Casimir Lindfors

Regularity of Minimizers

School of Science

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 29.4.2013

Thesis supervisor:

Prof. Juha Kinnunen

Thesis advisor:

D.Sc. (Tech.) Tuomo Kuusi

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
MASTER'S THESIS

Author: Casimir Lindfors

Title: Regularity of Minimizers

Date: 29.4.2013 Language: English Number of pages:6+51

Department of Mathematics and Systems Analysis

Professorship: Mathematics Code: T3020

Supervisor: Prof. Juha Kinnunen

Advisor: D.Sc. (Tech.) Tuomo Kuusi

In this thesis we study the variational problem

$$\min_{u \in \mathcal{A}} \int_{\Omega} F(Du)dx,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain, $\mathcal{A}$ the set of functions in $W^{1,2}(\Omega)$ with given boundary values, and F a smooth and strongly convex function. The aim is to show rigorously and in great detail that if we have a Lipschitz continuous minimizer, then it is, in fact, smooth. In order to prove the continuity of the first derivatives we use De Giorgi's method, and for the higher derivatives the classical Schauder theory is applied. The question whether variational minimizers are smooth is a slightly weaker version of Hilbert's 19th problem.

Keywords: Regularity, minimizer, calculus of variations, elliptic partial differential equation, Hilbert's problem, De Giorgi's method, Schauder theory

AALTO-YLIOPISTO
PERUSTIETEIDEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Casimir Lindfors

Työn nimi: Minimoijien säännöllisyys

Päivämäärä: 29.4.2013 Kieli: Englanti Sivumäärä:6+51

Matematiikan ja systeemianalyysin laitos

Professuuri: Matematiikka Koodi: T3020

Valvoja: Prof. Juha Kinnunen

Ohjaaja: TkT Tuomo Kuusi

Tässä työssä tutkitaan variaatio-ongelmaa

$$\min_{u \in \mathcal{A}} \int_{\Omega} F(Du)dx,$$

missä $\Omega \subset \mathbb{R}^n$ on rajoitettu alue, $\mathcal{A}$ joukko $W^{1,2}(\Omega)$-funktioita annetuilla reunaarvoilla ja F sileä ja vahvasti konveksi funktio. Tarkoituksena on näyttää täsmällisesti ja yksityiskohtaisesti, että jos meillä on Lipschitz-jatkuva minimoija, niin itse asiassa se on sileä. Ensimmäisten derivaattojen jatkuvuuden todistuksessa käytämme De Giorgin menetelmää, ja korkeampiin derivaattoihin sovellamme klassista Schauder-teoriaa. Kysymys variaatiominimoijien sileydestä on hieman heikompi versio Hilbertin 19:nnestä ongelmasta.

Avainsanat: Säännöllisyys, minimoija, variaatiolaskenta, elliptinen osittaisdifferentiaaliyhtälö, Hilbertin ongelma, De Giorgin menetelmä, Schauder-teoria

Preface

I would like to thank my advisor Tuomo Kuusi for his ever so helpful and patient guidance throughout the whole process of writing this thesis. It has truly been a privilege to learn from such an excellent mathematician.

I would also like to express my gratitude to my supervisor professor Juha Kinnunen for the useful discussions and all the material he has provided. Thanks are in order also to the rest of the nonlinear PDE research group at Aalto University for making me feel welcome and part of the group from the very beginning.

In addition, I am grateful to my family for their support and encouragement and for always being there when needed.

Last but not least, I wish to thank my fiancée, Katie, for her love and support and for bringing me joy and happiness also when mathematics has failed to do so.

Otaniemi, 29.4.2013

Casimir A. Lindfors

Contents

Abstract **ii**

Abstract (in Finnish) **iii**

Preface **iv**

Contents **v**

Basic Notation **vi**

1 Introduction **1**

2 Euler-Lagrange equation **6**
2.1 Preliminaries 6
2.2 Existence of a Lipschitz continuous minimizer 8
2.3 Second weak derivatives 10

3 Hölder continuity of first derivatives **18**
3.1 Weak maximum principle 18
3.2 Hölder continuity 23

4 Towards higher regularity **28**
4.1 Schauder theory 28
4.2 Campanato estimates 37

5 Smoothness of the minimizer **43**

6 Smoothness up to the boundary **48**

References **50**

Basic Notation

$\mathbb{R}^n$	n-dimensional Euclidean space
$\overline{A}$	closure of A
∂A	topological boundary of A
χ_A	characteristic function of A, 1 on A, otherwise 0
$\lvert A\rvert$	Lebesgue measure of $A \subset \mathbb{R}^n$
$A \setminus B$	$\{x \in A : x \notin B\}$
$\operatorname{dist}(A, B)$	Euclidean distance between sets A and B
$B \subset\subset A$	B is open and $\overline{B}$ is a compact subset of A
$\operatorname{supp} u$	support of u, closure of $\{x : u(x) \neq 0\}$
u_+	positive part of u, $\max\{u, 0\}$
$\operatorname{sgn}(u)$	sign of u, 1 when $u > 0$, -1 when $u < 0$
$\int_A u(x)dx$	integral of u on A with respect to Lebesgue measure
$(u)_A := \fint_A u(x)dx$	integral average, $\frac{1}{\lvert A\rvert}\int_A u(x)dx$
$L^p(A)$	functions u with $\int_A \lvert u\rvert^p\, dx < \infty$
$L^p_{\mathrm{loc}}(A)$	functions u with $\int_B \lvert u\rvert^p\, dx < \infty$ for every $B \subset\subset A$
$L^\infty(A)$	essentially bounded functions on A
$W^{k,p}(A)$	functions with weak derivatives up to order k in $L^p(A)$
$C^k(A)$	k times continuously differentiable functions on A
$C^\infty(A)$	functions in $C^k(A)$ for any k
$C^\infty_0(A)$	functions in $C^\infty(A)$ with compact support in A
$C^{k,\alpha}(A)$	functions with locally α-Hölder continuous derivatives up to order k on A
$C^{k,\alpha}(\overline{A})$	functions with globally α-Hölder continuous derivatives up to order k on A
$C^{0,1}(A)$	locally Lipschitz continuous functions on A
$\lvert\lvert u\rvert\rvert_{L^p(A)}$	L^p-norm of u, $\left(\int_A \lvert u\rvert^p\, dx\right)^{\frac{1}{p}}$
$\lvert\lvert u\rvert\rvert_{L^\infty(A)}$	$\operatorname{ess\,sup}_A \lvert u\rvert$
$B(x, r)$	open ball in $\mathbb{R}^n$ centered at x with radius r, $\{y \in \mathbb{R}^n : \lvert x - y\rvert < r\}$
$D^z_i u, \partial_{z_i} u$	partial derivative of u with respect to z_i, $\frac{\partial u}{\partial z_i}$
$D_i u$	$\frac{\partial u}{\partial x_i}$
Du	gradient of u, $(D_1 u, \ldots, D_n u)$
$D^\mu u$	$\frac{\partial^{\lvert\mu\rvert} u}{\partial x_1^{\mu_1} \cdots \partial x_n^{\mu_n}}$, where $\mu = (\mu_1, \ldots, \mu_n)$ is a multi-index and $\lvert\mu\rvert = \mu_1 + \ldots + \mu_n$
δ_{ij}	Kronecker delta, 1 when $i = j$, otherwise 0
$c = c(\cdot, \ldots, \cdot)$	positive constant depending only on the quantities in parentheses, may denote a different constant depending on the same arguments even within the same calculation

1 Introduction

One of the most important laws in physics is that nature strives to minimize the potential energy of any closed system. This energy can often be modeled with a functional of the type

$$\mathcal{F}(u) := \int_\Omega F(Du)dx,$$

where u is the state of the system, defined from the set Ω to real numbers, and F the so-called Lagrangian that is usually a convex function depending only on the gradient of u. Minimization of functionals of this type is central in a branch of analysis called the calculus of variations.

The purpose of this thesis is to give a detailed proof for the following: If u is a Lipschitz continuous solution of the minimization problem

$$\min_{u\in\mathcal{A}} \mathcal{F}(u), \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain, $\mathcal{A}$ the set of functions in $W^{1,2}(\Omega)$ with given boundary values u_0, and F a smooth and strongly convex function, then u is smooth as well. That is, we assume the existence of a solution in $W^{1,2}(\Omega)$ and, moreover, the boundedness of its gradient, and show that the solution has continuous derivatives up to any given order, provided the same is assumed of the Lagrangian F. Only local estimates are required, and thus regularity assumptions on the boundary are not necessary.

The calculus of variations can be said to have begun over 300 years ago, when Johann Bernoulli posed the brachistochrone curve problem [8], which asks for the curve of fastest descent under a gravitational field. The problem attracted the interest of many great mathematicians of the time. One of them was Leonhard Euler, whose work Elementa Calculi Variationum gave the science its name. Since then the calculus of variations has evolved into a significant field in mathematics, which not only has numerous applications, but is also of interest as such.

A typical example of an energy minimization problem is finding the minimum area of a surface with fixed boundary. The area of the graph of a function u on Ω is given by the functional

$$\int_\Omega \sqrt{1+|Du|^2}dx.$$

The Lagrangian $F(z) = \sqrt{1+|z|^2}$ is clearly strongly convex, and if the boundary $\partial\Omega$ is assumed smooth enough, a unique solution to the corresponding minimization problem can be found, provided u has fixed boundary values. Such a minimal surface is formed in nature, for example, by a soap film stretching over a solid frame. Two examples of minimal surfaces can be seen in Figure 1.

Another example of a variational problem is finding the path of shortest optical length, which a beam of light will follow according to Fermat's principle. In fact, Bernoulli applied this principle in his solution to the brachistochrone curve problem, showing that the curve of fastest decent is a cycloid [6]. A similar problem that can

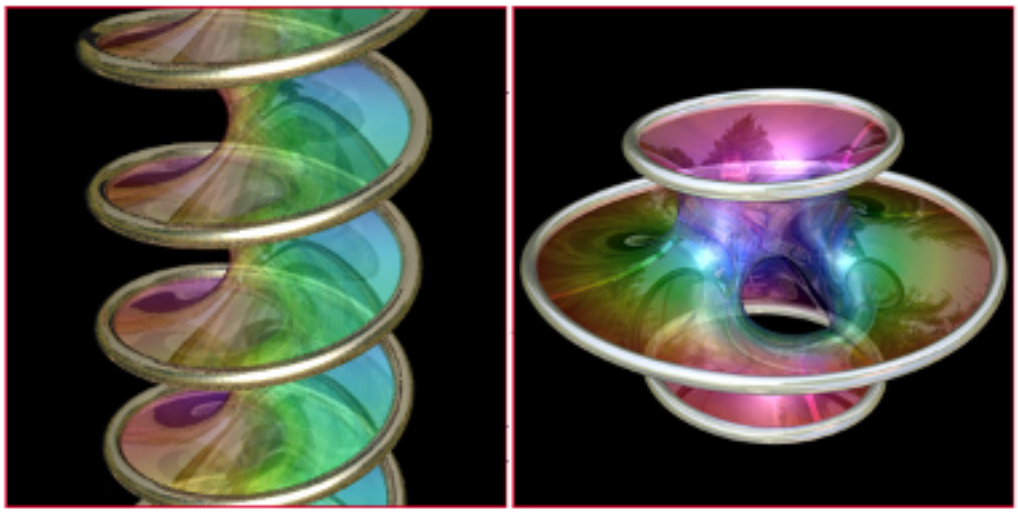

Figure 1: Two different minimal surfaces, a helicoid on the left and Costa's minimal surface on the right. (Graphics by Paul Nylander, `http://bugman123.com/`.)

also be solved using variational methods is determining the catenary curve, that is, the shape that a chain fixed at two points assumes under its own weight. Other applications for the calculus of variations are for instance isoperimetric problems, geodesics on manifolds, and optimal control theory.

In 1900 David Hilbert published his famous list of 23 mathematical problems [11], all of which unsolved at the time. Hilbert's 23 problems are generally considered the most influential compilation of open questions in mathematics, occupying a wide range of top mathematicians for over a century. Many of the problems have since been resolved, but some of them still remain open today. Knowing the importance in physics and the potential for applications that variational problems had, Hilbert included in his list two questions closely related to the calculus of variations, the 19th and 20th problems.

The 19th problem was among the ten problems, which Hilbert originally proposed at the International Congress of Mathematicians in Paris on August 8, 1900. His original formulation of the 19th problem was

> *Are the solutions of regular problems in the calculus of variations always necessarily analytic?*

The question is so broad it was solved in smaller pieces by several different mathematicians. The last piece in the puzzle was provided in 1957 by Ennio De Giorgi [5] and a year later by John Nash [16], who independently of each other showed that the solutions have Hölder continuous first derivatives. It was already known that this would be enough to prove real analyticity using Juliusz Schauder's estimates [17, 18], and on the other hand the existence of a Lipschitz continuous solution could be shown by applying direct methods in the calculus of variations, see [2] and the

references therein. A few years later Jürgen Moser gave a different proof for De Giorgi's and Nash's result [15]. He used a method now known as Moser iteration to prove a Harnack inequality, of which Hölder continuity is a simple consequence.

Hilbert originally presented his question in two dimensions, but the modern interpretation of the 19th problem is usually considered to be the following: Let $\Omega \subset \mathbb{R}^n$ be a bounded and suitably smooth domain and $\mathcal{A}$ the set of admissible functions. If F is a real analytic function with certain natural conditions such as convexity, are the solutions of the minimization problem

$$\min_{u \in \mathcal{A}} \int_\Omega F(Du)dx \tag{1.2}$$

real analytic as well? The set $\mathcal{A}$ typically consists of functions in a suitable function space satisfying certain boundary conditions, for example, $u = u_0$ on $\partial\Omega$, where u_0 is given. The problem is said to be regular, when the regularity assumptions on the Lagrangian F are satisfied.

How the set of admissible functions should be chosen was not clear to Hilbert. In his 20th problem he asks, whether all regular problems in the calculus of variations possess a solution, allowing the possibility to extend the notion of solution, if needed. It turned out this question could be answered positively, if the solution was sought in Sobolev spaces.

The problem considered in this work is basically a weaker version of Hilbert's 19th problem. To be precise, we prove the following theorem.

Theorem 1.1. *Let u be a Lipschitz continuous solution of problem* (1.1)*. Then $u \in C^\infty(\Omega)$.*

Achieving real analyticity would require a few additional arguments that are beyond the scope of this work. Moreover, showing the existence of a Lipschitz continuous solution requires extra assumptions on the boundary $\partial\Omega$ as well as the Lagrangian F. This can be done using direct methods in the calculus of variations, which we shall only briefly introduce in Section 2. A complete proof can be found, for example, in [10].

Variational minimization problems are directly related to partial differential equations through their corresponding Euler-Lagrange equations. The equation associated with problem (1.2) is

$$\sum_{i=1}^n D_i \partial_{z_i} F(Du) = 0, \tag{1.3}$$

in other words, any function u that solves (1.2) is a weak solution of the Euler-Lagrange equation (1.3) in Ω. Assuming u is smooth enough, this equation may be formally differentiated with respect to x_l for any $l = 1, \ldots, n$, which leads to the partial differential equation

$$\sum_{i,j=1}^n D_i(\partial_{z_i}\partial_{z_j} F(Du) D_j D_l u) = 0.$$

Thus, denoting $b_{ij}(x) := \partial_{z_i}\partial_{z_j}F(Du(x))$, we see that $w := D_l u$ satisfies the second order equation

$$\sum_{i,j=1}^{n} D_i(b_{ij}D_j w) = 0 \tag{1.4}$$

in Ω for every $l = 1, \ldots, n$. From the assumptions on F it follows that equation (1.4) is uniformly elliptic, and assuming the boundedness of Du it also has bounded coefficients. More precisely, for all $x \in \Omega$ and $\xi \in \mathbb{R}^n$ the coefficients b_{ij} satisfy

$$\sum_{i,j=1}^{n} b_{ij}(x)\xi_i\xi_j \geq \lambda\,|\xi|^2$$

and

$$\sum_{i,j=1}^{n} |b_{ij}(x)| \leq \Lambda$$

for some $0 < \lambda \leq \Lambda$.

Already in 1904 Sergei Bernstein showed in his doctoral thesis [1] that if the solutions of (1.4) are assumed three times continuously differentiable, then they are, indeed, necessarily real analytic. Although he only considered the two-dimensional case, at the time this was seen as the solution to Hilbert's 19th problem. However, such high regularity on the solution is not required in order to be able to state the problem. Moreover, the existence theory for the calculus of variations only gave a Lipschitz continuous solution, which at best could be shown to have second weak derivatives in L^2. Bernstein's results were also improved by several mathematicians lowering the a priori regularity assumptions on the solution, but the gap from $W^{2,2}$-solutions to continuous first derivatives could not be filled until the remarkable works of De Giorgi and Nash.

We now briefly describe the contents of this work. In Section 2 we introduce basic tools used in the thesis, such as Sobolev inequalities and the difference quotient operator. We also derive the Euler-Lagrange equation for problem (1.1) and apply it to prove the uniqueness of the minimizer u. As the main result of the section (Theorem 2.14) we show that the minimizer u has, in fact, second weak derivatives locally in $L^2(\Omega)$. This then implies in a straightforward manner that the derivatives of u satisfy equation (1.4).

Equation (1.4) is further studied in Section 3, eventually showing that the weak solutions of this equation are Hölder continuous (Theorem 3.6), which implies the continuity of the first derivatives of u. This is achieved by applying De Giorgi's iteration method, which together with a Caccioppoli inequality gives a so-called weak maximum principle. An oscillation estimate involving another iteration then yields the desired result.

In order to get our hands on the higher order derivatives of the minimizer u, equation (1.4) needs to be further differentiated. Repeating this process k times

leads to an equation of the type

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w^\mu) = \sum_{i=1}^{n} D_i g_i^\mu(x), \tag{1.5}$$

where w^μ is a kth order derivative of w and the functions g_i^μ depend on the derivatives of both w and the functions b_{ij} up to the same order. Weak solutions of this equation are studied in Section 4, and after some rather tedious lemmas, we are able to show that, under certain conditions, they belong to the space $C^{1,\alpha}$ (Theorem 4.6). The required conditions include the boundedness and Hölder continuity of the coefficients b_{ij} and g_i^μ. The proof is based on the observation that by freezing the coefficients at a given point, we obtain a much simpler equation, but at the same time the original equation acts locally as a perturbation of the new equation. The frozen equation can be further transformed into the Laplace's equation, which allows us to utilize the well-known properties of harmonic functions. This method of comparing the partial differential equation to one with constant coefficients is the essence of Schauder theory, named after Juliusz Schauder, who constructed it in [17] and [18]. However, Eberhard Hopf had already established the interior regularity for elliptic equations using similar ideas in [12] a couple of years before Schauder.

Finally, in Section 5 we combine the previous results so as to obtain the smoothness of the minimizer u. The main tool is the result proven in Section 4, Theorem 4.6, which we use as the inductive step. Equation (1.4), which is also of the type (1.5), acts as the basis of the induction. In order to show that the assumptions of Theorem 4.6 for this equation hold, we need Theorem 3.6, the crucial missing piece before the times of De Giorgi and Nash. In Section 6 we finish the thesis by briefly discussing the key steps required for extending the result up to the boundary.

2 Euler-Lagrange equation

In order to study the regularity of minimizers of a variational problem, it is useful to find the corresponding partial differential equation, often called the Euler-Lagrange equation. In this section we derive the Euler-Lagrange equation for the minimization problem (1.1) and prove that the minimizer u, which a priori is a $W^{1,2}(\Omega)$ function, in fact belongs to the space $W^{2,2}_{\text{loc}}(\Omega)$.

2.1 Preliminaries

Let us begin by introducing some notation and tools used in the thesis. More on basic Sobolev space theory can be read, for example, in [9], [10], and [20]. For the basic notation used in this thesis turn to page vi.

When dealing with variational problems, Sobolev spaces are the correct function spaces to work with. We say that a locally integrable scalar function u on $\Omega \subset \mathbb{R}^n$ is k times weakly differentiable, if it has derivatives up to order k in the sense of distributions. That is, for every multi-index μ with $|\mu| \leq k$ there exists a function $v^\mu \in L^1_{\text{loc}}(\Omega)$ such that

$$\int_\Omega u D^\mu \varphi dx = (-1)^{|\mu|} \int_\Omega v^\mu \varphi dx$$

for every $\varphi \in C_0^\infty(\Omega)$. We denote $v^\mu = D^\mu u$. If, moreover, all the weak derivatives together with u itself are in the space $L^p(\Omega)$ for some $1 \leq p \leq \infty$, u belongs to the Sobolev space $W^{k,p}(\Omega)$. The space $W^{k,p}(\Omega)$ can be shown to be a Banach space with the Sobolev norm

$$||u||_{W^{k,p}(\Omega)} := \sum_{|\mu| \leq k} ||D^\mu u||_{L^p(\Omega)} .$$

In the case $p = \infty$ the Sobolev space contains functions with essentially bounded derivatives. Furthermore, a function u is said to be in the space $W^{k,p}_{\text{loc}}(\Omega)$, if it belongs to $W^{k,p}(\Omega')$ for every $\Omega' \subset\subset \Omega$.

