. Throughout the section we shall always assume that A is a bounded open set in $\mathbb{R}^n$.

Proposition 20.21 (Comparison principle)**.** *Let $F : A \times \mathrm{Sym}^{n\times n} \to \mathbb{R}$ be continuous and satisfying, for some $\lambda > 0$, the strict monotonicity condition*

$$F(x, S + tI) \geq F(x, S) + \lambda t \qquad \forall t \geq 0$$

and the uniform continuity assumption

$$F(\cdot, S),\ S \in \mathrm{Sym}^{n\times n},\ \textit{are equi-continuous in } A.$$

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 equations of the form $G(\nabla^2 u) + f(x)$ with f continuous in A.

A direct consequence of the comparison principle (take $\underline{u} = \overline{u} = u$) is the following uniqueness result:

Theorem 20.22 (Uniqueness of continuous solutions)**.** *Let F be as in Proposition 20.21 and $h \in C(\partial A)$. Then the problem*

$$\begin{cases} -F(x, \nabla^2 u(x)) = 0 & \textit{in } A; \\ \\ u = h & \textit{on } \partial A \end{cases} \tag{20.13}$$

admits at most one viscosity solution $u \in C(\overline{A})$.

At the level of existence, we can exploit Theorem 20.12 to obtain the following result.

Theorem 20.23 (Existence of continuous solutions)**.** *Let F be as in Proposition 20.21 and let f, $g : A \to \mathbb{R}$ be respectively a subsolution and a supersolution of $-F(x, D^2 u) = 0$ 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 to* (20.13) *with $h = g^* = f_*$.*

In order to prove this last result, it suffices to take any solution u given by Perron's method (see Theorem 20.12), 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 will be devoted to the proof of the comparison principle, which uses besides doubling of variables, inf and sup-convolutions (see Definition 20.18) and Jensen's maximum principle (see Theorem 20.16).

Lemma 20.24. *Let F, $\underline{u}$ and $\overline{u}$ be as in Proposition 20.21 and set*

$$F_\gamma(x,S) := F(x, S-\gamma I) \le F(x,S) - \gamma\lambda \ ,$$

with $\gamma > 0$. For any $\delta > 0$, consider the function

$$v_{\delta,\gamma} := \underline{u} - \delta + \frac{\gamma}{2}|x|^2 \ .$$

Hence:

(i) $v_{\delta,\gamma}$ solves $-F_\gamma(x,\nabla^2 v_{\delta,\gamma}) \le 0$ in the viscosity sense;

(ii) if $\delta \ge \delta(\gamma,A)$ is large enough, then $v_{\delta,\gamma} \le \overline{u}$ 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

$$v_{\delta,\gamma} \le \overline{u} \quad \text{on } A,\ \forall\, \delta > \delta(\gamma,A) \ , \tag{20.14}$$

then $\underline{u} \le \overline{u}$ on A.

Proof. Statements (i) follows by the translation invariance w.r.t. u of the equation, and by $\nabla^2 v_\delta = \nabla^2 u + \gamma I$. Statement (ii) follows by the fact that $\underline{u} < \overline{u}$ on ∂A.
If (20.14) holds, then

$$\underline{u} - \delta \le v_{\delta,\gamma} \le \overline{u} \qquad \text{on } A \ ,$$

and the comparison principle for $\underline{u}$ follows letting $\gamma \downarrow 0$, which allows to choose arbitrarily small δ in view of (ii). ☐

Proof. (of Proposition 20.21) Thanks to Lemma 20.24, without loss of generality we can assume that $\underline{u}$ satisfies the stronger property

$$-F_\gamma(x,\nabla^2 u) \le 0$$

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 let us consider the sup convolution

$$u^\varepsilon(x) := \sup_{x' \in A}\left(\underline{u}(x') - \frac{1}{\varepsilon}|x-x'|^2\right) = \max_{x' \in \overline{A}}\left((\underline{u})^*(x') - \frac{1}{\varepsilon}|x-x'|^2\right), \tag{20.15}$$

of $\underline{u}$ and the inf convolution

$$u_\varepsilon(y) := \inf_{y' \in A}\left(\overline{u}(y') + \frac{1}{\varepsilon}|y-y'|^2\right) = \min_{y' \in \overline{A}}\left((\overline{u})_*(y') + \frac{1}{\varepsilon}|y-y'|^2\right), \tag{20.16}$$

of $\overline{u}$; since $u^\varepsilon \geq \underline{u}$ and $u_\varepsilon \leq \overline{u}$ we have

$$\max_{\overline{A}\times\overline{A}} \left(u^\varepsilon(x) - u_\varepsilon(y) - \frac{1}{4\varepsilon}|x-y|^4\right) \geq u^\varepsilon(x_0) - u_\varepsilon(x_0) \geq \underline{u}(x_0) - \overline{u}(x_0) = d_0$$

and we shall denote by $(x_\varepsilon, y_\varepsilon) \in \overline{A}\times\overline{A}$ a maximizing pair, so that

$$d_0 + \frac{1}{4\varepsilon}|x_\varepsilon - y_\varepsilon|^4 \leq u^\varepsilon(x_\varepsilon) - u_\varepsilon(y_\varepsilon) \leq \sup \underline{u} - \inf \overline{u} \ . \tag{20.17}$$

Also, we denote by $x'_\varepsilon \in \overline{A}$ and $y'_\varepsilon \in \overline{A}$ maximizers and minimizers respectively in (20.15) and (20.16).

Now 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) setting $M = \max\{\operatorname{osc}(\underline{i}), \operatorname{osc}(\overline{u})\}$, for ε small enough, the supremum in (20.15) 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 (20.16) 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 (20.17) 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 the supremum of $(\underline{u})^* - (\overline{u})_*$ is finite, this 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 semicontinuity gives $d_0 \leq (\underline{u})^*(\bar{x}) - (\overline{u})_*(\bar{x})$. By assumption $(\underline{u})^* \leq (\overline{u})_*$ on ∂A, therefore $\bar{x} \in A$ and this proves (a).

To prove (b), it suffices to choose, thanks to (a), $\varepsilon_0 > 0$ and $\delta_0 > 0$ small enough, so that $\operatorname{dist}(x_\varepsilon, \partial A) \geq \delta_0$ for $\varepsilon \in (0, \varepsilon_0)$. In general, for $x \in A$ we have

$$\underline{u}(x') - \frac{1}{\varepsilon}|x'-x|^2 \leq \underline{u}(x) \leq u^\varepsilon(x)$$

which implies that the supremum in the definition of $u^\varepsilon(x)$ is unchanged if we maximize in the ball $\overline{B}_x$ centered at x with radius $\sqrt{M\varepsilon}$. If $|x - x_\varepsilon| < \epsilon$ and $\varepsilon < \varepsilon_0$, since $\operatorname{dist}(x_\varepsilon, \partial A) \geq \delta_0$, this implies that the ball $\overline{B}_x$ 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[9] function

$$w(x,y) := u^\varepsilon(x) - u_\varepsilon(y) - \frac{1}{4\varepsilon}|x-y|^4$$

[9]The local semiconvexity of w follows from Proposition 20.19.

to find $z_n := (x_{\varepsilon,n}, y_{\varepsilon,n}) \to (x_\varepsilon, y_\varepsilon)$ and $\delta_n \downarrow 0$ such that w is pointwise second order differentiable at z_n, $\nabla w(z_n) \to 0$ and $\nabla^2 w(z_n) \le \delta_n I$. By statement (b) and Remark 20.20, for n large enough we have

$$-\sup_{|x-x_{\varepsilon,n}|^2 \le M\varepsilon} F_\gamma(x, \nabla^2 u^\varepsilon(x_{\varepsilon,n})) \le 0, \qquad -\inf_{|y-y_{\varepsilon,n}|^2 \le M\varepsilon} F(y_{\varepsilon,n}, \nabla^2 u_\varepsilon(y_{\varepsilon,n})) \ge 0. \quad (20.18)$$

