CENTRE DE RECERCA MATEMÀTICA

Title:	*Four Advanced Courses on Quasiconformal Mappings, PDE and Geometric Measure Theory Notes of the Courses*
Journal Information:	*CRM Preprints*,
Author(s):	Four Advanced Courses on Quasiconformal Mappings, PDE and Geometric Measure Theory Notes of the Courses.
Volume, pages:	*1-152*, DOI:[−−]

Four Advanced Courses on Quasiconformal Mappings, PDE and Geometric Measure Theory

Notes of the Courses

Centre de Recerca Matemàtica
Bellaterra (Spain)

June 3rd to 12th, 2009

Acknowledgements. The Four Advanced Courses on Quasiconformal Mappings, PDE and Geometric Measure Theory is supported by the AGAUR (ref. 2008ARCS200037), by the Ingenio Mathematica Programme (ref. C3-0126) and, by the Ministerio de Ciencia e Innovación (ref. MTM2007-30708-E)

Contents

Foreword v

Kari Astala
The Calderon's inverse problem 1

Giovanna Citti
The Heisenberg Group and Cortex Vision 43

David Preiss
Geometric Measure Theory 107

Xiao Zhong
Second order Operators in Divergence Form and Quasiconformal Mappings. 123

The notes contained in this booklet were printed directly from files supplied by the authors before the course.

Foreword

This set of notes corresponds to the Advanced Courses held in June 2009 at the Centre de Recerca Matemàtica (CRM) in Bellaterra (Barcelona). The Advanced School was one of the main activities of the Research Semestre *Harmonic Analysis, Geometric Measure Theory and Quasiconformal Mappings*, which took place in the CRM from February to July 2009. The courses were the following:

– *The Calderón's Inverse Problem* (10 hours), by Kari Astala (University of Helsinki).
– *The Heisenberg Group and Cortex Vision* (5 hours), by Giovanna Citti (Università di Bologna).
– *Geometric Measure Theory* (5 hours), by David Preiss (University of Warwick).
– *Second Order Operators in Divergence Form and Quasiconformal Mappings* (10 hours), by Xiao Zhong (University of Jyväskylä).

The four courses deal with questions which pertain to the flourishing field of the so called geometric analysis. The one by Kari Astala explains the relationship between quasiconformal mappings and impedance tomography. In this area one wishes to obtain information of the internal structure of a body from electrical measurements on its surface. As shown by A.P. Calderón, it turns out that this problem admits a precise mathematical formulation in terms of partial differential equations which can be studied using quasiconformal techniques.

The course by Giovanna Citti, on the relationship between the Heisenberg group and cortex vision, is another instance of successful application of techniques of geometric analysis to physical sciences.

We are also fortunate of having David Preiss in the CRM teaching a course on geometric measure theory. In the last years, geometric measure theory has become an essential ingredient of important results in analysis. For example, in the study of the structure of harmonic measure the so called tangent measures have proved to be a powerful tool; and in the last advances on analytic capacity, the use of techniques of "quantitative rectifiability" has been essential.

Lastly, the course by Xiao Zhong on quasiconformal mappings and second order divergence type operators focuses on the study of second order divergence type operators using techniques of quasiconformal mappings. This course and the one by Astala complement very well each other.

We wish to express our gratitude to the director and the staff of the CRM who helped us in the organization of these courses. We thank the Ingenio-Mathematica programme of the Spanish government and the Catalan Research Funding Agency (AGAUR) for providing financial support for the organization of this Advanced Courses.

The Coordinators

Joan Mateu, Joan Orobitg, Joan Verdera, and Xavier Tolsa

The Calderón's Inverse Problem

Kari Astala

THE CALDERON'S INVERSE PROBLEM

KARI ASTALA

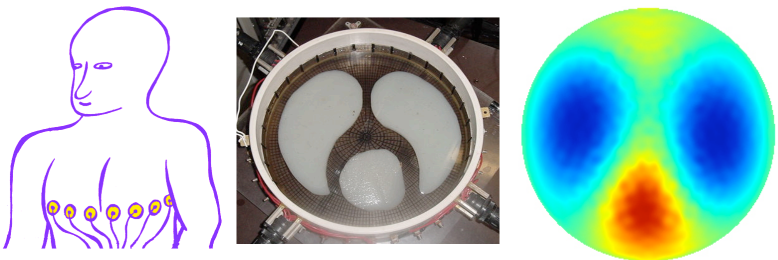
Advanced course at CRM, June 2009

The material is based on joint works with:

- Lassi Päivärinta and Matti Lassas (inverse problems)
- Tadeusz Iwaniec and Gaven Martin (elliptic equations and quasiconformal mappings)
- And many others !

1. Introduction

In impedance tomography one aims to determine the internal structure of a body from electrical measurements on its surface. Such methods have a variety of different applications for instance in engineering and medical diagnostics. For a general expository presentations see [12, 13], for medical applications see [15].



Medical imaging in two dimensions

In 1980 A.P. Calderón showed that the impedance tomography problem admits a clear and precise mathematical formulation. Indeed, suppose that $\Omega \subset \mathbb{R}^n$ is a bounded domain with connected complement and let $\sigma : \Omega \to (0, \infty)$ be a measurable function that is bounded away from zero and infinity.

Then the Dirichlet problem

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{in } \Omega, \tag{1}$$

$$u\big|_{\partial\Omega} = \phi \in W^{1/2,2}(\partial\Omega) \tag{2}$$

The author was supported by the Academy of Finland, by the Finnish Center of Exellence Analysis and Dynamics and project CODY of the European Commission.

admits a unique solution $u \in W^{1,2}(\Omega)$. Here

$$W^{1/2,2}(\partial\Omega) = H^{1/2}(\partial\Omega) = W^{1,2}(\Omega)/W_0^{1,2}(\Omega)$$

stands for the space of elements $\phi + W_0^{1,2}(\Omega)$, where $\phi \in W^{1,2}(\Omega)$. This is the most general space of functions that can possibly arise as Dirichlet boundary values or traces of general $W^{1,2}(\Omega)$-functions in a bounded domain Ω.

In terms of physics, if we charge the "body" Ω with an electric current, then $\phi = u|\partial\Omega$ represents the potential difference from $\partial\Omega$ to ∞. Furthermore, on the boundary the electric current J is equal to

$$J = (\sigma \nabla u)|_{\partial\Omega}$$

In practice, one can measure only the normal component of the current, $\sigma \partial u/\partial\nu$, with ν the unit outer normal to the boundary. For smooth σ this quantity is well defined pointwise, while for general bounded measurable σ we need to use the (equivalent) definition

$$\langle \sigma \frac{\partial u}{\partial \nu}, \psi \rangle = \int_\Omega \sigma \nabla u \cdot \nabla \psi \,, \qquad \psi \in W^{1,2}(\Omega), \tag{3}$$

as an element of the dual of $H^{1/2}(\partial\Omega) = W^{1/2,2}(\partial\Omega)$.

The inverse conductivity problem of Calderón asks if we can recover the pointwise conductivity $\sigma(x)$ *inside* the domain Ω from voltage/current measurements on the boundary $\partial\Omega$. In mathematical terms, the question is if the boundary data

$$\Big(u\big|_{\partial\Omega},\ \sigma \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega} \Big), \qquad \text{where } \ u\big|_{\partial\Omega} \in W^{1/2,2}(\partial\Omega),$$

determines the coefficient $\sigma(x)$ in (1) for all $x \in \Omega$. Equivalently, we may express the boundary data in terms of the operator

$$\Lambda_\sigma : \phi = u\big|_{\partial\Omega} \to \sigma \frac{\partial u}{\partial \nu}|_{\partial\Omega}, \tag{4}$$

the *Dirichlet-to-Neumann boundary map*, which can be considered as an operator from $W^{1/2,2}(\partial\Omega)$ to $W^{-1/2,2}(\partial\Omega)$. Thus the question we have is if the boundary operator Λ_σ determines the coefficient σ.

In this course we show how the Calderón can be solved in two dimensions. The presentation is based on the works [9, 5, 6] and the monograph [3], where more information can be found.

The physical background of the problem requires divergence type equations where the coefficients are not continuous, rather the assumption $\sigma \in L^\infty$ is the natural one. And it is the L^∞-setup that requires the quasiconformal techniques, in fact we shall make strong use of them.

There are a number of earlier results on this problem, assuming greater regularity, for instance by R. Brown, J. Sylvester, G. Uhlmann and A. Nachmann; [26, 27, 29]. in higher dimensions Calderón's problem remains open, unless some smoothness is assumed. In higher dimensions the usual method is to reduce, by substituting $v = \sigma^{1/2}u$, the conductivity equation (1) to the Schrödinger equation and then to apply the methods of scattering theory. Indeed, after such a substitution v satisfies

$$\Delta v - qv = 0,$$

where $q = \sigma^{-1/2}\Delta\sigma^{1/2}$. This substitution is possible only if σ has some smoothness. In the case $\sigma \in L^\infty$, relevant for practical applications, in general there is no smoothness and the reduction to the Schrödinger equation fails. Therefore one must turn to complex analytic tools.

However, what we adopt from the scattering theory type approaches is the use of exponentially growing solutions, the so called geometric optics solutions to the conductivity equation (1). These are specified by the condition

$$u(z,\xi) = e^{i\xi z}\left(1 + \mathcal{O}\left(\frac{1}{|z|}\right)\right) \quad \text{as } |z| \to \infty \tag{5}$$

Here we have set $\sigma \equiv 1$ outside Ω to get an equation defined globally. Studying the ξ-dependence of these solutions then gives rise to the basic concept of these notes, the **nonlinear Fourier transform** $\tau_\sigma(\xi)$. The detailed definition will be given Section 6.

Thus to start the study of $\tau_\sigma(\xi)$ we need first to establish the existence of exponential solutions, for $\sigma^\pm \in L^\infty(\mathbb{C})$ or even for degenerate σ's. Already here the quasiconformal techniques are essential. The themes of these notes can (should!) be considered as a study the non-linear Fourier transform: It is not difficult to show that the Dirichlet-to-Neumann boundary operator Λ_σ determines the nonlinear Fourier transforms $\tau_\sigma(\xi)$ for all $\xi \in \mathbb{C}$. Therefore the main difficulty, and our main strategy, is to invert the nonlinear Fourier transform, show that $\tau_\sigma(\xi)$ determines $\sigma(z)$ almost everywhere.

The properties of the nonlinear Fourier transform depend on the underlying differential equation. In one dimension the basic properties of the transform are fairly well understood, while deeper results such as analogs of Carleson's L^2-converge theorem remain open. The reader should consult the excellent lecture notes of Tao and Thiele [30] for an introduction to the one-dimensional theory.

For (1) with nonsmooth σ, many basic questions concerning the nonlinear Fourier transform, even such as finding a right version of the Plancherel formula, remain open. What we are able to show is that for $\sigma^{\pm 1} \in L^\infty$, with $\sigma \equiv 1$ near ∞, we have a Riemann-Lebesgue type result,

$$\tau_\sigma \in C_0(\mathbb{C}).$$

Indeed, this requires the asymptotic estimates of the solutions (5), and these are the key point and main technical part of our argument. For results on related equations, see [25].

To avoid some of the technical complications, in these notes we shall assume that the domain $\Omega = \mathbb{D}$, the unit disk. In fact, see [9], the reduction of general Ω to this case is not difficult. A main result of these notes is the following:

Theorem 1.1. [9] *Let $\sigma_j \in L^\infty(\mathbb{D})$, $j = 1, 2$. Suppose that there is a constant $c > 0$ such that $c^{-1} \le \sigma_j \le c$. If*

$$\Lambda_{\sigma_1} = \Lambda_{\sigma_2},$$

then $\sigma_1 = \sigma_2$ almost everywhere. Here Λ_{σ_i}, $i = 1, 2$, are defined by (4).

For the first steps in numerical implementation of our method see [7].

The proof and the necessary auxiliary results to this theorem will take the most of the time of this course. Towards the end we will also consider non-isotropic conductivities as well as degenerate ones, and study the limits of Calderón problem and impedance tomography.

Our approach will be based on quasiconformal methods, which also enables the use of tools from complex analysis. These are not available in higher dimensions, at least to the same extent, and this is one of the reasons why the problem is still open for L^∞-coefficients in $D \ge 3$. The complex analytic connection comes as follows: From Theorem 2.2 below we see that if $u \in W^{1,2}(\mathbb{D})$ is a real-valued solution of (1), then it has the σ-harmonic conjugate $v \in W^{1,2}(\mathbb{D})$ such that

$$\partial_x v = -\sigma \partial_y u \tag{6}$$

$$\partial_y v = \sigma \partial_x u \tag{7}$$

Equivalently (see (24)), the function $f = u + iv$ satisfies the $\mathbb{R}$-linear Beltrami equation

$$\frac{\partial f}{\partial \bar z} = \mu(z) \overline{\frac{\partial f}{\partial z}}, \tag{8}$$

where

$$\mu = \frac{1-\sigma}{1+\sigma}$$

In particular, note that μ is real-valued and that the assumptions on σ in Theorem 1.1 imply $\|\mu\|_{L^\infty} \le k < 1$. This reduction to the Beltrami equation and the complex analytic methods it provides will be the main tools in our analysis of the Dirichlet-to-Neumann map and the solutions to (1).

2. Linear and non-linear Beltrami equations

The most powerful tool for finding the exponential growing solutions to the conductivity equation (including degenerate conductivities) are given by the non-linear Beltrami equation. We therefore first review few of the basic facts here. For more details and results see [3].

We start with general facts on the linear divergence-type equation

$$\operatorname{div} A(z) \nabla u = 0, \qquad z \in \Omega \subset \mathbb{R}^2 \tag{9}$$

where we assume that $u \in W^{1,2}_{loc}(\Omega)$ and that the coefficient matrix

$$A = A(z) = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}, \qquad \alpha_{21} = \alpha_{12}, \tag{10}$$

is symmetric and elliptic,

$$\frac{1}{K(z)}|\xi|^2 \leq \langle A(z)\xi, \xi\rangle \leq K(z)|\xi|^2, \qquad \xi \in \mathbb{R}^2, \tag{11}$$

almost everywhere in Ω. For much of this course $A(z)$ is assumed to be isotropic, $A(z) = \sigma(z)\mathbf{I}$. The factor K here can be either a constant or a measurable function with $1 \leq K(z) < \infty$ almost everywhere. Until further notice we let K to be a constant.

For many purposes it is convenient to express the above ellipticity condition, equivalently, in terms of the following single inequality:

$$|\xi|^2 + |A(z)\xi|^2 \leq \left(K + \frac{1}{K}\right)\langle A(z)\xi, \xi\rangle \tag{12}$$

for almost every $z \in \Omega$ and all $\xi \in \mathbb{R}^2$. For the symmetric matrix $A(z)$ this is seen via construction of the eigenbasis.

In these notes we will study the divergence equation (9) by reducing it to the complex Beltrami system. For solutions to the divergence equation (9) a conjugate structure, similar to harmonic functions, is provided by the *Hodge star* operator $*$, which here really is nothing more than the (counterclockwise) rotation by 90 degrees,

$$* = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad ** = -\mathbf{I} \tag{13}$$

There are two vector fields associated with each solution to the homogeneous equation

$$\operatorname{div} A(z)\nabla u = 0, \qquad u \in W^{1,2}_{loc}(\Omega)$$

The first, $E = \nabla u$, has zero curl (in the sense of distributions, the curl of any gradient field is zero), while the second, $B = A(z)\nabla u$, is divergence-free as a solution to the equation.

It is the Hodge star $*$ operator that transforms curl-free fields into divergence-free fields, and vice versa. In particular, if

$$E = \nabla w = (w_x, w_y), \qquad w \in W^{1,1}_{loc}(\Omega),$$

then $*E = (-w_y, w_x)$ and hence

$$\operatorname{div}(*E) = \operatorname{div}(*\nabla w) = 0,$$

at least in the distributional sense. We recall here the following well-known fact from calculus (the Poincaré lemma).

Lemma 2.1. *Let $E \in L^p(\Omega, \mathbb{R}^2)$, $p \geq 1$, be a vector field defined on a simply connected domain Ω. If $\operatorname{Curl} E = 0$, then E is a gradient field; that is, there exists a real-valued function $u \in \mathbb{W}^{1,p}(\Omega)$ such that $\nabla u = E$.*

Thus in simply connected domains the A-harmonic equation $\operatorname{div} A(z)\nabla u = 0$ implies that the field $*A\nabla u$ is curl-free and may be rewritten as

$$\nabla v = *A(z)\nabla u, \tag{14}$$

where $v \in W^{1,2}_{loc}(\Omega)$ is some Sobolev function unique up to an additive constant. This function v we call the *A-harmonic conjugate* of u. Sometimes in the literature one also finds the term *stream function* used for v.

The ellipticity conditions for A can be equivalently formulated for the induced complex function $f = u + iv$. We arrive, after a lengthy but quite routine purely algebraic manipulation, at the equivalent complex first-order equation for $f = u + iv$, which we record in the following theorem.

Theorem 2.2. *Let Ω be a simply connected domain and let $u \in W^{1,1}_{loc}(\Omega)$ be a solution to*

$$\operatorname{div} A\,\nabla u = 0 \tag{15}$$

If $v \in W^{1,1}(\Omega)$ is a solution to the conjugate A-harmonic equation (14), then the function $f = u + iv$ satisfies the homogeneous Beltrami equation

$$\frac{\partial f}{\partial \bar z} - \mu(z)\frac{\partial f}{\partial z} - \nu(z)\,\overline{\frac{\partial f}{\partial z}} = 0 \tag{16}$$

The coefficients are given by

$$\mu = \frac{\alpha_{22} - \alpha_{11} - 2i\alpha_{12}}{1 + \mathrm{Trace}(A) + \det A}, \qquad \nu = \frac{1 - \det A}{1 + \mathrm{Trace}(A) + \det A}. \tag{17}$$

Conversely, if $f \in W^{1,1}_{loc}(\Omega, \mathbb{C})$ is a mapping satisfying (16), then $u = Re(f)$ and $v = Im(f)$ satisfy (14) with A given by solving the complex equations in (17),

$$\alpha_{11}(z) \;=\; \frac{|1-\mu|^2 - |\nu|^2}{|1+\nu|^2 - |\mu|^2} \tag{18}$$

$$\alpha_{22}(z) \;=\; \frac{|1+\mu|^2 - |\nu|^2}{|1+\nu|^2 - |\mu|^2} \tag{19}$$

$$\alpha_{12}(z) = \alpha_{21}(z) \;=\; \frac{-2\,Im\,(\mu)}{|1+\nu|^2 - |\mu|^2}, \tag{20}$$

The ellipticity of A can be explicitly measured in terms of μ and ν. The optimal ellipticity bound in (11) is

$$K(z) = \max\{\lambda_1(z), 1/\lambda_2(z)\}, \tag{21}$$

where $0 < \lambda_2(z) \le \lambda_1(z) < \infty$ are the eigenvalues of $A(z)$. With this choice we have pointwise

$$|\mu(z)| + |\nu(z)| = \frac{K(z)-1}{K(z)+1} < 1. \tag{22}$$

The (usual) notational convention:

$$k := \||\mu(z)| + |\nu(z)|\|_\infty, \quad K := \frac{1+k}{1-k}. \quad \text{(when } k<1\text{, equivalenty, } K<\infty\text{)}.$$

From (16)

$$\left|\frac{\partial f}{\partial \bar z}\right| \le k \left|\frac{\partial f}{\partial z}\right|$$

which is equivalent to

$$\|Df(z)\|^2 \le K J(z,f). \tag{23}$$

A mapping $f \in W^{1,2}_{loc}(\Omega)$ satisfying (23) is called a *K-quasiregular mappings.* If f is a homeomorphism, we call it $K-$*quasiconformal*. By Stoilow's factorization Theorem B.9 any K-quasiregular mapping is a composition of holomorphic function and a K-quasiconformal mapping.

Note the following:

- In this correspondence, ν is real valued if and only if the matrix A is symmetric.

- A has determinant 1 if and only if $\nu = 0$ (the $\mathbb{C}$-linear Beltrami equation).

- A is isotropic, meaning $A = \sigma(z)\mathbf{I}$ with $\sigma(z) \in \mathbb{R}$, if and only if $\mu = 0$. The complex equation now takes the form

$$\frac{\partial f}{\partial \bar{z}} - \frac{1-\sigma}{1+\sigma}\overline{\frac{\partial f}{\partial z}} = 0 \tag{24}$$

2.1. **Existence and uniqueness for non-linear Beltrami equations.** Solutions to the Beltrami equation conformal near infinity are particularly useful in solving the equation.

Principal Solutions:

When μ and ν as above have compact support and we have a $W^{1,2}_{loc}(\mathbb{C})$ solution to the Beltrami equation $f_{\bar{z}} = \mu f_z + \nu f_{\bar{z}}$ normalized by the condition

$$f(z) = z + \mathcal{O}(1/z)$$

near ∞, we call f a *principal solution.* Indeed, with the Cauchy and Beurling transform (see the appendix) we have the identities

$$\frac{\partial f}{\partial z} = 1 + \mathcal{S}\,\frac{\partial f}{\partial \bar{z}}, \tag{25}$$

and

$$f(z) = z + \mathcal{C}\left(\frac{\partial f}{\partial \bar{z}}\right)(z), \qquad z \in \mathbb{C} \tag{26}$$

Principal solutions are necessarily homeomorphisms. In fact we have the following fundamental Measurable Riemann Mapping Theorem,

Theorem 2.3. *Let $|\mu| \le k < 1$ be compactly supported and defined on $\mathbb{C}$. Then there is a unique principal solution to the Beltrami equation*

$$\frac{\partial f}{\partial \bar{z}} = \mu(z)\frac{\partial f}{\partial z} \quad \text{for almost every } z \in \mathbb{C},$$

and the solution $f \in W^{1,2}_{loc}(\mathbb{C})$ is a K-quasiconformal homeomorphism of $\mathbb{C}$.

The result holds also for the general Beltrami equation with coefficients μ and ν, see Theorem 2.4 below.

In constructing the exponentially growing solutions to the divergence and Beltrami equations, the most powerful approach is by non-linear Beltrami equations which we next discuss.

When one is looking for solutions to the general nonlinear elliptic systems

$$\frac{\partial f}{\partial \bar{z}} = H\big(z, f, \frac{\partial f}{\partial z}\big), \qquad z \in \mathbb{C} \tag{27}$$

there are necessarily some constraints to be placed on the function H that we now discuss. We write

$$H : \mathbb{C} \times \mathbb{C} \times \mathbb{C} \to \mathbb{C}.$$

We will not strive for full generality, but settle for the following special case. For the most general existence results, with very weak assumptions on H, see [3]. Here we assume:

(1) The homogeneity condition, that $f_{\bar{z}} = 0$ whenever $f_z = 0$, equivalently,

$$H(z, w, 0) \equiv 0, \quad \text{ for almost every } (z, w) \in \mathbb{C} \times \mathbb{C}$$

(2) The uniform ellipticity condition, that for almost every $z, w \in \mathbb{C}$ and all $\zeta, \xi \in \mathbb{C}$,

$$|H(z, w, \zeta) - H(z, w, \xi)| \leq k|\zeta - \xi|, \qquad 0 \leq k < 1$$

(3) The Lipschitz continuity in the function variable,

$$|H(z, w_1, \zeta) - H(z, w_2, \zeta)| \leq C|\zeta|\, |w_1 - w_2|$$

for some absolute constant C independent of z and ζ.

Theorem 2.4. *Suppose $H : \mathbb{C} \times \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ satisfies the conditions 1–3 above and is compactly supported in the z-variable. Then the uniformly elliptic nonlinear differential equation*

$$\frac{\partial f}{\partial \bar{z}} = H\Big(z, f, \frac{\partial f}{\partial z}\Big) \tag{28}$$

admits exactly one principal solution $f \in W^{1,2}_{loc}(\mathbb{C})$.

Proof. (*Sketch*). Uniqueness is easy. Suppose that both f and g are principal solutions to (28), so

$$\begin{aligned} \frac{\partial f}{\partial \bar{z}} &= H\Big(z, f, \frac{\partial f}{\partial z}\Big) \\ \frac{\partial g}{\partial \bar{z}} &= H\Big(z, g, \frac{\partial f}{\partial z}\Big) \end{aligned}$$

We set

$$F = f - g$$

and estimate

$$\begin{aligned} |F_{\bar{z}}| &= |H(z, f, f_z) - H(z, g, g_z)| \\ &\leq |H(z, f, f_z) - H(z, f, g_z)| + |H(z, f, g_z) - H(z, g, g_z)| \\ &\leq k|f_z - g_z| + C\chi_R |g_z||f - g|, \end{aligned}$$

where χ_R denotes the characteristic function of the disk $\mathbb{D}(0,R)$. Put briefly, F satisfies the differential inequality

$$|F_{\bar{z}}| \leq k|F_z| + C\chi_R|g_z||F|$$

By assumption, the principal solutions $f, g \in W^{1,2}_{loc}(\mathbb{C})$ with

$$\lim_{z\to\infty} f(z) - g(z) = 0$$

Once we observe that

$$\sigma = C\chi_R(z)|g_z| \in L^2(\mathbb{C})$$

and has compact support, Liouville type results such as Theorem B.8 in the Appendix shows us that $F \equiv 0$, as desired.

The proof of existence we only sketch, for details, in the more general setup of Lusin measurable H, see [3, Chapter 8].

We look for a solution f in the form

$$f(z) = z + \mathcal{C}\phi, \qquad \phi \in L^p(\mathbb{C}) \quad \text{of compact support,} \tag{29}$$

where the exponent $p > 2$. Note that

$$f_{\bar{z}} = \phi, \qquad f_z = 1 + \mathcal{S}\phi.$$

Thus we need to solve only the following integral equation:

$$\phi = H\big(z, z + \mathcal{C}\phi, 1 + \mathcal{S}\phi\big) \tag{30}$$

To solve this equation we first associate with every given $\phi \in L^p(\mathbb{C})$ an operator $\mathbf{R} : L^p(\mathbb{C}) \to L^p(\mathbb{C})$ defined by

$$\mathbf{R}\Phi = H\big(z, z + \mathcal{C}\phi, 1 + \mathcal{S}\Phi\big)$$

Through the ellipticity hypothesis we observe that $\mathbf{R}$ is a contractive operator on $L^p(\mathbb{C})$. Indeed, from (28) we have the pointwise inequality

$$|\mathbf{R}\Phi_1 - \mathbf{R}\Phi_2| \leq k\,|\mathcal{S}\Phi_1 - \mathcal{S}\Phi_2|$$

Hence

$$\|\mathbf{R}\Phi_1 - \mathbf{R}\Phi_2\|_p \leq k\,\mathbf{S}_p\|\Phi_1 - \Phi_2\|_p, \qquad k\mathbf{S}_p < 1,$$

for p sufficiently close to 2. By the Banach contraction principle, $\mathbf{R}$ has a unique fixed point $\Phi \in L^p(\mathbb{C})$. In other words, with each $\phi \in L^p(\mathbb{C})$ we can associate a unique function $\Phi \in L^p(\mathbb{C})$ such that

$$\Phi = H\big(z, z + \mathcal{C}\phi, 1 + \mathcal{S}\Phi\big) \tag{31}$$

In fact, the procedure (31), $\phi \mapsto \Phi$, gives a well-defined and nonlinear operator $\mathbf{T} : L^p(\mathbb{C}) \to L^p(\mathbb{C})$ by simply requiring that $\mathbf{T}\phi = \Phi$. Further, solving the original integral equation (30) means precisely that we have to find a fixed point for the operator $\mathbf{T}$. This, however, is more involved than in the case of the contraction $\mathbf{R}$, and one needs to invoke the celebrated Schauder fixed-point theorem, see [3, Chapter 8] for details. $\square$

3. Complex Geometric Optics Solutions

We will use the following convenient notation

$$e_\xi(z) = e^{i(z\xi + \bar z \bar\xi)}, \qquad z, \xi \in \mathbb{C} \tag{32}$$

The main emphasize in these notes is on isotropic conductivities, corresponding to the Beltrami equations of type (24). However, for later purposes it is useful to consider exponentially growing solutions to divergence equations with matrix coefficients, hence we are led to general Beltrami equations.

We will extend the coefficient matrix $A(z)$ to the entire plane $\mathbb{C}$ by requiring $A(z) \equiv \mathbf{I}$ when $|z| \geq 1$. Clearly, this keeps all ellipticity bounds. Moreover, then

$$\mu(z) \equiv \nu(z) \equiv 0, \qquad |z| \geq 1$$

As a first step toward Theorem 1.1, we establish the existence of a family of special solutions to (16). These, called the complex geometric optics solutions, are specified by having the asymptotics

$$f_{\mu,\nu}(z,\xi) = e^{i\xi z} M_{\mu,\nu}(z,\xi), \tag{33}$$

where

$$M_{\mu,\nu}(z,\xi) - 1 = \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } |z| \to \infty \tag{34}$$

Theorem 3.1. *For each parameter $\xi \in \mathbb{C}$ and for each $2 \leq p < 1 + 1/k$, the equation*

$$\frac{\partial f}{\partial \bar z} = \mu(z)\, \frac{\partial f}{\partial z} + \nu(z) \overline{\frac{\partial f}{\partial z}} \tag{35}$$

admits a unique solution $f = f_{\mu,\nu} \in W^{1,p}_{loc}(\mathbb{C})$ that has the form (33) with (34) holding. In particular, $f(z, 0) \equiv 1$.

Proof. Any solution to (35) is quasiregular. If $\xi = 0$, (33) and (34) imply that f is bounded, hence constant by the Liouville theorem.

If $\xi \neq 0$, look for a solution $f = f_{\mu,\nu}(z,\xi)$ in the form

$$f(z,\xi) = e^{i\xi \psi_\xi(z)}, \qquad \psi_\xi(z) = z + \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } |z| \to \infty \tag{36}$$

Substituting (36) into (35) indicates that ψ_ξ is the principal solution to the quasilinear equation

$$\frac{\partial}{\partial \bar z}\, \psi_\xi(z) = \mu(z) \frac{\partial}{\partial z}\, \psi_\xi(z) - \frac{\bar\xi}{\xi}\, e_{-\xi}\big(\psi_\xi(z)\big)\, \nu(z)\, \overline{\frac{\partial}{\partial z}\, \psi_\xi(z)} \tag{37}$$

The function $H(z,w,\zeta) = \mu(z)\zeta - (\bar\xi/\xi)\, \nu(z)\, e_\xi(w)\, \bar\zeta$ satisfies requirements 1–3 of Theorem 2.4. We thus obtain the existence and uniqueness of the principal solution ψ_ξ in $W^{1,2}_{loc}(\mathbb{C})$. Equation (37) together with Theorem B.5 yields $\psi_\xi \in W^{1,p}_{loc}(\mathbb{C})$ for all $p < 1 + 1/k$ since $|\mu(z)| \leq k$ and e_ξ is unimodular.

Finally, to see the uniqueness of the complex geometric optics solution $f_{\mu,\nu}$, let $f \in W^{1,2}_{loc}(\mathbb{C})$ be a solution to (35) satisfying

$$f = \alpha e^{i\xi z}\left(1 + \mathcal{O}\left(\frac{1}{z}\right)\right) \quad \text{as } |z| \to \infty \tag{38}$$

Denote then

$$\mu_1(z) = \mu(z)\frac{\overline{\partial_z f(z)}}{\partial_z f(z)}$$

where $\partial_z f(z) \neq 0$ and set $\mu_1 = 0$ elsewhere. Next, let φ be the unique principal solution to

$$\frac{\partial \varphi}{\partial \bar{z}} = \mu_1 \frac{\partial \varphi}{\partial \bar{z}} \tag{39}$$

Then the Stoilow factorization, Theorem B.9, gives $f = h \circ \varphi$, where $h : \mathbb{C} \to \mathbb{C}$ is an entire analytic function. But (38) shows that

$$\frac{h \circ \varphi(z)}{\exp(i\xi\varphi(z))} = \frac{f(z)}{\exp(i\xi\varphi(z))}$$

has the limit α when the variable $z \to \infty$. Thus

$$h(z) \equiv \alpha e^{i\xi z}$$

Therefore $f(z) = \alpha \exp(i\xi\varphi(z))$. In particular, If we have two solutions f_1, f_2 satisfying (33), (34), then the argument gives

$$f_\varepsilon := f_1 - (1+\varepsilon) f_2 = \varepsilon e^{i\xi\varphi(z)},$$

The principal solutions are homeomorphisms with$\phi(z) = z + \frac{1}{z}$ as $|z| \to \infty$, where the error term is uniformly bounded by Koebe distortion, Theorem B.6. Letting now $\varepsilon \to 0$ gives $f_1 = f_2$. $\square$

It is useful to note that if a function f satisfies (35), then if satisfies not the same equation but the equation where ν is replaced by $-\nu$. In terms of the real and imaginary parts of $f = u + iv$, we see that

$$\frac{\partial f}{\partial \bar{z}} = \mu(z)\,\frac{\partial f}{\partial z} + \nu(z)\overline{\frac{\partial f}{\partial z}} \quad \text{if and only if} \tag{40}$$

$$\nabla \cdot A(z)\,\nabla u = 0 \quad \text{and} \quad \nabla \cdot A^*(z)\,\nabla v = 0, \qquad \text{where } A^*(z) = *A(z)^{-1}* = \frac{1}{\det A}A.$$

In case $A(z) = \sigma(z)\mathbf{I}$ is isotropic ($\mu = 0$) we see that

$$\frac{\partial f}{\partial \bar{z}} = \frac{1\,-\,\sigma}{1\,+\,\sigma} \quad \Leftrightarrow \quad \nabla \cdot \sigma\,\nabla u = 0 \quad \text{and} \quad \nabla \cdot \frac{1}{\sigma}\,\nabla v = 0.$$

From these identities we obtain the complex geometric optics solutions also for the conductivity equation (1).

Corollary 3.2. *Suppose $A(z)$ is uniformly elliptic, so that (11) holds with $K \in L^\infty(\mathbb{D})$. Assume also that $A(z) = \mathbf{I}$ for $|z| \geq 1$.*

Then the equation $\nabla \cdot A(z)\nabla u(z) = 0$ admits a unique weak solution $u = u_\xi \in W^{1,2}_{loc}(\mathbb{C})$ such that

$$u(z,\xi) = e^{i\xi z}\left(1 + \mathcal{O}\left(\frac{1}{|z|}\right)\right) \text{ as } |z| \to \infty \tag{41}$$

Proof. For existence, in view of (40) the function

$$u(z,\xi) = \text{Re } f_{\mu,\nu} + i \text{ Im } f_{\mu,-\nu} \tag{42}$$

is precisely what we are looking for.

