# Lecture notes on Calculus of Variations

André Guerra

*ETH Zürich*
*andre.guerra@eth-its.ethz.ch*

January 31, 2025

# Table of Contents

# Preface

These notes are heavily inspired by the books [1, 3, 6].

The notes contain both *Exercises* and *Problems*; both are an integral part of the text. The exercises are contained in the bulk of the text and they consist of routine questions that should help the reader digest a definition, concept or proof; they are meant to be solved without too much effort. Problems, on the contrary, are displayed at the end of each section and are meant to be more challenging. They range from smaller questions that consist in generalizing a result covered in the text, to introducing other results or parts of the theory that were not covered due to time constraints.

# 0 Introduction

In these notes we will study one of the most basic problems in the Calculus of Variations: we consider *minimizers* of integral energies such as

$$\mathbb{E}[u] \equiv \int_\Omega F(x, u, \mathrm{D}u)\,\mathrm{d}x, \tag{0.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded smooth domain, $u\colon \Omega \to \mathbb{R}^m$, and $F\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is assumed to be smooth. At this point, it is not entirely clear what we mean by minimizer and, in general, this depends on the context. In these notes our focus is on *unconstrained* problems, and we will simply minimize $\mathbb{E}$ over those smooth maps $u\colon \overline{\Omega} \to \mathbb{R}^m$ such that

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

for a regular datum $g\colon \partial\Omega \to \mathbb{R}^m$.

## 0.1 The Euler–Lagrange system

It is worth noting that, to each minimization problem as above, one can associate a Partial Differential Equation (PDE): if $u$ is a minimizer, at least formally we must have

$$\mathbb{E}'[u](\varphi) \equiv \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathbb{E}[u + t\varphi] = 0 \qquad \forall \varphi \in C^\infty_c(\Omega, \mathbb{R}^m),$$

i.e. $u$ is a *critical point* of $\mathbb{E}$. Throughout these notes, we will use Greek letters $\alpha, \beta, \gamma, \dots$ to represent coordinates in the domain $\Omega \subset \mathbb{R}^n$, and Roman letters $i, j, k, \dots$ to represent coordinates in the target $\mathbb{R}^m$. A simple calculation then shows that

$$\mathbb{E}'[u](\varphi) = \int_\Omega \partial_{\xi^i_\alpha} F(x, u, \mathrm{D}u)\partial_\alpha \varphi^i + \partial_{u^i} F(x, u, \mathrm{D}u)\varphi^i\,\mathrm{d}x, \tag{0.2}$$

or, after integrating by parts (since $\varphi$ has compact support),

$$\mathbb{E}'[u](\varphi) = \int_\Omega \Big[-\partial_\alpha(\partial_{\xi^i_\alpha} F(x, u, \mathrm{D}u)) + \partial_{u^i} F(x, u, \mathrm{D}u)\Big]\,\varphi^i\,\mathrm{d}x.$$

Since $\varphi \in C^\infty_c(\Omega, \mathbb{R}^m)$ is arbitrary, we conclude that the term in square brackets must vanish identically, i.e.

$$\partial_\alpha(\partial_{\xi^i_\alpha} F(x, u, \mathrm{D}u)) = \partial_{u^i} F(x, u, \mathrm{D}u) \qquad i = 1, \dots, m. \tag{0.3}$$

This is known as the *Euler–Lagrange system*: it is a coupled system of $m$ quasilinear second order PDEs. Although the above derivation is formal, we will see later that it can be made precise under suitable growth assumptions on $F$.

The analysis of (0.3) is a daunting task, but there are two cases in which it becomes simpler: when $m = 1$, we have a *single* PDE; and when $n = 1$, we have a system of *ordinary differential equations*, which can be approached by much more elementary methods.

It is worth noting that, in general, (0.3) is just a *necessary condition* for $u$ to be a minimizer: there are solutions of (0.3) which do not minimize $\mathbb{E}$. However, whenever $F$ is convex in the last variables, then this necessary condition becomes sufficient, as we will see in Section 2. In fact, an important question at this stage is what kind of assumptions should we make on $F$ in order to hope for a reasonable theory. In order to have a feeling for what the answer could be, we now look at some classical examples.

## 0.2 Examples of variational problems

**Example 0.1** (Mechanical systems)**.** Consider a mechanical system composed of $N$ particles, with respective masses $m_i$, and whose positions at time $t$ are given by a vector $u_i(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$. The typical energy for such a system is

$$F(t, u(t), u'(t)) = \frac{1}{2}\sum_{i=1}^{N} m_i |u_i'(t)|^2 - U(t, u(t)),$$

where the first term is the *kinetic energy* of the system and the second term is the *potential energy*, whose specific form depends on the system under consideration. In this case we have $n = 1$ and $m = 3N$, and the Euler–Lagrange system is simply

$$m_i u_i''(t) = -\partial_{u_i} U(t, u(t)), \qquad i = 1, \dots, N.$$

These are *Newton's Laws of Motion.*

**Example 0.2** (Fermat's principle)**.** In 1662, Fermat proposed the following principle: the path taken by a ray of light between two given points is the path that can be traveled in the least amount of time. If the medium is isotropic and we write $u(t) = (x(t), y(t), z(t))$, then Fermat's principle asserts that the trajectories of light are minimizers for the *optical Lagrangian*

$$F(u(t), u'(t)) = r(u(t))|\dot{u}(t)|.$$

Here $r$ is the *refractive index*, i.e. the ratio between the speed of light in vacuum and in this inhomogeneous medium. In this case we have $n = m = 1$, in the above notation.

**Example 0.3** (Minimal surfaces)**.** The Plateau problem is a famous problem in the Calculus of Variations: find, among all $n$-dimensional surfaces in $\Sigma \subset \mathbb{R}^{n+m}$ with prescribed boundary, the one which has least area. This question, first raised by Lagrange 1760, can be studied experimentally by dipping a wire into soapy water: the resulting soap film is a *minimal surface.* The precise formulation of Plateau's problem depends on what we mean by "surface" and "area".

The simplest possible setting of Plateau's problem is when we only consider graphical $n$-surfaces:

$$\Sigma = \operatorname{graph}(u) \equiv \{(x, u(x)) \in \mathbb{R}^{n+m} : x \in \overline{\Omega}\},$$

where $u \colon \overline{\Omega} \to \mathbb{R}^m$. Here $m$ plays the role of *codimension.* In this case, the fact that the boundary of $\Sigma$ is fixed means simply that $u = g$ on $\partial\Omega$, for some fixed function $g \colon \partial\Omega \to \mathbb{R}$, and if $m = 1$ then the area is simply

$$\operatorname{Area}(\Sigma) = \int_\Omega F(\mathrm{D}u)\,\mathrm{d}x, \qquad F(\xi) = \sqrt{1 + |\xi|^2}. \tag{0.4}$$

This is known as the *nonparametric Plateau problem.*

Geometrically speaking, graphical surfaces are too restrictive: although each regular surface is locally a graph, this doesn't need to happen globally. Thus we are led to consider *parametric* surfaces, i.e.

$$\Sigma \equiv v(\overline{\Omega}), \qquad v \colon \overline{\Omega} \to \mathbb{R}^{n+m},$$

and the area of $\Sigma$ is defined in terms of $v$. The simplest case is clearly when $n = 1$: then we are simply considering simply curves, and we can write

$$\text{Length}\,(v([0,1])) = \int_0^1 |\dot v(t)|\,\mathrm{d}t, \qquad v(0)=a, v(1)=b \tag{0.5}$$

where $a, b \in \mathbb{R}^m$ are the given boundary conditions. This corresponds to the case $r = 1$ in Example 0.2 when $m = 1$.

Both the parametric and the non-parametric versions of Plateau's problem bring with them certain difficulties:

(i) In the non-parametric version, the integrand in (0.4) has *linear growth*, and this makes functional analytic methods harder to apply (in a nutshell, $L^1$ is not a reflexive space, but $L^p$ for $p > 1$ is). Moreover, it may be too restrictive to look for graphical solutions.

(ii) In the parametric version, although the surface may be unique the parametrization is in general non-unique. Indeed, in (0.5) clearly minimizers are obtained if and only if $v([0,1])$ is a straight line between $(0,a)$ and $(1,b)$. However, we cannot conclude that $v(t) = (b-a)t$, since the length is invariant under reparametrizations; in particular, one cannot infer any type of nice regularity properties on $v$. The difficulty here is due to the fact that $F(\xi) = |\xi|$ is not *strictly convex*.

Both of the above difficulties disappear when considering the Dirichlet integral:

**Example 0.4** (The Dirichlet integral)**.** We consider the Dirichlet energy

$$\mathbb{E}[u] \equiv \int_\Omega \frac{1}{2}|\mathrm{D}u|^2\,\mathrm{d}x,$$

for $u\colon \Omega \to \mathbb{R}$. In this case, the Euler–Lagrange equation is simply the Laplace equation

$$\Delta u = 0,$$

thus $u$ is a *harmonic function*. Note that, if $V \subset \Omega$ is a smooth region, then by the Divergence Theorem we have

$$0 = \int_V \Delta u\,\mathrm{d}x = \int_{\partial V} \mathrm{D}u \cdot \nu\,\mathrm{d}\sigma,$$

i.e. the *flux* of $u$ through any region is zero. This means that we can think of $u$ as representing any quantity in equilibrium (e.g. temperature or electrostatic potential).

One can think of the Dirichlet energy as a linearization of (0.4), since

$$\sqrt{1+|\xi|^2} = 1 + \frac{1}{2}|\xi|^2 + O(|\xi|^4) \quad \text{ as } |\xi| \to 0.$$

It is also worth noting that the non-uniqueness phenomena mentioned above are not present for the Dirichlet energy. For instance, if $n = 1$ then $\Delta u = u''$ and so the Euler–Lagrange equation does select the linear solution.

Note that, in the above examples, we always verified the assumption

$$\xi \mapsto F(x,u,\xi) \quad \text{ is convex for all } (x,u) \in \Omega \times \mathbb{R}^m.$$

Thinking about finite-dimensional problems, it is not surprising that convexity appears as a natural condition in our minimization problem. In these notes we will recurrently assume convexity-type conditions on $F$, which then translate into *ellipticity* conditions for the Euler–Lagrange system (0.3).

## 0.3 Hilbert's problems

Let us return to the general energy considered in (0.1). Our main concern in these notes is whether a minimizer of $\mathbb{E}$ exists and, if so, whether it is itself smooth. In fact, this is precisely the content of Hilbert's 19th and 20th problems, in his famous 1900 address at the International Congress of Mathematicians in Paris:

- 20th problem: *Has every regular variational problem a solution, provided certain assumptions regarding the given boundary conditions are satisfied, and provided also if need be that the notion of a solution shall be suitably extended?*

- 19th problem: *Are the solutions of regular problems in the Calculus of Variations always necessarily analytic?*

Hilbert was concerned with the case $n = 2$ (problems in two variables) and $m = 1$ (the *scalar* case). Whenever $m > 1$, we say that we are dealing with a *vectorial* problem. As we will see in these notes, scalar and vectorial problems differ quite dramatically, but we will try to treat them homogeneously as far as possible.

The list of contributions towards Hilbert's problems is enormous, and we refer the reader to any of the references [1, 6] for a more detailed account. These problems were the inspiration for an enormous amount of progress in the Calculus of Variations and Elliptic Partial Differential Equations (PDE) in the last century, and by now their answer is essentially understood in complete generality. We now briefly recall the answer to Hilbert's problems, as a way of outlining the contents of these notes.

Concerning the 20th problem, in Sections 1 and 2 we will take the modern approach and develop a functional-analytic framework to prove existence of *weak* solutions. In the spirit of the 20th problem, the notion of solution has to be suitably extended, and a priori solutions are just in a Sobolev space (i.e. a Banach space of weakly differentiable functions).

The existence of weak solutions brings us naturally to the 19th problem: are the solutions constructed by functional-analytic methods regular? Here the word regular typically means that we would like solutions to be as regular as the data of the problem allow, e.g. if the data is smooth we expect solutions to be smooth. We will not be so concerned with analiticity of the solutions per se, since in any case this ultimate form of regularity only holds in rather special circumstances.

Concerning the regularity of weak solutions, we will first have to develop, in Sections 3 and 4, the regularity theory for *linear* elliptic systems of the form

$$-\partial_\alpha(A_{ij}^{\alpha\beta}\partial_\beta u^j) = \partial_\alpha G_i^\alpha \quad \text{for } i = 1, \dots, m. \tag{0.6}$$

We note that linear elliptic systems arise as the Euler–Lagrange systems of quadratic energies $\mathbb{E}$. In order to have a successful linear theory, one needs in the very least that

$$A_{ij}^{\alpha\beta} \in C^0(\Omega).$$

One can then show that if $A_{ij}^{\alpha\beta}, G_\alpha^i \in C^\infty(\Omega)$ and $A_{ij}^{\alpha\beta}$ satisfy suitable *ellipticity conditions* then $u \in C^\infty(\Omega)$.

The typical approach to the regularity for the general nonlinear problem formulated in (0.3) is by linearization. For simplicity, let us assume that $F$ does not depend on $x$ or $u$, and let us also suppose that it is strongly convex and has bounded Hessian: there are $0 < \lambda < \Lambda$

such that

$$\lambda|\xi|^2 \leq \mathrm{D}^2 F[\xi, \xi] \leq \Lambda|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^{m\times n}, \tag{0.7}$$

Then, at least formally, if we differentiate (0.3), we see that

$$\partial_\gamma \partial_\alpha \left[\partial_{\xi^i_\alpha} F(\mathrm{D}u)\right] = \partial_\alpha \left[\partial_{\xi^i_\alpha} \partial_{\xi^j_\beta} F(\mathrm{D}u)\, \partial_\beta \partial_\gamma u^j\right] = 0.$$

In other words, $\partial_\gamma u^j$ solves a linear system as in (0.6), with coefficients

$$A^{\alpha\beta}_{ij} \equiv \partial_{\xi^i_\alpha} \partial_{\xi^j_\beta} F(\mathrm{D}u).$$

The tensor field $A = (A^{\alpha\beta}_{ij})$ inherits lower and upper bounds from (0.7) but a fundamental difficulty at this point is that $A$ is just *measurable*, since we do not know a priori that $\mathrm{D}u$ is even continuous. Hence the linear theory does not apply!

In the scalar case $m = 1$, the above gap was closed by what is now known as the De Giorgi–Nash theorem, first proved by De Giorgi and then almost simultaneously (and independently) by Nash in the more general parabolic setting between. The theorem asserts that, for a single elliptic equation with measurable coefficients, solutions are Hölder continuous, and so the linear theory can be applied in order to conclude that any minimizer of $\mathbb{E}$ is smooth.

In the vectorial case $m > 1$, the situation is fundamentally different, and around 10 years after his theorem, De Giorgi constructed an example for $n = m \geq 3$ showing that, for general measurable (and discontinuous) coefficients $A$, solutions to (0.6) need not even be bounded. As we will later see, the case $n = 2$ is special and there regularity does hold for other reasons.

## 0.4 Some notation

(i) $\Omega \subset \mathbb{R}^n$ is an open set. We will often assume that it is connected; in that case, we say it is a *domain*. Given another open set $\Omega' \Subset \mathbb{R}^n$, we write $\Omega \Subset \Omega'$ if $\overline{\Omega} \subset \Omega'$ and $\overline{\Omega}$ is compact; we then say that $\Omega$ is *compactly contained* in $\Omega'$.

(ii) $(u)_\Omega \equiv ⨍_\Omega u \,\mathrm{d}x$, where $⨍_\Omega \equiv \frac{1}{|\Omega|}\int_\Omega$. If $\Omega = B_r(x_0)$, then we write $(u)_{B_r(x_0)} \equiv (u)_{x_0,r}$.

(iii) We always write $\lambda$ for the lower bound in either Legendre or Legendre–Hadamard conditions of a tensor $A$, cf. (2.3) and (2.5) respectively. Likewise, we write $\Lambda \equiv \|A\|_{L^\infty}$.

(iv) We always assume that $1 \leq p \leq \infty$.

# 1 Sobolev spaces

In the introduction we encountered the Dirichlet energy as one of the main prototypical examples of energies that we will consider in these notes. To be more precise, we are concerned with the Dirichlet problem for the Poisson equation:

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.1}$$

Our goal is to prove existence of solutions to (1.1) using functional analytical methods. The following is a simple but natural approach:

**Theorem 1.1** (Abstract Dirichlet's principle)**.** *Let $H$ be a Hilbert space, let $\ell \in H^*$ and set*

$$E[u] \equiv \frac{1}{2}\|u\|_H^2 - \ell(u).$$

*Then $E$ attains a unique minimum $\bar{u}$ in $H$, and every minimizing sequence converges to it:*

$$E(u_j) \to \inf_{v \in H} E(v) \qquad \Longrightarrow \qquad u_j \to u \text{ in } H.$$

*The minimum $\bar{u}$ is characterized by*

$$0 = E'[u](v) \equiv \frac{\mathrm{d}}{\mathrm{d}t} E[u + tv]\Big|_{t=0} = \langle u, v\rangle - \ell(v) = 0 \qquad \text{for all } v \in H.$$

We will prove more general versions of Theorem 1.1 in Section 2, but it is instructive to give a quick proof of it here.

**Proof.** First note that $\inf_H E > -\infty$, since

$$E[u] \geq \frac{1}{2}\|u\|^2 - \|\ell\|\|u\| \geq -\frac{1}{2}\|\ell\|^2.$$

Recall the parallelogram identity: for all $u, v \in H$, we have

$$\frac{1}{2}\|u+v\|^2 + \frac{1}{2}\|u-v\|^2 = \|u\|^2 + \|v\|^2$$

which we can rewrite as

$$\frac{1}{4}\|u-v\|^2 = E[u] + E[v] - 2E\left[\frac{u+v}{2}\right],$$

and so we see that there is at most one minimizer. Applying this identity with $u = u_j$ and $v = u_k$, we also see that

$$\frac{1}{4}\|u_j - u_k\|^2 \leq E[u_j] + E[u_k] - 2\inf E.$$

Thus any minimizing sequence is a Cauchy sequence, so by completeness a minimum exists. The characterization of the minimum is part of Problem 1.1. □

In order to apply Theorem 1.1 to solve (1.1), what Hilbert space $H$ should we take? There is really only one option, as we would like to take

$$\|u\|_H^2 = \|\mathrm{D}u\|_{L^2}^2. \tag{1.2}$$

This choice leads us naturally to a certain Hilbert space, known as the *Sobolev space* $W_0^{1,2}(\Omega)$. Functions in this space possess certain weak differentiability properties, but in general can be rather irregular: for instance, in general they are discontinuous. However, we expect solutions of (1.1) to be rather smooth, if $f$ is sufficiently nice: intuitively, they should have two more derivatives than $f$. This expectation turns out to be correct, and it will be made precise in different manners in Sections 3 and 4.

In this section we recall the definitions and some relevant properties of Sobolev spaces; our exposition is minimalistic and most proofs will be omitted. We refer the reader to [3, §5] for a gentle introduction to Sobolev spaces and to [11] for a comprehensive development of their theory.

### 1.1 Weak derivatives

Sobolev spaces are Banach subspaces of $L^p(\Omega)$ whose elements have weak differentiability properties. In this subsection, we quickly recall the definition and some basic properties of weak derivatives. The basic idea is that one can talk about derivatives of very rough functions, if one integrates by parts:

**Definition 1.2** (Weak derivatives)**.** We say that $u \in L^1_{\mathrm{loc}}(\Omega)$ has $\alpha$-weak derivative equal to $g_\alpha \in L^1_{\mathrm{loc}}(\Omega)$ if

$$\int_\Omega u \partial_\alpha \varphi \, \mathrm{d}x = -\int_\Omega g_\alpha \varphi \, \mathrm{d}x \quad \forall \varphi \in C_c^\infty(\Omega).$$

If this holds we then write $\partial_\alpha u \equiv g_\alpha$, and more generally $\mathrm{D}u = (\partial_1 u, \dots, \partial_n u)$ if all weak derivatives exist.

Weak derivatives enjoy most of the formal properties of usual derivatives, such as the Leibniz rule and the fact that $\partial_\alpha \partial_\beta = \partial_\beta \partial_\alpha$. The next is a good exercise for the reader unfamiliar with weak derivatives:

**Exercise 1.3.** Check the following:

(i) the weak derivative of $x \mapsto x 1_{[0,1]} + 1_{(1,2)}$ is $1_{[0,1]}$;

(ii) the function $x \mapsto x 1_{[0,1]} + 2 \times 1_{(1,2)}$ does not have a weak derivative.

In general, a good way of proving properties of weak derivatives is through convolution, as the next two lemmas show.

**Lemma 1.4** (Uniqueness of weak derivatives)**.** *If $g_\alpha, h_\alpha \in L^1_{\mathrm{loc}}(\Omega)$ are $\alpha$-weak derivatives of $u \in L^1_{\mathrm{loc}}(\Omega)$, then $g_\alpha = h_\alpha$ a.e. in $\Omega$.*

**Proof.** The difference $w \equiv g_\alpha - h_\alpha \in L^1_{\mathrm{loc}}(\Omega)$ satisfies

$$\int_\Omega w(x)\varphi(x) \, \mathrm{d}x = 0 \quad \forall \varphi \in C_c^\infty(\Omega).$$

Let $\rho_\varepsilon$ be a mollifier, thus $\rho_\varepsilon \equiv \varepsilon^{-n} \rho(\cdot/\varepsilon)$ where $\rho \in C_c^\infty(B_1(0))$ has $\int_{\mathbb{R}^n} \rho \, \mathrm{d}x = 1$. Fix $x_0 \in \Omega$ and take $\varphi(x) \equiv \rho_\varepsilon(x_0 - x)$ above, with $\varepsilon \leq \operatorname{dist}(x_0, \partial\Omega)$, to get $0 = u * \rho_\varepsilon(x_0)$. Since $u * \rho_\varepsilon \to u$ in $L^1_{\mathrm{loc}}(\Omega)$, we conclude that $u = 0$ a.e. in $\Omega$. $\square$

**Lemma 1.5** (Constancy lemma)**.** *If $\Omega$ is a domain and $u \in L^1_{\mathrm{loc}}(\Omega)$ has $\mathrm{D}u = 0$ in the weak sense then there is $c \in \mathbb{R}$ with $u = c$ a.e. in $\Omega$.*

**Proof.** From the definition of weak derivatives, it is easy to see that if $\partial_\alpha u = g$ in the weak sense then

$$\partial_\alpha(u * \rho_\varepsilon) = g * \rho_\varepsilon \quad \text{ in } \Omega_\varepsilon \equiv \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) > \varepsilon\}.$$

Thus, when $g = 0$, the smooth functions $u * \rho_\varepsilon$ are (locally) constant, and as they converge in $L^1$ their limit $u$ is constant too. $\square$

Weak derivatives appear very naturally from a variational point of view: in (0.2) we saw that, when computing $\mathbb{E}'[u]$ for a general variational integral, one arrives at an expression which has just one derivative on $u$, instead of the two derivatives that one would expect. For instance, when $\mathbb{E}$ is the classical Dirichlet energy from Example 0.4, we have

$$\mathbb{E}'[u](\varphi) = \int_\Omega \mathrm{D}u \cdot \mathrm{D}\varphi \, \mathrm{d}x, \quad \varphi \in C^\infty_c(\Omega);$$

a similar expression appears of course in Theorem 1.1. Thus, even for functions $u \colon \Omega \to \mathbb{R}$ such that $\mathrm{D}u \in L^1_{\mathrm{loc}}(\Omega)$, we can still make sense of the equation $\Delta u = 0$ in the weak sense. This definition is consistent: you can check that, if $u$ is smooth, then $\Delta u = 0$ in the weak sense if and only if $\Delta u = 0$ in the classical sense.

## 1.2 Basic properties of Sobolev spaces

We now turn to the proper definition of Sobolev spaces.

**Definition 1.6.** For $p \in [1, \infty]$, we let $W^{1,p}(\Omega)$ be the set of those $L^p(\Omega)$ functions whose elements are weakly differentiable, with weak derivatives in $L^p(\Omega)$. We equip this space with the norm

$$\|u\|_{W^{1,p}(\Omega)} \equiv \|u\|_{L^p(\Omega)} + \|\mathrm{D}u\|_{L^p(\Omega)}.$$

We denote by $W^{1,p}_0(\Omega)$ the closure of $C^\infty_c(\Omega)$ in $\|\cdot\|_{W^{1,p}(\Omega)}$. We also define higher order spaces $W^{k,p}(\Omega)$ and $W^{k,p}_0(\Omega)$ in a similar way. The definitions also extend to vector-valued maps in an obvious way: for instance, we say that $u \in W^{k,p}(\Omega, \mathbb{R}^m)$ if $u^i \in W^{k,p}(\Omega)$ for all $i = 1, \dots, m$.

The above definitions admits clear *local* variants: for instance, we say that $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$ if, for every $x_0 \in \Omega$, there is $r > 0$ such that $u \in W^{1,p}(B_r(x_0))$. We adopt similar conventions for all other functions spaces in these notes.

Intuitively, $W^{1,p}_0(\Omega)$ consists of those functions in $W^{1,p}$ which vanish on $\partial\Omega$. This intuition can be made precise using *trace theory* which shows that, under appropriate smoothness assumptions on $\partial\Omega$, one can indeed make sense of $u|_{\partial\Omega}$ for $u \in W^{k,p}$. In these notes, however, we will not need or use any trace theory.

Essentially as a consequence of the fact that $L^p(\Omega)$ is a Banach space, we have:

**Theorem 1.7** (Sobolev spaces are function spaces)**.** *For each $k \in \mathbb{N}$ and $p \in [1, \infty]$, the space $W^{k,p}(\Omega)$ is a Banach space.*

The next theorem shows that Sobolev spaces also appear naturally as the closure (in the Sobolev norm) of smooth functions:

**Theorem 1.8** (Density of smooth functions)**.** *Let $\Omega$ be bounded, and suppose that $u \in W^{k,p}(\Omega)$ for some $p \in [1,\infty)$ and some $k \in \mathbb{N}$. There is a sequence $(u_j) \subset C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that $u_j \to u$ in $W^{k,p}(\Omega)$.*

Note that the approximating sequence is *not* in $C^\infty(\overline{\Omega})$: in order to construct such a sequence, we would need further regularity assumptions on $\partial\Omega$.

**Exercise 1.9.** Show that in Theorem 1.8 one cannot take $(u_j) \in C^\infty(\overline{\Omega})$ for any domain $\Omega$. (Hint: take $n = 2$ and consider the function $u(re^{i\theta}) = \theta$ defined away from the negative real axis. Then use Problems 1.2 and 1.3 to show that $u \notin W^{1,p}$.)

In these notes we will avoid discussing issues related to boundary behavior of functions, and therefore we will typically not need to impose any regularity on $\partial\Omega$. It is worth noting, however, that in general functions in $W^{1,p}(\Omega)$ can be extended to functions in $W^{1,p}(\mathbb{R}^n)$ whenever $\Omega$ is sufficiently smooth:

**Theorem 1.10** (Extension theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be a smooth domain and fix any other domain such that $\Omega \Subset \Omega'$. There is a bounded linear operator $E\colon W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ such that, for all $u \in W^{1,p}(\Omega)$, we have*

$$Eu = u \text{ a.e. in } \Omega, \qquad \operatorname{supp}(Eu) \subset \Omega'.$$

We call $Eu$ the *extension* of $u$. Thus Theorem 1.10 allows us to deduce many properties of $W^{1,p}(\Omega)$ from properties of $W^{1,p}(\mathbb{R}^n)$.

We now give a different characterization of Sobolev spaces, which will play quite an important role in Section 3, in the context of Nirenberg's method. The idea is to replace partial derivatives with *difference quotients*.