On the other hand, the upper bound on $\nabla^2 w(z_n)$ together with (20.9) give

$$\begin{cases} \nabla^2 u^\varepsilon(x_{\varepsilon,n}) - \frac{2}{\varepsilon}(x_{\varepsilon,n} - y_{\varepsilon,n}) \otimes (x_{\varepsilon,n} - y_{\varepsilon,n}) - \frac{1}{\varepsilon}|x_{\varepsilon,n} - y_{\varepsilon,n}|^2 I \le \delta_n I \ , \\ -\nabla^2 u_\varepsilon(y_{\varepsilon,n}) - \frac{2}{\varepsilon}(x_{\varepsilon,n} - y_{\varepsilon,n}) \otimes (x_{\varepsilon,n} - y_{\varepsilon,n}) - \frac{1}{\varepsilon}|x_{\varepsilon,n} - y_{\varepsilon,n}|^2 I \le \delta_n I \ . \end{cases} \quad (20.19)$$

By (20.19) we obtain that $\nabla^2 u^\varepsilon(x_{\varepsilon,n})$ are uniformly bounded above, and they are also uniformly bounded below, since u^ε is semiconvex. Since similar remarks apply to $\nabla^2 u_\varepsilon(y_{\varepsilon,n})$, we can assume with no loss of generality that $\nabla^2 u^\varepsilon(x_{\varepsilon,n}) \to X_\varepsilon$ and $\nabla^2 u_\varepsilon(y_{\varepsilon,n}) \to Y_\varepsilon$. 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

$$\langle \nabla^2 u_\varepsilon(x_{\varepsilon,n})\xi, \xi\rangle - \langle \nabla^2 u_\varepsilon(y_{\varepsilon,n})\xi, \xi\rangle \le 2\delta_n |\xi|^2 \ .$$

Taking limits, this proves that $X_\varepsilon \le Y_\varepsilon$. On the other hand, from (20.18) we get

$$-\sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, X_\varepsilon) \le 0 \quad \text{and} \quad -\inf_{y \in \overline{B}_{\sqrt{M\varepsilon}}(y_\varepsilon)} F(y, Y_\varepsilon) \ge 0 \ .$$

Now, the strict monotonicity of $F(x, \cdot)$ yields

$$\sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F(x, Y_\varepsilon) \ge \sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, Y_\varepsilon) + \lambda\gamma \ge \sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F_\gamma(x, X_\varepsilon) + \lambda\gamma \ge \lambda\gamma \ .$$

Hence

$$\sup_{x \in \overline{B}_{\sqrt{M\varepsilon}}(x_\varepsilon)} F(x, Y_\varepsilon) - \inf_{y \in \overline{B}_{\sqrt{M\varepsilon}}(y_\varepsilon)} F(y, Y_\varepsilon) \ge \lambda\gamma \ .$$

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 ε sufficiently small. □

20.6 Hölder regularity

Consider a paraboloid P, i.e. a second-order polynomial of the form

$$P(x) = c + \langle p, x\rangle + \frac{1}{2}\langle Sx, x\rangle$$

for some $c \in \mathbb{R}$, $p \in \mathbb{R}^n$ and $S \in \mathrm{Sym}^{n\times n}$. We say that P is a paraboloid with opening $M \in \mathbb{R}$ if $S = MI$, namely

$$P(x) = c + \langle p, x\rangle + \frac{M}{2}|x|^2 .$$

It will be occasionally convenient to center a paraboloid P with opening M at some point x_0, writing $P(x) = P(x_0) + \langle \nabla P(x_0), x - x_0\rangle + \frac{M}{2}|x - x_0|^2$.

Definition 20.25 (Tangent paraboloids)**.** *Given a function $u : \Omega \to \mathbb{R}$ and a subset $A \subset \Omega \subset \mathbb{R}^n$, we denote*

$$\overline{\theta}(x_0, A, u) := \inf\{M \,|\text{there exists } P \text{ with opening } M,\ u(x_0) = P(x_0) \text{ and } u \le P \text{ on } A\} .$$

Moreover, we set

$$\underline{\theta}(x_0, A, u) := \sup\{M \,|\text{there exists } P \text{ with opening } M,\ u(x_0) = P(x_0) \text{ and } u \ge P \text{ on } A\} ,$$

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

$$\theta(x_0, A, u) := \max\left\{\underline{\theta}^-(x_0, A, u), \overline{\theta}^+(x_0, A, u)\right\} \ge 0 .$$

Given a function $u : \Omega \to \mathbb{R}$ and $h > 0$, let us consider the symmetric difference quotient in the direction $\xi \in \mathbb{R}^n$

$$\Delta^2_{h,\xi} u(x_0) := \Delta_{h,\xi}(\Delta_{h,\xi} u)(x_0) = \frac{u(x_0 + h\xi) + u(x_0 - h\xi) - 2u(x_0)}{h^2} \sim \frac{\partial^2 u}{\partial \xi^2}(x_0) ,$$

well defined if $h|\xi| < \mathrm{dist}(x_0, \partial\Omega)$ and identically equal to M on paraboloids with opening M. Notice that the symmetric difference quotient satisfies, by applying twice the integration by parts formula for $\Delta_{h,\xi}$,

$$\int_\Omega u \Delta^2_{h,\xi} \phi \, dx = \int_\Omega \phi \Delta^2_{h,\xi} u \, dx \tag{20.20}$$

whenever $u \in L^1_{\mathrm{loc}}(\Omega)$, $\phi \in L^\infty(\Omega)$ has compact support, $|\xi| = 1$ and the h-neighbourhood of $\mathrm{supp}\,\phi$ is contained in Ω.

Remark 20.26 (Maximum principle for Δ^2_ξ)**.** If a paraboloid P with opening M "touches" u from above (i.e. $P(x_0) = u(x_0)$ and $P(x) \ge u(x)$ in some ball $B_r(x_0)$), then

$$\Delta^2_{h,\xi} u(x_0) \le \Delta^2_{h,\xi} P(x_0) = M \qquad \text{whenever } \ |\xi| = 1 \text{ and } |h| \le r ,$$

and a similar property holds for paraboloids touching from below. Thus, passing to the infimum from above and the supremum from below, we deduce the inequalities

$$\underline{\theta}(x_0, B_r(x_0), u) \le \Delta^2_{h,\xi} u(x_0) \le \overline{\theta}(x_0, B_r(x_0), u) \ \text{ whenever } \ |\xi| = 1 \text{ and } |h| \le r , \tag{20.21}$$

and

$$|\Delta^2_{h,\xi} u(x_0)| \le \theta(x_0, B_r(x_0), u) \qquad \text{whenever } \ |\xi| = 1 \text{ and } |h| \le r . \tag{20.22}$$

Proposition 20.27. *If $u : \Omega \to \mathbb{R}$ satisfies*

$$\theta_\varepsilon := \theta(\,\cdot\,, B_\varepsilon(\cdot) \cap \Omega, u) \in L^p(\Omega)$$

for some $\varepsilon > 0$ and $1 < p \le \infty$, then u belongs to $W^{2,p}(\Omega)$ and, more precisely,

$$\|\nabla^2_{\xi\xi} u\|_{L^p(\Omega)} \le \|\theta_\varepsilon\|_{L^p(\Omega)} \qquad \forall \xi \in \mathbf{S}^{n-1}. \tag{20.23}$$

Remark 20.28. By bilinearity it is possible to obtain, from (20.23), an estimate on mixed second derivatives:

$$\|\nabla^2_{\xi\eta} u\|_{L^p(\Omega)} \le |\xi||\eta|\|\theta_\varepsilon\|_{L^p(\Omega)} \qquad \forall \xi, \eta \in \mathbb{R}^n,\ \xi \perp \eta\ .$$