When it comes to uniqueness, if $u \in W^{1,2}_{loc}$ is any function satisfying the divergence equation $\nabla \cdot A(z)\nabla u(z) = 0$ with (41), then using Theorem 2.2 for the real and imaginary parts of u, we can write it as

$$u = \text{Re } f_+ + i\text{Im } f_- = \frac{1}{2}(f_+ + f_- + \overline{f_+} - \overline{f_-}),$$

where $f_\pm$ are quasiregular mappings with

$$\frac{\partial f_\pm}{\partial \bar{z}} = \mu(z)\,\frac{\partial f_\pm}{\partial z} \pm \nu(z)\overline{\frac{\partial f_\pm}{\partial z}}$$

and where μ, ν are given by (17). Given the asymptotics (41), it is not hard to see that both f_+ and f_- satisfy (33) with (34). Therefore $f_+ = f_{\mu,\nu}$ and $f_- = f_{\mu,-\nu}$. $\square$

The exponentially growing solutions of Corollary 3.2 can be considered σ-harmonic counterparts of the usual exponential functions $e^{i\xi z}$. They are the building blocks of the *nonlinear Fourier transform* to be discussed in more detail in Section 6.

4. The Hilbert Transform $\mathcal{H}_\sigma$

Until further notice *we will now assume* that A is isotropic,

$$A(z) = \sigma(z)\mathbf{I}, \quad \sigma(z) \text{ scalar}, \quad \text{and we let } \mu = \frac{1-\sigma(z)}{1+\sigma(z)}.$$

Assume that $u \in W^{1,2}(\mathbb{D})$ is a weak solution to $\nabla \cdot \sigma(z)\nabla u(z) = 0$. Then, by Theorem 2.2, u admits a conjugate function $v \in W^{1,2}(\mathbb{D})$ such that

$$\begin{aligned} \partial_x v &= -\sigma \partial_y u \\ \partial_y v &= \sigma \partial_x u \end{aligned}$$

Let us now elaborate on the relationship between u and v. Since the function v is defined only up to a constant, we will normalize it by assuming

$$\int_{\partial\mathbb{D}} v\, ds = 0 \tag{43}$$

This way we obtain a unique map $\mathcal{H}_\mu : W^{1/2,2}(\partial\mathbb{D}) \to W^{1/2,2}(\partial\mathbb{D})$ by setting

$$\mathcal{H}_\mu : u\big|_{\partial\mathbb{D}} \mapsto v\big|_{\partial\mathbb{D}} \tag{44}$$

In other words, $v = \mathcal{H}_\mu(u)$ if and only if $\int_{\partial\mathbb{D}} v\,ds = 0$, and $u + iv$ has a $W^{1,2}$-extension f to the disk $\mathbb{D}$ satisfying $f_{\bar{z}} = \mu\overline{f_z}$. We call $\mathcal{H}_\mu$ the *Hilbert transform* corresponding to (35).

Since the function $g = -if = v - iu$ satisfies $g_{\bar{z}} = -\mu\overline{g_z}$, we have

$$\mathcal{H}_\mu \circ \mathcal{H}_{-\mu} u = \mathcal{H}_{-\mu} \circ \mathcal{H}_\mu u = -u + \frac{1}{2\pi}\int_{\partial\mathbb{D}} u\,ds \tag{45}$$

So far we have defined $\mathcal{H}_\mu(u)$ only for real-valued functions u. By setting

$$\mathcal{H}_\mu(iu) = i\mathcal{H}_{-\mu}(u),$$

we extend the definition of $\mathcal{H}_\mu(\cdot)$ to all $\mathbb{C}$-valued functions in $W^{1/2,2}(\partial\mathbb{D})$. Note, however, that $\mathcal{H}_\mu$ still remains only $\mathbb{R}$-linear.

As in the case of analytic functions, the Hilbert transform defines a projection, now on the "μ-analytic" functions. That is, we define $Q_\mu : W^{1/2,2}(\partial\mathbb{D}) \to W^{1/2,2}(\partial\mathbb{D})$ by

$$Q_\mu(g) = \frac{1}{2}\left(g - i\mathcal{H}_\mu g\right) + \frac{1}{4\pi}\int_{\partial\mathbb{D}} g\,ds \tag{46}$$

Then it follows that $Q_\mu^2 = Q_\mu$. Furthermore, we have the following lemma.

Lemma 4.1. *If $g \in W^{1/2,2}(\partial\mathbb{D})$, the following conditions are equivalent:*

(a) $g = f\big|_{\partial\mathbb{D}}$, where $f \in W^{1,2}(\mathbb{D})$ satisfies $f_{\bar{z}} = \mu\overline{f_z}$

(b) $Q_\mu(g)$ is a constant

Proof. Condition (a) holds if and only if $g = u + i\mathcal{H}_\mu u + ic$ for some real-valued $u \in W^{1/2,2}(\partial\mathbb{D})$ and real constant c. If g has this representation, then $Q_\mu(g) = \frac{1}{4\pi}\int_{\partial\mathbb{D}} u ds + ic$. On the other hand, if $Q_\mu(g)$ is a constant, then we put $g = u + iw$ into (46) and use (45) to show that $w = \mathcal{H}_\mu u + \text{constant}$. This shows that (a) holds. □

The Dirichlet-to-Neumann map (4) and the Hilbert transform (44) are closely related, as the next lemma shows.

Theorem 4.2. *Choose the counterclockwise orientation for $\partial\mathbb{D}$ and denote by ∂_T the tangential (distributional) derivative on $\partial\mathbb{D}$ corresponding to this orientation. We then have*

$$\partial_T \mathcal{H}_\mu(u) = \Lambda_\sigma(u) \tag{47}$$

In particular, the Dirichlet-to-Neumann map Λ_σ uniquely determines $\mathcal{H}_\mu$, $\mathcal{H}_{-\mu}$ and $\Lambda_{1/\sigma}$.

Proof. By the definition of Λ_σ we have

$$\int_{\partial\mathbb{D}} \varphi\Lambda_\sigma u\,ds = \int_{\mathbb{D}} \nabla\varphi \cdot \sigma\nabla u\,, \qquad \varphi \in C^\infty(\overline{\mathbb{D}})$$

Thus, by (6) and (7) and integration by parts, we get

$$\int_{\partial\mathbb{D}} \varphi\Lambda_\sigma u\,ds = \int_{\mathbb{D}} (\partial_x\varphi\,\partial_y v - \partial_y\varphi\,\partial_x v) = -\int_{\partial\mathbb{D}} v\,\partial_T\varphi\,ds,$$

and (47) follows. Next,

$$-\mu = (1 - 1/\sigma)/(1 + 1/\sigma),$$

and so $\Lambda_{1/\sigma}(u) = \partial_T \mathcal{H}_{-\mu}(u)$. Since by (45) $\mathcal{H}_\mu$ uniquely determines $\mathcal{H}_{-\mu}$, the proof is complete. $\square$

With these identities we can now show that, for the points z that lie outside $\mathbb{D}$, the values of the complex geometric optics solutions $f_\mu(z,\xi)$ and $f_{-\mu}(z,\xi)$ are determined by the Dirichlet-to-Neumann operator Λ_σ.

Theorem 4.3. *Let σ and $\widetilde{\sigma}$ be two conductivities satisfying the assumptions of Theorem 1.1 and assume $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$. Then if μ and $\widetilde{\mu}$ are the corresponding Beltrami coefficients, we have*

$$f_\mu(z,\xi) = f_{\widetilde{\mu}}(z,\xi) \quad \text{and} \quad f_{-\mu}(z,\xi) = f_{-\widetilde{\mu}}(z,\xi) \tag{48}$$

for all $z \in \mathbb{C} \setminus \overline{\mathbb{D}}$ and $\xi \in \mathbb{C}$.

Proof. By Theorem 4.2 the condition $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$ implies that $\mathcal{H}_\mu = \mathcal{H}_{\widetilde{\mu}}$. In the same way Λ_σ determines $\Lambda_{\sigma^{-1}}$, and so it is enough to prove the first claim of (48).

Fix the value of the parameter $\xi \in \mathbb{C}$. From (46) we see that the projections $Q_\mu = Q_{\widetilde{\mu}}$, and thus by Lemma 4.1

$$Q_\mu(\widetilde{f}) = Q_{\widetilde{\mu}}(\widetilde{f}) \quad \text{is constant}$$

Here we have written

$$\widetilde{f} = (f_{\widetilde{\mu}})\big|_{\partial\mathbb{D}}$$

Using Lemma 4.1 again, we see that there exists a function $G \in W^{1,2}(\mathbb{D})$ such that $G_{\overline{z}} = \mu\overline{G_z}$ in $\mathbb{D}$ and

$$G\big|_{\partial\mathbb{D}} = \widetilde{f}$$

We then define $G(z) = f_{\widetilde{\mu}}(z,\xi)$ for z outside $\mathbb{D}$. Now $G \in W^{1,2}_{loc}(\mathbb{C})$, and it satisfies $G_{\overline{z}} = \mu\overline{G_z}$ in the whole plane. Thus it is quasiregular, and so $G \in W^{1,p}_{loc}(\mathbb{C})$ for all $2 \le p < 2 + 1/k$, $k = \|\mu\|_\infty$. But now G is a solution to (33) and (34). By the uniqueness part of Theorem 3.1, we obtain $G(z) \equiv f_\mu(z,\xi)$. $\square$

Similarly, the Dirichlet-to-Neumann operator determines the complex geometric optics solutions to the conductivity equation at every point z outside the disk $\mathbb{D}$.

Corollary 4.4. *Let σ and $\widetilde{\sigma}$ be two conductivities satisfying the assumptions of Theorem 1.1 and assume $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$.*

Then

$$u_\sigma(z,\xi) = u_{\widetilde{\sigma}}(z,\xi) \quad \text{for all} \ \ z \in \mathbb{C} \setminus \overline{\mathbb{D}} \ \ \text{and} \ \ \xi \in \mathbb{C}$$

Proof. The claim follows immediately from the previous theorem and the representation $u_\sigma(z,\xi) = \mathrm{Re}\ f_\mu(z,\xi) + i\mathrm{Im}\ f_{-\mu}(z,\xi)$. $\square$

5. Dependence on Parameters

Our strategy will be to extend the identities $f_\mu(z,\xi) = f_{\widetilde{\mu}}(z,\xi)$ and $u_\sigma(z,\xi) = u_{\widetilde{\sigma}}(z,\xi)$ from outside the disk to points z inside $\mathbb{D}$. Once we do that, Theorem 1.1 follows via the equation $f_{\overline{z}} = \mu\overline{f_z}$.

For this purpose we need to understand the ξ-dependence in $f_\mu(z,\xi)$ and the quantities controlling it. In particular, we will derive equations relating the solutions and their derivatives with respect to the ξ-variable. For this purpose we prove the following theorem.

Theorem 5.1. *The complex geometric optics solutions $u_\sigma(z,\xi)$ and $f_\mu(z,\xi)$ are (Hölder)-continuous in z and C^∞-smooth in the parameter ξ.*

The continuity in the z-variable is of course clear since f_μ is a quasiregular function of z. However, for analyzing the ξ-dependence we need to realize the solutions in a different manner, by identities involving linear operators that depend smoothly on the variable ξ.

Let $f_\mu(z,\xi) = e^{i\xi z} M_\mu(z,\xi)$ and $f_{-\mu}(z,\xi) = e^{i\xi z} M_{-\mu}(z,\xi)$ be the solutions of Theorem 3.1 corresponding to conductivities σ and σ^{-1}, respectively. We can write (8), (33) and (34) in the form

$$\frac{\partial}{\partial \bar z} M_\mu = \mu(z)\, \overline{\frac{\partial}{\partial z}\,(e_\xi M_\mu)}, \qquad M_\mu - 1 \in W^{1,p}(\mathbb{C}) \tag{49}$$

when $2 < p < 1 + 1/k$. By taking the Cauchy transform and introducing a $\mathbb{R}$-linear operator L_μ,

$$L_\mu g = \mathcal{C}\left(\mu \frac{\partial}{\partial \bar z}\left(e_{-\xi}\, \overline{g}\right)\right), \tag{50}$$

we see that (49) is equivalent to

$$(\mathbf{I} - L_\mu) M_\mu = 1 \tag{51}$$

Theorem 5.2. *Assume that $\xi \in \mathbb{C}$ and $\mu \in L^\infty(\mathbb{C})$ is compactly supported with $\|\mu\|_\infty \le k < 1$. Then for $2 < p < 1 + 1/k$ the operator*

$$\mathbf{I} - L_\mu : W^{1,p}(\mathbb{C}) \oplus \mathbb{C} \to W^{1,p}(\mathbb{C}) \oplus \mathbb{C}$$

is bounded and invertible.

Here we denote by $W^{1,p}(\mathbb{C}) \oplus \mathbb{C}$ the Banach space consisting of functions of the form $f = \text{constant} + f_0$, where $f_0 \in W^{1,p}(\mathbb{C})$.

Proof. We write $L_\mu(g)$ as

$$L_\mu(g) = \mathcal{C}\left(\mu\, e_{-\xi}\, \overline{g_z} - i\, \overline{\xi}\, \mu\, e_{-\xi}\, \overline{g}\right) \tag{52}$$

Then Theorem B.2 shows that

$$L_\mu : W^{1,p}(\mathbb{C}) \oplus \mathbb{C} \to W^{1,p}(\mathbb{C}) \tag{53}$$

is bounded. Thus we need only establish invertibility.

To this end let us assume $h \in W^{1,p}(\mathbb{C})$. Consider the equation

$$(\mathbf{I} - L_\mu)(g + C_0) = h + C_1, \tag{54}$$

where $g \in W^{1,p}(\mathbb{C})$ and C_0, C_1 are constants. Then

$$C_0 - C_1 = g - h - L_\mu(g + C_0),$$

which by (53) gives $C_0 = C_1$. By differentiating and rearranging we see that (54) is equivalent to $g_{\bar z} - \mu(e_{-\xi}\, \overline{g})_{\bar z} = h_{\bar z} + \mu(\overline{C}_0 e_{-\xi})_{\bar z}$, or in other words, to

$$g_{\bar z} - (\mathbf{I} - \mu e_{-\xi} \overline{\mathcal{S}})^{-1}\big(\mu (e_{-\xi})_{\bar z} \overline{g}\big) = (\mathbf{I} - \mu e_{-\xi} \overline{\mathcal{S}})^{-1}\big(h_{\bar z} + \mu(\overline{C}_0 e_{-\xi})_{\bar z}\big) \tag{55}$$

We are now faced with the operator R defined by

$$R(g) = \mathcal{C}\left(\mathbf{I} - \nu \overline{\mathcal{S}}\,\right)^{-1} (\alpha \overline{g}),$$

where $\nu(z) = \mu e_{-\xi}$ satisfies $|\nu(z)| \leq k\chi_{\mathbb{D}}(z)$ and α is defined by $\alpha = \mu(e_{-\xi})_{\overline{z}} = -i\overline{\xi}\,\mu\, e_{-\xi}$. According to Theorem B.4, $\mathbf{I} - \nu\overline{\mathcal{S}}$ is invertible in $L^p(\mathbb{C})$ when $1 + k < p < 1 + 1/k$, while the Cauchy transform requires $p > 2$. Therefore R is a well-defined and bounded operator on $L^p(\mathbb{C})$ for $2 < p < 1 + 1/k$.

Moreover, the right hand side of (55) belongs to $L^p(\mathbb{C})$ for each $h \in W^{1,p}(\mathbb{C})$. Hence this equation admits a unique solution $g \in W^{1,p}(\mathbb{C})$ if and only if the operator $\mathbf{I} - R$ is invertible in $L^p(\mathbb{C}),\ 2 < p < 1 + 1/k$.

To get this we will use Fredholm theory. First, Theorem B.3 shows that R is a compact operator on $L^p(\mathbb{C})$ when $2 < p < 1 + 1/k$. Therefore it suffices to show that $\mathbf{I} - R$ is injective. Suppose now that $g \in L^p(\mathbb{C})$ satisfies

$$g = Rg = \mathcal{C}\left(\mathbf{I} - \nu\overline{\mathcal{S}}\,\right)^{-1}(\alpha\overline{g})$$

Then $g \in W^{1,p}(\mathbb{C})$ by Theorem B.2 and $g_{\overline{z}} = \left(\mathbf{I} - \nu\overline{\mathcal{S}}\,\right)^{-1}(\alpha\overline{g})$. Equivalently

$$g_{\overline{z}} - \nu\overline{g_z} = \alpha\overline{g} \tag{56}$$

Thus the assumptions of Theorem B.8 are fulfilled, and we must have $g \equiv 0$. Therefore $\mathbf{I} - R$ is indeed injective on $L^p(\mathbb{C})$. As a Fredholm operator, it therefore is invertible in $L^p(\mathbb{C})$. Therefore the operator $\mathbf{I} - L_\mu$ is invertible in $W^{1,p}(\mathbb{C}),\ 2 < p < 1 + 1/k$. $\square$

A glance at (50) shows that $\xi \to L_\mu$ is an infinitely differentiable family of operators. Therefore, with Theorem 5.2, we see that $M_\mu = (\mathbf{I} - L_\mu)^{-1}1$ is C^∞-smooth in the parameter ξ. Thus we have obtained Theorem 5.1.

6. Nonlinear Fourier Transform

The idea of studying the $\overline{\xi}$-dependence of operators associated with complex geometric optics solutions was introduced by Beals and Coifman [10] in connection with the inverse scattering approach to KdV-equations. Here we will apply this method to the solutions u_σ to the conductivity equation (1) and show that they satisfy a simple $\overline{\partial}$-equation with respect to the parameter ξ.

We start with the representation $u_\sigma(z,\xi) = \mathrm{Re}\ f_\mu(z,\xi) + i\,\mathrm{Im}\ f_{-\mu}(z,\xi)$, where $f_{\pm\mu}$ are the solutions to the corresponding Beltrami equations; in particular, they are analytic outside the unit disk. Hence with the asymptotics (34) they admit the following power series development,

$$f_{\pm\mu}(z,\xi) = e^{i\xi z}\Big(1 + \sum_{n=1}^{\infty} b_n^{\pm}(\xi) z^{-n}\Big), \qquad |z| > 1, \tag{57}$$

where $b_n^+(\xi)$ and $b_n^-(\xi)$ are the coefficients of the series, depending on the parameter ξ. For the solutions to the conductivity equation, this gives

$$u_\sigma(z,\xi) = e^{i\xi z} + \frac{a(\xi)}{z}\, e^{i\xi z} + \frac{b(\xi)}{\bar{z}}\, e^{-i\bar{\xi}\bar{z}} + e^{i\xi z}\,\mathcal{O}(\frac{1}{|z|^2})$$

as $z \to \infty$, where

$$a(\xi) = \frac{b_1^+(\xi) + b_1^-(\xi)}{2}, \qquad b(\xi) = \frac{\overline{b_1^+(\xi)} - \overline{b_1^-(\xi)}}{2\,\bar{z}}. \tag{58}$$

Fixing the z-variable, we take the $\partial_{\bar{\xi}}$-derivative of $u_\sigma(z,\xi)$ and get

$$\partial_{\bar{\xi}}\, u_\sigma(z,\xi) = -i\tau_\sigma(\xi)\; e^{-i\bar{\xi}\bar{z}}\; \big(1+\mathcal{O}(\frac{1}{|z|})\,\big), \tag{59}$$

where the coefficient

$$\tau_\sigma(\xi) := \overline{b(\xi)} \tag{60}$$

However, the derivative $\partial_{\bar{\xi}}\, u_\sigma(z,\xi)$ is another solution to the conductivity equation! From the uniqueness of the complex geometric optics solutions under the given exponential asymptotics, Corollary 3.2, we therefore have the simple but important relation

$$\partial_{\bar{\xi}}\, u_\sigma(z,\xi) \;=\; -i\,\tau_\sigma(\xi)\,\overline{u_\sigma(z,\xi)} \quad \text{for all} \;\; \xi, z \in \mathbb{C} \tag{61}$$

The remarkable feature of this relation is that the coefficient τ_σ does not depend on the space variable z. Later, this phenomenon will become of crucial importance in the solution to the Calderón problem.

In analogy with the one-dimensional scattering theory of integrable systems and associated inverse problems (see [10, 26, 27]), we call τ_σ the **nonlinear Fourier transform** of σ.

To understand the basic properties of the nonlinear Fourier transform, we need to return to the Beltrami equation. We will first show that the Dirichlet-to-Neumann data determines τ_σ. This is straightforward. Then the later sections are devoted to showing that the nonlinear Fourier transform τ_σ determines the coefficient σ almost everywhere. There does not seem to be any direct method for this, rather we will have to show that from τ_σ we can determine the exponentially growing solutions $f_{\pm\mu}$ defined in the entire plane. From this information the coefficient μ, and hence σ, can be found.

In any case it seems that most properties of τ_σ are important and interesting in and of themselves. We have the usual transformation rules under scaling and and translation,

$$\sigma_1(z) = \sigma(Rz) \quad \Rightarrow \quad \tau_{\sigma_1}(\xi) = \frac{1}{R}\tau_\sigma(\xi/R),$$

$$\sigma_2(z) = \sigma(z+p) \quad \Rightarrow \quad \tau_{\sigma_2}(\xi) = e^{i(p\xi+\bar{p}\bar{\xi})}\tau_\sigma(\xi).$$

but not much is known concerning questions such as the possibility of a Plancherel formula.

However, in the first instance some simple bounds can be achieved. We will show that for σ as above, $\tau_\sigma \in L^\infty$. For this we need the following result, which is useful also elsewhere.

Here let $f_\mu(z,\xi) = e^{i\xi z}M_\mu(z,\xi)$ and $f_{-\mu}(z,\xi) = e^{i\xi z}M_{-\mu}(z,\xi)$ be the solutions of Theorem 3.1 corresponding to conductivities σ and σ^{-1}, respectively, holomorphic outside $\mathbb{D}$.

Theorem 6.1. *For every $\xi, z \in \mathbb{C}$ we have $M_{\pm\mu}(z,\xi) \neq 0$. Moreover,*

$$Re\left(\frac{M_\mu(z,\xi)}{M_{-\mu}(z,\xi)}\right) > 0 \tag{62}$$

Proof. First, note that (8) implies, for $M_{\pm\mu}$,

$$\frac{\partial}{\partial \bar{z}} M_{\pm\mu} \mp \mu e_{-\xi} \overline{\frac{\partial}{\partial z} M_{\pm\mu}} = \mp i\bar{\xi}\mu e_{-\xi}\overline{M_{\pm\mu}} \tag{63}$$

Thus we may apply Theorem B.8 to get

$$M_{\pm\mu}(z) = \exp(\eta_{\pm}(z)) \neq 0, \tag{64}$$

and consequently $M_\mu/M_{-\mu}$ is well defined. Second, if (62) is not true, the continuity of $M_{\pm\mu}$ and the fact $\lim_{z\to\infty} M_{\pm\mu}(z,\xi) = 1$ imply the existence of $z_0 \in \mathbb{C}$ such that

$$M_\mu(z_0, \xi) = it\, M_{-\mu}(z_0, \xi)$$

for some $t \in \mathbb{R} \setminus \{0\}$ and $\xi \in \mathbb{C}$. But then, $g = M_\mu - itM_{-\mu}$ satisfies

$$\begin{aligned} \frac{\partial}{\partial \bar{z}} g &= \mu(z)\overline{\frac{\partial}{\partial z} (e_\xi g)}, \\ g(z) &= 1 - it + \mathcal{O}\left(\frac{1}{z}\right), \quad \text{as } z \to \infty \end{aligned}$$

According to Theorem B.8, this implies

$$g(z) = (1 - it) \exp(\eta(z)) \neq 0,$$

contradicting the assumption $g(z_0) = 0$. □

The boundedness of the nonlinear Fourier transform is now a simple corollary of Schwarz's lemma.

Theorem 6.2. *The functions $f_{\pm\mu}(z,\xi) = e^{i\xi z} M_{\pm\mu}(z,\xi)$ satisfy, for $|z| > 1$ and for all $\xi \in \mathbb{C}$,*

$$\left| \frac{M_\mu(z,\xi) - M_{-\mu}(z,\xi)}{M_\mu(z,\xi) + M_{-\mu}(z,\xi)} \right| \leq \frac{1}{|z|} \tag{65}$$

Moreover, for the nonlinear Fourier transform τ_σ, we have

$$|\tau_\sigma(\xi)| \leq 1 \quad \text{for all } \xi \in \mathbb{C} \tag{66}$$

Proof. Fix the parameter $\xi \in \mathbb{C}$ and denote

$$m(z) = \frac{M_\mu(z,\xi) - M_{-\mu}(z,\xi)}{M_\mu(z,\xi) + M_{-\mu}(z,\xi)}$$

Then by Theorem 6.1, $|m(z)| < 1$ for all $z \in \mathbb{C}$. Moreover, m is holomorphic for $z \in \mathbb{C} \setminus \overline{\mathbb{D}}$, $m(\infty) = 0$, and thus by Schwarz's lemma we have $|m(z)| \leq 1/|z|$ for all $z \in \mathbb{C} \setminus \overline{\mathbb{D}}$.

On the other hand, from the development (57),

$$M_\mu(z,\xi) = 1 + \sum_{n=1}^{\infty} b_n(\xi) z^{-n} \quad \text{for } |z| > 1,$$

and similarly for $M_{-\mu}(z,\xi)$. We see that

$$\tau_\sigma(\xi) = \frac{1}{2}\left(\overline{b_1^+(\xi)} - \overline{b_1^-(\xi)}\right) = \lim_{z\to\infty} \overline{z\, m(z)}$$

Therefore the second claim also follows. □

With these results the Calderón problem reduces to the question whether we can invert the nonlinear Fourier transform.

Theorem 6.3. *The operator Λ_σ uniquely determines the nonlinear Fourier transform τ_σ.*

Proof. The claim follows immediately from Theorem 4.3, from the development (57) and from the definition (60) of τ_σ. □

From the relations $-\mu = (1-1/\sigma)/(1+1/\sigma)$, we have the symmetry

$$\tau_\sigma(\xi) = -\tau_{1/\sigma}(\xi)$$

It follows that the functions

$$u_1 = \mathrm{Re}\, f_\mu + i\,\mathrm{Im}\, f_{-\mu} = u_\sigma \quad \text{and} \quad u_2 = i\,\mathrm{Re}\, f_{-\mu} - \mathrm{Im}\, f_\mu = iu_{1/\sigma} \tag{67}$$

form a "primary pair" of complex geometric optics solutions.

Corollary 6.4. *The functions $u_1 = u_\sigma$ and $u_2 = iu_{1/\sigma}$ are complex-valued $W^{1,2}_{loc}(\mathbb{C})$-solutions to the conductivity equations*

$$\nabla\cdot\sigma\nabla u_1 = 0 \quad \textit{and} \quad \nabla\cdot\frac{1}{\sigma}\nabla u_2 = 0, \tag{68}$$

respectively. In the ξ-variable they are solutions to the same $\partial_{\overline{\xi}}$-equation,

$$\frac{\partial}{\partial\overline{\xi}}\, u_j(z,\xi) = -i\,\tau_\sigma(\xi)\,\overline{u_j}(z,\xi), \qquad j = 1,2, \tag{69}$$

and their asymptotics, as $|z|\to\infty$, are

$$u_\sigma(z,\xi) = e^{i\xi z}\left(1 + \mathcal{O}\left(\frac{1}{|z|}\right)\right), \qquad u_{1/\sigma}(z,\xi) = e^{i\xi z}\left(i + \mathcal{O}\left(\frac{1}{|z|}\right)\right)$$

7. Subexponential Growth

A basic obstacle in the solution to Calderón's problem is to find methods to control the asymptotic behavior in the parameter ξ for complex geometric optics solutions. If we knew that the assumptions of the Liouville type Theorem B.8 were valid in (69), then the equation, hence the Dirichlet-to-Neumann map, would uniquely determine $u_\sigma(z,\xi)$ with $u_{1/\sigma}(z,\xi)$, and we would be done. However, we only know from Theorem 6.2 that $\tau_\sigma(\xi)$ is bounded in ξ. It takes considerably more effort to prove the counterpart of the Riemann-Lebesgue lemma, that

$$\tau_\sigma(\xi)\to 0, \quad \text{as } \xi\to\infty.$$

Indeed, this will be one of the consequences of the results in the present section.

It is clear that some control of the parameter ξ is needed for $u_\sigma(z,\xi)$. Within the category of conductivity equations with L^∞-coefficients σ, the complex analytic and quasiconformal methods provide by far the most powerful methods. Therefore we return to the Beltrami equation. The purpose of this section is to study the ξ-behavior in the functions $f_\mu(z,\xi) = e^{i\xi z}M_\mu(z,\xi)$ and to show that for a fixed z, $M_\mu(z,\xi)$ grows at most subexponentially in ξ as $\xi\to\infty$. Subsequently, the result will be applied to $u_j(z,\xi)$.

For some later purposes we will also need to generalize the situation a bit by considering complex Beltrami coefficients μ_λ of the form $\mu_\lambda = \lambda\mu$, where the constant $\lambda \in \partial\mathbb{D}$ and μ is as before. Exactly as in Theorem 3.1, we can show the existence and uniqueness of $f_{\lambda\mu} \in W^{1,p}_{loc}(\mathbb{C})$ satisfying

$$\frac{\partial}{\partial \bar z} f_{\lambda\mu} = \lambda\mu \overline{\frac{\partial}{\partial z} f_{\lambda\mu}} \quad \text{and} \tag{70}$$

$$f_{\lambda\mu}(z,\xi) = e^{i\xi z}\left(1 + \mathcal{O}\left(\frac{1}{z}\right)\right) \quad \text{as } |z| \to \infty \tag{71}$$

In fact, we have that the function $f_{\lambda\mu}$ admits a representation of the form

$$f_{\lambda\mu}(z,\xi) = e^{i\,\xi\,\varphi_\lambda(z,\xi)}, \tag{72}$$

where for each fixed $\xi \in \mathbb{C} \setminus \{0\}$ and $\lambda \in \partial\mathbb{D}$, $\varphi_\lambda(z,\xi) = z + \mathcal{O}\left(\frac{1}{z}\right)$ for $z \to \infty$. The principal solution $\varphi = \varphi_\lambda(z,\xi)$ satisfies the nonlinear equation

$$\frac{\partial}{\partial \bar z}\varphi(z) = \kappa_{\lambda,\xi}\, e_{-\xi}\big(\varphi(z)\big)\,\mu(z)\,\overline{\frac{\partial}{\partial z}\varphi(z)} \tag{73}$$

where $\kappa = \kappa_{\lambda,\xi} = -\lambda\,\overline{\xi}^{\,2}|\xi|^{-2}$ is constant with $|\kappa_{\lambda,\xi}| = 1$.

The main goal of this section is to show the following theorem.

Theorem 7.1. *If $\varphi = \varphi_\lambda$ and $f_{\lambda\mu}$ are as in* (70)–(73)*, then*

$$\varphi_\lambda(z,\xi) \to z$$

uniformly in $z \in \mathbb{C}$ and $\lambda \in \partial\mathbb{D}$ as $\xi \to \infty$.

From the theorem we have the immediate consequence,

Corollary 7.2. *If $\sigma, \sigma^{-1} \in L^\infty(\mathbb{C})$ with $\sigma(z) = 1$ outside a compact set, then*

$$\tau_\sigma(\xi) \to 0, \qquad \text{as } \xi \to \infty.$$

Proof of Corollary 7.2. Let $\lambda = 1$. The principal solutions in (72) have the development

$$\varphi(z,\xi) = z + \sum_{n=1}^{\infty} \frac{c_n(\xi)}{z^n}, \qquad |z| > 1,$$

where by Cauchy integral formula and Theorem 7.1 we have

$$c_n(\xi) \to 0, \qquad \xi \to \infty,\ n \in \mathbb{N}.$$

Comparing now with (57)-(60) proves the claim. $\square$

It remains to prove Theorem 7.1, which will the rest of this section. We shall split the proof up into several lemmas.

Lemma 7.3. *Suppose $\varepsilon > 0$ is given. Suppose also that for $\mu_\lambda(z) = \lambda\mu(z)$, we have*

$$f_n = \mu_\lambda S_n \mu_\lambda S_{n-1} \mu_\lambda \cdots \mu_\lambda S_1 \mu_\lambda, \tag{74}$$

where $S_j : L^2(\mathbb{C}) \to L^2(\mathbb{C})$ are Fourier multiplier operators, each with a unimodular symbol. Then there is a number $R_n = R_n(k, \varepsilon)$ depending only on $k = \|\mu\|_\infty$, n and ε such that

$$|\widehat{f}_n(\eta)| < \varepsilon \quad \text{for } |\eta| > R_n \tag{75}$$

Proof. It is enough to prove the claim for $\lambda = 1$. By assumption,

$$\widehat{S_j g}(\eta) = m_j(\eta)\widehat{g}(\eta),$$

where $|m_j(\eta)| = 1$ for $\eta \in \mathbb{C}$. We have by (74),

$$\|f_n\|_{L^2} \leq \|\mu\|^n_{L^\infty} \|\mu\|_{L^2} \leq \sqrt{\pi} k^{n+1} \tag{76}$$

since $\operatorname{supp}(\mu) \subset \mathbb{D}$. Choose ρ_n so that

$$\int_{|\eta|>\rho_n} |\widehat{\mu}(\eta)|^2 \, d\eta < \varepsilon^2 \tag{77}$$

After this, choose $\rho_{n-1}, \rho_{n-2}, \ldots, \rho_1$ inductively so that for $l = n-1, \ldots, 1$,

$$\pi \int_{|\eta|>\rho_l} |\widehat{\mu}(\eta)|^2 \, d\eta \leq \varepsilon^2 \left(\prod_{j=l+1}^{n} \pi \rho_j \right)^{-2} \tag{78}$$

Finally, choose ρ_0 so that

$$|\widehat{\mu}(\eta)| < \varepsilon \pi^{-n} \left(\prod_{j=1}^{n} \rho_j \right)^{-1} \qquad \text{when} \quad |\eta| > \rho_0 \tag{79}$$

All these choices are possible since $\mu \in L^1 \cap L^2$.