**Definition 1.11.** For $u\colon \Omega \to \mathbb{R}$ and $\alpha \in \{1, \dots, n\}$ and $h > 0$, we define

$$\partial_{h,\alpha} u(x) \equiv \frac{\tau_{h,\alpha} u(x) - u(x)}{h} \equiv \frac{u(x + he_\alpha) - u(x)}{h}, \quad x \in \Omega_{s,\alpha} \equiv \{x \in \Omega : x + he_\alpha \in \Omega\}.$$

We encourage the reader to verify the following properties:

**Exercise 1.12** (Basic properties of difference quotients)**.** The following are easy to prove:

(i) Leibniz rule: $\partial_{h,\alpha}(uv) = \tau_{h,\alpha} u(\partial_{h,\alpha} v) + (\partial_{h,\alpha} u)v$;

(ii) Integration by parts: for any $\varphi \in C^1_c(\Omega)$ with $|h| < \operatorname{dist}(\operatorname{supp}(\varphi), \partial\Omega)$, we have

$$\int_\Omega \varphi \partial_{h,\alpha} u \,\mathrm{d}x = -\int_\Omega u \partial_{-h,\alpha} \varphi \,\mathrm{d}x.$$

Having the above exercise at our disposal, we can prove the following alternative characterization of Sobolev spaces:

**Proposition 1.13** (Weak derivatives vs difference quotients)**.** *Let $u \in L^p_{\mathrm{loc}}(\Omega)$ with $p \in (1,\infty)$. For $\alpha \in \{1, \dots, n\}$, we have $\partial_\alpha u \in L^p_{\mathrm{loc}}(\Omega)$ if and only if $(\partial_{h,\alpha} u)_{h \le h_0} \subset L^p_{\mathrm{loc}}$ is uniformly bounded.*

**Proof.** Let $\Omega' \Subset \Omega$ and suppose that $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$.

We first show $\implies$. Assume first that $u$ is smooth. For each $\alpha$, we have

$$|u(x + he_\alpha) - u(x)| = \left| h \int_0^1 \partial_\alpha u(x + the_\alpha)\, \mathrm{d}t \right| \le |h| \int_0^1 |\partial_\alpha u(x + the_\alpha)|\, \mathrm{d}t. \tag{1.3}$$

Thus, dividing by $|h|$ and then applying Jensen's inequality and Fubini's theorem,

$$\begin{aligned} \int_{\Omega'} |\partial_{h,\alpha} u(x)|^p \, \mathrm{d}x &\le \int_{\Omega'} \int_0^1 |\partial_\alpha u(x + the_\alpha)|^p \, \mathrm{d}t \, \mathrm{d}x \\ &= \int_0^1 \int_{\Omega'} |\partial_\alpha u(x + the_\alpha)|^p \, \mathrm{d}x \, \mathrm{d}t \le \int_\Omega |\partial_\alpha u(x)|^p \, \mathrm{d}x. \end{aligned}$$

By approximation, this estimate holds for an arbitrary $u \in W^{1,p}(\Omega)$. Note that this part of the argument in fact holds also for $p = 1$.

Conversely, by integration by parts (cf. Exercise 1.12), we have

$$\int_{\Omega'} u\, \partial_{h,\alpha} \varphi \, \mathrm{d}x = -\int_{\Omega'} \partial_{-h,\alpha} u\, \varphi \, \mathrm{d}x$$

whenever $\varphi \in C_c^\infty(\Omega')$. Recall that, for $p \in (1, \infty)$, bounded subsets of $L^p(\Omega)$ are weakly compact. Thus, as we assume that the family of difference quotients is bounded in $L^p(\Omega')$, for some $p > 1$, there is a function $v_\alpha$ and a sequence $h_j \to 0$ such that

$$\partial_{-h_j,\alpha} u \rightharpoonup v \text{ in } L^p(\Omega').$$

But then

$$\int_{\Omega'} u \partial_\alpha \varphi \, \mathrm{d}x = \lim_{h_j \to 0} \int_{\Omega'} u \partial_{h_j,\alpha} \varphi \, \mathrm{d}x = -\lim_{h_j \to 0} \int_{\Omega'} \partial_{h_j,\alpha} u \varphi \, \mathrm{d}x = -\int_{\Omega'} v_\alpha \varphi \, \mathrm{d}x$$

and so $v_\alpha = \partial_\alpha u$ in the weak sense, over $\Omega'$. In particular, $\partial_\alpha u \in L^p(\Omega')$ and so, as $u \in L^p(\Omega')$, we see that $u \in W^{1,p}(\Omega')$. Note that this part of the argument in fact holds also for $p = \infty$ (if we replace the weak topology with the weak$-*$ topology). $\square$

Inspecting the above argument in the case $p = \infty$ gives us a nice characterization of $W^{1,\infty}$, where we recall that $C^{0,1}(\Omega)$ is the space of bounded Lipschitz functions in $\Omega$:

**Theorem 1.14** ($W^{1,\infty}$)**.** *If $\Omega$ is either $\mathbb{R}^n$ or a smooth bounded domain then we have $C^{0,1}(\Omega) = W^{1,\infty}(\Omega)$.*

**Proof.** We give the proof when $\Omega = \mathbb{R}^n$; the general case follows from Theorem 1.10. Note that if $u \in C^{0,1}(\mathbb{R}^n)$ then $\|\partial_{-h,\alpha} u\|_{L^\infty} \le C$, uniformly in $h$, and so from the argument above when $p = \infty$ we deduce that $u \in W^{1,\infty}(\mathbb{R}^n)$. Conversely, if $u \in W^{1,\infty}(\mathbb{R}^n)$ and $u$ is smooth, then by (1.3) we have

$$h^{-1} \|u(\cdot + he_\alpha) - u\|_{L^\infty} = \|\partial_{h,\alpha} u\|_{L^\infty} \le \|\mathrm{D}u\|_{L^\infty(\mathbb{R}^n)}.$$

For any $u \in W^{1,\infty}$ we can find a sequence $u_j \in C^\infty(\mathbb{R}^n)$ with $u_j \to u$ uniformly and with $\|\mathrm{D}u_j\|_{L^\infty(\mathbb{R}^n)} \le \|\mathrm{D}u\|_{L^\infty(\mathbb{R}^n)}$ (indeed, it suffices to mollify $u$). Thus, for any $x \neq y$, we have

$$\frac{|u(x) - u(y)|}{|x - y|} = \lim_{j \to \infty} \frac{|u_j(x) - u_j(y)|}{|x - y|} \le \|\mathrm{D}u\|_{L^\infty(\mathbb{R}^n)}.$$

Hence the same estimate holds for general functions in $W^{1,\infty}(\mathbb{R}^n)$ which are not necessarily smooth, and we deduce that such functions are Lipschitz. $\square$

As the final result in this subsection, we will see that $W^{1,p}$ is a subspace of $L^p$ with good compactness properties. In order to make this precise, let $\mathbb{X}, \mathbb{Y}$ be Banach spaces and suppose that $\mathbb{X}$ is contained in $\mathbb{Y}$ as a set. Recall that:

(i) $\mathbb{X}$ is said to be *continuously embedded* in $\mathbb{Y}$, $\mathbb{X} \subset \mathbb{Y}$, if

$$\|u\|_{\mathbb{Y}} \leq C\|u\|_{\mathbb{X}} \quad \text{for all } u \in \mathbb{X};$$

(ii) $\mathbb{X}$ is said to be *compactly embedded* in $\mathbb{Y}$, $\mathbb{X} \Subset \mathbb{Y}$, if it is continuously embedded and each bounded sequence in $\mathbb{X}$ is pre-compact in $\mathbb{Y}$.

We have the following fundamental theorem:

**Theorem 1.15** (Rellich–Kondrachov Compactness Theorem)**.** *Assume $\Omega \subset \mathbb{R}^n$ is a bounded smooth domain and $p \in [1, \infty)$. Then $W^{1,p}(\Omega) \Subset L^p(\Omega)$. Thus any bounded sequence $(u_j) \subset W^{1,p}(\Omega)$ has a subsequence which is strongly convergent in $L^p(\Omega)$.*

The basic idea of the proof is again to use difference quotients in order to obtain uniform estimates for the differences $u(\cdot + he_\alpha) - u$.

## 1.3 Sobolev inequalities

In this subsection, we discuss inequalities relating $W^{1,p}$-spaces to other more familiar function spaces. The case $n = 1$ is rather simple (see Problem 1.2), and we are specially concerned with general dimension $n > 1$.

The next inequalities give continuous embeddings of $W^{1,p}(\mathbb{R}^n)$ for $p \in [1, \infty) \setminus \{n\}$. They also extend to smooth, bounded domains $\Omega \subset \mathbb{R}^n$. As usual, we define $p^* \equiv \frac{np}{n-p}$ for $p < n$.

**Theorem 1.16** (Sobolev inequalities)**.** *Let $u \in C^1(\mathbb{R}^n)$.*

(i) *Gagliardo–Nirenberg–Sobolev inequality: if $1 \leq p < n$ and $u$ has compact support, then*

$$\|u\|_{L^{p*}(\mathbb{R}^n)} \leq C\|\mathrm{D}u\|_{L^p(\mathbb{R}^n)}.$$

(ii) *Morrey's inequality: if $n < p$, then*

$$\|u\|_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}.$$

The exponent $p^*$ appearing in the left-hand side in Theorem 1.16(i), which may appear mysterious at first, is in fact the *only* exponent for which such an inequality can hold. This can be seen through simple scaling analysis: if $u_r(x) = u(x/r)$, then

$$\|u_r\|_{L^q(\mathbb{R}^n)} = r^{n/q}\|u\|_{L^q(\mathbb{R}^n)}, \qquad \|\mathrm{D}u_r\|_{L^p(\mathbb{R}^n)} = r^{n/p-1}\|\mathrm{D}u\|_{L^p(\mathbb{R}^n)} \tag{1.4}$$

and so, by sending $r \to 0$ or $r \to \infty$ we see that we must have $\frac{n}{q} = \frac{n}{p} - 1$, i.e. $q = p^*$. Note also that Theorem 1.16(i) shows (applying also Theorems 1.8 and 1.10) that $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ continuously whenever $\Omega$ is a smooth bounded domain. In (ii) we used the Hölder space norms, see already Definition 4.1 below.

From Theorem 1.16 one can also deduce (in an elementary but tedious fashion) more complicated inequalities for higher-order Sobolev spaces. We will not state these inequalities here in their full generality, but instead we simply note the following:

**Corollary 1.17.** *Let $u \in W^{k,p}(\mathbb{R}^n)$ and suppose that $k > \frac{n}{p}$. Then $u \in C^{k-1-[\frac{n}{p}]}(\mathbb{R}^n)$, where $[\frac{n}{p}]$ is the integer part of $\frac{n}{p}$.*

Thus, provided that $k$ is sufficiently large, any function in $W^{k,p}$ actually in some $C^{k'}$-space, where $k'$ can be determined explicitly in terms of $n, k, p$; the explicit dependence is typically not very important.

From Theorem 1.16 and the extension theorem, one deduces that $W^{1,n}(\Omega) \subset L^q(\Omega)$ for any $q < \infty$, whenever $\Omega$ is a bounded smooth domain. Thus one may hope that $W^{1,n}(\Omega) \subset L^\infty(\Omega)$. While this is true for $n = 1$ (see Problem 1.2) it fails for any other dimension:

**Exercise 1.18.** Check that $\log\log(1 + 1/|x|) \in W^{1,n}(B_1(0)) \setminus L^\infty(B_1(0))$, whenever $n > 1$.

In some sense, unbounded functions in $W^{1,n}$ look like the one in the previous exercise, in that they diverge logarithmically; this fact is made extremely precise by the Moser–Trudinger inequality [13]. For some further properties of functions in $W^{1,n}$, see Problem 4.2.

We conclude this section by discussing Poincaré inequalities. Loosely speaking, a Poincaré inequality is an inequality of the form

$$\|u\|_{L^q} \leq C\|\mathrm{D}u\|_{L^p}$$

for appropriate choices of exponents $p$ and $q$. Note that such an inequality cannot be quite true (take $u$ to be a constant), but essentially constants are the only obstruction to its validity.

**Theorem 1.19** (Poincaré's inequalities)**.** *Let $\Omega$ be a bounded, smooth domain and $p \in [1, \infty)$. There is a constant $C = C(p, n, \Omega)$ such that the following inequalities hold.*

(i) *If $p \in [1, n)$ and $u \in W^{1,p}_0(\Omega)$ then*

$$\|u\|_{L^{p^*}(\Omega)} \leq C\|\mathrm{D}u\|_{L^p(\Omega)}.$$

*In particular, we have*

$$\|u\|_{L^p(\Omega)} \leq C\|\mathrm{D}u\|_{L^p(\Omega)}. \tag{1.5}$$

(ii) *If $u \in W^{1,p}(\Omega)$ and we write $(u)_\Omega \equiv \frac{1}{|\Omega|}\int_\Omega u \,\mathrm{d}x$, then*

$$\|u - (u)_\Omega\|_{L^p(\Omega)} \leq C\|\mathrm{D}u\|_{L^p(\Omega)}.$$

**Proof.** We note that (i) is an immediate consequence of Theorem 1.16(i): since $u \in W^{1,p}_0(\Omega)$, there are $u_j \in C^\infty_c(\Omega)$ converging to $u$ in $W^{1,p}(\Omega)$, and we can apply Theorem 1.16(i) to each $u_j$. If we are just interested in (1.5), then it is possible to give a simple direct proof: if $\Omega \Subset [-L, L]^n$ and we write $\bar{x} = (x_2, \dots, x_n)$ then

$$u(x_1, \bar{x}) = \int_{-L}^{1} \partial_1 u(t, \bar{x}) \,\mathrm{d}t \quad \implies \quad |u(x)|^p \leq (2L)^{p-1}\int_{-L}^{L} |\partial_1 u(t, x)|^p \,\mathrm{d}t,$$

hence (1.5) follows by integrating the last inequality in $x$.

For (ii), we argue by a contradiction argument: if the inequality fails then for each $j$ there is $u_j \in W^{1,p}(\Omega)$ such that

$$\|u_j - (u_j)_\Omega\|_{L^p(\Omega)} \geq j\|\mathrm{D}u_j\|_{L^p(\Omega)}.$$

Thus, defining $v_j \equiv \frac{u_j-(u_j)_\Omega}{\|u_j-(u_j)_\Omega\|_{L^p(\Omega)}}$, we see that

$$\|v_j\|_{L^p(\Omega)} = 1, \qquad (v_j)_\Omega = 0, \qquad \|\mathrm{D}v_j\|_{L^p(\Omega)} \leq \frac{1}{j}.$$

In particular, by Theorem 1.15, we have $v_j \to v$ in $L^p(\Omega)$ and so $(v)_\Omega = 0$, $\|v\|_{L^p(\Omega)} = 1$. However, we also have $\mathrm{D}v = 0$: for any test function,

$$\int_\Omega v\,\mathrm{D}\varphi\,\mathrm{d}x = \lim_{j\to\infty}\int_\Omega v_j \mathrm{D}\varphi\,\mathrm{d}x = -\lim_{j\to\infty}\int_\Omega \mathrm{D}v_j\,\varphi\,\mathrm{d}x = 0.$$

Thus, by Lemma 1.5 $v$ is a constant, which is necessarily zero since $(v)_\Omega = 0$. But $\|v\|_{L^p} = 1$, a contradiction. □

Note that Theorem 1.19(i) shows that, in $W_0^{1,p}(\Omega)$, $\|\mathrm{D}u\|_{L^p}$ and $\|u\|_{W^{1,p}}$ are equivalent norms. Thus $W_0^{1,2}(\Omega)$ can be equivalently defined as the closure of test functions under the norm that we encountered in (1.2). This shows that $W_0^{1,2}(\Omega)$ is the smallest Hilbert space in which we can look for solutions to (1.1) using Theorem 1.1.

When $\Omega$ is a ball, for simplicity we will write

$$(u)_{B_r(x_0)} \equiv (u)_{x_0,r}.$$

By rescaling, we have the following corollary of Theorem 1.19(ii):

**Corollary 1.20** (Poincaré's inequality for a ball)**.** *For each $p \in [1,\infty)$ there is a constant $C = C(p,n)$ such that, for all $u \in W^{1,p}(B_r(x_0))$, we have*

$$\|u - (u)_{x_0,r}\|_{L^p(B_r(x_0))} \leq Cr\|\mathrm{D}u\|_{L^p(B_r(x_0))}.$$

**Proof.** The case where $r = 1$ follows from Theorem 1.19(ii). For general $r$, write $v(x) \equiv u(x_0 + rx)$, which defines an element in $W^{1,p}(B_1(0))$. Thus

$$\|v - (v)_{0,1}\|_{L^p(B_1(0))} \leq C\|\mathrm{D}v\|_{L^p(B_1(0))}$$

and changing variables we recover the desired inequality, cf. (1.4). □

## 1.4 Problems for Section 1

**Problem 1.1** (Riesz Representation Theorem)**.** Prove the characterization of the minimum in Theorem 1.1 and prove also that

$$\|u\|_H = \|\ell\|_{H^*} \equiv \sup_{\|v\|_H=1} \ell(v).$$

Deduce that there is an isometry $H \to H^*$.

**Problem 1.2** (Sobolev spaces when $n = 1$)**.** Let $p \in [1,\infty)$ and $u \in W^{1,p}(I)$ where $I = (0,1)$. Show that $u$ coincides a.e. with an absolutely continuous function whose derivative $u'$ (which exists a.e.) is in $L^p$.

**Problem 1.3** (Almost every sphere is a "good sphere")**.** If $u \in W^{1,p}(\Omega)$ show that for a.e. $r \in (0, \infty)$ we have

$$\int_{\Omega \cap \partial B_r(0)} [|u|^p + |\mathrm{D}u|^p] \, \mathrm{d}\omega < \infty.$$

Thus, for a.e. $r$, we have $u|_{\Omega \cap \partial B_r(0)} \in W^{1,p}(\Omega \cap \partial B_r(0))$.

**Problem 1.4** (De Giorgi isoperimetric inequality)**.** Let $u \in W^{1,2}(B_1)$ and consider the subsets

$$A \equiv \{u \leq 0\}, \qquad D \equiv \{u \geq \tfrac{1}{2}\}, \qquad E \equiv \{0 < u < \tfrac{1}{2}\}$$

of $B_1$. For $p > 1$, show that there is a constant $C = C(n, p)$ such that

$$|A||D| \leq C \|\mathrm{D}u\|_{L^p} |E|^{\frac{1}{p'}},$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. [*Hint: reduce to the case where $u = 0$ in $A$ and $u = \frac{1}{2}$ in $D$, and then start from the inequality $|A||D| \leq 2 \int_A \int_D |u(x) - u(y)| \, \mathrm{d}x \, \mathrm{d}y$.*] Conclude that, for $p > 1$, Sobolev functions cannot have arbitrarily fast jumps. Show that this fails for $p = 1$ by constructing a sequence bounded in $W^{1,1}$ which jumps arbitrarily quickly.

**Problem 1.5** (Ignoring singletons)**.** In this problem we show that weak derivatives are somewhat robust. We take $n \geq 2$.

(i) Show that for $\varphi \in C_c^\infty(B_1)$ there is $(\varphi_j) \subset C_c^\infty(B_1 \setminus \{0\})$ such that $\varphi_j \to \varphi$ in $W^{1,2}(B_1)$. [*Hint: let $\varphi_j = \varphi \, \psi_j$, where $\psi_j$ is a approximation of $x \mapsto |x|^{1/j}$.*]

(ii) Deduce that if $v \in L^2(B_1)$ and $\partial_\alpha v = 0$ weakly in $B_1 \setminus \{0\}$ then $\partial_\alpha v = 0$ weakly in $B_1$.

# 2 Existence theory for variational problems

## 2.1 Existence for linear elliptic systems

The purpose of this subsection is to construct weak solutions of the linear system

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

where repeated indices (i.e. $\alpha, \beta, j$) are always assumed to be summed. Here and in the rest of these notes, we will assume that $x \mapsto A^{ij}_{\alpha\beta}(x)$ is in measurable and bounded, with

$$\|A\|_{L^\infty(\Omega)} \leq \Lambda. \tag{2.2}$$

**Notation 2.1.** Sometimes it will be convenient to use a more compact notation. Note that for $x \in \Omega$ we have $A(x)\,\mathrm{D}u(x) \in \mathbb{R}^{m\times n}$, where its entries are $(A\,\mathrm{D}u)^\alpha_i = A^{\alpha\beta}_{ij}\partial_\beta u^j$. Thus we can write (2.1) more concisely as

$$-\operatorname{div}(A\,\mathrm{D}u) = f,$$

where the divergence is taken row-wise. Given two matrices $A, B \in \mathbb{R}^{m\times n}$, it is also convenient to denote by $\langle A, B\rangle = \operatorname{tr}(A^T B) = \operatorname{tr}(AB^T)$ their inner product; we use the same notation to denote inner products between vectors in $\mathbb{R}^m$. Likewise we write $|\cdot|$ for the Euclidean norm of both matrices and vectors.

Let us now give precisely the definition of weak solution:

**Definition 2.2.** A *weak solution* of (2.1) is a map $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ such that

$$\int_\Omega \langle A\,\mathrm{D}u, \mathrm{D}\varphi\rangle \,\mathrm{d}x \equiv \int_\Omega A^{\alpha\beta}_{ij}\partial_\beta u^j \partial_\alpha \varphi^i = \int_\Omega f_i \varphi^i \equiv \int_\Omega \langle f, \varphi\rangle \,\mathrm{d}x \qquad \forall \varphi \in C^\infty_c(\Omega, \mathbb{R}^m).$$

In other words, $u$ is a weak solution of (2.1) if the weak derivative of $A^{\alpha\beta}_{ij}\partial_\beta u^j$ is $f_i$. Note that this definition is very natural from the variational viewpoint. For instance, if $A$ is symmetric, i.e. if $A^{ij}_{\alpha\beta} = A^{ji}_{\beta\alpha}$, then linear systems as in (2.1) arise naturally as the Euler–Lagrange system for the energy

$$\mathbb{E}[u] \equiv \int_\Omega \frac{1}{2}\langle A\,\mathrm{D}u, \mathrm{D}u\rangle - \langle f, u\rangle \,\mathrm{d}x, \qquad u \colon \Omega \to \mathbb{R}^m.$$

Thus, in this case, the definition of a weak solution $u$ asserts precisely that $\mathbb{E}'[u](\varphi) = 0$ for all $\varphi \in C^\infty_c(\Omega, \mathbb{R}^m)$, i.e. it asserts that $u$ is a critical point of $\mathbb{E}$. It is convenient to give a name for the quadratic form on the left-hand side of the definition of weak solution:

$$a_A(u, \varphi) \equiv \int_\Omega \langle A\,\mathrm{D}u, \mathrm{D}\varphi\rangle \,\mathrm{d}x.$$

Let us now turn to the existence of weak solutions. We will first assume that the 4-tensor $A$ satisfies the following ellipticity condition:

**Definition 2.3.** We say that $A(x) = (A^{\alpha\beta}_{ij}(x))$ satisfies the *Legendre condition* if there is $\lambda > 0$ such that

$$\langle A\xi, \xi\rangle \geq \lambda|\xi|^2 \quad \text{for a.e. } x \text{ and for all } \xi \in \mathbb{R}^{m\times n}. \tag{2.3}$$

Note that, if $A$ is symmetric and it satisfies the Legendre condition, then $a_A$ is a scalar product in $W^{1,2}_0(\Omega, \mathbb{R}^m)$ which is equivalent to the standard one. In particular, under these assumptions, existence of a solution to (2.1) follows from Theorem 1.1. There is a more general result in this direction, which does not require any symmetry assumptions:

**Theorem 2.4** (Lax–Milgram)**.** *Let $H$ be a Hilbert space and let $a\colon H \times H \to \mathbb{R}$ be a bilinear form such that:*

(i) *$a$ is continuous: there is $\Lambda > 0$ such that $a(u,v) \leq \Lambda\|u\|\|v\|$ for all $u, v \in H$;*

(ii) *$a$ is coercive: there is $\lambda > 0$ such that $a(u,u) \geq \lambda\|u\|^2$ for all $u \in H$.*

*For any $\ell \in H^*$ there is a unique $u_\ell \in H$ with*

$$a(u_\ell, \cdot) = \ell.$$

**Proof.** For each fixed $u \in H$ our assumptions guarantee that $a(u, \cdot) \in H^*$. Thus, by the Riesz Representation Theorem, for each $u \in H$ there is a unique vector in $H$, which we denote by $Tu$, such that

$$a(u, \cdot) = \langle \cdot, Tu \rangle.$$

This clearly defines an operator $T\colon H \to H$, which is easily seen to be linear, since $a$ is bilinear. Since $a$ is bounded, so is $T$: we have

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

i.e. $\|Tu\| \leq \Lambda\|u\|$ for all $u \in H$. Since $a$ is coercive, so is $T$: we have

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

i.e. $\lambda\|u\|^2 \leq \|Tu\|$. This last inequality shows that $T$ is injective and has closed range. In fact, $T$ is onto $H$: if not, since its range is closed, there would be $w \neq 0$ orthogonal to the range of $T$, and so

$$\lambda\|w\|^2 \leq a(w,w) = \langle Tw, w \rangle = 0,$$

a contradiction. Thus $T\colon H \to H$ is a linear isomorphism, from which the conclusion follows: writing $\ell = \langle \cdot, w \rangle$ for some $w \in H$, we simply take $u_\ell = T^{-1}w$. $\square$

Note that, in view of (2.2), the continuity assumption on $a_A$ is always satisfied. Thus, in order to prove existence of a $W_0^{1,2}(\Omega, \mathbb{R}^m)$-solution of (2.1), we only need to ask for coercivity of the form $a_A$. Clearly the Legendre condition implies coercivity, but is this condition natural?

In order to answer this question, let us first consider the case where $A$ has *constant-coefficients*, thus we assume that $A = (A_{ij}^{\alpha\beta})$ is $x$-independent. We want to compare the two following coercivity conditions, one of which is *pointwise* (like in the Legendre condition), the other being *integral* (like in Lax–Milgram's theorem):

(i) $\langle A\xi, \xi \rangle \geq \lambda|\xi|^2$ for all $\xi \in \mathbb{R}^{m\times n}$;

(ii) $a_A(u,u) \geq \lambda \int_\Omega |\mathrm{D}u|^2 \,\mathrm{d}x$ for all $u \in W_0^{1,2}(\Omega, \mathbb{R}^m)$.

Clearly (i) $\Longrightarrow$ (ii), so we want to understand whether the converse holds. It turns out that the answer is in general no:

**Example 2.5.** Take $n = m = 2$ and consider

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

for some $\varepsilon > 0$. If we take $\xi = \operatorname{diag}(t, -t)$ then

$$\langle A\xi, \xi\rangle = -t^2 + 2t^2\varepsilon = t^2(2\varepsilon - 1) < 0$$

whenever $\varepsilon < \frac{1}{2}$. Hence, in this case, (i) fails even for $\lambda = 0$. However, it is easy to see that (ii) holds with $\lambda = \varepsilon$, since

$$\int_\Omega \det \mathrm{D}u \,\mathrm{d}x = 0 \qquad \text{for all } u \in W_0^{1,2}(\Omega, \mathbb{R}^2).$$

This last identity can be proved first for smooth functions $u \in C_c^\infty(\Omega, \mathbb{R}^2)$, the general case following by an approximation argument. For test functions, the identity follows from the divergence theorem, since

$$\det \mathrm{D}u = \partial_1 u^1 \partial_2 u^2 - \partial_2 u^1 \partial_1 u^2 = \partial_1(u^1 \partial_2 u^2) - \partial_2(u^1 \partial_1 u^2). \tag{2.4}$$

The previous example shows that integral coercivity conditions do not imply the Legendre condition. However, they do imply another (weaker) type of pointwise coercivity:

**Definition 2.6.** We say that $A(x) = (A_{ij}^{\alpha\beta}(x))$ satisfies the *Legendre–Hadamard condition* if there is $\lambda > 0$ such that, for any $a \in \mathbb{R}^m, b \in \mathbb{R}^n$,

$$\langle A(a \otimes b), a \otimes b\rangle \geq \lambda |a|^2 |b|^2 \quad \text{for a.e. } x \text{ in } \Omega. \tag{2.5}$$

Here $\xi = a \otimes b$ is a rank-one matrix, with entries $(a \otimes b)^i_\alpha = a^i b_\alpha$, thus (2.5) reads as

$$\langle A\xi, \xi\rangle = A_{ij}^{\alpha\beta} a^i a^j b_\alpha b_\beta \geq \lambda |a|^2 |b|^2 = \lambda |\xi|^2.$$

**Exercise 2.7.** Check that a matrix $\xi \in \mathbb{R}^{m\times n}$ has rank one if and only if it can be written in the form $\xi = a \otimes b$ for $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Then check that Example 2.5 satisfies (2.5).