Proof. For any $\varphi \in C^\infty_c(\Omega)$ one has

$$\begin{aligned} \left|\int_\Omega u(x)\frac{\partial^2\varphi}{\partial\xi^2}(x)\,dx\right| &= \left|\lim_{h\to 0}\int_\Omega u(x)\Delta^2_{h,\xi}\varphi(x)\,dx\right| \\ = &\left|\lim_{h\to 0}\int_\Omega (\Delta^2_{h,\xi}u(x))\varphi(x)\,dx\right| \le \|\theta_\varepsilon\|_{L^p(\Omega)}\|\varphi\|_{L^{p'}(\Omega)}\ , \end{aligned}$$

where we pass from the first to the second line with (20.20) and the inequality follows from (20.22). Thanks to Riesz representation theorem, we know that the map $\varphi \mapsto \int_\Omega u(x)\frac{\partial^2\varphi}{\partial\xi^2}(x)\,dx$ admits a representation with an element of $L^p(\Omega)$, which represents the derivative $\nabla^2_{\xi\xi}u$ in the sense of distributions and which satisfies (20.23). □

In the space of $n \times n$ matrices we will consider the operator norm $|\cdot|_{\mathcal{L}}$ and, in the subspace of symmetric matrices, the norm $\|\cdot\|$ provided by the largest modulus of the eigenvalues in the spectrum $\sigma(M)$. Obviously these two norms coincide on $\mathrm{Sym}^{n\times n}$. From (20.21) we get

$$\|\nabla^2 u(x_0)\| \le \theta(x_0, B_\varepsilon(x_0), u) \qquad \text{for all } \varepsilon > 0 \tag{20.24}$$

at any point x_0 where u has a second order Taylor expansion.

Corollary 20.29. *If $\Omega \subset \mathbb{R}^n$ is convex and $\theta_\varepsilon \in L^\infty(\Omega)$ for some $\varepsilon > 0$, then*

$$\mathrm{Lip}(\nabla u, \Omega) \le \|\theta_\varepsilon\|_{L^\infty(\Omega)}\ .$$

Proof. The previous proposition shows that $u \in W^{2,\infty}(\Omega)$ and (20.24) provides a pointwise control on $\nabla^2 u$ (recall that semiconvex/semiconcave functions have a second order Taylor expansion a.e.). We recall that since Ω is convex and v is scalar we have $\|\nabla v\|_{L^\infty(\Omega)} = \mathrm{Lip}(v, \Omega)$ (while, in general, $\|\nabla v\|_{L^\infty(\Omega)} \le \mathrm{Lip}(v, \Omega)$). If v takes values in $\mathbb{R}^n$ (in our case $v = \nabla u : \Omega \to \mathbb{R}^n$), then, by the same smoothing argument used in the scalar case, we can always show that

$$\||\nabla v|_{\mathcal{L}}\|_{L^\infty(\Omega)} = \mathrm{Lip}(v, \Omega) \tag{20.25}$$

because, when v is continuously differentiable, there holds

$$\Big| v(x) - v(y) \Big| = \left| \int_0^1 Dv((1-t)x + ty)(x-y)\, dt \right| \leq |x-y| \int_0^1 |\nabla v|_{\mathcal{L}}((1-t)x + ty)\, dt \; .$$

Therefore from (20.24) and (20.25) we conclude. □

At this point our aim is the study of a nonlinear PDE as

$$-F(\nabla^2 u(x)) + f(x) = 0 \tag{20.26}$$

with F non-decreasing on $\mathrm{Sym}^{n\times n}$ (the trace, corresponding to the Laplacian, for example).

Definition 20.30 (Ellipticity)**.** *In the problem* (20.26) *we have ellipticity with constants* $\Lambda \geq \lambda > 0$ *if*

$$\lambda \|N\| \leq F(M+N) - F(M) \leq \Lambda \|N\| \qquad \forall\, N \geq 0 \; . \tag{20.27}$$

Remark 20.31. Every symmetric matrix N admits a unique decomposition as a sum

$$N = N^+ - N^- \; ,$$

with N^+, $N^- \geq 0$ 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 $N^- = \sum_{\rho_i \leq 0} \rho_i e_i \otimes e_i$. Observing this, we are able to write the definition of elliptic problem replacing (20.27) with

$$F(M+N) - F(M) \leq \Lambda \|N^+\| - \lambda \|N^-\| \qquad \forall\, N \in \mathrm{Sym}^{n\times n} \; . \tag{20.28}$$

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 20.32. Consider the case

$$F(M) = \mathrm{tr}(BM)$$

where $B = (b_{ij})_{i,j=1,\dots,n}$ belongs to the set

$$\mathcal{A}_{\lambda,\Lambda} := \left\{ B \in \mathrm{Sym}^{n\times n} | \lambda I \leq B \leq \Lambda I \right\} \; .$$

Fix the symmetric matrix $N \geq 0$. To verify (20.27), we choose the coordinate system in which $N = \mathrm{diag}(\rho_1, \dots, \rho_n)$, thus (since $b_{ii} \geq \lambda$ and $\rho_i \geq 0$ for all $i = 1, \dots, n$)

$$F(M+N) - F(M) = \mathrm{tr}(BN) = \sum_{i=1}^n b_{ii}\rho_i \geq \lambda \sum_{i=1}^n \rho_i \geq \lambda \rho_{\max} \; .$$

Analogously, since $b_{ii} \le \Lambda$ one has

$$F(M+N) - F(M) = \mathrm{tr}(BN) = \sum_{i=1}^{n} b_{ii}\rho_i \le \Lambda \sum_{i=1}^{n} \rho_i \le n\Lambda\rho_{\max} \ .$$

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 the XIX Hilbert problem, the regularity will be deduced only from inequalities derived from ellipticity, without a specific attention to the original equation.

Definition 20.33 (Pucci's extremal operators)**.** *Given ellipticity constants $\Lambda \ge \lambda > 0$ and a symmetric matrix 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 20.34. Resuming Example 20.32, we can show that

$$\mathcal{M}^{-}_{\lambda,\Lambda}(M) = \inf_{B\in\mathcal{A}_{\lambda,\Lambda}} \mathrm{tr}(BM) \tag{20.29}$$

$$\mathcal{M}^{+}_{\lambda,\Lambda}(M) = \sup_{B\in\mathcal{A}_{\lambda,\Lambda}} \mathrm{tr}(BM) \ . \tag{20.30}$$

As a matter of fact, 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 $M = \mathrm{diag}(\rho_1, \ldots, \rho_n)$ we get

$$\mathrm{tr}(BM) = \sum_{i=1}^{n} b_{ii}\rho_i \ge \lambda \sum_{\rho_i>0} \rho_i + \Lambda \sum_{\rho_i<0} \rho_i \tag{20.31}$$

and the equality in (20.31) holds if

$$B = \sum_{\rho_i>0} \lambda e_i \otimes e_i + \sum_{\rho_i<0} \Lambda e_i \otimes e_i \ .$$

Remark 20.35. 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) for every M, N it is simple to obtain from (20.29) and (20.30) that

$$\mathcal{M}^+(M)+\mathcal{M}^-(N)\leq\mathcal{M}^+(M+N)\leq\mathcal{M}^+(M)+\mathcal{M}^+(N)$$

and, similarly,

$$\mathcal{M}^-(M)+\mathcal{M}^-(N)\leq\mathcal{M}^-(M+N)\leq\mathcal{M}^-(M)+\mathcal{M}^+(N)\ ;$$

(c) $\mathcal{M}^\pm$ are elliptic (i.e., they satisfy (20.27)) with constants $\lambda, n\Lambda$, because of Example 20.32 and (20.29), (20.30) which represent $\mathcal{M}^\pm$ as an envelope of a family of functionals with ellipticity constants $\lambda, n\Lambda$.

(d) thanks to (20.28), one has

$$\mathcal{M}^-_{\lambda,n\Lambda}(M)\leq F(M)\leq\mathcal{M}^+_{\lambda/n,\Lambda}(M)\qquad\forall M\in\mathrm{Sym}^{n\times n}\tag{20.32}$$

whenever F is elliptic with constants λ, Λ and $F(0)=0$.