Now, we set $R_n = \sum_{j=0}^n \rho_j$ and claim that (75) holds for this choice of R_n. Hence assume that $|\eta| > \sum_{j=0}^n \rho_j$. We have

$$\begin{aligned} |\widehat{f}_n(\eta)| \ \leq \ & \int_{|\eta-\zeta|\leq\rho_n} |\widehat{\mu}(\eta-\zeta)| \, |\widehat{f}_{n-1}(\zeta)| \, d\zeta \\ & + \int_{|\eta-\zeta|\geq\rho_n} |\widehat{\mu}(\eta-\zeta)| \, |\widehat{f}_{n-1}(\zeta)| \, d\zeta \end{aligned} \tag{80}$$

But if $|\eta - \zeta| \leq \rho_n$, then $|\zeta| > \sum_{j=0}^{n-1} \rho_j$. Thus, if we denote

$$\Delta_n = \sup \left\{ |\widehat{f}_n(\eta)| \ : \ |\eta| > \sum_{j=0}^{n} \rho_j \right\},$$

it follows from (80) and (76) that

$$\begin{aligned}\Delta_n &\leq \Delta_{n-1}(\pi\rho_n^2)^{1/2}\|\mu\|_{L^2} + \left(\int_{|\zeta|\geq\rho_n} |\widehat{\mu}(\zeta)|^2\, d\zeta\right)^{1/2} \|\widehat{f}_{n-1}\|_{L^2} \\ &\leq \pi\rho_n k\,\Delta_{n-1} + k^n\left(\pi\int_{|\zeta|\geq\rho_n} |\widehat{\mu}(\zeta)|^2\, d\zeta\right)^{1/2}\end{aligned}$$

for $n \geq 2$. Moreover, the same argument shows that

$$\Delta_1 \leq \pi\rho_1\, k\, \sup\{|\widehat{\mu}(\eta)| \,:\, |\eta| > \rho_0\} + k\left(\pi\int_{|\zeta|>\rho_1} |\widehat{\mu}(\zeta)|^2\, d\zeta\right)^{1/2}$$

In conclusion, after iteration we will have

$$\begin{aligned}\Delta_n &\leq (k\pi)^n\left(\prod_{j=1}^n \rho_j\right)\sup\{|\widehat{\mu}(\eta)| \,:\, |\eta| > \rho_0\} \\ &\quad + k^n\sum_{l=1}^n\left(\prod_{j=l+1}^n \pi\rho_j\right)\left(\pi\int_{|\zeta|>\rho_l} |\widehat{\mu}(\zeta)|^2\, d\zeta\right)^{1/2}\end{aligned}$$

With the choices (77)–(79), this leads to

$$\Delta_n \leq (n+1)k^n\varepsilon \leq \frac{\varepsilon}{1-k},$$

which proves the claim. □

Our next goal is to use Lemma 7.3 to prove the asymptotic result required in Theorem 7.1 for the solution of a closely related linear equation.

Theorem 7.4. *Suppose $\psi \in W^{1,2}_{loc}(\mathbb{C})$ satisfies*

$$\frac{\partial\psi}{\partial\bar{z}} = \kappa\,\mu(z)\,e_{-\xi}(z)\,\frac{\partial\psi}{\partial z} \quad \textit{and} \tag{81}$$

$$\psi(z) = z + \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } z \to \infty, \tag{82}$$

where κ is a constant with $|\kappa| = 1$.

Then $\psi(z,\xi) \to z$, uniformly in $z \in \mathbb{C}$ and $\kappa \in \partial\mathbb{D}$, as $\xi \to \infty$.

To prove Theorem 7.4 we need some preparation. First, since the L^p-norm of the Beurling transform $\mathbf{S}_p \to 1$ when $p \to 2$, we can choose a $\delta_k > 0$ so that $k\mathbf{S}_p < 1$ whenever $2 - \delta_k \leq p \leq 2 + \delta_k$. With this notation we then have the following lemma.

Lemma 7.5. *Let $\psi = \psi(\cdot,\xi)$ be the solution of (81) and let $\varepsilon > 0$. Then $\psi_{\bar{z}}$ can be decomposed as $\psi_{\bar{z}} = g + h$, where*

(1) $\|h(\cdot,\xi)\|_{L^p} < \varepsilon$ *for* $2 - \delta_k \leq p \leq 2 + \delta_k$ *uniformly in* ξ.
(2) $\|g(\cdot,\xi)\|_{L^p} \leq C_0 = C_0(k)$ *uniformly in* ξ.
(3) $\widehat{g}(\eta,\xi) \to 0$ *as* $\xi \to \infty$.

In statement 3 convergence is uniform on compact subsets of the η-plane and also uniform in $\kappa \in \partial\mathbb{D}$. The Fourier transform is with respect to the first variable only.

Proof. We may solve (81) using a Neumann series, which will converge in L^p,

$$\frac{\partial \psi}{\partial \bar{z}} = \sum_{n=0}^{\infty} (\kappa\, \mu\, e_{-\xi}\, \mathcal{S})^n\, (\kappa\, \mu\, e_{-\xi})$$

Let

$$h = \sum_{n=n_0}^{\infty} (\kappa\, \mu e_{-\xi}\, \mathcal{S})^n\, (\kappa\, \mu\, e_{-\xi})$$

Then

$$\|h\|_{L^p} \le \pi^{1/p} \frac{k^{n_0+1} \mathbf{S}_p^{n_0}}{1 - k\mathbf{S}_p}$$

We obtain the first statement by choosing n_0 large enough.

The remaining part clearly satisfies the second statement with a constant C_0 that is independent of ξ and λ. To prove statement 3 we first note that

$$\mathcal{S}(e_{-\xi}\, \phi) = e_{-\xi}\, S_\xi\, \phi,$$

where $\widehat{(S_\xi \phi)}(\eta) = m(\eta - \xi)\, \widehat{\phi}(\eta)$ and $m(\eta) = \eta/\overline{\eta}$. Consequently,

$$(\mu e_{-\xi}\, \mathcal{S})^n \mu\, e_{-\xi} = e_{-(n+1)\xi}\, \mu S_{n\xi}\, \mu S_{(n-1)\xi} \cdots \mu S_\xi\, \mu,$$

and so

$$g = \sum_{j=1}^{n_0} \kappa^j\, e_{-j\xi}\, \mu S_{(j-1)\xi}\, \mu \cdots \mu S_\xi \mu$$

Therefore

$$g = \sum_{j=1}^{n_0} e_{-j\xi} G_j,$$

where by Lemma 7.3, $|\widehat{G}_j(\eta)| < \widetilde{\varepsilon}$ whenever $|\eta| > R = \max_{j \le n_0} R_j$. As $\widehat{(e_{j\xi} G_j)}(\eta) = \widehat{G}_j(\eta + j\xi)$, for any fixed compact set K_0, we can take ξ so large that $j\xi + K_0 \subset \mathbb{C} \setminus \mathbb{D}(0, R)$ for each $1 \le j \le n_0$. Then

$$\sup_{\eta \in K_0} |\widehat{g}(\eta, \xi)| \le n_0 \widetilde{\varepsilon}$$

This proves the lemma. □

Proof of Theorem 7.4. We show first that when $\xi \to \infty$, $\psi_{\bar{z}} \to 0$ weakly in L^p, $2 - \delta_k \le p \le 2 + \delta_k$. For this suppose that $f_0 \in L^q$, $q = p/(p-1)$, is fixed and choose $\varepsilon > 0$. Then there exists $f \in C_0^\infty(\mathbb{C})$ such that $\|f_0 - f\|_{L^q} < \varepsilon$, and so by Lemma 7.5,

$$|\langle f_0, \psi_{\bar{z}} \rangle| \le \varepsilon C_1 + \Big| \int \widehat{f}(\eta) \widehat{g}(\eta, \xi)\, d\eta \Big|,$$

First choose R so large that

$$\int_{\mathbb{C} \setminus \mathbb{D}(0,R)} |\widehat{f}(\eta)|^2\, d\eta \le \varepsilon^2$$

and then $|\xi|$ so large that $|\widehat{g}(\eta,\xi)| \le \varepsilon/(\sqrt{\pi}R)$ for all $\eta \in \mathbb{D}(R)$. Now,

$$\left|\int \widehat{f}(\eta)\widehat{g}(\eta,\xi)\,d\eta\right| \le \int_{\mathbb{D}(R)} \widehat{f}(\eta)\widehat{g}(\eta,\xi)\,d\eta + \int_{\mathbb{C}\setminus\mathbb{D}(R)} \widehat{f}(\eta)\widehat{g}(\eta,\xi)\,d\eta$$

$$\le \varepsilon(\|f\|_{L^2} + \|g\|_{L^2}) \le C_2(f)\varepsilon \tag{83}$$

The bound is the same for all κ, hence

$$\sup_{\kappa\in\partial\mathbb{D}} |\langle f_0, \psi_{\bar z}\rangle| \to 0 \tag{84}$$

as $|\xi| \to \infty$.

To prove the uniform convergence of ψ itself, we write

$$\psi(z,\xi) = z - \frac{1}{\pi}\int_{\mathbb{D}} \frac{1}{\zeta - z}\frac{\partial}{\partial\bar\zeta}\psi(\zeta,\xi) \tag{85}$$

Here note that $\mathrm{supp}(\psi_{\bar z}) \subset \mathbb{D}$ and $\chi_{\mathbb{D}}(\zeta)/(\zeta - z) \in L^q$ for all $q < 2$. Thus by the weak convergence we have

$$\psi(z,\xi) \to z \quad \text{as } \xi \to \infty \tag{86}$$

for each fixed $z \in \mathbb{C}$, but uniformly in $\kappa \in \partial\mathbb{D}$. On the other hand, as

$$\sup_{\xi}\left\|\frac{\partial\psi}{\partial\bar z}\right\|_{L^p} \le C_0 = C_0(p, \|\mu\|_\infty) < \infty$$

for all z sufficiently large, $|\psi(z,\xi) - z| < \varepsilon$, uniformly in $\xi \in \mathbb{C}$ and $\kappa \in \partial\mathbb{D}$. Moreover, (85) shows also that the family $\{\psi(\cdot,\xi) \,:\, \xi \in \mathbb{C},\ \kappa \in \partial\mathbb{D}\}$ is equicontinuous. Combining all these observations shows that the convergence in (86) is uniform in $z \in \mathbb{C}$ and $\kappa \in \partial\mathbb{D}$. $\square$

Finally, we proceed to the nonlinear case: Assume that φ_λ satisfies (70) and (72). Since φ is a (quasiconformal) homeomorphism, we may consider its inverse $\psi_\lambda : \mathbb{C} \to \mathbb{C}$,

$$\psi_\lambda \circ \varphi_\lambda(z) = z, \tag{87}$$

which also is quasiconformal. By differentiating (87) with respect to z and $\bar z$ we find that ψ satisfies

$$\frac{\partial}{\partial\bar z}\psi_\lambda = -\frac{\bar\xi}{\xi}\lambda\,(\mu\circ\psi_\lambda)\,e_{-\xi}\,\frac{\partial}{\partial z}\,\psi_\lambda \quad \text{and} \tag{88}$$

$$\psi_\lambda(z,\xi) = z + \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } z \to \infty \tag{89}$$

Proof of Theorem 7.1. It is enough to show that

$$\psi_\lambda(z,\xi) \to z \tag{90}$$

uniformly in z and λ as $\xi \to \infty$. For this we introduce the notation

$$\Sigma_k = \{g \in W^{1,2}_{loc}(\mathbb{C}) : g_{\bar z} = \nu\, g_z,\ |\nu| \le k\chi_{\mathbb{D}(2)} \quad \text{and} \quad g = z + \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } z \to \infty\} \tag{91}$$

Note that all mappings $g \in \Sigma_k$ are principal solutions and hence homeomorphisms.

The support of the coefficient $\mu \circ \psi_\lambda$ in (88) need no longer be contained in $\mathbb{D}$. However, by Koebe distortion theorem, see e.g. [3, p. 44], $\varphi_\lambda(\mathbb{D}) \subset \mathbb{D}(0,2)$ and thus $\operatorname{supp}(\mu\circ\psi_\lambda) \subset \mathbb{D}(0,2)$. Accordingly, $\psi_\lambda \in \Sigma_k$.

Since normalized quasiconformal mappings form a normal family, we see that the family Σ_k is compact in the topology of uniform convergence. Given sequences $\xi_n \to \infty$ and $\lambda_n \in \partial\mathbb{D}$, we may pass to a subsequence and assume that $\kappa_{\lambda_n,\xi_n} = -\lambda_n \,\overline{\xi_n}^{\,2} |\xi_n|^{-2} \to \kappa \in \partial\mathbb{D}$ and that the corresponding mapping $\psi_{\lambda_n}(\cdot,\xi_n) \to \psi_\infty$ uniformly, with $\psi_\infty \in \Sigma_k$. To prove Theorem 7.1 it is enough to show that for any such sequence $\psi_\infty(z) \equiv z$.

Hence we assume that there is such a limit function ψ_∞. We consider the $W^{1,2}_{loc}$-solution $\Phi(z) = \Phi_\lambda(z,\xi)$ of

$$\begin{aligned}\frac{\partial\Phi}{\partial\bar z} &= \kappa\,(\mu\circ\psi_\infty)\,e_{-\xi}\,\frac{\partial\Phi}{\partial\bar z}\\ \Phi(z) &= z + \mathcal{O}\left(\frac{1}{z}\right) \quad \text{as } z\to\infty\end{aligned}$$

This is now a linear Beltrami equation which, by Theorem 2.4 has a unique solution $\Phi \in \Sigma_k$ for each $\xi \in \mathbb{C}$ and $|\lambda| = 1$. According to Theorem 7.4,

$$\Phi_\lambda(z,\xi) \to z \quad \text{as } \xi\to\infty \tag{92}$$

Further, when $2 < p < 1 + 1/k$, by Lemma B.7,

$$\begin{aligned} &|\psi_{\lambda_n}(z,\xi_n) - \Phi_\lambda(z,\xi_n)|\\ &\quad= \frac{1}{\pi}\left|\int_{\mathbb{D}} \frac{1}{\zeta - z}\frac{\partial}{\partial\bar z}\big(\psi_{\lambda_n}(\zeta,\xi_n) - \Phi_\lambda(\zeta,\xi_n)\big)\,d\zeta\right|\\ &\quad\le C_1\left\|\frac{\partial}{\partial\bar z}\big(\psi_{\lambda_n}(\zeta,\xi_n) - \Phi_\lambda(\zeta,\xi_n)\big)\right\|_{L^p}\\ &\quad\le C_2\big|\kappa_{\lambda_n,\xi_n} - \kappa\big|\\ &\qquad + C_2\left(\int_{2\mathbb{D}}\big|\mu\left(\psi_{\lambda_n}(\zeta,\xi_n)\right) - \mu\left(\psi_\infty(\zeta)\right)\big|^{\frac{p(1+\varepsilon)}{\varepsilon}}\,d\zeta\right)^{\frac{\varepsilon}{p(1+\varepsilon)}} \end{aligned} \tag{93}$$

Finally, we apply our higher-integrability results, such as Theorem B.5. Thus for all $2 < p < 1 + 1/k$ and for all $g = \psi^{-1}$, $\psi \in \Sigma_k$, we have the estimate for the Jacobian $J(z,g)$:

$$\int_{\mathbb{D}} J(z,g)^{p/2} \le \int_{\mathbb{D}}\left|\frac{\partial g}{\partial z}\right|^p \le C(k) < \infty, \tag{94}$$

where $C(k)$ depends only on k. We use this estimate in the cases $\psi(z) = \psi_{\lambda_n}(z,\xi_n)$ and $\psi = \psi_\infty$. Namely, we have for each $\gamma \in C_0^\infty(\mathbb{D})$ that

$$\begin{aligned}\int_{2\mathbb{D}} |\mu(\psi) - \gamma(\psi)|^{\frac{p(1+\varepsilon)}{\varepsilon}} &= \int_{\mathbb{D}} |\mu-\gamma|^{\frac{p(1+\varepsilon)}{\varepsilon}}\,J_g\\ &\le \left(\int_{\mathbb{D}}|\mu-\gamma|^{\frac{p^2(1+\varepsilon)}{\varepsilon(p-2)}}\right)^{(p-2)/p}\left(\int_{\mathbb{D}} J_g^{p/2}\right)^{2/p}\end{aligned}$$

Since μ can be approximated in the mean by smooth γ, the last term can be made arbitrarily small. By uniform convergence $\gamma(\psi_{\lambda_n}(z,\xi_n)) \to \gamma(\psi_\infty(z))$, and so we see that the last bound in (93) converges to zero as $\kappa_{\lambda_n,\xi_n} \to \kappa$. In view of (92) and (93), we have established that

$$\psi_{\lambda_n}(z, \xi_n) \to z$$

and that $\psi_\infty(z) \equiv z$. The theorem is proved. $\square$

8. The Solution to Calderón's Problem

The Jacobian $J(z, f)$ of a quasiregular map can vanish only on a set of Lebesque measure zero. Since $J(z, f) = |f_z|^2 - |f_{\bar z}|^2 \le |f_z|^2$, this implies that once we know the values $f_\mu(z, \xi)$ for every $z \in \mathbb{C}$, then we can recover the values $\mu(z)$ and hence $\sigma(z)$ almost everywhere, from f_μ by the formulas

$$\frac{\partial f_\mu}{\partial \bar z} = \mu(z)\overline{\frac{\partial f_\mu}{\partial z}} \quad \text{and} \quad \sigma = \frac{1-\mu}{1+\mu} \tag{95}$$

On the other hand, considering the functions

$$u_1 = u_\sigma = \text{Re } f_\mu + i \,\text{Im } f_{-\mu} \quad \text{and} \quad u_2 = i u_{1/\sigma} = i \,\text{Re } f_{-\mu} - \text{ Im } f_\mu$$

that were described in Corollary 6.4, it is clear that the pair $\{u_1(z, \xi), u_2(z, \xi)\}$ determines the pair $\{f_\mu(z, \xi), f_{-\mu}(z, \xi)\}$, and vice versa. Therefore to prove Theorem 1.1 it will suffice to establish the following result.

Theorem 8.1. *Assume that $\Lambda_\sigma = \Lambda_{\widetilde\sigma}$ for two L^∞-conductivities σ and $\widetilde\sigma$. Then for all $z, \xi \in \mathbb{C}$,*

$$u_\sigma(z, \xi) = u_{\widetilde\sigma}(z, \xi) \quad \textit{and} \quad u_{1/\sigma}(z, \xi) = u_{1/\widetilde\sigma}(z, \xi)$$

For the proof of the theorem, our first task it to determine the asymptotic behavior of $u_\sigma(z, \xi)$. We state this as a separate result.

Lemma 8.2. *We have $u_\sigma(z, \xi) \ne 0$ for every $(z, \xi) \in \mathbb{C} \times \mathbb{C}$. Furthermore, for each fixed $\xi \ne 0$, we have with respect to z*

$$u_\sigma(z, \xi) = \exp(i\xi z + v(z)),$$

where $v = v_\xi \in L^\infty(\mathbb{C})$. On the other hand, for each fixed z we have with respect to ξ

$$u_\sigma(z, \xi) = \exp(i\xi z + \xi\varepsilon(\xi)), \tag{96}$$

where $\varepsilon(\xi) \to 0$ as $\xi \to \infty$.

Proof. For the first claim we write

$$\begin{aligned} u_\sigma &= \frac{1}{2}\left(f_\mu + f_{-\mu} + \overline{f_\mu} - \overline{f_{-\mu}}\right) \\ &= f_\mu \left(1 + \frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}}\right)^{-1} \left(1 + \frac{\overline{f_\mu} - \overline{f_{-\mu}}}{f_\mu + f_{-\mu}}\right) \end{aligned}$$

Each factor in the product is continuous and nonvanishing in z by Theorem 6.1. Taking the logarithm and using $f_{\pm\mu}(z, \xi) = e^{i\xi z}(1 + \mathcal{O}_\xi(1/z))$ leads to

$$u_\sigma(z, \xi) = \exp\left(i\xi z + \mathcal{O}_\xi\left(\frac{1}{z}\right)\right)$$

For the ξ-asymptotics we apply Theorem 7.1, which governs the growth of the functions f_μ for $\xi \to \infty$. We see that for (96) it is enough to show that

$$\inf_t \left| \frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}} + e^{it} \right| \geq e^{-|\xi|\varepsilon(\xi)} \tag{97}$$

For this, define

$$\Phi_t = e^{-it/2}(f_\mu \cos t/2 + i f_{-\mu} \sin t/2)$$

Then for each fixed ξ,

$$\Phi_t(z,\xi) = e^{i\xi z}\left(1 + \mathcal{O}_\xi\left(\frac{1}{z}\right)\right) \quad \text{as } z \to \infty,$$

and

$$\frac{\partial}{\partial \bar z}\,\Phi_t = \mu e^{-it}\,\overline{\frac{\partial}{\partial z}\,\Phi_t}$$

Thus for $\lambda = e^{-it}$, the mapping $\Phi_t = f_{\lambda\mu}$ is precisely the exponentially growing solution from (70) and (71). But

$$\frac{f_\mu - f_{-\mu}}{f_\mu + f_{-\mu}} + e^{it} = \frac{2e^{it}\,\Phi_t}{f_\mu + f_{-\mu}} = \frac{f_{\lambda\mu}}{f_\mu}\,\frac{2e^{it}}{1 + M_{-\mu}/M_\mu} \tag{98}$$

By Theorem 7.1,

$$e^{-|\xi|\varepsilon_1(\xi)} \leq |M_{\pm\mu}(z,\xi)| \leq e^{|\xi|\varepsilon_1(\xi)} \tag{99}$$

and

$$e^{-|\xi|\varepsilon_2(\xi)} \leq \inf_{\lambda\in\partial\mathbb{D}} \left|\frac{f_{\lambda\mu}(z,\xi)}{f_\mu(z,\xi)}\right| \leq \sup_{\lambda\in\partial\mathbb{D}} \left|\frac{f_{\lambda\mu}(z,\xi)}{f_\mu(z,\xi)}\right| \leq e^{|\xi|\varepsilon_2(\xi)}, \tag{100}$$

where $\varepsilon_j(\xi) \to 0$ as $\xi \to \infty$. Since $\mathrm{Re}\,(M_{-\mu}/M_\mu) > 0$, the inequality (97) follows, completing the proof of the lemma. □

As discussed earlier, the functions $u_1 = u_\sigma$ and $u_2 = i u_{1/\sigma}$ satisfy a $\partial_{\bar\xi}$-equation as a function of the parameter ξ, but it is clear that for a fixed z the asymptotics in (96) are not strong enough to determine the individual solution $u_j(z,\xi)$. However, if we consider the entire family $\{u_j(z,\xi) : z \in \mathbb{C}\}$, then, somewhat surprisingly, uniqueness properties do arise.

To prove this assume that the Dirichlet-to-Neumann operators are equal for the conductivities σ and $\widetilde\sigma$. By Lemma 8.2, we have $u_\sigma(z,\xi), u_{\widetilde\sigma}(z,\xi) \neq 0$ at every point (z,ξ). Therefore we can take their logarithms δ_σ and $\delta_{\widetilde\sigma}$, respectively, where for each fixed $z \in \mathbb{C}$,

$$\delta_\sigma(z,\xi) = \log u_\sigma(z,\xi) = i\xi z + \xi\varepsilon_1(\xi) \tag{101}$$

$$\delta_{\widetilde\sigma}(z,\xi) = \log u_{\widetilde\sigma}(z,\xi) = i\xi z + \xi\varepsilon_2(\xi) \tag{102}$$

Here, for $|\xi| \to \infty$, $\varepsilon_j(\xi) \to 0$. Moreover, by Theorem 3.1,

$$\delta_\sigma(z,0) \equiv \delta_{\widetilde\sigma}(z,0) \equiv 0$$

for all $z \in \mathbb{C}$.

In addition, $z \mapsto \delta_\sigma(z,\xi)$ is continuous, and we have

$$\delta_\sigma(z,\xi) = i\xi z \left(1 + \frac{v_\xi(z)}{i\xi z}\right), \qquad \xi \neq 0, \tag{103}$$

where by Lemma 8.2, $v_\xi \in L^\infty(\mathbb{C})$ for each fixed $\xi \in \mathbb{C}$. Since δ_σ is close to a multiple of the identity for $|z|$ large, an elementary topological argument shows that $z \mapsto \delta_\sigma(z,\xi)$ is surjective $\mathbb{C} \to \mathbb{C}$.

To prove the theorem it suffices to show that, if $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$, then

$$\delta_{\widetilde{\sigma}}(z,\xi) \neq \delta_\sigma(w,\xi) \quad \text{for } z \neq w \text{ and } \xi \neq 0 \tag{104}$$

If this property is established, then (104) and the surjectivity of $z \mapsto \delta_\sigma(z,\xi)$ show that we necessarily have $\delta_\sigma(z,\xi) = \delta_{\widetilde{\sigma}}(z,\xi)$ for all $\xi, z \in \mathbb{C}$. Hence $u_{\widetilde{\sigma}}(z,\xi) = u_\sigma(z,\xi)$.

We are now at a point where the $\partial_{\overline{\xi}}$-method and (69) can be applied. Substituting $u_\sigma = \exp(\delta_\sigma)$ in this identity shows that $\xi \to \delta_\sigma(z,\xi)$ and $\xi \to \delta_{\widetilde{\sigma}}(w,\xi)$ both satisfy the $\partial_{\overline{\xi}}$-equation

$$\frac{\partial \delta}{\partial \overline{\xi}} = -i\tau(\xi)\, e^{(\overline{\delta} - \delta)}, \qquad \xi \in \mathbb{C}, \tag{105}$$

where by Theorem 4.3 and the assumption $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$, the coefficient $\tau(\xi)$ is the same for both functions δ_σ and $\delta_{\widetilde{\sigma}}$. The difference

$$g(\xi) := \delta_{\widetilde{\sigma}}(w,\xi) - \delta_\sigma(z,\xi)$$

thus satisfies the identity

$$\frac{\partial g}{\partial \overline{\xi}} = -i\tau(\xi)\, e^{(\overline{\delta} - \delta)} \left[e^{(\overline{g} - g)} - 1\right]$$

In particular,

$$\left|\frac{\partial g}{\partial \overline{\xi}}\right| \leq |\overline{g} - g| \leq 2|g| \tag{106}$$

From (101) we have $g(\xi) = i(w-z)\xi + \xi\varepsilon(\xi)$. Now we only need to apply Theorem A.1 (with respect to ξ) to see that for $w \neq z$ the function g vanishes only at $\xi = 0$. This establishes (104).

According to Theorem 4.2 (or by the identity $\tau_\sigma = -\tau_{1/\sigma}$), if $\Lambda_\sigma = \Lambda_{\widetilde{\sigma}}$, the same argument works to show that $u_{1/\widetilde{\sigma}}(z,\xi) = u_{1/\sigma}(z,\xi)$ as well. Theorem 8.1 is thus proved. As the pair $\{u_1(z,\xi), u_2(z,\xi)\}$ pointwise determines the pair $\{f_\mu(z,\xi), f_{-\mu}(z,\xi)\}$, we find via (95) that $\sigma \equiv \widetilde{\sigma}$. Therefore the proof of Theorem 1.1 is complete. $\square$

9. Non-isotropic conductivities

Let us then consider the anisotropic conductivity equation in two dimensions

$$\begin{aligned} \nabla \cdot \sigma \nabla u = \sum_{j,k=1}^{2} \frac{\partial}{\partial x^j}\sigma_{jk}(x)\frac{\partial}{\partial x^k} u &= 0 \text{ in } \Omega, \\ u|_{\partial\Omega} &= \phi. \end{aligned} \tag{107}$$

The conductivity $\sigma = [\sigma_{jk}]_{j,k=1}^2$ is a now symmetric, positive definite matrix function.

Applying the divergence theorem, we have

$$Q_\sigma(\phi) := \int_\Omega \sum_{j,k=1}^{n} \sigma_{jk}(x)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k} dx = \int_{\partial\Omega} \Lambda_\sigma(\phi)\phi \, dS, \tag{108}$$

where dS denotes arc length on $\partial\Omega$. The quantity $Q_\sigma(\phi)$ represents the power needed to maintain the potential ϕ on $\partial\Omega$. By symmetry of Λ_σ, knowing Q_σ is equivalent with knowing Λ_σ; thus for general Ω and $\sigma \in L^\infty(\Omega)$, formula (108) may be used to define the map Λ_σ. Moreover, the Dirichlet-to-Neumann quadratic form corresponding to (Ω, σ) is

$$Q_\sigma[\phi] = \inf A_\sigma[u], \quad \text{where } A_\sigma[u] = \int_\Omega \sigma(x)\nabla u \cdot \nabla u \, dx, \tag{109}$$

and the infimum is taken over $u \in L^1(\Omega)$ such that $\nabla u \in L^1(\Omega)^3$ and $u|_{\partial\Omega} = \phi$. In the case where $A[u]$ reaches its minimum at some u, we say that u is a $W^{1,1}(\Omega)$ solution of the conductivity problem. This definition actually generalizes to degenerate conductivities to be discussed in the next section.

If $F : \Omega \to \Omega$, $\quad F(x) = (F^1(x), F^2(x))$, is a diffeomorphism with $F|_{\partial\Omega} = \text{Identity}$, then by making the change of variables $y = F(x)$ and setting $v = u \circ F^{-1}$ in the first integral in (108), we obtain

$$\nabla\cdot (F_*\sigma)\nabla v = 0 \quad \text{in } \Omega,$$

where

$$(F_*\sigma)^{jk}(y) = \frac{1}{\det\left[\frac{\partial F^j}{\partial x^k}(x)\right]} \sum_{p,q=1}^{n} \frac{\partial F^j}{\partial x^p}(x)\,\frac{\partial F^k}{\partial x^q}(x)\sigma_{pq}(x)\Bigg|_{x=F^{-1}(y)}, \tag{110}$$

or

$$F_*\sigma(y) = \frac{1}{J_F(x)} DF(x)\,\sigma(x)\,DF(x)^t\Bigg|_{x=F^{-1}(y)}, \tag{111}$$

is the push-forward of the conductivity σ by F. Moreover, since F is identity at $\partial\Omega$, we obtain from (108) that

$$\Lambda_{F_*\sigma} = \Lambda_\sigma, \tag{112}$$

Thus, the change of coordinates shows that there is a large class of conductivities which give rise to the same electrical measurements at the boundary.

We consider here the converse question, that if we have two conductivities which have the same Dirichlet-to-Neumann map, is it the case that each can be obtained by pushing forward the other, by a mapping that is the identity on the boundary.

If $\Omega \subset \mathbb{R}^2$ is a bounded domain, it is convenient to consider the class of matrix functions $\sigma = [\sigma_{jk}]$ for which (11) holds for some constant $1 \le K < \infty$.

$$[\sigma_{ij}] \in L^\infty(\Omega; \mathbb{R}^{2\times 2}), \quad [\sigma_{ij}]^t = [\sigma_{ij}], \quad K^{-1} I \le [\sigma_{ij}] \le KI \tag{113}$$

In sequel, the minimal possible value of K is denoted by $K(\sigma)$. We use the notation

$$\Sigma(\Omega) = \{\sigma \in L^\infty(\Omega; \mathbb{R}^{2\times 2}) \quad | \quad K(\sigma) < \infty\}.$$

Note that it is necessary to require some control on the distortion as otherwise there would be counterexamples showing that even the equivalence class of the conductivity can not be recovered.

The main goal of this section is to show that an anisotropic L^∞–conductivity can be determined up to a $W^{1,2}$-diffeomorphism:

Theorem 9.1. [5] *Let $\Omega \subset \mathbb{R}^2$ be a simply connected bounded domain and $\sigma \in L^\infty(\Omega; \mathbb{R}^{2\times 2})$. Suppose that the assumptions (113) are valid. Then the Dirichlet-to-Neumann map Λ_σ determines the equivalence class*

$$E_\sigma = \{\sigma_1 \in \Sigma(\Omega) \quad | \quad \sigma_1 = F_*\sigma,\ F : \Omega \to \Omega \textit{ is } W^{1,2}\textit{-diffeomorphism and } F|_{\partial\Omega} = I\}.$$

Note that the $W^{1,2}$–diffeomorphisms F preserving the class $\Sigma(\Omega)$ are precisely the quasiconformal mappings. Namely, if $\sigma_0 \in \Sigma(\Omega)$ and $\sigma_1 = F_*(\sigma_0) \in \Sigma(F(\Omega))$ then

$$\frac{1}{K_0} ||DF(x)||^2 I \le DF(x)\, \sigma_0(x)\, DF(x)^t \le K_1 J_F(x) I \tag{114}$$

and we obtain

$$||DF(x)||^2 \le K J_F(x), \quad a.e.\ x \in \Omega \tag{115}$$

where $K = K_1 K_0 < \infty$. Clearly the converse is also true.

Proof of Theorem 9.1 (*Sketch*) From (42) we have the complex geometric optics solutions $u(z,\xi) = \operatorname{Re} f_{\mu,\nu} + i \operatorname{Im} f_{\mu,-\nu}$ where the f's are exponentially growing solutions to the appropriate Beltrami equations. One can find a quasiconformal change of coordinates F such that $F_*\sigma$ is isotropic, see e.g. [3, Theorem 10.2.1].

Now, if we have two (anisotropic) conductivities such that the Dirichlet-to-Neumann operators $\Lambda_{\sigma_1} = \Lambda_{\sigma_2}$, the by (112) and Theorem 1.1 the corresponding isotropic conductivites, obtained by a quasiconformal coordinate change, satisfy

$$(F_1)_*\sigma_1 = (F_2)_*\sigma_1$$

It remains to show that $F_1 = F_2$ on the boundary $\partial\mathbb{D}$. Here we return to the exponentially growing solutions; similarly as in Section 4 we see that *outside the disk* these are determined by Λ_σ. On the other hand, from the uniqueness of the exponential solutions,

$$u_{(F_j)_{*\sigma_j}} \circ F_j = u_{\sigma_j}, \quad j = 1, 2.$$

Therefore from the subexponential growth, Lemma 8.2,

$$F_1(z) = \lim_{\xi\to\infty} \frac{\log u_{(F_1)_{*\sigma_1}} \circ F_1(z)}{i\xi} = \lim_{\xi\to\infty} \frac{\log u_{(F_2)_{*\sigma_2}} \circ F_2(z)}{i\xi} = F_2(z), \quad |z| > 1.$$

$\square$

10. Degenerate conductivities and limits visibility and invisibility

When we allow degenerate conductivities, the results of the previous section begin to break. But on the positive side, we can create invisibility cloakings, see the next example from [18]. In the "quasiconformal world", a similar idea was used by J. Ball to find examples of mapping creating cavities.

These ideas have now been greatly generalized, see e.g. [17].

10.1. **Counterexample 1:** Cloaking of an arbitrary body using a very degenerate conductivity.