It should be clear that if either $m = 1$ or if $n = 1$ then the Legendre and the Legendre–Hadamard conditions are equivalent, while Example 2.5 shows that they differ if $m, n > 1$.

We are yet to justify that the Legendre–Hadamard condition is the correct condition in general. Let us first show that it is a necessary condition:

**Proposition 2.8.** *Suppose that $A$ is constant and that*

$$a_A(u, u) \geq \lambda \int_\Omega |\mathrm{D}u|^2 \,\mathrm{d}x \qquad \forall u \in W_0^{1,2}(\Omega, \mathbb{R}^m). \tag{2.6}$$

*Then $A$ satisfies the Legendre–Hadamard condition* (2.5)*.*

**Proof.** We begin by extending $a_A$ to complex-valued maps, through

$$a_A(u, v) \equiv \int_\Omega A_{ij}^{\alpha\beta} \partial_\alpha u^i \overline{\partial_\beta v^j} \,\mathrm{d}x = \int_\Omega \langle A\mathrm{D}u, \overline{\mathrm{D}v}\rangle \,\mathrm{d}x.$$

In general, we have

$$a_A(u, u) = a_A(\operatorname{Re} u, \operatorname{Re} u) + a_A(\operatorname{Im} u, \operatorname{Im} u) + i\left[a_A(\operatorname{Im} u, \operatorname{Re} u) - a_A(\operatorname{Re} u, \operatorname{Im} u)\right],$$

where the second term vanishes whenever $A$ is symmetric. We now fix $\varphi \in C_c^\infty(\Omega)$, $v \in \mathbb{R}^n, w \in \mathbb{R}^m$. The idea is to take $u_t(x) = \varphi(x) e^{itx\cdot v} w$ and send $t \to \infty$. Note that

$$\mathrm{D}u_t(x) = e^{itx\cdot v}\left(\mathrm{D}\varphi(x) \otimes w\right) + it\varphi(x) e^{itx\cdot v}\left(v \otimes w\right),$$

and as the last term is a scalar multiple of a fixed matrix $v \otimes w$, we see that

$$a_A(\operatorname{Im} u_t, \operatorname{Re} u_t) - a_A(\operatorname{Re} u_t, \operatorname{Im} u_t) = o(t^2),$$

which means that $a_A(u_t, u_t)$ behaves as though $A$ is symmetric. Thus

$$a_A(\operatorname{Re} u_t, \operatorname{Re} u_t)+a_A(\operatorname{Im} u_t, \operatorname{Im} u_t) = a_A(u_t, u_t)+o(t^2) = t^2 \int_\Omega \varphi(x)^2 \langle A(v\otimes w), v\otimes w\rangle \,\mathrm{d}x+o(t^2),$$

and so, after rearranging, since $A$ is constant,

$$\begin{aligned}\langle A(v \otimes w), v \otimes w\rangle \int_\Omega \varphi^2 \,\mathrm{d}x &= \frac{1}{t^2}\left[a_A(\operatorname{Re} u_t, \operatorname{Re} u_t) + a_A(\operatorname{Im} u_t, \operatorname{Im} u_t)\right] + o(1) \\ &\geq \frac{\lambda}{t^2} \int_\Omega |\mathrm{D}(\operatorname{Re} u_t)|^2 + |\mathrm{D}(\operatorname{Im} u_t)|^2 \,\mathrm{d}x \\ &= \lambda\, |v \otimes w|^2 \int_\Omega \varphi^2 \,\mathrm{d}x + o(1),\end{aligned}$$

since in general $|\mathrm{D}u|^2 = |\mathrm{D}(\operatorname{Re} u)|^2 + |\mathrm{D}(\operatorname{Im} u)|^2$ for a complex-valued map. The conclusion follows by choosing any non-zero $\varphi$. $\square$

We note that there is a generalization of Proposition 2.8 to the case where $A$ has variable coefficients, see Problem 2.1. We now turn to the converse of Proposition 2.8, which asserts the sufficiency of the Legendre–Hadamard condition for coercivity:

**Theorem 2.9** (Gårding's inequality)**.** *Let $A$ be constant and satisfy the Legendre–Hadamard condition. Then* (2.6) *holds.*

**Proof.** The proof relies on the Fourier transform, which makes transparent why the Legendre–Hadamard condition is useful. Recall that, for $f \in L^2(\mathbb{R}^n)$, its Fourier transform is

$$\widehat{f}(\xi) \equiv \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi} \,\mathrm{d}x.$$

The Fourier transform is an isometry in $L^2$ (Parseval's identity) and satisfies $\widehat{\partial_\alpha f}(\xi) = 2\pi i \xi_\alpha \widehat{f}$ (derivatives are converted into multiplication and conversely). Using these properties, and extending $u \in W^{1,2}_0(\Omega)$ by zero outside $\Omega$, we have

$$\begin{aligned}a_A(u, u) = A^{\alpha\beta}_{ij} \int_{\mathbb{R}^n} \widehat{\partial_\alpha u^i}\, \overline{\widehat{\partial_\beta u^j}} \,\mathrm{d}\xi &= (2\pi)^2 A^{\alpha\beta}_{ij} \int_{\mathbb{R}^n} \xi_\alpha \xi_\beta \widehat{u}^i \overline{\widehat{u}^j} \,\mathrm{d}\xi \\ &\geq (2\pi)^2 \lambda \int_{\mathbb{R}^n} |\xi|^2 |\widehat{u}|^2 \,\mathrm{d}\xi = \lambda \int_{\mathbb{R}^n} |\widehat{\mathrm{D}u}|^2 \,\mathrm{d}\xi = \lambda \int_{\mathbb{R}^n} |\mathrm{D}u|^2 \,\mathrm{d}x,\end{aligned}$$

as wished. $\square$

**Corollary 2.10** (Existence)**.** *Let $A$ have constant coefficients and satisfy the Legendre–Hadamard condition. Then for any $f \in L^2(\Omega, \mathbb{R}^{m\times n})$ there is a unique weak solution to the Dirichlet problem*

$$\begin{cases} \partial_\alpha(A^{\alpha\beta}_{ij} \partial_\beta u^j) = \partial_\alpha f^\alpha_i & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

**Proof.** Note that the weak formulation of the PDE is that, for all $\varphi \in C_c^\infty(\Omega, \mathbb{R}^m)$,

$$\int_\Omega A_{ij}^{\alpha\beta} \partial_\beta u^j \partial_\alpha \varphi^i \, \mathrm{d}x = \int_\Omega f_\alpha^i \partial_\alpha \varphi^i \, \mathrm{d}x.$$

The right-hand side in this identity defines a linear functional $\ell(\varphi)$, which is bounded in $W^{1,2}(\Omega, \mathbb{R}^m)$, since

$$|\ell(\varphi)| = \left| \int_\Omega f_\alpha^i \partial_\alpha \varphi^i \, \mathrm{d}x \right| \leq \|f\|_{L^2(\Omega)} \|\mathrm{D}\varphi\|_{L^2(\Omega)}$$

by Hölder's inequality. Since the bilinear form $a_A$ is coercive thanks to Gårding's inequality, and bounded by assumption (2.2), the conclusion follows from the Lax–Milgram theorem. $\square$

We now generalize Gårding's inequality to the case of variable coefficients. The basic idea, which is ubiquitous in PDE, is to treat the oscillations in the coefficients *perturbatively.*

**Corollary 2.11.** *Let $A = A(x) \in C^0(\overline{\Omega})$ satisfy the Legendre–Hadamard condition. Then $a_A$ is weakly coercive: there are $\lambda_0 > 0, \lambda_1 \geq 0$ such that*

$$a_A(u,u) \geq \lambda_0 \int_\Omega |\mathrm{D}u|^2 \, \mathrm{d}x - \lambda_1 \int_\Omega |u|^2 \, \mathrm{d}x \quad \textit{for all } u \in W_0^{1,2}(\Omega, \mathbb{R}^m).$$

**Proof.** Take $x_0 \in \Omega$ and $r$ small enough. Let us first suppose that $u \in W_0^{1,2}(B_r(x_0))$. Then, by Theorem 2.9, we have

$$\begin{aligned} a_A(u,u) &= A_{ij}^{\alpha\beta}(x_0) \int_\Omega \partial_\alpha u^i \partial_\beta u^j \, \mathrm{d}x + \int_\Omega [A_{ij}^{\alpha\beta}(x) - A_{ij}^{\alpha\beta}(x_0)] \partial_\alpha u^i \partial_\beta u^j \, \mathrm{d}x \\ &\geq (\lambda - \omega(r)) \int_\Omega |\mathrm{D}u|^2 \, \mathrm{d}x, \end{aligned}$$

where $\omega(r) \equiv \sup\{\max_{\alpha,\beta,i,j} |A_{ij}^{\alpha\beta}(x) - A_{ij}^{\alpha\beta}(y)| : x, y \in \Omega, |x-y| \leq r\}$ is the modulus of continuity of $A$.

In order to deal with general maps, fix $r$ so small that $\lambda_0 \equiv \lambda - \omega(r) > 0$, and cover $\overline{\Omega}$ with a finite number of balls $\{B_r(x_k)\}$. Let $\{\varphi_k^2\}$ be a partition of unity associated with this covering, i.e. $\varphi_k \in C_c^\infty(B_r(x_k))$ and $\sum_k \varphi_k^2 = 1$ on $\overline{\Omega}$. Thus

$$\begin{aligned} a_A(u,u) = \int_\Omega A_{ij}^{\alpha\beta} \sum_k \varphi_k^2 \partial_\alpha u^i \partial_\beta u^j \, \mathrm{d}x &= \sum_k \left[ \int_\Omega A_{ij}^{\alpha\beta} \partial_\alpha(\varphi_k u^i) \partial_\beta(\varphi_k u^j) \, \mathrm{d}x - \int_\Omega \mathrm{error}_k \, \mathrm{d}x \right] \\ &\geq \sum_k \left[ \lambda_0 \int_\Omega |\mathrm{D}(\varphi_k u)|^2 \, \mathrm{d}x - \int_\Omega \mathrm{error}_k \, \mathrm{d}x \right], \end{aligned}$$

where

$$\mathrm{error}_k = A_{ij}^{\alpha\beta} u^i u^j \partial_\alpha \varphi_k \partial_\beta \varphi_k + A_{ij}^{\alpha\beta} \varphi_k u^j \partial_\alpha u^i \partial_\beta \varphi_k + A_{ij}^{\alpha\beta} \varphi_k u^i \partial_\alpha \varphi_k \partial_\beta u^j.$$

Since $\mathrm{D}(\varphi_k u) = \varphi_k \mathrm{D}u + u \otimes \mathrm{D}\varphi_k$, we have

$$|\mathrm{D}(\varphi_k u)|^2 = \varphi_k^2 |\mathrm{D}u|^2 + |u \otimes \mathrm{D}\varphi_k|^2 + 2\varphi_k \langle \mathrm{D}u, u \otimes \mathrm{D}\varphi_k \rangle$$

Applying Young's inequality $|2ab| \leq \varepsilon a^2 + \varepsilon^{-1} b^2$ we get

$$2\varphi_k \langle \mathrm{D}u, u \otimes \mathrm{D}\varphi \rangle \geq -\varepsilon \varphi_k^2 |\mathrm{D}u|^2 - \frac{1}{2\varepsilon} |u|^2 |\mathrm{D}\varphi_k|^2$$

and so, for a constant $C_\varepsilon = C(\varepsilon, \{\varphi_k\}_k)$, we have

$$a_A(u,u) \geq (\lambda_0 - \varepsilon)\int_\Omega |\mathrm{D}u|^2 - C_\varepsilon \int_\Omega |u|^2 \,\mathrm{d}x - \int_\Omega \sum_k \mathrm{error}_k \,\mathrm{d}x.$$

It remains to deal with the error terms, which we estimate similarly: e.g. we have

$$\left|A_{ij}^{\alpha\beta} u^i u^j \partial_\alpha \varphi_k \partial_\beta \varphi_k\right| \leq C\Lambda |u|^2, \qquad \left|A_{ij}^{\alpha\beta} \varphi_k u^j \partial_\alpha u^i \partial_\beta \varphi_k\right| \leq \varepsilon \varphi_k^2 |\mathrm{D}u|^2 + C\Lambda\varepsilon^{-1}|u|^2|\mathrm{D}\varphi_k|^2.$$

Thus, all in all, we get

$$a_A(u,u) \geq (\lambda_0 - 2\varepsilon)\int_\Omega |\mathrm{D}u|^2 - C_\varepsilon \int_\Omega |u|^2 \,\mathrm{d}x,$$

and so we can now choose $\varepsilon = \lambda_0/4$ to get the conclusion. □

**Remark 2.12.** Corollary 2.11 is essentially optimal. If $A = A(x) \in L^\infty$ is not continuous then the above weak coercivity estimate fails [15] and if $A \in C^0(\overline{\Omega})$ is continuous (but does not have small BMO-norm) then $a_A$ is not strongly coercive [16].

## 2.2 Existence for nonlinear variational problems: the Direct Method

As discussed in the introduction, the classical problem in the Calculus of Variations is to consider critical points of integral energies such as

$$\mathbb{E}[u] \equiv \int_\Omega F(x, u, \mathrm{D}u)\,\mathrm{d}x, \tag{2.7}$$

where $u\colon \Omega \to \mathbb{R}^m$ as usual, and $F\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is smooth. If $F$ is quadratic in $\mathrm{D}u$ then we obtain a linear elliptic system as in (2.1), but in general a formal application of the chain rule leads us to the *quasilinear* Euler–Lagrange system

$$\partial_\alpha(\partial_{\xi^i_\alpha} F(x,u,\mathrm{D}u)) = \partial_{u^i} F(x,u,\mathrm{D}u) \qquad i = 1,\ldots,m. \tag{2.8}$$

We now want to give precise conditions under which this formal calculation can be made rigorous. To do so, let us introduce the following definition, which will play a key role in these notes:

**Definition 2.13.** We say that $u \in W^{1,p}(\Omega,\mathbb{R}^m)$ is a *minimizer* for $\mathbb{E}$ in $W^{1,p}(\Omega,\mathbb{R}^m)$ if

$$\mathbb{E}[u] \leq \mathbb{E}[u+\varphi] \quad \text{for all } \varphi \in W^{1,p}_0(\Omega,\mathbb{R}^m).$$

Equivalently, $u$ minimizes $\mathbb{E}$ in its own Dirichlet class $W^{1,p}_u(\Omega,\mathbb{R}^m) \equiv u + W^{1,p}_0(\Omega,\mathbb{R}^m)$.

Under suitable growth conditions on the integrand $F$, one can derive the Euler–Lagrange system:

**Proposition 2.14** (Euler–Lagrange system)**.** *Let $F\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be smooth and suppose that there is $g \in L^{p'}(\Omega)$ and $C \geq 0$ such that, at all points $(x,u,\xi) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$,*

$$|F(x,u,\xi)| \leq g(x) + C(|u|^p + |\xi|^p),$$
$$|\partial_u F(x,u,\xi)| + |\partial_\xi F(x,u,\xi)| \leq g(x) + C(|u|^{p-1} + |\xi|^{p-1}),$$

*for some $p \in [1,\infty)$; here $\frac{1}{p}+\frac{1}{p'}=1$. If $u \in W^{1,p}(\Omega,\mathbb{R}^m)$ is a minimizer for $\mathbb{E}$ then $u$ is a weak solution of (2.8), in the sense that, for all $\varphi \in W^{1,p}_0(\Omega,\mathbb{R}^m)$, we have*

$$\int_\Omega -\langle \partial_\xi F(x,u,\mathrm{D}u), \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle \partial_u F(x,u,\mathrm{D}u), \varphi\rangle \,\mathrm{d}x.$$

**Proof.** Fix $\varphi \in W^{1,p}_0(\Omega,\mathbb{R}^m)$ and let $e(t) \equiv \mathbb{E}[u+t\varphi]$. Note that, by the growth condition on $F$ itself, $e\colon \mathbb{R}\to\mathbb{R}$ is finite. We can then write

$$\frac{e(t)-e(0)}{t} = \frac{1}{t}\int_\Omega \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} F(x,u+s\varphi,\mathrm{D}u+s\mathrm{D}\varphi)\,\mathrm{d}s\,\mathrm{d}x = \int_\Omega h(x,t)\,\mathrm{d}x,$$

where we set

$$h(x,t) \equiv \frac{1}{t}\int_0^t \langle \partial_u F(x,u+s\varphi,\mathrm{D}u+s\mathrm{D}\varphi),\varphi\rangle + \langle \partial_\xi F(x,u+s\varphi,\mathrm{D}u+s\mathrm{D}\varphi),\mathrm{D}\varphi\rangle \,\mathrm{d}s$$

where the brackets denote the natural inner products between vectors or matrices. Since $u$ is a minimizer, we have

$$0 = \lim_{t\to 0}\frac{e(t)-e(0)}{t} = \lim_{t\to 0}\int_\Omega h(x,t)\,\mathrm{d}x.$$

provided the limits exist. For a.e. $x\in\Omega$ we clearly have

$$\lim_{t\to 0} h(x,t) = \langle \partial_u F(x,u,\mathrm{D}u),\varphi\rangle + \langle \partial_\xi F(x,u,\mathrm{D}u),\mathrm{D}\varphi\rangle$$

and so we only need to justify that we can apply the Dominated Convergence Theorem. This in turn follows from the growth conditions on the derivatives of $F$. Indeed, we have

$$\begin{aligned}
|\langle \partial_u F(x,u+s\varphi,\mathrm{D}u+s\mathrm{D}\varphi),\varphi\rangle| &\le g(x)|\mathrm{D}\varphi(x)| + C(|u+s\varphi|^{p-1}+|\mathrm{D}u+s\mathrm{D}\varphi|^{p-1})|\varphi|,\\
|\langle \partial_\xi F(x,u+s\varphi,\mathrm{D}u+s\mathrm{D}\varphi),\mathrm{D}\varphi\rangle| &\le g(x)|\mathrm{D}\varphi(x)| + C(|u+s\varphi|^{p-1}+|\mathrm{D}u+s\mathrm{D}\varphi|^{p-1})|\mathrm{D}\varphi|
\end{aligned}$$

and so, for $t\le 1$,

$$|h(t,\cdot)| \le g|\mathrm{D}\varphi| + C(|u|^{p-1}+|\mathrm{D}u|^{p-1}+|\varphi|^{p-1}+|\mathrm{D}\varphi|^{p-1})(|\varphi|+|\mathrm{D}\varphi|).$$

By Hölder's inequality the right-hand side is in $L^1$ and so the Dominated Convergence Theorem is applicable. $\square$

In general, there is no systematic form of constructing *all* solutions to (2.8). A special case occurs whenever the integrand is convex, and we then have the following converse to Proposition 2.14:

**Proposition 2.15.** *Assume that $F\colon \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n}\to\mathbb{R}$ is smooth and that*

$$(u,\xi)\mapsto F(x,u,\xi) \text{ is convex.}$$

*Then each weak solution $u\in W^{1,p}(\Omega,\mathbb{R}^m)$ of (2.8) is a minimizer of $\mathbb{E}$ in $W^{1,p}$.*

**Proof.** By the convexity assumption, we have for a.e. $x$ the inequality

$$F(x,v,\mathrm{D}v) \ge F(x,u,\mathrm{D}u) + \langle \partial_\xi F(x,u,\mathrm{D}u),\mathrm{D}v-\mathrm{D}u\rangle + \langle \partial_u F(x,u,\mathrm{D}u), v-u\rangle$$

whenever $v$ is a Sobolev map. Taking $v\in u+W^{1,p}_0(\Omega,\mathbb{R}^m)$ and integrating the above inequality over $\Omega$, we find

$$\mathbb{E}[v] \ge \mathbb{E}[u] + \int_\Omega \Big[\langle \partial_\xi F(x,u,\mathrm{D}u),\mathrm{D}(v-u)\rangle + \langle \partial_u F(x,u,\mathrm{D}u),v-u\rangle\Big]\,\mathrm{d}x,$$

and the last integral vanishes due to the boundary conditions on $v$ and the fact that $u$ is a solution of (2.8). Thus $\mathbb{E}[v] \geq \mathbb{E}[u]$, as wished. $\square$

As in finite-dimensional problems, for general integrands which are not convex in $(u, \xi)$, there are solutions of (2.8) which are not minimizers. Thus one may ask for a procedure to find *all* solutions to (2.8). In general this is a difficult question: we already saw, in Corollary 2.11, that there are difficulties in solving the Euler–Lagrange system even when it is linear and just $x$-dependent. Instead of looking for all solutions of (2.8), we will be content with constructing minimizers of (2.7): unlike the case of critical points, there is a systematic approach for constructing minimizers, which is based on the so-called Direct Method of the Calculus of Variations:

**Theorem 2.16** (Direct method). *Let $\mathbb{X}$ be a reflexive Banach space and let $\mathbb{Y} \subset \mathbb{X}$ be a weakly closed set. Suppose that $\mathbb{E}\colon \mathbb{X} \to \mathbb{R}$ is a (non-linear) functional satisfying the following conditions:*

(i) *coercivity: $\mathbb{E}[v] \to \infty$ if $\|v\| \to \infty$.*

(ii) *sequential weak lower semicontinuity: if $(u_j) \subset \mathbb{Y}$ and $u_j \rightharpoonup u$ in $\mathbb{X}$, then*

$$\mathbb{E}[u] \leq \liminf_{j\to\infty} \mathbb{E}[u_j].$$

*Then $\mathbb{E}$ is bounded from below on $\mathbb{Y}$ and attains its infimum on $\mathbb{Y}$.*

**Proof.** Let $(u_j) \subset \mathbb{Y}$ be a minimizing sequence in $\mathbb{Y}$, i.e.

$$\lim_{j\to\infty} \mathbb{E}[u_j] = \inf_{v\in\mathbb{Y}} \mathbb{E}[v].$$

By the coercivity condition, $(u_j)$ is a bounded sequence and so, up to a subsequence, $u_j \rightharpoonup u$ for some $u \in \mathbb{X}$, since $\mathbb{X}$ is reflexive. Since $\mathbb{Y}$ is weakly closed $u \in \mathbb{Y}$. By weak lsc,

$$\mathbb{E}[u] \leq \liminf_{j\to\infty} \mathbb{E}[u_j],$$

and the conclusion follows. $\square$

The coercivity condition holds e.g. if we impose suitable pointwise conditions on $F$. Here we are mostly concerned with the problem of finding conditions under which sequential wlsc holds; we are of course interested in taking $\mathbb{X}$ to be a Sobolev space. For $p \in [1, \infty)$, we say that a sequence $u_j$ converges *weakly* to $u$ in $W^{1,p}(\Omega)$, denoted by $u_j \rightharpoonup u$ in $W^{1,p}$, if $u_j \rightharpoonup u$ in $L^p(\Omega)$ and $\mathrm{D}u_j \rightharpoonup \mathrm{D}u$ in $L^p(\Omega)$; this definition coincides with the usual functional analytic definition of weak convergence [11, Exercise 11.64]. The following classical theorem gives us a sufficient condition for sequential wlsc:

**Theorem 2.17** (Tonelli). *Assume that $F\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is smooth, bounded from below, and that*

$$\xi \mapsto F(x, u, \xi) \text{ is convex.}$$

*Then $\mathbb{E}$ is sequentially wlsc in $W^{1,p}(\Omega, \mathbb{R}^m)$ for any $p \in [1, \infty)$.*

**Proof.** Without loss of generality we can assume that $F \geq 0$, otherwise we add a large constant to $F$. Let $u_j \rightharpoonup u$ in $W^{1,p}(\Omega, \mathbb{R}^m)$; we aim to show that

$$\mathbb{E}[u] \leq \liminf_{j\to\infty} \mathbb{E}[u_j].$$

By passing to a subsequence, which we do not relabel, we can replace the lim inf with an actual limit. By the Rellich–Kondrachov compactness theorem (Theorem 1.15) we see that, up to a subsequence, $u_j \to u$ in $L^p(\Omega)$ and so, up to a further subsequence, $u_j \to u$ a.e. in $\Omega$. We can now apply Egorov's theorem: for each $\varepsilon > 0$ there is a set $E_\varepsilon$ with $|\Omega \setminus E_\varepsilon| < \varepsilon$ and such that $u_j \to u$ uniformly on $E_\varepsilon$. We may assume that $E_{\varepsilon'} \subseteq E_\varepsilon$ if $\varepsilon' \leq \varepsilon$. Let us take the "good set" of points

$$G_\varepsilon \equiv \left\{ x \in E_\varepsilon : |u(x)| + |\mathrm{D}u(x)| \leq \frac{1}{\varepsilon} \right\},$$

and note that $|\Omega \setminus G_\varepsilon| \to 0$ as $\varepsilon \to 0$.

We now apply the convexity of $F$ in the last variable, together with $F \geq 0$, to estimate

$$\begin{aligned} \mathbb{E}[u_j] &\geq \int_{G_\varepsilon} F(x, u_j, \mathrm{D}u_j)\,\mathrm{d}x \\ &\geq \int_{G_\varepsilon} F(x, u_j, \mathrm{D}u)\,\mathrm{d}x + \int_{G_\varepsilon} \partial_\xi F(x, u_j, \mathrm{D}u) \cdot (\mathrm{D}u_j - \mathrm{D}u)\,\mathrm{d}x. \end{aligned}$$

Now we inspect each of the terms separately. By construction of the set $G_\varepsilon$, we have

$$\lim_{j\to\infty} \int_{G_\varepsilon} F(x, u_j, \mathrm{D}u)\,\mathrm{d}x = \int_{G_\varepsilon} F(x, u, \mathrm{D}u)\,\mathrm{d}x.$$

Moreover, since $\partial_\xi F(x, u_j, \mathrm{D}u) \to \partial_\xi F(x, u, \mathrm{D}u)$ uniformly on $G_\varepsilon$ and $\mathrm{D}u_j \rightharpoonup \mathrm{D}u$ in $L^p$, we have

$$\lim_{j\to\infty} \int_{G_\varepsilon} \partial_\xi F(x, u_j, \mathrm{D}u) \cdot (\mathrm{D}u_j - \mathrm{D}u)\,\mathrm{d}x = 0.$$

Thus we see that

$$\lim_{j\to\infty} \mathbb{E}[u_j] \geq \int_{G_\varepsilon} F(x, u, \mathrm{D}u)\,\mathrm{d}x.$$

Since $F \geq 0$ and $\varepsilon > 0$ is arbitrary, the conclusion follows from the Monotone Convergence Theorem. $\square$

**Corollary 2.18** (Existence of minimizers)**.** *Let $F$ be as in Theorem 2.17 and assume in addition that, for some $p \in (1, \infty)$ and $c_1 > 0, c_2 \geq 0$, we have*

$$F(x, u, \xi) \geq c_1 |\xi|^p - c_2 \qquad \textit{for all } (x, u, \xi) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$$

*Then for each $g \in C^\infty(\overline{\Omega}, \mathbb{R}^m)$, $\mathbb{E}$ has a minimizer in $W^{1,p}_g(\Omega, \mathbb{R}^m) \equiv g + W^{1,p}_0(\Omega, \mathbb{R}^m)$.*

**Proof.** We want to apply Theorem 2.16 with $\mathbb{X} = W^{1,p}(\Omega, \mathbb{R}^m)$ and $\mathbb{Y} = W^{1,p}_g(\Omega, \mathbb{R}^m)$; the assumption $p \in (1, \infty)$ guarantees that $\mathbb{X}$ is reflexive. Note that, by Mazur's Theorem, $\mathbb{Y}$ is a weakly closed space. Hence, by Theorem 2.17, it suffices to check the coercivity of $\mathbb{E}$.