Definition 20.36. *With the previous notations, we will denote*

$$\begin{aligned}\mathrm{Sub}_{\lambda,\Lambda}(f)&:=\left\{u:\Omega\to\mathbb{R}\,\middle|\,-\mathcal{M}^+_{\lambda,\Lambda}(\nabla^2u)+f\leq0\ \textit{in}\ \Omega\right\}\\\mathrm{Sup}_{\lambda,\Lambda}(f)&:=\left\{u:\Omega\to\mathbb{R}\,\middle|\,-\mathcal{M}^-_{\lambda,\Lambda}(\nabla^2u)+f\geq0\ \textit{in}\ \Omega\right\}\ .\end{aligned}$$

We also set

$$\mathrm{Sol}_{\lambda,\Lambda}(f):=\mathrm{Sub}_{\lambda/n,\Lambda}(-|f|)\cap\mathrm{Sup}_{\lambda,n\Lambda}(|f|)\ .\tag{20.33}$$

Remark 20.37. Roughly speaking, the classes defined above correspond to De Giorgi's classes $\mathrm{DG}_\pm(\Omega)$, since u being a solution to (20.26) with F having ellipticity constants λ and Λ implies $u\in\mathrm{Sol}_{\lambda,\Lambda}(f)$; thus, if we are able to infer regularity of functions in $\mathrm{Sol}_{\lambda,\Lambda}(f)$ then we can "forget" thanks to Remark 20.35(d) the specific equation.

21 Regularity theory for viscosity solutions

21.1 The Alexandrov-Bakelman-Pucci estimate

Let us recall the notation from the previous section:

$$\begin{aligned}\mathrm{Sub}(f)&:=\left\{u:\Omega\to\mathbb{R}\,\middle|\,-\mathcal{M}^+(\nabla^2u)+f\leq0\text{ in }\Omega\right\}\\\mathrm{Sup}(f)&:=\left\{u:\Omega\to\mathbb{R}\,\middle|\,-\mathcal{M}^-(\nabla^2u)+f\geq0\text{ in }\Omega\right\}\ ,\end{aligned}$$

where $\mathcal{M}^\pm$ are Pucci's extremal operators, and we shall not emphasize from now on the dependence on the ellipticity coefficients λ and Λ. Notice that, since $\mathcal{M}^+\geq\mathcal{M}^-$, the intersection of the two sets can be nonempty.

The estimate we want to prove is named after Alexandrov, Bakelman and Pucci and is therefore called *ABP weak maximum principle*. It plays the role in this regularity theory played by the Caccioppoli inequality in the standard linear elliptic theory.

In the sequel we call "universal" a constant which depends only on the space dimension n and on the ellipticity constants λ, Λ.

Theorem 21.1 (Alexandrov-Bakelman-Pucci weak maximum principle)**.** *Let u be in $\mathrm{Sup}(f) \cap C(\overline{B}_r)$ with $u \geq 0$ on ∂B_r and $f \in C(\overline{B}_r)$. Then*

$$\max_{\overline{B}_r} u^- \leq Cr \left(\int_{\{u=\Gamma_u\}} \left(f^+\right)^n \, dx \right)^{1/n},$$

where C is universal and Γ_u is defined below.

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 21.2 (Definition of Γ_u)**.** *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*

$$\Gamma_u(x) = \sup \left\{ L(x) \,\middle|\, L \text{ affine, } L \leq -u^- \text{ on } \overline{B}_{2r} \right\}.$$

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 the convex analysis part: firstly $-M \leq \Gamma_u \leq 0$, as a consequence $\Gamma_u \in W^{1,\infty}_{\mathrm{loc}}(B_{2r})$ and finally since Γ_u is differentiable a.e. 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 $\overline{B}_r$.

We need some preliminary results, here is the first one.

Theorem 21.3. *Assume $u \in C(\overline{B}_r)$, $u \geq 0$ on ∂B_r and $\Gamma_u \in C^{1,1}(B_r)$. 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 $x_1 \in B_r$ be such that $u^-(x_1) = M$. 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 $\alpha \gg 1$, then the corresponding hyperplane lies below the graph of $-u^-$ and there is a minimum value of α such that this happens, that is $-u^- \geq L_\alpha$ on $\overline{B}_{2r}$. The graph of $-u^-$ will then meet the corresponding hyperplane 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 $-u^- \geq \Gamma_u \geq L_\alpha$ 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 $c = 3\omega_n^{-1/n}$. □

Remark 21.4. The previous theorem implies the ABP estimate, provided we show that

- $\Gamma_u \in C^{1,1}(B_r)$, as a consequence of $u \in \mathrm{Sup}(f)$;
- $\mathscr{L}^n$-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 $\{u = \Gamma_u\}$ (the so-called *contact region*) one has $\det \nabla^2\Gamma_u \leq C(f^+)^n$, with C universal.

Let us now come to the next steps. The next theorem shows that regularity, 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 21.5 (Propagation of regularity)**.** *Let $u \in C(\overline{B}_r)$ and suppose there exist $\varepsilon \in (0, r]$ and $M \geq 0$ such that, for all $x_0 \in \overline{B}_r \cap \{u = \Gamma_u\}$, 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 $\Gamma_u \in C^{1,1}(\overline{B}_r)$ and $\det \nabla^2\Gamma_u = 0$ a.e. on $\{u > \Gamma_u\}$.*

With the notation introduced before, the assumption of Theorem 21.5 means

$$\overline{\theta}(x_0, B_\varepsilon(x_0), \Gamma_u) \leq M \qquad \forall x_0 \in \overline{B}_r \cap \{u = \Gamma_u\} .$$

Since Γ_u is convex, the corresponding quantity $\underline{\theta}$ is null. Recall also that we have already proved that $\overline{\theta}$, $\underline{\theta} \in L^\infty$ implies $u \in C^{1,1}$ in Corollary 20.29.

Theorem 21.6 (Regularity at contact points)**.** *Consider $v \in \mathrm{Sup}(f)$ in B_δ, φ convex in B_δ with $0 \leq \varphi \leq v$ and $v(0) = \varphi(0) = 0$. Then $\varphi(x) \leq C\left(\sup_{B_\delta} f^+\right)|x|^2$ in $B_{\nu\delta}$, where ν and C are universal constants.*

We can get a naive interpretation of this lemma (or, better, of its infinitesimal version as $\delta \downarrow 0$) by this formal argument: $v - \varphi$ having a local minimum at 0 implies, by the assumption $v \in \mathrm{Sup}(f)$ $\mathcal{M}^-(\nabla^2\varphi(0)) \leq f(0)$. Formally, $\mathcal{M}^-(\nabla^2\varphi(0)) \leq \mathcal{M}^-(\nabla^2 v(0)) \leq f(0)$.

Now it is possible to see how these tools allow to prove the ABP estimate.

Proof. [of Theorem 21.1] Pick a point $x_0 \in B_r \cap \{u = \Gamma_u\}$ and let L be a supporting hyperplane for Γ_u at x_0, so that $\Gamma_u \geq L$ and $\Gamma_u(x_0) = L(x_0)$. Recalling Theorem 21.6, define $\varphi := \Gamma_u - L$, $v := -u^- - L$ (and notice that v is a supersolution because $v \in \mathrm{Sup}(f\chi_{B_r})$). Now, $\varphi(x_0) = v(x_0)$ implies, by means of Theorem 21.6,

$$\overline{\theta}(x_0, B_{\nu\delta}(x_0), \varphi) \leq C \sup_{B_\delta(x_0)} f^+ \qquad \forall x_0 \in \overline{B}_r \tag{21.1}$$

with ν and C universal, for all $\delta \in (0, r)$. Hence

$$\overline{\theta}(x_0, B_{\nu\delta}(x_0), \Gamma_u) \leq C \sup_{B_\delta(x_0)} f^+ \, . \tag{21.2}$$

By Theorem 21.5 we get $\Gamma_u \in C^{1,1}$ and $\det \nabla^2\Gamma_u = 0$ a.e. in the non-contact region. Finally, in order to get the desired estimate, we have to show that a.e. in the contact region one has $\det \nabla^2\Gamma_u \leq c(f^+)^n$. But this comes at once by passing to the limit as $\delta \to 0$ in (21.2) at any differentiability point x_0 of Γ_u. In fact, all the eigenvalues of $\nabla^2\Gamma_u(x_0)$ do not exceed $Cf^+(x_0)$ and the conclusion follows. $\square$

Now we prove Theorem 21.6.