Let $B(\rho)$ be an open 2-dimensional disc of radius ρ and center zero. Consider the map

$$F : B(2) \setminus \{0\} \to B(2) \setminus \overline{B}(1), \quad F(x) = (\frac{|x|}{2} + 1)\frac{x}{|x|}. \tag{116}$$

Let us apply the push forward operation for conductivities, defined in (111). Using the map F defined in (116), we can define a singular conductivity

$$\tilde{\sigma} = \begin{cases} F_*\sigma(x) & \text{for } x \in B(2) \setminus \overline{B}(1), \\ a^{jk} & \text{for } x \in \overline{B}(1), \end{cases} \tag{117}$$

where $A = [a^{jk}]$ is any symmetric measurable matrix satisfying $c_1 I \leq A(x) \leq c_2 I$ with $c_1, c_2 > 0$. This conductivity is the so-called *cloaking conductivity.*

Proposition 10.1. *Let $\tilde{\sigma}$ be the conductivity defined in (116) and $\gamma \equiv 1$. Then the boundary measurements for $\tilde{\sigma}$ and γ coincide in the sense that $Q_{\tilde{\sigma}} = Q_\gamma$.*

Proof. For $0 \leq r \leq 2$ and a conductivity η we define the quadratic form $A^r_\eta : W^{1,1}(B(2)) \to \mathbb{R}_+ \cup \{0, \infty\}$,

$$A^r_\eta[u] = \int_{B(2)\setminus B(r)} \eta(x)\nabla u \cdot \nabla u \, dx.$$

Considering F as a change of variables, we see that

$$A^r_{\tilde{\sigma}}[u] = A^\rho_\gamma[v], \quad u = v \circ F, \ \rho = 2(r-1), \ r > 1.$$

Now for the conductivity $\gamma = 1$ the minimization problem (109) is solved by the unique minimizer u satisfying

$$\Delta u = 0 \quad \text{in } B(2), \quad u|_{\partial B(2)} = f.$$

The solution u is C^∞-smooth in $B(2)$ and we see that $v = u \circ F$ is a $W^{1,1}$-function on $B(2) \setminus \overline{B}(0,1)$ which trace on $\partial B(0,1)$ is equal to the constant function $h(x) = u(0)$ on $\partial B(1)$. Defining function $\tilde{v}$ that is v in $B(2) \setminus \overline{B}(0,1)$ and $u(0)$ in $\overline{B}(1)$ we obtain a $W^{1,1}(B(2))$ function for which

$$Q_{\tilde{\sigma}}[f] \leq A^1_{\tilde{\sigma}}[v] = \lim_{r\to 1} A^r_{\tilde{\sigma}}[v] = \lim_{\rho\to 0} A^\rho_\gamma[u] = Q_\gamma[f]. \tag{118}$$

To construct an inequality opposite to (118), let η_ρ be a conductivity which coincides with $\tilde{\sigma}$ in $B(2) \setminus B(\rho)$ and is 0 in $B(\rho)$. For this conductivity the minimization problem

(109) has a minimizer that in $B(2) \setminus \overline{B}(\rho)$ coincide with the solution of the boundary value problem

$$\Delta u = 0 \quad \text{in } B(2) \setminus \overline{B}(\rho), \quad u|_{\partial B(2)} = f, \quad \partial_\nu u|_{\partial B(\rho)} = 0$$

and is arbitrary $W^{1,1}$-smooth extension of u to $B(\rho)$. Then $\eta_\rho(x) \le \tilde\sigma(x)$ for all $x \in B(2)$ and thus

$$Q_{\eta_\rho}[f] \le Q_{\tilde\sigma}[f]$$

It is not difficult to see that

$$\lim_{\rho \to 0} Q_{\eta_\rho}[f] = Q_\gamma[f],$$

that is, the effect of an insulating disc of radius ρ in the boundary measurements vanishes as $\rho \to 0$. These and (118) yield $Q_{\tilde\sigma}[f] = Q_\gamma[f]$. □

As eigenvalues of the cloaking conductivity $\tilde\sigma$ in $B(2) \setminus \overline{B}(1)$ behave asymptotically as $(|x| - 1)$ and $(|x| - 1)^{-1}$ as $|x| \to 1$, the cloaking conductivity has so strong degeneracy that

$$\mathrm{Trace}(\tilde\sigma) \notin L^1(B(2) \setminus \overline{B}(1))$$

even though $\det(\tilde\sigma)$ is identically 1 in $B(2) \setminus \overline{B}(1)$.

10.2. **Counterexample 2:** Illusion of a non-existing obstacle showing that even the topology of the domain can not be determined with relatively weakly degenerate conductivities.

Consider again the map $F : B(2)\setminus\{0\} \to B(2)\setminus\overline{B}(1)$. Instead of considering the above conductivity $\tilde\sigma = F_*\gamma$, $\gamma = 1$, and change our point of view by looking for a conductivity σ_0 satisfying $F_*\sigma_0 = \gamma$. That is, σ_0 satisfies satisfies

$$DF(x)\sigma_0(x)DF(x)^t = [\det DF(x)]\, Id.$$

A simple computation shows that

$$\mathrm{Trace}(\sigma_0) \in L^p(\Omega),\ p < 2.$$

Let us now consider other possible maps F that might produce a cavity with as little degeneracy in distrotion as possible. Here Iwaniec and Martin [20] have basically identified the extreme behaviour. Assume that $\mathcal{A}(t)$, $t \ge 0$ is an increasing positive function that is sub-linear in the sense that

$$\int_1^\infty \frac{\mathcal{A}(t)}{t^2} < \infty. \tag{119}$$

Then there exists a $W^{1,1}$-homeomorphism $F_1 : B(2)\setminus\{0\} \to B(2)\setminus\overline{B}(1)$ and a conductivity σ such that

$$DF_1(x)\sigma(x)DF_1(x)^t = [\det DF_1(x)]\, Id, \tag{120}$$

$$\det(\sigma) = 1, \tag{121}$$

$$\int_\Omega \exp(\mathcal{A}(\mathrm{Trace}(\sigma) + \mathrm{Trace}(\sigma^{-1})))\, dx < \infty. \tag{122}$$

Note here that in view of (21), we have

$$K(z,F) \leq \text{Trace}(\sigma) + \text{Trace}(\sigma^{-1}) \leq 4K(z,F)$$

Let us now consider a conductivity $\sigma_1 = (F_1^{-1})_*\gamma_1$ on $B(2)\setminus\{0\}$ that is considered as an a.e. defined measurable function on $B(2)$ and the conductivity γ_1 that is identically 1 in $B(2)\setminus\overline{B}(1)$ and zero in $B(1)$. Note that γ_1 corresponds the case when $B(1)$ as a perfect insulator. Then we have

Theorem 10.2. *Let $\mathcal{A}$ be a function satisfying (119). Then there is a conductivity σ_1 satisfying (121-122) such that the boundary measurements corresponding to σ_1 and γ_1 coincide, that is,*

$$Q_{\sigma_1} = Q_{\gamma_1}, \quad \textit{i.e. } \Lambda_{\sigma_1} = \Lambda_{\sigma_2}. \tag{123}$$

Proof. Following [20, sect. 11.2.1], there is $k(s)$ satisfies the relation

$$k(s)e^{\mathcal{A}(k(s))} = \frac{e}{s^2}, \quad 0 < s < 1$$

that is strictly decreasing function and satisfies $k(s) \leq s^{-1}$ and $k(1) = 1$. Then

$$\rho(t) = \exp\left(\int_0^t \frac{ds}{sk(s)}\right)$$

is a function for which $\rho(0) = 1$. Then by defining fucntion the maps $h(t) = 2\rho(\frac{|x|}{2})/\rho(1)$ and

$$F_h : B(2)\setminus\{0\} \to B(2)\setminus\overline{B}(1), \quad F_h(x) = h(t)\frac{x}{|x|} \tag{124}$$

and $\sigma_1 = (F_h)_*\gamma_1$, we obtain a conductivity that satisfies conditions (121-122).

Finally, the proof of (123) follows similarly as in Proposition 10.1. □

We note that above in (**??**) in the case where f and thus u are strictly positive functions, we have $Dv \notin L^2(B(2)\setminus\{0\})$. This shows that in the minimization problem (109) the minima is not obtained in the class $W^{1,2}(\Omega)$ for all conductivities.

Theorem 10.2 can be interpreted by saying that there is a relatively weakly degenerated conductivity satisfying integrability condition (120) that creates in the boundary observations an illusion of an obstacle that does not exists. Thus the conductivity can be considered as "electromagnetic hologram". As the obstacle can be considered as a "hole" in the domain, we can say also that even the topology of the domain can not be detected.

11. Calderon's inverse problem at the borderline.

The example of Iwaniec and Martin essentially presents the smallest degenerate distortion for a mapping to create a cavity. This suggest that if the integral in (119) diverges, then Calderon's might be solvable. With little extra technical conditions on $\mathcal{A}$ this appears to be the case. We consider gauges $\mathcal{A}$ controlling the distortion function, with the following properties:

(1) $\mathcal{A} : [1,\infty) \to [0,\infty)$ is a smooth increasing function with $\mathcal{A}(1) = 0$.

(2)
$$\int_1^\infty \frac{\mathcal{A}(t)}{t^2}\,dt = \infty$$

(3) $t\,\mathcal{A}'(t) \geq 5$ for large values of t

Note in particular that linear $\mathcal{A}(t)$ satisfy these conditions.

Theorem 11.1. [6] *Let $\sigma_1, \sigma_2 \in \Sigma_{\mathcal{A}}$ where $\mathcal{A}$ satisfies (1)-(3). Suppose*
$$\det \sigma_j, (\det \sigma_j)^{-1} \in L^\infty(\Omega), \quad j = 1, 2.$$
and assume that
$$Q_{\sigma_1} = Q_{\sigma_2}.$$
Then
$$\sigma_1 = F_* \sigma_2$$
with a $W^{1,1}_{loc}$-homeomorphism $F : \Omega \to \Omega$ satisfying $F|_{\partial\Omega} = \mathrm{id}$.

Proof. The idea of the argument is similar to that one in Theorem 9.1. To construct the complex geometric solutions for σ we now need to solve the non-linear Beltrami equation in the degenerate setting, and here estimates from Theorem B.11 are valuable. Similarly it allows a change of variables to obtain isotropic conductivities. Finally the proof is reduced to Theorem 1.1. □

In [6] Theorem 1.1 is also generalized to certain unbounded isotropic conductivities.

By Theorem 10.2 the condition (1)-(3) cannot be weakened. We have thus identified "the borderline between visibility and invisibility" !!!

Appendix A. Argument Principle

The solution to the Calderón problem combines analysis with topological arguments that are specific to two dimensions. For instance, we need a version of the argument principle, which we here consider.

Theorem A.1. *Let $F \in W^{1,p}_{\text{loc}}(\mathbb{C})$ and $\gamma \in L^p_{\text{loc}}(\mathbb{C})$ for some $p > 2$. Suppose that, for some constant $0 \le k < 1$, the differential inequality*

$$\left|\frac{\partial F}{\partial \bar{z}}\right| \le k \left|\frac{\partial F}{\partial z}\right| + \gamma(z)\left|F(z)\right| \tag{125}$$

holds for almost every $z \in \mathbb{C}$ and assume that, for large z, $F(z) = \lambda z + \varepsilon(z) z$, where the constant $\lambda \neq 0$ and $\varepsilon(z) \to 0$ as $|z| \to \infty$.

Then $F(z) = 0$ at exactly one point, $z = z_0 \in \mathbb{C}$.

Proof. The continuity of $F(z) = \lambda z + \varepsilon(z)z$ and an elementary topological argument show that F is surjective, and consequently there exists at least one point $z_0 \in \mathbb{C}$ such that $F(z_0) = 0$.

To show that F cannot have more zeros, let $z_1 \in \mathbb{C}$ and choose a large disk $B = \mathbb{D}(0, R)$ containing both z_1 and z_0. If R is so large that $\varepsilon(z) < \lambda/2$ for $|z| = R$, then $F\big|_{\{|z|=R\}}$ is homotopic to the identity relative to $\mathbb{C} \setminus \{0\}$. Next, we express (125) in the form

$$\frac{\partial F}{\partial \bar{z}} = \nu(z)\frac{\partial F}{\partial z} + A(z)\, F, \tag{126}$$

where $|\nu(z)| \le k < 1$ and $|A(z)| \le \gamma(z)$ for almost every $z \in \mathbb{C}$. Now $A\chi_B \in L^r(\mathbb{C})$ for all $2 \le r \le p' = \min\{p, 1+1/k\}$, and we obtain from Theorem B.4 that $(\mathbf{I} - \nu\mathcal{S})^{-1}(A\chi_B) \in L^r$ for all $p'/(p'-1) < r < p'$.

Next, we define $\eta = \mathcal{C}\big((\mathbf{I} - \nu\mathcal{S})^{-1}(A\chi_B)\big)$. Then by Theorem B.3 we have $\eta \in C_0(\mathbb{C})$, and we also have

$$\frac{\partial \eta}{\partial \bar{z}} - \nu\frac{\partial \eta}{\partial z} = A(z), \qquad z \in B \tag{127}$$

Therefore simply by differentation we see that the function

$$g = e^{-\eta} F \tag{128}$$

satisfies

$$\frac{\partial g}{\partial \bar{z}} - \nu\frac{\partial g}{\partial z} = 0, \qquad z \in B \tag{129}$$

Since η has derivatives in $L^r(\mathbb{C})$, we have $g \in W^{1,r}_{loc}(\mathbb{C})$. As $r \ge 2$, the mapping g is quasiregular in B. The Stoilow factorization theorem gives $g = h \circ \psi$, where $\psi : B \to B$ is a quasiconformal homeomorphism and h is holomorphic, both continuous up to the boundary.

Since η is continuous, (128) shows that $g\big|_{|z|=R}$ is homotopic to the identity relative to $\mathbb{C} \setminus \{0\}$, as is the holomorphic function h. Therefore the argument principle shows that h has precisely one zero in $B = \mathbb{D}(0, R)$. Already, $h(\psi(z_0)) = e^{-\eta(z_0)} F(z_0) = 0$, and there can be no further zeros for F either. This finishes the proof. $\square$

Appendix B. Some Background in complex analysis and quasiconformal mappings.

Here we collect, without proof, some basic facts related to quasiconformal mappings. The proofs can be found e.g. in [3].

We start with harmonic analysis, where we often need refine estimates of the Cauchy transform.

Definition B.1. *The Cauchy transform is defined by the rule*

$$(\mathcal{C}\phi)(z) := \frac{1}{\pi}\int_{\mathbb{C}} \frac{\phi(\tau)}{z-\tau}\, d\tau \tag{130}$$

Theorem B.2. *Let $1 < p < \infty$. If $\phi \in L^p(\mathbb{C})$ and $\phi(\tau) = 0$ for $|\tau| \geq R$, then*

- $\|\mathcal{C}\phi\|_{L^p(D_{2R})} \leq 6R\,\|\phi\|_p$
- $\|\mathcal{C}\phi(z) - \frac{1}{\pi z}\int \phi\|_{L^p(\mathbb{C}\setminus D_{2R})} \leq \frac{2\,R}{(p-1)^{1/p}}\|\phi\|_p$

Thus, in particular, for $1 < p \leq 2$,

$$\|\mathcal{C}\phi\|_{L^p(\mathbb{C})} \leq \frac{8\,R}{(p-1)^{1/p}}\|\phi\|_p \qquad \textit{provided} \int \phi = 0$$

For $p > 2$ this vanishing condition is not needed, and we have

$$\|\mathcal{C}\phi\|_{L^p(\mathbb{C})} \leq (6 + 3(p-2)^{-1/p})\, R\,\|\phi\|_p, \qquad p > 2$$

Concerning compactness, we have

Theorem B.3. *Let Ω be a bounded measurable subset of $\mathbb{C}$. Then the following operators are compact.*

- *For $2 < p \leq \infty$,*

$$\chi_\Omega \circ \mathcal{C} : L^p(\mathbb{C}) \to C^\alpha(\Omega), \qquad 0 \leq \alpha < 1 - \frac{2}{p}$$

- *For $1 \leq p \leq 2$,*

$$\chi_\Omega \circ \mathcal{C} : L^p(\mathbb{C}) \to L^s(\mathbb{C}), \qquad 1 \leq s < \frac{2p}{2-p}$$

The fundamental operator in the theory of planar quasiconformal mappings id the Beurling transform,

$$(\mathcal{S}\phi)(z) := -\frac{1}{\pi}\int_{\mathbb{C}} \frac{\phi(\tau)}{(z-\tau)^2}\, d\tau \tag{131}$$

The importance of the Beurling transform in complex analysis is furnished by the identity

$$\mathcal{S} \circ \frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial z}, \tag{132}$$

initially valid for functions contained in the space $C_0^\infty(\mathbb{C})$. Moreover, $\mathcal{S}$ extends to abounded operator on $L^p(\mathbb{C})$, $1 < p < \infty$; on $L^2(\mathbb{C})$ it is an isometry. We denote by

$$\mathbf{S}_p := \|\mathcal{S} : L^p(\mathbb{C}) \to L^p(\mathbb{C})\|$$

the norm of this operator. By Riesz-Thorin interpolation, $\mathbf{S}_p \to 1$ as $p \to 2$.

In other words, $\mathcal{S}$ intertwines the Cauchy-Riemann operators $\frac{\partial}{\partial \bar{z}}$ and $\frac{\partial}{\partial z}$, a fact that explains the importance of the operator in complex analysis. For instance we have [3, p.363] the following result.

Theorem B.4. *Let μ be measurable with $\|\mu\|_\infty \le k < 1$. Then the operator*

$$\mathbf{I} - \mu\mathcal{S}$$

is invertible on $L^p(\mathbb{C})$ whenever $\|\mu\|_\infty \le k < 1$ and $1 + k < p < 1 + 1/k$.

The result has important consequences on the regularity of elliptic systems. In fact, it is equivalent to the improved Sobolev regularity of quasiregular mappings.

Theorem B.5. *Let $\mu, \nu \in L^\infty(\mathbb{C})$ with $|\mu| + |\nu| \le k < 1$ almost everywhere. Then the equation*

$$\frac{\partial f}{\partial \bar{z}} - \mu(z)\frac{\partial f}{\partial z} - \nu(z)\overline{\frac{\partial f}{\partial z}} = h(z)$$

has a solution f, locally integrable with gradient in $L^p(\mathbb{C})$, whenever $1 + k < p < 1 + 1/k$ and $h \in L^p(\mathbb{C})$. Further, f is unique up to an additive constant.

We will also need a simple version of the Koebe distortion theorem.

Lemma B.6. [3, p. 42] *If $f \in W^{1,1}_{loc}(\mathbb{C})$ is a homeomorphism analytic outside the disk $\mathbb{D}(0, r)$ with $|f(z) - z| = o(1)$ at ∞, then*

$$|f(z)| < |z| + 3r, \quad \textit{for all } z \in \mathbb{C} \tag{133}$$

Next, we have the continuous dependence of the quasiconformal mappings on the complex dilatation.

Lemma B.7. *Suppose $|\mu|, |\nu| \le k\chi_{\mathbb{D}_r}$, where $0 \le k < 1$. Let $f, g \in W^{1,2}_{loc}(\mathbb{C})$ be the principal solutions to the equations*

$$\frac{\partial f}{\partial \bar{z}} = \mu(z)\frac{\partial f}{\partial z}, \qquad \frac{\partial g}{\partial \bar{z}} = \nu(z)\frac{\partial g}{\partial z}$$

If for a number s we have $2 \le p < ps < P(k)$, then

$$\|f_{\bar{z}} - g_{\bar{z}}\|_{L^p(\mathbb{C})} \le C(p, s, k)\, r^{2/ps}\, \|\mu - \nu\|_{L^{ps/(s-1)}(\mathbb{C})}$$

To prove uniqueness, Liouville type result are often valuable. Here we have collected a number of such results.

Theorem B.8. *Suppose that $F \in W^{1,q}_{loc}(\mathbb{C})$ satisfies the distortion inequality*

$$|F_{\bar{z}}| \le k|F_z| + \sigma(z)|F|, \qquad 0 \le k < 1, \tag{134}$$

where $\sigma \in L^2(\mathbb{C})$ and the Sobolev regularity exponent q lies in the critical interval $1 + k < q < 1 + 1/k$. Then $F = e^\theta g$, where g is quasiregular and $\theta \in VMO$. If $\sigma \in L^{2\pm}(\mathbb{C})$, then θ is continuous, and if furthermore F is bounded, then $F = C_1 e^\theta$.

In addition, if one of the following additional hypotheses holds,

(1) *σ has compact support and* $\lim_{z\to\infty} F(z) = 0$, *or*
(2) *$F \in L^p(\mathbb{C})$ for some $p > 0$ and* $\limsup_{z\to\infty} |F(z)| < \infty$,
then $F \equiv 0$.

Here we used the notation

$$L^{2\pm}(\mathbb{C}) = \{f : f \in L^s(\mathbb{C}) \cap L^t(\mathbb{C}) \text{ for some } s < 2 < t\}.$$

Theorem B.9. (Stoilow Factorization) *Let $f : \Omega \to \Omega'$ be a homeomorphic solution to the Beltrami equation*

$$\frac{\partial f}{\partial \bar z} = \mu(z)\frac{\partial f}{\partial z}, \qquad f \in W^{1,1}_{loc}(\Omega), \tag{135}$$

with $|\mu(z)| \le k < 1$ almost everywhere in Ω.

Suppose $g \in W^{1,2}_{loc}(\Omega)$ is any other solution to (135) on Ω. Then there exists a holomorphic function $\Phi : \Omega' \to \mathbb{C}$ such that

$$g(z) = \Phi(f(z)), \qquad z \in \Omega \tag{136}$$

Conversely, if Φ is holomorphic on Ω', then the composition $\Phi \circ f$ is a $W^{1,2}_{loc}$-solution to (135) in the domain Ω.

Stoilow factorization generalizes little bit outside $W^{1,2}_{loc}$. We will assume that $f \in W^{1,Q}_{loc}(\Omega)$ where

$$Q(t) \ = \ \frac{t^2}{\log(e+t)}$$

Theorem B.10. *Suppose we are given a homeomorphic solution $f \in W^{1,Q}_{loc}(\Omega)$ to the Beltrami equation*

$$\frac{\partial f}{\partial \bar z} = \mu(z)\,\frac{\partial f}{\partial \bar z}, \qquad z \in \Omega, \tag{137}$$

where $|\mu(z)| < 1$ almost everywhere. Then every other solution $h \in W^{1,Q}_{loc}(\Omega)$ to (137) takes the form

$$h(z) = \phi(f(z)), \qquad z \in \Omega,$$

where $\phi : f(\Omega) \to \mathbb{C}$ is holomorphic.

Slightly further steps will be possible using general Orlicz space. Let us recall the notation

$$W^{1,P}(\Omega) = \{f \in W^{1,1}_{loc}(\Omega) : \int_\Omega P(|Df|) < \infty\}$$

Consider gauges $\mathcal{A}$ controlling the distortion function, with the following properties:

(1) $\mathcal{A} : [1,\infty) \to [0,\infty)$ is a smooth increasing function with $\mathcal{A}(1) = 0$.
(2)
$$\int_1^\infty \frac{\mathcal{A}(t)}{t^2}\, dt = \infty$$
(3) $\quad t\,\mathcal{A}'(t) \ge 5 \qquad$ for large values of t

In view of the above integral condition, the last easily verifiable condition involves an insignificant loss of generality as the function $t\mathcal{A}'(t)$ behaves more or less like $\mathcal{A}(t)$, at least in the typical examples we have in mind for applications. However, for technical reasons it is necessary. For the Sobolev regularity we require that f is contained in the Orlicz-Sobolev class $W^{1,P}_{loc}(\Omega)$, where

$$P(t) = \begin{cases} t^2, & 0 \leq t \leq 1 \\ \frac{t^2}{\mathcal{A}^{-1}(\log t^2)}\,, & t \geq 1 \end{cases} \tag{138}$$

Theorem B.11. (Existence and Uniqueness) *Let $\mathcal{A} = \mathcal{A}(t)$ satisfy the above conditions* 1 – 3*. Suppose the Beltrami coefficient, with $|\mu(z)| < 1$ almost everywhere, is compactly supported and the associated distortion function $K(z) = \frac{1+|\mu(z)|}{1-|\mu(z)|}$ satisfies*

$$e^{\mathcal{A}(K(z))} \in L^1_{loc}(\mathbb{C}) \tag{139}$$

Then the Beltrami equation $f_{\bar{z}}(z) = \mu(z)\ f_z(z)$ admits a unique principal solution $f \in W^{1,P}_{loc}(\mathbb{C})$ with $P(t)$ as in (138). Moreover, any solution $h \in W^{1,P}_{loc}(\Omega)$ to this Beltrami equation in a domain $\Omega \subset \mathbb{C}$ admits a factorization

$$h = \phi \circ f,$$

where ϕ is holomorphic in $f(\Omega)$.

References

[1] L.V. Ahlfors, *Lectures on quasiconformal mappings*, Van Nostrand, Princeton 1966; Reprinted by American Mathematical Society, 2006.

[2] K. Astala, *Area distortion of quasiconformal mappings*, Acta Math., **173**, (1994), 37–60.

[3] K. Astala, T. Iwaniec and G. Martin, *Elliptic Partial Differential Equations and Quasiconformal Mappings in the Plane*, Princeton Mathematical Series 48, Princeton University Press, 2009.

[4] K. Astala, T. Iwaniec and E. Saksman, *Beltrami operators in the plane*, Duke Math. J. 107 (2001), 27–56.

[5] K. Astala, M. Lassas and L. Päivärinta, *Calderón's inverse conductivity problem in the plane*, Comm. Partial Differential Equations, **30** (2005), 207–224.

[6] K. Astala, M. Lassas and L. Päivärinta, *Limits of visibility and invisibility for Calderón's inverse problem in the plane.* Preprint in preparation.

[7] K. Astala, J.E. Mueller, L. Päivärinta and S. Siltanen, *Numerical computation of complex geometrical optics solutions to the conductivity equation*, (Preprint).

[8] K. Astala and V. Nesi, *Composites and quasiconformal mappings: new optimal bounds in two dimensions,* Calculus of Variations and Partial Differential Equations, **18**, (2003), 335–355.

[9] K. Astala and L. Päivärinta, *Calderón's inverse problem for anisotropic conductivity in the plane*, Ann. of Math., **163** (2006), 265–299.

[10] R. Beals and R. Coifman, *The spectral problem for the Davey-Stewartson and Ishimori hierarchies*, Nonlinear Evolution Equations: Integrability and Spectral Methods, Manchester University Press, 1988, pp. 15–23.

[11] B. Bojarski, *Homeomorphic solutions of Beltrami systems*, Dokl. Akad. Nauk. SSSR, **102**, (1955), 661–664.

[12] L. Borcea, *Electrical impedance tomography*, Inverse problems 18 (2002) R99-R136.

[13] M. Cheney, D. Isaacson and J. Newell, *Electrical impedance tomography.* SIAM Review, **41**, (1999), 85-101.

[14] G. David, *Solutions de l'equation de Beltrami avec $\|\mu\|_\infty = 1$*, Ann. Acad. Sci. Fenn. Ser. AI Math., **13**, (1988), 25–70.

[15] A. Dijkstra, B. Brown, N. Harris, D. Barber and D. Endbrooke, *Review: Clinical applications of electrical impedance tomography*, J. Med. Eng. Technol., **17** (1993), 89–98.

[16] D. Faraco, P. Koskela and X.Zhong, *Mappings of finite distortion: the degree of regularity,* Adv. Math. 190 (2005) 300–318.

[17] A. Greenleaf, Y. Kurylev, M. Lassas, G. Uhlmann: Electromagnetic wormholes and virtual magnetic monopoles from metamaterials. *Phys. Rev. Lett.* **99**, 183901 (2007).

[18] A. Greenleaf, M. Lassas, and G. Uhlmann, On nonuniqueness for Calderón's inverse problem, *Math. Res. Lett.* **10** (2003), no. 5-6, 685-693.

[19] T. Iwaniec and G.J. Martin *Geometric Function Theory and Nonlinear Analysis*, Oxford University Press, 2001.

[20] T. Iwaniec and G.J. Martin *The Beltrami equation,* Mem. Amer. Math. Soc. 191, 2008.

[21] T. Iwaniec, T, P. Koskela and G.Martin *Mappings of BMO-distortion and Beltrami-type operators.* J. Anal. Math. 88, (2002), 337–381.

[22] T. Iwaniec, T, P. Koskela, G. Martin and C. Sbordone, *Mappings of finite distortion: $L^n \log^\chi L$-integrability*, J. London Math. Soc. (2) 67, (2003), 123–136.

[23] O. Lehto, K. Virtanen, *Quasiconformal mappings in the plane*, Second edition, Springer-Verlag, 1973.

[24] O. Martio and J. Maly, *Lusin's condition (N) and mappings of the class $W^{1,n}$,* J. Reine Angew. Math. 458, (1995), 19–36.

[25] R. Brown, *Estimates for the scattering map associated with a two-dimensional first-order system*, J. Nonlinear Sci., **11** (2001), 459–471.

[26] R. Brown and G. Uhlmann, *Uniqueness in the inverse conductivity problem for nonsmooth conductivities in two dimensions*, Comm. Partial Differential Equations, **22** (1997), 1009–1027.

[27] A. Nachman, *Global uniqueness for a two-dimensional inverse boundary value problem.* Ann. of Math., **143** (1996), 71–96.

[28] L. Päivärinta, A. Panchenko and G. Uhlmann, *Complex geometrical optics solutions for Lipschitz conductivities*, Rev. Math. Iberoamericana, **19** (2003), 57–72.

[29] J. Sylvester and G. Uhlmann, *A global uniqueness theorem for an inverse boundary value problem*, Ann. of Math., **125** (1987), 153–169.

[30] T. Tao and Ch. Thiele, *Nonlinear Fourier analysis*, Lecture Notes; LAS Park City Summer School, 2003.

Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68 (Gustaf Hällstrmin katu 2b), FI-00014 University of Helsinki, Finland.

E-mail address: `kari.astala@helsinki.fi`

The Heisenberg Group and Cortex Vision

Giovanna Citti

G. Citti, A. Sarti

April, 2009

Contents

1 Models of the visual cortex in Lie groups **1**
1.1 Introduction 1
1.2 perceptual completion phenomena 2
1.2.1 Associations fields 3
1.2.2 Higher order operators and elastica 5
1.3 The functional structure of the visual cortex 6
1.3.1 The retinotopic structure 8
1.3.2 The hypercolumnar structure 8
1.3.3 The neural circuitry 9
1.4 The visual cortex as a Lie group 10
1.4.1 A first model in the Heisenberg group 10
1.4.2 A subriemannian model in the rototraslation group 11
1.4.3 Hörmander vector fields and Sub-Riemannian structures. . 16
1.4.4 connectivity property 19
1.4.5 Control distance 22
1.4.6 Riemannian approximation of the metric 23
1.4.7 geodesics and elastica 23
1.5 Activity propagation and differential operators in Lie groups . . . 24
1.5.1 Integral curves, Association fields, and the experiment of Bosking 24
1.5.2 Differential calculus in subriemannian setting 25
1.5.3 Subriemannian differential operators 28
1.6 Regular surfaces in subriemannian setting 30
1.6.1 Maximum selectivity and lifting images to regular surfaces . 30
1.6.2 Definition of regular surface 31
1.6.3 Implicit function theorem 32
1.6.4 Non regular and non linear vector fields 36
1.7 Completion and minimal surfaces 37
1.7.1 A Completion process 37
1.7.2 Minimal surfaces in the Heisenberg group 38
1.7.3 Uniform regularity for the Riemannian approximating minimal graph 40

1.7.4 Regularity of the viscosity minimal surface 49
1.7.5 Foliation of minimal surfaces and completion result 49

Chapter 1

Models of the visual cortex in Lie groups

1.1 Introduction

The most classical and exhaustive theory which states and studies the phenomenological laws of visual reconstruction is Gestalt theory [51, 52]. It formalize visual perceptual phenomena in terms of geometric concepts, as good continuation, orientation, or vicinity. Consequently phenomenological models of vision have been expressed in terms of geometrical instruments and minima of calculus of variation ([67], [34], [6]). On the other hand the recent progress of medical imaging and integrative neuroscience allows to study neurological structures related to perception of space and motion. The first results which use instruments of differential geometry to model the cortex and justify the macroscopical visual phenomena in terms of the microscopical behavior of the cortex, are due to Hoffmann [49], and Petitot, Tondut [70]. More recently G.Citti, A.Sarti [28], modeled the visual cortex as a Lie group with a sub-Riemannian metric. Other models in Lie groups are due to Zucker [82], Duits, [33], [37]. We refer to these papers for a complete description of these type of problems.

Here we will simply give an exhaustive presentation of the model of Citti Sarti, together with the instruments of sub-Riemannian differential geometry necessary for its description, and the results which support the model. The main goal is to underline who the sub-Riemannian geometry is a natural instrument for the description of the visual cortex.

In Section 2 and 3 we will describe the problem of perceptual completion, and give a short description of the functional architecture of the visual cortex.

In Section 4 we describe the functional geometry of the visual cortex as a sub-Riemannian structure, and give the principal definition and properties of a sub-Riemannian space.

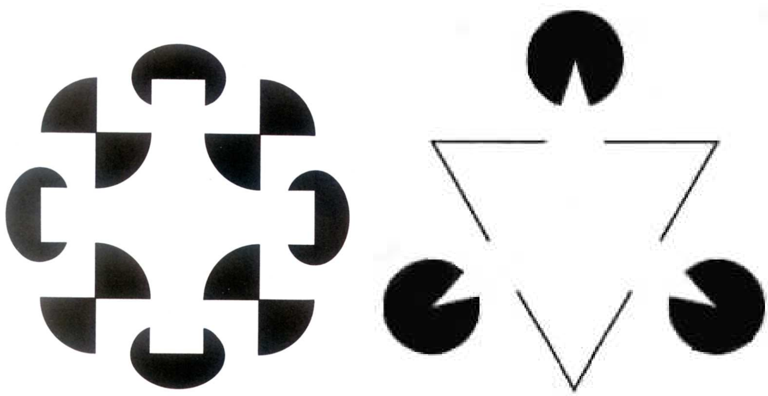


Figure 1.1: Images proposed by Kanitza

In Section 5 we give an introduction of differential calculus in Lie groups define an uniformly sub-Riemannian operator, and its time-dependent counterpart. Then we show that these operator can model the propagation of the visual signal in the cortex.

In Section 6 we study the regular surfaces of the structure and prove that the neural mechanism of non maxima suppression generates regular surfaces in the cortical space.

Finally in Section 7 we prove that the two mechanisms of propagation of the visual signal, and non maxima suppression, generates a diffusion driven motion by curvature. The perceptual completion is then obtained through a minimal surface. Hence we will study its regularity and foliation properties.

1.2 perceptual completion phenomena

Gaetano Kanizsa in [51, 52] provided a taxonomy of perceptual completion phenomena and outlined that they are interesting test to understand how the visual system interpolates existing information and builds the perceived units.

He discriminated between modal completion and amodal completion. In the first one the interpolated parts of the image are perceived with the full modality of the vision and are phenomenally undistinguishable from real stimuli (this happens for example in the formation of illusory contours and surfaces). In amodal presence the configuration is perceived without any sensorial counterpart. Amodal completion is evoked every time one reconstructs the shape of a partially occluded object. Thus it is at the base of the most primitive perceptual configuration that is the segmentation of figure and ground. Mathematical models of perceptual completion take into account main phenomenological properties as described by psychology of Gestalt.