Clearly our hypotheses implies the coercivity inequality $\mathbb{E}[v] \geq c_1 \|\mathrm{D}v\|^p_{L^p(\Omega)} - c_2|\Omega|$. By the Poincaré inequality, for each $v \in W^{1,p}_g(\Omega, \mathbb{R}^m)$ we have

$$\|v\|_{L^p} \leq \|v - g\|_{L^p} + \|g\|_{L^p} \leq C\|\mathrm{D}v - \mathrm{D}g\|_{L^p} + \|g\|_{L^p} \leq C(\|\mathrm{D}v\|_{L^p} + \|g\|_{W^{1,p}}).$$

Thus $\|v\|_{W^{1,p}} \leq C(\|\mathrm{D}v\|_{L^p} + 1)$ and so $\mathbb{E}[v] \to \infty$ when $\|v\|_{W^{1,p}} \to \infty$, as wished. $\square$

The assumptions in Corollary 2.10 are fairly optimal. We first give a classical example which shows that minimizers need not exist can fail even if coercivity fails at single point:

**Example 2.19** (Weierstrass)**.** Take $n = m = 1$, let $F\colon (0,1)\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be defined by $F(x,u,\xi) = x\xi^2$ and consider the problem

$$\inf\left\{\mathbb{E}[u] : u \in W^{1,2}(0,1) \text{ and } u(0) = 1, u(1) = 0\right\}.$$

Note that this example verifies all of the conditions of Corollary 2.10, except that the coercivity fails at a single point. We claim that the above infimum is zero: indeed,

$$u_k(x) = \begin{cases} 1 & \text{if } x \in [0, \frac{1}{k}], \\ -\frac{\log x}{\log k} & \text{if } x \in (\frac{1}{k}, 1], \end{cases}$$

satisfies $\mathbb{E}[u_k] = \frac{1}{\log k} \to 0$ as $k \to \infty$. Clearly there is no admissible function with $\mathbb{E}[u] = 0$ (otherwise we would have that $u$ is constant), and so the infimum is not attained.

The next example shows that, in the absence of convexity, minimizers need not exist:

**Example 2.20** (Bolza)**.** Let again $n = m = 1$, take $F(x,u,\xi) = (\xi^2-1)^2 + u^4$ and consider the problem

$$\inf\left\{\mathbb{E}[u] : u \in W^{1,4}(0,1) \text{ and } u(0) = 0, u(1) = 0\right\}.$$

We claim that this infimum is zero, and so there is no minimizer, since $\mathbb{E}[u] = 0$ forces simultaneously $u = 0$ and $|u'| = 1$, which is impossible. To see this, we simply construct a sequence $(u_k) \in W^{1,4}_0(0,1)$ with $|u_k'| = 1$ and $|u_k| \leq 1/(2k)$: this can be achieved by setting

$$u_k(x) = \begin{cases} x & \text{if } x \in [0, \frac{1}{2k}], \\ \frac{1}{k} - x & \text{if } x \in [\frac{1}{2k}, \frac{1}{k}], \end{cases}$$

and then extending $u_k$ periodically. We then have

$$0 \leq \mathbb{E}[u_k] \leq \frac{1}{(2k)^4} \to 0,$$

as wished.

Example 2.20 is scalar and indeed, in the scalar case, convexity is essentially a necessary condition for existence of minimizers. In the vectorial case, however, this is not the case. This is illustrated most clearly in the case of quadratic energies, such as the ones we discussed in Section 2.1; we leave for the reader the next simple exercise:

**Exercise 2.21.** Define $F\colon \mathbb{R}^{m\times n}\to\mathbb{R}$ to be the quadratic energy

$$F(\xi) \equiv \langle A\xi, \xi\rangle = A^{ij}_{\alpha\beta}\xi^\alpha_i\xi^\beta_j$$

where $A$ is $x$-independent. Show that $F$ is convex if and only if $F \geq 0$, i.e. if and only if it satisfies the Legendre condition with $\lambda = 0$. Conclude that the quadratic form in Example 2.5 is not convex.

In the general vectorial case, the correct convexity-type condition is called *quasiconvexity*, but this is outside the scope of these notes. Nonetheless, we refer the reader to Problem 2.4 for a general family of vectorial non-convex examples where one can still prove existence of minimizers.

### 2.3 Problems for Section 2

**Problem 2.1.** Generalize Proposition 2.8 to variable coefficients $A = A(x) \in L^\infty(\Omega)$ by using Lebesgue's differentiation theorem and choosing $\varphi$ suitably. For simplicity you may assume that $A$ is symmetric in $(i, j)$ and in $(\alpha, \beta)$.

**Problem 2.2** (Korn's inequality)**.** Use Gårding's inequality to deduce *Korn's inequality*:

$$\lambda \int_\Omega |\mathrm{D}u|^2 \,\mathrm{d}x \le \int_\Omega \left| \frac{\mathrm{D}u + (\mathrm{D}u)^T}{2} \right|^2 \mathrm{d}x \quad \text{for all } u \in W_0^{1,2}(\Omega, \mathbb{R}^n).$$

Then give a direct proof of Korn's inequality by writing

$$2\left| \frac{\mathrm{D}u + (\mathrm{D}u)^T}{2} \right|^2 = |\mathrm{D}u|^2 + (\operatorname{div} u)^2 + L(\mathrm{D}u, \mathrm{D}u),$$

where $\int_\Omega L(\mathrm{D}u, \mathrm{D}u)\,\mathrm{d}x = 0$ for all $u \in W_0^{1,2}(\Omega, \mathbb{R}^n)$.

**Problem 2.3** (Mazur's theorem and lsc)**.** Recall Mazur's Theorem: if $X$ is a normed space with $u_j \rightharpoonup u$ in $X$, then there is a sequence $(v_k)$ contained in the convex hull of $(u_j)_{j=1}^\infty$ such that $v_k \to u$ in $X$. Use Mazur's Theorem and Fatou's Lemma to give a different proof of Tonelli's Theorem in the case where $F$ is $u$-independent.

**Problem 2.4** (Polyconvexity)**.** Let $2 = n = m < p$.

(i) Use (2.4) to show that if $u_j \rightharpoonup u$ in $W^{1,p}(\Omega, \mathbb{R}^2)$ then $\det \mathrm{D}u_j \rightharpoonup \det \mathrm{D}u$ in $L^{p/2}(\Omega)$.

(ii) Follow the strategy of Tonelli's theorem to prove that if $F \colon \Omega \times \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \times \mathbb{R} \to \mathbb{R}$ is smooth, bounded from below, and $(\xi, \delta) \mapsto F(x, u, \xi, \delta)$ is convex for all $(x, u)$, then

$$\mathbb{E}[u] = \int_\Omega F(x, u, \mathrm{D}u, \det \mathrm{D}u)\,\mathrm{d}x$$

is sequentially weakly lower semicontinuous in $W^{1,p}$. Such integrands are called *polyconvex* [2].

(iii) Give an example of $F = F(\xi, \delta)$ as in (ii) which is not convex and for which one can prove existence of minimizers in $W_g^{1,p}(\Omega, \mathbb{R}^2)$ for $g \in C^\infty(\overline{\Omega}, \mathbb{R}^2)$.

# 3 $L^2$-theory for linear elliptic systems

## 3.1 The model case: harmonic functions

Recall that a weakly harmonic function $u \in W^{1,2}(\Omega)$ is a function for which

$$\mathbb{E}'[u](\varphi) = \int_\Omega \langle \mathrm{D}u, \mathrm{D}\varphi \rangle \,\mathrm{d}x = 0 \quad \forall \varphi \in W_0^{1,2}(\Omega), \tag{3.1}$$

where $\mathbb{E}$ is the Dirichlet energy from Example 0.4. In this subsection we sketch a proof that weakly harmonic functions are smooth. There are of course many different ways of proving this result: for instance, one can give a short proof by using the fact that harmonic functions have the mean value property. The proof we give here is instead more complicated but also much more robust: we will be able to generalize it without too much difficulty to very general linear PDEs below. The following is the crucial step in the approach we take here:

**Theorem 3.1** (Caccioppoli's inequality)**.** *Let $u \in W^{1,2}(\Omega)$ be such that $\Delta u = 0$ weakly. For any ball $B_R(x_0) \subset \Omega$ we have*

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \le \frac{16}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x$$

*for all $u_0 \in \mathbb{R}$ and all $r < R$.*

**Proof.** Let $\eta \in C_c^\infty(B_R(x_0))$ be a cut-off function such that $0 \le \eta \le 1$, $\eta = 1$ in $B_r(x_0)$ and $|\mathrm{D}\eta| \le \frac{2}{R-r}$. We test the weak formulation of $\Delta u = 0$ with $\varphi = \eta^2(u-u_0) \in W_0^{1,2}(\Omega)$ to obtain

$$0 = \int_\Omega \langle \mathrm{D}u, \mathrm{D}\varphi \rangle \,\mathrm{d}x = \int_\Omega \eta^2 |\mathrm{D}u|^2 + (u-u_0) 2\eta \langle \mathrm{D}u, \mathrm{D}\eta \rangle \,\mathrm{d}x.$$

Using Hölder's inequality, we find

$$\begin{aligned}\int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x &\le \int_{B_R(x_0)} 2\eta |\mathrm{D}u||u-u_0||\mathrm{D}\eta| \,\mathrm{d}x \\ &\le \left(\int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(4\int_{B_R(x_0)} |u-u_0|^2 |\mathrm{D}\eta|^2 \,\mathrm{d}x\right)^{\frac{1}{2}}\end{aligned}$$

and so, rearranging, we obtain

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \le \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x \le \frac{16}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x,$$

as claimed. □

Since derivatives of smooth harmonic functions are themselves harmonic, by applying Theorem 3.1 in suitable annuli, one can deduce *a priori* estimates for higher order derivatives:

**Corollary 3.2.** *Let $u \in C^\infty(\Omega)$ be a harmonic function, i.e. $\Delta u = 0$. For all $k \ge 1$ and all balls $B_R(x_0) \subset \Omega$, we have*

$$\int_{B_{\frac{R}{2}}(x_0)} |\mathrm{D}^k u|^2 \,\mathrm{d}x \le C(k,r) \int_{B_R(x_0)} |u|^2 \,\mathrm{d}x.$$

**Exercise 3.3.** Prove Corollary 3.2.

We call the estimates in Corollary 3.2 *a priori* since, in the statement, we are already assuming that $u \in C^\infty$, which is what we want to prove! However, since the Laplace equation is linear with constant coefficients, the a priori estimate from Corollary 3.2, together with the Sobolev embedding theorem, implies the following:

**Theorem 3.4.** *Let $u \in W^{1,2}(\Omega)$ be a weakly harmonic function. Then $u \in C^\infty(\Omega)$.*

**Proof.** By Theorem 1.16, it suffices to prove that $u \in W^{k,2}(\Omega)$ for all $k \geq 1$. Consider the mollified functions $u_\varepsilon \equiv u * \rho_\varepsilon$, which are defined on $\Omega_\varepsilon \equiv \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\}$. We claim that, on $\Omega_\varepsilon$, we have $\Delta u_\varepsilon = 0$: indeed, for $\varphi \in C_c^\infty(\Omega_\varepsilon)$, we have

$$\int_\Omega \langle \mathrm{D}u_\varepsilon, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle \mathrm{D}u * \rho_\varepsilon, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle \mathrm{D}u, \mathrm{D}\varphi * \rho_\varepsilon\rangle \,\mathrm{d}x = \int_\Omega \langle \mathrm{D}u, \mathrm{D}(\varphi * \rho_\varepsilon)\rangle \,\mathrm{d}x = 0,$$

since $\varphi * \rho_\varepsilon \in C_c^\infty(\Omega)$. Thus $u_\varepsilon$ is weakly harmonic on $\Omega_\varepsilon$, but since it is smooth it is then actually harmonic. Now let us fix a ball $B_R(x_0) \Subset \Omega$. Since

$$\int_{B_R(x_0)} |u_\varepsilon|^2 \,\mathrm{d}x \leq \int_{B_{R+\varepsilon}(x_0)} |u|^2 \,\mathrm{d}x,$$

by Corollary 3.2 we see that $(u_\varepsilon)_{\varepsilon>0} \subset W^{k,2}(B_{R/2}(x_0))$ is bounded. Since $u_\varepsilon \to u$ in $W^{1,2}(B_{R/2}(x_0))$, we deduce that also $u \in W^{k,2}(B_{R/2}(x_0))$: this can be seen from the fact that $(u_\varepsilon)$ is weakly pre-compact in $W^{k,2}(B_{R/2}(x_0))$ and so, by uniqueness of limits, $u$ is the only accumulation point of the sequence. □

We defined weakly harmonic functions through the variational formulation (3.1). This definition is natural and it extends in a straightforward way to critical points of the general variational integral (0.1). However, in the case of the Dirichlet energy, one can consider an even weaker notion of critical point, by integrating by parts again in (3.1): let us say that $u \in L^1_{\mathrm{loc}}(\Omega)$ is *harmonic in the sense of distributions* if

$$\int_\Omega u \Delta\varphi \,\mathrm{d}x = 0 \qquad \forall \varphi \in C_c^\infty(\Omega).$$

The approach outlined in this section can be used to give a proof of Weyl's lemma: any $L^1_{\mathrm{loc}}$ function which is harmonic in the sense of distributions is smooth, cf. Problem 3.1.

## 3.2 Regularity for linear elliptic systems

Having studied in some detail the case of harmonic functions, we now turn to the general divergence-form linear elliptic systems

$$\partial_\alpha(A_{ij}^{\alpha\beta}(x)\partial_\beta u^j) = \partial_\alpha F_i^\alpha, \qquad i = 1, \dots, m, \tag{3.2}$$

which we encountered in Section 2.1. We refer the reader to Problems 3.2 and 3.3 for more general results in the case where the right-hand side is not in the divergence form.

We begin with the general version of Theorem 3.1:

**Theorem 3.5** (Caccioppoli's inequality)**.** *Let $F \in L^2(\Omega, \mathbb{R}^{m\times n})$ and let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.2)*. Assume that:*

(i) *either $A_{ij}^{\alpha\beta} \in L^\infty(\Omega)$ satisfy the Legendre condition* (2.3)*;*

(ii) *or* $A_{ij}^{\alpha\beta} \in C^0(\overline{\Omega})$ *satisfy the Legendre–Hadamard condition* (2.5).

*For any ball* $B_R(x_0) \subset \Omega$, *with* $R$ *small enough in case* (ii), *and any* $u_0 \in \mathbb{R}^m$, *we have*

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \le C \left( \frac{1}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x + \int_{B_R(x_0)} |F|^2 \,\mathrm{d}x \right)$$

*for all* $r < R$. *The constant* $C$ *depends on* $\lambda, \Lambda \equiv \|A\|_\infty$ *and on the modulus of continuity of* $A$ *in case* (ii).

**Proof.** We will only deal with case (i), which is slightly simpler. As in the proof of Theorem 3.1, the idea is to test the weak formulation of (3.2), that is

$$\int_\Omega \langle A\mathrm{D}u, \mathrm{D}\varphi \rangle \,\mathrm{d}x = \int_\Omega \langle F, \mathrm{D}\varphi \rangle \,\mathrm{d}x,$$

with $\varphi = (u-u_0)\eta^2$, where $\eta \in C_c^\infty(B_R(x_0))$ is a cut-off function such that $0 \le \eta \le 1$, $\eta = 1$ in $B_r(x_0)$ and $|\mathrm{D}\eta| \le \frac{2}{R-r}$. Since $\mathrm{D}\varphi = \eta^2 \mathrm{D}u + \eta(u-u_0)\otimes \mathrm{D}\eta$, we obtain

$$\begin{aligned}\int_{B_R(x_0)} \eta^2 \langle A\mathrm{D}u, \mathrm{D}u \rangle \,\mathrm{d}x + 2\int_{B_R(x_0)} \eta \langle A\mathrm{D}u, (u-u_0)\otimes \mathrm{D}\eta \rangle \,\mathrm{d}x = \\ = \int_{B_R(x_0)} \eta^2 \langle F, \mathrm{D}u \rangle \,\mathrm{d}x + 2\int_{B_R(x_0)} \eta \langle F, (u-u_0)\otimes \mathrm{D}\eta \rangle \,\mathrm{d}x.\end{aligned}$$

We now treat each of the terms separately. The main term is the first one, which we can estimate using (2.3):

$$\lambda \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x \le \int_{B_R(x_0)} \eta^2 \langle A\mathrm{D}u, \mathrm{D}u \rangle \,\mathrm{d}x.$$

As for the second term, we use Cauchy–Schwarz and Young's inequality, together with the properties of $\eta$:

$$\begin{aligned} 2\left| \int_{B_R(x_0)} \eta \langle A\mathrm{D}u, (u-u_0)\otimes \mathrm{D}\eta \rangle \,\mathrm{d}x \right| &\le 2\int_{B_R(x_0)} |A| (\eta|\mathrm{D}u|)(|u-u_0||\mathrm{D}\eta|) \,\mathrm{d}x \\ &\le \varepsilon\Lambda \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x + \frac{4\Lambda}{\varepsilon(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x.\end{aligned}$$

The terms on the right-hand side are dealt with similarly: since $\eta \le 1$,

$$\begin{aligned} &\int_{B_R(x_0)} \eta^2 \langle F, \mathrm{D}u \rangle \,\mathrm{d}x \le \varepsilon\Lambda \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x + \frac{1}{4\varepsilon\Lambda} \int_{B_R(x_0)} |F|^2 \,\mathrm{d}x, \\ &2\int_{B_R(x_0)} \eta \langle F, (u-u_0)\otimes \mathrm{D}\eta \rangle \,\mathrm{d}x \le 4\int_{B_R(x_0)} |F|^2 \,\mathrm{d}x + \frac{4}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x.\end{aligned}$$

Thus, combining the above estimates, we obtain

$$\begin{aligned} \lambda \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x &\le 2\varepsilon\Lambda \int_{B_R(x_0)} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x \\ &\quad + \frac{4(\Lambda/\varepsilon + 1)}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u-u_0|^2 \,\mathrm{d}x + \left(4 + \frac{1}{4\varepsilon\Lambda}\right) \int_{B_R(x_0)} |F|^2 \,\mathrm{d}x.\end{aligned}$$

We choose $\varepsilon = \lambda/(4\Lambda)$ in order to reabsorb the $\eta^2|\mathrm{D}u|^2$ term and obtain the conclusion. □

**Exercise 3.6.** Prove Theorem 3.5 in case (ii) using Gårding's inequality in Corollary 2.11. In fact, recall from that proof that, if $R$ is sufficiently small, then

$$\lambda_0 \int_{B_R(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \int_{B_R(x_0)} \langle A\mathrm{D}u, \mathrm{D}u\rangle \,\mathrm{d}x$$

for all $u \in W_0^{1,2}(B_R(x_0), \mathbb{R}^m)$.

**Remark 3.7.** Similarly to Remark 2.12, the continuity assumption in case (ii) cannot be dropped [7].

We now want to proceed in a similar way as we did for harmonic functions, i.e. by proving higher order estimates and then iterating this result. There is, however, a key difference: recall that Corollary 3.2 is a consequence of the fact that the derivative of a harmonic function is harmonic, since $\partial_\alpha$ and $\Delta$ commute. However, in the general $x$-dependent case we consider here, this is no longer true, and differentiating the equation would lead to extra terms on the right-hand side. Thus, instead of differentiating, we use Nirenberg's method by considering difference quotients $\partial_{h,\alpha}$ and proving estimates uniform in $h$.

**Theorem 3.8** ($W^{2,2}$-estimate)**.** *Let $F \in W^{1,2}(\Omega, \mathbb{R}^{m\times n})$ and let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.2)*. Assume that $A_{ij}^{\alpha\beta} \in C^{0,1}(\Omega)$ satisfy the Legendre–Hadamard condition. Then we have $u \in W^{2,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ and, for any $\Omega' \Subset \Omega$,*

$$\|\mathrm{D}^2 u\|_{L^2(\Omega')} \leq C(\Omega, \Omega', A)\left(\|u\|_{L^2(\Omega)} + \|F\|_{W^{1,2}(\Omega)}\right).$$

**Proof.** By a covering argument it is enough to prove the result when $\Omega' = B_R(x_0)$ and $\Omega = B_{4R}(x_0)$. We also assume that $x_0 = 0$ without loss of generality, and so we suppress the center of the balls from the notation.

Recall that (3.2) means that, for any $\varphi \in C_c^\infty(\Omega, \mathbb{R}^m)$, we have

$$\int_\Omega \langle A\mathrm{D}u, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle F, \mathrm{D}\varphi\rangle.$$

For $h$ small enough (depending on $\varphi$), we can consider as another test function $\tau_{-h,\alpha}\varphi$, to get

$$\int_\Omega \langle (\tau_{h,\alpha}A)\mathrm{D}(\tau_{h,\alpha}u), \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle \tau_{h,\alpha}F, \mathrm{D}\varphi\rangle \,\mathrm{d}x,$$

since of course D and $\tau_{h,\alpha}$ commute, and where we used also Exercise 1.12. Subtracting the previous two identities and dividing by $h$, we get

$$\int_\Omega \langle \tau_{h,\alpha}A\,\mathrm{D}(\partial_{h,\alpha}u), \mathrm{D}\varphi\rangle \,\mathrm{d}x + \int_\Omega \langle \partial_{h,\alpha}A\,\mathrm{D}u, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_\Omega \langle \partial_{h,\alpha}F, \mathrm{D}\varphi\rangle \,\mathrm{d}x.$$

Thus, we see that $\partial_{h,\alpha}u$ is a weak solution of a system identical to (3.2): we can thus apply Theorem 3.5 to get

$$\int_{B_R} |\partial_{h,\alpha}\mathrm{D}u|^2 \,\mathrm{d}x \lesssim \frac{1}{R^2}\int_{B_{2R}} |\partial_{h,\alpha}u|^2 \,\mathrm{d}x + \int_{B_{2R}} |\partial_{h,\alpha}A|^2 |\mathrm{D}u|^2 \,\mathrm{d}x + \int_{B_{2R}} |\partial_{h,\alpha}F|^2 \,\mathrm{d}x.$$

Applying Proposition 1.13, we see that all of the terms on the right-hand side are bounded as $h \to 0$, and so $\mathrm{D}u \in W^{1,2}(B_R, \mathbb{R}^{m\times n})$. Thus, sending $h \to 0$, and summing the resulting estimate over $\alpha$, we obtain

$$\int_{B_R} |\mathrm{D}^2 u|^2 \,\mathrm{d}x \leq C(\lambda, \Lambda) \left( \frac{1}{R^2} \int_{B_{2R}} |\mathrm{D}u|^2 \,\mathrm{d}x + \|\mathrm{D}A\|_{L^\infty}^2 \int_{B_{2R}} |\mathrm{D}u|^2 \,\mathrm{d}x + \int_{B_{2R}} |\mathrm{D}F|^2 \,\mathrm{d}x \right)$$
$$\leq C(R, A) \left( \int_{B_{4R}} |u|^2 + \int_{B_{2R}} |F|^2 \,\mathrm{d}x + \int_{B_{2R}} |\mathrm{D}F|^2 \,\mathrm{d}x \right),$$

where the last line follows by applying Theorem 3.5 again. □

As before, we can iterate Theorem 3.8 in a straight-forward manner to obtain estimates of arbitrarily high order:

**Theorem 3.9** (Higher order regularity)**.** *Let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.2)*. Assume that $A_{ij}^{\alpha\beta} \in C^{k,1}(\Omega)$ satisfy the Legendre–Hadamard condition for some $k \geq 0$. If $F \in W^{k+1,2}(\Omega)$ then $u \in W^{k+2,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, and for every $\Omega' \Subset \Omega$ we have*

$$\|\mathrm{D}^{k+2} u\|_{L^2(\Omega')} \leq C(k, \Omega', \Omega, \lambda, \|A\|_{C^{k,1}}) \left( \|u\|_{L^2(\Omega)} + \|F\|_{W^{k+1,2}(\Omega)} \right).$$

**Proof.** See Problem 3.5. □

As a consequence of Theorem 3.9 we see that, when the data of (3.2) is smooth, so is the solution.

**Corollary 3.10.** *Let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.2)*. Assume that $A$ satisfies the Legendre–Hadamard condition, and that $A_{ij}^{\alpha\beta}, F_i^\alpha \in C^\infty(\Omega)$. Then $u \in C^\infty(\Omega, \mathbb{R}^m)$.*

**Proof.** By Theorem 3.9, $u \in W^{k,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ for all $k \geq 0$, and so by the Sobolev embedding theorem we have $u \in C^\infty(\Omega)$. □

### 3.3 Decay estimates

We conclude this section by returning to the case of systems with constant coefficients; we will prove in particular two simple decay estimates that will play a very important role in the next section. We will always consider the *homogeneous* problem

$$\partial_\alpha (A_{ij}^{\alpha\beta} \partial_\beta u^j) = 0, \quad i = 1, \ldots, m \tag{3.3}$$

where moreover

$$A \text{ is constant and satisfies the Legendre–Hadamard condition.} \tag{3.4}$$

**Remark 3.11.** As motivation for the estiamtes in this section, let us inspect Caccioppoli's inequality. There are two things which are yet to take advantage of:

(i) we have freedom in choosing $u_0 \in \mathbb{R}^m$, and so far we have only taken $u_0 = 0$;

(ii) the integral of $|u - u_0|^2$ on the right-hand side is over an annulus $B_R(x_0) \setminus B_r(x_0)$, while so far we have only applied the (weaker) estimate

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \frac{C}{(R-r)^2} \int_{B_R(x_0)} |u - u_0|^2 \,\mathrm{d}x.$$

We will see in this section that both points above lead to improved information.

Let us begin with point (i) in the above remark. Clearly Caccioppoli's estimate is strongest when $u_0$ is chosen to minimize the right-hand side. The next exercise shows that the optimal choice of $u_0$ is given by the average of $u$:

**Exercise 3.12.** Let $1 \leq p < \infty$ and consider, for $u \in L^p(\Omega, \mathbb{R}^m)$, the problem

$$\inf_{u_0 \in \mathbb{R}^m} \int_\Omega |u(x) - u_0|^p \,\mathrm{d}x.$$

(i) Show that, when $p = 2$, the infimum is attained at $(u)_\Omega \equiv \frac{1}{|\Omega|} \int_\Omega u \,\mathrm{d}x$.