Proof. Let $r \in (0, \delta/4)$ and call $\overline{c} := \left(\sup_{\overline{B}_r} \varphi\right)/r^2$. 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). By means of 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 observing that $\varphi(y) - \varphi(ry/|y|) = o(|y - \bar{x}|)$, 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

$$p(x) := \frac{\overline{c}}{8}(x_n + r)^2 - 4\frac{\overline{c}}{\delta^2} r^2 |x'|^2$$

and notice that the following properties hold:

(a) on A_1, $p(x) \le \overline{c}/(2r^2) = \varphi(\bar{x})/2 \le \varphi(x)/2$;

(b) on A_2, $p(x) \le 0 \le \varphi(x)$ (and in particular $p(x) \le v(x)$);

(c) on A_3, $\delta^2/4 = |x'|^2 + x_n^2 \le |x'|^2 + r^2 \le |x'|^2 + \delta^2/16$, which implies $|x'|^2 \ge (3/16)\delta^2$. By means of the last estimate we get $p(x) \le (\overline{c}/2)r^2 - (3/4)\overline{c}r^2 \le 0 \le \varphi$.

Combining (a), (b), (c) above we get $p \le v$ on ∂A. Since $p(0) = \overline{c}r^2/8 > 0 = \varphi(0)$ 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 $p \le v$ on ∂A, the point y is internal to 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)

$$\lambda\frac{\overline{c}}{4} - 8(n-1)\Lambda\overline{c}\frac{r^2}{\delta^2} \le \sup_{B_\delta} f.$$

But now we can fix r such that $8(n-1)\Lambda\overline{c}r^2/\delta^2 \le \lambda\overline{c}/8$ (it is done by taking r so that $8r \le \delta\sqrt{\lambda/((n-1)\Lambda)}$): we have therefore $\overline{c} \le \frac{8}{\lambda}\sup_{B_\delta} f$. The statement then follows with $C = 8/\lambda$ and $\nu := \frac{1}{8}\sqrt{\lambda/((n-1)\Lambda)}$. $\square$

It remains to prove Theorem 21.5.

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 20.27, we are able to obtain $C^{1,1}$ regularity of Γ_u as soon we are able to propagate this estimate also to non-contact points.

Consider now any point $x_0 \in \overline{B}_r \cap \{u > \Gamma_u\}$ and call L a supporting hyperplane for Γ_u at x_0. Notice that $x_0 \in \{-u^- = L\} \subset \{u = \Gamma_u\}$. We claim that:

(a) There exist $n+1$ points $x_1, \ldots, x_{n+1}$ such that $x_0 \in S := \mathrm{co}(x_1, \ldots, x_{n+1})$ (here and in the sequel co stands for convex hull) and, moreover, all such points belong to $\overline{B}_r \cap \{-u^- = L\}$ 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 $x_1 \in \{u = \Gamma_u\} \cap \overline{B_r}$ 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

$$\begin{aligned}\Gamma_u(x_0+h) &\leq t_1\Gamma_u\left(x_1+h/t_1\right)+t_2\Gamma_u(x_2)+\cdots+t_{n+1}\Gamma_u(x_{n+1})\\ &\leq t_1\left[L(x_1)+M\left|\frac{h}{t_1}\right|^2\right]+t_2L(x_2)+\cdots+t_{n+1}L(x_{n+1})\\ &= L(x_0)+M|h|^2/t_1\leq\Gamma_u(x_0)+3nM|h|^2\end{aligned}$$

and this estimate is clearly uniform since we only require $|h/t_1|\leq\varepsilon$, which is implied by $|h|\leq\varepsilon/(3n)$.

Hence, the problem is reduced to prove the two claims above. This is primarily based on a standard result in convex analysis (first proved by Carathéodory for closed sets), which is recalled here for completeness.

Theorem 21.7 (Carathéodory)**.** *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*

$$x=\sum_{i=1}^{n+1}t_ix_i\qquad\text{and}\qquad\sum_{i=1}^{n+1}t_i=1\ .$$

Set then $C':=\left\{x\in\overline{B}_{2r}\,|\,L(x)=-u^-(x)\right\}$ and $C=\mathrm{co}(C')$. We immediately notice that $C'\neq\emptyset$. 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 $(L+\delta L')(x_0)>L(x_0)$ and, at the same time,

$$L+\delta L'\leq-u^-\quad\text{on }\overline{B}_{2r}\ ,$$

which contradicts the maximality of L.

Thanks to Carathéodory's theorem, we can write $x_0=\sum_{i=1}^{n+1}t_ix_i$ 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 $\Gamma_u\equiv0$ 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 $S := \mathrm{co}(x_1, \ldots, x_{n+1})$. The implication $\geq$ 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 $i = 1, \ldots, n$ implies $t_{n+1} > 2/3$ and therefore

$$r \geq |x_0| \geq 2t_{n+1}r - \sum_{i=1}^{n} t_i|x_i| > \frac{4}{3}r - \frac{n}{3n}r = r\ .$$

□

21.2 The Harnack inequality

In this section we shall prove the Harnack inequality for functions in the class $\mathrm{Sol}(f) := \mathrm{Sub}(-|f|) \cap \mathrm{Sup}(|f|)$ where, according to Definition 20.36, the sets $\mathrm{Sup}(|f|)$ and $\mathrm{Sub}(-|f|)$ are defined through Pucci's extremal operators (with fixed ellipticity constants $0 < \lambda \leq \Lambda$):[10], in the sense of viscosity solutions,

$$u \in \mathrm{Sub}(-|f|) \qquad \Longleftrightarrow \qquad -\mathcal{M}^+(\nabla^2 u) - |f| \leq 0\ ; \tag{21.3}$$

$$u \in \mathrm{Sup}(|f|) \qquad \Longleftrightarrow \qquad -\mathcal{M}^-(\nabla^2 u) + |f| \geq 0\ . \tag{21.4}$$

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 21.13 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.

Theorem 21.8. *Consider a function $u : Q_1 \to \mathbb{R}$ with $u \geq 0$ and $u \in \mathrm{Sol}(f) \cap C(Q_1)$. There exists a universal constant C_H such that*

$$\sup_{Q_{1/2}} u \leq C_H \left(\inf_{Q_{1/2}} u + \|f\|_{L^n(Q_1)} \right)\ . \tag{21.5}$$

Let us show how (21.5) leads to the Hölder regularity result for viscosity solutions of the fully nonlinear elliptic PDE

$$-F(\nabla^2 u(x)) + f(x) = 0\ . \tag{21.6}$$

Step 1. As usual, we need to control the oscillation (now on cubes), defined by

$$\omega_r := M_r - m_r$$

[10]Notice that $\mathrm{Sup}(f) \subset \mathrm{Sup}(|f|)$ and $\mathrm{Sub}(f) \subset \mathrm{Sub}(-|f|)$.

with $M_r := \sup_{Q_r} u$ and $m_r := \inf_{Q_r} u$.
With the same notation of Theorem 21.8, there exists a universal constant $\mu \in (0,1)$ such that
$$\omega_{1/2} \leq \mu\, \omega_1 + 2\|f\|_{L^n(Q_1)} \ . \tag{21.7}$$
Indeed, we apply the Harnack inequality (21.5)

- to the function $u - m_1$, so that
$$M_{1/2} - m_1 \leq C_H \left(m_{1/2} - m_1 + \|f\|_{L^n(Q_1)} \right) \ ; \tag{21.8}$$
- to the function $M_1 - u$, so that
$$M_1 - m_{1/2} \leq C_H \left(M_1 - M_{1/2} + \|f\|_{L^n(Q_1)} \right) \ . \tag{21.9}$$