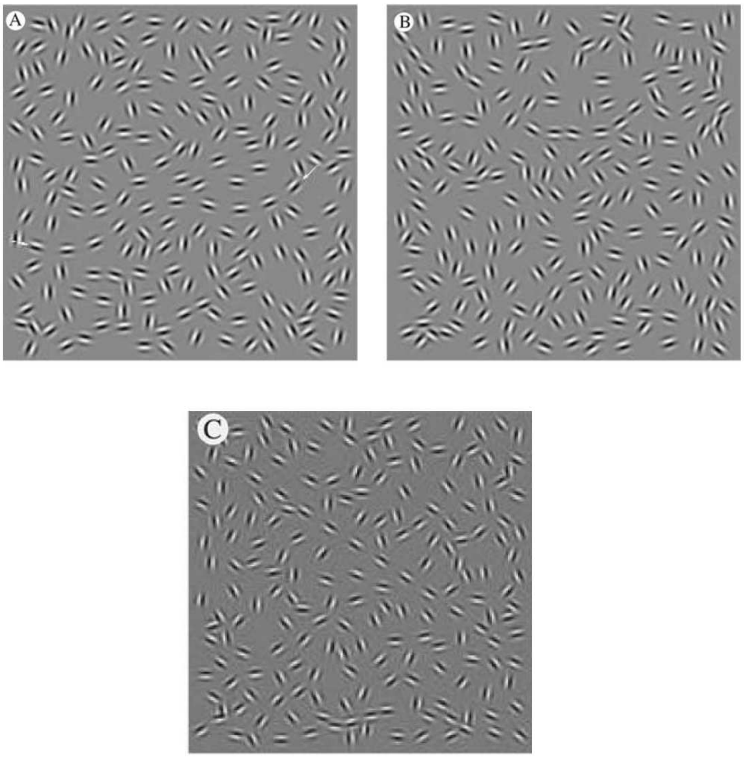


Figure 1.2: The experiment of Field, Heyes and Hess

1.2.1 Associations fields

The history of studies on contour integration is a long one, stretching back to the Gestalt psychologists who formulated rules for perceptually significant image structure, including contour continuity: the Gestalt law of good continuation. Field, Hayes and Hess [47] developed a new approach to psychophysically investigating how the visual system codes contour continuity by using contours of varying curvature made up of spatial frequency narrowband elements. The contour stimulus is shown in Fig. 1.2. Within a field of evenly spaced, randomly oriented, Gabor elements, a subset of the elements is aligned in orientation and position along a notional contour (Fig. 1.2 A). This stimulus is paired with an similar stimulus (Fig.

1.2 B), where all of the elements are unaligned (called the background elements). The observer was asked to recognize structures and alignments in the stimulus, and to discriminate the two stimula. From a simple informational point of view Fig. 1.2A and B are equivalent, so a difference in their detectability reflects the ability of human observers to detect the contour and constraints imposed by the visual system. In particular it is interesting to note that contours composed of elements whose local orientation was orthogonal to the contour are far less detectable.

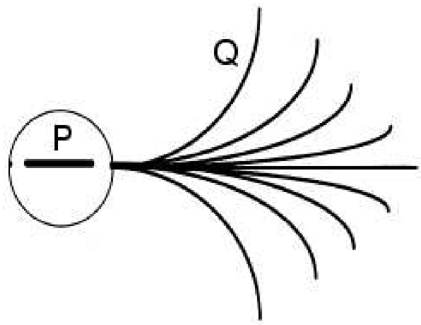


Figure 1.3: Association fields

Another finding of this study was the human ability to detect increasingly curved contours. A good performance for contour detection was possible even in presence of curvature of the contour, suggesting that the output of cells with similar, but not necessarily equal orientation preference are being integrated. Fig. 1.2 C shows another stimulus manipulation that reinforces the notion that the task of contour integration reflects the action of a network rather than that of single neurons interaction. Here the polarity of every other Gabor element is flipped. The contour (and background) is now composed of Gabor elements alternating in their contrast polarity. The visibility of the contour in Fig. 1.2 A and C is similar. Psychophysical measurement shows that although there is a small decrement in performance in the alternating polarity condition, curved contours are still readily detectable when composed of elements of alternating polarity.

This model of cellular interaction and contour completion has been summarized by Field Hayes and Hess in terms of an association field which is depicted in Fig. 1.3. The stimulus in the central position can be jointed with other stimula tangent to the lines in figure, but can not be joined with stimula with different direction.

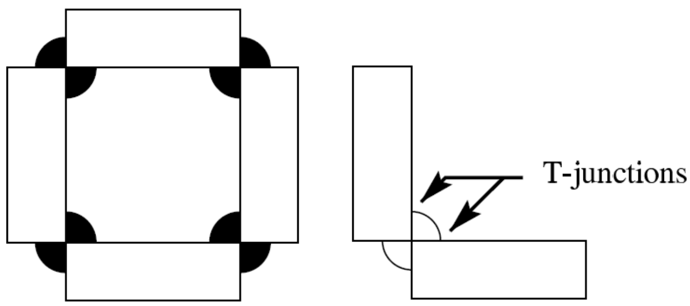


Figure 1.4: an example of T-junction

1.2.2 Higher order operators and elastica

Since subjective boundaries could be linear or curvilinear, their reconstruction is classically performed minimizing the elastica functional

$$\int_\gamma (1 + k^2) ds, \tag{1.1}$$

where the integral is computed on the missed boundary, and k is its curvature (see [67]). The minimum of the elastica functional is taken on all the curves with fixed endpoints and with fixed directions at the endpoints. It appears that continuation of objects boundaries plays a central role in the disocclusion process. This continuation is performed between T-junctions, which are points where image edges intersect orthogonally as illustrated in Figure 1.4.

In [67] Nitzberg, Mumford and Shiota deduced from the amodal completion principles a method for detecting and recovering occluded objects in a still image within the framework of a segmentation and depth computing algorithm.

Approximation in the sense of Γ convergence by elliptic functionals have been proposed by De Giorgi in [32] (the conjecture is still open). Bellettini and Paolini [11] proposed and proved a new approximation, of Modica Mortola type. They also proved that functional (1.1) does not allow non regular completion, which on the contrary can occur (see Figure 1.5) and propose to modify the functional, with a new functional

$$\int_\gamma (1 + \phi(k^2)) ds. \tag{1.2}$$

When ϕ has linear growth at the origin and behave as a square root at infinity, completion with kinks is allowed.

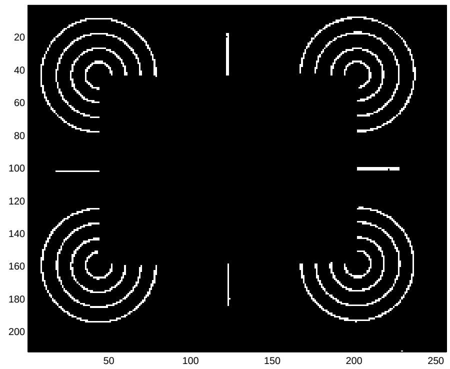


Figure 1.5: non regular completion

The extension of the elastica functional to the level set of the image I, has been applied in problems of inpainting (that can be considered a particular case of modal completion) by [58], [1] :

$$\int_\Omega |\nabla I|\Big(1+\Big|div\Big(\frac{\nabla I}{|\nabla I|}\Big)\Big|^2\Big)dx, \quad x\in\Omega\subset\mathbb{R}^2 \tag{1.3}$$

where the integral is extended to the domain of the image. In this way each level line of the image is completed either linearly or curvilinearly as elastica curve.

In order to make occluded and occluding objects present at the same time in the image, in [67] (and then in [10], [34]) a third dimension is introduced, and the objects present in the image are represented as a stack of sets, ordered by depth. In [77] the third added dimension is represented by the time, and the algorithm first detects occluding objects, then occluded ones. In [6] the associated evolution equation was split in two equations, each one of the first order, and depending on two different variables: the image I, and the direction of its gradient $\nu = \nabla I/|\nabla I|$.

1.3 The functional structure of the visual cortex

From the neurophysiological point of view the acquisition of the visual system is performed in the retina that, after a preprocessing, projects the information to the lateral geniculate nucleus and to the primary visual cortex in which signal is deeply processed. In particular the primary visual cortex V1 process the orientation of contours by means of the so called simple cells and other features of the visual signal by means of complex cells (stereoscopic vision, estimation of motion direction, detection of angles.). Every cell is characterized by its receptive field,

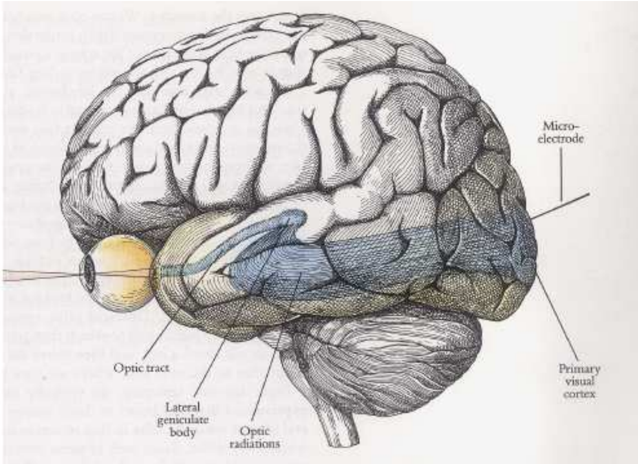


Figure 1.6: The visual path

that's the domain of the retinal plane to which the cell is connected with neural synapses of the retinal-geniculate-cortical path. When the domain is stimulated by a visual signal the cell respond generating spikes.

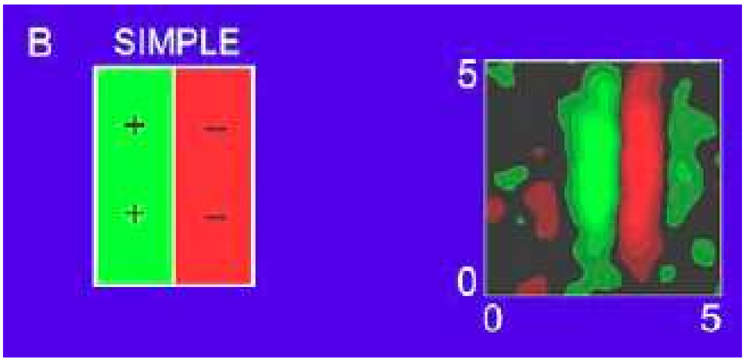


Figure 1.7: receptive profiles

Classically a receptive profile is subdivided in "on" and "off" areas. The area is considered "on" if the cell spikes responding to a positive signal and "off" if it spikes responding to a negative signal. The receptive profile is mathematically described by a function Ψ_0, defined on the retinal plane. This function models the neural output of the cell in response to a punctual stimulus in the 2 dimensional point x. Simple cells have directional receptive profiles as it is shown in Figure 1.7 and they are sensitive to the boundaries of images.

To understand the processing of the image operated by these cells, it is necessary to consider the functional structures of the primary visual cortex: the retinotopic organization, the hypercolumnar structure with intracortical circuitry and the connectivity structure between hypercolumns.

1.3.1 The retinotopic structure

The retinotopic structure is a mapping between the retina and the primary visual cortices that preserves the retinal topology and it is mathematically described by a logarithmic conformal mapping. From the image processing point of view, the retinotopic mapping introduces a simple deformation of the stimulus image that will be neglected in the present study.

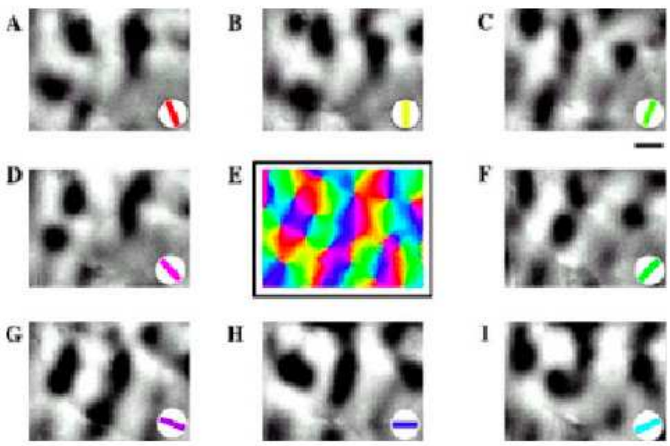


Figure 1.8: Representation of Bosking. Wihin an hypercolumn the cells sensible to different orientations is represented in different colours.

1.3.2 The hypercolumnar structure

The hypercolumnar structure organizes the cortical cells in columns corresponding to parameters like orientation, ocular dominance, color etc. For the simple cells (sensitive to orientation) columnar structure means that to every retinal position is associated a set of cells (hypercolumn) sensitive to all the possible orientations. The visual cortex is indeed two-dimensional and then the third dimension collapses onto the plane giving rise to the fascinating pinwheels configuration observed by William Bosking et al. With optical imaging techniques. In Figures 1.8 the orientation preference of cells is coded by colors and every hypercolumn is represented by a pinwheel.

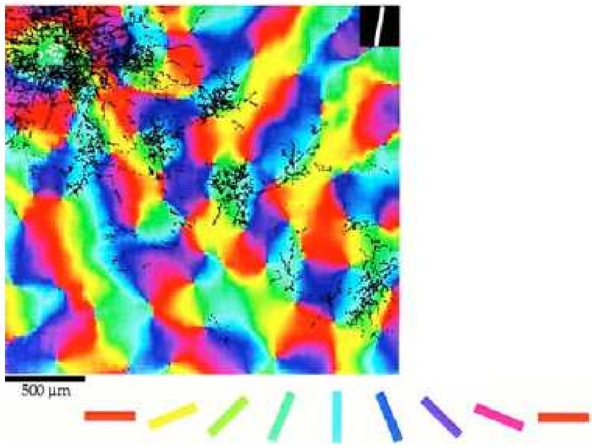


Figure 1.9: A marker is injected in the cortex, in a specific point, and it diffuses mainly in regions with the same orientation as the point of injection (marked with the same color in figure).

1.3.3 The neural circuitry

The intracortical circuitry is able to select the orientation of maximum output of the hypercolumn in response to a visual stimulus and to suppress all the others. The mechanism able to produce this selection is called non-maximal suppression or orientation selection, and its deep functioning is still controversial, even if many models have been proposed (see [59, 73, 65]).

The connectivity structure, also called horizontal or cortico-cortical connectivity is the structure of the visual cortex which ensures connectivity between hypercolumns. The horizontal connections connect cells with the same orientation belonging to different hypercolumns. Historically correlation techniques have been used to estimate the relation between connectivity and preferred orientation of cells [81]. Only recently techniques of optical imaging associated to tracers allowed a large-scale observation of neural signal propagation via cortico-cortical connectivity. These tests have shown that the propagation is highly anisotropic and almost collinear to the preferred orientation of the cell (see figure 1.9 and the study of Bosking [16]). It is already confirmed that this connectivity allows the integration process, that is at the base of the formation of regular and illusory contours and of subjective surfaces [71]. Obviously the functional architecture of the visual cortex is much richer of the schemata we have delineated, just think to the high percentage of feedback connectivity from superior cortical areas, but for now we will try to propose a model of low level vision, aiming to mathematically model correctly the functional structures we have described and able to show that theses are at the base of perceptual completion of contours.

1.4 The visual cortex as a Lie group

1.4.1 A first model in the Heisenberg group

Petitot and Tondut in [71] proposed a new approach to the problem, which is particularly interesting because the perceptual completion problem is considered as a problem of naturalizing phenomenological models on the basis of biological and neurophysiological evidence. Let us recall here their model

Retinotopic and (hyper)columnar structure

The main structures of the cortex: retinotopic and (hyper)columnar can be modeled as follows.

- The retinotopy means that there exist mappings from the retina to the cortical layers which preserve retinal topography. If we identify the retinal structure with a plane R the retina and by M the cortical layer, the retinotopy is then described by a map $q : R \to M$ which is an isomorphism. Hence we will identify the two planes, and call M both of them.
- The columnar and hypercolumnar structure organizes the cells of $V1$ in columns corresponding orientation. Due to their RP they detect preferred orientations, that is points (x, u) where x denote a 2 dimensional (retinal) position and u denotes the direction of a boundary of an image mapped on the retina at the point x.

  The hypercolumnar organization means essentially that to each position x of the retina there exists a full fibre of possible orientations u at x.

Contour detection an lifting

Formally at a retinal point $x = (x_1, x_2)$, we consider edges of images as regular curves of the form

$$x_2 = f(x_1).$$

The orientation at the point x is then $u = f'(x_1)$. The tangent line to the considered edge at the point x has the expression

$$X_u = \partial_1 + u(x_1)\partial_2. \tag{1.4}$$

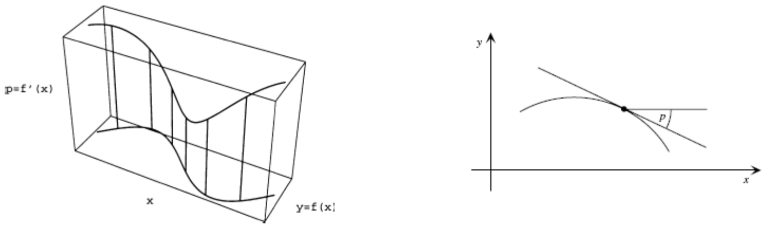


Figure 1.10: curves lifted in the cortical contact structure

In presence of the visual stimulus all the hypercolumn over the point x is activated, and the simple cell sensible to the direction u has the maximal response. The retinal point x is lifted to the cortical point (x, u), the whole curve is then lifted to the curve

$$(x_1, f(x_1), u(x_1))$$

in a 3-dimensional space R^3 endowed with the constraint $f' = u$. Formally this is a constraint on the tangent space TR^3 at every point. We can define a $1-$ form

$$\omega = dx_2 - udx_1,$$

and note that all the lifted curves lie in the kernel of ω. This formal constraint can be expressed saying that we consider a subset of the tangent plane, kernel of the $1-$form ω,

$$HT = \{\alpha X_1 + \beta X_2\},$$

where

$$X_1 = \partial_1 + u\partial_2, \quad X_2 = \partial_u. \tag{1.5}$$

The lifted curves have to be integral curves of the vector fields X_1, X_2.

1.4.2 A subriemannian model in the rototraslation group

The previous model can describe only images with equi-oriented boundaries. This can be easily overcame in the $E(2)$- group of motion of the plane. In [28] we recognize the previously described structure as a subriemannian structure. Besides we will focus on level lines representation, instead of edge detection. Indeed if $I(x)$ is a gray level image, the family of level lines is a complete representation of I, from which I can be reconstructed. This model is compatible with the functionality of the simple cells and their orientation sensitivity.

Lifting in $EO(2)$ - a purely perceptual description We now consider a real stimulus, represented as an image I. We can assume that cells over each point x can code the direction of the level lines of I, without a preferred direction. Hence the eingrafted variable in the hypercolumn will be an angle, and we will assume that the cell which give the maximal response is sensible to the direction $\theta(x) = -arctan(I_1/I_2)$, $\theta \in [0, \pi]$. This means that the vector field

$$X_\theta = cos(\theta(x))\partial_1 + sin(\theta(x))\partial_2 \tag{1.6}$$

is tangent to the level lines of I at the point x. As before this process associates to each retinal point x the three dimensional cortical point $(x, \theta) \in R^2 \times S^1$. Since the process is repeated at each point, each level line is lifted to a new curve in the three dimensional space. The tangent vector to the lifted curve can be represented

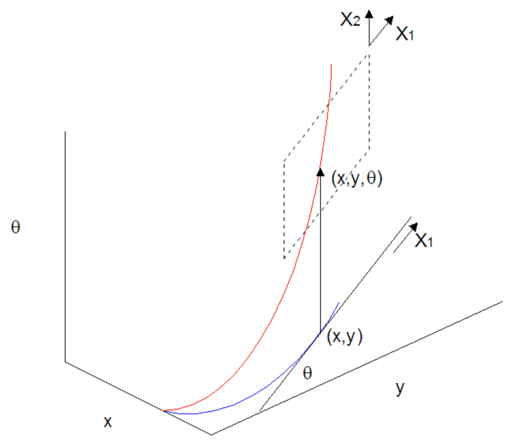


Figure 1.11: a lifted surface, foliated in lifted curves

as a linear combination of the vectors

$$X_1 = cos(\theta)\partial_1 + sin(\theta)\partial_2 \quad X_2 = \partial_\theta. \tag{1.7}$$

The set of vectors

$$a_1X_1 + a_2X_2$$

defines a plane and every lifted curve is tangent to a vector of the plane.

The lifting process - a neurophisiological description Neural evidence supports this model of the cortex. When a visual stimulus of intensity $I(x)$ activates

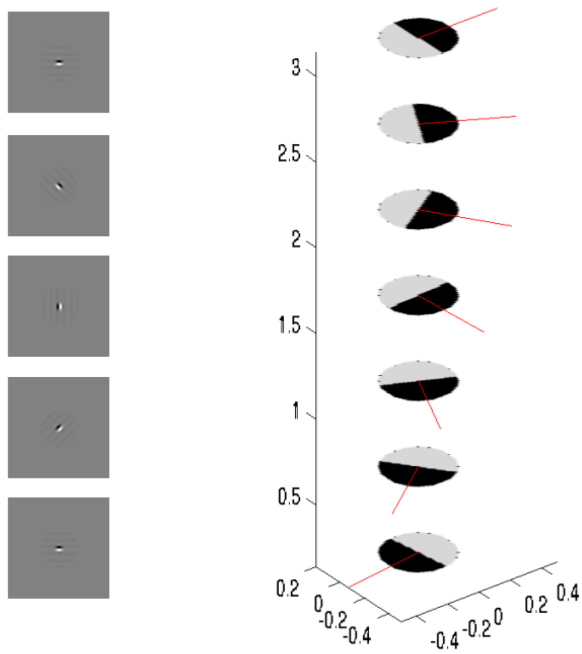


Figure 1.12: Odd part of Gabor filters with different orientations (left) and Schemata of odd simple cells arranged in a hypercolumn of orientations.

the retinal layer of photoreceptors $M \subset \mathbb{R}^2$, the cells centered at every point x of M process in parallel the retinal stimulus with their receptive profile which is a function defined on M.

Each RP depends upon a preferred direction θ and it has been observed experimentally that the set of simple cells RPs is obtained via translations and rotations from a unique profile, of Gabor type (see for example Jones and Palmer [50], Daugman [31], Marcelja [57]). This means that there exists a mother profile Ψ_0 from which all the observed profiles can be deduced by rigid transformation.

A good formula for Ψ_0 seems to be (see Figure 1.13 and compare with Figure 1.7)

$$\Psi_0(x) = \partial_2 e^{-|x|^2}.$$

Therefore by rotation all the observed profiles can be modeled as

$$\Psi_{x,\theta}(\widetilde{x}, \theta)(x) =$$

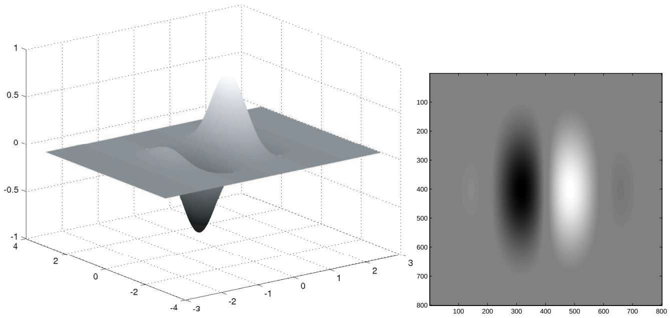


Figure 1.13: the shape of the Gabor filter and a schematic representation of it - compare with the in vivo registration - Figure 1.7

$$= \Psi_0\Big(\widetilde{x}_1 - x_1) \cos\theta + (\widetilde{x}_2 - x_2) \sin\theta, -(\widetilde{x}_1 - x_2) \sin\theta + (\widetilde{x}_2 - x_2) \cos\theta\Big).$$

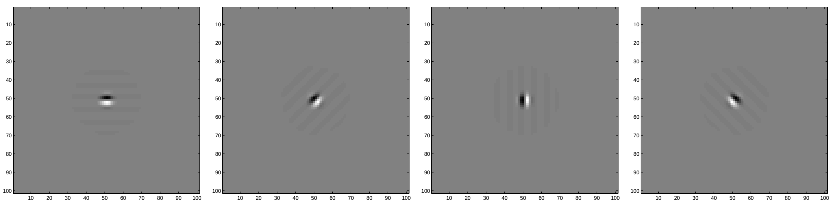


Figure 1.14: Odd part of Gabor filters with different orientations $\theta = 0$, $\theta = \pi/4$, $\theta = \pi/2$,$\theta = 3/2\pi$

In the rotation of an angle θ, the derivative ∂_2 becomes

$$X_3 = -\sin(\theta)\partial_1 + \cos(\theta)\partial_2. \tag{1.8}$$

Hence

$$\Psi_\theta(x) = X_3 e^{-|x|^2}.$$

With this notation the filtering can be described as the convolution with the image I and generates a function

$$O(x, \theta) = -X_3 exp(-|x|^2) * I = -X_3(\theta) I_s \tag{1.9}$$

where we have denoted I_s the convolution of I with a smoothing kernel:

$$I_s = I * exp(-|x|^2).$$

This function O is the output of the cells, and measure their activity. Note that $O(x, \theta)$ depends on the orientation θ. Due to the expression of the Gabor filter, the function O exponentially decays from its maxima. Hence for θ fixed it selects a neighborhood of the points where the component of the gradient in the direction $(-sin(\theta), cos(\theta))$, is sufficiently big (see Figure 1.15).

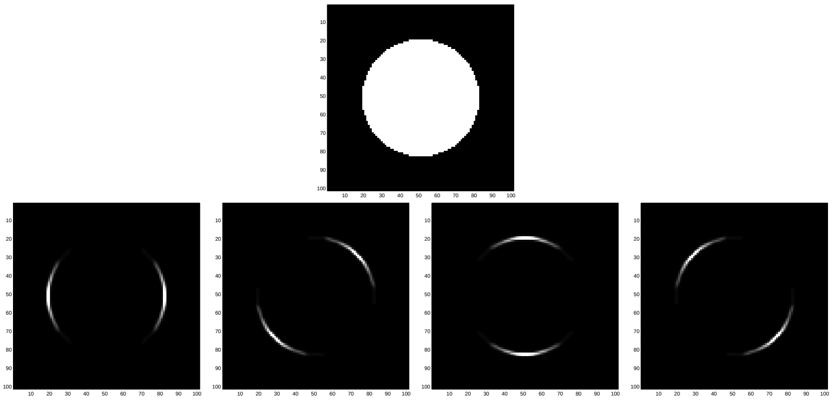


Figure 1.15: The original image showing a white disk (upper) and a sequence of convolutions with different orientations Gabor filters.

The convolution mechanism (1.9) is insufficient to explain the strong orientation tuning exhibited by most simple cells. For these reasons, the classic feed-forward mechanism must be integrated with additional mechanisms, in order to provide the sharp tuning experimentally observed. The basic mechanism is controversial and in the past years several models have been presented to explain the emergence of orientation selectivity in the primary visual cortex: ("push-pull" models [59, 73], "emergent" models [65], "recurrent" models [80] only to cite a few). Nevertheless it is evident that the intracortical circuitry is able to filter out all the spurious directions and to strictly keep the direction of maximum response of the simple cells.

We will then define

$$O(x, \bar{\theta}) = \max_{\theta} O(x, \theta).$$

This maximality condition can be mathematically expressed requiring that the derivative of O with respect to the variables θ vanishes at the point $(x, \bar{\theta})$:

$$\partial_\theta O(x, \bar{\theta}) = 0.$$

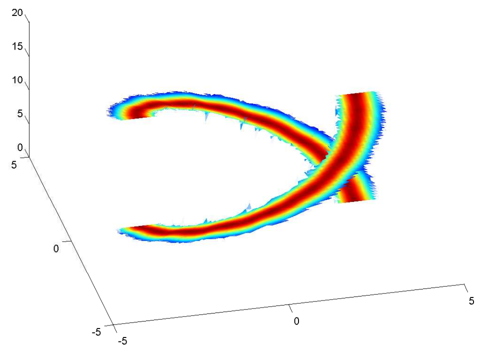


Figure 1.16: the resulting surface after non maximal suppression, called lifted surface (right).

At the maximum point $\bar\theta$ the derivative with respect of θ vanishes, and we have

$$0 = \frac{\partial}{\partial\theta} O(x, \bar\theta) = \frac{\partial}{\partial\theta} X_3(\bar\theta) I = -X_1(\bar\theta) I = - < X_1(\bar\theta), \nabla I > .$$

As a direct consequence we can deduce that the lifted curves are tangent to the plane generated by the vector X_1 and X_2.

1.4.3 Hörmander vector fields and Sub-Riemannian structures.

In the standard Euclidean setting, the tangent space to R^n has dimension n at every point. In the geometric setting arising from the model of the cortex the dimension of the space is 3, but we have selected at every point a 2 dimension subspace of the tangent space, and verified that all admissible curves are tangent to this subspace at every point. We will see that these are examples of sub-Riemannian structures.

In general we will denote ξ the points in $\mathbb{R}^n$, and we will choose m first order smooth differential operators

$$X_j = \sum_{k=1}^{n} a_{jk} \partial_k \qquad j = 1 \cdots m,$$

in $\mathbb{R}^n$ with $m < n$ and a_{jk} of class C^∞. We will call Horizontal tangent space at the point $\xi \in R^n$ the vector space $HH_{|\xi}$ spanned by these vector fields at the point ξ. The distribution of planes defined in this way is called horizontal tangent bundle and it is a subbundle of the tangent one. A differential operator X is called horizontal, if it belongs to the horizontal bundle HH.

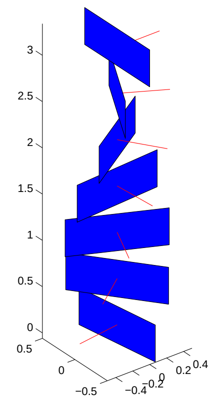


Figure 1.17: The contact planes at every point, and the orthogonal vector $X_3 I$

Definition 1.4.1. We will call horizontal norm, and horizontal scalar product and denote them respectively $\langle \cdot, \cdot \rangle_H$ and $|\cdot|_H$ the scalar product and the norm, defined on the Horizontal bundle which makes the basis $X_1, \ldots, X_m$ an orthonormal basis.

The Horizontal tangent bundle is naturally endowed with a structure of algebra, through the bracket.

Definition 1.4.2. If X, Y are first order regular differential operators their commutator (or bracket) is defined as

$$[X, Y] = XY - YX,$$

and it is a first order differential operator. We call Lie algebra generated by $X_1, \cdots, X_m$ and denote it as

$$\mathcal{L}(X_1, \cdots, X_m)$$

the linear span of the operators $X_1, \cdots, X_m$ and their commutators of any order.

We will say that the vectors

$$X_1 \cdots X_m \text{ have degree } 1$$

$$[X_i, X_j] \text{ have degree } 2,$$

and define in an analogous way higher order commutators.

Example 1. In general the degree is not unique. Indeed, if we consider the vector fields introduced in (1.7), the vector X_1 has degree 1, but it also have degree 3, since in that specific example $X_1 = -[X_2, [X_2, X_1]]$.

Hence we will call minimum degree of $X_j \in \mathcal{L}(X_1, \cdots, X_m)$, and denote it

$$deg(X_j) = min\{i : X_j \text{ has degree } i\}.$$

Since $m < n$, in general

$$\mathcal{L}(X_1, \cdots, X_m)$$

will not coincide with the Euclidean tangent plane. If these two spaces coincide, we will say that the Hörmander condition is satisfied:

Definition 1.4.3. Let $\Omega \subset \mathbb{R}^n$ be an open set, and let (X_j), with $j = 1, \cdots, m$ be a family of smooth vector fields defined on Ω. If the condition

$$rank(\mathcal{L}(X_1, \cdots, X_m))(\xi) = n,$$

for every $\xi \in \mathbb{R}^n$ is satisfied we say that the vector fields $(X_j)_{j=1..m}$ satisfy the Hörmander rank condition.

If this condition is satisfied, at every point ξ we can find a number s such that $(X_j)_{i=1..m}$ and their commutators of degree smaller or equal to s span the space at ξ. If s is the smallest of such natural numbers, we will say that the space has step s at the point ξ. At every point we can select a basis $\{X_j : j = 1 \cdots n\}$ of the space made out of commutators of the vector fields $\{X_j : j = 1 \cdots m\}$. In general the choice of the basis will not be unique, but we will choice a basis such that for every point

$$Q = \sum_{j=1}^{n} deg(X_j) \tag{1.10}$$

is minima. The value of Q is called local homogeneous dimension of the space. In general it is not constant, but by simplicity in the sequel we will assume that

$$s \text{ and } Q \tag{1.11}$$

are constant in the considered open set. This assumption is always satisfied in a Lie group.

Example 2. The simplest example of family of vector fields is the Euclidean one: $X_i = \partial_i \; i = 1 \cdots m$ in $\mathbb{R}^n$. If $m = n$, then the Hörmander condition is satisfied while it is violated if $m < n$.

Example 3. Let us consider the family of vector fields introduced in (1.5). In that example the point of $\mathbb{R}^3$, are denoted $\xi = (x_1, x_2, u)$ and

$$X_1 = \partial_1 + u\partial_2 \quad X_2 = \partial_u.$$

Since

$$[X_1, X_2] = -\partial_2,$$

then the Hörmander condition is satisfied.

Example 4. In (1.7) we denote $\xi = (x_1, x_2, \theta)$ a point in $\mathbb{R}^2 \times S^1$ and denote

$$X_1 = \cos(\theta)\partial_1 + \sin(\theta)\partial_2, \quad X_2 = \partial_\theta$$

the generators of the Lie algebra. The commutator is

$$X_3 = [X_2, X_1] = -\sin(\theta)\partial_1 + \cos(\theta)\partial_2,$$

which is linearly independent of X_1 and X_2.

1.4.4 connectivity property

If X is a smooth first order differential operator, $X = \sum_{k=1}^n a_k \partial_k$ and I is the identity map $I(\xi) = \xi$, then it is possible to represent the vector field with the same components as the differential operator X in the form

$$XI(\xi) = (a_1, \cdots a_n).$$

Sometimes the vector and the differential operator are identified, but we will keep them distinct here for reader convenience.

We will call integral curve of the vector field XI starting at ξ_0 a curve γ such that

$$\gamma' = XI(\gamma), \quad \gamma(0) = \xi_0$$

the curve will also be denoted

$$\gamma(t) = \exp(tX)(\xi_0).$$

If X is horizontal we will call Horizontal curves its integral curves.

The Carnot Carathéodory distance in the space, is defined in terms of horizontal integral curves, in analogy with the well known Riemannian distance. Since in the subriemannian setting we will allow only integral curves of horizontal vector fields, we need to ensure that it is possible to connect any couple of points ξ and ξ_0 through an horizontal integral curve.