(ii) Show that, for general $p$, we still have for any $u_0 \in \mathbb{R}^m$ the inequality

$$\int_\Omega |u - (u)_\Omega|^p \,\mathrm{d}x \leq 2^p \int_\Omega |u - u_0|^p \,\mathrm{d}x.$$

Taking $u_0$ to be an average will allow us to use Poincaré's inequality. Recall that, when $\Omega = B_r(x_0)$, we write $(u)_\Omega \equiv (u)_{x_0,r}$. We then have:

**Lemma 3.13** (Decay estimates for systems with constant coefficients)**.** *Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.3)*, where $A$ satisfies* (3.4)*. We have*

$$\int_{B_r(x_0)} |u|^2 \,\mathrm{d}x \leq C\left(\frac{r}{R}\right)^n \int_{B_R(x_0)} |u|^2 \,\mathrm{d}x, \tag{3.5}$$

$$\int_{B_r(x_0)} |u - (u)_{x_0,r}|^2 \,\mathrm{d}x \leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u - (u)_{x_0,R}|^2 \,\mathrm{d}x, \tag{3.6}$$

*for any $B_r(x_0) \subset B_R(x_0) \Subset \Omega$, where $C = C(n, \lambda, \Lambda)$.*

**Proof.** Let us first prove (3.5). By rescaling and translating (i.e. by considering $u_{x_0,R} \equiv u(x_0 + R\cdot)$ instead of $u$) we can assume that $x_0 = 0, R = 1$. Let $k$ be an integer such that $k > \frac{n}{2}$, i.e. such that $W^{k,2} \subset C^0$. By Caccioppoli's inequality, we have

$$\int_{B_{\frac{1}{2}}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq C(\lambda, \Lambda) \int_{B_1} |u|^2 \,\mathrm{d}x.$$

Since the coefficients are constant, each derivative of $u$ also solves (2.1), and so by iterating this inequality (or alternatively by applying Theorem 3.9) we see that

$$\|u\|_{L^\infty(B_{2^{-k}})} \leq C\|u\|_{W^{k,2}(B_{2^{-k}})} \leq C(\lambda, \Lambda)\|u\|_{L^2(B_1)}.$$

Thus, if $r \leq 2^{-k}$, then

$$\int_{B_r} |u|^2 \,\mathrm{d}x \leq Cr^n \|u\|^2_{L^\infty(B_{2^{-k}})} \leq Cr^n \int_{B_1} |u|^2 \,\mathrm{d}x.$$

This is the only interesting case, because if $r > 2^{-k}$ then the inequality holds trivially with $C = (2^k)^n$, since $\int_{B_r} |u|^2 \,\mathrm{d}x \leq \int_{B_1} |u|^2 \,\mathrm{d}x$.

We now prove (3.6), which follows by applying the previous inequality to the derivatives of $u$, which are also solutions of (2.1), since $A$ is constant. Let us first assume that $r \leq \frac{R}{2}$; as before, this is the only interesting case. Applying the Poincaré inequality (Corollary 1.20), (3.5) and Caccioppoli's inequality, we have

$$\int_{B_r(x_0)} |u - (u)_{x_0,r}|^2 \,\mathrm{d}x \leq Cr^2 \int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x$$

$$
\begin{aligned}
&\leq Cr^2\left(\frac{r}{R}\right)^n \int_{B_{R/2}(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \\
&\leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u-(u)_{x_0,R}|^2 \,\mathrm{d}x.
\end{aligned}
$$

When $r \geq \frac{R}{2}$, we can simply estimate

$$
\int_{B_r(x_0)} |u-(u)_{x_0,r}|^2 \,\mathrm{d}x \leq \int_{B_r(x_0)} |u-(u)_{x_0,R}|^2 \,\mathrm{d}x \leq 2^{n+2}\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |u-(u)_{x_0,r}|^2 \,\mathrm{d}x,
$$

where in the first inequality we used Exercise 3.12. □

**Exercise 3.14.** Fill in the details for the "scaling argument" used in the proof above.

Let us now return to Remark 3.11, in particular to point (ii). We will use both the fact that we should choose $u_0$ to be the average of $u$ over the annulus, together with the fact that the integral is over an annulus rather than a ball.

**Lemma 3.15** (Widman's hole-filing technique)**.** *Let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (3.3)*, where $A$ satisfies* (3.4)*. Then there is an exponent $\alpha \in (0,1)$ such that*

$$
\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \leq C\left(\frac{r}{R}\right)^{\alpha} \int_{B_R(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x,
$$

*whenever $B_r(x_0) \subset B_R(x_0) \Subset \Omega$, where $C = C(n, \lambda, \Lambda)$.*

**Proof.** As before, by replacing $u$ with $u_{x_0,R} \equiv u(x_0 + R\cdot)$, we can assume that $R = 1, x_0 = 0$. Applying Caccioppoli's inequality with $u_0 = (u)_{B_1 \setminus B_{1/2}}$, together with Poincaré's inequality, we get

$$
\int_{B_{1/2}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \frac{C}{R^2} \int_{B_1 \setminus B_{1/2}} |u-u_0|^2 \,\mathrm{d}x \leq C_0 \int_{B_1 \setminus B_{1/2}} |\mathrm{D}u|^2 \,\mathrm{d}x.
$$

We now add $C \int_{B_{1/2}} |\mathrm{D}u|^2 \,\mathrm{d}x$ to both sides (i.e. we "fill the hole" on the integral in the right-hand side), to find that

$$
(C_0+1) \int_{B_{1/2}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq C_0 \int_{B_1} |\mathrm{D}u|^2 \,\mathrm{d}x
$$

or, rearranging and writing $\theta \equiv \frac{C_0}{C_0+1} < 1$, we obtain the decay inequality

$$
\int_{B_{1/2}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \theta \int_{B_1} |\mathrm{D}u|^2 \,\mathrm{d}x.
$$

This inequality is striking: it asserts that there is always at most a *universal* fraction of the energy of $u$ concentrated on a ball of half the radius. We iterate this estimate $k$ times to get

$$
\int_{B_{2^{-k}}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \theta^k \int_{B_1} |\mathrm{D}u|^2 \,\mathrm{d}x.
$$

Now fix $r > 0$, and let $k$ is the unique integer such that $2^{-k-1} \leq r < 2^{-k}$. Writing $\alpha \equiv \log_2(1/\theta) > 0$, we have $\theta^k = 2^{k \log_2 \theta} = (2^{-k})^\alpha \leq (2r)^\alpha$, and so

$$
\int_{B_r} |\mathrm{D}u|^2 \,\mathrm{d}x \leq \int_{B_{2^{-k}}} |\mathrm{D}u|^2 \,\mathrm{d}x \leq 2^\alpha r^\alpha \int_{B_1} |\mathrm{D}u|^2 \,\mathrm{d}x
$$

as wished. $\square$

**Remark 3.16.** Since the proof only relies on the Caccioppoli inequality, the result also holds in the case where $A_{ij}^{\alpha\beta} \in L^\infty(\Omega)$ satisfies the Legendre condition. It is special interesting in that case, since if $A$ is constant Lemma 3.13 provides a stronger result.

As we will see in the next section, when $n = 2$ Lemma 3.15 already implies that $u \in C^{0,\alpha/2}$. This type of estimate will also be useful in Section 5.1, when we study nonlinear problems in the plane. For now, we simply note the following:

**Corollary 3.17.** *Assume* (3.4) *holds and let* $u\colon \mathbb{R}^n \to \mathbb{R}^m$ *be an entire solution of* (3.3) *with*

$$\int_{\mathbb{R}^n} |\mathrm{D}u|^2 \,\mathrm{d}x < \infty.$$

*Then* $u$ *is constant.*

**Proof.** Indeed, for any $0 < r < R$, we have

$$\int_{B_r(0)} |\mathrm{D}u|^2 \le C\left(\frac{r}{R}\right)^\alpha \int_{B_R(0)} |\mathrm{D}u|^2 \,\mathrm{d}x \le C\left(\frac{r}{R}\right)^\alpha \int_{\mathbb{R}^n} |\mathrm{D}u|^2 \,\mathrm{d}x \to 0$$

as $R \to \infty$. Thus $\int_{B_r(0)} |\mathrm{D}u|^2 \,\mathrm{d}x = 0$, for any $r > 0$, so $u$ is constant. $\square$

In particular, when $n = 2$, from the Caccioppoli inequality we obtain

$$\int_{B_R(0)} |\mathrm{D}u|^2 \le \frac{C}{R^2} \int_{B_{2R}(0)} |u|^2 \,\mathrm{d}x \le C\|u\|^2_{L^\infty(\mathbb{R}^2)}.$$

Thus Corollary 3.17 implies the Liouville theorem: if $u \in L^\infty(\mathbb{R}^2, \mathbb{R}^m)$ is an entire solution of (3.2) then $u$ is constant. In fact, we have the following more general version of this result:

**Theorem 3.18** (Liouville's Theorem)**.** *Assume* (3.4) *holds and let* $u\colon \mathbb{R}^n \to \mathbb{R}^m$ *be an entire solution of* (3.3) *such that*

$$|u(x)| \le C(1 + |x|^k) \quad \text{for all } x \in \mathbb{R}^n,$$

*for some* $k \in \mathbb{N}$*. Then* $u$ *is a polynomial of degree at most* $k$*.*

**Proof.** We will show that $\mathrm{D}^{k+1}u = 0$; since $u$ is smooth by Corollary 3.10, this implies[1] that $u$ is a polynomial of degree at most $k$. Fix $0 < r \le 2R$, and applying (3.5) and Caccioppoli's inequality $(k + 1)$-times, we obtain

$$\begin{aligned}
\int_{B_r(0)} |\mathrm{D}^{k+1}u|^2 \,\mathrm{d}x &\le C\left(\frac{r}{R}\right)^n \int_{B_R(0)} |\mathrm{D}^{k+1}u|^2 \,\mathrm{d}x \\
&\le C\left(\frac{r}{R}\right)^n \frac{1}{R^{2k+2}} \int_{B_{2^{k+1}R}(0)} |u|^2 \,\mathrm{d}x \\
&\le C\left(\frac{r}{R}\right)^n \frac{R^{2k+n}}{R^{2k+2}} = C\frac{r^n}{R^2}.
\end{aligned}$$

Thus, sending $R \to \infty$, we conclude that $\int_{B_r(0)} |\mathrm{D}^{k+1}u|^2 \,\mathrm{d}x = 0$. Since $r > 0$ is arbitrary, the conclusion follows. $\square$

[1] The same is true even for $L^1_{\mathrm{loc}}$ functions which satisfy $\mathrm{D}^{k+1}u = 0$ weakly, cf. Lemma 1.5 for the case $k = 0$.

### 3.4 Problems for Section 3

**Problem 3.1** (Weyl's lemma)**.** Use the strategy of Section 3.1 to give a proof of Weyl's lemma: any function $u \in L^2_{\mathrm{loc}}(\Omega)$ which is harmonic in the sense of distributions is smooth.

**Problem 3.2** (Poisson equation)**.** Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega)$ be a weak solution of

$$-\Delta u = f. \tag{3.7}$$

The goal of this problem is to prove that if $f \in L^2(\Omega)$ then $u \in W^{2,2}_{\mathrm{loc}}(\Omega)$.

(i) Integrate by parts to prove that, for a solution $u \in C^\infty_c(\Omega)$ of (3.7),

$$\int_\Omega |\mathrm{D}^2 u|^2 \,\mathrm{d}x = \int_\Omega |f|^2 \,\mathrm{d}x.$$

(ii) Now let $u \in C^\infty(\Omega)$ be a solution of (3.7). By deriving the equation for $\eta u$, where $\eta \in C^\infty_c(\Omega)$ is a cutoff function, show that for any $\Omega_0 \Subset \Omega_1 \Subset \Omega$ we have

$$\int_{\Omega_0} |\mathrm{D}^2 u|^2 \,\mathrm{d}x \le C(\Omega_0, \Omega_1) \int_{\Omega_1} \Big(|f|^2 + |u|^2 + |\mathrm{D}u|^2\Big) \,\mathrm{d}x.$$

(iii) Deduce from the above a priori estimate that, if $u \in W^{1,2}_{\mathrm{loc}}(\Omega)$ is a weak solution of (3.7), then $u \in W^{2,2}_{\mathrm{loc}}(\Omega)$.

(iv) Give a different proof of the fact that $u \in W^{2,2}_{\mathrm{loc}}(\Omega)$ by testing the weak formulation of (3.7) against $\partial_{-h,\alpha}(\eta^2 \partial_{h,\alpha} u)$.

**Problem 3.3** (Non-divergence form right-hand side)**.** Let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of the system

$$-\partial_\alpha(A^{\alpha\beta}_{ij}(x)\partial_\beta u^j) = f_i, \qquad i = 1, \ldots, m,$$

where $f \in L^2(\Omega, \mathbb{R}^m)$ and $A$ is as in Theorem 3.5.

(i) Whenever $B_R(x_0) \subset \Omega$ and $u_0 \in \mathbb{R}^m$, prove the Caccioppoli estimate

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \le C \left( \frac{1}{(R-r)^2} \int_{B_R(x_0)\setminus B_r(x_0)} |u - u_0|^2 \,\mathrm{d}x + R^2 \int_{B_R(x_0)} |f|^2 \,\mathrm{d}x \right).$$

[*Hint: reduce this to Theorem 3.5 by constructing $F^1 \in L^2(\Omega, \mathbb{R}^m)$ with $f_i = \partial_1 F^1_i$.*]

(ii) Using the test function $-\partial_{-h,\alpha}(\eta^2 \partial_{h,\alpha} u)$ and part (i), prove that $u \in W^{2,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$.

**Problem 3.4** (Boundary regularity)**.** Let $\Omega = B_1^+(0) \equiv B_1(0) \cap \mathbb{R}^n_+$, where

$$\mathbb{R}^n_+ \equiv \{x \in \mathbb{R}^n : x_n > 0\}.$$

For $g \in W^{1,2}(\Omega, \mathbb{R}^n)$, consider the weak solution $u \in W^{1,2}_0(\Omega)$ of the Dirichlet problem

$$\begin{cases} \Delta u = \operatorname{div} g & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

The goal of this problem is to show that $u \in W^{2,2}(B^+_{1/2}(0))$.

(i) Test the equation with $\partial_{-h,\alpha}(\eta^2 \partial_{h,\alpha} u)$ where $\alpha \neq n$ and $\eta \in C^\infty_c(\Omega)$, and obtain estimates showing that $\partial_\alpha u \in W^{1,2}(B^+_{1/2}(0))$.

(ii) Use the equation $\Delta u = \operatorname{div} g$ to deduce that $\partial_{nn} u \in W^{1,2}(B^+_{1/2}(0))$ as well.

**Problem 3.5** (Higher order regularity)**.** Use an induction argument to prove Theorem 3.9.

# 4 Schauder theory for linear elliptic systems

## 4.1 Hölder, Morrey and Campanato spaces

In this section we develop the classical perturbation theory of Schauder for linear elliptic systems in divergence form

$$\partial_\alpha(A_{ij}^{\alpha\beta}\partial_\beta u^j) = \partial_\alpha F_i^\alpha \quad \text{for } i = 1, \dots, m, \tag{4.1}$$

that we already studied in the previous two sections. In Section 3 we studied how the Sobolev regularity of the data $A, F$ is related to the regularity of the solution $u$. In Schauder theory one instead measures this regularity in a different scale of function spaces, which we begin by introducing.

**Definition 4.1.** For $\alpha \in (0, 1]$, a map $u\colon \Omega \to \mathbb{R}^m$ is said to be $\alpha$-Hölder continuous, and we write $u \in C^{0,\alpha}(\Omega, \mathbb{R}^m)$ if $[u]_{C^{0,\alpha}(\Omega,\mathbb{R}^m)} < \infty$, where

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

The space $C^{0,\alpha}(\Omega, \mathbb{R}^m)$ is a Banach space with the norm

$$\|u\|_{C^{0,\alpha}} \equiv \|u\|_{L^\infty} + [u]_{C^{0,\alpha}}$$

As usual, we say that $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ if any point in $\Omega$ has a neighborhood $\Omega'$ where $u \in C^{0,\alpha}(\Omega', \mathbb{R}^m)$. We also define for $k \in \mathbb{N}$ the spaces $C^{k,\alpha}(\Omega)$ in a similar way.

We refer the reader to Problem 4.1 for some basic properties of Hölder spaces.

In these notes, we will follow Campanato's approach to Schauder theory. For this approach, we will introduce the following spaces, which are naturally motivated by the two decay estimates of Lemma 3.13:

**Definition 4.2.** Let $\Omega(x_0, r) \equiv \Omega \cap B_r(x_0)$. For $1 \le p < \infty$ and $\mu \ge 0$ we define:

(i) the *Morrey space* $L^{p,\mu}(\Omega)$ as the space of those functions in $L^p(\Omega)$ such that

$$\|u\|^p_{L^{p,\mu}(\Omega)} \equiv \sup_{x_0\in\Omega, 0<r<\mathrm{diam}(\Omega)} r^{-\mu} \int_{\Omega(x_0,r)} |u|^p \,\mathrm{d}x < \infty;$$

(ii) the *Campanato space* $\mathcal{L}^{p,\mu}(\Omega)$ as the space of those functions in $L^p(\Omega)$ such that

$$[u]^p_{\mathcal{L}^{p,\mu}(\Omega)} \equiv \sup_{x_0\in\Omega, 0<r<\mathrm{diam}(\Omega)} r^{-\mu} \int_{\Omega(x_0,r)} |u - (u)_{x_0,r}|^p \,\mathrm{d}x < \infty,$$

where we exceptionally write $(u)_{x_0,r} \equiv \frac{1}{|\Omega(x_0,r)|} \int_{\Omega(x_0,r)} u \,\mathrm{d}x$. We write

$$\|u\|_{\mathcal{L}^{p,\mu}(\Omega)} \equiv \|u\|_{L^p(\Omega)} + [u]_{\mathcal{L}^{p,\mu}}.$$

Both Morrey and Campanato spaces are Banach spaces, although we will not use this fact.

Let us first look quickly into Morrey spaces. They are only interesting when $\mu \in [0, n]$, and they serve as a scale of function spaces between $L^p$ and $L^\infty$:

$$L^{p,0} = L^p, \qquad L^{p,n} = L^\infty, \qquad L^{p,\mu} = \{0\} \text{ if } \mu > n,$$

as can be easily checked using Lebesgue's differentiation theorem. Hölder's inequality shows that $L^{q,\mu}$ are ordered, i.e. $L^{q,\mu} \subseteq L^{p,\mu}$ whenever $q \geq p$.

We now investigate in more detail Campanato spaces. Here and in the rest of this section, we implicitly assume that $\Omega$ satisfies the following mild regularity condition, which is always satisfied for smooth or even Lipschitz domains:

$$|\Omega(r, x_0)| \geq cr^n \quad \text{for all } x_0 \in \overline{\Omega}, r \in (0, \operatorname{diam}(\Omega)). \tag{4.2}$$

We will split our analysis in the cases $\mu < n$ and $\mu > n$; the critical case $\mu = n$ is more subtle, and will be addressed in Problem 4.2.

**Theorem 4.3** (Characterization of Campanato spaces)**.** *Let $p \in [1, \infty)$.*

(i) $0 \leq \mu < n$*: then* $\mathcal{L}^{p,\mu}(\Omega)$ *and* $L^{p,\mu}(\Omega)$ *are equivalent spaces.*

(ii) $n < \mu \leq n + p$*: then* $\mathcal{L}^{p,\mu}(\Omega)$ *and* $C^{0,\alpha}(\Omega)$*, with* $\alpha = \frac{\mu - n}{p}$*, are equivalent spaces.*

**Proof.** Let us first dispense with the easy inclusions. For each $\mu \geq 0$, we have $L^{p,\mu} \subseteq \mathcal{L}^{p,\mu}$: using $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ and Jensen's inequality, we have

$$\int_{\Omega(x_0,r)} |u - (u)_{x_0,r}|^p \,\mathrm{d}x \leq 2^{p-1} \left[\int_{\Omega(x_0,r)} |u|^p \,\mathrm{d}x + |\Omega(x_0,r)||(u)_{x_0,r}|^p\right] \leq 2^p \int_{\Omega(x_0,r)} |u|^p \,\mathrm{d}x.$$

The inclusion $C^{0,\alpha} \subset \mathcal{L}^{p,n+\alpha p}$ is also easy, because we can take averages of the pointwise estimate $|u(x) - u(y)| \leq [u]_{C^{0,\alpha}}(2r)^\alpha$, which holds for all $x, y \in B_r(x_0)$. Thus

$$|u(x) - (u)_{x_0,r}| = \left|\frac{1}{|\Omega(x_0,r)|} \int_{\Omega(x_0,r)} [u(x) - u(y)] \,\mathrm{d}y\right| \leq [u]_{C^{0,\alpha}}(2r)^\alpha$$

and so, integrating in $x$, we get

$$\int_{\Omega(x_0,r)} |u(x) - (u)_{x_0,r}|^p \,\mathrm{d}x \leq [u]^p_{C^{0,\alpha}}(2r)^{\alpha p}|B_r(x_0)| \leq C[u]^p_{C^{0,\alpha}} r^{n+\alpha p}.$$

Hence the rest of the proof is dedicated to prove that $\mathcal{L}^{p,\mu}$ is contained either in a Morrey space in case (i), or in a Hölder space in case (ii). Throughout the proof, we let $C$ denote a generic constant depending on $n, p, \mu$ and $\Omega$.

The key point in the proof, in either of the two cases, is to have good estimates on the growth of the averages $f(r) \equiv |(u)_{x_0,r}|^p$, as $r \to 0$. To do so, we want to estimate how much $f$ can change from one dyadic scale to the next. Given $0 < r < R$ and $x, x_0 \in \Omega$, we have

$$|(u)_{x_0,r} - (u)_{x_0,R}|^p \leq 2^{p-1} \left(|u(x) - (u)_{x_0,R}|^p + |u(x) - (u)_{x_0,r}|^p\right)$$

and so, integrating in $x$ and using (4.2), we get

$$\begin{aligned} |(u)_{x_0,r} - (u)_{x_0,R}|^p &\leq \frac{2^{p-1}}{cr^n} \left(\int_{\Omega(x_0,R)} |u - (u)_{x_0,R}|^p \,\mathrm{d}x + \int_{\Omega(x_0,r)} |u - (u)_{x_0,r}|^p \,\mathrm{d}x\right) \\ &\leq \frac{C}{r^n} \left(R^\mu + r^\mu\right) [u]^p_{\mathcal{L}^{p,\mu}} \\ &\leq C \frac{R^\mu}{r^n} [u]^p_{\mathcal{L}^{p,\mu}}, \end{aligned}$$

since $r < R$. Thus, taking the $p$-th rooth, we get

$$|(u)_{x_0,r} - (u)_{x_0,R}| \leq CR^{\frac{\mu}{p}} r^{-\frac{n}{p}} [u]_{\mathcal{L}^{p,\mu}} = C\left(\frac{R}{r}\right)^{\frac{n}{p}} R^{\frac{\mu-n}{p}} [u]_{\mathcal{L}^{p,\mu}}.$$

Now we assume that $r, R$ are in adjacent dyadic scales. To be precise, fix $0 < \rho < \operatorname{diam}(\Omega)$, which we think of as a fixed constant, corresponding to the initial dyadic scale, and take

$$r \equiv 2^{-(k+1)}\rho, \qquad R \equiv 2^{-k}\rho.$$

With these choices, the previous estimate takes the form

$$|(u)_{x_0,\rho/2^{k+1}} - (u)_{x_0,\rho/2^k}| \leq C[u]_{\mathcal{L}^{p,\mu}} \left(\frac{\rho}{2^k}\right)^{\frac{\mu-n}{p}} \tag{4.3}$$

and notice how the behavior of the right-hand side will change depending on whether we are in case (i) or (ii). Let us sum (4.3) from the initial scale $k = 0$ to some large scale $k = N - 1$ to get

$$|(u)_{x_0,\rho/2^N} - (u)_{x_0,\rho}| \leq C[u]_{\mathcal{L}^{p,\mu}} \rho^{\frac{\mu-n}{p}} \left(\frac{2^{N\frac{n-\mu}{p}} - 1}{2^{\frac{n-\mu}{p}} - 1}\right). \tag{4.4}$$

At this point, we split the analysis depending on whether $\mu < n$ or $\mu > n$.

Let us first deal with the case (i), where $\mu < n$; in this case, the term in parentheses in (4.4) is comparable to $2^{N\frac{n-\mu}{p}}$, and so we obtain

$$|(u)_{x_0,\rho/2^N} - (u)_{x_0,\rho}| \leq C[u]_{\mathcal{L}^{p,\mu}} \left(\frac{\rho}{2^N}\right)^{\frac{\mu-n}{p}}. \tag{4.5}$$

Now let $r \in (0, \operatorname{diam}(\Omega))$ be arbitrary. We can thus find a unique $N \in \mathbb{N}$ and a unique $\frac{1}{2}\operatorname{diam}(\Omega) \leq \rho \leq \operatorname{diam}(\Omega)$ such that $r = \rho/2^N$. Thus (4.5) yields

$$\begin{aligned} |(u)_{x_0,r}|^p &\leq 2^{p-1}\left(|(u)_{x_0,\rho}|^p + |(u)_{x_0,r} - (u)_{x_0,\rho}|^p\right) \\ &\leq 2^{p-1}\left(C\operatorname{diam}(\Omega)^{-n}\|u\|_{L^p}^p + Cr^{\mu-n}[u]_{\mathcal{L}^{p,\mu}}^p\right). \end{aligned} \tag{4.6}$$

This is the desired estimate for $|(u)_{x_0,r}|^p$, which easily implies the conclusion. Indeed, by the triangle inequality, we have

$$\int_{\Omega(x_0,r)} |u|^p \,\mathrm{d}x \leq 2^{p-1}\int_{\Omega(x_0,r)} |u - (u)_{x_0,r}|^p \,\mathrm{d}x + cr^n|(u)_{x_0,r}|^p \leq Cr^\mu[u]_{\mathcal{L}^{p,\mu}}^p + Cr^n|(u)_{x_0,r}|^p,$$

and so inserting (4.6) into this estimate, and using the fact that $r \leq \operatorname{diam}(\Omega)$, we get

$$r^{-\mu}\int_{\Omega(x_0,r)} |u|^p \,\mathrm{d}x \leq C[u]_{\mathcal{L}^{p,\mu}}^p + C\operatorname{diam}(\Omega)^{-\mu}\|u\|_{L^p}^p$$

as wished.

Now let us deal with case (ii), so we assume that $\mu > n$. Going back to (4.3), note that since $\mu > n$ this inequality asserts that, for fixed $\rho$ and $x_0$, the sequence $\left((u)_{x_0,\rho/2^k}\right)_k$ is Cauchy and so it has a limit, say $\tilde{u}(x_0)$. By Lebesgue's differentiation theorem, we must have

$$u(x_0) = \tilde{u}(x_0) \quad \text{for a.e. } x_0 \text{ in } \Omega.$$

We now claim that $\tilde{u}$ is a continuous representative for $u$. Sending $N \to \infty$ in (4.4), as $\mu > n$, we obtain

$$|\tilde{u}(x_0) - (u)_{x_0,\rho}| \leq C[u]_{\mathcal{L}^{p,\mu}} \rho^{\frac{\mu-n}{p}}. \tag{4.7}$$

Thus we see that $(u)_{x,\rho} \to \tilde{u}(x)$ *uniformly* as $\rho \to 0$; since $x \mapsto (u)_{x,\rho}$ is continuous, we deduce that also $\tilde{u}$ is continuous. Thus we can identify $u$ with its continuous representative $\tilde{u}$.

We now prove that, in fact, $u$ is Hölder continuous: given $x, y \in \Omega$ and writing $\rho = |x-y|$, using (4.7) we have

$$\begin{aligned} |u(x) - u(y)| &\le |(u)_{x,2\rho} - u(x)| + |(u)_{x,2\rho} - (u)_{y,2\rho}| + |(u)_{y,2\rho} - u(y)| \\ &\le C[u]_{\mathcal{L}^{p,\mu}} \rho^{\frac{\mu-n}{p}} + |(u)_{x,2\rho} - (u)_{y,2\rho}|. \end{aligned} \tag{4.8}$$

Thus it remains to deal with the last term. Since $\Omega(x,\rho) \subset \Omega(y,2\rho)$, we have

$$\begin{aligned} |(u)_{x,2\rho} - (u)_{y,2\rho}|^p &\le C\rho^{-n} \int_{\Omega(x,\rho)} |(u)_{x,2\rho} - (u)_{y,2\rho}|^p \,\mathrm{d}z \\ &\le 2^{p-1}\rho^{-n} \left( \int_{\Omega(x,2\rho)} |u(z) - (u)_{x,2\rho}|^p \,\mathrm{d}z + \int_{\Omega(y,2\rho)} |u(z) - (u)_{y,2\rho}|^p \,\mathrm{d}z \right) \\ &\le 2^p \rho^{\mu-n} [u]_{\mathcal{L}^{p,\mu}} \end{aligned}$$

This, combined with (4.8), yields the conclusion. □

Theorem 4.3(ii) is interesting in that it gives an *integral* (rather than pointwise) characterization of Hölder spaces. Note that Hölder spaces are only interesting for $\alpha \le 1$, cf. Problem 4.1, so the restriction $\mu \le n + p$ is natural, although not strictly necessary for the above argument.