Adding (21.8) and (21.9) we get
$$\omega_1 + \omega_{1/2} \leq C_H \left(\omega_1 - \omega_{1/2} + 2\|f\|_{L^n(Q_1)} \right) \ ,$$
which proves (21.7) because
$$\omega_{1/2} \leq \frac{C_H - 1}{C_H + 1}\, \omega_1 + 2\frac{C_H}{C_H + 1}\, \|f\|_{L^n(Q_1)} < \frac{C_H - 1}{C_H + 1}\, \omega_1 + 2\, \|f\|_{L^n(Q_1)} \ .$$
We spend a line to remark that $\mu = (C_H - 1)/(C_H + 1)$, C_H being the universal constant in (21.5). It is crucial for the decay of the oscillation that $\mu < 1$.
Step 2. Thanks to a rescaling argument (which we will be hugely used also in the proof of the Harnack inequality), we can generalize (21.7). Fix a radius $0 < r \leq 1$ and put
$$u_r(y) := \frac{u(ry)}{r^2}, \qquad f_r(y) = f(ry) \qquad \text{with } y \in Q_1 \ .$$
Notice that (21.7) holds also for u_r (with the corresponding source f_r) because Pucci's operators are positively 1-homogeneous. Moreover, passing to a smaller scale, the L^n-norm improves.
For simplicity we keep the notation ω_r for the oscillation of the function u, we use $\mathrm{osc}(\cdot, Q_r)$ otherwise. We can estimate
$$\begin{aligned} \omega_{r/2} &= r^2 \mathrm{osc}(u_r, Q_{1/2}) \leq \mu r^2 \mathrm{osc}(u_r, Q_1) + 2r^2 \|f_r\|_{L^n(Q_1)} \\ &= \mu \omega_r + 2r\|f\|_{L^n(Q_r)} \leq \mu \omega_r + 2r \|f\|_{L^n(Q_1)} \ . \end{aligned}$$
Step 3. By the iteration lemmas we used so frequently in the elliptic regularity chapters[11], we are immediately able to conclude that
$$\omega_r \leq C \omega_1 r^{\min\{1,\alpha\}} \quad \forall r \in (0,1] \qquad \text{with } \left(\frac{1}{2}\right)^{\alpha} = \mu \ ,$$

[11]See, for instance, Lemma 9.2.

and with C depending only on μ and $\|f\|_{L^n(Q_1)}$, thus we have Hölder regularity.

In order to prove the Harnack inequality, we will pass through the following reformulation of Theorem 21.8.

Theorem 21.9. *There exist universal positive constants ε_0, C such that if $u : Q_{4\sqrt{n}} \to [0,\infty)$ belongs to $\mathrm{Sol}(f) \cap C(Q_{4\sqrt{n}})$ on $Q_{4\sqrt{n}}$, then*

$$\inf_{Q_{1/4}} u \le 1 \qquad \Longrightarrow \qquad \sup_{Q_{1/4}} u \le C \tag{21.10}$$

provided

$$\|f\|_{L^n(Q_{4\sqrt{n}})} \le \varepsilon_0 \ .$$

Remark 21.10. Theorem 21.8 and Theorem 21.9 are easily seen to be equivalent: since we will prove the second one, it is more important for us to check that Theorem 21.8 follows from Theorem 21.9.
For some positive $\delta > 0$ (needed to avoid a potential division by 0) consider the function

$$v := \frac{u}{\delta + \inf_{Q_{1/4}} u + \|f\|_{L^n(Q_{4\sqrt{n}})}/\varepsilon_0} \ .$$

Denoting by f_v the source term associated with v, the homogenity of Pucci's operators gives $\|f_v\|_{L^n(Q_{4\sqrt{n}})} \le \varepsilon_0$. Since $\inf_{Q_{1/4}} v \le 1$ we have $\sup_{Q_{1/4}} v \le C$, hence

$$\sup_{Q_{1/4}} u \le C \left(\inf_{Q_{1/4}} u + \delta + \|f\|_{L^n(Q_{4\sqrt{n}})}/\varepsilon_0 \right) \ .$$

We let $\delta \to 0$ and we obtain Harnack inequality with the cubes $Q_{1/4}$, $Q_{4\sqrt{n}}$; by the same scaling argument we already used, this means

$$\sup_{Q_r(x_0)} u \le C \left(\inf_{Q_r(x_0)} u + r\|f\|_{L^n(Q_{16r\sqrt{n}}(x_0))} \right) \ . \tag{21.11}$$

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 $x \in Q_{1/2}$, $y \in Q_1$ 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 (21.11) we get (21.5) with $C_H \sim C^N$.

We describe the strategy of the proof of Theorem 21.9, even if the full proof will be completed at the end of this section.
We will 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 $\mathscr{L}^n(\{u > t\} \cap Q_1) \le dt^{-\varepsilon}$, thanks to the fact that $u \in \mathrm{Sup}(|f|)$ (see Lemma 21.13),
- the full thesis of Theorem 21.9 using the fact that $u \in \mathrm{Sol}(f) \subset \mathrm{Sub}(-|f|)$, too.

The first goal will be achieved using the Alexandrov-Bakelman-Pucci inequality of the previous section. The structure of the proof remembers that of De Giorgi's regularity theorem, as we said, and we will complete it through the following lemmas and remarks.

The first lemma is a particular case of Calderón-Zygmund decomposition.

Lemma 21.11 (Dyadic Lemma). *Consider Borel sets $A \subset B \subset Q_1$ with $\mathscr{L}^n(A) \le \delta < 1$. If the implication*

$$\mathscr{L}^n(A \cap Q) > \delta \mathscr{L}^n(Q) \implies \tilde{Q} \subset B \ , \tag{21.12}$$

holds for any dyadic cube $Q \subset Q_1$, with $\tilde{Q}$ being the predecessor of Q, then

$$\mathscr{L}^n(A) \le \delta \mathscr{L}^n(B) \ .$$

Proof. We apply the construction of Calderón-Zygmund (seen in the proof of Theorem 14.1) to $f = \chi_A$ with $\alpha = \delta$: there exists a countable family of cubes $\{Q_i\}_{i \in I}$, pairwise disjoint, such that

$$\chi_A \le \delta \qquad \mathscr{L}^n\text{-a.e. on } Q_1 \setminus \bigcup_{i \in I} Q_i \tag{21.13}$$

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, (21.13) 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 (21.12) we get $\tilde{Q}_i \subset B$ for all i and

$$\mathscr{L}^n(A \cap \tilde{Q}_i) \le \delta \mathscr{L}^n(\tilde{Q}_i) \qquad \forall\, i \in I \ . \tag{21.14}$$

This is due to the fact that a cube Q, in the Calderón-Zygmund construction, is divided in subcubes as long as $\mathscr{L}^n(A \cap Q) \le \delta \mathscr{L}^n(Q)$. Thus (note that we sum on $\tilde{Q}_i$ rather than on i, because different cubes might have the same predecessor)

$$\mathscr{L}^n(A) \le \sum_{\tilde{Q}_i} \mathscr{L}^n(A \cap \tilde{Q}_i) \le \sum_{\tilde{Q}_i} \delta \mathscr{L}^n(\tilde{Q}_i) \le \delta \mathscr{L}^n(B) \ .$$

□

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 Theorem.

Lemma 21.12 (Truncation Lemma). *There exists a universal function $\varphi \in C^\infty(\mathbb{R}^n)$ 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) $\mathcal{M}^+(\nabla^2\varphi) \leq C_\varphi \chi_{Q_1}$ *on* $\mathbb{R}^n$*.*

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 $Q_3 \setminus B_{1/2}$.