For $1 \leq p < \infty$ the Sobolev space $W^{k,p}(\Omega)$ can be equivalently characterized as the completion of $C^\infty(\Omega)$ with respect to the norm $||\cdot||_{W^{k,p}(\Omega)}$ [14]. Analogously, the Sobolev space with zero boundary values, denoted $W^{k,p}_0(\Omega)$, may be defined as the completion of $C_0^\infty(\Omega)$ with respect to the corresponding Sobolev norm. As a special case, functions with locally essentially bounded first derivatives, that is, functions in $W^{1,\infty}_{\text{loc}}(\Omega)$, are precisely the locally Lipschitz continuous functions on Ω [7, p. 131–132].

Sobolev functions have the following basic properties. Suppose $u, v \in W^{k,p}(\Omega)$ and let μ and ν be multi-indices such that $|\mu| + |\nu| \leq k$. Then

(i) $D^\mu u \in W^{k-|\mu|,p}(\Omega)$,

(ii) $D^\nu(D^\mu u) = D^\mu(D^\nu u)$,

(iii) $\alpha u + \beta v \in W^{k,p}(\Omega)$ for all $\alpha, \beta \in \mathbb{R}$,

(iv) $u \in W^{k,p}(\Omega')$ for all $\Omega' \subset\subset \Omega$,

(v) $u_+ := \max\{u, 0\} \in W^{k,p}(\Omega)$,

(vi) $\eta u \in W_0^{k,p}(\Omega)$ for all $\eta \in C_0^\infty(\Omega)$ and we have the generalized Leibniz rule

$$D^\mu(\eta u) = \sum_{\nu \le \mu} \binom{\mu}{\nu} D^{\mu-\nu}\eta D^\nu u,$$

where

$$\binom{\mu}{\nu} = \frac{\mu!}{\nu!(\mu-\nu)!}, \quad \mu! = \mu_1! \cdots \mu_n!,$$

and $\nu \le \mu$ whenever $\nu_j \le \mu_j$ for every $j = 1, \dots, n$.

Sobolev embeddings and Poincaré inequalities are essential in the theory of partial differential equations. We present here two versions suitable for our purposes without proofs. The Sobolev exponent for $1 \le p < n$ is denoted by

$$p^* := \frac{np}{n-p}.$$

Lemma 2.1. (Sobolev inequality) *Let $\Omega \subset \mathbb{R}^n$ be a bounded set, $1 \le p < n$, and $u \in W_0^{1,p}(\Omega)$. Then there exists a constant $c = c(n,p)$ such that*

$$\left(\int_\Omega |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} \le c\left(\int_\Omega |Du|^p\,dx\right)^{\frac{1}{p}}.$$

Proof. See for example [9, p. 155–157]. □

Remark 2.2. In the case $p = n$ we have exponential integrability for u [9, p. 162]. A straightforward calculation using the Taylor expansion of the exponential function shows that

$$\left(\int_\Omega |u|^q\,dx\right)^{\frac{1}{q}} \le c(n,q)\,|\Omega|^{\frac{1}{q}}\left(\int_\Omega |Du|^n\,dx\right)^{\frac{1}{n}} \tag{2.1}$$

for any $q \ge 1$.

Remark 2.3. If we define $p^* = 2n$, when $p = n$, then combining Lemma 2.1 and Remark 2.2 with $\Omega = B_r$ implies

$$\left(\fint_{B_r} |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} \le c(n,p)r\left(\fint_{B_r} |Du|^p\,dx\right)^{\frac{1}{p}} \tag{2.2}$$

for every $1 \le p \le n$.

Lemma 2.4. (Sobolev-Poincaré inequality) *Let $1 \le p < n$ and $u \in W^{1,p}_{loc}(\Omega)$. Then there exists a constant $c = c(n,p)$ such that*

$$\left(\int_{B_r} |u - (u)_{B_r}|^{p^*}\,dx\right)^{\frac{1}{p^*}} \le c\left(\int_{B_r} |Du|^p\,dx\right)^{\frac{1}{p}}$$

for every $B_r \subset\subset \Omega$.

Proof. See for example [10, p. 101–102]. □

2.2 Existence of a Lipschitz continuous minimizer

Let Ω be a bounded domain in $\mathbb{R}^n, n \geq 2$, and denote

$$\mathcal{A} := \left\{ u \in W^{1,2}(\Omega) : u - u_0 \in W_0^{1,2}(\Omega) \right\},$$

where $u_0 \in W^{1,2}(\Omega)$ is given. That is, $\mathcal{A}$ is the set of functions in $W^{1,2}(\Omega)$ with fixed boundary values u_0. We define the functional $\mathcal{F} : \mathcal{A} \to \mathbb{R}$ as

$$\mathcal{F}(u) := \int_\Omega F(Du)dx,$$

where $F : \mathbb{R}^n \to \mathbb{R}$ is a smooth and strongly convex function. By smooth we mean that it is continuously differentiable for any given order, in other words $F \in C^\infty(\mathbb{R}^n)$. Strong convexity is, as the name suggests, a stronger version of convexity, in fact, even stronger than strict convexity. To be precise, F is strongly convex with modulus $\lambda > 0$, if

$$F(tz_1 + (1-t)z_2) \leq tF(z_1) + (1-t)F(z_2) - \frac{1}{2}\lambda t(1-t)\left|z_1 - z_2\right|^2$$

for all $z_1, z_2 \in \mathbb{R}^n$ and $0 \leq t \leq 1$. Clearly any strongly convex function is also strictly convex, and when $\lambda \to 0$, the definition of strong convexity approaches that of convexity.

For each $z \in \mathbb{R}^n$ let $A(z)$ be the $n \times n$ matrix formed by the second derivatives of F at z, denoted by

$$a_{ij}(z) := \partial_{z_i}\partial_{z_j}F(z),$$

where $1 \leq i, j \leq n$. It can be shown that F is strongly convex with modulus λ if and only if the smallest eigenvalue of $A(z)$ is at least λ for every $z \in \mathbb{R}^n$. Thus, the functions a_{ij} satisfy the uniform ellipticity condition

$$\sum_{i,j=1}^{n} a_{ij}(z)\xi_i\xi_j \geq \lambda\left|\xi\right|^2 \tag{2.3}$$

for all $z, \xi \in \mathbb{R}^n$. Observe that $A(z)$ is, by definition, positive definite and, since F is smooth, symmetric for all $z \in \mathbb{R}^n$.

The reason for assuming strong convexity instead of strict convexity is that the latter does not guarantee uniform ellipticity. This can be seen by choosing, for example, $F(z) = |z|^4$, which is a strictly convex function, but $\partial_{z_i}\partial_{z_j}F(0) = 0$ for all $1 \leq i, j \leq n$.

In order to be able to talk about minimizing the functional $\mathcal{F}$, we need the following definition.

Definition 2.5. A function $u \in \mathcal{A}$ is a *minimizer* of the functional $\mathcal{F}$, if

$$\mathcal{F}(u) \leq \mathcal{F}(w)$$

for every $w \in \mathcal{A}$.

We prove the regularity of minimizers of the functional $\mathcal{F}$, Theorem 1.1, starting from the assumption that we have a minimizer, which is not only in $W^{1,2}(\Omega)$, but also Lipschitz continuous. This is far from a trivial assumption, and thus we shall briefly justify it here. The existence of a Lipschitz continuous solution with given boundary values can, indeed, be shown using the so-called direct methods in the calculus of variations. However, some regularity on the boundary of Ω and also on the boundary values must be assumed. We follow closely the proof shown in [10] and only highlight the main ideas and necessary assumptions.

First we need a couple of definitions.

Definition 2.6. A function $v \in C^{0,1}(\Omega)$ is a *super(sub)-minimum* for the functional $\mathcal{F}$ in Ω, if for all $w \in C^{0,1}(\Omega)$ with $w \geq v$ ($w \leq v$) and $w = v$ on $\partial\Omega$ we have

$$\mathcal{F}(v) \leq \mathcal{F}(w).$$

Let $d(x)$ denote the distance of $x \in \Omega$ from the boundary $\partial\Omega$ and set for $t > 0$

$$\Sigma_t := \{x \in \Omega : d(x) < t\},$$
$$\Gamma_t := \{x \in \Omega : d(x) = t\}.$$

Definition 2.7. A function v^+ is an *upper barrier* relative to the functional $\mathcal{F}$, if $v^+ = u_0$ on $\partial\Omega$, v^+ is a super-minimum in Σ_t, and $v^+ \geq \sup_{\partial\Omega} u_0$ on Γ_t for some $t > 0$. Similarly, v^- is a *lower barrier*, if $v^- = u_0$ on $\partial\Omega$, v^- is a sub-minimum in Σ_t, and $v^- \leq \inf_{\partial\Omega} u_0$ on Γ_t.

In order to prove that there exists a Lipschitz continuous solution for the minimization problem with boundary values u_0, it suffices to find an upper barrier and a lower barrier relative to $\mathcal{F}$. The only question is how these barriers can be constructed.

To this end, we assume that the boundary of Ω is twice continuously differentiable. Moreover, the function u_0 is assumed to be the restriction of a function in $C^2(\mathbb{R}^n)$ to Ω, which we also denote by u_0. If we further make the technical assumption

$$\limsup_{|z|\to\infty} \frac{|z|\,\Lambda(z)}{\mathcal{E}(z)} < \infty,$$

where $\mathcal{E}(z) := \sum_{i,j=1}^n a_{ij}(z) z_i z_j$ and $\Lambda(z)$ is the largest eigenvalue of the matrix $A(z)$, it is possible to show the existence of upper and lower barriers, and hence that of a Lipschitz continuous minimizer of $\mathcal{F}$ with boundary values u_0. A concrete upper barrier can be obtained by choosing

$$v^+(x) = u_0(x) + c\log(1 + \sigma d(x)),$$

where c and σ are suitable constants. A lower barrier may be constructed similarly. Therefore, assuming that we have a Lipschitz continuous minimizer of the functional $\mathcal{F}$ is justified.

As mentioned above, locally Lipschitz continuous functions are the same as $W^{1,\infty}_{\mathrm{loc}}$-functions. Globally this does not hold in general. However, since we only

aim for proving local smoothness, we may assume without losing generality that the gradient of the minimizer u is bounded up to the boundary. This can be seen as follows.

Suppose the result holds for minimizers in $W^{1,\infty}(\Omega)$. Now if u is only locally Lipschitz continuous in Ω, or equivalently belongs to $W^{1,\infty}_{\text{loc}}(\Omega)$, we have $u \in W^{1,\infty}(\Omega')$ for any $\Omega' \subset\subset \Omega$. Therefore, we may apply the result for Ω' to infer that $u \in C^\infty(\Omega')$. The arbitrariness of Ω' then implies that u is locally smooth over the whole domain Ω. Thus, from here on we shall assume that the minimizer u belongs to $W^{1,\infty}(\Omega)$, unless otherwise stated. Moreover, we shall always implicitly assume that $u \in \mathcal{A}$.

Let us denote

$$\Xi := \overline{B}\left(0, ||Du||_{L^\infty(\Omega)}\right).$$

Since F is smooth, the functions a_{ij} are continuous and thus obtain their maximum and minimum on the compact set Ξ. We denote

$$\Lambda := \max_{z\in\Xi} \sum_{i,j=1}^{n} |a_{ij}(z)|,$$

and by calculating

$$\sum_{i,j=1}^{n} a_{ij}(z)\xi_i\xi_j \leq \sum_{i,j=1}^{n} |a_{ij}(z)|\,|\xi|^2 \leq \Lambda\,|\xi|^2$$

we see that the eigenvalues of the matrix $A(z)$ are bounded also from above by Λ for all $z \in \Xi$.

2.3 Second weak derivatives

A necessary condition for a function to have a minimum at a certain point is that its first derivatives vanish at that point. The same applies when minimizing a functional, but in this case we require the first variation of the functional to vanish. This leads to the corresponding Euler-Lagrange equation.

Since a priori the minimizer u is only weakly differentiable, the Euler-Lagrange equation must be understood in a weak sense. Hence we need the concept of weak solutions. We give the definition for general partial differential equations of divergence type.

Definition 2.8. Let $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be a Carathéodory function, that is, measurable with respect to the first variable and continuous with respect to the second. Moreover, suppose $\mathcal{L}$ satisfies the growth condition

$$|\mathcal{L}(x,z)| \leq C\,(|z|+1)$$

for some constant C and for all $x \in \Omega, z \in \mathbb{R}^n$. A function $u \in W^{1,2}_{\text{loc}}(\Omega)$ is a *weak solution* of the equation

$$\sum_{i=1}^{n} D_i\mathcal{L}_i(x, Du) = 0$$

in Ω, if it satisfies

$$\int_\Omega \sum_{i=1}^n \mathcal{L}_i(x, Du) D_i v dx = 0 \tag{2.4}$$

for all test functions $v \in W_0^{1,2}(\Omega)$.

Let us then derive the Euler-Lagrange equation associated with $\mathcal{F}$.

Lemma 2.9. *Let $u \in W^{1,\infty}(\Omega)$ be a minimizer of the functional $\mathcal{F}$. Then it is a weak solution of the Euler-Lagrange equation*

$$\sum_{i=1}^n D_i \partial_{z_i} F(Du) = 0 \tag{2.5}$$

in Ω.

Proof. Since u is a minimizer of $\mathcal{F}$, we have

$$\int_\Omega F(Du) dx \leq \int_\Omega F(D(u + \varepsilon v)) dx$$

for every $v \in W_0^{1,2}(\Omega)$ and $\varepsilon \in \mathbb{R}$. Thus, the derivative

$$\frac{d}{d\varepsilon} \int_\Omega F(D(u + \varepsilon v)) dx = \int_\Omega \sum_{i=1}^n \partial_{z_i} F(D(u + \varepsilon v)) D_i v dx$$

must vanish at $\varepsilon = 0$ and we get

$$\int_\Omega \sum_{i=1}^n \partial_{z_i} F(Du) D_i v dx = 0 \tag{2.6}$$

for all $v \in W_0^{1,2}(\Omega)$. Hence, u is a weak solution of (2.5) in Ω. □

The Euler-Lagrange equation can be used to show the uniqueness of a minimizer with fixed boundary values.

Theorem 2.10. *A minimizer of the functional $\mathcal{F}$ in $W^{1,\infty}(\Omega)$ with given boundary values is unique.*

Proof. Let $u_1, u_2 \in W^{1,\infty}(\Omega)$ be minimizers of $\mathcal{F}$ such that $u_1 - u_2 \in W_0^{1,2}(\Omega)$. Thus, they are both weak solutions of the Euler-Lagrange equation, and by subtracting equation (2.6) for u_2 from that for u_1 we have

$$\int_\Omega \sum_{i=1}^n \left(\partial_{z_i} F(Du_1) - \partial_{z_i} F(Du_2) \right) D_i v dx = 0$$

for every $v \in W_0^{1,2}(\Omega)$. We write

$$\begin{aligned} \partial_{z_i} F(Du_1) - \partial_{z_i} F(Du_2) &= \int_0^1 \frac{d}{dt} \partial_{z_i} F(tDu_1 + (1-t)Du_2) dt \\ &= \int_0^1 \sum_{j=1}^n \partial_{z_i} \partial_{z_j} F(tDu_1 + (1-t)Du_2) D_j (u_1 - u_2) dt \end{aligned}$$

and choose $v = u_1 - u_2 \in W_0^{1,2}(\Omega)$ as the test function. This together with the ellipticity condition (2.3) and Sobolev inequality, Lemma 2.1, yields

$$\begin{aligned}
0 &= \int_\Omega \sum_{i=1}^n \int_0^1 \sum_{j=1}^n \partial_{z_i}\partial_{z_j} F(tDu_1 + (1-t)Du_2) D_j(u_1 - u_2) dt D_i(u_1 - u_2) dx \\
&= \int_\Omega \int_0^1 \sum_{i,j=1}^n a_{ij}(tDu_1 + (1-t)Du_2) D_i(u_1 - u_2) D_j(u_1 - u_2) dt dx \\
&\geq \int_\Omega \int_0^1 \lambda \left|D(u_1 - u_2)\right|^2 dt dx \\
&= \lambda \int_\Omega \left|D(u_1 - u_2)\right|^2 dx \\
&\geq \frac{\lambda}{c} \left(\int_\Omega |u_1 - u_2|^{2^*} dx \right)^{\frac{2}{2^*}}
\end{aligned}$$

for $n > 2$. When $n = 2$, we replace the last inequality with (2.1) and obtain

$$\left(\int_\Omega |u_1 - u_2|^q dx \right)^{\frac{2}{q}} \leq 0$$

for any $q \geq 1$. It follows that $||u_1 - u_2||_{L^{2^*}(\Omega)} = 0$, if we define $2^* = q$ when $n = 2$, and therefore we must have $u_1 = u_2$. □

Next we prove the main result of the section, that is, we show that the minimizer u actually belongs to $W^{2,2}_{\text{loc}}(\Omega)$. We use a similar technique as in proving uniqueness together with difference quotients. To this end, define for $f : \Omega \to \mathbb{R}$

$$\Delta_m^h f(x) := \frac{f(x + he_m) - f(x)}{h}$$

for all $x \in \Omega_{|h|} := \{x \in \Omega : \text{dist}(x, \partial\Omega) > |h|\}$, where e_m is the unit vector in the x_m direction and $h \neq 0$. The difference quotient operator Δ_m^h has the following useful properties.

Lemma 2.11. *(i) If $f \in W^{1,2}(\Omega)$, then $\Delta_m^h f \in W^{1,2}(\Omega_{|h|})$ and we have*

$$D_i(\Delta_m^h f) = \Delta_m^h (D_i f)$$

for $i = 1, \ldots, n$.

(ii) If $f, g \in L^2(\Omega)$ such that $\operatorname{supp} g \subset \Omega_{|h|}$, we have

$$\int_\Omega g \Delta_m^h f dx = -\int_\Omega f \Delta_m^{-h} g dx.$$

(iii) We have the Leibniz rule

$$\Delta_m^h (fg)(x) = f(x + he_m)\Delta_m^h g(x) + \Delta_m^h f(x) g(x).$$

Remark 2.12. We interpret

$$\int_\Omega g\Delta_m^h f dx = \int_{\Omega_{|h|}} g\Delta_m^h f dx,$$

whenever $\operatorname{supp} g \subset \Omega_{|h|}$. That is, we can write the integral over the whole domain Ω, even though $\Delta_m^h f$ is not defined near the boundary.

Proof of Lemma 2.11. (i) Let $f \in W^{1,2}(\Omega)$ and calculate using the linearity of the weak derivative

$$D_i(\Delta_m^h f(x)) = D_i \frac{f(x+he_m)-f(x)}{h} = \frac{D_i f(x+he_m) - D_i f(x)}{h} = \Delta_m^h(D_i f(x)).$$

(ii) Let $f, g \in L^2(\Omega)$ such that $\operatorname{supp} g \subset \Omega_{|h|}$. Then

$$\begin{aligned}
\int_\Omega g(x)\Delta_m^h f(x)dx &= \int_{\Omega_{|h|}} g(x)\frac{f(x+he_m)-f(x)}{h}dx \\
&= \frac{1}{h}\left(\int_{\Omega_{|h|}} f(x+he_m)g(x)dx - \int_{\Omega_{|h|}} f(x)g(x)dx\right) \\
&= \frac{1}{h}\left(\int_{\{x\in\Omega : x-he_m\in\Omega_{|h|}\}} f(x)g(x-he_m)dx - \int_{\Omega_{|h|}} f(x)g(x)dx\right) \\
&= -\int_\Omega f(x)\frac{g(x-he_m)-g(x)}{-h}dx \\
&= -\int_\Omega f(x)\Delta_m^{-h} g(x)dx.
\end{aligned}$$

(iii) A direct calculation gives

$$\begin{aligned}
\Delta_m^h(fg)(x) &= \frac{1}{h}(f(x+he_m)g(x+he_m) - f(x)g(x)) \\
&= \frac{1}{h}(f(x+he_m)(g(x+he_m)-g(x)) + (f(x+he_m)-f(x))g(x)) \\
&= f(x+he_m)\Delta_m^h g(x) + \Delta_m^h f(x)g(x).
\end{aligned}$$

□

We will also need the following standard lemma in the proof of the next theorem.