We now observe that Theorem 4.3 actually implies Morrey's embedding from Theorem 1.16(ii). To see this, we first state the following:

**Lemma 4.4.** *Let $p \in (1, \infty)$ and $\mu \ge 0$. If $|\mathrm{D}u| \in L^{p,\mu}_{\mathrm{loc}}(\Omega)$ then $u \in \mathcal{L}^{p,\mu+p}_{\mathrm{loc}}(\Omega)$.*

**Proof.** We have

$$\frac{1}{r^{\mu+p}} \int_{B_r(x_0)} |u - (u)_{x_0,r}|^p \,\mathrm{d}x \le C \frac{1}{r^\mu} \int_{B_r(x_0)} |\mathrm{D}u|^p \,\mathrm{d}x \le \|u\|^p_{L^{p,\mu}(\Omega)},$$

by the Poincaré inequality in Corollary 1.20. □

**Corollary 4.5.** *Let $p > n$. If $u \in W^{1,p}(\Omega)$ then $u \in C^{0,\alpha}_{\mathrm{loc}}(\Omega)$ with $\alpha = 1 - \frac{n}{p}$.*

**Proof.** By Lemma 4.4, if $u \in W^{1,p}(\Omega)$ then $\mathrm{D}u \in L^p(\Omega) = L^{p,0}(\Omega)$ and so $u \in \mathcal{L}^{p,p}_{\mathrm{loc}}(\Omega)$. The conclusion then follows from Theorem 4.3(ii). □

We conclude this section with an elementary but very useful technical lemma, that we will use several times in the sequel.

**Lemma 4.6** (Iteration lemma)**.** *Consider a non-decreasing function $\phi\colon (0, R_0] \to [0, +\infty)$ which satisfies, for some constants $A, B, \varepsilon > 0$ and $0 < \beta < \alpha$,*

$$\phi(r) \le A\left[\left(\frac{r}{R}\right)^\alpha + \varepsilon\right]\phi(R) + BR^\beta \quad \textit{for all } 0 < r \le R \le R_0.$$

*Then there is $C = C(\alpha, \beta, A)$ and $\varepsilon_0 = \varepsilon_0(\alpha, \beta, A)$ such that, if $\varepsilon \le \varepsilon_0$, we have*

$$\phi(r) \le C\left[\frac{\phi(R)}{R^\beta} + B\right] r^\beta \quad \textit{for all } 0 < r \le R \le R_0.$$

**Proof.** Since $\varepsilon \in (0,1)$ let us assume without loss of generality that $A > \frac{1}{2}$. Let us take $\gamma \equiv \frac{\alpha+\beta}{2}$. We choose $\tau \in (0,1)$ with $2A\tau^\alpha = \tau^\gamma$ and $\varepsilon_0$ so that $\varepsilon_0 \leq \tau^\alpha$. Thus

$$\phi(\tau R) \leq A(\tau^\alpha + \varepsilon_0)\phi(R) + BR^\beta \leq 2A\tau^\alpha \phi(R) + BR^\beta = \tau^\gamma \phi(R) + BR^\beta.$$

Thus, iterating once, we get

$$\phi(\tau^2 R) \leq \tau^\gamma \phi(\tau R) + B\tau^\beta R^\beta \leq \tau^{2\gamma}\phi(R) + \tau^\gamma BR^\beta + B\tau^\beta R^\beta = \tau^{2\gamma}\phi(R) + BR^\beta \tau^\beta (1 + \tau^{\gamma-\beta}).$$

Hence, iterating the first estimate $k$ times, we get

$$\begin{aligned}\phi(\tau^k R) &\leq \tau^{k\gamma}\phi(R) + BR^\beta \tau^{(k-1)\beta} \sum_{i=0}^{k-1} \tau^{i(\gamma-\beta)} \\ &= \tau^{k\gamma}\phi(R) + BR^\beta \tau^{(k-1)\beta} \frac{1-\tau^{k(\gamma-\beta)}}{1-\tau^{\gamma-\beta}} \leq C\tau^{(k+1)\beta}[\phi(R) + BR^\beta].\end{aligned}$$

Now for $0 < r < R$, let $k \in \mathbb{N}$ be such that $\tau^{k+1}R < r \leq \tau^k R$. Then

$$\phi(r) \leq \phi(\tau^k R) \leq C\tau^{(k+1)\beta}[\phi(R) + BR^\beta] \leq C\left[\phi(R) + BR^\beta\right]\left(\frac{r}{R}\right)^\beta,$$

as wished. □

## 4.2 Interior Schauder estimates

The fundamental idea in Schauder theory is to reduce the study of (4.1) to the constant coefficient case, so we begin by proving estimates in that case.

**Theorem 4.7** (Constant coefficients)**.** *Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a weak solution of (4.1), where $A$ is constant and satisfies the Legendre–Hadamard condition. If $F \in \mathcal{L}^{2,\mu}_{\mathrm{loc}}(\Omega, \mathbb{R}^{m\times n})$ for $\mu \in [0, n+2)$ then $\mathrm{D}u \in \mathcal{L}^{2,\mu}_{\mathrm{loc}}(\Omega, \mathbb{R}^{m\times n})$, with the estimate*

$$\|\mathrm{D}u\|_{\mathcal{L}^{2,\mu}(\Omega'')} \leq C(\|\mathrm{D}u\|_{L^2(\Omega')} + [F]_{\mathcal{L}^{2,\mu}(\Omega')}),$$

*where $\Omega'' \Subset \Omega' \Subset \Omega$ and $C = C(n, m, \Omega', \Omega'', \lambda, \Lambda, \mu)$.*

**Proof.** For a ball $B_R(x_0) \Subset \Omega$, let us write $u = v + w$, where $v$ is the unique solution to

$$\begin{cases} \operatorname{div}(A\,\mathrm{D}v) = 0 & \text{in } B_R(x_0), \\ v = u & \text{on } \partial B_R(x_0); \end{cases}$$

note that the solution exists by Corollary 2.10. By Lemma 3.13, we get

$$\int_{B_r(x_0)} |\mathrm{D}v - (\mathrm{D}v)_{x_0,r}|^2 \,\mathrm{d}x \leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |\mathrm{D}v - (\mathrm{D}v)_{x_0,R}|^2 \,\mathrm{d}x$$

and so, using the triangle inequality and Exercise 3.12, we get

$$\begin{aligned}\int_{B_r(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,r}|^2 &\leq 2\int_{B_r(x_0)} |\mathrm{D}v - (\mathrm{D}v)_{x_0,r}|^2 + 2\int_{B_r(x_0)} |\mathrm{D}w - (\mathrm{D}w)_{x_0,r}|^2 \\ &\leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |\mathrm{D}v - (\mathrm{D}v)_{x_0,R}|^2 + 2\int_{B_r(x_0)} |\mathrm{D}w - (\mathrm{D}w)_{x_0,R}|^2 \qquad (4.9) \\ &\leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,R}|^2 + C\int_{B_R(x_0)} |\mathrm{D}w - (\mathrm{D}w)_{x_0,R}|^2.\end{aligned}$$

We now want to estimate the last term, using Exercise 3.12 to write

$$\int_{B_R(x_0)} |\mathrm{D}w - (\mathrm{D}w)_{x_0,R}|^2 \,\mathrm{d}x \le \int_{B_R(x_0)} |\mathrm{D}w|^2 \,\mathrm{d}x.$$

Clearly, by its definition, $w \in W_0^{1,2}(B_R(x_0), \mathbb{R}^m)$ is a weak solution of

$$\begin{cases} \operatorname{div}(A\,\mathrm{D}w) = \operatorname{div} F & \text{in } B_R(x_0), \\ w = 0 & \text{on } \partial B_R(x_0), \end{cases}$$

and so we have

$$\int_{B_R(x_0)} \langle A\,\mathrm{D}w, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_{B_R(x_0)} \langle F, \mathrm{D}\varphi\rangle \,\mathrm{d}x = \int_{B_R(x_0)} \langle F - (F)_{x_0,R}, \mathrm{D}\varphi\rangle \,\mathrm{d}x$$

for any $\varphi \in W_0^{1,2}(B_R(x_0), \mathbb{R}^m)$. Thus, taking $\varphi = w$ and using Gårding's inequality, we get

$$\begin{aligned} \lambda \int_{B_R(x_0)} |\mathrm{D}w|^2 &\le \int_{B_R(x_0)} \langle A\,\mathrm{D}w, \mathrm{D}w\rangle \,\mathrm{d}x \\ &\le \left( \int_{B_R(x_0)} |F - (F)_{x_0,R}|^2 \,\mathrm{d}x \right)^{1/2} \left( \int_{B_R(x_0)} |\mathrm{D}w|^2 \,\mathrm{d}x \right)^{1/2}. \end{aligned}$$

So, combining the previous estimates,

$$\int_{B_R(x_0)} |\mathrm{D}w - (\mathrm{D}w)_{x_0,r}|^2 \,\mathrm{d}x \le C[F]^2_{\mathcal{L}^{2,\mu}} R^\mu.$$

Inserting this estimate into (4.9), we have thus shown that

$$\phi(r) \equiv \int_{B_r(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,r}|^2 \,\mathrm{d}x \le C \left(\frac{r}{R}\right)^{n+2} \phi(R) + C[F]^2_{\mathcal{L}^{2,\mu}} R^\mu$$

and so Lemma 4.6, applied with $\alpha = n+2$ and $\beta = \mu$ yields

$$\phi(r) \le C \left( \left(\frac{r}{R}\right)^\mu \phi(R) + C[F]^2_{\mathcal{L}^{2,\mu}} r^\mu \right) \le C \left( \|\mathrm{D}u\|^2_{L^2} \left(\frac{r}{R}\right)^\mu + [F]^2_{\mathcal{L}^{2,\mu}} r^\mu \right),$$

which yields the conclusion. □

We now treat the case of variable coefficients.

**Theorem 4.8** (Variable coefficients). *Fix $\sigma \in (0,1)$. Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a weak solution of (4.1), where $F \in C^{0,\sigma}_{\mathrm{loc}}(\Omega, \mathbb{R}^{m\times n})$ and $A^{ij}_{\alpha\beta} \in C^{0,\sigma}_{\mathrm{loc}}(\Omega)$ satisfies the Legendre–Hadamard condition. Then $\mathrm{D}u \in C^{0,\sigma}_{\mathrm{loc}}(\Omega, \mathbb{R}^{m\times n})$, with the estimate*

$$\|\mathrm{D}u\|_{C^{0,\sigma}(\Omega'')} \le C(\|\mathrm{D}u\|_{L^2(\Omega')} + \|F\|_{C^{0,\sigma}(\Omega')}),$$

*where $\Omega'' \Subset \Omega' \Subset \Omega$ and $C = C(n, m, \Omega', \Omega'', \lambda, \Lambda, \sigma, [A]_{C^{0,\sigma}})$.*

**Proof.** As in the proof of Corollary 2.11, the idea is to treat the case of variable coefficients as a perturbation of the constant coefficient case. In particular, we use Korn's trick:

$$\operatorname{div}(A(x_0)\mathrm{D}u) = \operatorname{div} G, \qquad G(x) \equiv -\left(A(x) - A(x_0)\right)\mathrm{D}u(x) + F(x) - (F)_{x_0,R}.$$

We now argue as in the proof of Theorem 4.7, and we let $v$ be the solution of

$$\begin{cases} \operatorname{div}(A(x_0)\mathrm{D}v) = 0 & \text{in } B_R(x_0), \\ v = u & \text{on } \partial B_R(x_0). \end{cases}$$

Arguing exactly as in the proof of Theorem 4.7 (i.e. using (3.6)), with $w = u - v$, we get

$$\int_{B_r(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,r}|^2 \leq C\left(\frac{r}{R}\right)^{n+2} \int_{B_R(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,R}|^2 + C\int_{B_R(x_0)} |\mathrm{D}w|^2. \tag{4.10}$$

Also as before, we have an estimate for the last term, by plugging in $w$ as test function in its own equation, and using the fact that $A \in C^{0,\sigma}$:

$$\begin{aligned} \int_{B_R(x_0)} |\mathrm{D}w|^2 \leq C\int_{B_R(x_0)} |G|^2 &\leq C\int_{B_R(x_0)} |F - (F)_{x_0,R}|^2 + CR^{2\sigma}\int_{B_R(x_0)} |\mathrm{D}u|^2 \\ &\leq C[F]^2_{\mathcal{L}^{2,n+2\sigma}} R^{n+2\sigma} + CR^{2\sigma}\int_{B_R(x_0)} |\mathrm{D}u|^2. \end{aligned} \tag{4.11}$$

Thus, writing

$$\phi(r) \equiv \int_{B_r(x_0)} |\mathrm{D}u - (\mathrm{D}u)_{x_0,r}|^2 \,\mathrm{d}x,$$

we arrive at

$$\phi(r) \leq C\left(\frac{r}{R}\right)^{n+2} \phi(R) + C[F]^2_{\mathcal{L}^{2,n+2\sigma}} R^{n+2\sigma} + CR^{2\sigma}\int_{B_R(x_0)} |\mathrm{D}u|^2. \tag{4.12}$$

In order to be able to conclude, we need an estimate on the last term in (4.12). For this, we argue exactly as for (4.10), but we use (3.5) instead of (3.6): this yields

$$\begin{aligned} \int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x &\leq C\left(\frac{r}{R}\right)^{n} \int_{B_R(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x + C\int_{B_R(x_0)} |\mathrm{D}w|^2 \,\mathrm{d}x \\ &\leq C\left[\left(\frac{r}{R}\right)^{n} + R^{2\sigma}\right] \int_{B_R(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x + C[F]^2_{\mathcal{L}^{2,n+2\sigma}} R^{n+2\sigma}, \end{aligned}$$

where the last line follows from (4.11). Let $\varepsilon > 0$ be arbitrarily small and suppose without loss of generality that $R \leq R_0 \leq 1$ is sufficiently small. Since $R^{n+2\sigma} \leq R^{n-\varepsilon}$, if we set $\psi(r) = \int_{B_r(x_0)} |\mathrm{D}u|^2$, we can apply Lemma 4.6 to $\psi$, with $\alpha = n$ and $\beta = n - \varepsilon$, to get

$$\int_{B_r(x_0)} |\mathrm{D}u|^2 \,\mathrm{d}x \leq C(\varepsilon, R_0)\left(\|\mathrm{D}u\|^2_{L^2(B_{R_0}(x_0))} + [F]^2_{\mathcal{L}^{2,n+2\sigma}}\right) r^{n-\varepsilon}.$$

Recall that $[F]_{\mathcal{L}^{2,n+2\sigma}}$ is comparable to $[F]_{C^{0,\sigma}}$ by Theorem 4.3.

We can now complete the proof. Returning to (4.12), we see that

$$\phi(r) \leq C\left(\left(\frac{r}{R}\right)^{n+2} \phi(R) + [F]^2_{\mathcal{L}^{2,n+2\sigma}} R^{n+2\sigma} + R^{2\sigma} R^{n-\varepsilon}\right)$$

and so, by Lemma 4.6, we have that $\mathrm{D}u \in \mathcal{L}^{2,n+2\sigma-\varepsilon}_{\mathrm{loc}}(\Omega) \cong C^{0,\sigma-\varepsilon/2}_{\mathrm{loc}}$, by Theorem 4.3(ii). Thus in particular $\mathrm{D}u$ is bounded and so actually (4.12) improves to

$$\phi(r) \leq C\left(\left(\frac{r}{R}\right)^{n+2} \phi(R) + ([F]^2_{\mathcal{L}^{2,n+2\sigma}} + 1)R^{n+2\sigma}\right)$$

which, again by Lemma 4.6, yields $\phi(r) \le Cr^{n+2\sigma}$. Thus $\mathrm{D}u \in \mathcal{L}^{2,n+2\sigma}_{\mathrm{loc}}(\Omega) \cong C^{0,\sigma}_{\mathrm{loc}}$ as wished. □

By iterating the above result, we obtain the following analogue of Theorem 3.9:

**Theorem 4.9** (Higher-order regularity)**.** *Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a weak solution of* (4.1)*. If $A^{\alpha\beta}_{ij} \in C^{k,\sigma}_{\mathrm{loc}}(\Omega)$ satisfy the Legendre–Hadamard condition and $F^\alpha_i \in C^{k,\sigma}_{\mathrm{loc}}(\Omega)$ then we have $u \in C^{k+1,\sigma}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$.*

**Proof.** Differentiate the system and argue inductively, see Problem 4.5. □

Thus Schauder theory allows us to reprove Corollary 3.10 in a rather different way.

## 4.3 Schauder theory as an existence theory

Recall that our strategy in Sections 2–3 was to first construct, using Functional Analytic methods, weak solutions of (4.1) and then to prove regularity of such solutions. In Schauder theory, instead, one can construct regular solutions directly without passing through some generalized notion of solution. For this reason, Schauder theory is not just a regularity theory, but an *existence theory* as well, and in fact it is substantially older than the approach based on Sobolev spaces.

In this subsection we sketch how to use Schauder theory to obtain existence of solutions over a smooth domain $\Omega$. There are two crucial ingredients in the method:

(i) *Solvability of the Poisson equation*: for $F \in C^{1,\alpha}(\Omega, \mathbb{R}^m)$ there is a unique solution $u \in C^{2,\alpha}(\Omega, \mathbb{R}^m)$ of

$$\begin{cases} \Delta u = \operatorname{div} F & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

(ii) *Global Schauder estimates*: For $F \in C^{1,\alpha}(\overline{\Omega}, \mathbb{R}^m)$, any solution of (4.1) satisfies

$$\|u\|_{C^{2,\alpha}(\overline{\Omega})} \le C(\Omega, m, \lambda, \Lambda, \|A\|_{C^{1,\alpha}(\overline{\Omega})}) \|F\|_{C^{1,\alpha}(\overline{\Omega})}.$$

There are several ways of obtaining (i) without passing through the theory of Sobolev spaces. For instance, if $\Omega$ is a ball then one can write the unique solution of the Poisson equation explicitly, through the corresponding Green's function. If $\Omega$ is a more general domain then one needs to use other ideas, such as barriers and Perron's method.

Concerning (ii), note that this is simply a *global* (i.e. up to the boundary) version of Theorem 4.9. Since the focus of these notes is on *interior* regularity, and in any case the main ideas in the proof of (ii) are already contained in the proof of Theorem 4.9, we do not present the proof of (ii) here.

In any case, let us now see how (i) and (ii) combined yield an existence theorem for linear elliptic systems satisfying the Legendre–Hadamard condition. We consider two operators

$$L_0 = \Delta \colon \mathbb{X} \to \mathbb{Y}, \qquad L_1 = \operatorname{div}(A\,\mathrm{D}\cdot) \colon \mathbb{X} \to \mathbb{Y},$$

where

$$\mathbb{X} = C^{2,\alpha}(\overline{\Omega}, \mathbb{R}^m) \cap \{u|_{\partial\Omega} = 0\}, \qquad \mathbb{Y} = C^{1,\alpha}(\overline{\Omega}).$$

From (i) we already know that $L_0$ is surjective, and our goal is now to show that $L_1$ is surjective as well. That this is the case follows from the a priori estimate (ii) together with the following abstract result:

**Theorem 4.10** (Continuity method)**.** *Let $L_0, L_1 \colon \mathbb{X} \to \mathbb{Y}$ be bounded linear operators between Banach spaces. Set*

$$L_t \equiv (1-t)L_0 + tL_1, \quad t \in [0,1].$$

*Suppose that*

$$\|u\|_{\mathbb{X}} \le C\|L_t u\|_{\mathbb{Y}} \qquad \textit{for all } u \in \mathbb{X},\ \textit{all } t \in [0,1]. \tag{4.13}$$

*If $L_0$ is surjective, then so is $L_1$.*

**Proof.** Suppose that $L_s$ is surjective, for some $s \in [0,1]$. By (4.13), $L_s$ is injective as well and hence we have a bounded inverse $L_s^{-1} \colon \mathbb{Y} \to \mathbb{X}$. We now rewrite the equation $L_t u = f$ as

$$L_s u = f + (L_s - L_t)u = f + (t-s)(L_0 u - L_1 u)$$

or, in yet another way,

$$u = L_s^{-1} f + (t-s)L_s^{-1}(L_0 u - L_1 u) \equiv Tu.$$

Thus we need to find a fixed point of $T \colon \mathbb{X} \to \mathbb{X}$. We estimate

$$\|Tu - Tv\| \le \|L_s^{-1}\|(\|L_0\| + \|L_1\|)|t-s|\|u-v\|.$$

According to (4.13) we have $\|L_s^{-1}\| \le C$ and so if $|t-s| \le \frac{1}{2}(C(\|L_0\| + \|L_1\|))^{-1} \equiv c$, we can apply the contraction mapping theorem to find a fixed point of $T$. Thus, if $L_s$ is surjective then so is $L_t$ whenever $|t-s| \le c$, and the conclusion follows. □

## 4.4 Problems for Section 4

**Problem 4.1** (Properties of Hölder spaces)**.** Let $\alpha \in (0,1]$.

(i) Show that $C^{0,\alpha}(\Omega, \mathbb{R}^m)$ is a Banach space.

(ii) Show that, if $\alpha > 1$ and $\Omega$ is connected, $C^{0,\alpha}(\Omega, \mathbb{R}^m)$ contains only constants.

(iii) Prove that the embedding $C^{0,\alpha}(\Omega, \mathbb{R}^m) \subset C^{0,\beta}(\Omega, \mathbb{R}^m)$, for $0 < \beta < \alpha \le 1$, is compact.

(iv) Let $\Omega = (-1,1)$. Show that the function $x \mapsto |x|^{\frac{1}{2}}$ is in $C^{0,1/2}(\Omega)$ but that it cannot be approximated in $C^{0,1/2}$ by smooth functions.

**Problem 4.2** (Functions of bounded mean oscillation)**.** The space $\mathcal{L}^{1,n}$ is a very important function space, both in elliptic PDE and in Harmonic Analysis, and it is more commonly known as

$$\mathrm{BMO}(\Omega) \equiv \mathcal{L}^{1,n}(\Omega).$$

Let us take $\Omega$ to be a ball.

(i) Use the Poincaré inequality to show that $W^{1,n}(\Omega) \subset \mathrm{BMO}(\Omega)$.

(ii) Show that $L^\infty(\Omega) \subset \mathrm{BMO}(\Omega)$.

(iii) Prove that the previous inclusions are strict, by showing that $\log \in \mathrm{BMO}(0,1)$.

In fact, a famous result in Harmonic Analysis due to John and Nirenberg asserts that, for all $p \in [1, \infty)$, the spaces $\mathcal{L}^{p,n}$ are equivalent to BMO.

**Problem 4.3** (Continuous coefficients)**.** Let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a weak solution of (4.1), where $A^{\alpha\beta}_{ij} \in C^0(\Omega)$ satisfies the Legendre–Hadamard condition. Use Korn's trick to prove that if $F \in L^{2,\mu}_{\mathrm{loc}}(\Omega, \mathbb{R}^{m\times n})$ for some $\mu \in [0, n)$ then $\mathrm{D}u \in L^{2,\mu}_{\mathrm{loc}}(\Omega)$. Conclude from Lemma 4.4 that if $\mu \in (n-2, n)$ then $u$ is Hölder continuous.

**Problem 4.4** (Non-divergence form right-hand side)**.** Let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ solve

$$-\partial_\alpha(A^{\alpha\beta}_{ij}\partial_\beta u^j) = f_i \quad \text{for } i = 1, \ldots, m,$$

where $A$ is constant and $f \in L^{2,\mu}(\Omega, \mathbb{R}^m)$ for some $\mu \in [0, n)$. Following the proof of Theorem 4.7, show that $\mathrm{D}u \in \mathcal{L}^{2,\mu+2}_{\mathrm{loc}}(\Omega)$.

**Problem 4.5** (Higher order regularity)**.** Prove Theorem 4.9.

# 5 Regularity for nonlinear convex problems

In this section we study regularity properties of minimizers of the energy

$$\mathbb{E}[u] \equiv \int_\Omega F(\mathrm{D}u)\,\mathrm{d}x,$$

where we assume that $F\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is smooth and satisfies

$$\lambda|\xi|^2 \leq \partial_{\xi_i^\alpha}\partial_{\xi_j^\beta} F(\eta)\xi_i^\alpha\xi_j^\beta = F''(\eta)[\xi,\xi] \leq \Lambda|\xi|^2 \tag{5.1}$$

for all $\eta, \xi \in \mathbb{R}^{m\times n}$, thus $F$ is *strongly convex* and has bounded Hessian.

**Exercise 5.1.** Show that (5.1) implies that $\lambda|\xi| \leq |F''(\eta)\xi| \leq \Lambda|\xi|$ for any $\eta, \xi \in \mathbb{R}^{m\times n}$.

We begin by showing that minimizers of $\mathbb{E}$ exist, and are unique.

**Theorem 5.2.** *Assume* (5.1) *holds. For any $g \in C^\infty(\overline{\Omega}, \mathbb{R}^m)$, $\mathbb{E}$ has a unique minimizer $u \in W^{1,2}_g(\Omega, \mathbb{R}^m)$. This minimizer is characterized as being the unique weak solution in $W^{1,2}_g(\Omega, \mathbb{R}^m)$ of the nonlinear system*

$$\operatorname{div}(F'(\mathrm{D}u)) = 0. \tag{5.2}$$

**Proof.** Let us begin by noting a few simple consequences of (5.1). Given $\xi_1, \xi_2 \in \mathbb{R}^{m\times n}$, by the fundamental theorem of calculus we have

$$\langle F'(\xi_1) - F'(\xi_2), \xi_1 - \xi_2\rangle = \int_0^1 F''(t\xi_1 + (1-t)\xi_2)[\xi_1 - \xi_2, \xi_1 - \xi_2]\,\mathrm{d}t$$

and so, by (5.1) and Exercise 5.1, we have

$$\lambda|\xi_1 - \xi_2| \leq |F'(\xi_1) - F'(\xi_2)| \leq \Lambda|\xi_1 - \xi_2|.$$

By adding an affine function to $F$, we may assume that $F(0) = 0$ and $F'(0) = 0$; then, applying again the fundamental theorem of calculus, we see that

$$\frac{\lambda}{2}|\xi|^2 \leq |F(\xi)| \leq \frac{\Lambda}{2}|\xi|^2.$$

Thus (5.1) implies conditions on $F$ and its derivatives which ensure that Proposition 2.14 and Corollary 2.18 are applicable, and so $\mathbb{E}$ has a minimizer in $W^{1,2}_g(\Omega, \mathbb{R}^m)$, which solves (5.2). Since (5.1) implies that $F$ is convex, we may also apply Proposition 2.15 to deduce that any weak solution of (5.2) is a minimizer of $\mathbb{E}$.

It remains to show the uniqueness of the minimizer, which follows from the strong convexity of $F$: for any distinct $\xi_1, \xi_2 \in \mathbb{R}^{m\times n}$, we have

$$\frac{1}{2}F(\xi_1) + \frac{1}{2}F(\xi_2) > F\left(\frac{\xi_1 + \xi_2}{2}\right).$$

If $u_1, u_2$ are different minimizers of $\mathbb{E}$, then the set $U \equiv \{\mathrm{D}u_1 \neq \mathrm{D}u_2\}$ has positive measure. But then the map $v = \frac{1}{2}(u_1 + u_2) \in W^{1,2}_g(\Omega, \mathbb{R}^m)$ has less energy than $u_1, u_2$, since

$$\begin{aligned}
\mathbb{E}[u_1] &= \frac{\mathbb{E}[u_1] + \mathbb{E}[u_2]}{2} \\
&= \int_U \frac{F(\mathrm{D}u_1) + F(\mathrm{D}u_2)}{2}\,\mathrm{d}x + \int_{\Omega\setminus U} \frac{F(\mathrm{D}u_1) + F(\mathrm{D}u_2)}{2}\,\mathrm{d}x
\end{aligned}$$

$$> \int_U F(\mathrm{D}v)\,\mathrm{d}x + \int_{\Omega\setminus U} F(\mathrm{D}v)\,\mathrm{d}x = \int_\Omega F(\mathrm{D}v)\,\mathrm{d}x,$$

which is impossible. □

The above theorem gives an answer to Hilbert's 20th problem. The focus of the rest of this section is thus on the 19th problem: is it the case that $u$ is smooth?