After choosing a smooth extension for φ on $B_{1/2}$, still less than -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 . We compute

$$\nabla^2\left(|x|^{-\alpha}\right) = -\frac{\alpha}{|x|^{\alpha+2}}I + \alpha(\alpha+2)\frac{x \otimes x}{|x|^{\alpha+4}} \ ,$$

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 $|x| \geq 1/4$ 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 $\alpha = \alpha(n, \lambda, \Lambda) \gg 1$. Since $B_{1/2} \subset Q_1$ and φ is smooth, we conclude that (iii) holds for a suitable constant C. □

Lemma 21.13 (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* $\|f\|_{L^n(Q_{4\sqrt{n}})} \leq \varepsilon_0$*, then for every integer* $k \geq 1$

$$\mathscr{L}^n\left(\{u > M^k\} \cap Q_1\right) \leq (1-\mu)^k \ . \tag{21.15}$$

Proof. We prove the first step, that is

$$\mathscr{L}^n(\{u > M\} \cap Q_1) \leq (1 - \mu) \,, \tag{21.16}$$

with $M := \max \varphi^-$, φ given by Lemma 21.12, and μ and ε_0 are respectively given by

$$\mu := (2C_{ABP}C_\varphi)^{-n}, \qquad \varepsilon_0 = \frac{1}{2C_{ABP}} \,, \tag{21.17}$$

where C_{ABP} is the universal constant of the Alexandrov-Bakelman-Pucci estimate of Theorem 21.1. Since u is nonnegative, 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

(i)
$$w \geq 0 \quad \text{on } \partial B_{2\sqrt{n}} \tag{21.18}$$
because $u \geq 0$ on $Q_{4\sqrt{n}} \supset B_{2\sqrt{n}}$ and $\varphi \geq 0$ on $\mathbb{R}^n \setminus B_{2\sqrt{n}}$;

(ii)
$$\inf_{B_{2\sqrt{n}}} w \leq \inf_{Q_3} w \leq -1 \tag{21.19}$$
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 $\inf_{Q_3} u \leq 1$;

(iii) directly from the definition of $\mathrm{Sup}(|f|)$ we get $-\mathcal{M}^-(\nabla^2 u) + |f| \geq 0$, moreover $\mathcal{M}^+(\nabla^2\varphi) \leq C_\varphi \chi_{Q_1}$. Since in general $\mathcal{M}^-(A + B) \leq \mathcal{M}^-(A) + \mathcal{M}^+(B)$ (see Remark 20.35), then
$$-\mathcal{M}^-(\nabla^2 w) + (|f| + C_\varphi\chi_{Q_1}) \geq (-\mathcal{M}^-(\nabla^2 u) + |f|) + (-\mathcal{M}^+(\nabla^2\varphi) + C_\varphi\chi_{Q_1}) \geq 0 \,. \tag{21.20}$$
The inequality (21.20) means that $w \in \mathrm{Sup}(|f| + C_\varphi\chi_{Q_1})$.

Thanks to the ABP estimate (which we can apply to w thanks to (21.18) and (21.20)) we get

$$\max_{x \in \overline{B}_{2\sqrt{n}}} w^-(x) \leq C_{ABP} \left(\int_{\{w = \Gamma_w\}} \left(|f(y)| + C_\varphi \chi_{\overline{Q}_1}(y) \right)^n dy \right)^{1/n} . \tag{21.21}$$

Now, remembering that (21.19) holds and that, by definition, $\{w = \Gamma_w\} \subset \{w \leq 0\}$, we can expand (21.21) with

$$\begin{aligned}
1 \;\leq\; & \max_{x \in \overline{B}_{2\sqrt{n}}} w^-(x) \leq C_{ABP} \left(\int_{\{w \leq 0\}} \left(|f| + C_\varphi \chi_{\overline{Q}_1} \right)^n dy \right)^{1/n} && (21.22)\\
\leq\; & C_{ABP} \|f\|_{L^n(Q_{4\sqrt{n}})} + C_{ABP} C_\varphi \mathscr{L}^n \left(Q_1 \cap \{w \leq 0\} \right)^{1/n} && (21.23)\\
\leq\; & C_{ABP} \|f\|_{L^n(Q_{4\sqrt{n}})} + C_{ABP} C_\varphi \mathscr{L}^n \left(Q_1 \cap \{u \leq M\} \right)^{1/n} \,, && (21.24)
\end{aligned}$$

where we pass from line (21.22) to line (21.23) by Minkowski inequality and from line (21.23) to line (21.24) because, if $w(x) \leq 0$, then $u(x) \leq -\varphi(x)$ and then $u(x) \leq M$.
Using our choice of ε_0 we obtain from (21.24) the lower bound

$$\mathscr{L}^n \left(Q_1 \cap \{u \leq M\}\right)^{1/n} \geq \frac{1}{2C_{ABP}C_\varphi} \ . \tag{21.25}$$

Thus, if μ is given by (21.17), we obtain (21.16).
We prove the inductive step: suppose that (21.15) holds for every $j \leq k-1$. We exploit the dyadic Lemma 21.11 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 (21.12) 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 (21.12), suppose by contradiction that for some dyadic cube $Q \subset Q_1$ we have that

$$\mathscr{L}^n(A \cap Q) > \delta \mathscr{L}^n(Q) \tag{21.26}$$

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 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 intention of this definition of $\tilde{f}$ is to ensure that $\tilde{u} \in \mathrm{Sup}(|\tilde{f}|)$, in fact

$$-\mathcal{M}^-(\nabla^2 \tilde{u}) + |\tilde{f}| = \frac{1}{2^{2i}M^{k-1}} \left(-\mathcal{M}^-(\nabla^2 u) + |f|\right) \geq 0 \ .$$

Since the point corresponding to z belongs to Q_3, we ge

$$\inf_{y \in Q_3} \tilde{u}(y) \leq \frac{u(z)}{M^{k-1}} \leq 1 \ .$$

If $\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}})} \leq \varepsilon_0$, then, applying what we already saw in (21.25) 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,

$$\mathscr{L}^n \left(\{u > M^k\} \cap Q\right) \leq (1-\mu)\mathscr{L}^n(Q) \ ,$$

which contradicts (21.26).
In order to complete our proof, we show that effectively $\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}})} \leq \varepsilon_0$. In general, let us remark that the rescaling technique does not cause any problem at the level of the source term f. Indeed

$$\|\tilde{f}\|_{L^n(Q_{4\sqrt{n}})} = \frac{1}{M^{k-1}2^i}\|f\|_{L^n(Q_{4\sqrt{n}/2^i}(x_0))} \leq \|f\|_{L^n(Q_{4\sqrt{n}})}\varepsilon_0 \ .$$

□

Corollary 21.14. *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 $\|f\|_{L^n(Q_{4\sqrt{n}})} \leq \varepsilon_0$, then*

$$\mathscr{L}^n\left(\{u > t\} \cap Q_1\right) \leq dt^{-\varepsilon} \qquad \forall\, t > 0 \ . \tag{21.27}$$

Proof. This corollary is obtained by Lemma 21.13 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 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 (21.31), 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 21.15. *Suppose that $u \in \mathrm{Sub}(-|f|)$ is nonnegative on $Q_{4\sqrt{n}}$ and $\|f\|_{L^n(Q_{4\sqrt{n}})} \leq \varepsilon_0$, with ε_0 given by the decay Lemma 21.13. Assume that (21.27) holds. Then there exist universal constants $M_0 > 1$ and $\sigma > 0$ such that if*

$$x_0 \in Q_{1/2} \quad \text{and} \quad u(x_0) \geq M_0\nu^{j-1} \text{ for some } j \geq 1 \ ,$$

then

$$\exists\, x_1 \in \overline{Q}_{l_j}(x_0) \quad \text{such that} \quad u(x_1) \geq M_0\nu^j \ ,$$

where $\nu := 2M_0/(2M_0-1) > 1$ and $l_j := \sigma M_0^{-\varepsilon/n}\nu^{-\varepsilon j/n}$.