Lemma 2.13. (i) *Let $f \in W^{1,2}(\Omega)$ and $\Omega' \subset\subset \Omega$. Then*

$$\left|\left|\Delta_m^h f\right|\right|_{L^2(\Omega')} \leq ||D_m f||_{L^2(\Omega)}$$

for all $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$.

(ii) *Let $f \in L^2(\Omega)$ and $\Omega' \subset\subset \Omega$. If there exist constants $0 < h_0 \leq \operatorname{dist}(\Omega', \partial\Omega)$ and K such that*

$$\left|\left|\Delta_m^h f\right|\right|_{L^2(\Omega')} \leq K$$

for all $0 < |h| < h_0$, then $D_m f \in L^2(\Omega')$ and

$$||D_m f||_{L^2(\Omega')} \leq K.$$

Proof. (i) Let $\Omega' \subset\subset \Omega$ and fix h such that $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. Assume first that $f \in C^\infty(\Omega) \cap W^{1,2}(\Omega)$. Then we may write

$$\begin{aligned}
\Delta_m^h f(x) &= \frac{f(x + he_m) - f(x)}{h} \\
&= \frac{1}{h} \int_0^{|h|} \frac{d}{dt} f(x + t\operatorname{sgn}(h)e_m)dt \\
&= \frac{1}{h} \int_0^{|h|} \sum_{i=1}^n D_i f(x + t\operatorname{sgn}(h)e_m)\operatorname{sgn}(h)\delta_{mi}dt \\
&= \frac{1}{|h|} \int_0^{|h|} D_m f(x + t\operatorname{sgn}(h)e_m)dt
\end{aligned}$$

for all $x \in \Omega'$. Hence, by Hölder's inequality and Fubini's theorem

$$\begin{aligned}
\int_{\Omega'} \left|\Delta_m^h f(x)\right|^2 dx &\leq \int_{\Omega'} \left(\frac{1}{|h|} \int_0^{|h|} \left|D_m f(x + t\operatorname{sgn}(h)e_m)\right| dt \right)^2 dx \\
&\leq \frac{1}{|h|} \int_{\Omega'} \int_0^{|h|} \left|D_m f(x + t\operatorname{sgn}(h)e_m)\right|^2 dtdx \\
&= \frac{1}{|h|} \int_0^{|h|} \int_{\Omega'} \left|D_m f(x + t\operatorname{sgn}(h)e_m)\right|^2 dxdt \\
&\leq \frac{1}{|h|} \int_0^{|h|} \int_{\Omega} \left|D_m f(x)\right|^2 dxdt \\
&= \int_{\Omega} \left|D_m f(x)\right|^2 dx.
\end{aligned}$$

For the case $f \in W^{1,2}(\Omega)$ the result follows from the fact that $C^\infty(\Omega) \cap W^{1,2}(\Omega)$ is dense in $W^{1,2}(\Omega)$.

(ii) Let $f \in L^2(\Omega)$ and $\Omega' \subset\subset \Omega$. The uniform boundedness of $\Delta_m^h f$ in $L^2(\Omega')$ for $0 < |h| < h_0$ and the reflexivity of L^2 imply that there exists a function $g \in L^2(\Omega')$ with $||g||_{L^2(\Omega')} \leq K$ and a sequence $\{h_i\}$ tending to zero such that

$$\int_{\Omega'} \varphi \Delta_m^{h_i} f dx \to \int_{\Omega'} \varphi g dx$$

for every $\varphi \in C_0^\infty(\Omega')$ as $i \to \infty$. Therefore, by Lebesgue's dominated convergence theorem and Lemma 2.11 part (ii)

$$\int_{\Omega'} f D_m \varphi dx = \lim_{i\to\infty} \int_{\Omega'} f \Delta_m^{-h_i} \varphi dx = -\lim_{i\to\infty} \int_{\Omega'} \varphi \Delta_m^{h_i} f dx = -\int_{\Omega'} \varphi g dx.$$

Hence $g = D_m f$. □

Theorem 2.14. *Let $u \in W^{1,\infty}(\Omega)$ be a minimizer of the functional $\mathcal{F}$. Then $u \in W^{2,2}_{loc}(\Omega)$ and there exists a constant c depending only on λ, Λ and $\operatorname{dist}(\Omega', \partial\Omega)$ such that*

$$||D_m Du||_{L^2(\Omega')} \leq c\, ||Du||_{L^2(\Omega)}$$

for every $m = 1, \ldots, n$ and $\Omega' \subset\subset \Omega$.

Proof. Fix an arbitrary $\Omega' \subset\subset \Omega$ and choose $\Omega' \subset\subset \Omega'' \subset\subset \Omega$ such that

$$\operatorname{dist}(\Omega', \partial\Omega'') \geq \frac{1}{2}\operatorname{dist}(\Omega', \partial\Omega).$$

Take a cut-off function $\eta \in C_0^\infty(\Omega'')$ such that $0 \leq \eta \leq 1, \eta \equiv 1$ in Ω', and

$$|D\eta| \leq \frac{2}{\operatorname{dist}(\Omega', \partial\Omega'')} \leq \frac{4}{\operatorname{dist}(\Omega', \partial\Omega)}.$$

Since u is a minimizer of $\mathcal{F}$, it solves equation (2.6) for every $v \in W_0^{1,2}(\Omega)$. Now fix h such that $0 < |h| < \operatorname{dist}(\Omega'', \partial\Omega)$ and choose for $m = 1, \ldots, n$

$$v = -\Delta_m^{-h}(\eta^2 \Delta_m^h u) \in W_0^{1,2}(\Omega)$$

as the test function in (2.6). Since $\operatorname{supp}\eta \subset \Omega'' \subset \Omega_{|h|}$, we then have by Lemma 2.11

$$\begin{aligned} 0 &= \int_\Omega \sum_{i=1}^n \partial_{z_i} F(Du) D_i(-\Delta_m^{-h}(\eta^2 \Delta_m^h u))dx \\ &= -\int_\Omega \sum_{i=1}^n \partial_{z_i} F(Du) \Delta_m^{-h} D_i(\eta^2 \Delta_m^h u)dx \\ &= \int_\Omega \sum_{i=1}^n \Delta_m^h \partial_{z_i} F(Du) D_i(\eta^2 \Delta_m^h u)dx. \end{aligned} \tag{2.7}$$

We can further write

$$\begin{aligned} \Delta_m^h \partial_{z_i} F(Du(x)) &= \frac{1}{h}(\partial_{z_i} F(Du(x + he_m)) - \partial_{z_i} F(Du(x))) \\ &= \frac{1}{h}\int_0^1 \frac{d}{dt}\partial_{z_i} F(tDu(x + he_m) + (1-t)Du(x))dt \\ &= \frac{1}{h}\int_0^1 \sum_{j=1}^n \partial_{z_i}\partial_{z_j} F(tDu(x + he_m) + (1-t)Du(x)) D_j(u(x + he_m) - u(x))dt \\ &= \int_0^1 \sum_{j=1}^n a_{ij}(tDu(x + he_m) + (1-t)Du(x)) D_j \Delta_m^h u(x)dt, \end{aligned}$$

which combined with (2.7) gives

$$\int_\Omega \int_0^1 \sum_{i,j=1}^n a_{ij}(z_t) D_j \Delta_m^h u(x) D_i(\eta^2 \Delta_m^h u)dtdx = 0,$$

where $z_t := tDu(x + he_m) + (1-t)Du(x)$.

Together with the Leibniz rule and the ellipticity condition (2.3) we now get

$$\begin{aligned}
\lambda \int_\Omega \left|D\Delta_m^h u\right|^2 \eta^2 dx &= \int_\Omega \int_0^1 \lambda \left|D\Delta_m^h u\right|^2 \eta^2 dt dx \\
&\leq \int_\Omega \int_0^1 \sum_{i,j=1}^n a_{ij}(z_t) D_j \Delta_m^h u D_i \Delta_m^h u \eta^2 dt dx \\
&= -2 \int_\Omega \int_0^1 \sum_{i,j=1}^n a_{ij}(z_t) D_j \Delta_m^h u \eta D_i \eta \Delta_m^h u dt dx \\
&\leq 2 \int_\Omega \int_0^1 \sum_{i,j=1}^n |a_{ij}(z_t)| \left|D_j \Delta_m^h u\right| \eta \left|D_i \eta\right| \left|\Delta_m^h u\right| dt dx \\
&\leq 2\Lambda \int_\Omega \left|D\Delta_m^h u\right| \eta \left|D\eta\right| \left|\Delta_m^h u\right| dx.
\end{aligned}$$

For the last inequality we have used the fact that Ξ is convex, and thus for all $0 \leq t \leq 1$ and $x \in \Omega''$ we have $z_t \in \Xi$, so that

$$\sum_{i,j=1}^n |a_{ij}(z_t)| \leq \Lambda.$$

Next we use Young's inequality with ε, which yields

$$\int_\Omega \left|D\Delta_m^h u\right| \eta \left|D\eta\right| \left|\Delta_m^h u\right| dx \leq \frac{\varepsilon}{2} \int_\Omega \left|D\Delta_m^h u\right|^2 \eta^2 dx + \frac{1}{2\varepsilon} \int_\Omega \left|D\eta\right|^2 \left|\Delta_m^h u\right|^2 dx,$$

and by choosing $\varepsilon = \frac{\lambda}{2\Lambda}$ we arrive at

$$\lambda \int_\Omega \left|D\Delta_m^h u\right|^2 \eta^2 dx \leq \frac{\lambda}{2} \int_\Omega \left|D\Delta_m^h u\right|^2 \eta^2 dx + \frac{2\Lambda^2}{\lambda} \int_\Omega \left|D\eta\right|^2 \left|\Delta_m^h u\right|^2 dx.$$

This implies

$$\int_\Omega \left|D\Delta_m^h u\right|^2 \eta^2 dx \leq \frac{4\Lambda^2}{\lambda^2} \int_\Omega \left|D\eta\right|^2 \left|\Delta_m^h u\right|^2 dx,$$

and by using the properties of the cut-off function we deduce

$$\int_{\Omega'} \left|D\Delta_m^h u\right|^2 dx \leq \frac{64\Lambda^2}{\lambda^2 \operatorname{dist}(\Omega', \partial\Omega)^2} \int_{\Omega''} \left|\Delta_m^h u\right|^2 dx,$$

or $\left|\left|D\Delta_m^h u\right|\right|_{L^2(\Omega')} \leq c \left|\left|\Delta_m^h u\right|\right|_{L^2(\Omega'')}$, where $c := \frac{8\Lambda}{\lambda \operatorname{dist}(\Omega', \partial\Omega)}$.

Now by Lemma 2.13 part (i) we have

$$\left|\left|\Delta_m^h u\right|\right|_{L^2(\Omega'')} \leq \left|\left|D_m u\right|\right|_{L^2(\Omega)} \leq \left|\left|Du\right|\right|_{L^2(\Omega)},$$

and by Lemma 2.11 part (i)

$$D\Delta_m^h u = \Delta_m^h Du.$$

Together these give $\left|\left|\Delta_m^h Du\right|\right|_{L^2(\Omega')} \leq K$ for all $0 < |h| < h_0 := \operatorname{dist}(\Omega'', \partial\Omega)$, where $K := c\,||Du||_{L^2(\Omega)}$. Thus, we may apply part (ii) of Lemma 2.13 to $Du \in L^2(\Omega)$ to deduce $D_m Du \in L^2(\Omega')$ and

$$||D_m Du||_{L^2(\Omega')} \leq c\,||Du||_{L^2(\Omega)}\,.$$

This holds for every $m = 1, \ldots, n$, and since $\Omega' \subset\subset \Omega$ was arbitrary, we have $u \in W^{2,2}_{\text{loc}}(\Omega)$. □

Now that we have established the existence of the second weak derivatives for the minimizer u, we can show that the first weak derivatives satisfy a certain partial differential equation, which will be studied more in the following section.

Corollary 2.15. *Let $u \in W^{1,\infty}(\Omega)$ be a minimizer of the functional $\mathcal{F}$. Then $w := D_l u \in W^{1,2}_{loc}(\Omega)$ is a weak solution of the equation*

$$\sum_{i,j=1}^{n} D_i(a_{ij}(Du)D_j w) = 0 \tag{2.8}$$

in Ω for every $l = 1, \ldots, n$.

Proof. Take any $\varphi \in C_0^\infty(\Omega)$ and $l = 1, \ldots, n$ and test equation (2.6) with $v = -D_l\varphi$. Since $u \in W^{2,2}_{\text{loc}}(\Omega)$by Theorem 2.14, $w \in W^{1,2}_{\text{loc}}(\Omega)$ and we may integrate by parts, which gives

$$\begin{aligned}
0 &= -\int_\Omega \sum_{i=1}^{n} \partial_{z_i} F(Du) D_i D_l \varphi dx \\
&= \int_\Omega \sum_{i=1}^{n} \partial_{z_i} D_l F(Du) D_i \varphi dx \\
&= \int_\Omega \sum_{i=1}^{n} \partial_{z_i} \sum_{j=1}^{n} \partial_{z_j} F(Du) D_l D_j u D_i \varphi dx \\
&= \int_\Omega \sum_{i,j=1}^{n} \partial_{z_i}\partial_{z_j} F(Du) D_j D_l u D_i \varphi dx \\
&= \int_\Omega \sum_{i,j=1}^{n} a_{ij}(Du) D_j w D_i \varphi dx.
\end{aligned}$$

Since this holds for all $\varphi \in C_0^\infty(\Omega)$ and $C_0^\infty(\Omega)$ is dense in $W_0^{1,2}(\Omega)$, we obtain the result. □

3 Hölder continuity of first derivatives

In this section we prove that the first weak derivatives of the minimizer u are Hölder continuous using De Giorgi's iteration technique.

3.1 Weak maximum principle

By the result derived in the end of the previous section, Corollary 2.15, it suffices to consider weak solutions of the equation

$$\sum_{i,j=1}^{n} D_i(a_{ij}(Du)D_j w) = 0 \tag{3.1}$$

in Ω. Since $u \in W^{1,\infty}(\Omega)$, we may assume that $w \in L^\infty(\Omega)$.

To emphasize the dependence on x rather than Du, we write

$$b_{ij}(x) := a_{ij}(Du(x)),$$

and whenever no confusion can arise, we also omit the argument x. Now equation (3.1) can be written as

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w) = 0 \tag{3.2}$$

in Ω, and due to the ellipticity condition (2.3) and the boundedness of the functions a_{ij} on the set Ξ, the coefficients b_{ij} satisfy

$$\sum_{i,j=1}^{n} b_{ij}(x)\xi_i\xi_j \geq \lambda \left|\xi\right|^2 \tag{3.3}$$

and

$$\sum_{i,j=1}^{n} \left|b_{ij}(x)\right| \leq \Lambda \tag{3.4}$$

for all $\xi \in \mathbb{R}^n$ and $x \in \Omega$.

Let us first derive a Caccioppoli type estimate. In fact, this will be the only step where the equation is used directly.

Lemma 3.1. *Let $w \in W^{1,2}_{loc}(\Omega) \cap L^\infty(\Omega)$ be a weak solution of* (3.2)*. Then for every $k \in \mathbb{R}$ and $\eta \in C_0^\infty(B_\rho)$ we have*

$$\left(\fint_{B_\rho} \left|D(w-k)_+\right|^2 \eta^2 dx\right)^{\frac{1}{2}} \leq \frac{2\Lambda}{\lambda}\left(\fint_{B_\rho} \left|D\eta\right|^2 (w-k)_+^2 dx\right)^{\frac{1}{2}} \tag{3.5}$$

whenever $B_\rho \subset\subset \Omega$.

Remark 3.2. We denote $B_\rho := B(x_0, \rho)$, when the center $x_0 \in \Omega$ is not relevant.

Proof of Lemma 3.1. Let $\rho > 0$ be such that $B_\rho \subset\subset \Omega$ and let $k \in \mathbb{R}$ and $\eta \in C_0^\infty(B_\rho)$. Since w is a weak solution of (3.2), $w - k$ is as well and we have

$$\int_\Omega \sum_{i,j=1}^n b_{ij} D_j(w-k) D_i v dx = 0 \tag{3.6}$$

for every $v \in W_0^{1,2}(\Omega)$.

If we now choose $v = \eta^2(w-k)_+ \in W_0^{1,2}(B_\rho)$ as the test function, we see that v is non-zero only when $w > k$, and thus we may write the factor $D_j(w-k)$ as $D_j(w-k)_+$ in (3.6). The Leibniz rule and the ellipticity and boundedness assumptions (3.3) and (3.4) then yield

$$\begin{aligned}
\lambda \int_{B_\rho} |D(w-k)_+|^2 \eta^2 dx &\leq \int_\Omega \sum_{i,j=1}^n b_{ij} D_j(w-k)_+ D_i(w-k)_+ \eta^2 dx \\
&= -2 \int_\Omega \sum_{i,j=1}^n b_{ij} D_j(w-k)_+ \eta D_i \eta (w-k)_+ dx \\
&\leq 2 \int_\Omega \sum_{i,j=1}^n |b_{ij}| \, |D_j(w-k)_+| \, \eta \, |D_i \eta| \, (w-k)_+ dx \\
&\leq 2\Lambda \int_\Omega |D(w-k)_+| \, \eta \, |D\eta| \, (w-k)_+ dx.
\end{aligned}$$

Applying Young's inequality with ε, this can be further estimated from above by

$$\begin{aligned}
&2\Lambda \left(\frac{\varepsilon}{2} \int_\Omega |D(w-k)_+|^2 \eta^2 dx + \frac{1}{2\varepsilon} \int_\Omega |D\eta|^2 (w-k)_+^2 dx \right) \\
&= \frac{\lambda}{2} \int_{B_\rho} |D(w-k)_+|^2 \eta^2 dx + \frac{2\Lambda^2}{\lambda} \int_{B_\rho} |D\eta|^2 (w-k)_+^2 dx,
\end{aligned}$$

where we have chosen $\varepsilon = \frac{\lambda}{2\Lambda}$. Now the first term can be absorbed to the left hand side, which leads to

$$\int_{B_\rho} |D(w-k)_+|^2 \eta^2 dx \leq \frac{4\Lambda^2}{\lambda^2} \int_{B_\rho} |D\eta|^2 (w-k)_+^2 dx,$$

and by dividing by the measure of B_ρ and taking square roots, we are done. □

For the next lemma we introduce some useful notation. For every $k \in \mathbb{R}$ and $\rho > 0$ define

$$A(k,\rho) := \{x \in B_\rho : w(x) > k\}$$

and

$$\psi(k,\rho) := \left(\fint_{B_\rho} (w-k)_+^2 dx \right)^{\frac{1}{2}}.$$

The proof of the next lemma as well as the following theorem closely follow the proof of Theorem 1 in [19].

Lemma 3.3. *Let $w \in W^{1,2}_{loc}(\Omega) \cap L^\infty(\Omega)$ be a weak solution of (3.2) and let $r > 0$ be such that $B_{2r} \subset\subset \Omega$. Then there exist constants $\theta = \theta(n) > 0$ and $c = c(n, \lambda, \Lambda)$ such that the inequality*

$$\psi(k', \rho') \leq c \frac{\rho}{\rho - \rho'} \frac{1}{(k' - k)^\theta} \psi(k, \rho)^{1+\theta} \tag{3.7}$$

holds for every $k < k'$ and $r \leq \rho' < \rho \leq 2r$.