## 5.1 The $W^{2,2}$-estimate

Before addressing the full regularity of $u$, we show that $\mathrm{D}u$ is weakly differentiable and we determine the system it solves; essentially we follow the strategy of the proof of Theorem 3.8. Let us write

$$A_{ij}^{\alpha\beta}(x) \equiv \partial_{\xi_i^\alpha}\partial_{\xi_j^\beta} F(\mathrm{D}u(x)), \tag{5.3}$$

and note that (5.1) guarantees that $A$ satisfies both (2.2) and the Legendre condition.

**Proposition 5.3.** *Let $F$ satisfy* (5.1) *and let $u \in W^{1,2}(\Omega,\mathbb{R}^m)$ be a solution of* (5.2)*. Then $u \in W^{2,2}_{\mathrm{loc}}(\Omega,\mathbb{R}^m)$ and, for any $\gamma$, $w \equiv \partial_\gamma u$ is a weak solution of*

$$\partial_\alpha(A_{ij}^{\alpha\beta}\partial_\beta w^j) = 0 \quad \textit{for } i = 1,\dots,m.$$

**Proof.** Given $\varphi \in C^\infty(\Omega,\mathbb{R}^m)$, for all $h$ sufficiently small we have $\tau_{-h,\gamma}\varphi \in C^\infty(\Omega,\mathbb{R}^m)$ and so, since $u$ is a weak solution of (5.2),

$$\int_\Omega \langle F'(\mathrm{D}u(x+he_\gamma)) - F'(\mathrm{D}u(x)), \mathrm{D}\varphi\rangle\,\mathrm{d}x = 0.$$

Using the Fundamental Theorem of Calculus, we have

$$\begin{aligned} F'(\mathrm{D}u(x+he_\gamma)) - F'(\mathrm{D}u(x)) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} F'(t\mathrm{D}u(x+he_\gamma) + (1-t)\mathrm{D}u(x))\,\mathrm{d}t \\ &= \left(\int_0^1 F''\left(t\mathrm{D}u(x+he_\gamma) + (1-t)\mathrm{D}u(x)\right)\mathrm{d}t\right)\left(\mathrm{D}u(x+he_\gamma) - \mathrm{D}u(x)\right). \end{aligned}$$

Thus, writing $A_h(x)$ for this integral, and dividing by $h$, we get

$$\int_\Omega \langle A_h \partial_{h,\gamma}\mathrm{D}u, \mathrm{D}\varphi\rangle\,\mathrm{d}x = 0, \tag{5.4}$$

where we note that, for all $\xi \in \mathbb{R}^{m\times n}$ and for a.e. $x \in \Omega$,

$$\lambda|\xi|^2 \leq \langle A_h(x)\xi,\xi\rangle \leq \Lambda|\xi|^2.$$

Thus, by Caccioppoli's inequality, whenever $B_{2R}(x_0)$ and $h$ is small enough, we obtain

$$\int_{B_{R/2}(x_0)} |\partial_{h,\gamma}\mathrm{D}u|^2\,\mathrm{d}x \leq \frac{C}{R^2}\int_{B_R(x_0)} |\partial_{h,\gamma}u|^2 \leq \frac{C}{R^2}\int_\Omega |\mathrm{D}u|^2\,\mathrm{d}x,$$

where the last inequality follows from Proposition 1.13. Thus, using again the same proposition, we see that $\mathrm{D}u \in W^{1,2}_{\mathrm{loc}}(\Omega,\mathbb{R}^{m\times n})$. To derive the PDE for $\partial_\gamma u$, we pass to the limit as $h \to 0$ in (5.4): we claim that, up to subsequences, we have

$$\partial_{h,\gamma}\mathrm{D}u \rightharpoonup \partial_\gamma \mathrm{D}u \text{ in } L^2_{\mathrm{loc}}(\Omega), \qquad A_h \to A \text{ in } L^2_{\mathrm{loc}}(\Omega),$$

so that $A_h \partial_{h,\gamma} \mathrm{D}u \rightharpoonup u$ in $L^2_{\mathrm{loc}}(\Omega)$. The first convergence follows from the above estimate. For the second convergence, we argue as follows: for any $\varepsilon > 0$ we have

$$\int_{\{|\mathrm{D}u|>\varepsilon^{-1}\}} |A_h(x) - A(x)|^2 \,\mathrm{d}x \leq 2 \int_{\{|\mathrm{D}u|>\varepsilon^{-1}\}} \left(|A_h|^2 + |A|^2\right) \mathrm{d}x \leq C(\Lambda)\varepsilon^2 \int_\Omega |\mathrm{D}u|^2 \,\mathrm{d}x.$$

Since $A$ is smooth, in the ball $\{|\mathrm{D}u| \leq \varepsilon^{-1}\}$ there is a constant $C_\varepsilon$ such that $|A'| \leq C_\varepsilon$. Hence we can estimate

$$\int_{B_R(x_0)} |A_h(x) - A(x)|^2 \,\mathrm{d}x \leq C\varepsilon^2 + C_\varepsilon \int_0^1 \int_{\{|\mathrm{D}u|\leq\varepsilon^{-1}\}\cap B_R(x_0)} t^2 |\mathrm{D}u(x + he_\gamma) - \mathrm{D}u(x)|^2 \,\mathrm{d}x \,\mathrm{d}t.$$

Thus, by continuity of translations in $L^2(B_R(x_0))$, we have

$$\lim_{h\to 0} \int_{B_R(x_0)} |A_h(x) - A(x)|^2 \,\mathrm{d}x \leq C\varepsilon^2$$

and, as $\varepsilon > 0$ is arbitrary, the claim follows. □

**Corollary 5.4.** *In the setting of Proposition 5.3, suppose that $u \in C^{1,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ for some $\alpha > 0$. Then $u \in C^\infty(\Omega, \mathbb{R}^m)$.*

**Proof.** If $u \in C^{1,\alpha}$ then the coefficient tensor $A$ defined in (5.3) is Hölder continuous, so we can apply Theorem 4.8 to deduce that $w \in C^{1,\alpha}_{\mathrm{loc}}(\Omega)$, and so in fact $u \in C^{2,\alpha}_{\mathrm{loc}}(\Omega)$. But then $A$ is $C^{1,\alpha}_{\mathrm{loc}}(\Omega)$, and so we can apply Theorem 4.9 to deduce that $u \in C^{2,\alpha}_{\mathrm{loc}}(\Omega)$. Iterating this reasoning we conclude that $u \in C^\infty(\Omega)$. □

We note that Corollary 5.4 holds even when $\alpha = 0$, by applying Problem 4.3. Thus the key step in deducing smoothness of solutions to (5.2) is to show that $\mathrm{D}u$ is continuous: if this is the case, then the previous approach shows that essentially one can linearize (5.2).

When $n = 2$, note that $W^{2,2}$ almost embeds into $C^1$. Thus, in this case, Proposition 5.3 gives us hope that $u$ is smooth. And indeed, we have:

**Theorem 5.5.** *Let $F$ be a smooth integrand satisfying* (5.1) *and let $u \in W^{1,2}(\Omega, \mathbb{R}^m)$ be a weak solution of* (5.2)*. If $n = 2$ then $u \in C^\infty(\Omega, \mathbb{R}^m)$.*

**Proof.** By Proposition 5.3 we know that $w = \partial_\gamma u$ is a solution of a linear elliptic system with *measurable* coefficients. By the hole-filling technique of Lemma 3.15 (see also Remark 3.16), we find that $\mathrm{D}w \in L^{2,\alpha}_{\mathrm{loc}}(\Omega)$ for some $\alpha > 0$, and so by Lemma 4.4 we have $w \in \mathcal{L}^{2,\alpha+2}_{\mathrm{loc}}(\Omega)$. So far the argument is valid in all dimensions. When $n = 2$, we see from Theorem 4.3(i) that in fact $w \in C^{0,\alpha/2}_{\mathrm{loc}}(\Omega)$, and so $u \in C^{1,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ since $\gamma$ is arbitrary. We can then conclude from Corollary 5.4. □

Theorem 5.5 gives a positive answer to the regularity problem in the plane. However, at this stage it is not clear how to proceed in general dimension. In fact it turns out that the answer to Hilbert's 19th problem depends fundamentally on whether we are in the *scalar case* $m = 1$ or in the *vectorial case* $m > 1$. Let us begin with the former.

## 5.2 The scalar case: De Giorgi–Nash Theorem

We now take $m = 1$ and we consider weak solutions of the elliptic equation

$$Lu \equiv -\partial_\alpha(A^{\alpha\beta}\partial_\beta u) = 0, \tag{5.5}$$

where the coefficients $A$ are measurable and satisfy for a.e. $x$ in $\Omega$ the bounds

$$\lambda|\xi|^2 \leq \langle A(x)\xi, \xi\rangle \leq \Lambda|\xi|^2. \tag{5.6}$$

The next theorem asserts that, despite the roughness of the coefficients, any solution of (5.5) is Hölder continuous:

**Theorem 5.6** (De Giorgi–Nash)**.** *Let $u \in W^{1,2}(\Omega)$ be a weak solution of* (5.5)*, where $A$ is measurable and satisfies* (5.6)*. There is $\alpha = \alpha(n, \lambda, \Lambda) > 0$ such that, whenever $\Omega' \Subset \Omega$,*

$$\|u\|_{C^{0,\alpha}(\Omega')} \leq C\|u\|_{L^2(\Omega)},$$

*for a constant $C = C(\Omega, \Omega', n, \lambda, \Lambda)$.*

Clearly, applying Proposition 5.3 and Corollary 5.4, we obtain:

**Corollary 5.7.** *When $m = 1$, the minimizers of Theorem 5.2 are smooth.*

Theorem 5.6 is deeply remarkable at least in two aspects: firstly, it is a *genuinely scalar result*, as the conclusion fails completely in the vectorial setting, see already Problem 5.1; secondly, although it is a statement about a *linear* equation, the resulting estimate is actually *nonlinear*, as we will see during the proof.

A big part of the argument actually holds not just for solutions of (5.5) but also for subsolutions, and this is often useful in applications. Precisely, we have:

**Definition 5.8.** We say that $u \in W^{1,2}(\Omega)$ is a *subsolution* of (5.5), and we write $Lu \leq 0$, if

$$\int_\Omega \langle A\, \mathrm{D}u, \mathrm{D}\varphi\rangle \, \mathrm{d}x \leq 0 \qquad \forall \varphi \in W^{1,2}_0(\Omega) \text{ with } \varphi \geq 0.$$

When $A = \mathrm{Id}$, we say that $u$ is *subharmonic*.

Note that the definition of subsolution already uses crucially the fact that $m = 1$. Let us note the following basic property of subsolutions:

**Lemma 5.9.** *If $u \in W^{1,2}(\Omega)$ be a subsolution of* (5.5) *and* (5.6) *holds, then:*

(i) *if $f \in C^2(\mathbb{R})$ is convex, non-decreasing with $0 \leq f' \in L^\infty(\mathbb{R})$ then $f \circ u$ is also a subsolution of* (5.5)*;*

(ii) *the function $u_+ \equiv \max\{u, 0\}$ is also a subsolution of* (5.5).

**Proof.** To prove (i), let $0 \leq \varphi \in W^{1,2}_0(\Omega)$ be arbitrary and take $\psi = (f' \circ u)\varphi \in W^{1,2}_0(\Omega)$. Then, by (5.6),

$$\begin{aligned}\langle A\, \mathrm{D}u, \mathrm{D}\psi\rangle &= \langle A\, \mathrm{D}u, (f' \circ u)\mathrm{D}\varphi\rangle + \varphi\langle A\, \mathrm{D}u, \mathrm{D}(f' \circ u)\rangle \\ &= \langle A\, \mathrm{D}(f \circ u), \mathrm{D}\varphi\rangle + \varphi f''(u)\langle A\, \mathrm{D}u, \mathrm{D}u\rangle \geq \langle A\, \mathrm{D}(f \circ u), \mathrm{D}\varphi\rangle\end{aligned}$$

for a.e. $x$ in $\Omega$. The conclusion follows by integrating.

To prove (ii), consider the family of functions $f_\varepsilon(t) \equiv \Big((t^4 + \varepsilon^4)^{1/4} - \varepsilon\Big)\, 1_{\{t \geq 0\}}$, which are such that $f_0(t) = t_+$. Applying (i) to $f_\varepsilon(u)$, we can use the Dominated Convergence Theorem to conclude that $u_+ = f_0(u)$ is also a subsolution of (5.5). □

Although we will not use it explicitly in the sequel, we also observe that subsolutions satisfy the maximum principle over any domain $\Omega \subset \mathbb{R}^n$:

**Lemma 5.10** (Maximum principle). *If $u \in W^{1,2}(\Omega)$ has $Lu \le 0$, and $A$ satisfies* (5.6)*, then*

$$\sup_{\Omega} u = \sup_{\partial\Omega} u.$$

To be precise, the supremum on the left-hand side is the essential supremum of $u$, while on the right-hand side we define $\sup_{\partial\Omega} u \equiv \inf\{c : (u-c)^+ \in W^{1,2}_0(\Omega)\}$, and we assume that $\sup_{\partial\Omega} u > -\infty$.

**Proof.** According to Lemma 5.9, $0 \le \varphi = (u - \sup_{\partial\Omega} u)_+ \in W^{1,2}_0(\Omega)$ is also a subsolution of (5.5). Thus, testing the subsolution inequality against $\varphi$ itself, we see that

$$0 = \int_{\Omega} \langle A\, \mathrm{D}\varphi, \mathrm{D}\varphi\rangle \,\mathrm{d}x \ge \lambda \int_{\Omega} |\mathrm{D}\varphi|^2 \,\mathrm{d}x,$$

and so $\mathrm{D}\varphi = 0$ a.e. in $\Omega$. Thus by Lemma 1.5 we see that $\varphi$ is a.e. constant, i.e. we have $u \le \sup_{\partial\Omega} u$ a.e. in $\Omega$. □

In fact, Lemma 5.10 holds, with the same proof, under even weaker assumptions on $A$: the lower bound in (5.6) can be relaxed to $\langle A(x)\xi, \xi\rangle > 0$ for all $\xi \in \mathbb{R}^n$ and a.e. $x \in \Omega$. For sufficiently regular subsolutions one can also prove a stronger version of Lemma 5.10, which asserts that the maximum of a subsolution $u$ is never achieved in the interior, unless $u$ is constant:

**Lemma 5.11** (Strong maximum principle). *If $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfies $Lu \le 0$, then*

$$u \ne \text{constant} \qquad \Longrightarrow \qquad \max_{\Omega'} u < \max_{\partial\Omega} u,$$

*whenever $\Omega' \Subset \Omega$.*

We omit the proof of Lemma 5.11, as we will not need it.

We next state Caccioppoli's inequality, which we proved in Section 3 in the general vectorial case. Here we restate this result, noting that when $m = 1$ the same proof applies also to subsolutions.

**Lemma 5.12** (Caccioppoli's inequality). *Let $u \in W^{1,2}(B_1)$ be such that $Lu \le 0$. Then*

$$\int_{B_1} |\mathrm{D}(\eta u)|^2 \,\mathrm{d}x \le C \int_{B_1} u^2 |\mathrm{D}\eta|^2 \,\mathrm{d}x$$

*for all $\eta \in C^\infty_c(B_1)$.*

**Proof.** Clearly the desired inequality is equivalent to

$$\int_{B_1} \eta^2 |\mathrm{D}u|^2 \,\mathrm{d}x \le C \int_{B_1} u^2 |\mathrm{D}\eta|^2 \,\mathrm{d}x,$$

which is the standard Caccioppoli inequality that we proved (implicitly) in Theorem 3.5. □

We now follow essentially the strategy of De Giorgi to prove Theorem 5.6. The proof is split into two key steps: we first show that solutions are necessarily bounded, and then we show that bounded solutions are in fact Hölder continuous. We emphasize that the conclusion of neither step is true in the general vectorial case: the crucial point of $m = 1$ is essentially that one can look at *sublevel sets* of $u$. Let us begin with the first step.

**Theorem 5.13** (From $L^2$ to $L^\infty$). *Let $u \in W^{1,2}(B_1)$ be such that $Lu \le 0$. Then*

$$\|u_+\|_{L^\infty(B_{1/2})} \le C(n,\Lambda,\lambda)\|u_+\|_{L^2(B_1)}.$$

**Proof.** Let $\delta > 0$ be arbitrary, to be chosen later depending only on $n, \Lambda, \lambda$. By multiplying $u$ by a suitable constant, it is enough to show that

$$\|u_+\|_{L^2(B_1)} \le \delta \quad \Longrightarrow \quad \|u\|_{L^\infty(B_{1/2})} \le 1. \tag{5.7}$$

Let us set $u_k \equiv (u-(1-2^{-k}))_+$, and also $B_k \equiv B_{\frac{1}{2}+2^{-k}}$, $a_k \equiv \int_{B_k} u_k^2 \,\mathrm{d}x$. Note that, according to Lemma 5.9, $Lu_k \le 0$. Our goal is to derive an iteration formula for $a_k$.

Let $\eta_k$ be a cutoff function with $0 \le \eta_k \le 1$, $\eta_k = 1$ in $B_k$, $\eta_k = 0$ outside $B_{k-1}$ and $|\mathrm{D}\eta_k| \le C2^k$. First, by Lemma 5.12, since $u_k \le u_{k-1}$, we have

$$\int_{B_{k-1}} |\mathrm{D}(\eta_k u_k)|^2 \,\mathrm{d}x \le C2^{2k} a_{k-1}.$$

By Sobolev's inequality, we also have

$$\left(\int_{B_k} |u_k|^{2^*} \,\mathrm{d}x\right)^{2/2^*} \le \left(\int_{B_{k-1}} |\eta_k u_k|^{2^*} \,\mathrm{d}x\right)^{2/2^*} \le \int_{B_{k-1}} |\mathrm{D}(\eta_k u_k)|^2 \,\mathrm{d}x;$$

here we assume that $n \ge 3$, as we already dealt with $n = 2$ in Theorem 5.5. On the other hand, by Hölder's inequality we have

$$a_k = \int_{B_k} u_k^2 \,\mathrm{d}x \le \left(\int_{B_k} |u_k|^{2^*} \,\mathrm{d}x\right)^{2/2^*} \left|\{u_k > 0\} \cap B_k\right|^{2/n}.$$

Finally, by Chebyshev's inequality, and since $B_k \subset B_{k-1}$, we have

$$\left|\{u_k > 0\} \cap B_k\right| \le \left|\{u_{k-1} > 2^{-k}\} \cap B_{k-1}\right| \le 2^{2k} \int_{B_{k-1}} |u_{k-1}|^2 \,\mathrm{d}x = 2^{2k} a_{k-1}.$$

Thus, once we combine the above estimates, we get

$$a_k \le C2^{(2+4/n)k} a_{k-1}^{1+2/n}.$$

The crucial point in this iteration estimate is that the exponent of $a_{k-1}$ is larger than one, which guarantees that $a_k$ converges super-exponentially:

**Lemma 5.14.** *Suppose that there are constants $C, \gamma > 1$ such that*

$$a_k \le C^k a_{k-1}^\gamma \quad \textit{for all } k \in \mathbb{N}.$$

*Then $\lim_{k\to\infty} a_k = 0$ provided that $a_0$ is small enough, depending only on $C, \gamma$.*

**Proof.** Indeed, we have

$$a_k \le C^k a_{k-1}^\gamma \le C^{k+(k-1)\gamma} a_{k-2}^{\gamma^2} \le \cdots \le C^{\sum_{i=0}^{k-1} \gamma^i (k-i)} a_0^{\gamma^k} = a_0^{\gamma^k} C^{\gamma^k \sum_{i=0}^{k-1} \frac{k-i}{\gamma^{k-i}}} \le (a_0 C_\gamma^M)^{\gamma^k},$$

where $M_\gamma \equiv \sum_{i=0}^\infty \frac{k-i}{\gamma^{k-i}} < \infty$ as $\gamma > 1$. The claim follows by choosing $a_0$ so that $a_0 C^M < 1$. □

Hence we deduce that, by choosing $\int_{B_1} u_+^2 \,\mathrm{d}x = a_0 \le \delta$ suitably small, we have

$$\int_{B_{1/2}} (u-1)_+^2 \,\mathrm{d}x \le \lim_{k\to\infty} \int_{B_k} u_k^2 \,\mathrm{d}x = 0$$

and so $u \leq 1$ in $B_{1/2}$, proving (5.7). $\square$

We now proceed to the second step of De Giorgi's theorem, and we show that bounded solutions are Hölder continuous.

**Theorem 5.15** (From $L^\infty$ to $C^{0,\alpha}$)**.** *Let $u \in W^{1,2}(B_1)$ be such that $Lu = 0$. Then*

$$\|u\|_{C^{0,\alpha}(B_{1/2})} \leq C(n,\lambda,\Lambda)\|u\|_{L^\infty(B_1)}$$

*for some exponent $\alpha = \alpha(n,\lambda,\Lambda) > 0$.*

As we will see in a moment, Theorem 5.15 is a relatively straightforward consequence of the following key result:

**Lemma 5.16** (Oscillation lemma)**.** *Let $u \in W^{1,2}(B_2)$ be such that*

$$u \leq 1 \text{ in } B_2, \qquad Lu \leq 0 \text{ in } B_2.$$

*There is $\theta = \theta(\mu,n,\lambda,\Lambda) > 0$ such that*

$$0 < \mu \leq |\{u \leq 0\} \cap B_1| \quad \Longrightarrow \quad \sup_{B_{1/2}} u \leq 1-\theta.$$

**Proof.** Similarly to the proof of Theorem 5.13, we consider the sequence of subsolutions of $L$ given by

$$u_k \equiv 2^k(u-(1-2^{-k}))_+.$$

Note that $0 \leq u_k \leq 1$ in $B_2$, since $u \leq 1$, that $u_0 = u_+$, and moreover let us register three basic properties of this sequence:

(i) the sets $\{0 < u_k < \frac{1}{2}\}$ are disjoint for all $k \in \mathbb{N}$;

(ii) $\{u_{k+1} > 0\} \subseteq \{u_k > \frac{1}{2}\}$;

(iii) $\{u \leq 0\} \subset \{u_k = 0\}$.

Let us fix $\delta > 0$, to be chosen later, and suppose that $0 < \mu \leq |\{u \leq 0\} \cap B_1|$. We claim that there is $k_0 = k_0(n,\delta,\mu)$ such that

$$\int_{B_1} u_{k_0}^2 \,\mathrm{d}x \leq \delta^2. \tag{5.8}$$

Indeed, suppose that $u_0,\ldots,u_{k+1}$ do not satisfy (5.8). Then from (ii) we obtain

$$|\{u_k \geq \tfrac{1}{2}\} \cap B_1| \geq |\{u_{k+1} > 0\} \cap B_1| \geq \int_{B_1} u_{k+1}^2 \,\mathrm{d}x \geq \delta^2$$

while from (iii) we obtain

$$|\{u_k = 0\} \cap B_1| \geq \mu,$$

and so we can use Problem 1.4 to conclude that

$$\|\mathrm{D}u_k\|_{L^2(B_1)}^2 |\{0 < u_k < \tfrac{1}{2}\} \cap B_1| \geq c(n)\delta^4\mu^2.$$

Caccioppoli's inequality gives us an upper bound on the energy of the sequence, since

$$\int_{B_1} |\mathrm{D}u_k|^2 \,\mathrm{d}x \leq C \int_{B_2} u_k^2 \,\mathrm{d}x \leq C.$$

The last two inequalities thus show that

$$|\{0 < u_k < \tfrac{1}{2}\} \cap B_1| \geq c(n, \lambda, \Lambda, \mu)\delta^4,$$

an estimate which is uniform in $k$. Thus from (i) we see that for $k_0$ sufficiently large, depending only on $c$ and $\delta$, (5.8) must hold.

Now let $C$ be the constant in Theorem 5.13 and let us choose $\delta \leq \frac{1}{2C}$. Applying Theorem 5.13 and (5.8), we find that

$$\|u_{k_0}\|_{L^\infty(B_{1/2})} \leq C\|u_{k_0}\|_{L^2(B_1)} \leq C\delta \leq \frac{1}{2}.$$

Thus, recalling the definition of $u_k$, we have

$$u \leq \frac{1}{2}2^{-k_0} + (1 - 2^{-k_0}) = 1 - 2^{-(k_0+1)} \equiv 1 - \theta,$$

as wished. □

**Remark 5.17.** Note that Lemma 5.16 can be seen as a *quantitative* form of the strong maximum principle from Lemma 5.11. Indeed, suppose that $u$ is a smooth non-constant subsolution such that $\sup_{B_2} u = 1$. Then while Lemma 5.11 says that $1 - \sup_{B_{1/2}} u > 0$, Lemma 5.16 asserts that one can in fact quantify the difference $1 - \sup_{B_{1/2}} u \geq \theta > 0$ just in terms of the ellipticity constants $\lambda, \Lambda$ and in terms of $\mu = |\{u \leq 0\}|$, which is a measure of how far $u$ is from being identically 1.

**Proof of Theorem 5.15.** We claim the following *oscillation decay estimate*: writing as usual $\mathrm{osc}_B u \equiv \sup_B u - \inf_B u$, there is a constant $\theta = \theta(n, \lambda, \Lambda) > 0$ with

$$\mathrm{osc}_{B_{1/2}} u \leq (1 - \theta)\,\mathrm{osc}_{B_2} u. \tag{5.9}$$

Once this is proved, the conclusion follows the same type of principle that we used in the proof of Lemma 3.15. Indeed, by translating $u$ it is enough to prove that it is Hölder continuous at 0, so let $k \in \mathbb{N}$ be such that $2^{-(k+1)} \leq |x| < 2^{-k}$. Then, by iterating (5.9) $k$-times, we have

$$\begin{aligned}
|u(x) - u(0)| &\leq \mathrm{osc}_{B_{2^{-k}}} u \\
&\leq (1 - \theta)^k \mathrm{osc}_{B_1} u \\
&\leq (1 - \theta)^k \left(2\|u\|_{L^\infty(B_1)}\right) \\
&= 2^{-\alpha k}\left(2\|u\|_{L^\infty(B_1)}\right) \leq 2\|u\|_{L^\infty(B_1)}(2|x|)^\alpha,
\end{aligned}$$

provided we set $\alpha \equiv -\log_2(1 - \theta)$.

In turn, (5.9) is a simple consequence of Lemma 5.16. Indeed, consider the function

$$v(x) \equiv \frac{2}{\mathrm{osc}_{B_2} u}\left(u(x) - \frac{\sup_{B_2} u + \inf_{B_2} u}{2}\right),$$

which is defined so that $\sup_{B_2} v = 1, \inf_{B_2} v = -1$. Hence we can assume (up to switching the sign of $v$) that $|\{v \leq 0\} \cap B_1| \geq \frac{1}{2}|B_1|$ and so, applying Lemma 5.16 with $v$ instead of $u$, we find that

$$v \leq 1 - \theta \text{ in } B_{1/2} \quad \Longrightarrow \quad \mathrm{osc}_{B_{1/2}} v \leq 2 - \theta.$$

Rewriting this inequality in terms of $u$ yields $\mathrm{osc}_{B_{1/2}} u \leq (1 - \frac{\theta}{2})\,\mathrm{osc}_{B_2} u$, which is exactly (5.9) up to redefining $\theta$. □

## 5.3 Partial regularity in the vectorial case

The techniques used in the proof of Theorem 5.6 are extremely particular to the case $m = 1$, and do not carry over to the vectorial setting. It turns out that, in fact, Theorem 5.6 does not hold when $m > 1$, and there are many examples to that effect in the literature, going back to De Giorgi's example in 1968; see Problem 5.1. When $m > 1$, the minimizer of Theorem 5.2 is not $C^1$ in general, see [10] for an example in the lowest possible dimensions $n = 3, m = 2$. If $n \geq 5$ and $m \gg 1$, then one can even construct solutions of (5.2) which are *unbounded* [14]; hence even the weaker assertion proved in Theorem 5.13 fails when $m > 1$.