Theorem 1.4.4. Chow theorem *If the Hörmander condition, is satisfied, then any couple of points in $\mathbb{R}^n$ can be joint with a piecewise C^1 horizontal curve.*

Let us postpone the proof after a few examples of vector fields satisfying the connectivity condition. We will consider the same examples as before

Example 5. In the Euclidean case considered in example 2, section 1.4.3, if $m = n$, then the Hörmander condition is satisfied, and it is clear that any couple of points can be joint with an Euclidean integral curve. If $m < n$, when the Hörmander condition is violated, it is clear that also the connectivity condition fails. Indeed if we start from the origin, with an integral curve of the vectors $X_i = \partial_i$ $i = 1 \cdots m$, we can reach only points with the last $n - m$ identically 0.

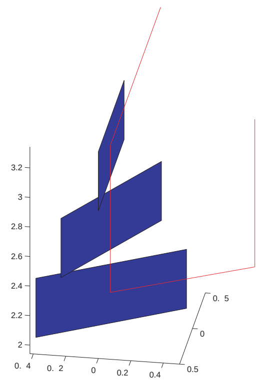


Figure 1.18: piecewise constant integral curves of the structure

Example 6. In the example 3 section 1.4.3, the Hörmander condition is satisfied. On the other side, it is easy to see that we can connect any point (x, u) with the origin through a piecewise regular horizontal curve. Indeed we can call $\widetilde{u} = x_2/x_1$, consider the segment $[(0,0),(0,\widetilde{u})]$, which is an integral curve of X_2. Then the segment $[(0,\widetilde{u}),(x,\widetilde{u})]$ is an integral curve of X_1. Finally the segment $[(x,\widetilde{u}),(x,u)]$ is an integral curve of X_2.

Example 7. We already verified that the vector fields described in example 4 section 1.4.3, satisfy the Hörmander condition. On the other hand also in this case it is possible to verify directly that any couple of points can be connected by a piecewise regular path (see Figure 1.18).

We follow the approach of [15] of the proof of Chow theorem. It is based on the following lemma:

Lemma 1.4.5. *Let X be of class C^2, then the following estimation holds:*

$$C(t)(\xi) = e^{-tY}e^{-tX}e^{tY}e^{tX}(\xi) = \xi + t^2(YX - XY)I(\xi) + o(t^2) = \tag{1.12}$$

$$exp(t^2[X,Y](\xi) + o(t^2))(\xi).$$

If the coefficients of the vector field X be of class C^h, we can define inductively

$$C(t, X_1, \cdots X_h)(\xi) = e^{-tX_1}C(t, -X_2, \dots X_m)e^{tX_1}C(t, X_2, \dots X_h)(\xi) \tag{1.13}$$

In this case we have:

$$C(t, X_1, \cdots X_h) = exp(t^h[[[[X_1, X_2]\cdots X_h] + o(t^h))(\xi).$$

Proof Let us prove the first assertion. The Taylor expansions ensures that

$$e^{tX}(\xi) = \xi + tXI(\xi) + \frac{t^2}{2}X^2I(\xi) + o(t^2).$$

Also note that, by definition of Lie derivative

$$YI(\xi+tXI(\xi)+\frac{t^2}{2}X^2I(\xi)+o(t^2)) = YI(e^{tX}(\xi)+o(t^2)) = YI(\xi)+tYXI(\xi)+o(t).$$

Hence

$$e^{tY}e^{tX}(\xi) = e^{tY}(\xi + tXI(\xi) + \frac{t^2}{2}X^2I(\xi) + o(t^2)) =$$

$$= \xi + tXI(\xi) + \frac{t^2}{2}X^2I(\xi) + tYI(\xi + tXI(\xi) + o(t)) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) =$$

$$= \xi + tXI(\xi) + \frac{t^2}{2}X^2I(\xi) + tYI(\xi) + t^2XYI(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) =$$

$$= \xi + t(XI(\xi) + YI(\xi)) + \frac{t^2}{2}(X^2I(\xi) + 2XYI(\xi) + Y^2I(\xi)) + o(t^2).$$

Applying e^{-tX} we obtain

$$e^{-tX}e^{tY}e^{tX}(\xi) = \xi + tYI(\xi) + \frac{t^2}{2}(2[X,Y]I(\xi) + Y^2I(\xi)) + o(t^2).$$

Finally

$$e^{-tY}e^{-tX}e^{tY}e^{tX}(\xi) = \xi + t^2[X,Y]I(\xi) + o(t^2).$$

The second assertion can be proved by induction, using the same ideas.

Proof of connectivity property We make the choice of basis described in (1.10), and assume that

$$X_i = [X_{j_1}[\cdots[X_{j_i}]].$$

for suitable indices j_i.

Let us call

$$C_i(t) = C(t^{1/deg(X_i)}, X_{j_1}\cdots X_{j_i}). \tag{1.14}$$

By the previous lemma

$$\frac{d}{dt}C_i(t)_{|t=0} = X_i.$$

Now for every $e \in \mathbb{R}^n$ $\xi \in \Omega$ we define

$$C_p(e)(\xi) = \prod_{i=1}^{n} C_i(e_i)(\xi). \tag{1.15}$$

The Jacobian determinant of C_p with respect to e is the determinant of X_i. So that it is different from 0. Hence the map $C_p(e)$ is a local homeomorphism, and the connectivity property is locally proved. A connectness argument conclude the proof.

1.4.5 Control distance

If the connectivity property is satisfied, it is possible to give the definition of distance of the space. We have chosen the Euclidean metric on the contact planes, so that we can call length of any horizontal curve γ

$$\lambda(\gamma) = \int_0^1 |\gamma'(t)|dt.$$

Consequently we can define a distance as:

$$d(\xi, \xi_0) = \inf\{\lambda(\gamma) : \gamma \text{ is an horizontal curve connecting } \xi \text{ and } \xi_0\}. \tag{1.16}$$

Parameterizing the curve by arc length we deduce

$$d(\xi, \xi_0) = \inf\{T : \gamma' = \sum_{j=1}^{m} e_j X_j, \gamma(0) = \xi_0, \gamma(T) = \xi, \sqrt{\sum_{j=1}^{m} |e_j|^2} = 1\} =$$

$$= \inf\{T : \gamma' = \sum_{j=1}^{m} e_j X_j, \gamma(0) = \xi_0, \gamma(T) = \xi, \sqrt{\sum_{j=1}^{m} |e_j|^2} \leq 1\}.$$

As a consequence of Hörmander condition we can represent any vector in the form

$$X = \sum_{j=1}^{n} e_j X_j.$$

The norm $\sqrt{\sum_{j=1}^{m} |e_j|^2}$ is the horizontal norm defined in Definition 1.4.1. We can extend it as a homogeneous norm on the whole space setting:

$$||e|| = (\sum_{j=1}^{n} |e_j|^{Q/deg(X_j)})^{1/Q}, \tag{1.17}$$

where Q has been defined in (1.10).

Since the exponential mapping is a local diffeomorphism, we will define

Definition 1.4.6. If $\xi_0 \in \Omega$ is fixed, we define canonical coordinates of ξ around a fixed point ξ_0, the coefficients e such that

$$\xi = exp(\sum_{j=1}^{n} e_j X_j)(\xi_0).$$

These representation will be used to give an other characterization of the distance

Proposition 1.4.7. *[64] The distance defined in (1.16) is locally equivalent to*

$$d_1(\xi, \xi_0) = ||e||,$$

where e are the canonical coordinates of ξ around ξ_0 and $||.||$ is the homogeneous norm, defined in (1.17).

1.4.6 Riemannian approximation of the metric

In Definition 1.4.1 we introduced an horizonal norm only on the horizontal tangent plane. We can extend it to a Riemannian norm all the tangent space as follows: for every $\epsilon > 0$ we define

$$X_j^\epsilon = X_j \quad j = 1 \cdots m \tag{1.18}$$
$$X_j^\epsilon = \epsilon X_j \quad j > m.$$

The family X_j^ϵ $j = 1 \cdots n$ formally tends to the family X_j $j = 1 \cdots m$ as $\epsilon \to 0$. We call Riemannian approximation of the metric g the Riemannian metric g_ε which makes the vector fields orthonormal. Clearly g_ε restricted to the horizontal plane coincide with the Horizontal metric. The geodesic distance associated to g_ε is denoted d_ε, while the ball in this metrics of center ξ_0 and radius r will be denoted

$$B_\varepsilon(\xi_0, r) = \{\xi : d_\varepsilon(\xi, \xi_0) < \varepsilon\}. \tag{1.19}$$

The distance d_ε tends to the distance d defined in (1.16) as ε goes to 0. We refer to [18] and the references therein for a complete treatment of this topic.

1.4.7 geodesics and elastica

The curve which minimize the distance is called geodesics. We refer to the book of Montgomery [60] for reference to this topic. We do study this problem here but we only recognize the relation between geodesics of $EO(2)$, and elastica. A 2D curve

$$\widetilde{\gamma} = x(t)$$

can be represented in arc length coordinates

$$x'(t) = (\cos(\theta(t)), \sin(\theta(t)))$$

at every point, where θ denotes the direction of the curve at the point $x(t)$. In section 1.4.2 we lifted it to a 3D curve $\gamma(t) = (x(t), \theta(t))$. By the properties of the arch length parametrization

$$\theta' = k,$$

where k is the Euclidean curvature of $\widetilde{\gamma}$.

The length of the lifted curve is:

$$\int \sqrt{x'^2 + \theta'^2} = \int \sqrt{x'^2}\sqrt{1 + k^2}.$$

We see that the length of γ is the elastica functional, evaluated on $\widetilde{\gamma}$. In this sense this model can be considered as a neurological motivation of the existing higher order models of modified elastica (see section 1.2.2).

1.5 Activity propagation and differential operators in Lie groups

1.5.1 Integral curves, Association fields, and the experiment of Bosking

Let us go back to the problem of the description of the cortex. Up to now we have built up a geometric space inspired by the functional geometry of the primary visual cortex. Let us focus on the model in the group $EO(2)$. In the cortex neural activity develops and propagates itself in this-subriemannian space. For seek of simplicity, in this study we consider an extremely simple model of activity propagation, i.e. a simple linear diffusion along the integral curves of the structure.

This integrative process allows to connect local tangent vectors to form integral curves and is at the base of both regular contours and illusory contours formation [71].

This countour formation has been described by the association field (Field [47]). The anatomical network of horizontal long-range connections has been proposed as the implementation of association fields, and the experiments of Bosking (see section 1.3.3) prove that the diffusion of a marker in the cortex are in perfect agreement with the association fields.

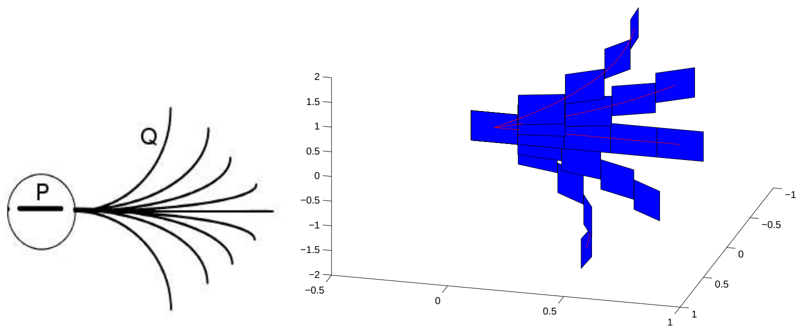


Figure 1.19: the association fields and the integral curves of the subriemannian structure

We propose to interpret these lines as a family of integral curves of the generators of the $EO(2)$, the vector fields X_1 and X_2, starting at a fixed point $\xi = (x, \theta)$:

$$\gamma'(t) = X_1 I(\gamma) + k X_2 I(\gamma), \quad \gamma(0) = (x, \theta), \tag{1.20}$$

obtained by varying the parameter k in $\mathbb{R}$ (fig. (1.19)).

Long-range connections can consequently be modeled as admissible curves with piecewise constant coefficients k.

1.5.2 Differential calculus in subriemannian setting

In order to describe the diffusion of the visual signal we need to introduce the main instruments of differential calculus in a subriemannian setting.

Definition 1.5.1. Let X be a fixed vector field we call Lie derivative of f in the direction of the vector X on the tangent space to $\mathbb{R}^n$ at a point ξ the derivative with respect to t in $t = 0$ of the function $f \circ exp(tX)(\xi)$.

Clearly if f is C^1, then the Lie derivative coincides with directional derivative, but the Lie derivative can exist even though the directional derivatives does not exist.

Definition 1.5.2. Let $\Omega \subset \mathbb{R}^n$ be an open set, let (X_j), $j = 1 \cdots m$ be a family of smooth vector fields defined on Ω, and let $f : \Omega \to \mathbb{R}$. If there exist $X_j f$ for every $j = 1 \cdots m$ we call horizontal gradient of a function f

$$\nabla_H f = (X_1 \cdots X_m).$$

A function f is of class C^1_H if $\nabla_H f$ is continuous, with respect of the distance defined in (1.16). A function f is of class C^2_H is $\nabla_H f$ is of class C^1_H, and by induction all C^k_H classes are defined.

Note that a C^1_H function is not differentiable with respect to X_j if $j > m$. It follows that a function of class C^1_H is not of class C^1_E, in the standard Euclidean sense. If the vector fields (X_j), $j = 1 \cdots m$ have step s, a function f of class C^s_H is C^1_E.

Remark 1.5.3. If the vector fields (X_j), $j = 1 \cdots m$ satisfy the Hörmander condition, f is C^∞_H if and only if is a function is of class C^∞_E in a standard sense

Remark 1.5.4. The Heisenberg group, and the group $EO(2)$, with the choice of vector fields made in examples 3 and 4 section 1.4.3, are of step 2. Hence, if a function f is of class C^k_H in one of these structures it is of class $C^{k/2}_E$ in the standard sense.

From the definition of Lie derivative, and the properties of integral curve, the following result follows:

Proposition 1.5.5. *Let $\Omega \subset \mathbb{R}^n$, let X and Y be horizontal vector fields defined on Ω and let $f : \Omega \to \mathbb{R}$. Assume that at every point in Ω there exist $Xf(\xi)$ and $Yf(\xi)$, and these derivatives are continuous. If $\gamma(t) = exp(tX)(exp(tY)(\xi))$, then there exists*

$$(f \circ \gamma)'(0) = Xf(\xi) + Yf(\xi).$$

Proof

$$\frac{1}{t}\Big(f(\gamma(t)) - f(\gamma(0))\Big) =$$

$$= \frac{1}{t}\Big(f(exp(tX)(exp(tY)(\xi)) - f((exp(tY)(\xi)))\Big) +$$

$$+\frac{1}{t}\Big(f(exp(tY)(\xi)) - f((\xi)))\Big) =$$

by the mean value theorem

$$Xf(exp(t_1X)(exp(tY)(\xi))) + Yf(exp(t_2Y)(\xi)))$$

$$\to Xf(\xi) + Yf(\xi)$$

as $t \to 0$. □

From the previous proposition we immediately deduce the corollary:

Remark 1.5.6. If C is the function defined in Lemma 1.4.5,

$$C(t) = exp(-tY)exp(-tX)exp(tY)exp(tX)(\xi),$$

and $f \in C^1_H(\Omega)$, then there exists

$$\frac{d}{dt}(f \circ C)(0) = 0.$$

Proposition 1.5.7. *Let $\Omega \subset \mathbb{R}^n$, and assume that on Ω is defined a family of vector fields (X_j) $j = 1 \cdots m$, satisfying the Hörmander condition (see Definition 1.4.3). If f is of class $C^1_H(\Omega)$, then*

- *f is continuous in Ω*
- *if C_p is the function defined in (1.15), the function f satisfies*

$$f(C_p(e)(\xi)) - f(\xi) = \sum_{j=1}^{m} e_j X_j + o(||e||),$$

where $||.||$ is the homogeneous norm defined in (1.17).

The second assertion is a direct consequence of the previous remark and proposition, together with the definition of C_p. The fact that f is continuous follows from the fact that C_p is a local diffeomorphism (see the proof of connectivity).

Proposition 1.5.8. *Let $\Omega \subset \mathbb{R}^n$, let $f : \Omega \to \mathbb{R}$ be a continuous function such that there exist the Lie derivatives Xf and Yf and they are continuous functions. Then there also exists $(X+Y)f = Xf + Yf$ in Ω.*

Proof Arguing as in Lemma 1.4.5, we immediately see that

$$|exp(tX)exp(tY)(\xi) - exp(t(X+Y))(\xi)| = O(t^2),$$

locally uniformly in ξ. It follows that

$$\frac{1}{t}\Big(f(exp(t(X+Y))(\xi)) - f(\xi)\Big) =$$

$$= \frac{1}{t}\Big(f(exp(tX)(exp(tY)(\xi)) - f((\xi)))\Big) + O(t)$$

$$\to Xf(\xi) + Yf(\xi),$$

as $t \to 0$ by Proposition 1.5.5.

Definition 1.5.9. Let Ω be an open set in $\mathbb{R}^n$, and assume that on Ω is defined a family of vector fields X_j $j = 1 \cdots m$, satisfying the Hörmander condition. A function $f : \Omega \to \mathbb{R}$ is differentiable at a point $\xi \in \Omega$ in the intrinsic sense if

$$f(\sum_{j=1}^{n} \exp(e_j X_j)(\xi)) - f(\xi) = \sum_{j=1}^{m} e_j X_j f(\xi) + o(||e||)$$

as $||e|| \to 0$. Note that only vector fields of degree 1 appear in the definition.

As a direct consequence of the previous propositions we have:

Proposition 1.5.10. *Let $\Omega \subset \mathbb{R}^n$, and assume that on Ω is defined a family of vector fields X_j $j = 1 \cdots m$, satisfying the Hörmander condition. If f is of class $C^1_H(\Omega)$, then it is differentiable.*

The previous result implies in particular that,

Remark 1.5.11. Let $\Omega \subset \mathbb{R}^n$ be an open set and let $f \in C^1_H(\Omega)$.
If $\gamma(t) = exp\Big(\sum_{j=1}^n t^{deg(X_j)} e_j X_j\Big)(\xi_0)$, then

$$\exists \lim_{t \to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t} = \sum_{j=1}^{m} e_j X_j f(\xi_0)$$

locally uniformly on Ω and with respect to e.

1.5.3 Subriemannian differential operators

Definition 1.5.12. If $\phi = (\phi_1 \cdots \phi_m)$ is a C^1_H section of the horizontal tangent plane, we call divergence of ϕ

$$div_H(\phi) = \sum_{j=1}^{m} X_j^* \phi_j,$$

where X_j^* is the formal adjoint of the vector field X_j.

From now on we will assume that for every j the vector fields

$$X_j \text{ is self adjoint }, \tag{1.21}$$

and denote X_j^* the adjoint operator of X_j.

Accordingly we will define Sublaplacian operator as

$$\Delta_H = div_H(\nabla_H).$$

An uniformly subelliptic operator minic the structure of uniformly elliptic operators. An $m \times m$ matrix (A_{ij}) is an uniformly elliptic matrix, is there exist two real numbers λ, Λ such that

$$\lambda |\xi|^2 \leq \sum_{j=1}^{m} A_{ij} \xi_i \xi_j \leq \Lambda |\xi|^2.$$

Accordingly the operator

$$L_A = \sum_{ij=1}^{m} A_{ij} X_i X_j \tag{1.22}$$

is called uniformly subelliptic.

We will define subcaloric equation, the natural analogous of the heat equation, expressed in terms of the subelliptic operator:

$$\partial_t = L_A.$$

Example 8. Note that the solution of a sum of squares of 2 vector fields in $\mathbb{R}^3$ is not in general regular. Indeed any function of the variable ξ_3 is a solution of

$$\partial_1^2 + \partial_2^2 = 0 \quad \text{in } \mathbb{R}^3.$$

Theorem 1.5.13. Hörmander theorem *If $X_1 \cdots X_m$ satisfy the Hörmander rank condition, then the associated supelliptic operator and the heat operator are hypoelliptic operators.*

These operators admit a fundamental solution Γ, of class C^∞. Existence and local estimates of the fundamental solution it terms of the control distance have been first proved by Folland Stein [39] Rothshild Stein [76], Nagel, Stein, Weinger[64].

Precisely they proved that fundamental solution can be locally estimated as

$$|\Gamma(\xi, \xi_0)| \leq C \frac{d^2(\xi, \xi_0)}{|B(x, d(\xi, \xi_0))|},$$

for every ξ, ξ_0 in a neighborhood of a fixed point, and for a suitable constant C. Gaussian estimates, local and global of the fundamental solution have been investigated by many authors. We refer to the book [14] for an exaustive presentation of the topic.

In the application to the cortex it is necessary to study elliptic regularization of this type of operators. This means that the vector fields X_j will be replaced by the vectors X_j^ε, introduced in (1.18). The matrix A_{ij} will be extended to a $n \times n$ matrix A_{ij}^ε uniformly elliptic. Then the riemannian approximating operator of the operator (1.22) is

$$L_\varepsilon = \sum_{ij=1}^{n} A_{ij}^\varepsilon X_i^\varepsilon X_j^\varepsilon. \tag{1.23}$$

This operator is clearly uniformly elliptic in Ω, but the ellipticity constant tends to $+\infty$ with ϵ, since the limit operator is not elliptic. On the contrary for the fundamental solution of this operator it is possible to prove subelliptic estimates uniform in ε (see [26]).

Theorem 1.5.14. *For every compact set $K \subset \Omega$ and for every choice of vector fields in the basis $X_{j_1}^\epsilon \cdots X_{j_k}^\epsilon$ there exist two positive constants C, C_k independent of ε such that for every $\xi, \xi_0 \in K$ with $\xi \neq \xi_0$,*

$$|X_{j_1}^\epsilon \cdots X_{j_k}^\epsilon \Gamma_\varepsilon(\xi, \xi_0)| \leq C_k \frac{d_\varepsilon^{2-k}(\xi, \xi_0)}{|B_\varepsilon(\xi, d_\varepsilon(\xi, \xi_0))|}, \tag{1.24}$$

where $B_\varepsilon(\xi, r)$ is the ball in the approximating riemannian metic defined in (1.19).
$\square$

This theorem provides uniform estimates of fundamental solution of an operator, in terms of its control distance. Letting ε goes to 0, it allows to deduce from regularity results known in the elliptic case, analogous results for the subelliptic situation. In general this approach allows to work with smooth solutions of an elliptic problem $L_\varepsilon u_\varepsilon = f$ in order to obtain uniform estimates for the limit equation.

A first consequence of this result is the regularity in the intrinsic Sobolev spaces Let $\Omega_0 \subset \Omega$, and $W_\varepsilon^{k,p}(\Omega_0)$ be the set of functions $f \in L^p(\Omega_0)$ such that

$$X_{i_1}^\varepsilon \cdots X_{i_k}^\varepsilon f \in L^p(\Omega_0), \quad i_1, \ldots, i_k \in \{1, \ldots, n\},$$

with natural norm

$$||f||_{W^{k,p}_\varepsilon(\Omega_0)} = \sum_{i_1,\dots,i_k\in\{1,\dots,n\}} ||X^\varepsilon_{i_1}\dots X^\varepsilon_{i_k} f||_{L^p(\Omega_0)}.$$

Let us make some example of applications. Assume that Q is the homogeneous dimension of the limit operator. Then the following Sobolev type inequality holds:

Corollary 1.5.15. *If $u \in W^{1p}_\varepsilon$ and is compactly supported in an open set Ω, then there exist a constant C independent of ε such that*

$$||u||_{L^r(\Omega)} \leq C||u||_{W^{k,p}_\varepsilon(\Omega)}$$

where $r = Qp/(Q - kp)$.

Corollary 1.5.16. *Assume that $u \in L^q_{loc}(\Omega)$ is a solution of*

$$L_\varepsilon u = f \ in\ \Omega,$$

with $f \in W^{p,q}_{\varepsilon,X}(\Omega)$ and let $K_1 \subset\subset K_2 \subset\subset \Omega$. Then there exists a constant C independent of ε such that

$$||u||_{W^{p+2,q}_{\varepsilon,X}(K_1)} \leq C||f||_{W^{p,q}_{\varepsilon,X}(K_2)},$$

for every $p \geq 1$.

1.6 Regular surfaces in subriemannian setting

1.6.1 Maximum selectivity and lifting images to regular surfaces

The mechanism of non maxima suppression does not lift each level lines independently, but is applied to the whole image. If O is the output of the simple cells, the maximum of O over the fiber is taken:

$$|O(x,\bar\theta)| = max_\theta |O(x,\theta)|. \tag{1.25}$$

In this process each point x in the 2D domain of the image is lifted to the point $(x,\bar\theta(x))$, and the whole image domain is lifted to the graph of the function $\bar\theta$:

$$\Sigma = \{(x,\theta) : \theta = \bar\theta(x)\}. \tag{1.26}$$

This lifted set corresponds to the maximum of activity of the output of the simple cells. Setting $f(x,\theta) = \partial_\theta O(x,\theta)$, and considering only strict maxima are considered the surface becomes:

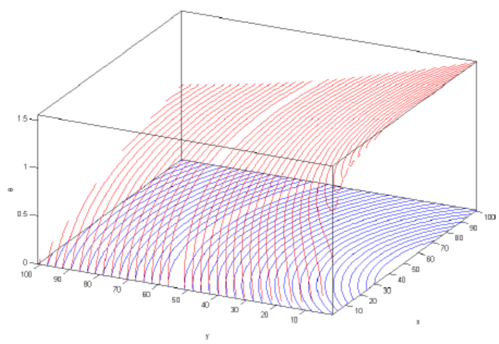


Figure 1.20: lifting of level lines of an image

$$\Sigma = \{(x,\theta) : f(x,\theta) = 0, \partial_\theta f(x,\theta) > 0\}. \tag{1.27}$$

where the vector ∂_θ is an horizontal vector.

We recall that on the domain of $\bar{\theta}$ only one vector field was defined (see (1.6)):

$$X_{\bar{\theta}} = \cos(\bar{\theta}(x))\partial_1 + \sin(\bar{\theta}(x))\partial_2 \tag{1.28}$$

tangent to the level lines of I.

We will see that Σ is a regular surface in the subriemannian structure, and that in any subriemannian structure the implicit function $\bar{\theta}$ is regular with respect to non linear vector fields, depending on $\bar{\theta}$.

1.6.2 Definition of regular surface

In this setting the notion of regular surface in not completely clear. The first definition, given by Federer in [36] was that a regular surface is the image of a open set of R^{n-1} through a lipschitz continuous function. However the Heisenberg group turn out to be completely non rectifiable in this sense ([3]). A more natural definition of regular surface has been given by Franchi Serapioni and Serracassano and investigated in a long series of papers: [40, 41, 42, 43, 44].

Definition 1.6.1. A regular surface is a subset Σ of $\mathbb{R}^n$ which can be locally represented as the zero level set of a function $f \in C^1_H$ such that $\nabla_H f(\xi) \neq 0$. If the vector $\nabla_H f(\xi)$ vanishes at a point ξ, this point is called characteristic. If it does not vanish, we define intrinsic normal of Σ

$$\nu_H = \frac{\nabla_H f(\xi)}{|\nabla_H f(\xi)|}.$$

In other words the vector ν_H takes the place of normal vector in this setting. It can be recovered through a Blow up procedure similar to the De Giorgi method for the Euclidean proof of rectifiability. We refer to [40] for the proof in the Heisenberg setting and to [24] for the proof in general setting.

Example 9. The generators of the Heisenberg algebra introduced in (1.5) are

$$X_1 = \partial_1 + u\partial_2 \quad X_2 = \partial_u.$$

in R^3, whose points are denoted $\xi = (x, u)$. The plane

$$u = 0$$

has as intrinsic normal

$$\nu_H = \partial_u.$$

The intrinsic normal of the plane $y = 0$ is

$$(u, 0) = uX_1.$$

Hence the point $(x_1, 0, 0)$ are characteristic for this plane.

Example 10. We provide an example of characteristic surface in the group $EO(2)$, defined in example 4, in section 1.4.3. The points of the space will be denoted (x, θ) as before. Let us denote $\widetilde{\gamma}$ a curve in the plane x and let us consider the surface

$$\Sigma = \{(x, \theta) : x \in \widetilde{\gamma},\ \theta \in [0, 2\pi]\}.$$

In section 1.4.7 we pointed out that the lifting of the curve $\widetilde{\gamma}$ is a new curve γ, whose tangent vector is $X_1 + kX_2$, where k is the Euclidean curvature of $\widetilde{\gamma}$. Hence at every point of the lifted curve γ the surface Σ has two horizontal tangent vectors: $X_1 + kX_2$ and X_2. Consequently all these point are characteristic.

1.6.3 Implicit function theorem

Regular surface in this setting are not regular in the Euclidean sense. An example of intrinsic regular surface, which has a fractal structure has been provided by [53]. However a first proof of the Dini theorem for hypersurfaces have been given by [40] in the Heisenberg group. A much simpler proof in a general subriemannian structure has been proved in [25]. Indeed, due to the structure of the vector fields, the implicit function u found in [40] is not a graph in standard sense. The problem is related to the fact that the definition of graph is not completely intrinsic, but it assigns a different role to the first variable, lying in the image of u with respect to the other $n-1$ variables, belonging to the domain of u.

Hence we choose a suitable change of variables. In the new variables it is possible to represent the generators of the Lie algebra in the following way:

$$X_1 = \partial_1, \quad X_j = \sum_{k=2}^{n} a_{jk}(\xi)\partial_k, \;\; j = 2, \ldots, m. \tag{1.29}$$

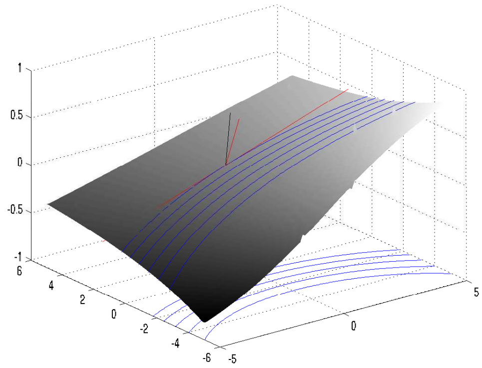


Figure 1.21: a regular surfaces, foliated in horizontal curves.

Let us note that the explicit expression of the vector fields appearing in the model of the cortex is of this type.

In order to simplify the notations we will represents the points of the space in the form

$$\xi = (\xi_1, x)$$

where $\xi_1 \in \mathbb{R}$, $x \in \mathbb{R}^{n-1}$.

In these new variables from the classical implicit function theorem we immediately deduce the following

Lemma 1.6.2. *Let $\Omega \subset \mathbb{R}^n$ be an open set. Let $0 \in \Omega$ and $f \in C^1_X(\Omega)$ be such that*

$$\partial_1 f(0) > 0, \quad f(0) = 0.$$

If

$$\Sigma = \{\xi \in \Omega : \, f(\xi) = 0\},$$

then there exist neighborhoods of 0 $I \subset \mathbb{R}^{n-1}$, $J \subset \mathbb{R}$ and a continuous function $u : I \to J$ such that

$$\Sigma \cap (J \times I) = \{(u(x), x) : \, x \in I\}.$$

Proof The existence of the function u is standard. We recall here only the proof of the continuity of u in order to point out that in this part of the proof we only need the continuity of the derivative $\partial_1 f$, which here is continuous by assumption, since it is horizontal.

$$0 = f(u(x), x) - f(u(x_0), x_0) =$$

$$= f(u(x), x) - f(u(x_0), x) + f(u(x_0), x) - f(u(x_0), x_0) =$$

(by the mean value theorem)

$$\partial_1 f(s, x)(u(x) - u(x_0)) + f(u(x_0), x) - f(u(x_0), x_0).$$

Then

$$|u(x) - u(x_0)| = |\frac{f(u(x_0), x) - f(u(x_0), x_0)}{\partial_1 f(s, x)}| = o(1)$$

since the denominator is bounded away from 0 by assumption, and f is continuous. □

In order to study the regularity of the function u we will need to project on its domain the vector fields X_j. To this end we define a projection on $\mathbb{R}^n$:

$$\pi(\xi) = x,$$

and a projection on its tangent plane:

$$\pi_u(\sum_{k=1}^{n} a_k(\xi)\partial_k) = \sum_{k=2}^{n} a_k(u(x), x)\partial_k.$$

Accordingly we will define

$$X_{ju} = \pi_u(X_j).$$

In particular

$$X_{1u} = 0,$$

and, since $X_j = \sum_{k=2}^{n} a_{jk}\partial_k$ their projection will be

$$X_{ju} = \sum_{k=2}^{n} a_{jk}(u(x), x)\partial_k. \tag{1.30}$$

Definition 1.6.3. Let $I \subset P$ be an open set. We say that a continuous function $u : I \to \mathbb{R}$ is of class $C^1_u(I)$ if for every $x \in I$

$$\exists X_{ju}(x), \quad \text{for} \quad j = 2, \dots, m$$

and they are continuous. We will call intrinsic gradient

$$\nabla_u u = (X_{2u}u, \cdots, X_{mu}u).$$

Theorem 1.6.4. *If the assumptions of lemma 1.6.2 are satisfied, the implicit function u is of class $\mathbb{C}^1_u$, and*

$$\nabla_u u(x_0) = -\frac{(X_2 f(\xi_0), \cdots, X_m f(\xi_0))}{\partial_1 f(\xi_0)}.$$

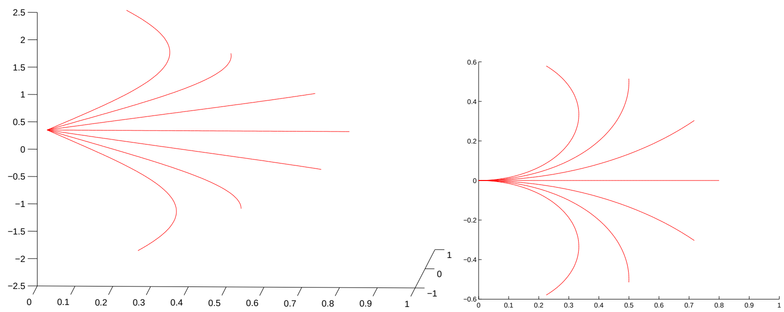


Figure 1.22: integral curves of the vector fields and their $n-1$D projection

Proof Let us consider the vector X_{ju}, a point x_0 and let us call $\gamma_u(t) = exp(tX_{ju})(x_0)$. We also call $\gamma(t) = exp(tX_j)(u(x_0), x_0)$ and $\gamma_\pi(t) = \pi(exp(tX_j)(u(x_0), x_0))$. Then by definition of Σ,

$$0 = f(u(\gamma_u(t), \gamma_u(t)) - f(u(\gamma_u(0)), \gamma_u(0)) =$$

$$f(u(\gamma_u(t)), \gamma_u(t)) - f(u(\gamma_u(0)), \gamma_u(t)) +$$

$$+f(u(\gamma_\pi(0)), \gamma_u(t)) - f(u(\gamma_\pi(0)), \gamma_\pi(t)) + f(u(\gamma_\pi(0)), \gamma_\pi(t)) - f(\gamma(0)) =$$

by the classical mean value theorem there exist y and c such that

$$= \partial_1 f(c, \gamma_u(t))\Big(u(\gamma_u(t)) - u(\gamma_\pi(0))\Big) +$$

$$+\partial_1 f(u(\gamma_\pi(0)), y)\Big(\gamma_u(t)) - \gamma_\pi(t)\Big) - (f \circ \gamma)(t) - (f \circ \gamma)(0).$$

(note that the curve γ has the first component constant, so that $\gamma(t) = (u(\gamma_\pi(0)), \gamma_\pi(t))$. Dividing by t and letting t go to 0 we obtain:

$$0 = \partial_1 f(\xi_0) X_{ju} u(x_0) + X_j f(\xi_0).$$

Then

$$X_{ju} u(x_0) = -\frac{X_j f(\xi_0)}{\partial_1 f(\xi_0)}.$$

$\square$

1.6.4 Non regular and non linear vector fields

A consequence of the Dini Theorem is the fact that, if we start with a regular surface of class C^1_H, its implicit function u is differentiable with respect to nonlinear the vector fields (X_{ju}). This open a large spectrum of problems, since these new vector fields are non regular, and in general satisfy conditions different from the initial vector fields.