Proof. Fix $r > 0$ such that $B_{2r} \subset\subset \Omega$ and take any $r \leq \rho' < \rho \leq 2r$ and $k < k'$. From the definitions above it immediately follows that $A(k', \rho) \subset A(k, \rho)$ and

$$\begin{aligned} |A(k', \rho)| &= \frac{1}{(k' - k)^2} \int_{A(k', \rho)} (k' - k)^2 dx \\ &\leq \frac{1}{(k' - k)^2} \int_{A(k', \rho)} (w - k)^2 dx \\ &\leq \frac{1}{(k' - k)^2} \int_{A(k, \rho)} (w - k)^2 dx \\ &= \frac{|B_\rho|}{(k' - k)^2} \psi(k, \rho)^2. \end{aligned} \tag{3.8}$$

Choose a cut-off function $\eta \in C_0^\infty(B_\rho)$ such that $0 \leq \eta \leq 1, \eta \equiv 1$ in $B_{\rho'}$ and $|D\eta| \leq \frac{2}{\rho - \rho'}$. Denote

$$2^* := \begin{cases} \frac{2n}{n-2}, & n > 2 \\ 4, & n = 2 \end{cases},$$

and note that $|B_\rho| \leq 2^n |B_{\rho'}|$, since $\rho \leq 2r$ and $\rho' \geq r$. Hölder's inequality, Remark 2.3, and (3.8) then give

$$\begin{aligned} \psi(k', \rho') &= \left(\fint_{B_{\rho'}} (w - k')_+^2 dx \right)^{\frac{1}{2}} \\ &\leq |B_{\rho'}|^{-\frac{1}{2}} \left(\int_{B_\rho} ((w - k')_+ \eta)^2 dx \right)^{\frac{1}{2}} \\ &\leq |B_{\rho'}|^{-\frac{1}{2}} \left(\int_{B_\rho} ((w - k')_+ \eta)^{2^*} dx \right)^{\frac{1}{2^*}} |A(k', \rho)|^{\frac{1}{2} - \frac{1}{2^*}} \\ &= \left(\frac{|B_\rho|}{|B_{\rho'}|} \right)^{\frac{1}{2}} \left(\fint_{B_\rho} ((w - k')_+ \eta)^{2^*} dx \right)^{\frac{1}{2^*}} \left(\frac{|A(k', \rho)|}{|B_\rho|} \right)^{\frac{1}{2} - \frac{1}{2^*}} \\ &\leq c(n) \rho \left(\fint_{B_\rho} |D((w - k')_+ \eta)|^2 dx \right)^{\frac{1}{2}} \frac{1}{(k' - k)^{1 - \frac{2}{2^*}}} \psi(k, \rho)^{1 - \frac{2}{2^*}}. \end{aligned} \tag{3.9}$$

The last integral in (3.9) can be estimated by using the Leibniz rule, Minkowski's inequality, and the Caccioppoli estimate, Lemma 3.1. Thus,

$$\begin{aligned}
&\left(\fint_{B_\rho} |D((w-k')_+\eta)|^2\,dx\right)^{\frac{1}{2}} \\
&\quad\le \left(\fint_{B_\rho} |D(w-k')_+|^2\,\eta^2 dx\right)^{\frac{1}{2}} + \left(\fint_{B_\rho} |D\eta|^2\,(w-k')_+^2 dx\right)^{\frac{1}{2}} \\
&\quad\le \left(\frac{2\Lambda}{\lambda}+1\right)\left(\fint_{B_\rho} |D\eta|^2\,(w-k')_+^2 dx\right)^{\frac{1}{2}} \\
&\quad\le \left(\frac{2\Lambda}{\lambda}+1\right)\frac{2}{\rho-\rho'}\left(\fint_{B_\rho} (w-k)_+^2 dx\right)^{\frac{1}{2}} \\
&\quad= \frac{c(\lambda,\Lambda)}{\rho-\rho'}\psi(k,\rho).
\end{aligned} \tag{3.10}$$

We also used the fact that $k < k'$ implies $(w-k')_+ \le (w-k)_+$.

Denote

$$\theta(n) := 1 - \frac{2}{2^*} = \begin{cases} \frac{2}{n}, & n > 2 \\ \frac{1}{2}, & n = 2 \end{cases}.$$

Then combining (3.9) and (3.10) yields

$$\begin{aligned}
\psi(k',\rho') &\le c(n)\rho\frac{c(\lambda,\Lambda)}{\rho-\rho'}\psi(k,\rho)\frac{1}{(k'-k)^{1-\frac{2}{2^*}}}\psi(k,\rho)^{1-\frac{2}{2^*}} \\
&= c(n,\lambda,\Lambda)\frac{\rho}{\rho-\rho'}\frac{1}{(k'-k)^{\theta}}\psi(k,\rho)^{1+\theta},
\end{aligned}$$

as required. □

Now we are ready to apply De Giorgi's iteration scheme. This will give us a result often called the weak maximum principle. Observe that had we not already assumed the boundedness of w, it would also follow from this result.

Theorem 3.4. *Let $w \in W^{1,2}_{loc}(\Omega) \cap L^\infty(\Omega)$ be a weak solution of (3.2). Then there exists a constant $c = c(n,\lambda,\Lambda)$ such that*

$$\operatorname*{ess\,sup}_{B_r} w \le \tilde{k} + c\left(\fint_{B_{2r}} (w-\tilde{k})_+^2 dx\right)^{\frac{1}{2}} \tag{3.11}$$

for every $\tilde{k} \in \mathbb{R}$ and $r > 0$ such that $B_{2r} \subset\subset \Omega$.

Proof. Let $r > 0$ be such that $B_{2r} \subset\subset \Omega$ and fix $\tilde{k} \in \mathbb{R}$. Set for $m = 0, 1, 2, \ldots$

$$\begin{aligned}
k_m &= \tilde{k} + (1-2^{-m})d, \\
\rho_m &= (1+2^{-m})r,
\end{aligned}$$

where d is to be determined later. We observe that $k_0 = \tilde{k}$ and $\rho_0 = 2r$ and, moreover, k_m increases to $\tilde{k} + d$ and ρ_m decreases to r as m tends to infinity. Now applying Lemma 3.3 with $k = k_m, k' = k_{m+1}, \rho = \rho_m$, and $\rho' = \rho_{m+1}$ leads to

$$\begin{aligned}\psi(k_{m+1}, \rho_{m+1}) &\leq c\frac{\rho_m}{\rho_m - \rho_{m+1}}\frac{1}{(k_{m+1} - k_m)^\theta}\psi(k_m, \rho_m)^{1+\theta} \\ &\leq c2^{m+2}\frac{2^{\theta(m+1)}}{d^\theta}\psi(k_m, \rho_m)^{1+\theta} \\ &= c\frac{2^{(1+\theta)m}}{d^\theta}\psi(k_m, \rho_m)^{1+\theta}.\end{aligned} \tag{3.12}$$

Next we show by induction that for a suitably chosen d

$$\psi(k_m, \rho_m) \leq \frac{\psi(\tilde{k}, 2r)}{\sigma^m} \tag{3.13}$$

for all $m = 0, 1, 2 \ldots$ and some $\sigma > 1$ that will also be chosen shortly. Clearly the claim holds when $m = 0$. Assume then that it holds for m. By (3.12) and the induction assumption we may write

$$\begin{aligned}\psi(k_{m+1}, \rho_{m+1}) &\leq c\frac{2^{(1+\theta)m}}{d^\theta}\left(\frac{\psi(\tilde{k}, 2r)}{\sigma^m}\right)^{1+\theta} \\ &= c\sigma\left(\frac{\psi(\tilde{k}, 2r)}{d}\right)^\theta\left(\frac{2^{1+\theta}}{\sigma^\theta}\right)^m\frac{\psi(\tilde{k}, 2r)}{\sigma^{m+1}}.\end{aligned}$$

If we now choose $\sigma = 2^{1+\frac{1}{\theta}} > 1$ and d such that

$$c\sigma\left(\frac{\psi(\tilde{k}, 2r)}{d}\right)^\theta = 1,$$

or $d = c\psi(\tilde{k}, 2r)$, we see that the claim holds also for $m + 1$. Thus, it holds for all $m = 0, 1, 2, \ldots$ and by letting $m \to \infty$ in (3.13) we get $\lim_{m\to\infty}\psi(k_m, \rho_m) \leq 0$. Moreover, by Fatou's lemma

$$\begin{aligned}\left|\left|(w - (\tilde{k} + d))_+\right|\right|_{L^2(B_r)} &= |B_r|^{\frac{1}{2}}\left(\frac{1}{|B_r|}\int_\Omega (w - (\tilde{k} + d))_+^2\chi_{B_r}dx\right)^{\frac{1}{2}} \\ &\leq |B_r|^{\frac{1}{2}}\liminf_{m\to\infty}\left(\frac{1}{|B_{\rho_m}|}\int_\Omega (w - k_m)_+^2\chi_{B_{\rho_m}}dx\right)^{\frac{1}{2}} \\ &= |B_r|^{\frac{1}{2}}\lim_{m\to\infty}\psi(k_m, \rho_m).\end{aligned}$$

Therefore, $\left|\left|(w - (\tilde{k} + d))_+\right|\right|_{L^2(B_r)} = 0$, which implies $w \leq \tilde{k} + d$ almost everywhere in B_r. Recalling the choice of d then yields

$$\operatorname*{ess\,sup}_{B_r} w \leq \tilde{k} + c\left(\fint_{B_{2r}} (w - \tilde{k})_+^2 dx\right)^{\frac{1}{2}}. \qquad \square$$

3.2 Hölder continuity

Next we aim to prove that weak solutions of (3.2) are locally Hölder continuous. To this end, we denote

$$M(r) := \operatorname*{ess\,sup}_{B_r} w \quad \text{and} \quad m(r) := \operatorname*{ess\,inf}_{B_r} w,$$

both of which are finite numbers, since $w \in L^\infty(\Omega)$. First we prove an estimate for the oscillation of w in a ball, denoted by

$$\operatorname*{osc}_{B(x,r)} w := \operatorname*{ess\,sup}_{B(x,r)} w - \operatorname*{ess\,inf}_{B(x,r)} w.$$

Lemma 3.5. *Let $w \in W^{1,2}_{loc}(\Omega) \cap L^\infty(\Omega)$ be a weak solution of (3.2) and let $r > 0$ be such that $B_{4r} \subset\subset \Omega$. Then there exists a constant $\frac{7}{8} \le \gamma < 1$ depending only on n, λ, and Λ such that*

$$\operatorname*{osc}_{B_r} w \le \gamma \operatorname*{osc}_{B_{4r}} w. \tag{3.14}$$

Proof. We follow the proof of Lemma 2.107 in [13]. Let $r > 0$ be such that $B_{4r} \subset\subset \Omega$ and let $k' > k \ge k_0$, where

$$k_0 := \frac{1}{2}(M(4r) + m(4r)).$$

Assume $|A(k_0, 2r)| \le \frac{1}{2}|B_{2r}|$, where $A(k,\rho) := \{x \in B_\rho : w(x) > k\}$ as before. If not, we can write

$$\begin{aligned} |\{x \in B_{2r} : -w(x) > -k_0\}| &\le |\{x \in B_{2r} : w(x) \le k_0\}| \\ &= |B_{2r}| - |A(k_0, 2r)| < \frac{1}{2}|B_{2r}|. \end{aligned}$$

Thus, if $|A(k_0, 2r)| > \frac{1}{2}|B_{2r}|$, instead of w we may consider $-w$, which is also a weak solution of (3.2) and has the same oscillation as w.

Define

$$v := \begin{cases} k' - k, & w \ge k' \\ w - k, & k < w < k' \\ 0, & w \le k \end{cases}.$$

Using the above assumption we have

$$\begin{aligned} |\{x \in B_{2r} : v(x) = 0\}| &= |\{x \in B_{2r} : w(x) \le k\}| \\ &\ge |\{x \in B_{2r} : w(x) \le k_0\}| \\ &\ge \frac{1}{2}|B_{2r}|, \end{aligned}$$

and further

$$\begin{aligned} k' - k &= \frac{1}{|\{x \in B_{2r} : v(x) = 0\}|} \int_{\{x \in B_{2r} : v(x) = 0\}} (k' - k - v)dx \\ &\le \frac{2}{|B_{2r}|} \int_{B_{2r}} (k' - k - v)dx \\ &= 2(k' - k - (v)_{B_{2r}}). \end{aligned}$$

Now the definition of v and Hölder's inequality give

$$\begin{aligned}(k'-k)\,|A(k',2r)| &\le 2\int_{A(k',2r)}(k'-k-(v)_{B_{2r}})dx\\ &= 2\int_{A(k',2r)}(v-(v)_{B_{2r}})dx\\ &\le 2\int_{B_{2r}}|v-(v)_{B_{2r}}|dx\\ &\le 2\left(\int_{B_{2r}}|v-(v)_{B_{2r}}|^{\frac{n}{n-1}}dx\right)^{\frac{n-1}{n}}|B_{2r}|^{\frac{1}{n}}.\end{aligned}$$

Since $w\in W^{1,2}_{\mathrm{loc}}(\Omega)\subset W^{1,1}_{\mathrm{loc}}(\Omega)$, we may use Sobolev-Poincaré inequality, Lemma 2.4, in the case $p=1$. Therefore, with $|B_{2r}|^{\frac{1}{n}}=c(n)r$ we have

$$\begin{aligned}(k'-k)\,|A(k',2r)| &\le c(n)r\int_{B_{2r}}|Dv|\,dx\\ &= c(n)r\int_{A(k,2r)\setminus A(k',2r)}|Dw|\,dx\\ &\le c(n)r\left(\int_{A(k,2r)\setminus A(k',2r)}|Dw|^2\,dx\right)^{\frac{1}{2}}|A(k,2r)\setminus A(k',2r)|^{\frac{1}{2}}\\ &\le c(n)r\left(\int_{A(k,2r)}|Dw|^2\,dx\right)^{\frac{1}{2}}(|A(k,2r)|-|A(k',2r)|)^{\frac{1}{2}}.\end{aligned}\tag{3.15}$$

Here we have also used Hölder's inequality and the fact that $Dv=Dw$, when $k<w<k'$, and zero elsewhere.

Next we choose a cut-off function $\eta\in C_0^\infty(B_{4r})$ such that $0\le\eta\le 1,\eta\equiv 1$ in B_{2r}, and $|D\eta|\le\frac{1}{r}$. Lemma 3.1 then yields

$$\begin{aligned}\int_{A(k,2r)}|Dw|^2\,dx &= \int_{B_{2r}}|D(w-k)_+|^2\,dx\\ &\le \int_{B_{4r}}|D(w-k)_+|^2\,\eta^2dx\\ &\le \frac{4\Lambda^2}{\lambda^2}\int_{B_{4r}}|D\eta|^2\,(w-k)_+^2dx\\ &\le \frac{4\Lambda^2}{\lambda^2}\frac{1}{r^2}\int_{B_{4r}}(w-k)_+^2dx\\ &\le \frac{c(\lambda,\Lambda)}{r^2}\,(M(4r)-k)^2\,|B_{4r}|\\ &= c(n,\lambda,\Lambda)r^{n-2}(M(4r)-k)^2.\end{aligned}$$

Combining this with (3.15) gives

$$(k'-k)^2\,|A(k',2r)|^2\le cr^n(M(4r)-k)^2\,(|A(k,2r)|-|A(k',2r)|)\,,\tag{3.16}$$

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

Now define for $j = 0, 1, 2, \ldots$

$$k_j := M(4r) - 2^{-j-1} \operatorname*{osc}_{B_{4r}} w.$$

For $j \geq 1$ choose $k = k_{j-1}$ and $k' = k_j$. Then we have

$$k' - k = 2^{-j-1} \operatorname*{osc}_{B_{4r}} w$$

and

$$M(4r) - k = 2^{-j} \operatorname*{osc}_{B_{4r}} w,$$

and by plugging these into (3.16) we obtain

$$\left(2^{-j-1} \operatorname*{osc}_{B_{4r}} w\right)^2 |A(k_j, 2r)|^2 \leq cr^n \left(2^{-j} \operatorname*{osc}_{B_{4r}} w\right)^2 \left(|A(k_{j-1}, 2r)| - |A(k_j, 2r)|\right).$$

This implies

$$|A(k_j, 2r)|^2 \leq cr^n \left(|A(k_{j-1}, 2r)| - |A(k_j, 2r)|\right)$$

and summing over j up to an arbitrary integer $l \geq 1$ gives

$$\begin{aligned}
l\,|A(k_l, 2r)|^2 = \sum_{j=1}^{l} |A(k_l, 2r)|^2 &\leq \sum_{j=1}^{l} |A(k_j, 2r)|^2 \\
&\leq cr^n \sum_{j=1}^{l} \left(|A(k_{j-1}, 2r)| - |A(k_j, 2r)|\right) \\
&= cr^n \left(|A(k_0, 2r)| - |A(k_l, 2r)|\right) \\
&\leq cr^n\,|B_{2r}| = cr^{2n}.
\end{aligned}$$

Here the first inequality follows from the fact that $|A(k, \rho)|$ is non-increasing with respect to k. Thus, for every $l = 1, 2, \ldots$ we have the inequality

$$|A(k_l, 2r)| \leq cr^n l^{-\frac{1}{2}}.$$

Now we are ready to apply Theorem 3.4. By replacing $\tilde{k}$ with k_l we have

$$\begin{aligned}
M(r) &= \operatorname*{ess\,sup}_{B_r} w \\
&\leq k_l + c \left(\fint_{B_{2r}} (w - k_l)_+^2 dx \right)^{\frac{1}{2}} \\
&= k_l + c \left(\frac{1}{|B_{2r}|} \int_{A(k_l, 2r)} (w - k_l)^2 dx \right)^{\frac{1}{2}} \\
&\leq k_l + cr^{-\frac{n}{2}} (M(4r) - k_l)\,|A(k_l, 2r)|^{\frac{1}{2}} \\
&\leq k_l + cl^{-\frac{1}{4}} (M(4r) - k_l).
\end{aligned}$$

Then choose $l \geq 16c^4$ and recall the definition of k_j. This leads to

$$\begin{aligned} M(r) &\leq \frac{1}{2}\left(k_l + M(4r)\right) \\ &= \frac{1}{2}\left(M(4r) - 2^{-l-1} \operatorname*{osc}_{B_{4r}} w + M(4r)\right) \\ &= M(4r) - 2^{-l-2} \operatorname*{osc}_{B_{4r}} w, \end{aligned}$$

and we finally get our result by calculating

$$\begin{aligned} \operatorname*{osc}_{B_r} w &= M(r) - m(r) \\ &\leq M(r) - m(4r) \\ &\leq M(4r) - 2^{-l-2} \operatorname*{osc}_{B_{4r}} w - m(4r) \\ &= (1 - 2^{-l-2}) \operatorname*{osc}_{B_{4r}} w. \end{aligned}$$

Note that $\gamma := (1 - 2^{-l-2}) < 1$ depends only on l, which in turn depends on $c = c(n, \lambda, \Lambda)$. Thus, we have $\gamma = \gamma(n, \lambda, \Lambda)$. Moreover, $\gamma \geq \frac{7}{8}$ since $l \geq 1$. □

We may now prove the local Hölder continuity of weak solutions of (3.2) by iterating the previous result. Here and also in the following sections we denote $d_{x,y} := \min\{d_x, d_y\}$, where $d_x := \min\{1, \operatorname{dist}(x, \partial\Omega)\}$.

Theorem 3.6. *Let $w \in W^{1,2}_{loc}(\Omega) \cap L^\infty(\Omega)$ be a weak solution of (3.2). Then there exist constants $0 < \alpha < 1$ and c depending only on n, λ and Λ such that*

$$|w(x) - w(y)| \leq c\, ||w||_{L^\infty(\Omega)}\, d_{x,y}^{-\alpha}\, |x - y|^\alpha \tag{3.17}$$

for all $x, y \in \Omega$, after possibly redefining w on a set of measure zero.

Proof. Let us first generalize the previous result by considering $0 < r < R$ such that $B_R \subset\subset \Omega$. Choose a positive integer m such that

$$4^{m-1} < \frac{R}{r} \leq 4^m \tag{3.18}$$

and iterate (3.14) to obtain

$$\operatorname*{osc}_{B_r} w \leq \gamma^{m-1} \operatorname*{osc}_{B_{4^{m-1}r}} w \leq \gamma^{m-1} \operatorname*{osc}_{B_R} w.$$

From (3.18) we see that $m \geq \frac{\log \frac{R}{r}}{\log 4}$, and since $\gamma < 1$, we have

$$\gamma^{m-1} \leq \gamma^{\frac{\log \frac{R}{r}}{\log 4} - 1} = e^{\frac{\log \gamma}{\log 4}\left(\log \frac{R}{r} - \log 4\right)} = e^{-\frac{\log \gamma}{\log 4} \log \frac{4r}{R}} = \left(\frac{4r}{R}\right)^{-\frac{\log \gamma}{\log 4}}.$$

Therefore,

$$\operatorname*{osc}_{B_r} w \leq 4^\alpha \left(\frac{r}{R}\right)^\alpha \operatorname*{osc}_{B_R} w, \tag{3.19}$$

where $\alpha := -\frac{\log\gamma}{\log 4} > 0$. Moreover, due to the lower bound of γ, we have

$$\alpha = \frac{\log\frac{1}{\gamma}}{\log 4} \leq \frac{\log\frac{8}{7}}{\log 4} < 1.$$

Next take $x \in \Omega$, $y \in B(x, \frac{d_x}{4})$, and choose $r = 2\,|x-y|$ and $R = \frac{d_x}{2}$. Clearly $r < R$, so that we may apply (3.19), and since $\operatorname{osc}_\Omega w \leq 2\,||w||_{L^\infty(\Omega)}$ and $d_{x,y} \leq d_x$, we have

$$|w(x) - w(y)| \leq \operatorname*{osc}_{B(x,r)} w \leq 4^\alpha \left(\frac{r}{R}\right)^\alpha \operatorname*{osc}_{B(x,R)} w \leq 2 \cdot 16^\alpha\,||w||_{L^\infty(\Omega)}\, d_{x,y}^{-\alpha}\,|x-y|^\alpha$$

for almost every $x \in \Omega, y \in B(x, \frac{d_x}{4})$.