Thus, in the vectorial case, singularities are present and we cannot hope for full regularity of minimizers. Instead, we can prove *partial regularity* theorems. In order to explain what partial regularity means, we introduce the following definition:

**Definition 5.18.** For $u\colon \Omega \to \mathbb{R}^m$ we define the *regular set* by

$$\Omega_{\mathrm{reg}} \equiv \{x \in \Omega : \exists r > 0 \text{ such that } B_r(x) \subset \Omega \text{ and } u \in C^1(B_r(x), \mathbb{R}^m)\}$$

and the *singular set* by

$$\Sigma(u) \equiv \Omega \setminus \Omega_{\mathrm{reg}}.$$

Note that $\Omega_{\mathrm{reg}}$ is open and that, by Corollary 5.4 and the discussion after it, if $u$ is the minimizer obtained in Theorem 5.2 then we have

$$u \in C^\infty(\Omega_{\mathrm{reg}}),$$

justifying our choice of terminology.

By the counter-examples mentioned above, in general we have $\Sigma(u) \neq \emptyset$. A partial regularity theorem asserts that, nonetheless, this set is not too large, e.g. it asserts that $\Sigma(u)$ is a null set [4, 8, 12]. This is still a relatively weak statement, since null sets can be large in a topological sense. In order to make more precise statements, let us recall here the definition of Hausdorff measure.

**Definition 5.19.** For a subset $E \subset \mathbb{R}^n$ and $k \geq 0$, we define the *Hausdorff measure* of $E$ as

$$\mathscr{H}^k(E) \equiv c_k \lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{\infty} \operatorname{diam}(U_i)^k : E \subseteq \bigcup_{i=1}^{\infty} U_i, \operatorname{diam}(U_i) \leq \delta \right\},$$

where $c_k \equiv \omega_k / 2^k$ and $\omega_k \equiv \pi^{k/2}\Gamma(1 + k/2)$ is a generalization of the formula for the volume of $B_1 \subset \mathbb{R}^k$ to the case where $k$ is not necessarily a natural number.

The goal of this subsection is to prove the following:

**Theorem 5.20** (Partial regularity)**.** *Assume* (5.1) *holds and that* $F\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ *is smooth. If* $u$ *is the minimizer of Theorem 5.2 then its regular set can be characterized as*

$$\Omega_{\mathrm{reg}} = \left\{ x \in \Omega : \lim_{r \to 0} (\mathrm{D}u)_{x,r} = \mathrm{D}u(x) \text{ and } \lim_{r \to 0} \fint_{B_r(x)} |\mathrm{D}u - (\mathrm{D}u)_{x,r}|^2 = 0 \right\}. \tag{5.10}$$

*In particular, for any* $\varepsilon > 0$ *we have*

$$\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0.$$

Before proceeding with the proof of Theorem 5.20, we make two remarks.

**Remark 5.21.** Any function $f \in L^1_{\mathrm{loc}}(\Omega)$ admits a privileged representative, known as the *precise representative*:

$$f^*(x) = \begin{cases} \lim_{r\to 0} \fint_{B_r(x)} f \,\mathrm{d}y & \text{if this limit exists,} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly if $f = g$ a.e. in $\Omega$ we have $f^* = g^*$, thus this is indeed a meaningful definition. If moreover $f \in W^{1,p}_{\mathrm{loc}}(\Omega)$ is Sobolev, then we have finer information: there is a Borel set $E \subset \Omega$ such that

$$\Omega \setminus E = \left\{ x \in \Omega : \lim_{r\to 0} (f)_{x,r} = f^*(x) \text{ and } \lim_{r\to 0} \fint_{B_r(x)} |f - f^*|^p \,\mathrm{d}y = 0 \right\}.$$

and, moreover,

$$\mathscr{H}^{n-p+\varepsilon}(E) = 0 \qquad \forall \varepsilon > 0.$$

This is a general measure-theoretic fact about Sobolev functions, see [5, §4.8, Theorem 1 and §4.7, Theorem 4].

**Remark 5.22.** Let us call $\Omega_0$ the set on the right-hand side of (5.10). By the previous remark, the bound $\mathscr{H}^{n-2+\varepsilon}(\Sigma(u)) = 0$ follows from (5.10) and the fact that $\mathrm{D}u \in W^{2,2}_{\mathrm{loc}}(\Omega)$, according to Theorem 3.8.

Recalling that $\mathscr{H}^0$ is the counting measure, when $n = 2$ Theorem 5.20 implies that

$$\mathscr{H}^{\varepsilon}(\Sigma(u)) = 0 \quad \forall \varepsilon > 0,$$

while we already know from Theorem 5.5 that, in this case, $\Sigma(u) = \emptyset$. Thus Theorem 5.20 is optimal up to allowing $\varepsilon = 0$. This can in fact be done (one can even take $\varepsilon < 0$ not too small), but it is substantially more difficult.

A simple but key ingredient in the proof of Theorem 5.20 is the following nonlinear analogue of Caccioppoli's inequality from Theorem 3.5:

**Lemma 5.23** (Nonlinear Caccioppoli's inequality)**.** *Let $F$ satisfy* (5.1) *and let $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ be a solution of* (5.2)*. For any $\xi_0 \in \mathbb{R}^{m\times n}, u_0 \in \mathbb{R}^m$, we have*

$$\int_{B_r(x_0)} |\mathrm{D}u - \xi_0|^2 \,\mathrm{d}x \le \frac{C(\lambda, \Lambda)}{(R-r)^2} \int_{B_R(x_0)} |u(x) - u_0 - \xi_0 x|^2 \,\mathrm{d}x.$$

*whenever $r < R$ and $B_R(x_0) \subset \Omega$.*

**Proof.** Recall from the proof of Theorem 5.2 that (5.1) implies the following *monotonicity property* of $F'$:

$$\lambda |\xi_1 - \xi_2|^2 \le \langle F'(\xi_1) - F'(\xi_2), \xi_1 - \xi_2 \rangle.$$

Now let $\eta \in C^\infty_c(B_R(x_0))$ be such that $|\mathrm{D}\eta| \le 2/(R-r)$ and $\eta = 1$ in $B_r(x_0)$. Then

$$\begin{aligned} \int_{B_R(x_0)} \eta^2 |\mathrm{D}u - \xi_0|^2 \,\mathrm{d}x &\le \frac{1}{\lambda} \int_{B_R(x_0)} \eta^2 \langle F'(\mathrm{D}u) - F'(\xi_0), \mathrm{D}u - \xi_0 \rangle \,\mathrm{d}x \\ &= \frac{1}{\lambda} \int_{B_R(x_0)} \eta^2 \langle F'(\mathrm{D}u) - F'(\xi_0), \mathrm{D}[u(x) - u_0 - \xi_0 x] \rangle \,\mathrm{d}x \\ &= \frac{-2}{\lambda} \int_{B_R(x_0)} \eta \langle F'(\mathrm{D}u) - F'(\xi_0), [u(x) - u_0 - \xi_0 x] \otimes \mathrm{D}\eta \rangle \,\mathrm{d}x \end{aligned}$$

$$\le \frac{2\Lambda}{\lambda}\int_{B_R(x_0)} \eta|\mathrm{D}u-\xi_0||u(x)-u_0-\xi_0 x||\mathrm{D}\eta|\,\mathrm{d}x$$
$$\le C\left(\int_{B_R(x_0)} \eta^2|\mathrm{D}u-\xi_0|^2\,\mathrm{d}x\right)^{\frac12}\left(\int_{B_R(x_0)} \frac{|u(x)-u_0-\xi_0 x|^2}{(R-r)^2}\,\mathrm{d}x\right)^{\frac12}$$

where we used (5.2) when integrating by parts. Rearranging yields the conclusion. □

A partial regularity theorem, as in Theorem 5.20, is better thought of as an *$\varepsilon$-regularity theorem*. By this we simply mean that there is a suitable quantity which, if smaller than a fixed constant $\varepsilon$ in a ball of radius $r$, ensures that the solution is regular in the ball of radius $r/2$, with a corresponding estimate. In the context of Theorem 5.20, regularity means $C^{1,\alpha}$-regularity (since we can then bootstrap) and so, as in Section 4, the relevant quantity is

$$\mathrm{Exc}(u,x,r) \equiv \left(\fint_{B_r(x)} |\mathrm{D}u-(\mathrm{D}u)_{x,r}|^2\,\mathrm{d}y\right)^{1/2},$$

known as the excess of $u$. Thus the next lemma is the key step towards Theorem 5.20: if the excess is small then it decays.

**Lemma 5.24** (Excess decay)**.** *Let $u$ be the minimizer obtained in Theorem 5.2. For every $\tau \in (0,\frac14)$ and $M \ge 0$ there are $C=C(n,m,F)$ and $\varepsilon_0=\varepsilon_0(n,m,F,M,\tau)$ such that*

$$\left.\begin{aligned} &\mathrm{Exc}(u,x,r)\le\varepsilon_0\\ &|(\mathrm{D}u)_{x,r}|\le M\end{aligned}\right\}\quad\Longrightarrow\quad \mathrm{Exc}(u,x,\tau r)\le C\tau\,\mathrm{Exc}(u,x,r).$$

**Proof. Step 1.** We claim that the inequality holds with a constant $C_*$, which will be chosen later independently of $\tau, M$. If not, then there is a fixed pair $(\tau, M)$ and there are sequences of minimizers $u_k$ and balls $B_{r_k}(x_k)\subset\Omega$ such that

$$\varepsilon_k \equiv \mathrm{Exc}(u_k,x_k,r_k)\to 0,\qquad \mathrm{Exc}(u_k,x_k,\tau r_k) > C_*\tau\varepsilon_k, \tag{5.11}$$

where moreover we have

$$|\xi_k|\le M,\qquad \xi_k\equiv(\mathrm{D}u_k)_{x_k,r_k}.$$

Similarly to the proof of Theorem 1.19, let us define a new sequence

$$v_k(x)\equiv\frac{u(x_k+r_k x)-(u)_{x_k,r_k}-r_k\xi_k x}{\varepsilon_k r_k},\qquad \mathrm{D}v_k(x)=\frac{\mathrm{D}u(x_k+r_k x)-\xi_k}{\varepsilon_k},$$

which is normalized so that

$$(v_k)_{1,0}=0,\qquad (\mathrm{D}v_k)_{1,0}=0,\qquad \fint_{B_1(0)}|\mathrm{D}v_k|^2\,\mathrm{d}x=1.$$

In particular, by the Poincaré inequality we have $\int_{B_1(0)}|v_k|^2\,\mathrm{d}x\le C$, while (5.11) yields

$$\mathrm{Exc}(v_k,0,\tau)>C_*\tau. \tag{5.12}$$

By passing to subsequences, we can further assume that

$$\xi_k\to\xi,\qquad v_k\to v \text{ in } L^2(B_1,\mathbb{R}^m),\qquad \mathrm{D}v_k\rightharpoonup \mathrm{D}v \text{ in } L^2(B_1,\mathbb{R}^{m\times n}).$$

Finally, observe that since $u$ solves (5.2), $v_k$ satisfy

$$\frac{1}{\varepsilon_k}\int_{B_1(0)}\langle F'(\xi_k+\varepsilon_k \mathrm{D}v_k)-F'(\xi_k), \mathrm{D}\varphi\rangle\,\mathrm{d}x=0 \tag{5.13}$$

for any test function $\varphi\in C_c^\infty(B_1,\mathbb{R}^m)$.

**Step 2.** We now claim that the limit $v$ is a weak solution of

$$\operatorname{div}(A\,\mathrm{D}v)=0, \qquad A\equiv F''(\xi). \tag{5.14}$$

Indeed, by writing the difference quotient in (5.13) in the form

$$\frac{F'(\xi_k+\varepsilon_k\mathrm{D}v_k)-F'(\xi_k)}{\varepsilon_k}=\int_0^1 F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)\mathrm{D}v_k\,\mathrm{d}t$$

the claim follows from (5.13) through Fubini's theorem, provided that we show that

$$\lim_{k\to\infty}\int_{B_1}|F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)-A|^2\,\mathrm{d}x=0 \tag{5.15}$$

for any $t\in(0,1)$, since by the weak convergence $\mathrm{D}v_k\rightharpoonup \mathrm{D}v$ in $L^2$ we have

$$\begin{aligned}
&\left|\int_{B_1(0)}\langle F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)\mathrm{D}v_k-A\mathrm{D}v,\mathrm{D}\varphi\rangle\,\mathrm{d}x\right|\\
&\quad\le\left|\int_{B_1(0)}\langle[F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)-A]\mathrm{D}v_k,\mathrm{D}\varphi\rangle\,\mathrm{d}x\right|+\left|\int_{B_1(0)}\langle A(\mathrm{D}v_k-\mathrm{D}v),\mathrm{D}\varphi\rangle\right|\\
&\quad\le C\left(\int_{B_1(0)}|F''(\xi_k+t\varepsilon_k+\mathrm{D}v_k)-A|^2\,\mathrm{d}x\right)^{1/2}+o(1).
\end{aligned}$$

Now (5.15) can be seen by choosing a large constant $\ell$ and looking separately at the regions $\{|\mathrm{D}v_k|\le\ell\}$ and $\{|\mathrm{D}v_k|>\ell\}$. Clearly we have

$$\lim_{k\to\infty}\int_{\{|\mathrm{D}v_k|\le\ell\}\cap B_1}|F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)-A|^2\,\mathrm{d}x=0,$$

while for the complementary region we can estimate

$$\int_{\{|\mathrm{D}v_k|>\ell\}\cap B_1}|F''(\xi_k+t\varepsilon_k\mathrm{D}v_k)-A|^2\,\mathrm{d}x\le(\Lambda+|A|)\left|\{|\mathrm{D}v_k|>\ell\}\cap B_1\right|\le C\ell^{-2},$$

since $⨍_{B_1(0)}|\mathrm{D}v_k|^2\,\mathrm{d}x=1$. Thus the above integral goes to 0 uniformly in $k$ as $\ell\to\infty$, and so the claim follows.

**Step 3.** Through the above blow-up procedure, we have now obtained a solution $v$ of the constant-coefficient system (5.14). Hence, by Lemma 3.13, we find that

$$⨍_{B_{2\tau}}|\mathrm{D}v-(\mathrm{D}v)_{0,2\tau}|^2\,\mathrm{d}x\le 4C\tau^2⨍_{B_1}|\mathrm{D}v|^2\,\mathrm{d}x=4C\tau^2. \tag{5.16}$$

On the other hand, by (5.12), Exercise 3.12 and Lemma 5.23 we have

$$C_*^2\tau^2\le\operatorname{Exc}(v_k,0,\tau)^2\le⨍_{B_\tau}|\mathrm{D}v_k-(\mathrm{D}v_k)_{0,2\tau}|^2\,\mathrm{d}x\le\frac{C}{\tau^2}⨍_{B_{2\tau}}|v_k(x)-(v_k)_{0,2\tau}-(\mathrm{D}v_k)_{0,2\tau}x|^2\,\mathrm{d}x.$$

Thus, sending $k \to \infty$ and using the Poincaré inequality and (5.16), we find that

$$C_*^2\tau^2 \le \frac{C}{\tau^2} \fint_{B_{2\tau}} |v(x) - (v)_{0,2\tau} - (\mathrm{D}v)_{0,2\tau}x|^2 \,\mathrm{d}x \le C \fint_{B_{2\tau}} |\mathrm{D}v - (\mathrm{D}v)_{0,2\tau}|^2 \,\mathrm{d}x \le C\tau^2.$$

By choosing $C_*$ larger than the square-root of this final constant $C$, we obtain the desired contradiction. □

**Proof of Theorem 5.20.** Let $C_*$ be the constant in Lemma 5.24 and fix $\tau \in (0, \frac{1}{4})$ so that $C_*\tau \le \frac{1}{2}$. Then, for each $M > 0$, let $\varepsilon_0 = \varepsilon_0(M)$ be as in Lemma 5.24, and let

$$\Omega_M \equiv \left\{ x \in \Omega : |(\mathrm{D}u)_{x,r}| \le \frac{M}{2} \text{ and } \operatorname{Exc}(u, x, r) \le \varepsilon(M) \text{ for some } B_r(x) \Subset \Omega \right\},$$

where $\varepsilon(M) \le \varepsilon_0(M)$ will be chosen later. Writing $\Omega_0$ for the set on the right-hand side of (5.10), we clearly have

$$\Omega_{\mathrm{reg}} \subseteq \Omega_0 = \bigcup_{M>0} \Omega_M,$$

and so we ought to prove that $\Omega_M \subset \Omega_{\mathrm{reg}}$ for each fixed $M > 0$. To this end, let $x \in \Omega_M$ and let $B_r(x) \Subset \Omega$ be the corresponding ball; we will show that $B_{r/2}(x) \subset \Omega_{\mathrm{reg}}$.

Given $y \in B_{r/2}(x)$, by Exercise 3.12 we have

$$\operatorname{Exc}(u, y, \tfrac{r}{2}) \le \left( \fint_{B_{r/2}(y)} |\mathrm{D}u - (\mathrm{D}u)_{x,r}|^2 \,\mathrm{d}z \right)^{\frac{1}{2}} \le 2^{n/2} \operatorname{Exc}(u, x, r) \le 2^{n/2}\varepsilon.$$

Since $B_{r/2}(y) \subset B_r(x)$, we can also estimate

$$\begin{aligned} |(\mathrm{D}u)_{y,\frac{r}{2}}| &\le \fint_{B_{r/2}(y)} |\mathrm{D}u - (\mathrm{D}u)_{x,r}| \,\mathrm{d}z + |(\mathrm{D}u)_{x,r}| \\ &\le 2^{n/2} \left( \fint_{B_r(x)} |\mathrm{D}u - (\mathrm{D}u)_{x,r}|^2 \,\mathrm{d}z \right)^{\frac{1}{2}} + |(\mathrm{D}u)_{x,r}| \\ &\le 2^{n/2} \operatorname{Exc}(u, x, r) + |(\mathrm{D}u)_{x,r}| \\ &\le 2^{n/2}\varepsilon + \frac{M}{2}. \end{aligned} \tag{5.17}$$

Thus, provided that we choose $\varepsilon$ small enough so that $2^{n/2}\varepsilon \le \varepsilon_0$ and $2^{n/2}\varepsilon + \frac{M}{2} \le M$, we can apply Lemma 5.24 in $B_{r/2}(y)$ to infer that

$$\operatorname{Exc}(u, y, \tau\tfrac{r}{2}) \le \frac{1}{2} \operatorname{Exc}(u, y, \tfrac{r}{2}) \le 2^{n/2}\varepsilon. \tag{5.18}$$

With an argument identical to (5.17), we have

$$|(\mathrm{D}u)_{y,\tau\frac{r}{2}}| \le \tau^{-n/2} \operatorname{Exc}(u, y, \tfrac{r}{2}) + |(\mathrm{D}u)_{y,\frac{r}{2}}| \le \left(\frac{2}{\tau}\right)^{n/2} \varepsilon + 2^{n/2}\varepsilon + \frac{M}{2},$$

and so by (5.18) we can apply Lemma 5.24 in $B_{\tau\frac{r}{2}}(y)$, provided we choose $\varepsilon$ even smaller.

As usual, we now want to iterate the last inequality, applying Lemma 5.24 consecutively. We proceed inductively, assuming that we can apply Lemma 5.24 in $B_{\tau^k\frac{r}{2}}, \dots, B_{\tau\frac{r}{2}}$. From

(5.18) we have

$$\operatorname{Exc}(u, y, \tau^k \tfrac{r}{2}) \leq 2^{-k} \operatorname{Exc}(u, y, \tau \tfrac{r}{2}) \leq 2^{-k} 2^{n/2} \varepsilon, \tag{5.19}$$

while arguing once again as in (5.17) we have

$$|(\mathrm{D}u)_{y,\tau^{k+1}\frac{r}{2}}| \leq \tau^{-n/2} \operatorname{Exc}(u, y, \tau^k \tfrac{r}{2}) + |(\mathrm{D}u)_{y,\tau^k \frac{r}{2}}| \leq 2^{-k} \Big(\frac{2}{\tau}\Big)^{n/2} \varepsilon + C(\tau, n)\varepsilon + \frac{M}{2}$$

and so, by choosing $\varepsilon = \varepsilon(M, \tau, n)$ sufficiently small, we can keep applying Lemma 5.24 at all scales. In particular we find that (5.19) holds for all $k \in \mathbb{N}$.

Having proved (5.19), we argue as we have done already many times in these notes: for any $s \in (0, \frac{r}{2})$, we can find $k \in \mathbb{N}$ so that $\tau^{k+1}\frac{r}{2} \leq s \leq \tau^k \frac{r}{2}$, and by applying (5.18) we deduce that

$$\operatorname{Exc}(u, y, s) \leq C(\tau, n) \Big(\frac{s}{r/2}\Big)^{\mu} \operatorname{Exc}(u, y, \tau \tfrac{r}{2}) \leq C(\tau, n, r) s^{\mu},$$

where $\mu \equiv -1/\log_2 \tau > 0$. Thus $\mathrm{D}u \in \mathcal{L}^{2,n+2\mu}(B_{r/2}(x), \mathbb{R}^{m\times n})$ and so $u \in C^{1,\mu}(B_{r/2}, \mathbb{R}^m)$ by Theorem 4.3(ii). This shows that $B_{r/2}(x) \subset \Omega_{\mathrm{reg}}$, as wished. □

## 5.4 Problems for Section 5

**Problem 5.1** (De Giorgi's counterexample)**.** The goal of this problem is to prove that, when $n \geq 3$ and $m > 1$, the conclusion of Theorem 5.6 does not hold; note that, when $n = 2$, by the hole-filling technique, we do have Hölder continuity of solutions.

Let $n = m$, $u_\gamma(x) \equiv |x|^\gamma x$, and consider a symmetric coefficient tensor $A$ defined by

$$\langle A(x)\xi, \xi\rangle = \bigg((n-2)\operatorname{tr}\xi + n\Big\langle \xi, \frac{x \otimes x}{|x|^2}\Big\rangle\bigg)^2 + |\xi|^2$$

for all $x \neq 0$ and all $\xi \in \mathbb{R}^{n\times n}$. Note that $A$ is bounded and satisfies the Legendre condition.

(i) Show that $u_\gamma \in C^\infty(\mathbb{R}^n \setminus \{0\})$ is a classical solution of $\operatorname{div}(A\mathrm{D}u_\gamma) = 0$ in $B_1(0) \setminus \{0\}$ if and only if

$$(2n-2)^2(\gamma + \tfrac{n}{2})^2 + \gamma n + \gamma^2 = 0. \tag{5.20}$$

(ii) Show that there is a solution $\gamma$ of (5.20) with $\gamma > -\frac{n}{2}$ and conclude that, for that solution, we have $u_\gamma \in W^{1,2}(B_1, \mathbb{R}^n)$.

(iii) Deduce from the previous two points and Problem 1.5 that $u_\gamma$ is a weak solution of $\operatorname{div}(A\mathrm{D}u_\gamma) = 0$. Then check that $u_\gamma \notin L^\infty(B_1)$ when $n \geq 3$ and that $u_\gamma \notin C^{1,1}(B_1)$ when $n = 2$.

**Problem 5.2** (Moser's iteration)**.** The goal of this problem is to give an alternative approach to Theorem 5.13, due to Moser. In fact this approach can also be used to obtain the full result in Theorem 5.6.

Let $n \geq 3$ and $u \in W^{1,2}(B_1)$ be such that $Lu \leq 0$, where $L$ is as in (5.5)–(5.6).

(i) By testing the equation against $\eta^2 u_+^\beta$, for $\beta \geq 1$, show that

$$\bigg(\int_{B_r(0)} u_+^{p\gamma}\,\mathrm{d}x\bigg)^{\frac{1}{\gamma}} \leq \frac{C(n,\lambda,\Lambda)}{(R-r)^2} \int_{B_R(0)} u_+^p\,\mathrm{d}x,$$

where $p \equiv \beta + 1 \geq 2$ and $\gamma \equiv 2^*/2 = \frac{n}{n-2} > 1$. This is a reverse Hölder inequality: an $L^p$-norm is controlling an $L^q$-norm for $q > p$!

(ii) As in the proof of Theorem 5.6, iterate the inequality in (i). More precisely, let

$$r_k \equiv \frac{1}{2} + \frac{1}{2^k}, \qquad p_k = 2\gamma^k,$$

and note that (i) yields

$$\|u_+\|_{L^{p_{k+1}}(B_{r_{k+1}})} \leq C(n, \lambda, \Lambda)^{1/\gamma^k} \|u_+\|_{L^{p_k}(B_{r_k})}.$$

Iterate this inequality from $k = 0$ to $k = \infty$ to conclude that

$$\|u_+\|_{L^\infty(B_{1/2})} \leq C\|u_+\|_{L^2(B_{3/2})}.$$

# Bibliography


[1] L. Ambrosio, A. Carlotto, and A. Massaccesi. *Lectures on Elliptic Partial Differential Equations.* Scuola Normale Superiore, Pisa, 2018.

[2] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 63(4):337–403, 1977.

[3] L. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, Rhode Island, 2010.

[4] L. C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Arch. Ration. Mech. Anal.*, 95(3):227–252, 1986.

[5] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition.* Chapman and Hall/CRC, New York, 1st edition, 2015.

[6] M. Giaquinta and L. Martinazzi. *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs.* Scuola Normale Superiore, Pisa, 2012.

[7] M. Giaquinta and J. Souček. Caccioppoli's inequality and Legendre-Hadamard condition. *Math. Ann.*, 270(1):105–107, 1985.

[8] E. Giusti and M. Miranda. Sulla regolarità delle soluzioni deboli di una classe di sistemi ellittici quasi-lineari. *Arch. Ration. Mech. Anal.*, 31(3):173–184, 1968.

[9] F. Gmeineder and J. Kristensen. Partial Regularity for BV Minimizers. *Arch. Ration. Mech. Anal.*, 232(3):1429–1473, 2019.

[10] G. Kitavtsev, G. Lauteri, S. Luckhaus, and A. Rüland. A Compactness and Structure Result for a Discrete Multi-well Problem with SO(n) Symmetry in Arbitrary Dimension. *Arch. Ration. Mech. Anal.*, 232(1):531–555, 2019.

[11] G. Leoni. *A First Course in Sobolev Spaces*, volume 181 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, Rhode Island, 2017.

[12] C. B. Morrey. Partial Regularity Results for Non-Linear Elliptic Systems. *J. Math. Mech.*, 17(7):649–670, 1968.

[13] J. Moser. A Sharp Form of an Inequality by N. Trudinger. *Indiana Univ. Math. J.*, 20(11):1077–1092, 1971.

[14] V. Šverák and X. Yan. Non-lipschitz minimizers of smooth uniformly convex functionals. *Proc. Natl. Acad. Sci. U. S. A.*, 99(24):15269–15276, 2002.

[15] K. Zhang. A counterexample in the theory of coerciveness for elliptic systems. *J. Partial Differ. Equations*, 2(3):79–82, 1989.

[16] K. Zhang. On the coercivity of elliptic systems in two dimensional spaces. *Bull. Aust. Math. Soc.*, 54(3):423–430, 1996.