Proof. First of all, we fix a large universal constant $\sigma > 0$ such that

$$\frac{1}{2}\Big(\frac{\sigma}{4\sqrt{n}}\Big)^n > d2^{\varepsilon} \tag{21.28}$$

and then we choose another universal constant M_0 so large that

$$dM_0^{-\varepsilon} < \frac{1}{2} \tag{21.29}$$

and

$$\sigma M_0^{-\varepsilon/n} < 2\sqrt{n}\ . \tag{21.30}$$

We first estimate the superlevels

$$\begin{aligned}
\mathscr{L}^n\left(\{u \geq \nu^j M_0/2\} \cap Q_{l_j/(4\sqrt{n})}(x_0)\right) &\leq \mathscr{L}^n\left(\{u \geq \nu^j M_0/2\} \cap Q_1\right) \\
&\leq d\left(\nu^j M_0/2\right)^{-\varepsilon} < \frac{1}{2}\left(\frac{\sigma}{4\sqrt{n}}\right)^n \nu^{-j\varepsilon} M_0^{-\varepsilon} \\
&= \frac{1}{2}\left(\frac{l_j}{4\sqrt{n}}\right)^n\ ,
\end{aligned} \tag{21.31}$$

where we used condition (21.28) 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

$$\max_{\overline{Q}_{l_j}(x_0)} u < M_0 \nu^j\ . \tag{21.32}$$

Under this assumption, we claim that the superlevel can be estimated as follows:

$$\mathscr{L}^n\left(\{u < \nu^j M_0/2\} \cap Q_{l_j/(4\sqrt{n})}(x_0)\right) < \frac{1}{2}\mathscr{L}^n\left(Q_{l_j/(4\sqrt{n})}\right)\ . \tag{21.33}$$

Obviously the validity of (21.31) and (21.33) is the contradiction that will conclude the proof, so we need only to show (21.33).

Define the auxiliary function

$$v(y) := \frac{\nu M_0 - u(x)\nu^{-(j-1)}}{(\nu-1)M_0} = 2\big(M_0 - \frac{u(x)}{\nu^j}\big)\ ,$$

where $x = x_0 + \frac{l_j}{4\sqrt{n}}y$ and the second equality is a consequence of the relation $M_0 = \nu/[2(\nu-1)]$. Since $y \in Q_{4\sqrt{n}} \iff x \in Q_{l_j}(x_0)$, by (21.32) 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 $l_j/(4\sqrt{n})$, we get

$$f_v(y) = \frac{2l_j^2}{\nu^j} f(x)$$

so that

$$\|f_v\|_{L^n(Q_{4\sqrt{n}})} = \frac{2l_j}{4\sqrt{n}\nu^j} \|f\|_{L^n(Q_{l_j}(x_0))} \leq \varepsilon_0 \tag{21.34}$$

because

$$\frac{2l_j}{4\sqrt{n}\nu^j} = \frac{\sigma M_0^{-\varepsilon/n}}{2\sqrt{n}} \nu^{-\varepsilon j/n - j} < 1$$

thanks to (21.30). The estimate in (21.34) allows us to use Corollary 21.14 for v, that is

$$\mathscr{L}^n \left(\{v > M_0\} \cap Q_1 \right) \leq d M_0^{-\varepsilon} \ ,$$

and we can use this, together with (21.29), to obtain that (21.33) holds:

$$\mathscr{L}^n \left(\{u < \nu^j M_0/2\} \cap Q_{l_j/(4\sqrt{n})}(x_0) \right) \leq d M_0^{-\varepsilon} \mathscr{L}^n \left(Q_{l_j/(4\sqrt{n})} \right) < \frac{1}{2} \mathscr{L}^n \left(Q_{l_j/(4\sqrt{n})} \right) \ .$$

□

We can now complete the proof of Theorem 21.9, using Lemma 21.15. Notice that in Theorem 21.9 we made all assumptions needed to apply Lemma 21.15, taking also Corollary 21.14 into account, which ensures the validity of (21.27).
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 21.15, we should be able to find recursively a sequence (x_j) with the property that

$$u(x_j) \geq M_0 \nu^j \quad \text{and} \quad x_{j+1} \in Q_{l_j}(x_j);$$

since $\sum_j l_j < \infty$, the sequence (x_j) admits a converging subsequence, and in the limit point we find a contradiction. However, in order to iterate Lemma 21.15 we have to confine the 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 cubes).
To achieve this, we fix a universal positive integer j_0 such that $\sum_{j \geq j_0} l_j < 1/4$ 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 21.15 is contained in $Q_{1/2}$ and

$$u(x_k) \geq M_0 \nu^{j_0 + k - 1} \ . \tag{21.35}$$

When $k \to \infty$ in (21.35) we obtain the contradiction. This way, we obtained also an "explicit" expression of the universal constant in (21.10), in fact we proved that

$$\sup_{x \in Q_{1/4}} u(x) \leq M_0 \nu^{j_0 - 1} \ .$$

References

[1] R.Adams: *Sobolev spaces.* Academic Press, 1975.

[2] G.Alberti, L.Ambrosio: *A geometric approach to monotone functions in* $\mathbb{R}^n$. Math. Z, **230** (1999), 259–316.

[3] L.Ambrosio, N.Fusco, D.Pallara: *Functions of bounded variation and free discontinuity problems.* Oxford University Press, 2000.

[4] H.Brezis: *Analyse Fonctionelle: Théorie et applications.* Masson, Paris, 1983.

[5] L.A.Caffarelli, X.Cabré: *Fully nonlinear elliptic equations.* Colloquium Publications, **43** (1995), American Mathematical Society.

[6] E.De Giorgi: *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.

[7] E.De Giorgi: *Frontiere orientate di misura minima.* Seminario di Matematica della Scuola Normale Superiore di Pisa, (1960-61), Editrice Tecnico Scientifica, Pisa.

[8] E.De Giorgi: *Un esempio di estremali discontinue per un problema variazionale di tipo ellittico.* Boll. Un. Mat. Ital. (4), **1** (1968), 135–137.

[9] E.De Giorgi: *Sulla differenziabilità e l'analicità degli estremali degli integrali multipli regolari.* Mem. Acc. Sc. Torino, 3 (1957), 25–43.

[10] L.C.Evans: *Quasiconvexity and partial regularity in the calculus of variations.* Arch. Rational Mech. Anal. 95, **3** (1986), 227–252.

[11] L.C.Evans, R.F.Gariepy: *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, 1992.

[12] H.Federer: *Geometric Measure Theory.* Die Grundlehren der mathematischen Wissenschaft, Band 153, Springer-Verlag New York Inc., 1969.

[13] E.Gagliardo: *Caratterizzazione delle tracce sulla frontiera relative ad alcune class di funzioni in piú variabili.* Rend. Sem. Mat. Univ. Padova, **27** (1957), 284–305.

[14] D.Gilbarg, N.S. Trudinger: *Elliptic Partial Differential Equations of Second Order.* Springer Verlag, 1983.

[15] E.Giusti: *Minimal surfaces and functions of bounded variation.* Birhkhäuser, Boston, 1994.

[16] M.Giaquinta, E.Giusti: *On the regularity of the minima of variational integrals.* Acta Math. 148, (1982), 31–46.

[17] M.Giaquinta, E.Giusti: *Quasiminima.* Ann. Inst. H. Poincaré Anal. Non Linéaire 1, **2** (1984), 79–107.

[18] M.Giaquinta, E.Giusti: *The singular set of the minima of certain quadratic functionals.* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 11, **1** (1984), 45–55.

[19] M.Giaquinta: *Multiple integrals in the Calculus of Variations and Nonlinear elliptic systems.* Princeton University Press, 1983.

[20] E.Hopf: *Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung.* Math. Zeitschrift, Band 34 (1932), 194–233.

[21] F.John, L.Nirenberg: *On Functions of Bounded Mean Oscillation.* Comm. on Pure and Applied Math., Vol. XIV (1961), 415–426.

[22] J.Kristensen, G.Mingione: *The singular set of minima of integral functionals.* Arch. Ration. Mech. Anal., **180** (2006), 331–398.

[23] *The singular set of lipschitzian minima of multiple integrals.* Arch. Ration. Mech. Anal. **184** (2007), 341–369.

[24] N.G.Meyers, J.Serrin: $H = W$. Proc. Nat. Acad. Sci. U.S.A., **51** (1964), 1055–1056.

[25] S.Müller, V.Sverak: *Convex integration for Lipschitz mappings and counterexamples to regularity.* Ann. of Math., **157** (2003), 715–742.

[26] E.M.Stein, G.Weiss: *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, 1971.

[27] N.Trudinger: *On embedding into Orlicz spaces and some applications.* J. Math. Mech., **17** (1967), 473–483.

[28] K.Yosida: *Functional Analysis.* Mathematical surveys and monographs, **62**, American Mathematical Society, 1998.