If $y \in \Omega \setminus B(x, \frac{d_x}{4})$, we have $d_x \leq 4\,|x-y|$, and therefore almost everywhere

$$|w(x) - w(y)| \leq 2\max\{|w(x)|\,, |w(y)|\} \leq 2 \cdot 4^\alpha\,||w||_{L^\infty(\Omega)}\, d_{x,y}^{-\alpha}\,|x-y|^\alpha\,.$$

Thus, (3.17) holds for almost every $x, y \in \Omega$ and hence Hölder continuity points are dense in Ω. Therefore, for each discontinuity point x we may choose a sequence (x_i) of Hölder continuous points such that $x_i \to x$. The sequence (x_i) is Cauchy, and since

$$|w(x_i) - w(x_j)| \leq c\,||w||_{L^\infty(\Omega)}\, d_{x_i,x_j}^{-\alpha}\,|x_i - x_j|^\alpha \to 0,$$

as $i, j \to \infty$, we see that also $(w(x_i))$ is Cauchy. Thus, we may redefine w at x such that

$$w(x) := \lim_{i\to\infty} w(x_i).$$

Now take any $x, y \in \Omega$ and corresponding Hölder continuous sequences (x_i) and (y_i). Then

$$\begin{aligned} |w(x) - w(y)| &\leq |w(x) - w(x_i)| + |w(x_i) - w(y_i)| + |w(y_i) - w(y)| \\ &\leq |w(x) - w(x_i)| + c\,||w||_{L^\infty(\Omega)}\, d_{x_i,y_i}^{-\alpha}\,|x_i - y_i|^\alpha + |w(y_i) - w(y)|\,, \end{aligned}$$

and by letting $i \to \infty$, we obtain the result. □

The following is an immediate consequence of Theorem 3.6.

Corollary 3.7. *Let $u \in W^{1,\infty}(\Omega)$ be a minimizer of the functional $\mathcal{F}$. Then $u \in C^{1,\alpha}(\Omega)$ for some $0 < \alpha < 1$.*

4 Towards higher regularity

In this section we study weak solutions of slightly more general equations, namely equations of the form

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w) = \sum_{i=1}^{n} D_i g_i(x) \tag{4.1}$$

in Ω, where $w \in W^{1,2}_{\mathrm{loc}}(\Omega)$ and the function $g : \Omega \to \mathbb{R}^n$ is assumed to be locally bounded and Hölder continuous. The coefficients b_{ij} play the same role as in the previous section and may thus be assumed to satisfy conditions (3.3) and (3.4). Moreover, since F was assumed to be smooth and $Du \in C^{0,\alpha}(\Omega)$ by Corollary 3.7, a simple application of the mean value theorem shows that $\partial_{z_i}\partial_{z_j}F(Du) \in C^{0,\alpha}(\Omega)$. Hence it is reasonable to assume that also the coefficients b_{ij} are Hölder continuous. The aim is to prove local boundedness and Hölder continuity for the gradient of w, and then in the following section use this result repeatedly in order to prove the smoothness of the solutions of our original minimization problem.

4.1 Schauder theory

To begin with, let us write the above assumptions more precisely. As in the previous section, we denote $d_x := \min\{1, \mathrm{dist}(x, \partial\Omega)\}$ and $d_{x,y} := \min\{d_x, d_y\}$. We assume there exist $0 < \alpha < 1, \beta \geq 0$, and $M \geq 1$ such that

$$|w(x)|, |g(x)| \leq M d_x^{-\beta}, \tag{4.2}$$

$$\sum_{i,j=1}^{n} |b_{ij}(x) - b_{ij}(y)| \leq M d_{x,y}^{-\alpha} |x-y|^{\alpha}, \tag{4.3}$$

and

$$|g(x) - g(y)| \leq M d_{x,y}^{-\alpha-\beta} |x-y|^{\alpha} \tag{4.4}$$

for all $x, y \in \Omega$.

Fix a point $x_0 \in \Omega$ and denote $d := \frac{1}{4}d_{x_0}$, so that $B(x_0, 4d) \subset \Omega$ and for all $x, y \in B(x_0, 2d)$ we have

$$|w(x)|, |g(x)| \leq M d^{-\beta}, \tag{4.5}$$

$$\sum_{i,j=1}^{n} |b_{ij}(x) - b_{ij}(y)| \leq M d^{-\alpha} |x-y|^{\alpha}, \tag{4.6}$$

and

$$|g(x) - g(y)| \leq M d^{-\alpha-\beta} |x-y|^{\alpha}. \tag{4.7}$$

Before proving the main result of the section, we need a few lemmas. First of them is another Caccioppoli estimate. Again, we use the shorthand notation $B_\rho := B(x_0, \rho)$.

Lemma 4.1. *Let $w \in W^{1,2}_{loc}(\Omega)$ be a weak solution of (4.1). Then there exists a constant $c = c(n, \lambda, \Lambda, M)$ such that*

$$\left(\fint_{B_d} |Dw|^2 \, dx \right)^{\frac{1}{2}} \leq c d^{-1-\beta}.$$

Proof. Since w is a weak solution of (4.1), we have

$$\int_\Omega \sum_{i,j=1}^n b_{ij} D_j w D_i v dx = \int_\Omega \sum_{i=1}^n g_i D_i v dx$$

for every $v \in W^{1,2}_0(\Omega)$. Take a cut-off function $\eta \in C^\infty_0(B_{2d})$ with $0 \leq \eta \leq 1, \eta \equiv 1$ in B_d, and $|D\eta| \leq \frac{2}{d}$, and choose $v = \eta^2 w \in W^{1,2}_0(\Omega)$ as the test function.

Now using assumptions (3.3) and (3.4) and Cauchy-Schwarz inequality yields

$$\begin{aligned}
\lambda \int_\Omega |Dw|^2 \eta^2 dx &\leq \int_\Omega \sum_{i,j=1}^n b_{ij} D_j w D_i w \eta^2 dx \\
&= \int_\Omega \sum_{i=1}^n g_i D_i w \eta^2 dx + 2 \int_\Omega \sum_{i=1}^n g_i \eta D_i \eta w dx - 2 \int_\Omega \sum_{i,j=1}^n b_{ij} D_j w \eta D_i \eta w dx \\
&\leq \int_\Omega |g| \, |Dw| \, \eta^2 dx + 2 \int_\Omega |g| \, \eta \, |D\eta| \, |w| \, dx + 2\Lambda \int_\Omega |Dw| \, \eta \, |D\eta| \, |w| \, dx.
\end{aligned}$$

Next we apply Young's inequality to the middle term and Young's inequality with ε to the others, which gives

$$\begin{aligned}
&\lambda \int_\Omega |Dw|^2 \eta^2 dx \\
&\quad \leq \frac{\varepsilon_1}{2} \int_\Omega |Dw|^2 \eta^2 dx + \frac{1}{2\varepsilon_1} \int_\Omega |g|^2 \eta^2 dx + \int_\Omega |g|^2 \eta^2 dx \\
&\quad + \int_\Omega |D\eta|^2 |w|^2 dx + \varepsilon_2 \Lambda \int_\Omega |Dw|^2 \eta^2 dx + \frac{\Lambda}{\varepsilon_2} \int_\Omega |D\eta|^2 |w|^2 dx \\
&\quad = \frac{\lambda}{2} \int_\Omega |Dw|^2 \eta^2 dx + \left(\frac{1}{\lambda} + 1 \right) \int_\Omega |g|^2 \eta^2 dx + \left(\frac{4\Lambda^2}{\lambda} + 1 \right) \int_\Omega |D\eta|^2 |w|^2 dx,
\end{aligned}$$

where we have chosen $\varepsilon_1 = \frac{\lambda}{2}$ and $\varepsilon_2 = \frac{\lambda}{4\Lambda}$.

Using the properties of the cut-off function and assumption (4.5) it now follows that

$$\begin{aligned}
\int_{B_d} |Dw|^2 dx &\leq \int_\Omega |Dw|^2 \eta^2 dx \\
&\leq \frac{2(\lambda+1)}{\lambda^2} \int_\Omega |g|^2 \eta^2 dx + \frac{2(4\Lambda^2+\lambda)}{\lambda^2} \int_\Omega |D\eta|^2 |w|^2 dx \\
&\leq \frac{2(\lambda+1)}{\lambda^2} \int_{B_{2d}} |g|^2 dx + \frac{8(4\Lambda^2+\lambda)}{\lambda^2} d^{-2} \int_{B_{2d}} |w|^2 dx \\
&\leq \frac{2(16\Lambda^2+5\lambda+1)}{\lambda^2} M^2 d^{-2-2\beta} 2^n |B_d| .
\end{aligned}$$

Here we have used the fact that $d \leq 1$ or $d^{-2} \geq 1$. Dividing by $|B_d|$ and taking square roots completes the proof. $\square$

Next we apply the so-called freezing technique to equation (4.1). In order to get the desired result for the gradient of w, first we have to put quite a lot of effort into deriving estimates for the corresponding equation with constant coefficients.

Thus, consider equation (4.1) in the ball $B_\rho, 0 < \rho \leq d$, with the functions b_{ij} and g fixed at the point x_0. Let $w_\rho \in W^{1,2}(B_\rho)$ be a weak solution of this equation and, moreover, assume that $w_\rho = w$ on the boundary of B_ρ. That is, w_ρ satisfies the equation

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x_0)D_j w_\rho) = \sum_{i=1}^{n} D_i g_i(x_0) \tag{4.8}$$

in B_ρ such that $w - w_\rho \in W_0^{1,2}(B_\rho)$. The right hand side of (4.8) is clearly zero, but for the next lemma we shall keep it in the above form.

Lemma 4.2. *Let $0 < \rho \leq d$ and let $w \in W^{1,2}_{loc}(\Omega)$ and $w_\rho \in W^{1,2}(B_\rho)$ be weak solutions of* (4.1) *and* (4.8)*, respectively. Then there exists a constant $c = c(\lambda, M)$ such that*

$$\left(\fint_{B_\rho} |D(w - w_\rho)|^2 \, dx \right)^{\frac{1}{2}} \leq c d^{-\alpha} \rho^\alpha \left(\left(\fint_{B_\rho} |Dw|^2 \, dx \right)^{\frac{1}{2}} + d^{-\beta} \right).$$

Proof. Since w_ρ is a weak solution of (4.8), we have

$$\int_{B_\rho} \sum_{i,j=1}^{n} b_{ij}(x_0) D_j w_\rho D_i v dx = \int_{B_\rho} \sum_{i=1}^{n} g_i(x_0) D_i v dx \tag{4.9}$$

for every $v \in W_0^{1,2}(B_\rho)$. By only considering test functions whose support is in B_ρ, we get a similar equation for w, that is,

$$\int_{B_\rho} \sum_{i,j=1}^{n} b_{ij}(x) D_j w D_i v dx = \int_{B_\rho} \sum_{i=1}^{n} g_i(x) D_i v dx \tag{4.10}$$

for every $v \in W_0^{1,2}(B_\rho)$.

Subtract (4.9) from (4.10) to obtain

$$\int_{B_\rho} \sum_{i,j=1}^{n} \left(b_{ij}(x) D_j w - b_{ij}(x_0) D_j w_\rho \right) D_i v dx = \int_{B_\rho} \sum_{i=1}^{n} (g_i(x) - g_i(x_0)) D_i v dx,$$

after which adding and subtracting the term $b_{ij}(x_0) D_j w$ leads to

$$\begin{aligned} &\int_{B_\rho} \sum_{i,j=1}^{n} b_{ij}(x_0) D_j (w - w_\rho) D_i v dx \\ &\quad = \int_{B_\rho} \sum_{i,j=1}^{n} (b_{ij}(x_0) - b_{ij}(x)) D_j w D_i v dx + \int_{B_\rho} \sum_{i=1}^{n} (g_i(x) - g_i(x_0)) D_i v dx \end{aligned}$$

for every $v \in W_0^{1,2}(B_\rho)$. Now, by choosing $v = w - w_\rho \in W_0^{1,2}(B_\rho)$ and using the ellipticity condition (3.3) and Cauchy-Schwarz inequality, we get

$$\begin{aligned}
\lambda \int_{B_\rho} |D(w - w_\rho)|^2 \, dx &\leq \int_{B_\rho} \sum_{i,j=1}^{n} b_{ij}(x_0) D_j(w - w_\rho) D_i(w - w_\rho) dx \\
&= \int_{B_\rho} \sum_{i,j=1}^{n} (b_{ij}(x_0) - b_{ij}(x)) D_j w D_i(w - w_\rho) dx \\
&\quad + \int_{B_\rho} \sum_{i=1}^{n} (g_i(x) - g_i(x_0)) D_i(w - w_\rho) dx \\
&\leq \int_{B_\rho} \sum_{i,j=1}^{n} |b_{ij}(x_0) - b_{ij}(x)| \, |Dw| \, |D(w - w_\rho)| \, dx \\
&\quad + \int_{B_\rho} |g(x) - g(x_0)| \, |D(w - w_\rho)| \, dx.
\end{aligned}$$

Once again we apply Young's inequality with ε, this time with $\varepsilon = \frac{\lambda}{2}$ for both terms, and arrive at

$$\begin{aligned}
\lambda \int_{B_\rho} |D(w - w_\rho)|^2 \, dx &\leq \frac{\lambda}{2} \int_{B_\rho} |D(w - w_\rho)|^2 \, dx \\
&\quad + \frac{1}{\lambda} \int_{B_\rho} \left(\sum_{i,j=1}^{n} |b_{ij}(x_0) - b_{ij}(x)| \right)^2 |Dw|^2 \, dx + \frac{1}{\lambda} \int_{B_\rho} |g(x) - g(x_0)|^2 \, dx.
\end{aligned}$$

Next we use assumptions (4.6) and (4.7) and notice that $|x - x_0| < \rho$ for all $x \in B_\rho$. Thus

$$\int_{B_\rho} |D(w - w_\rho)|^2 \, dx \leq \frac{2M^2}{\lambda^2} \rho^{2\alpha} \left(d^{-2\alpha} \int_{B_\rho} |Dw|^2 \, dx + d^{-2\alpha - 2\beta} \, |B_\rho| \right),$$

which implies

$$\left(\fint_{B_\rho} |D(w - w_\rho)|^2 \, dx \right)^{\frac{1}{2}} \leq c(\lambda, M) d^{-\alpha} \rho^{\alpha} \left(\left(\fint_{B_\rho} |Dw|^2 \, dx \right)^{\frac{1}{2}} + d^{-\beta} \right). \qquad \square$$

The next simple lemma proves to be useful in what follows.

Lemma 4.3. *Let $A \subset B$, $|A| > 0$, and $|B| < \infty$. Then a function $f \in L^2(B)$ satisfies*

$$\left(\fint_A |f - (f)_A|^2 \, dx \right)^{\frac{1}{2}} \leq \left(\frac{|B|}{|A|} \right)^{\frac{1}{2}} \left(\fint_B |f - (f)_B|^2 \, dx \right)^{\frac{1}{2}}.$$

Proof. Using the Hilbert space structure of L^2 we may define an inner product

$$< f, g > := \fint_A fg dx$$

for all $f, g \in L^2(A)$. A simple calculation shows that $(f)_A$ minimizes the function $h(a) :=< f - a, f - a >$, and therefore

$$\left(\fint_A |f - (f)_A|^2 \, dx\right)^{\frac{1}{2}} \leq \left(\fint_A |f - (f)_B|^2 \, dx\right)^{\frac{1}{2}} \leq \left(\frac{|B|}{|A|}\right)^{\frac{1}{2}} \left(\fint_B |f - (f)_B|^2 \, dx\right)^{\frac{1}{2}}.$$

□

In the following lemma we exploit the fact that the coefficients in (4.8) are constants. This enables us to derive a useful estimate for w_ρ by changing variables and using well-known properties of harmonic functions. The proof uses some ideas from [9, p. 87–88].

Lemma 4.4. *Let $0 < \rho \leq d$ and let $w_\rho \in W^{1,2}(B_\rho)$ be a weak solution of (4.8). Then there exists a constant $c = c(n, \lambda, \Lambda)$ such that*

$$\left(\fint_{B_{\delta\rho}} \left|Dw_\rho - (Dw_\rho)_{B_{\delta\rho}}\right|^2 dx\right)^{\frac{1}{2}} \leq c\delta \left(\fint_{B_\rho} \left|Dw_\rho - (Dw_\rho)_{B_\rho}\right|^2 dx\right)^{\frac{1}{2}}$$

for every $0 < \delta < 1$.

Proof. Since $b_{ij}(x_0)$ and $g(x_0)$ are constants, equation (4.8) can be written as

$$\sum_{i,j=1}^{n} b_{ij}(x_0) D_i D_j w_\rho(x) = 0. \tag{4.11}$$

Let B be the $n \times n$ matrix formed by the coefficients $b_{ij}(x_0)$ scaled by Λ, that is,

$$(B)_{ij} := \frac{b_{ij}(x_0)}{\Lambda}$$

for $1 \leq i, j \leq n$. Defined this way the eigenvalues of B are on the interval $\left[\frac{\lambda}{\Lambda}, 1\right]$. Clearly B is symmetric and positive definite and hence possesses an inverse matrix B^{-1}, which is also symmetric and positive definite. Thus, Cholesky decomposition gives a unique positive definite upper triangular matrix H such that $H^T H = B^{-1}$ or equivalently $HBH^T = I$. Denoting the $(i, j)^{\text{th}}$ element of H by h_{ij} this can be written as

$$\sum_{i,j=1}^{n} b_{ij}(x_0) h_{kj} h_{mi} = \delta_{mk} \Lambda \tag{4.12}$$

for every $1 \leq k, m \leq n$.

Next we change variables to $y = H(x - x_0)$ and define $w_\rho(x) =: \overline{w}_\rho(y)$. Using the chain rule we calculate

$$D_j w_\rho(x) = D_j \overline{w}_\rho(y) = \sum_{k=1}^{n} D_j y_k D_k^y \overline{w}_\rho(y) = \sum_{k=1}^{n} h_{kj} D_k^y \overline{w}_\rho(y), \tag{4.13}$$

where D^y denotes the weak gradient with respect to y. Differentiating once more yields

$$D_i D_j w_\rho(x) = \sum_{k=1}^{n} h_{kj} D_k^y D_i \overline{w}_\rho(y) = \sum_{k,m=1}^{n} h_{kj} h_{mi} D_k^y D_m^y \overline{w}_\rho(y),$$

so that using identity (4.12) equation (4.11) becomes

$$\begin{aligned}
0 &= \sum_{i,j=1}^{n} b_{ij}(x_0) D_i D_j w_\rho(x) = \sum_{i,j=1}^{n} b_{ij}(x_0) \sum_{k,m=1}^{n} h_{kj} h_{mi} D_k^y D_m^y \overline{w}_\rho(y) \\
&= \sum_{k,m=1}^{n} D_k^y D_m^y \overline{w}_\rho(y) \sum_{i,j=1}^{n} b_{ij}(x_0) h_{kj} h_{mi} = \sum_{k,m=1}^{n} D_k^y D_m^y \overline{w}_\rho(y) \delta_{mk} \Lambda \\
&= \Lambda \sum_{k=1}^{n} D_k^y D_k^y \overline{w}_\rho(y).
\end{aligned}$$

Therefore, $\overline{w}_\rho$ is a harmonic function in the new coordinates.