Let us make some examples:

Example 11. Let us now consider an Heisenberg group of higher dimension. This is R^5, with the choice of vector fields

$$X_1 = \partial_1 \quad X_2 = \partial_2 + \xi_1\partial_5 \quad X_3 = \partial_3 \quad X_4 = \partial_4 + \xi_3\partial_5 \quad \in R^5.$$

Since

$$[X_3, X_4] = \partial_5, \tag{1.31}$$

then these vector fields satisfy the Hörmander rank condition. The associated non linear vector fields will be obtained setting

$$x = (\xi_2, \cdots \xi_5),\ \xi_1 = u(x).$$

We then obtain the following vectors, in the tangent space to $\mathbb{R}^4$:

$$X_{2u} = \partial_2 + u\partial_5 \quad X_3 = \partial_3 \quad X_4 = \partial_4 + \xi_3\partial_5 \quad \in R^4$$

It is clear that, if u is smooth, these are Hörmander vector fields, by condition (1.31). However in general the solution u will be only C^1_u, and the difficulty in handling these vectors are the lack or regularity. We will say that a weak Hörmander condition is verified.

In this situation there is reasonable hope to prove Poicaré inequalities, estimates of fundamental solution, and mimic in this non regular situation results known in the smooth setting. A first a Poincaré inequality for non regular vector fields have been established in [54]. After that such an inequality of this type has been proved in [61] for vector fields of class C^2 and step 2. A similar inequality requires C^{s+1} regularity for vector fields of step s. [17], [62]. Very recently a Poincaré inequality for Heisenberg non linear vector fields of class C^1 has been proved by Manfredini in [55]. From this a Sobolev inequality with optimal exponent follows. Estimates for the fundamental solution for non linear vector fields have been proved in [56].

Example 12. In the case of the Heisenberg group of dimension 1, (see example 3 in section 1.4.3, we have a Lie algebra with 2 generators in a 3D space. The vector fields X_1, X_2 projected on the plane x, reduce to only one vector field:

$$X_{1u} = \partial_1 + u(x)\partial_2. \tag{1.32}$$

In this case we have an unique non linear vector field in R^2. It is clear that this vector field does not satisfy the Hörmander condition, not even when u is smooth

The same thing happens in the group $EO(2)$, where the unique vector field defined on $\mathbb{R}^2$ is

$$X_{1u} = cos(u(x))\partial_1 + sin(u(x))\partial_2.$$

In this low dimensional case, a few results are known only in the Heseinberg group. More recently Ambrosio, Serra Cassano, Vittone gave a characterisation of implicit functions in [4], while Bigolin, Serra Cassano, started the study of the set of C^1_u functions in [13]. In this case there is no hope to prove an estimate of the fundamental solution, of linear operators defined in terms of non linear vector fields. For these operators the riemannian approximation can be extremely useful. Indeed using the estimate of the approximating fundamental solution, Citti Capogna Manfredini proved a Sobolev estimate for the linearized operator:

$$\sum_{ij} A_{ij} X^{\varepsilon}_{ju} X^{\varepsilon}_{ju} z = 0, \tag{1.33}$$

where A_{ij} is positive defined,

$$X^{\varepsilon}_{1u} = X_{1u}, \quad X^{\varepsilon}_{2u} = \varepsilon X_{2u}, \quad \nabla^{\varepsilon}_{u} = (X^{\varepsilon}_{1u}, X^{\varepsilon}_{2u}). \tag{1.34}$$

The result in [20] reads as follows:

Theorem 1.6.5. *Let us assume that z is a classical solution of the approximated problem (1.33): where u is a smooth function. Assume that there exists a constant C independent of ε such that*

$$||A_{ij}||_{C^{\alpha}(K)} + ||u||_{C^{1,\alpha}(K)} + ||\partial_2 z||_{L^p(K)} + ||\partial_2 X_u z||_{L^q(K)} + ||(\nabla^{\varepsilon}_u)^2 z||_{L^2(K)} \leq C.$$

Then for any compact set $K_1 \subset\subset K$, there exists a constant C_1 only dependent on K, C, such that

$$||z||_{W^{2,r}_{\varepsilon}(K_1)} \leq C_1,$$

where $r = min(5q/(5-(1+\alpha q)), 5p/(5-\alpha p))$.

The proof is based on the estimates of the fundamental solution uniform in ε stated in Theorem 1.5.14. The exponent r is reminiscent of a Sobolev exponent, modeled on a homogeneous dimension $Q = 5$. However it is not optimal, since the coefficients are not regular.

1.7 Completion and minimal surfaces

1.7.1 A Completion process

The joint work of subriemannian diffusion (Section 1.5) and non maximal suppression (Section 1.6) allows to propagate existing information and then to complete boundaries and surfaces. Starting from the lifted surface the two mechanisms are simultaneously applied until the completion is reached. To take into account the simultaneous work of diffusion and non maximal suppression we consider iteratively diffusion in a finite time interval followed by non maximal suppression, and we compute the limit when the time interval tends to 0.

The algorithm is an extension of the diffusion driven motion by curvature introduced by J. Bence, B. Merriman, S. Osher in [12]. It is described by induction as follows: given a function u_n, whose maxima in a given direction are attained on a surface Σ_n, we diffuse in an interval of length h

$$v_t = \Delta_H v, \quad v_{t=0} = v_{\Sigma_n} t \in [nh, (n+1)h] \tag{1.35}$$

At time $(n+1)h$ the solution defines a new function v_{n+1}, and we built a new surface, through the non maxima suppression.

$$\Sigma_{n+1}((n+1)h) = \{\partial_{\nu_{\Sigma_n}} v_{n+1} = 0, \partial^2_{\nu_{\Sigma_n}} v_{n+1} < 0\}$$

If we fix a time T, we can choose intervals of length $h = T/(n+1)$, and we get the two sequences: $v_{n+1}(\cdot, T)$, $\Sigma_{n+1}(T)$. We expect the convergence of the two sequences $\Sigma_n(T)$ and $u_n(T)$ respectively to mean curvature flow $\Sigma(T)$ of the surface Σ_0 and the Beltrami flow on Σ. For $T \to +\infty$ the function $\Sigma(T)$ should converge to a minimal surface in the rototraslation space, in the sense that its curvature identically vanishes.

The formal proof of the convergence of diffusion driven motion by curvature in the Euclidean setting is due to Evans [35] and G.Barles, C. Georgelin, [8]. The proof of the analogous assertion in this context is still work in progress. Indeed we have a preliminary result regarding the motion by curvature [19]. We are now attacking the problem of convergence of the motion by curvature to the minimal surface equation.

By now we have studied properties of minimal surfaces and verified that they have the properties required by the completion model.

1.7.2 Minimal surfaces in the Heisenberg group

Several equivalent notions of *horizontal mean curvature* H_0 for a regular C^2_H surface $M \subset \mathbb{H}^1$ (outside characteristic points) have been given in the literature. To quote a few: H_0 can be defined in terms of the first variation of the area functional [29, 48, 22, 75, 79, 63] as horizontal divergence of the horizontal unit normal. As such the expression of the curvature of a surface level set of a function f becomes:

$$H_0 f = \sum_{j=1}^{m} X_j \left(\frac{X_j f}{|\nabla_H f|} \right). \tag{1.36}$$

A different, but equivalent notion of curvature, in term of a notion of a metric normal has been given by [5]. In [18] it has been recognized that the curvature can be obtained as limit of the mean curvatures H_ε in the Riemannian metrics g_ε, defined in section 1.4.6. The definition of H_ε can be given in terms of the vector fields X_j^ε defined in (1.18) as follows:

$$H_\varepsilon = \sum_{j=1}^{m} X_j^\varepsilon \left(\frac{X_j^\varepsilon f}{|\nabla_\varepsilon f|} \right). \tag{1.37}$$

Here ∇_ε denotes the approximated gradient

$$\nabla_\varepsilon = (X_1^\varepsilon \cdots X_n^\varepsilon).$$

In the particular case of intrinsic graphs it can be expressed in terms of the vector fields (X_{ju}) defined in section 1.6. As we already noted the regularity theory for intrinsic minimal surfaces is completely different if a weak Hörmander type condition is satisfied or not. In $\mathbb{H}^n$ with $n > 1$ this condition is satisfied and the problem has been afforded in [21].

Hence here we focus on the low dimensional case, which naturally arises from the application to the visual cortex. By simplicity we restrict to the monodimensional Heisenberg group. The extension to general Lie algebras with two generators, step 2 and dimension 3 is due to [7]. Through the implicit function theorem we have defined in (1.32) an unique vector field X_{1u} on $\mathbb{R}^2$.

The curvature operator for intrinsic graphs reduces to:

$$X_{1u} \left(\frac{X_{1u} u}{\sqrt{1 + |X_{1u} u|^2}} \right) = f, \text{ for } x \in \Omega \subset \mathbb{R}^2. \tag{1.38}$$

Properties of regular minimal surfaces have been studied in [46], [68], [22], [23], [45], [30], [9] and [66]. The Riemannian approximating vector fields have been defined in (1.34), while the Riemannian approximating operator reads:

$$L_\varepsilon u = \sum_{i=1}^{2} X_{iu}^\varepsilon \left(\frac{X_{iu}^\varepsilon u}{\sqrt{1 + |\nabla_u^\varepsilon u|^2}} \right) = f, \text{ for } x \in \Omega \subset \mathbb{R}^2, \tag{1.39}$$

Using this approximation, we can give the definition of vanishing viscosity solution

Definition 1.7.1. If C_E^1 denotes the standard Euclidean C^1 norm, we will say that an Euclidean Lipschitz continuous function u is a *vanishing viscosity solution* of (1.38) in an open set Ω, if there exists a sequence $\epsilon_j \to 0$ as $j \to +\infty$, and a sequence (u_j) of smooth solutions of (1.39) in Ω such that for every compact set $K \subset \Omega$

- $||u_j||_{C_E^1(K)} \leq C$ for every j;
- $u_j \to u$ as $j \to +\infty$ pointwise a.e. in Ω.

Existence of viscosity solutions has been proved by J. H. Cheng, J. F. Hwang, P. Yangin in [23], while the problem of regularity of minimal surfaces has been afforded in [20]. We present here the regularity result

Theorem 1.7.2. *The Lipschitz continuous vanishing viscosity solutions of* (1.38) *are intrinsically smooth functions.*

This theorem highlight a very general idea: any positive semi-definite operator of second order regularizes in the direction of its positive eigenvalues. However, in general, this does not imply smoothness of solutions, since regularity can be expected only in the directions of the non vanishing eigenvalues. Indeed the following foliation result holds for minimal graphs:

Corollary 1.7.3. *Let $\{x_3 = u(x), x \in \Omega\}$ be a Lipschitz continuous vanishing viscosity minimal graph. The flow of the vector $X_{1u}u$ yields a foliation of the domain Ω by polynomial curves γ of degree two. For every fixed $x_0 \in \Omega$ denote by γ the unique leaf passing through that fixed point. The function u is differentiable at x_0 in the Lie sense along γ and the equation* (1.38) *reduces to* $\frac{d^2}{dt^2}(u(\gamma(t)) = 0$.

Remark 1.7.4. To better understand the notion of intrinsic regularity we consider to the non-smooth minimal graph $u(x) = \frac{x_2}{x_1 - sgn\,(x_2)}$. Although this function is not C^1 in the Euclidean sense, observe that $X_{1u}u = 0$ for every $x \in \Omega$. Hence, this is an example of a minimal surface which is not smooth but which can be differentiated indefinitely in the direction of the Legendrian foliation. An other example of non regular minimal surface has been provided in [69].

1.7.3 Uniform regularity for the Riemannian approximating minimal graph

In this section we fix a solution of the Riemannian approximating equation, and establish a priori estimates, uniform in ε. To this end we assume that f is a fixed smooth functions defined on an open set Ω of $\mathbb{R}^2$, and that u is a solution of the (1.39) in Ω. We also assume that

$$M = ||u||_{L^\infty(\Omega)} + ||\nabla_u^\varepsilon u||_{L^\infty(\Omega)} + ||\partial_2 u||_{L^\infty(\Omega)} < \infty. \tag{1.40}$$

The necessary estimates will be provided in suitable Sobolev spaces defined in terms of the vector fields.

Definition 1.7.5. We will say that $\phi \in W_\varepsilon^{1,p}(\Omega)$, $p > 1$ if

$$\phi, \nabla_u^\varepsilon \phi \in L^p(\Omega).$$

In this case we will set

$$||\phi||_{W_\varepsilon^{1,p}(\Omega)} = ||\phi||_{L^p(\Omega)} + ||\nabla_u^\varepsilon \phi||_{L^p(\Omega)}.$$

We will say that $\phi \in W^{k,p}_\varepsilon(\Omega)$ if $\phi \in L^p$, $\nabla^\varepsilon_u \phi \in W^{k-1,p}_\varepsilon(\Omega)$.

If $\varepsilon = 0$ we give analogous definition of Sobolev spaces in the Subriemannian setting. We will denote by $W^{k,p}_0(\Omega)$ the space of $L^p(\Omega)$ functions ϕ such that

$$X^\varepsilon_{1u}\phi, \cdots, (X^\varepsilon_{1u})^k \phi \in L^p(\Omega).$$

Using in full strength the nonlinearity of the operator L_ε, we prove here some Caccioppoli-type inequalities for the intrinsic gradient of u, and for the derivative $\partial_2 u$.

We first prove that if u is a smooth solution of equation (1.39) then its derivatives $\partial_2 u$ and $X^\varepsilon_{ku} u$ are solution of new second order equation, defined in terms of vector fields:

$$M_\varepsilon z = \sum_{ij=1}^{2} X^\varepsilon_{iu}\Big(\frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}} X^\varepsilon_{ju} z\Big) \quad \text{where} \quad A_{ij}(p) = \delta_{ij} - \frac{p_i p_j}{1+|p|^2}. \tag{1.41}$$

We first observe that

$$\partial_2 X^\varepsilon_{iu} u = -(X^\varepsilon_{iu})^* \partial_2 u,$$

where $(X^\varepsilon_{iu})^*$ is the L^2- adjoint of the differential operator X^ε_{iu} and

$$(X^\varepsilon_{1u})^* = -X^\varepsilon_{1u} - \partial_2 u, \quad (X^\varepsilon_{2u})^* = -X^\varepsilon_{2u}. \tag{1.42}$$

Lemma 1.7.6. *If u is a smooth solution of (1.39) then $v = \partial_2 u$ is a solution of the equation*

$$\sum_{i,j} (X^\varepsilon_{iu})^* \Big(\frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}} (X^\varepsilon_{ju})^* v\Big) = 0, \tag{1.43}$$

where A_{ij} are defined in (1.41). This equation can be equivalently represented as

$$M_\varepsilon z = f_1(\nabla^\varepsilon_u u) v^3 + f_{2,i}(\nabla^\varepsilon_u u) v X^\varepsilon_{iu} v^2 + X_i\Big(f_{3,i}(\nabla^\varepsilon_u u) v^2\Big), \tag{1.44}$$

for suitable smooth functions f_1 and $f_{j,i}$. Analogously the function $z = X^\varepsilon_{ku} u$ with $k \leq 2$ is a solution of the equation

$$M_\varepsilon z = f_1(\nabla^\varepsilon_u u) v^2 + f_{2,i}(\nabla^\varepsilon_u u) X^\varepsilon_{iu} v^2 + X_i\Big(f_{3,i}(\nabla^\varepsilon_u u) v\Big). \tag{1.45}$$

Proof. Let us prove the first assertion. Differentiating the equation (1.39) with respect to ∂_2 we obtain

$$\partial_2\Big(X^\varepsilon_{iu}\Big(\frac{X^\varepsilon_{iu} u}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}\Big)\Big) = 0$$

Using (1.42)

$$(X_{iu}^{\varepsilon})^{*}\Big(\partial_2\Big(\frac{X_{iu}^{\varepsilon}u}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}\Big)\Big)=0$$

Note that

$$\partial_2\Big(\frac{X_{iu}^{\varepsilon}u}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}\Big)=\frac{\partial_2 X_{iu}^{\varepsilon}u}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}-\frac{X_{iu}^{\varepsilon}u\,X_{ju}^{\varepsilon}u\,\partial_2 X_{ju}^{\varepsilon}u}{(1+|\nabla_u^{\varepsilon}u|^2)^{3/2}}$$

$$=-\frac{(X_{iu}^{\varepsilon})^{*}\partial_2 u}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}+\frac{X_{iu}^{\varepsilon}u\,X_{ju}^{\varepsilon}u\,(X_{ju}^{\varepsilon})^{*}\partial_2 u}{(1+|\nabla_u^{\varepsilon}u|^2)^{3/2}}=\frac{A_{ij}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}(X_{ju}^{\varepsilon})^{*}v$$

The first assertion is proved.

Assertion (1.44) follows from (1.42) and (1.43). Indeed

$$0=\sum_{i,j}X_{iu}^{\varepsilon}\Big(\frac{A_{ij}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}X_{ju}^{\varepsilon}v\Big)+\sum_{i}X_{iu}^{\varepsilon}\Big(\frac{A_{i1}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}v^2\Big)+$$

$$\sum_{j}\frac{A_{1j}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}vX_{ju}^{\varepsilon}v+\frac{A_{11}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}v^3.$$

We omit the proof of (1.45), which is a similar direct verification. □

Since the operator M_ε in (1.41) is in divergence form, it is quite standard to prove the following intrinsic Cacciopoli type inequalities:

Proposition 1.7.7. *(Intrinsic Cacciopoli type inequality) Let u be a smooth solution of (1.39), satisfying (1.40). Let us denote*

$$z=X_{uk}^{\varepsilon}u+2M,\quad v=\partial_2 u+2M,$$

where M is the constant in (1.40). Then for every p there exists a constant C, only dependent on p and M in such that for every $\phi\in C_0^{\infty}$

$$\int|\nabla_u^{\varepsilon}v|^2z^{p-2}\phi^2\le C\int z^p(\phi^2+|\nabla_u^{\varepsilon}\phi|^2)+\int|\nabla_u^{\varepsilon}z|^2z^{p-2}\phi^2,$$

$$\int|\nabla_u^{\varepsilon}z|^2z^{p-2}\phi^2\le C\int z^p(\phi^2+|\nabla_u^{\varepsilon}\phi|^2).$$

Proof. Since A_{ij} is uniformly elliptic, we have

$$\int|\nabla_u^{\varepsilon}v|^2z^{p-2}\phi^2\le C\int\frac{A_{ij}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}X_{iu}^{\varepsilon}vX_{ju}^{\varepsilon}vz^{p-2}\phi^2=$$

(using the expression (1.42) of the formal adjoint)

$$=-C\int\frac{A_{ij}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}(X_{iu}^{\varepsilon})^{*}v\,X_{ju}^{\varepsilon}v\,z^{p-2}\,\phi^2+C\int\frac{A_{1j}(\nabla_u^{\varepsilon}u)}{\sqrt{1+|\nabla_u^{\varepsilon}u|^2}}X_{ju}^{\varepsilon}v\,v\partial_2 u\,z^{p-2}\phi^2=$$

(integrating by parts X^ε_{ju} in the first integral)

$$= C\int (X^\varepsilon_{ju})^*\Big(\frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}(X^\varepsilon_{iu})^* v\Big) v z^{p-2}\phi^2 +$$

$$+(p-2)C\int \frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}(X^\varepsilon_{iu})^* v\ v X^\varepsilon_{ju} z\, z^{p-3}\phi^2 +$$

$$+2C\int \frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}(X^\varepsilon_{iu})^* v\ v z^{p-2}\phi X^\varepsilon_{ju}\phi +$$

$$+C\int \frac{A_{i\,1}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}(X^\varepsilon_{iu})^* v\, v\partial_2 u\, z^{p-2}\phi^2 + C\int \frac{A_{1\,j}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}} X^\varepsilon_{ju} v\, v\partial_2 u z^{p-2}\phi^2.$$

The first integral vanishes by Lemma 1.7.6. In the other integrals we can use the fact that

$$\Big|\frac{A_{ij}(\nabla^\varepsilon_u u)}{\sqrt{1+|\nabla^\varepsilon_u u|^2}}\Big| \leq 1, \quad |v| \leq M, \quad \text{and} \quad |(X^\varepsilon_{iu})^* v| \leq (M^2 + |\nabla^\varepsilon_u v|)$$

where M is defined in (1.40). Then (eventually changing the constant C)

$$\int |\nabla^\varepsilon_u v|^2 z^{p-2}\phi^2 \leq C\Big(\int |\nabla^\varepsilon_u v||\nabla^\varepsilon_u z| z^{p-3}\phi^2 + \int |\nabla^\varepsilon_u v| z^{p-2}(\phi^2 + |\phi\nabla^\varepsilon_u\phi|)\Big)$$

(by Hölder inequality and the fact that z is uniformly bounded away from 0)

$$\leq \delta\int |\nabla^\varepsilon_u v|^2 z^{p-2}\phi^2 + C(\delta)\int |\nabla^\varepsilon_u z|^2 z^{p-2}\phi^2 + C(\delta)\int z^p(\phi^2 + |\nabla^\varepsilon_u\phi|^2).$$

For δ sufficiently small this implies that

$$\int |\nabla^\varepsilon_u v|^2 z^{p-2}\phi^2 \leq C\int |\nabla^\varepsilon_u z|^2 z^{p-2}\phi^2 + C\int z^p(\phi^2 + |\nabla^\varepsilon_u\phi|^2). \tag{1.46}$$

This prove the first inequality. We omit the proof of the second, which is completely analogous, and can be founded in [20]. □

We want to prove the $C^{1\alpha}$ regularity of z. The classical proof is based on the Moser procedure. This method requires two ingredients: the Sobolev embedding and the Caccioppoli inequality. Here we have proved an intrinsic Caccioppoli type inequality, but we can not prove the intrinsic Sobolev embedding for vector fields with non regular coefficients. This is why we will establish now an Euclidean Caccioppoli inequality, and use the the standard, Euclidean procedure for a first gain of regularity:

Proposition 1.7.8. *Let u be a solution of equation (1.39) satisfying (1.40). For every compact set $K \subset\subset \Omega$ then there exist a real number α and a constant C, only dependent on the constant M in (1.40) such that*

$$||u||_{W^{2,2}_\varepsilon(K)} + ||\partial_2 u||_{W^{1,2}_\varepsilon(K)} + ||u||_{C^{1,\alpha}_u(K)} \leq C.$$

Proof The first part of the thesis

$$||u||_{W^{2,2}_\varepsilon(K)} + ||\partial_2 u||_{W^{1,2}_\varepsilon(K)} \leq C,$$

is a proved in Proposition 1.7.7. Let us now establish an Euclidean Cacciopoli type inequality for $z = X^\varepsilon_{ku}u$. We observe that the Euclidean gradient can be estimated as follows:

$$\begin{aligned} |\nabla_E z|^2 \leq |X^\varepsilon_{1u} z - u\partial_2 z|^2 + |\partial_2 z|^2 \leq \\ \leq |X^\varepsilon_{1u} z|^2 + C|\partial_2 (X^\varepsilon_{1u} u)|^2 = \\ = |X^\varepsilon_{1u} z|^2 + C|(X^\varepsilon_{1u})^* v|^2 \leq |\nabla^\varepsilon_u z|^2 + |\nabla^\varepsilon_u v|^2 + C. \end{aligned} \tag{1.47}$$

From Proposition 1.7.7 it follows that for every $p \neq 1$ there exists a constant C, only dependent on p such that for every $\phi \in C^\infty_0$

$$\int |\nabla_E z|^2 z^{p-2} \phi^2 \leq C \int z^p (\phi^2 + |\nabla_E \phi|^2). \tag{1.48}$$

Now the thesis follows via the classical Euclidean Moser technique.

With this better regularity of the coefficients, we can prove use the Sobolev type Theorem 1.6.5 for vector fields with $C^{1,\alpha}$ coefficients, to obtain a further gain of regularity.

Proposition 1.7.9. *Let u be a solution of equation (1.39) satisfying (1.40). For every compact set $K \subset\subset \Omega$ then there exist a real number α and a constant C, only dependent on the constant M in (1.40) such that*

$$||u||_{W^{2,10/3}_\varepsilon(K)} + ||\partial_2 u||_{W^{1,2}_\varepsilon(K)} + ||u||_{C^{1,\alpha}_u(K)} \leq C. \tag{1.49}$$

Proof We first note that equation (1.39) can be as well written in divergence form:

$$L_\varepsilon = \sum_{ij} A_{ij}(\nabla^\varepsilon_u u) X^\varepsilon_{iu} X^\varepsilon_{ju}$$

where A_{ij} are the coefficients defined in (1.41). Since the function u satisfies uniform $C^{1\alpha}$ estimates, the coefficients $A_{ij}(\nabla^\varepsilon_u u)$ satisfy uniform C^α estimates. Then we can apply Theorem 1.6.5 using the fact that for every p

$$||\partial_2 u||_{L^p(K)} + ||\nabla^\varepsilon_u \partial_2 u||_{L^2(K)} \leq C.$$

It follows that

$$||u||_{W^{2,r}_\varepsilon(K)} \le C$$

where $r = 10/(5-2(1-\alpha))$. Since we do not have an estimate for α, we will set $\alpha = 0$, and obtain $r = 10/3$.

Due to the fact that our Sobolev inequality is not optimal, we will also need an interpolation property, which is completely intrinsic, and can take the place of a Sobolev inequality:

Proposition 1.7.10. *For every $p \ge 3$, for every function $z \in C^\infty(\Omega)$ there exists a constant C_p, dependent on p, the constant M in (1.40) such that and for every $\phi \in C^\infty_0(\Omega)$, and every $\delta > 0$*

$$\int |X^\varepsilon_{iu} z|^{p+1}\phi^{2p} \le C\int \Big(z^{p+1}\phi^{2p} +$$

$$z^2|X^\varepsilon_{iu}z|^{p-1}\phi^{2p-2}|X^\varepsilon_{iu}\phi|^2\Big) + C\int |(X^\varepsilon_{iu})^2 z|^2 |X^\varepsilon_{iu}z|^{p-3}z^2\phi^{2p},$$

where i can be either 1 *or* 2.

Proof We have

$$\int |X^\varepsilon_{iu} z|^{p+1}\phi^{2p} = \int X^\varepsilon_{iu}z|X^\varepsilon_{iu}z|^p \mathrm{sign}(X^\varepsilon_{iu}z)\phi^{2p} =$$

(integrating by parts, using (1.42)) and the Kroneker function δ_{ij}

$$= -\delta_{1i}\int \partial_2 uz|X^\varepsilon_{iu}z|^p\mathrm{sign}(X^\varepsilon_{iu}z)\phi^{2p} - p\int z(X^\varepsilon_{iu})^2 z|X^\varepsilon_{iu}z|^{p-1}\phi^{2p} \tag{1.50}$$

$$-2p\int z|X^\varepsilon_{iu}z|^p\mathrm{sign}(X^\varepsilon_{iu}z)\phi^{2p-1}X^\varepsilon_{iu}\phi \le$$

(by Hölder inequality)

$$\le \frac{C}{\delta}\int \Big(z^{p+1}\phi^{2p} + z^2|X^\varepsilon_{iu}z|^{p-1}\phi^{2p-2}|X^\varepsilon_{iu}\phi|^2\Big) +$$

$$\delta\int |X^\varepsilon_{iu}z|^{p+1}\phi^{2p} + \frac{C}{\delta}\int z^2|(X^\varepsilon_{iu})^2 z|^2|X^\varepsilon_{iu}z|^{p-3}\phi^{2p},$$

choosing δ sufficiently small we obtain the desired inequality.

Next step is to iterate the previous argument, and obtain the higher integrability of the Hessian of u. The proof goes as before: we establish two intrinsic Cacciopoli type inequalities, for the derivatives of $z = X^\varepsilon_{iu}\nabla^\varepsilon_u u$ and $v = \partial_2\nabla^\varepsilon_u u$. From here we deduce that u belongs belong to a better class of Hölder continuous functions. Then the intrinsic Sobolev inequality Theorem 1.6.5 gave the desired estimate of the second derivatives in the natural Sobolev spaces.

Lemma 1.7.11. *Let $p \geq 3$ be fixed, let $f \in C^\infty(\Omega)$, let u be a function satisfying the bound* (1.40) *and let z be a smooth solution of equation $M_\varepsilon z = f$. There exist a constant C which depend on p and the constant M in (1.40) but are independent of ε and z such for every $\phi \in C_0^\infty(\Omega)$, $\phi > 0$,*

$$\int |\nabla_u^\varepsilon (|\nabla_u^\varepsilon z|^{(p-1)/2})|^2 \phi^{2p} \leq$$

$$C\Big(\int \big(|\nabla_u^\varepsilon \phi|^2 + \phi^2\big)^p + \int |\nabla_u^\varepsilon z|^{p+1/2}\phi^{2p} + \int |X_{2u}^\varepsilon(\partial_2 u)|^p \phi^{2p} \tag{1.51}$$
$$+ \int |f|^{2p}\big(|\nabla_u^\varepsilon \phi|^2 + \phi^2\big)\phi^{2p-2} + \int |(\nabla_u^\varepsilon)^2 u||\nabla_u^\varepsilon z|^{p-1}\phi^{2p} +$$
$$\int |(\nabla_u^\varepsilon)^2 u|^2 |\nabla_u^\varepsilon z|^{p-1}\phi^{2p} + \int |(\nabla_u^\varepsilon)^2 u||\nabla_u^\varepsilon z|^{p-1}\phi^{2p-1}|\nabla_u^\varepsilon \phi|\Big).$$

Lemma 1.7.12. *Let u be a smooth solution of equation (1.39) satisfying* (1.40) *and denote $v = \partial_2 u$. For every open set $\Omega_1 \subset\subset \Omega$, for every $p \geq 1$ there exists a positive constant C which depends on Ω_1, p, and on M in (1.40), but is independent of ε such that*

$$||\nabla_u^\varepsilon u||_{C_E^{1/2}} + ||\nabla_u^\varepsilon v||_{L^4(\Omega_1)}^4 \leq C.$$

Proof. We can apply Lemma 1.7.11 with $p = 3$ to the function $v = \partial_2 u$ and deduce that

$$\int |(\nabla_u^\varepsilon)^2 v|^2 \phi^6 \leq C_1 + C_2\Big(\int |\nabla_u^\varepsilon v|^{3+1/2}\phi^6 + \tag{1.52}$$
$$\int (1 + |\nabla_u^\varepsilon v| + |(\nabla_u^\varepsilon)^2 u|)^{7/5}\phi^{23/5}(|\nabla_u^\varepsilon \phi| + \phi)^{7/5} +$$
$$+ \int |(\nabla_u^\varepsilon)^2 u||\nabla_u^\varepsilon v|^2\phi^6 + \int |(\nabla_u^\varepsilon)^2 u|^2|\nabla_u^\varepsilon v|^2\phi^6 + \int |(\nabla_u^\varepsilon)^2 u||\nabla_u^\varepsilon v|^2\phi^5|\nabla_u^\varepsilon \phi|\Big).$$

It follows that

$$\int |(\nabla_u^\varepsilon)^2 v|^2\phi^6 \leq \frac{C_2}{\delta}\int |(\nabla_u^\varepsilon)^2 u|^4\phi^6 + \delta\int |\nabla_u^\varepsilon v|^4\phi^6 + \frac{C_1}{\delta}. \tag{1.53}$$

Analogously, if we set $z = X_{1u}^\varepsilon u$, or $z = X_{2u}^\varepsilon u$, we have

$$\int |(\nabla_u^\varepsilon)^2 z|^2\phi^6 \leq \frac{C_2}{\delta}\int |(\nabla_u^\varepsilon)^2 u|^4\phi^6 + \frac{C_1}{\delta} + C_2\int |\nabla_u^\varepsilon v|^3\phi^6 \tag{1.54}$$

Using Lemma 1.7.10, (1.53) and (1.49), we obtain immediately

$$\int |\nabla_u^\varepsilon v|^4\phi^6 \leq C_1 + C_2\int |(\nabla_u^\varepsilon)^2 v|^2\phi^6 \leq C_1 + \frac{C_2}{\delta}\int |(\nabla_u^\varepsilon)^2 u|^4\phi^6 + \delta\int |\nabla_u^\varepsilon v|^4\phi^6$$

Hence
$$\int |\nabla_u^\varepsilon v|^4 \phi^6 \leq C_1 + C_2 \int |(\nabla_u^\varepsilon)^2 u|^4 \phi^6 \tag{1.55}$$

Consequently, from the latter and (1.54) we deduce that

$$\int |(\nabla_u^\varepsilon)^2 z|^4 \phi^6 \leq C_1 + C_2 \int |(\nabla_u^\varepsilon)^2 u|^4 \phi^6 \tag{1.56}$$

Next, from the intrinsic Cacciopoli inequalities (1.55) and (1.56) we deduce an Euclidean Cacciopoli inequality: Note that

$$|\nabla_E X_{1u}^\varepsilon z| \leq |(X_{1u}^\varepsilon)^2 z| + C_2 |\partial_2 (X_{1u}^\varepsilon z)| \leq |(X_{1u}^\varepsilon)^2 z| + C_2 |v \partial_2 z| + C_2 |X_{1u}^\varepsilon \partial_2 z| \leq$$

(since $\partial_2 z = \partial_2 X_{1u}^\varepsilon u = v^2 + X_{1u}^\varepsilon v$)

$$|(\nabla_u^\varepsilon)^2 z| + C_2 |(\nabla_u^\varepsilon)^2 v| + C_2 |\nabla_u^\varepsilon v| + C_2.$$

From the latter and (1.55) and (1.56) we infer

$$\int |\nabla_E (\nabla_u^\varepsilon) z|^2 \phi^6 \leq C_2 \Big(\int |(\nabla_u^\varepsilon)^2 v|^2 \phi^6 + \tag{1.57}$$

$$\int |(\nabla_u^\varepsilon)^2 z|^2 \phi^6 + 1 \Big) \leq C_2 \int |\nabla_u^\varepsilon z|^4 \phi^6 + C_1$$

Now we can apply the standard Euclidean Sobolev inequality in $\mathbb{R}^2$ and obtain

$$\Big(\int (|\nabla_u^\varepsilon z| \phi^3)^6 \Big)^{1/3} \leq C_2 \int |\nabla_E (\nabla_u^\varepsilon z \phi^3)|^2 \leq C_2 \int |\nabla_u^\varepsilon z|^4 \phi^6 + C_1 \leq$$

(using Hölder inequality)

$$\leq C_2 \Big(\int (|\nabla_u^\varepsilon z| \phi^3)^6 \Big)^{1/3} \Big(\int_{supp(\phi)} |\nabla_u^\varepsilon z|^3 \Big)^{2/3} + C_1.$$

By (1.49) and the fact that $|\nabla_u^\varepsilon z| \leq |\nabla_\varepsilon^2 u|$, we already know that $|\nabla_u^\varepsilon z| \in L^3_{loc}$. In fact

$$\Big(\int_{supp(\phi)} |\nabla_u^\varepsilon z|^3 \Big)^{2/3} \leq \Big(\int_{supp(\phi)} |\nabla_u^\varepsilon z|^{10/3} \Big)^{3/5} |supp(\phi)|^{1/15}.$$

Recall that C_2 does not depend on $|\nabla_u^\varepsilon \phi|$. If we choose the support of ϕ sufficiently small, we can assume that the integral $\int_{supp(\phi)} |\nabla_u^\varepsilon z|^3$ is arbitrarily small. It follows that

$$\Big(\int (|\nabla_u^\varepsilon z| \phi^3)^6 \Big)^{1/3} \leq C_1$$

and consequently, by (1.55)

$$\int |\nabla_u^\varepsilon v|^4 \phi^6 \leq C_1$$

But this implies that $|\nabla_E(\nabla^\varepsilon_u u)| \leq |(\nabla^\varepsilon_u)^2 u| + |\nabla^\varepsilon_u v| + v^2 \in L^4_{loc}$. This implies, buy the standard Euclidean Sobolev Morrey inequality in $\mathbb{R}^2$ that

$$\nabla^\varepsilon_u u \in C^{1/2}_E.$$

□

Lemma 1.7.13. *Let u be a smooth solution of equation (1.39) satisfying* (1.40) *and denote $v = \partial_2 u$. For every open set $\Omega_1 \subset\subset \Omega$, for every $p \geq 1$ there exists a positive constant C which depends on Ω_1, p, and on M in (1.40), but is independent of ε such that*

$$||u||_{W^{2,p}_\varepsilon(\Omega_1)} \leq C.$$

Proof. We have already noted that equation (1.39) can be as well written in divergence form:

$$L_\varepsilon u = \sum_{ij} A_{ij}(\nabla^\varepsilon_u u) X^\varepsilon_{iu} X^\varepsilon_{ju} u = 0.$$

Now the function u satisfy uniform $C^{1,1/2}$ estimates, the coefficients $A_{ij}(\nabla^\varepsilon_u u)$ satisfy uniform $C^{1/2}$ estimates. Then we can apply Theorem 1.6.5 using the fact that for every p

$$||\partial_2 u||_{L^p(\Omega_1)} + ||\nabla^\varepsilon_u \partial_2 u||_{L^4(\Omega_1)} \leq C.$$

If follows that for every $r > 1$ there exists a constant $C > 0$ independent of ε such that

$$||u||_{W^{2,r}_\varepsilon(\Omega_1)} \leq C.$$

□

Using a bootstrap argument, we can now deduce the same result for derivative of any order:

Theorem 1.7.14. *Let u be a smooth solution of equation (1.38), satisfying* (1.40). *For every open set $\Omega_1 \subset\subset \Omega$, for every $p \geq 3$, and every integer $k \geq 2$ there exists a constant C which depends on p, k Ω_1 and on M in (1.40), but is independent of ε such that the following estimates holds*

$$||u||_{W^{k,p}_\varepsilon(\Omega_1)} + ||\partial_2 u||_{W^{k,p}_\varepsilon(\Omega_1)} \leq C. \tag{1.58}$$

Corollary 1.7.15. *Let u be a smooth solution of equation (1.38), satisfying* (1.40). *For every open set $\Omega_1 \subset\subset \Omega$, for every $p \geq 3$, $\alpha < 1$ and every integer $k \geq 2$ there exists a constant C which depends on p, k Ω_1 and on M in (1.40), but is independent of ε such that the following estimates holds*

$$||(\nabla^\varepsilon_u)^{k+1} u||_{L^p(\Omega_1)} + ||\partial_2 (\nabla^\varepsilon_u)^k u||_{L^p(\Omega_1)} + ||(\nabla^\varepsilon_u)^k u||_{C^\alpha_E(\Omega_1)} \leq C. \tag{1.59}$$

1.7.4 Regularity of the viscosity minimal surface

In this section we turn our attention to the proof of regularity for vanishing viscosity solutions u of equation (1.38). The regularity rests on the *a priori* estimates proved in the previous section in the limit $\varepsilon \to 0$.