Let us next consider how the change of variables affects the domain. Assume $y \in B(0, \tilde{\rho})$ for some $0 < \tilde{\rho} \leq \rho$. Then x belongs to the set

$$\begin{aligned}
E(x_0, \tilde{\rho}) &:= x_0 + H^{-1} B(0, \tilde{\rho}) \\
&= x_0 + \{H^{-1} z : z^T z < \tilde{\rho}^2\} \\
&= x_0 + \{z : z^T H^T H z < \tilde{\rho}^2\} \\
&= \{x : (x - x_0)^T (\tilde{\rho}^2 B)^{-1} (x - x_0) < 1\},
\end{aligned}$$

which, in fact, is an ellipsoid due to the fact that B is positive definite. The center of the ellipsoid $E(x_0, \tilde{\rho})$ is x_0 and the lengths of its semi-axes are given by the square roots of the eigenvalues of the matrix $\tilde{\rho}^2 B$. Therefore, we have

$$B(x_0, \theta\tilde{\rho}) \subset E(x_0, \tilde{\rho}) \subset B(x_0, \tilde{\rho}), \tag{4.14}$$

where $\theta := \left(\frac{\lambda}{\Lambda}\right)^{\frac{1}{2}} \leq 1$. Similarly, when $x \in B(x_0, \rho)$, we have

$$y \in \tilde{E}(0, \rho) := HB(0, \rho) = \{y : y^T (\rho^2 H H^T)^{-1} y < 1\}.$$

Thus, the domain of $\overline{w}_\rho$ is the ellipsoid $\tilde{E}(0, \rho)$, whose semiaxes have lengths bounded from below by ρ and from above by $\theta^{-1}\rho$.

Let us assume $0 < \delta < \frac{\theta}{4}$. For if $\frac{\theta}{4} \leq \delta < 1$, the result follows by using Lemma 4.3 and calculating

$$\begin{aligned}
\left(\fint_{B_{\delta\rho}} \left| Dw_\rho - (Dw_\rho)_{B_{\delta\rho}} \right|^2 dx \right)^{\frac{1}{2}} &\leq \left(\frac{|B_\rho|}{|B_{\delta\rho}|} \right)^{\frac{1}{2}} \left(\fint_{B_\rho} \left| Dw_\rho - (Dw_\rho)_{B_\rho} \right|^2 dx \right)^{\frac{1}{2}} \\
&= \delta^{-\frac{n}{2}} \left(\fint_{B_\rho} \left| Dw_\rho - (Dw_\rho)_{B_\rho} \right|^2 dx \right)^{\frac{1}{2}} \\
&\leq \left(\frac{\theta}{4} \right)^{-\frac{n}{2}-1} \delta \left(\fint_{B_\rho} \left| Dw_\rho - (Dw_\rho)_{B_\rho} \right|^2 dx \right)^{\frac{1}{2}}.
\end{aligned}$$

By choosing $\tilde{\rho} = \theta^{-1}\delta\rho < \rho$ in (4.14) we get

$$B(x_0, \delta\rho) \subset E(x_0, \theta^{-1}\delta\rho) \subset B(x_0, \theta^{-1}\delta\rho).$$

Thus by Lemma 4.3, identity (4.13), and Minkowski's inequality we obtain

$$
\begin{aligned}
&\left(\fint_{B(x_0,\delta\rho)} \left|Dw_\rho(x) - (Dw_\rho)_{B(x_0,\delta\rho)}\right|^2 dx\right)^{\frac{1}{2}} \\
&\leq \left(\frac{|E(x_0,\theta^{-1}\delta\rho)|}{|B(x_0,\delta\rho)|}\right)^{\frac{1}{2}} \left(\fint_{E(x_0,\theta^{-1}\delta\rho)} \left|Dw_\rho(x) - (Dw_\rho)_{E(x_0,\theta^{-1}\delta\rho)}\right|^2 dx\right)^{\frac{1}{2}} \\
&\leq \left(\frac{|B(x_0,\theta^{-1}\delta\rho)|}{|B(x_0,\delta\rho)|}\right)^{\frac{1}{2}} \sum_{j=1}^{n} \left(\fint_{E(x_0,\theta^{-1}\delta\rho)} \left|D_j w_\rho(x) - (D_j w_\rho)_{E(x_0,\theta^{-1}\delta\rho)}\right|^2 dx\right)^{\frac{1}{2}} \\
&= \theta^{-\frac{n}{2}} \sum_{j=1}^{n} \left(\fint_{B(0,\theta^{-1}\delta\rho)} \left|\sum_{k=1}^{n} h_{kj} \left(D_k^y \overline{w}_\rho(y) - (D_k^y \overline{w}_\rho)_{B(0,\theta^{-1}\delta\rho)}\right)\right|^2 dy\right)^{\frac{1}{2}} \\
&\leq \theta^{-\frac{n}{2}} \sum_{j,k=1}^{n} |h_{kj}| \left(\fint_{B(0,\theta^{-1}\delta\rho)} \left|D_k^y \overline{w}_\rho(y) - (D_k^y \overline{w}_\rho)_{B(0,\theta^{-1}\delta\rho)}\right|^2 dy\right)^{\frac{1}{2}}.
\end{aligned}
\tag{4.15}
$$

To illustrate how the change of variables is done above we calculate

$$
\begin{aligned}
(D_j w_\rho)_{E(x_0,\theta^{-1}\delta\rho)} &= \frac{1}{\int_{E(x_0,\theta^{-1}\delta\rho)} dx} \int_{E(x_0,\theta^{-1}\delta\rho)} D_j w_\rho(x) dx \\
&= \frac{1}{\int_{B(0,\theta^{-1}\delta\rho)} \det(H^{-1}) dy} \int_{B(0,\theta^{-1}\delta\rho)} D_j \overline{w}_\rho(y) \det(H^{-1}) dy \\
&= \fint_{B(0,\theta^{-1}\delta\rho)} \sum_{k=1}^{n} h_{kj} D_k^y \overline{w}_\rho(y) dy \\
&= \sum_{k=1}^{n} h_{kj} (D_k^y \overline{w}_\rho)_{B(0,\theta^{-1}\delta\rho)},
\end{aligned}
$$

where $\det(H^{-1}) > 0$ is the determinant of the positive definite matrix H^{-1}. A very similar calculation to (4.15) shows that

$$
\begin{aligned}
&\left(\fint_{B(0,\rho)} \left|D_k^y \overline{w}_\rho(y) - (D_k^y \overline{w}_\rho)_{B(0,\rho)}\right|^2 dy\right)^{\frac{1}{2}} \\
&\qquad \leq \theta^{-\frac{n}{2}} \sum_{i=1}^{n} \left|h^{ik}\right| \left(\fint_{B(x_0,\rho)} \left|Dw_\rho(x) - (Dw_\rho)_{B(x_0,\rho)}\right|^2 dx\right)^{\frac{1}{2}},
\end{aligned}
\tag{4.16}
$$

where h^{ik} denotes the $(i,k)^{\text{th}}$ element of H^{-1}.

Let us now define for $k = 1, \ldots, n$

$$
\psi_k(y) := D_k^y \overline{w}_\rho(y) - (D_k^y \overline{w}_\rho)_{B(0,\rho)}.
$$

Clearly ψ_k is harmonic in $\tilde{E}(0,\rho)$ and, moreover,

$$\begin{aligned}\left|D_k^y\overline{w}_\rho(y) - (D_k^y\overline{w}_\rho)_{B(0,\theta^{-1}\delta\rho)}\right| &\leq \fint_{B(0,\theta^{-1}\delta\rho)} |D_k^y\overline{w}_\rho(y) - D_k^y\overline{w}_\rho(z)|\,dz \\ &= \fint_{B(0,\theta^{-1}\delta\rho)} |\psi_k(y) - \psi_k(z)|\,dz.\end{aligned} \tag{4.17}$$

Harmonic functions can be characterized by the mean value principle, that is,

$$\psi_k(z) = \fint_{B(z,r)} \psi_k(y)dy$$

whenever $B(z,r) \subset \tilde{E}(0,\rho)$. Thus, with $z_1, z_2 \in B(0,\theta^{-1}\delta\rho)$ and $r = \frac{\rho}{4}$ we may write

$$\begin{aligned}&|\psi_k(z_1) - \psi_k(z_2)| = \left|\fint_{B(z_1,r)} \psi_k(y)dy - \fint_{B(z_2,r)} \psi_k(y)dy\right| \\ &= \left|\frac{1}{|B_r|}\int_{B(z_1,r)\cup B(z_2,r)} \psi_k(y)\chi_{B(z_1,r)}dy - \frac{1}{|B_r|}\int_{B(z_1,r)\cup B(z_2,r)} \psi_k(y)\chi_{B(z_2,r)}dy\right| \\ &\leq \frac{1}{|B_r|}\int_{B(z_1,r)\cup B(z_2,r)} |\psi_k(y)|\left|\chi_{B(z_1,r)} - \chi_{B(z_2,r)}\right|dy \\ &\leq \frac{1}{|B_r|}\operatorname*{ess\,sup}_{B(z_1,r)\cup B(z_2,r)} |\psi_k| \int_{B(z_1,r)\cup B(z_2,r)} \left|\chi_{B(z_1,r)} - \chi_{B(z_2,r)}\right|dy.\end{aligned}$$

The remaining integral is, in fact, the measure of the symmetric difference of the two balls. This can be estimated from above by

$$|\partial B_r|\,|z_1 - z_2| = n\,|B_r|\,\frac{|z_1 - z_2|}{r},$$

where $|\partial B_r|$ is the $(n-1)$-dimensional Lebesgue measure of the boundary of B_r. Moreover, we have $B(z_1,r) \cup B(z_2,r) \subset B(0,2r)$, since $r = \frac{\rho}{4} > \theta^{-1}\delta\rho$, so that

$$\begin{aligned}\operatorname*{ess\,sup}_{B(z_1,r)\cup B(z_2,r)} |\psi_k| &\leq \operatorname*{ess\,sup}_{B(0,2r)} |\psi_k| \\ &\leq \operatorname*{ess\,sup}_{z\in B(0,2r)} \fint_{B(z,2r)} |\psi_k(y)|\,dy \\ &\leq \operatorname*{ess\,sup}_{z\in B(0,2r)} \frac{1}{|B_{2r}|}\int_{B(0,4r)} |\psi_k(y)|\,dy \\ &= 2^n \fint_{B(0,\rho)} |\psi_k(y)|\,dy.\end{aligned}$$

Combining these estimates and the fact that $|z_1 - z_2| < 2\theta^{-1}\delta\rho$ gives

$$|\psi_k(z_1) - \psi_k(z_2)| \leq 2^{n+3}n\theta^{-1}\delta \fint_{B(0,\rho)} |\psi_k(y)|\,dy \tag{4.18}$$

for all $z_1, z_2 \in B(0,\theta^{-1}\delta\rho)$.

Now estimates (4.15)–(4.18) yield

$$
\begin{aligned}
&\left(\fint_{B(x_0,\delta\rho)} \left|Dw_\rho(x) - (Dw_\rho)_{B(x_0,\delta\rho)}\right|^2 dx\right)^{\frac{1}{2}} \\
&\le \theta^{-\frac{n}{2}} \sum_{j,k=1}^{n} |h_{kj}| \left(\fint_{B(0,\theta^{-1}\delta\rho)} \left(\fint_{B(0,\theta^{-1}\delta\rho)} |\psi_k(y) - \psi_k(z)|\, dz\right)^2 dy\right)^{\frac{1}{2}} \\
&\le 2^{n+3} n \theta^{-\frac{n}{2}-1} \sum_{j,k=1}^{n} |h_{kj}|\, \delta \fint_{B(0,\rho)} |\psi_k(y)|\, dy \\
&\le 2^{n+3} n \theta^{-\frac{n}{2}-1} \sum_{j,k=1}^{n} |h_{kj}|\, \delta \left(\fint_{B(0,\rho)} \left|D_k^y \overline{w}_\rho(y) - (D_k^y \overline{w}_\rho)_{B(0,\rho)}\right|^2 dy\right)^{\frac{1}{2}} \\
&\le 2^{n+3} n \theta^{-n-1} \sum_{i,j,k=1}^{n} \left|h^{ik} h_{kj}\right| \delta \left(\fint_{B(x_0,\rho)} \left|Dw_\rho(x) - (Dw_\rho)_{B(x_0,\rho)}\right|^2 dx\right)^{\frac{1}{2}}.
\end{aligned}
$$

Denote the trace of an $n \times n$ matrix A by

$$
\operatorname{tr}(A) := \sum_{i=1}^{n} (A)_{ii} = \sum_{i=1}^{n} \lambda_i(A),
$$

where $\lambda_i(A)$ is the ith eigenvalue of A. By Cauchy-Schwartz inequality and the bounds of the eigenvalues of the matrix B we have

$$
\begin{aligned}
\sum_{i,j,k=1}^{n} \left|h^{ik} h_{kj}\right| &\le \left(\sum_{i,j,k=1}^{n} \left|h^{ik}\right|^2\right)^{\frac{1}{2}} \left(\sum_{i,j,k=1}^{n} |h_{kj}|^2\right)^{\frac{1}{2}} \\
&= n \left(\sum_{i,k=1}^{n} \left|h^{ik}\right|^2 \sum_{j,k=1}^{n} |h_{kj}|^2\right)^{\frac{1}{2}} \\
&= n \left(\operatorname{tr}\left(H^{-1}\left(H^{-1}\right)^T\right) \operatorname{tr}\left(H^T H\right)\right)^{\frac{1}{2}} \\
&= n \left(\operatorname{tr}(B) \operatorname{tr}\left(B^{-1}\right)\right)^{\frac{1}{2}} \\
&= n \left(\sum_{i=1}^{n} \lambda_i(B) \sum_{j=1}^{n} \lambda_j\left(B^{-1}\right)\right)^{\frac{1}{2}} \\
&\le n^2 \left(\frac{\Lambda}{\lambda}\right)^{\frac{1}{2}}.
\end{aligned}
$$

This completes the proof. $\square$

4.2 Campanato estimates

Let us now combine the previous lemmas in order to show that the gradient of w belongs to a certain Campanato space. Campanato spaces were first introduced by Sergio Campanato in [3] and [4], and have since proven to be a useful tool in showing Hölder continuity for solutions of partial differential equations.

Lemma 4.5. *Let $w \in W^{1,2}_{loc}(\Omega)$ be a weak solution of (4.1) and denote $\bar{\beta} := 1+\alpha+\beta$. Then there exists a constant $c = c(n, \lambda, \Lambda, \alpha, M)$ such that*

$$\left(\fint_{B_\rho} \left| Dw - (Dw)_{B_\rho} \right|^2 dx \right)^{\frac{1}{2}} \leq c d^{-\bar{\beta}} \rho^\alpha$$

for every $0 < \rho \leq d$.

Proof. For $j = 0, 1, 2, \ldots$ define $\rho_j := \delta^j r$ and $B_j := B(x_0, \rho_j)$, where $0 < \delta < 1$ and $0 < r \leq d$ are constants to be fixed in the course of the proof. To somewhat shorten the notation we denote

$$E_j := \left(\fint_{B_j} \left| Dw - (Dw)_{B_j} \right|^2 dx \right)^{\frac{1}{2}}.$$

Let $w_{\rho_j} \in W^{1,2}(B_j)$ be a weak solution of (4.8) such that $w - w_{\rho_j} \in W^{1,2}_0(B_j)$. Since

$$\begin{aligned} \left| (Dw)_{B_{j+1}} - (Dw_{\rho_j})_{B_{j+1}} \right| &\leq \fint_{B_{j+1}} \left| Dw - Dw_{\rho_j} \right| dx \\ &\leq \left(\fint_{B_{j+1}} \left| Dw - Dw_{\rho_j} \right|^2 dx \right)^{\frac{1}{2}} \\ &= \delta^{-\frac{n}{2}} \left(\fint_{B_j} \left| Dw - Dw_{\rho_j} \right|^2 dx \right)^{\frac{1}{2}}, \end{aligned}$$

we get by Minkowski's inequality

$$\begin{aligned} E_{j+1} &= \left(\fint_{B_{j+1}} \left| Dw - (Dw)_{B_{j+1}} \right|^2 dx \right)^{\frac{1}{2}} \\ &\leq \left(\fint_{B_{j+1}} \left| Dw - Dw_{\rho_j} \right|^2 dx \right)^{\frac{1}{2}} + \left(\fint_{B_{j+1}} \left| Dw_{\rho_j} - (Dw_{\rho_j})_{B_{j+1}} \right|^2 dx \right)^{\frac{1}{2}} \\ &\quad + \left(\fint_{B_{j+1}} \left| (Dw_{\rho_j})_{B_{j+1}} - (Dw)_{B_{j+1}} \right|^2 dx \right)^{\frac{1}{2}} \\ &\leq \left(\fint_{B_{j+1}} \left| Dw_{\rho_j} - (Dw_{\rho_j})_{B_{j+1}} \right|^2 dx \right)^{\frac{1}{2}} + 2\delta^{-\frac{n}{2}} \left(\fint_{B_j} \left| Dw - Dw_{\rho_j} \right|^2 dx \right)^{\frac{1}{2}}. \end{aligned} \tag{4.19}$$

The first term can be further estimated by first applying Lemma 4.4 and then Minkowski's inequality again. Thus we obtain

$$\begin{aligned}
&\left(\fint_{B_{j+1}} \left|Dw_{\rho_j} - (Dw_{\rho_j})_{B_{j+1}}\right|^2 dx\right)^{\frac{1}{2}} \le c\delta \left(\fint_{B_j} \left|Dw_{\rho_j} - (Dw_{\rho_j})_{B_j}\right|^2 dx\right)^{\frac{1}{2}} \\
&\le c\delta \left[\left(\fint_{B_j} \left|Dw_{\rho_j} - Dw\right|^2 dx\right)^{\frac{1}{2}} + \left(\fint_{B_j} \left|Dw - (Dw)_{B_j}\right|^2 dx\right)^{\frac{1}{2}}\right. \\
&\qquad \left. + \left(\fint_{B_j} \left|(Dw)_{B_j} - (Dw_{\rho_j})_{B_j}\right|^2 dx\right)^{\frac{1}{2}}\right] \\
&\le c\delta \left[E_j + 2\left(\fint_{B_j} \left|Dw_{\rho_j} - Dw\right|^2 dx\right)^{\frac{1}{2}}\right].
\end{aligned}$$

Combining this with (4.19) and then using Lemma 4.2 yields

$$\begin{aligned}
E_{j+1} &\le c\delta E_j + \left(2c\delta + 2\delta^{-\frac{n}{2}}\right)\left(\fint_{B_j} \left|Dw - Dw_{\rho_j}\right|^2 dx\right)^{\frac{1}{2}} \\
&\le c\delta E_j + \left(2c\delta + 2\delta^{-\frac{n}{2}}\right) cd^{-\alpha}\rho_j^{\alpha}\left(\left(\fint_{B_j} |Dw|^2 dx\right)^{\frac{1}{2}} + d^{-\beta}\right) \\
&\le c\delta E_j + c\delta\left(1 + \delta^{-\frac{n}{2}-1}\right) d^{-\alpha}\rho_j^{\alpha}\left(E_j + \left|(Dw)_{B_j}\right| + d^{-\beta}\right) \\
&\le c\delta\left(1 + \left(1 + \delta^{-\frac{n}{2}-1}\right) d^{-\alpha} r^{\alpha}\right) E_j \\
&\qquad + c\delta\left(1 + \delta^{-\frac{n}{2}-1}\right) d^{-\alpha}\rho_j^{\alpha}\left(\left|(Dw)_{B_j}\right| + d^{-\beta}\right).
\end{aligned}$$

Here we used the fact that $\rho_j \le r$ for all $j = 0, 1, 2, \ldots$ and

$$\begin{aligned}
\left(\fint_{B_j} |Dw|^2 dx\right)^{\frac{1}{2}} &\le \left(\fint_{B_j} \left|Dw - (Dw)_{B_j}\right|^2 dx\right)^{\frac{1}{2}} + \left(\fint_{B_j} \left|(Dw)_{B_j}\right|^2 dx\right)^{\frac{1}{2}} \\
&= E_j + \left|(Dw)_{B_j}\right|.
\end{aligned}$$

Choose δ to be the largest number on the interval $\left(0, \frac{1}{4}\right]$ such that $c\delta \le \frac{1}{2}\delta^{\frac{1+\alpha}{2}}$, in other words $\delta = \min\left\{\frac{1}{4}, (2c)^{-\frac{2}{1-\alpha}}\right\}$. If we further assume r to satisfy

$$\left(1 + \delta^{-\frac{n}{2}-1}\right) d^{-\alpha} r^{\alpha} \le 1,$$

or $r \le \left(1 + \delta^{-\frac{n}{2}-1}\right)^{-\frac{1}{\alpha}} d$, we have

$$E_{j+1} \le \delta^{\frac{1+\alpha}{2}} E_j + cd^{-\alpha}\rho_j^{\alpha}\left(\left|(Dw)_{B_j}\right| + d^{-\beta}\right). \tag{4.20}$$

Here the constant c depends on δ, which in turn depends on $n, \lambda, \Lambda, \alpha$ and M, so we have $c = c(n, \lambda, \Lambda, \alpha, M)$.

Next we show that $\left|(Dw)_{B_j}\right|$ is uniformly bounded with an estimate

$$\sup_{j=0,1,2,\dots} \left|(Dw)_{B_j}\right| \le c \left(\fint_{B_0} |Dw|^2 \, dx \right)^{\frac{1}{2}} + d^{-\beta}. \tag{4.21}$$

For a given integer k, summing up (4.20) yields

$$\sum_{j=0}^{k} E_j = E_0 + \sum_{j=0}^{k-1} E_{j+1} \le E_0 + \delta^{\frac{1+\alpha}{2}} \sum_{j=0}^{k-1} E_j + cd^{-\alpha} \sum_{j=0}^{k-1} \rho_j^{\alpha} \left(\left|(Dw)_{B_j}\right| + d^{-\beta} \right).$$

Since $\delta \le \frac{1}{4}$ and E_j is positive for all j, we have

$$\delta^{\frac{1+\alpha}{2}} \sum_{j=0}^{k-1} E_j \le \frac{1}{2} \sum_{j=0}^{k} E_j,$$

and therefore

$$\sum_{j=0}^{k} E_j \le 2E_0 + 2cd^{-\alpha} \sum_{j=0}^{k-1} \rho_j^{\alpha} \left(\left|(Dw)_{B_j}\right| + d^{-\beta} \right).$$

It follows that

$$\begin{aligned} &|(Dw)_{B_k} - (Dw)_{B_0}| \\ &\quad \le \sum_{j=0}^{k-1} \left|(Dw)_{B_{j+1}} - (Dw)_{B_j}\right| \le \sum_{j=0}^{k-1} \fint_{B_{j+1}} \left|Dw - (Dw)_{B_j}\right| dx \\ &\quad \le \delta^{-\frac{n}{2}} \sum_{j=0}^{k-1} E_j \le 2\delta^{-\frac{n}{2}} E_0 + 2c\delta^{-\frac{n}{2}} d^{-\alpha} \sum_{j=0}^{k-2} \rho_j^{\alpha} \left(\left|(Dw)_{B_j}\right| + d^{-\beta} \right). \end{aligned}$$

Let $k^* \in \{0, 1, \dots, k\}$ be such that $\left|(Dw)_{B_j}\right| \le |(Dw)_{B_{k^*}}|$ for all $j = 0, 1, \dots, k$. Then

$$\begin{aligned} |(Dw)_{B_{k^*}}| &\le |(Dw)_{B_0}| + |(Dw)_{B_{k^*}} - (Dw)_{B_0}| \\ &\le |(Dw)_{B_0}| + 2\delta^{-\frac{n}{2}} E_0 + 2c\delta^{-\frac{n}{2}} d^{-\alpha} r^{\alpha} \sum_{j=0}^{\infty} (\delta^{\alpha})^j \left(|(Dw)_{B_{k^*}}| + d^{-\beta} \right). \end{aligned}$$

Since $\delta^{\alpha} < 1$, the above series converges to $\frac{1}{1-\delta^{\alpha}}$. Now we choose r such that $2c\frac{\delta^{-\frac{n}{2}}}{1-\delta^{\alpha}} d^{-\alpha} r^{\alpha} \le \frac{1}{2}$, that is, by taking into account the previous assumptions on r,

$$r = \min \left\{ 1, \left(1 + \delta^{-\frac{n}{2}-1}\right)^{-\frac{1}{\alpha}}, \left(\frac{4c\delta^{-\frac{n}{2}}}{1-\delta^{\alpha}} \right)^{-\frac{1}{\alpha}} \right\} d =: \gamma d.$$

Moreover, by Hölder's and Minkowski's inequalities we clearly have the estimates $|(Dw)_{B_0}| \le \left(\fint_{B_0} |Dw|^2 \, dx \right)^{\frac{1}{2}}$ and $E_0 \le 2 \left(\fint_{B_0} |Dw|^2 \, dx \right)^{\frac{1}{2}}$. Thus, we obtain

$$|(Dw)_{B_{k^*}}| \le |(Dw)_{B_0}| + 2\delta^{-\frac{n}{2}} E_0 + \frac{1}{2} |(Dw)_{B_{k^*}}| + \frac{1}{2} d^{-\beta},$$

which implies

$$|(Dw)_{B_k}| \le |(Dw)_{B_{k^*}}| \le 2\,|(Dw)_{B_0}| + 4\delta^{-\frac{n}{2}} E_0 + d^{-\beta} \le c\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} + d^{-\beta}.$$

This establishes (4.21).

Now (4.20) and (4.21) yield

$$E_{j+1} \le \delta^{\frac{1+\alpha}{2}} E_j + cd^{-\alpha}\rho_j^\alpha \left(\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} + d^{-\beta}\right).$$

If we then denote $\overline{\delta} := \delta^{\frac{1+\alpha}{2}}$ and $\Psi := cd^{-\alpha}\left(\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} + d^{-\beta}\right)$ and iterate the previous inequality, we obtain

$$\begin{aligned}
E_{j+1} &\le \overline{\delta} E_j + \rho_j^\alpha \Psi \\
&\le \overline{\delta}^2 E_{j-1} + \delta^{\frac{1+\alpha}{2}} \delta^{\alpha(j-1)} r^\alpha \Psi + \rho_j^\alpha \Psi \\
&= \overline{\delta}^2 E_{j-1} + \left(1 + \delta^{\frac{1-\alpha}{2}}\right) \rho_j^\alpha \Psi \\
&\le \overline{\delta}^3 E_{j-2} + \delta^{2\frac{1+\alpha}{2}} \delta^{\alpha(j-2)} r^\alpha \Psi + \left(1 + \delta^{\frac{1-\alpha}{2}}\right) \rho_j^\alpha \Psi \\
&= \overline{\delta}^3 E_{j-2} + \left(1 + \delta^{\frac{1-\alpha}{2}} + \delta^{2\frac{1-\alpha}{2}}\right) \rho_j^\alpha \Psi \\
&\le \overline{\delta}^{j+1} E_0 + \sum_{i=0}^{\infty} \delta^{i\frac{1-\alpha}{2}} \rho_j^\alpha \Psi \\
&= \overline{\delta}^{j+1} E_0 + \frac{\delta^{-\alpha}}{1 - \delta^{\frac{1-\alpha}{2}}} \rho_{j+1}^\alpha \Psi.
\end{aligned}$$

The series converges, since $\delta, \alpha < 1$ implies $\delta^{\frac{1-\alpha}{2}} < 1$. For the same reason we also have $\overline{\delta} < \delta^\alpha$, and therefore $\overline{\delta}^j < \delta^{\alpha j} = r^{-\alpha}\rho_j^\alpha = cd^{-\alpha}\rho_j^\alpha$. With the above estimate for E_0 we then obtain for $j = 1, 2, \ldots$

$$E_j \le \overline{\delta}^j E_0 + c\rho_j^\alpha \Psi \le cd^{-\alpha}\rho_j^\alpha \left(\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} + d^{-\beta}\right).$$

Using the fact that $B_0 \subset B_d$ and the Caccioppoli estimate, Lemma 4.1, gives

$$\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} \le \gamma^{-\frac{n}{2}} \left(\fint_{B_d} |Dw|^2\,dx\right)^{\frac{1}{2}} \le cd^{-1-\beta}, \tag{4.22}$$

and thus, since $d^{-1} \ge 1$, we have

$$E_j \le cd^{-\alpha}\rho_j^\alpha \left(d^{-1-\beta} + d^{-\beta}\right) \le cd^{-\bar{\beta}}\rho_j^\alpha. \tag{4.23}$$

When $j = 0$, this holds trivially, since

$$E_0 \le 2\left(\fint_{B_0} |Dw|^2\,dx\right)^{\frac{1}{2}} \le cd^{-1-\alpha-\beta} d^\alpha = c\gamma^{-\alpha} d^{-\bar{\beta}} \rho_0^\alpha.$$

Now let $0 < \rho \leq \gamma d$ and choose j such that $\rho_{j+1} < \rho \leq \rho_j$. The result is then obtained by applying Lemma 4.3 and estimate (4.23) and calculating

$$\begin{aligned}\left(\fint_{B_\rho} \left|Dw - (Dw)_{B_\rho}\right|^2 dx\right)^{\frac{1}{2}} &\leq \left(\frac{|B_j|}{|B_\rho|}\right)^{\frac{1}{2}} \left(\fint_{B_j} \left|Dw - (Dw)_{B_j}\right|^2 dx\right)^{\frac{1}{2}} \\ &= \left(\frac{\rho_j}{\rho}\right)^{\frac{n}{2}} E_j \leq \delta^{-\frac{n}{2}} c d^{-\bar{\beta}} \rho_j^\alpha \leq \delta^{-\frac{n}{2}-\alpha} c d^{-\bar{\beta}} \rho^\alpha.\end{aligned}$$

When $\gamma d < \rho \leq d$, we have by Lemma 4.1

$$\begin{aligned}\left(\fint_{B_\rho} \left|Dw - (Dw)_{B_\rho}\right|^2 dx\right)^{\frac{1}{2}} &\leq 2\left(\fint_{B_\rho} |Dw|^2 dx\right)^{\frac{1}{2}} \\ &\leq 2\left(\frac{|B_d|}{|B_\rho|}\right)^{\frac{1}{2}} \left(\fint_{B_d} |Dw|^2 dx\right)^{\frac{1}{2}} \\ &\leq 2\gamma^{-\frac{n}{2}} c d^{-1-\beta} \\ &\leq 2\gamma^{-\frac{n}{2}-\alpha} c d^{-\bar{\beta}} \rho^\alpha,\end{aligned}$$

and we are done. □

Now we have the required tools to prove the main result of the section, that is, that the weak solutions of equation (4.1) belong to $C^{1,\alpha}(\Omega)$.

Theorem 4.6. *Let $w \in W^{1,2}_{loc}(\Omega)$ be a weak solution of (4.1). Then there exists a constant $c = c(n, \lambda, \Lambda, \alpha, \beta, M)$ such that*

$$|Dw(x)| \leq c d_x^{-1-\beta} \tag{4.24}$$

and

$$|Dw(x) - Dw(y)| \leq c d_{x,y}^{-\bar{\beta}} |x - y|^\alpha \tag{4.25}$$

for all $x, y \in \Omega$, after possibly redefining Dw on a set of measure zero. Again, we denote $\bar{\beta} = 1 + \alpha + \beta$.

Proof. Assuming x_0 is a Lebesgue point, Lebesgue's differentiation theorem and estimates (4.21) and (4.22) give

$$|Dw(x_0)| = \lim_{j\to\infty} \left|(Dw)_{B_j}\right| \leq c\left(\fint_{B_0} |Dw|^2 dx\right)^{\frac{1}{2}} + d^{-\beta} \leq 4^\beta c d_{x_0}^{-1-\beta}. \tag{4.26}$$

If we now replace x_0 by any Lebesgue point $x \in \Omega$, we see that (4.24) holds for almost every $x \in \Omega$.

To show that (4.24) is true for every $x \in \Omega$, let us assume for now that (4.25) holds. For any $x \in \Omega$ take a sequence of Lebesgue points (x_i) such that $x_i \to x$ as $i \to \infty$. Now

$$|Dw(x)| \leq |Dw(x) - Dw(x_i)| + |Dw(x_i)| \leq c d_{x,x_i}^{-\bar{\beta}} |x - x_i|^\alpha + 4^\beta c d_{x_i}^{-1-\beta},$$

and by taking the limit we obtain the result.

Let us then prove (4.25). Take a Lebesgue point $x_1 \in B(x_0, \frac{d}{4})$ and denote $\rho := |x_0 - x_1|$ and $B_i := B(x_0, 2^{-i+1}\rho)$. By Lemma 4.5 we have

$$\begin{aligned}\left|(Dw)_{B_{i+1}} - (Dw)_{B_i}\right| &\leq \fint_{B_{i+1}} |Dw - (Dw)_{B_i}|\, dx \\ &\leq 2^n \left(\fint_{B_i} |Dw - (Dw)_{B_i}|^2\, dx\right)^{\frac{1}{2}} \\ &\leq cd^{-\bar{\beta}} 2^{(-i+1)\alpha} \rho^\alpha,\end{aligned}$$

and thus, for $k = 1, 2, \ldots$

$$|(Dw)_{B_k} - (Dw)_{B_0}| \leq \sum_{i=0}^{k-1} \left|(Dw)_{B_{i+1}} - (Dw)_{B_i}\right| \leq cd^{-\bar{\beta}}\rho^\alpha \sum_{i=0}^{\infty} 2^{-i\alpha} = cd^{-\bar{\beta}}\rho^\alpha.$$

Since x_0 is a Lebesgue point, Lebesgue's differentiation theorem yields

$$\left|Dw(x_0) - (Dw)_{B(x_0,2\rho)}\right| = \lim_{k\to\infty} |(Dw)_{B_k} - (Dw)_{B_0}| \leq cd^{-\bar{\beta}}\rho^\alpha = 4^\beta c d_{x_0}^{-\bar{\beta}}\rho^\alpha.$$

By replacing x_0 with x_1, we obtain

$$\left|Dw(x_1) - (Dw)_{B(x_1,2\rho)}\right| \leq 4^\beta c d_{x_1}^{-\bar{\beta}}\rho^\alpha.$$

Moreover, since $B(x_1, 2\rho) \subset B(x_0, 4\rho)$, we have

$$\begin{aligned}&\left|(Dw)_{B(x_0,2\rho)} - (Dw)_{B(x_1,2\rho)}\right| \\ &\quad\leq \left|(Dw)_{B(x_0,2\rho)} - (Dw)_{B(x_0,4\rho)}\right| + \left|(Dw)_{B(x_0,4\rho)} - (Dw)_{B(x_1,2\rho)}\right| \\ &\quad\leq \fint_{B(x_0,2\rho)} \left|Dw - (Dw)_{B(x_0,4\rho)}\right| dx + \fint_{B(x_1,2\rho)} \left|Dw - (Dw)_{B(x_0,4\rho)}\right| dx \\ &\quad\leq 2^{n+1} \left(\fint_{B(x_0,4\rho)} \left|Dw - (Dw)_{B(x_0,4\rho)}\right|^2 dx\right)^{\frac{1}{2}} \\ &\quad\leq 4^\beta c d_{x_0}^{-\bar{\beta}}\rho^\alpha\end{aligned}$$

by Lemma 4.5. Combining the previous estimates and the fact that $d_{x_0,x_1} \leq d_{x_0}, d_{x_1}$ now gives

$$\begin{aligned}|Dw(x_0) - Dw(x_1)| &\leq \left|Dw(x_0) - (Dw)_{B(x_0,2\rho)}\right| + \left|(Dw)_{B(x_0,2\rho)} - (Dw)_{B(x_1,2\rho)}\right| \\ &\qquad + \left|(Dw)_{B(x_1,2\rho)} - Dw(x_1)\right| \\ &\leq 4^\beta c d_{x_0,x_1}^{-\bar{\beta}}\rho^\alpha.\end{aligned}$$

If $x_1 \in \Omega \setminus B\left(x_0, \frac{d}{4}\right)$, we have $d_{x_0,x_1} \leq 4d \leq 16\rho$ and by (4.26) we may estimate

$$\begin{aligned}|Dw(x_0) - Dw(x_1)| &\leq 2\max\left\{|Dw(x_0)|, |Dw(x_1)|\right\} \\ &\leq 4^\beta c \max\left\{d_{x_0}^{-1-\beta}, d_{x_1}^{-1-\beta}\right\} \\ &= 4^\beta c d_{x_0,x_1}^{-1-\beta} \\ &\leq 4^\beta c d_{x_0,x_1}^{-\bar{\beta}}\rho^\alpha.\end{aligned}$$

Replacing x_0 and x_1 by any Lebesgue points x and y in Ω shows that (4.25) holds almost everywhere. A representative of Dw that satisfies (4.25) for all $x, y \in \Omega$ can be found in the exact same way as shown in the end of the previous section. $\square$

5 Smoothness of the minimizer

In this section we prove the main result of the thesis, that is, the smoothness of the minimizer u. We start by considering equation (3.2) and using the main result of the previous section, Theorem 4.6, to show that $u \in C^{2,\alpha}(\Omega)$. Then, by utilizing difference quotients as in Section 2, we show that the second derivatives of u belong to $W^{1,2}_{\text{loc}}(\Omega)$ and solve a slightly different equation of the type (4.1) in a weak sense. After that we use Theorem 4.6 again and repeat the process by induction.

Let us first prove a result that makes Theorem 4.6 so powerful. We show that, under suitable assumptions, any higher order derivative of a solution of equation (3.2) is a weak solution of an equation of the type (4.1). The idea behind the proof can be seen by formally differentiating equation (3.2) repeatedly and moving the leftover terms on the right hand side.

Lemma 5.1. *Let $w \in W^{k+1,2}_{loc}(\Omega)$ be a weak solution of (3.2) for a given positive integer k. Moreover, let $b_{ij} \in W^{k,\infty}_{loc}(\Omega)$. Then $w^\mu := D^\mu w$ is a weak solution of the equation*

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w^\mu) = \sum_{i=1}^{n} D_i g_i^\mu(x) \tag{5.1}$$

in Ω, where

$$g_i^\mu(x) := -\sum_{j=1}^{n}\sum_{\nu<\mu}\binom{\mu}{\nu} D^{\mu-\nu} b_{ij}(x) D_j D^\nu w(x)$$

and μ and ν are n-dimensional multi-indices with $|\mu| = k$.

Proof. Since w is a weak solution of (3.2), we have

$$\int_\Omega \sum_{i,j=1}^{n} b_{ij} D_j w D_i v dx = 0$$

for every $v \in W^{1,2}_0(\Omega)$. Take any $\varphi \in C^\infty_0(\Omega)$ and choose $v = D^\mu \varphi$ as the test function. By the generalized Leibniz rule we have

$$D^\mu(b_{ij}D_j w) = \sum_{\nu\le\mu}\binom{\mu}{\nu} D^{\mu-\nu} b_{ij} D^\nu D_j w.$$

Here all the derivatives of the functions b_{ij} are locally bounded by assumption. Furthermore, $D^\nu D_j w \in L^2_{\text{loc}}(\Omega)$ for all $\nu \le \mu$. Therefore, $b_{ij}D_j w \in W^{k,2}_{\text{loc}}(\Omega)$ and we may integrate by parts k times. This leads to

$$\begin{aligned}
0 &= \int_\Omega \sum_{i,j=1}^{n} b_{ij} D_j w D_i D^\mu \varphi dx \\
&= (-1)^k \int_\Omega \sum_{i,j=1}^{n} D^\mu(b_{ij}D_j w) D_i \varphi dx \\
&= (-1)^k \int_\Omega \sum_{i,j=1}^{n}\sum_{\nu\le\mu}\binom{\mu}{\nu} D^{\mu-\nu} b_{ij} D^\nu D_j w D_i \varphi dx.
\end{aligned}$$

Moving all the terms with $\nu < \mu$ to the other side yields

$$\int_\Omega \sum_{i,j=1}^n b_{ij} D_j D^\mu w D_i \varphi dx = -\int_\Omega \sum_{i,j=1}^n \sum_{\nu<\mu} \binom{\mu}{\nu} D^{\mu-\nu} b_{ij} D_j D^\nu w D_i \varphi dx,$$

and the result follows, since $C_0^\infty(\Omega)$ is dense in $W_0^{1,2}(\Omega)$. □

In Section 3 we saw that the first derivatives of the minimizer u belong to the local Hölder space $C^{0,\alpha}(\Omega)$. Let us now prove that, in fact, this holds for any order derivatives of u.

Theorem 5.2. *Let $u \in W^{1,\infty}(\Omega)$ be a minimizer of the functional $\mathcal{F}$. Then for every $k = 0, 1, 2, \ldots$ there exist constants $0 < \alpha < 1$ and $c = c\left(k, n, \lambda, \Lambda, ||Du||_{L^\infty(\Omega)}, F\right)$ such that for each $l = 1, \ldots, n$ $w := D_l u$ satisfies*

$$|DD^\mu w(x)| \leq c d_x^{-1-2k} \tag{5.2}$$

and

$$|DD^\mu w(x) - DD^\mu w(y)| \leq c d_{x,y}^{-1-\alpha-2k} |x-y|^\alpha \tag{5.3}$$

for every multi-index μ with $|\mu| = k$ and for all $x, y \in \Omega$.

Proof. The proof is by strong induction on k. For the base case $k = 0$ we consider equation (3.2), which is of the type (4.1) with $g \equiv 0$. By Corollary 2.15 $w \in W^{1,2}_{\text{loc}}(\Omega)$ is a weak solution of (3.2), so in order to apply Theorem 4.6 we only need to check that the assumptions hold.

Since the functions $a_{ij} = \partial_{z_i}\partial_{z_j} F$ are smooth and $Du(x) \in \Xi = \overline{B}\left(0, ||Du||_{L^\infty(\Omega)}\right)$ for every $x \in \Omega$, we obtain

$$|b_{ij}(x) - b_{ij}(y)| = |a_{ij}(Du(x)) - a_{ij}(Du(y))| \leq ||D^z a_{ij}||_{L^\infty(\Xi)} |Du(x) - Du(y)|$$

by the mean value theorem. Applying Theorem 3.6 to $D_l u$ yields

$$|Du(x) - Du(y)| = \left(\sum_{l=1}^n |D_l u(x) - D_l u(y)|^2\right)^{\frac{1}{2}} \leq c\, ||Du||_{L^\infty(\Omega)} d_{x,y}^{-\alpha} |x-y|^\alpha,$$

and summing over i and j leads to

$$\sum_{i,j=1}^n |b_{ij}(x) - b_{ij}(y)| \leq c\, ||Du||_{L^\infty(\Omega)} \sum_{i,j=1}^n ||D^z a_{ij}||_{L^\infty(\Xi)} d_{x,y}^{-\alpha} |x-y|^\alpha,$$

where $0 < \alpha < 1$ and c depend only on n, λ, and Λ. Since $u \in W^{1,\infty}(\Omega)$ and w is continuous by Theorem 3.6, we have $|w(x)| \leq ||Du||_{L^\infty(\Omega)}$ for all $x \in \Omega$. Therefore, we see that assumptions (4.2)-(4.4) hold if we choose $\beta = 0$ and

$$M = \max\left\{1, ||Du||_{L^\infty(\Omega)}, c\, ||Du||_{L^\infty(\Omega)} \sum_{i,j=1}^n ||D^z a_{ij}||_{L^\infty(\Xi)}\right\}.$$

Now Theorem 4.6 yields

$$|Dw(x)| \leq c d_x^{-1}$$

and

$$|Dw(x) - Dw(y)| \leq c d_{x,y}^{-1-\alpha} |x-y|^\alpha$$

for all $x, y \in \Omega$, where c depends only on $n, \lambda, \Lambda, ||Du||_{L^\infty(\Omega)}$ and F. This is exactly the claim when $k = 0$.

Let us next assume that the claim holds up to some index k. Then $w \in C^{k+1,\alpha}(\Omega)$, which implies $w \in W^{k+1,2}_{\text{loc}}(\Omega)$. Moreover, the coefficients b_{ij} inherit continuity properties from Du by the mean value theorem, and thus also from w, so we have $b_{ij} \in C^{k+1,\alpha}(\Omega) \subset W^{k+1,\infty}_{\text{loc}}(\Omega)$. Therefore, by Lemma 5.1 $D^\mu w =: w^\mu$ is a weak solution of equation (5.1), and in a similar way as in the proof of Theorem 2.14 we may use difference quotients to show that, in fact, $w \in W^{k+2,2}_{\text{loc}}(\Omega)$. Due to the similarities we shall omit some details.

Take any $\Omega' \subset\subset \Omega$ and some $\Omega' \subset\subset \tilde{\Omega} \subset\subset \Omega$. The set $\tilde{\Omega}$ will play the same role as Ω in the proof of Theorem 2.14. This is necessary, since this time the functions in question belong to their respective Sobolev spaces only locally.

Consider equation (5.1) in $\tilde{\Omega}$. Let $\Omega' \subset\subset \Omega'' \subset\subset \tilde{\Omega}$ be such that

$$\frac{1}{2} \operatorname{dist}(\Omega', \partial\tilde{\Omega}) \leq \operatorname{dist}(\Omega', \partial\Omega'') \leq \frac{3}{4} \operatorname{dist}(\Omega', \partial\tilde{\Omega}).$$

Choose a cut-off function $\eta \in C_0^\infty(\Omega'')$ as in the proof of Theorem 2.14 (with Ω replaced by $\tilde{\Omega}$) and h such that $0 < |h| < \frac{1}{2} \operatorname{dist}(\Omega'', \partial\tilde{\Omega})$. Then write (5.1) in the variational form and choose $-\Delta_m^{-h}(\eta^2 \Delta_m^h w^\mu) \in W_0^{1,2}(\tilde{\Omega})$ as the test function.

Now, using Lemma 2.11 part (ii) on both sides, we have

$$\int_{\tilde{\Omega}} \sum_{i,j=1}^n \Delta_m^h (b_{ij} D_j w^\mu) D_i(\eta^2 \Delta_m^h w^\mu) dx = \int_{\tilde{\Omega}} \sum_{i=1}^n \Delta_m^h g_i^\mu D_i(\eta^2 \Delta_m^h w^\mu) dx.$$

Then we use Lemma 2.11 part (iii) and the Leibniz rule and rearrange terms to obtain the estimate

$$\begin{aligned}
&\int_{\tilde{\Omega}} \sum_{i,j=1}^n b_{ij}(x + he_m) D_j \Delta_m^h w^\mu D_i \Delta_m^h w^\mu \eta^2 dx \\
&\qquad \leq \int_{\tilde{\Omega}} \left|\Delta_m^h g^\mu\right| \left|D\Delta_m^h w^\mu\right| \eta^2 dx + 2 \int_{\tilde{\Omega}} \left|\Delta_m^h g^\mu\right| \eta \left|D\eta\right| \left|\Delta_m^h w^\mu\right| dx \\
&\qquad\quad + 2 \int_{\tilde{\Omega}} \sum_{i,j=1}^n \left|b_{ij}(x + he_m)\right| \left|D\Delta_m^h w^\mu\right| \eta \left|D\eta\right| \left|\Delta_m^h w^\mu\right| dx \\
&\qquad\quad + \int_{\tilde{\Omega}} \sum_{i,j=1}^n \left|\Delta_m^h b_{ij}\right| \left|Dw^\mu\right| \left|D\Delta_m^h w^\mu\right| \eta^2 dx \\
&\qquad\quad + 2 \int_{\tilde{\Omega}} \sum_{i,j=1}^n \left|\Delta_m^h b_{ij}\right| \left|Dw^\mu\right| \eta \left|D\eta\right| \left|\Delta_m^h w^\mu\right| dx.
\end{aligned}$$

Above we saw that $b_{ij} \in C^{k+1,\alpha}(\Omega)$, which implies $b_{ij} \in C^{0,1}(\Omega)$. A small calculation exploiting the choices of Ω'' and h shows that the Lipschitz constant of b_{ij} in $\{x : \operatorname{dist}(x, \Omega'') < |h|\}$, say L_{ij}, depends only on $n, \lambda, \Lambda, ||Du||_{L^\infty(\Omega)}, F$, and $\operatorname{dist}(\Omega', \partial\tilde{\Omega})$. Thus, for any $x \in \Omega''$ we have

$$\sum_{i,j=1}^{n} \left|\Delta_m^h b_{ij}(x)\right| = \frac{1}{|h|} \sum_{i,j=1}^{n} |b_{ij}(x + he_m) - b_{ij}(x)| \leq \sum_{i,j=1}^{n} L_{ij} =: L.$$

This together with assumptions (3.3) and (3.4), Young's inequality with ε, and some more arranging yields

$$\begin{aligned}\lambda \int_{\tilde{\Omega}} \left|D\Delta_m^h w^\mu\right|^2 \eta^2 dx \leq \frac{\lambda}{2} \int_{\tilde{\Omega}} \left|D\Delta_m^h w^\mu\right|^2 \eta^2 dx + \left(\frac{3}{2\lambda} + 1\right) \int_{\tilde{\Omega}} \left|\Delta_m^h g^\mu\right|^2 \eta^2 dx \\ + \left(\frac{3L^2}{2\lambda} + L\right) \int_{\tilde{\Omega}} |Dw^\mu|^2 \eta^2 dx + \left(\frac{6\Lambda^2}{\lambda} + L + 1\right) \int_{\tilde{\Omega}} |D\eta|^2 \left|\Delta_m^h w^\mu\right|^2 dx.\end{aligned}$$

Now, after using the properties of the cut-off function, we arrive at a constant $c = c\left(n, \lambda, \Lambda, ||Du||_{L^\infty(\Omega)}, F, \operatorname{dist}(\Omega', \partial\tilde{\Omega})\right)$ such that

$$\left|\left|D\Delta_m^h w^\mu\right|\right|_{L^2(\Omega')} \leq c\left(\left|\left|\Delta_m^h g^\mu\right|\right|_{L^2(\Omega'')} + ||Dw^\mu||_{L^2(\Omega'')} + \left|\left|\Delta_m^h w^\mu\right|\right|_{L^2(\Omega'')}\right).$$

Since $w, b_{ij} \in C^{k+1,\alpha}(\Omega)$, the function g^μ belongs to the space $C^{1,\alpha}(\Omega)$ as a sum of products of derivatives up to order k of these two. It follows that $g^\mu \in W^{1,2}(\tilde{\Omega})$ and we can apply Lemma 2.13 part (i) for g^μ as well as for w^μ. Part (ii) of the same lemma then gives $D_m D w^\mu \in L^2(\Omega')$ with the estimate

$$||D_m D w^\mu||_{L^2(\Omega')} \leq c\left(||Dg^\mu||_{L^2(\tilde{\Omega})} + ||Dw^\mu||_{L^2(\tilde{\Omega})}\right).$$

Since $\Omega' \subset\subset \Omega$ and $m = 1, \ldots, n$ were arbitrary, we have shown that $w \in W^{k+2,2}_{\text{loc}}(\Omega)$.

Now we are in the position to use Lemma 5.1 for $k+1$. Thus, for a given multi-index $\tilde{\mu}$ with $|\tilde{\mu}| = k+1$, the function $w^{\tilde{\mu}} \in W^{1,2}_{\text{loc}}(\Omega)$ is a weak solution of equation (5.1). Since b_{ij} stays the same in every round, assumption (4.3) holds, as shown above. Moreover, due to the induction assumption, we have $\left|w^{\tilde{\mu}}(x)\right| \leq c d_x^{-1-2k}$ for every $x \in \Omega$. Both w and b_{ij} belong to $C^{k+1,\alpha}(\Omega)$, from which we deduce, similarly as above, that $g^{\tilde{\mu}} \in C^{0,\alpha}(\Omega)$. In fact, a rather tedious calculation shows that

$$\left|g^{\tilde{\mu}}(x)\right| \leq c d_x^{-2(k+1)}$$

and

$$\left|g^{\tilde{\mu}}(x) - g^{\tilde{\mu}}(y)\right| \leq c d_{x,y}^{-\alpha-2(k+1)} |x-y|^\alpha$$

for all $x, y \in \Omega$. Therefore, also assumptions (4.2) and (4.4) are satisfied, if we choose $\beta = 2(k+1)$.

By Theorem 4.6 we now obtain

$$\left|Dw^{\tilde{\mu}}(x)\right| \leq c d_x^{-1-2(k+1)}$$

and

$$\left|Dw^{\tilde{\mu}}(x) - Dw^{\tilde{\mu}}(y)\right| \leq c d_{x,y}^{-1-\alpha-2(k+1)} \left|x-y\right|^{\alpha}$$

for all $x, y \in \Omega$, where the constant c depends only on $k, n, \lambda, \Lambda, ||Du||_{L^\infty(\Omega)}$ and F. This concludes the proof. □

The smoothness of the minimizer u follows immediately.

Proof of Theorem 1.1. We may assume that $u \in W^{1,\infty}(\Omega)$ as seen in Section 2.2. By Theorem 5.2 we have $u \in C^{k+2,\alpha}(\Omega)$ for any $k = 0, 1, 2, \ldots$ and some $0 < \alpha < 1$. Therefore $u \in C^\infty(\Omega)$. □

6 Smoothness up to the boundary

After establishing the local smoothness of the minimizer u, the next question that naturally arises is how to generalize the result to cover global smoothness. This requires taking the estimates up to the boundary of Ω, and therefore, it is necessary to assume that the boundary together with the boundary values u_0 are smooth as well. Moreover, the minimizer u has to be assumed globally Lipschitz continuous, which implies $u \in W^{1,\infty}(\Omega)$. The existence of such a minimizer can be shown using direct methods in the calculus of variations as we saw in Section 2.2.

We shall only sketch the necessary steps for the proof. More details can be found, for example, in [9].

Step 1: De Giorgi's method up to the boundary

Suppose the minimizer u belongs to $W^{2,2}(\Omega)$. This can be shown by applying the reflection technique described below and the difference quotient argument as in Section 2. Let x_0 be a point on the boundary and choose a ball $B_\rho := B(x_0, \rho)$. Mimicking Section 3 we consider solutions $w \in W^{1,2}(\Omega)$ of the equation

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w) = 0$$

in $\Omega \cap B_\rho$ such that $w = w_0 := D_l u_0$ on $\partial\Omega$, where the boundary values u_0 are assumed to be the restriction of a smooth function in $\mathbb{R}^n$ to Ω. By replacing w with $w - w(x_0)$, if necessary, we may assume that $w(x_0) = 0$.

In order to prove the Caccioppoli estimate, Lemma 3.1, we have to assume that $k \geq \operatorname{osc}_{B_\rho} w_0$. Then the test function $\eta^2(w-k)_+$ vanishes on the boundary of Ω and we may extend it to be zero outside of Ω. Now the rest of the arguments work as such until Lemma 3.5, where we need to add $\operatorname{osc}_{B_{2r}} w_0$ in the definition of k_j due to the above restriction on k. But this causes no difficulties, since we can always estimate $\operatorname{osc}_{B_{2r}} w_0 \leq cr^\alpha$. Therefore, we deduce that $u \in C^{1,\alpha}(\overline{\Omega})$.

Step 2: Flattening the boundary

For the higher order derivatives consider the equation we studied in Section 4,

$$\sum_{i,j=1}^{n} D_i(b_{ij}(x)D_j w) = \sum_{i=1}^{n} D_i g_i(x), \tag{6.1}$$

in $B_\rho \cap \Omega$, where $w \in W^{1,2}(\Omega) \cap C^{0,\alpha}(\overline{\Omega})$ and $b_{ij}, g_i \in C^{0,\alpha}(\overline{\Omega})$. We may assume without loss of generality that the boundary values $w_0 = 0$, for if not, we replace g_i with $g_i - \sum_{j=1}^{n} b_{ij} D_j w_0$ and consider $w - w_0$ instead of w. Moreover, we may assume that $x_0 = 0$ and the hyperplane $x_n = 0$ is tangential to the boundary at x_0, since rotation and translation do not change the structure of (6.1).

Consider $B_\rho \cap \partial\Omega$ as the graph of a smooth function $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ such that $x_n = \psi(x_1, \ldots, x_{n-1})$ when $x \in \partial\Omega$. Denote $x' = (x_1, \ldots, x_{n-1})$ and change variables

to $y = \Psi(x) = (x', x_n - \psi(x'))$. The Jacobian determinant of Ψ is 1, so it is invertible. The idea behind the transformation is that it takes $B_\rho \cap \partial\Omega$ into a subset of the hyperplane $y_n = 0$, therefore flattening the boundary.

Let $\tilde{w}(y) = w(x)$ under the mapping $y = \Psi(x)$, that is, $\tilde{w} := w \circ \Psi^{-1}$. Then $\tilde{w}$ is a solution of an equation of the type (6.1) in the new variables, namely

$$\sum_{i,j=1}^{n} D_i^y(\tilde{b}_{ij}(y) D_j^y \tilde{w}) = \sum_{i=1}^{n} D_i^y \tilde{g}_i(y) \tag{6.2}$$

in $\Omega' := \Psi(B_\rho \cap \Omega)$, where

$$\begin{aligned}\tilde{b}_{ij}(y) := b_{ij}(x) - \delta_{in} \sum_{k=1}^{n-1} D_k\psi(x') b_{kj}(x) - \delta_{jn} \sum_{m=1}^{n-1} D_m\psi(x') b_{im}(x) \\ + \delta_{in}\delta_{jn} \sum_{k,m=1}^{n-1} D_k\psi(x') D_m\psi(x') b_{km}(x),\end{aligned}$$

$$\tilde{g}_i(y) := g_i(x) - \delta_{in} \sum_{k=1}^{n-1} D_k\psi(x') g_k(x).$$

Since ψ is smooth, we clearly have $\tilde{b}_{ij}, \tilde{g}_i \in C^{0,\alpha}(\overline{\Omega'})$, and also the ellipticity and boundedness conditions, (3.3) and (3.4), can be shown to hold for $\tilde{b}_{ij}$, provided the radius ρ is chosen small enough.

Step 3: Freezing and reflecting

Choose a radius, again denoted by ρ to match the notation in Section 4, and a half ball $B_\rho^+ := \{y \in B_\rho : y_n > 0\} \subset \Omega'$. Since $\tilde{w}$ vanishes on $y_n = 0$, the Caccioppoli estimate, Lemma 4.1, works as such, if the test function is defined to be zero in $B_\rho \setminus B_\rho^+$. The same applies for Lemma 4.2.

Lemma 4.4 requires a bit more work. The idea is to first change variables (back to x for notational convenience) such that y_n stays fixed and (6.2) frozen at 0 gets the form

$$\sum_{i,j=1}^{n} \hat{b}_{ij} D_i D_j w_\rho = 0, \tag{6.3}$$

where the elements $\hat{b}_{in}$ and $\hat{b}_{nj}$ are zero for $i, j < n$. Then we reflect w_ρ onto $B_\rho \setminus B_\rho^+$ by defining

$$w_\rho^*(x) := \begin{cases} w_\rho(x', x_n), & x_n \geq 0 \\ -w_\rho(x', -x_n), & x_n < 0 \end{cases}.$$

Now w_ρ^* satisfies (6.3) in the entire B_ρ due to the structure of the matrix $\hat{b}$. Then one more change of variables, identical to that in Lemma 4.4, gives us a harmonic function and the rest is clear. Also the rest of Section 4 works with minor modifications.

Again, collecting the results we may proceed to show the global smoothness of the minimizer u by using arguments similar to those in the previous section.

References

[1] Bernstein S.: Sur la nature analytique des solutions des équations aux dérivés partielles du second ordre. *Mathematische Annalen* 59 (1904), 20–76.

[2] Bernstein S.: Sur les équations du calcul des variations. *Ann. Sci. École Norm. Sup.* 29 (1912), 431–485.

[3] Campanato S.: Proprietà di hölderianità di alcune classi di funzioni. *Ann. Sc. Norm. Sup. Pisa.* 17 (1963), 175–188.

[4] Campanato S.: Proprietà di una familia di spazi funzionali. *Ann. Sc. Norm. Sup. Pisa.* 18 (1964), 137–160.

[5] De Giorgi E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* 3 (1957), 25–43.

[6] Erlichson H.: Johann Bernoulli's brachistochrone solution using Fermat's principle of least time. *Eur. J. Phys.* 20 (1999), 299–304.

[7] Evans L. C. & Gariepy R. F.: *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton, 1992.

[8] Gelfand I. M. & Fomin S. V.: *Calculus of variations.* Dover Publications, New York, 2000.

[9] Gilbarg D. & Trudinger N. S.: *Elliptic partial differential equations of second order.* Second Edition, Springer, Berlin-Heidelberg, 1983.

[10] Giusti E.: *Direct Methods in the Calculus of Variations.* World Scientific, Singapore, 2003.

[11] Hilbert D.: Mathematische Probleme. *Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen* (1900), 253–297.

[12] Hopf E.: Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung. *Mathematische Zeitschrift.* 34 (1932), 194–233.

[13] Malý J. & Ziemer W. P.: *Fine Regularity of Solutions of Elliptic Partial Differential Equations.* Amer. Math. Soc., Rhode Island, 1997.

[14] Meyers N. G. & Serrin J.: H=W. *Proc. Natl. Acad. Sci. U.S.A.* 51 (1964), 1055–1056.

[15] Moser J.: A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations. *Comm. Pure Appl. Math.* 13 (1960), 457–468.

[16] Nash J.: Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* 80 (1958), 931–954.

[17] Schauder J.: Über lineare elliptische Differentialgleichungen zweiter Ordnung. *Mathematische Zeitschrift.* 38 (1934), 257–282.

[18] Schauder J.: Numerische Abschätzungen in elliptischen linearen Differentialgleichungen. *Studia Mathematica.* 5 (1935), 34–42.

[19] Urbano J. M.: Regularity for Partial Differential Equations: from De Giorgi-Nash-Moser Theory to Intrinsic Scaling. *Centro Internacional de Matemática Bulletin.* 12 (2002), 8–14.

[20] Ziemer W. P.: *Weakly Differentiable Functions.* Springer-Verlag, New York, 1989.