Theorem 1.7.16. *Let $u \in Lip(\Omega)$ be a vanishing viscosity solution of* (1.38), *then equation (1.38) can be represented as $X_{1u}^2 u = 0$ and is satisfied weakly in the Sobolev sense, and hence, pointwise a.e. in Ω, i.e.*

$$\int_\Omega X_{1u} u X_{1u}^* \phi = 0 \textit{ for all } \phi \in C_0^\infty(\Omega).$$

Moreover forevery $\alpha < 1$, for every $p > 1$, for every natural k

$$(\nabla_u u)^k \in C_E^\alpha \quad \partial_2(\nabla_u u)^k \in W_0^{1,p}(B(R)) \tag{1.60}$$

Proof. Let (u_j) denote the sequence approximating u, as defined in Definition 1.7.1. For each ε_j the function u_j is a solution of (1.39). Hence, by corollary (1.7.15) the sequence

$$(\nabla_{u_j}^{\varepsilon_j} u_j)_j$$

is bounded in C_E^α for every α. Evetually extracting a subsequence we see that it weakly converges to $(X_{1u}u, 0)$. Hence this is limit in C_E^α norm. On the other hand $\partial_2 u_j$ is weakly convergent to $\partial_2 u$. Hence letting j go to ∞ in the divergence form equation we conclude that $X_{1u}^2 u = 0$ in the weak Sobolev sense. The other part of the thesis always follows from Corollary 1.7.15. $\square$

1.7.5 Foliation of minimal surfaces and completion result

If the weak derivative of a function f is sufficiently regular, they are Lie derivatives.

Proposition 1.7.17. *If $f \in C_{loc}^\alpha(\Omega)$ for some $\alpha \in]0,1[$ and its weak derivatives $X_{1u}f \in C_{loc}^\alpha(\Omega), \partial_2 f \in L_{loc}^p(\Omega)$ with $p > 1/\alpha$, then for all $\xi \in \Omega$ the Lie-derivatives $X_{1u}f(\xi)$ exist and coincide with the weak ones.*

We are now ready to prove the result concerning the foliation

Proof of Corollary 1.7.3 First note that, by Proposition 1.7.17 the derivatives of u are Lie derivatives. The equation $\gamma' = X_{1u}I(\gamma)$ has an unique solution, of the form

$$\gamma(x) = (x, y(x)),$$

where $y'(x) = u(x, y(x))$. In view of the regularity of u and of the previous proposition then $y''(x) = Xu(x, y(x))$, and $y'''(x) = X^2 u(x, y(x)) = 0$. This shows that γ is a polynomial of order 2 and concludes the proof. $\square$

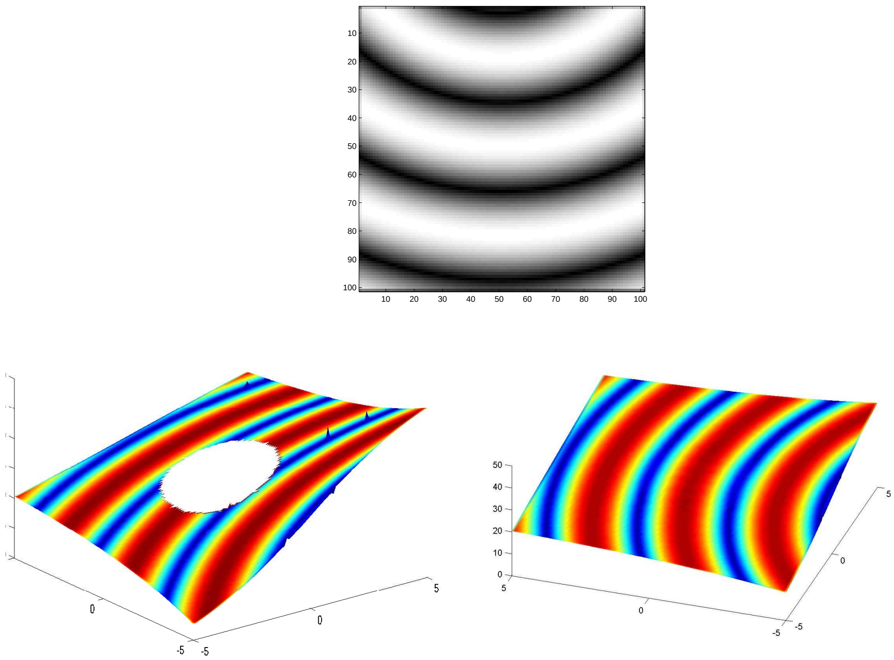


Figure 1.23: The original image (top left) is lifted in the rototranslation space with missing information in the center, like in the phenomenon of macula cieca (top right). The surface is completed by the algorithm (bottom).

Let us now present some computational results, applied to well known images.

The minimal surface which perform the completion is foliated in geodesics. This implies that each level lines of the image is completed independently through an elastica, and this is compatible with the phenomenological evidence. We consider here the completion of a figure that has been only partially lifted in the roto-translation space. This example mimics the missing information due to the presence of the macula cieca (blind spot) that is modally completed by the human visual system, as outlined in [51]. The original image (see Figure 1.23), top left) is lifted in the rotranslation space with missing information in the center (top right). The lifted surface is completed by iteratively applying eqs until a steady state is achieved. The final surface is minimal with respects to the sub-Riemannian metric.

Bibliography

[1] L. Ambrosio, S. Masnou, *A direct variational approach to a problem arising in image recostruction* Interfaces Free Bound. **5** (2003), 1, 63–81.

[2] L. Ambrosio, F. Serra Cassano, D. Vittone, *Intrinsic regular hypersurfaces in Heisenberg groups* J. Geom. Anal. **16** (2006), 2, 187–232.

[3] L. Ambrosio, B. Kirchheim, *Currents in metric spaces* Acta Math. **185** (2000), 1, 1–80.

[4] L. Ambrosio, F. Serra Cassano, D. Vittone, *Intrinsic regular hypersurfaces in Heisenberg groups* Journal of geometric analysis **16** (2006), 2, 187-232.

[5] N. Arcozzi, F. Ferrari, *Metric normal and distance function in the Heisenberg group* Math. Z. **256** (2007), 3, 661–684.

[6] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, J. Verdera, *Filling-in by joint interpolation of vector fields and gray levels* IEEE Trans. Image Process. **10** (2001), 8, 1200–1211.

[7] D. Barbieri, G. Citti, *Regularity of non charachteristic minimal graphs in Lie groups of step 2 and dimension 3*, preprint

[8] G. Barles, C. Georgelin, *A simple proof for the convergence for an approximation wcheme for computing motions by mean curvature* SIAM J. Numerical Analysis **32** (1995), 2, 484–500.

[9] V. Barone Adesi, F. Serra Cassano, D. Vittone, *The Bernstein problem for intrinsic graphs in Heisenberg groups and calibrations* Calc. Var. Partial Differential Equations **30** (2007), 1, 17–49.

[10] G. Bellettini, R. March, *An image segmentation variational model with free discontinuities and contour curvature* Math. Models Methods Appl. Sci. **14** (2004), 1, 1–45.

[11] G. Bellettini, M.Paolini, *Approssimazione variazionale di funzionali con curvatura* Seminario di analisi matematica, Dip. di Mat. dell'Univ. di Bologna, 1993.

[12] J. Bence, B. Merriman, S. Osher, *Diffusion generated motion by mena curvature* Computational Crystal Growers Workshop, J. Taylor Sel. Taylor Ed.

[13] F. Bigolin, F. Serra Cassano, *Intrinsic regular graphs in Heisenberg groups vs. weak solutions of nonlinear first-order PDEs*, preprint 2007.

[14] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians.* Springer Monographs in Mathematics, Springer, Berlin, 2007.

[15] J. M. Bony, *Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés* Ann. Inst. Fourier (Grenoble) **19** (1969), 277–304.

[16] W. Bosking, Y. Zhang, B. Schofield, D. Fitzpatrick, *Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex* J. Neurosi. **17** (1997), 6, 2112–2127.

[17] , M. Bramanti, L. Brandolini, M. Pedroni, *Basic properties of nonsmooth Hrmander's vector fields and Poincar's inequality* preprint 2008.

[18] L.Capogna, D. Danielli, S. Pauls, J. Tyson, *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem* Progress in Mathematics **259**, Birkhäuser, 2007.

[19] L. Capogna, G. Citti, *Generalized mean curvature flow in Carnot groups* to appear in Communications in PDE.

[20] L. Capogna, G. Citti, M. Manfredini, *Regularity of non-characteristic minimal graphs in the Heisenberg group* H^1 to appear in Indiana University Mathematical Journal, arXiv:0804.3406

[21] L. Capogna, G. Citti, M. Manfredini, *Smoothness of Lipschitz intrinsic minimal graphs in the Heisenberg group* H^n, $n > 1$ to appear in in Crelle's Journal.

[22] J.-H. Cheng, J.-F. Hwang, A. Malchiodi, P. Yang, *Minimal surfaces in pseudohermitian geometry*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **4** (2005), 1, 129–177.

[23] J. H. Cheng, J. F. Hwang, P. Yang, *Existence and uniqueness for p-area minimizers in the Heisenberg group* Math. Ann. **337** (2007), 2, 253–293.

[24] G. Citti, M. Manfredini, *Blow-up in non homogeneous Lie groups and rectifiability* Houston J. Math. **31** (2005), 2, 333–353.

[25] G. Citti, M. Manfredini, *Implicit function theorem in Carnot-Carathodory spaces* Commun. Contemp. Math. **8** (2006), 5, 657–680.

[26] G. Citti, M. Manfredini, *Uniform estimates of the fundamental solution for a family of hypoelliptic operators* Potential Anal. **25** (2006), 2, 147–164.

[27] G. Citti, E. Lanconelli, A. Montanari, *Smoothness of Lipchitz-continuous graphs with nonvanishing Levi curvature* Acta Math. **188** (2002), 1, 87–128.

[28] G. Citti, A. Sarti, *A cortical based model of perceptual completion in the roto-translation space* J. Math. Imaging Vision **24** (2006), 3, 307–326.

[29] D. Danielli, N. Garofalo, D.-M. Nhieu, *Sub-Riemannian calculus on hypersurfaces in Carnot groups* Adv. Math. **215** (2007), 1, 292–378.

[30] D. Danielli, N. Garofalo, D.-N. Nhieu, *A notable family of entire intrinsic minimal graphs in the Heisenberg group which are not perimeter minimizing* Amer. J. Math. **130** (2008), 2, 317–339.

[31] J. G. Daugman, *Uncertainty - relation for resolution in sapce, spatial frequency and orientation optimized by two dimensional visual cortical filters* J. Opt. Soc. Amer. **7** (1985), 2, 1160–1169.

[32] E. De Giorgi, *Some remarks on* Γ *convergence and least square methods* in "Composite media and homogeniziation Theory", G. Dal Maso and G. F. Dell'Antonio (Eds.), Birkhauser, 1991, 153-142.

[33] R. Duits, E. Franken. *Left-invariant stochastic evolution equations on se(2) and its applications to contour enhancement and contour com- pletion via invertible orientation scores.* ArXiv e-prints, November 2007.

[34] S. Esedoglu, R. March, *Segmentation with Deph but without detecting junctions* Journal of Mathematical Imaging and Vision **18** (2003), 7–15.

[35] L. Evans, *Convergence of an Algorithm for mean curvature motion* Indiana Univ. Math J. **42** (1993), 2, 553–557 .

[36] H. Federer, *Geometric measure theory*, Springer, Berlin, 1969.

[37] E. Franken, R. Duits, B. M. ter Haar Romeny. *Nonlinear diffusion on the 2d euclidean motion group.* In Fiorella Sgallari, Almerico Murli, and Nikos Paragios, editors, SSVM, volume 4485 of Lecture Notes in Computer Science, pages 461-472. Springer, 2007.

[38] G.B. Folland, *Subelliptic estimates and function spaces on nilpotent Lie groups* Ark. Mat. (1975) **13**, 161–207.

[39] G.B. Folland, E.M. Stein, *Estimates for the* ∂_b *Complex and Analysis on the Heisenberg Group* Comm. Pure Appl. Math. **20** (1974), 429–522.

[40] B. Franchi, R. Serapioni, F. Serra Cassano, *Rectifiability and Perimeter in the Heisenberg Group* Math.Ann. **321** (2001), 479–531.

[41] B. Franchi, R. Serapioni, F. Serra Cassano, *Rectifiability and perimeter in step 2 groups* Mathematica Bohemica **127** (2002), 2, 219–228.

[42] B. Franchi, R. Serapioni, F. Serra Cassano, *Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups* Comm. Anal. Geom. **11** (2003), 5, 909–944.

[43] B. Franchi, R. Serapioni, F. Serra Cassano, *On the structure of finite perimeter sets in step 2 Carnot groups* Journal of geometric analysis **13** (2003), 3, 421–466.

[44] B. Franchi, R. Serapioni, F. Serra Cassano, *Regular hypersurfaces, Intrinsic perimeter and implicit function theorem in Carnot groups* Communications in analysis and geometry **11** (2003), 909–944.

[45] N. Garofalo, S. Pauls, *The Bernstein problem in the Heisenberg group*, preprint 2003.

[46] N. Garofalo, D.N. Nhieu, *Isoperimetric and Sobolev inequalities for Carnot-Carathédory spaces and the existence of minimal surfaces* Commun. Pure Appl. Math. **49** (1996), 10, 1081–1144.

[47] A. Field, R. F. Heyes, Hess, *Contour integration by the human visual system: evidence for a local Association Field* Vision Research **33** (1993), 173–193.

[48] R. K. Hladky, S. D. Pauls, *Constant mean curvature surfaces in sub-Riemannian geometry* J. Differential Geom. **79** (2008), 1, 111–139.

[49] W.C. Hoffman, *The visual cortex is a contact bundle*, Applied Mathematics and Computation, vol 32, (89), 137-167.

[50] J. P. Jones, L. A. Palmer, *An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex* J. Neurophysiology **58** (1987), 1233–1258.

[51] G. Kanizsa, *Grammatica del vedere*, Il Mulino, Bologna, 1980.

[52] G. Kanizsa, *Organization in Vision*, Hardcover, 1979.

[53] B. Kirchheim, F. Serra Cassano, *Rectifiability and parametrization of intrinsic regular surfaces in the Heisenberg group*, Annali della Scuola normale superiore di Pisa. Classe di scienze **4** (2004), 3, 4, 871–896.

[54] E. Lanconelli, D. Morbidelli *On the Poincare' inequality for vector fields*, Arkiv för Matematik, 38 (2000), 327-342.

[55] M. Manfredini, *A note on the Poincaré inequality for non smooth vector fields* preprint.

[56] M. Manfredini, *Fundamental solution for sum of squares of vector fields with $C^{1,\alpha}$ coefficients* preprint.

[57] S. Marcelja, *Mathematical description of the response of simple cortical cells* J. Opt. Soc. Amer. **70** (1980), 1297–1300.

[58] S. Masnou, J.M. Morel, *Level lines based disocclusion* Proc. 5th. IEEE International Conference on Image Processing, Chicago, Illinois, October 4-7, 1998.

[59] K. D. Miller, A. Kayser, N. J. Priebe, *Contrast-dependent nonlinearities arise locally in a model of contrast-invariant orientation tuning* J. Neurophysiol. **85** (2001), 2130–2149.

[60] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs **91**, American Mathematical Society, Providence, RI,2002.

[61] A. Montanari, D. Morbidelli, *Balls defined by vector fields and the Poincar inequality* Ann. Inst. Fourier (Grenoble) **54** (2004), 431–452.

[62] A. Montanari, D. Morbidelli, *Nonsmooth Hörmander vector fields and their control balls*, preprint 2008.

[63] F. Montefalcone, *Hypersurfaces and variational formulas in sub-Riemannian Carnot groups* J. Math. Pures Appl.(9) **87** , (2007), 5, 453–494.

[64] A. Nagel, E. M. Stein, S. Wainger, *Balls and metrics defined by vector fields I: Basic properties* Acta Math. **155** (1985), 103–147.

[65] S.B. Nelson. M. Sur D.C. Somers, *An emergent model of orientation selectivity in cat visual cortical simples cells* J. Neurosci. **15** (1995), 5448–5465.

[66] Y. Ni, *Sub-Riemannian constant mean curvature surfaces in the Heisenberg group as limits* Ann. Mat. Pura Appl. **183** (2004), 4, 555–570.

[67] M. Nitzberg, D. Mumford, T. Shiota, *Filtering, Segmentation and Deph*, Springer-Verlag, Berlin, 1993.

[68] S. D. Pauls, *Minimal surfaces in the Heisenberg group*, Geom. Dedicata **104** (2004), 201–231.

[69] S. D. Pauls, *H-minimal graphs of low regularity in H^1*, Comment. Math. Helv. **81** (2006), 2, 337–381.

[70] J. Petitot, *Phenomenology of Perception, Qualitative Physics and Sheaf Mereology, Proceedings of the 16th International Wittgenstein Symposium*, Vienna, Verlag Hlder-Pichler-Tempsky (1994) 387–408.

[71] J. Petitot, Y. Tondut, *Vers une Neuro-geometrie. Fibrations corticales, structures de contact et contours subjectifs modaux, Mathmatiques, Informatique et Sciences Humaines*, EHESS, Paris, 145 (1998) 5-101.

[72] J. Petitot, *Morphological Eidetics for Phenomenology of Perception, in Naturalizing Phenomenology* Issues in Contemporary Phenomenology and Cognitive Science, (Petitot J., Varela F.J., Roy J.-M., Pachoud B. , eds.), Stanford, Stanford University Press, (1998), 330-371.

[73] N. J. Priebe, K. D. Miller, T. W. Troyer, A. E. Krukowsky, *Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity* J. Neurosci. **18** (1998), 5908–5927.

[74] M. Ritoré, C. Rosales, *Rotationally invariant hypersurfaces with constant mean curvature in the Heisenberg group* H^n J. Geom. Anal. **16** (2006), 4, 703–720.

[75] M. Ritoré, C. Rosales, *Area stationary surfaces in the Heisenberg group* H^1 Adv. Math. **219** (2008), 2, 633–671.

[76] L. Rothschild, E. M. Stein, *Hypoelliptic differential operators and nihilpotent Lie groups*, Acta Math. **137** (1977), 247-320.

[77] A. Sarti, R. Malladi, J.A. Sethian, *Subjective surfaces: A Method for Completion of Missing Boundaries, Proceedings of the National Academy of Sciences of the United States of America* **12** (2000), 97, 6258–6263.

[78] A. Sarti, G. Citti, J. Petitot, *The symplectic structure of the primary visual cortex* Biol. Cybernet. **98** (2008), 1, 33–48.

[79] N. Sherbakova, *Minimal surfaces in contact subriemannian manifolds*, preprint 2006.

[80] M. Shelley, D. J. Wielaard, D. McLaughlin, R. Shapley, *A neuronal network model of macaque primary visual cortex (v1): orientation selectivity and dynamics in the input layer 4calpha* Proc. Natl. Acad. Sci. U.S.A. **97** (2000), 8087-8092.

[81] D. Ts'o, C. D. Gilbert, T.N. Diesel, *Relationship between horizontal interactions and functional architecturein cat striate cortex as revealed by cross-correlation analysis*, J. Neurosc. **6** (1986), 4, 1160–1170

[82] S. W. Zucker. *The curve indicator random field: curve organization via edge correlation.* In Perceptual Organization for Artificial Vision Systems, Kluwer Academic, 2000 265- 288.

Geometric Measure Theory

David Preiss

Some problems of geometric measure theory: Part I: Structure of Lebesgue null sets

based on joint work of G. Alberti[1], M. Csornyei[2] and D. Preiss[3]

[1]University of Pisa
[2]University College London
[3]University of Warwick

Barcelona, June 2009

Lebesgue null sets

We all know what Lebesgue null sets are:
$E \subset \mathbb{R}^n$ is null if for every $\varepsilon > 0$ there is a cover of E by balls $B(x_i, r_i)$ such that $\sum_{i=1}^{\infty} r_i^n < \varepsilon$.

Of course, this is by far not the only description of null sets. For example, Fubini's theorem tells us that a Borel set $E \subset \mathbb{R}^2$ is null if and only if almost every line parallel to the x axis meets it in a set of (one dimensional) measure zero.

At least in $\mathbb{R}$, a number of very different descriptions is easy to prove. Although this was not our original motivation, much of what we do may be understood as attempts to find out whether such descriptions remain valid in $\mathbb{R}^n$, $n > 1$.

We start by looking at two examples.

Non-differentiability in $\mathbb{R}$

Observation. *For every Lebesgue null set $E \subset \mathbb{R}$ there is a Lipschitz $f : \mathbb{R} \to \mathbb{R}$ which is non-differentiable at any point of E.*

Proof. Recursively find open $\mathbb{R} = G_0 \supset G_1 \supset G_2 \supset \cdots \supset E$ so small that G_{k+1} is small in every component of G_k. (For example, $|G_{k+1} \cap C| < 2^{-k-1}|C|$ for any component C of G_k.)

Let $\psi(x) = (-1)^k$ where k is the least index such that $x \in G_k$. Define $f(x) = \int \psi(x)\,dx$.

If $x \in E$ and (a, b) is a component of G_k containing x, then

$$\left| \frac{f(b) - f(a)}{b - a} - (-1)^k \right| \le 2^{-k},$$

so f is not differentiable at x.

What about non-differentiability in $\mathbb{R}^2$?

Theorem. *For $n \ge 2$ there is a Lebesgue null set $E \subset \mathbb{R}^n$ such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point of E.*

How one defines such E?

That's easy: order all lines passing through two rational points into a sequence L_i, choose small $\varepsilon_i > 0$ and let

$$E = \bigcap_{i=1}^{\infty} \bigcup_{j=i}^{\infty} B(L_j, \varepsilon_j)$$

How one proves the theorem?

That's not so easy and will not be done here (unless there is a huge popular demand).

Notice that E can have Hausdorff dimension one; and that it cannot have smaller Hausdorff dimension.

Recent results of Michael Dore and Olga Maleva

Already proved:

Theorem. *For $n \geq 2$ there is a compact null set $E \subset \mathbb{R}^n$ such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point of E.*

Will be (almost surely) proved soon:

Theorem. *For $n \geq 2$ there is a compact $E \subset \mathbb{R}^n$ of Hausdorff dimension one such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point of E.*

Theorem. *For $n \geq 3$ there is a compact $E \subset \mathbb{R}^n$ of Hausdorff dimension two such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^2$ is differentiable at a point of E.*

Problem (Lars Olsen). What about the packing dimension?

Non-differentiable vector-valued maps

The key problem with repeating the construction from the real line is that we need to modify a function to get derivative zero on (small) balls, and that cannot be done unless we modify it also elsewhere.

Even if we consider just sets that can be covered by narrow strips, ie

$$E = \bigcap_{i=1}^{\infty} \bigcup_{j=i}^{\infty} B(L_j, \varepsilon_j) \qquad (||)$$

we also arrive to difficulties (and we know we have to).

But the vector-valued case is different!

Observation. *Let L_i be a sequence of lines in $\mathbb{R}^2$ that pass through two rational points. Then for sufficiently small $\varepsilon_i > 0$ there is a Lipschitz $f : \mathbb{R}^2 \to \mathbb{R}^2$ which in nondifferentiable at any point of the set $(||)$.*

Proof. Fold paper.

Copying from $\mathbb{R}$ to $\mathbb{R}^2$

Using essentially the same construction as in $\mathbb{R}$, we can find, given any set $E \subset \mathbb{R}^2$ of measure zero, a bounded measurable $\psi : \mathbb{R}^2 \to \mathbb{R}$ so that for every $x \in E$ there are arbitrarily small radii r, s for which

$$\int_{\|y-x\|<r} \psi(y)\, dy > r^2 \quad \text{and} \quad \int_{\|y-x\|<s} \psi(y)\, dy < -s^2.$$

Of course, we cannot even dream of a function $f : \mathbb{R}^2 \to \mathbb{R}$ whose derivative is ψ. V. Šverák suggested to consider a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\mathrm{div}(f) = \psi$; this would give the nondifferentiability results with maps to $\mathbb{R}^2$.

Exercise. Show that f is non-differentiable at any point of E.

Unfortunately, for our null set E, the equation $\mathrm{div}(f) = \psi$ has no Lipschitz solution. (This problem attracted some attention independently of differentiability; different constructions of such ψ are due to Burago–Kleiner, McMullen, ...)

Let's try something else

Fact. *Every compact set $K \subset \mathbb{R}$ of positive Lebesgue measure may be contracted onto an interval.*

Proof. Take $f(x) = |(-\infty, x) \cap K|$.

Problem (Laczkovich). Is it true that every compact $K \subset \mathbb{R}^n$ of positive Lebesgue measure may be contracted onto a set with nonempty interior (equivalently, onto a ball, onto a square, etc).

I believe that Laczkovich asked this originally only for $\mathbb{R}^2$; and only twenty years later we found out (thanks to Peter Jones) that this was answered ten years before it was asked.

Theorem (Uy; Khruscev). Given a compact $K \subset \mathbb{C}$ of positive measure, there is a non-constant bounded Lipschitz $f : \mathbb{C} \to \mathbb{C}$ which is holomorphic outside K.

Exercise. Show that $f(K) = f(\mathbb{C})$ has nonempty interior.

Matoušek's solution of Laczkovich's problem in $\mathbb{R}^2$

Erdős-Szekeres Theorem. *For any planar set Q having q^2 points there is a function $\psi : \mathbb{R} \to \mathbb{R}$ with Lip$(\psi) \le 1$ such that one of the sets*

$$\{(x, y) \in Q : y = \psi(x)\} \quad \text{or} \quad \{(x, y) \in Q : x = \psi(y)\}$$

has at least q points.

I understand that in Hungary the proof is given as an exercise to 5 years old kids.

Exercise. Show that the natural higher dimensional version ($Q \subset \mathbb{R}^3$ has q^3 points, $\psi : \mathbb{R}^2 \to \mathbb{R}$, there are three possible sets to consider and one of them should have at least cq^2 points) is false. (Hint: put the points on a plane that meets all faces of a cube.)

Alberti and Csörnyei even managed to show that this is false if we replace 1 in the estimate of Lip(ψ) by an arbitrary constant.

Matoušek's conjecture

Conjecture. *For any set $Q \subset \mathbb{R}^n$ having q^n points there is a function $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ with Lip$(\psi) \le C_n$ and an orthonormal system of coordinates such that the set*

$$\{(x_1, \dots, x_n) \in Q : x_n = \psi(x_1, \dots, x_{n-1})\}$$

has at least $c_n q^{n-1}$ points.

This problem is open. We only know that, unlike in the plane, the coordinate systems cannot be restricted to permutations of the standard coordinate system.

As far as we know, positive answer to this question would not immediately lead to an answer to Laczkovich's problem. (And, indeed, Laczkovich's problem is open for $n \ge 3$.) But it would lead to an answer to differentiability problems (this will be indicated later).

Matoušek's construction

Suppose $K \subset Q = [0, 1]^2$ and $G := Q \setminus K$ has small measure.

Idea. Map Q to a rectangle so that G is mapped to a null set.

Cutting map. Cutting away a Lipschitz strip, ie the set

$$\{(x, y) : \varphi(x) < y < \varphi(x) + h\}$$

is a Lipschitz map of ℓ_∞ to ℓ_∞ with Lipschitz constant one.

Procedure. Approximate G by a finite mesh of squares of side δ. Use Erdös Szekeres (for the centres of these squares) repeatedly to cut away strips of with $c\delta$ which meet G in a set of measure at least $c\sqrt{|G|}\delta$. Do this till what remains of G has measure $|G|/2$. The sides of Q have diminished by at most $(c|G|/\sqrt{|G|}\delta)\delta = c\sqrt{|G|}$.

Start again with the new G and repeat. The square will be diminished by $c\sum_{k=0}^{\infty} \sqrt{|G|/2^k} < 1$, so we are left with a rectangle and G is mapped on a null set.

Non-differentiability and directional non-differentiability

Recall that the directional (or partial) derivative of $f : \mathbb{R}^n \to \mathbb{R}^m$ at $x \in \mathbb{R}^n$ in direction $u \in \mathbb{R}^n$ is

$$f'(x; u) := \lim_{t \to 0} \frac{f(x + tu) - f(x)}{t}$$

provided that the limit exists.

Although differentiability is not the same as existence of enough partial derivatives, the set of points at which these two notions differ is relatively easy to control. An example showing how this is achieved is given by the following statement which is valid also in infinite dimensional separable Banach spaces.

Theorem. *[Zajíček, Preiss 2001] The set of points at which a Lipschitz function is differentiable at a spanning set of directions but is not differentiable is σ-porous.*

Definition of σ-porous sets

- A set E in X is said to be c-porous at $x \in E$, $0 < c < 1$, if for every $\varepsilon > 0$ there is a $z \in X \setminus E$ such that $\|x - z\| < \varepsilon$ and $B(z, c\|x - z\|) \cap E = \emptyset$.
- A set is porous if it is porous at each of its points.
- A countable union of porous sets is called σ-porous.

A remark particularly relevant to the content of this talk is that E is porous at $x \in E$ if and only if the function $x \to \mathrm{dist}(x, E)$ is non-differentiable at x.

Notice however that σ-porous sets are of the first category, and so they do not fully describe non-differentiability sets of Lipschitz functions (already on $\mathbb{R}$).

There are even compact subsets of $\mathbb{R}$ of measure zero (and so also of the first category) that are not σ-porous. (Zajíček 1976)

Differentiability in no direction

In what follows, we may pretend that differentiability and differentiability in a spanning set of directions are equivalent.

Although the treatment of the exceptional σ-porous set is not always trivial, it is a minor problem compared to other difficulties one encounters in this area.

There is another type of sets that look very negligible, sets of points at which a Lipschitz function may be differentiable in no direction. These sets form a σ-ideal of "super-negligible" sets for differentiability problems. We will call them uniformly purely unrectifiable.

Notice that uniformly purely unrectifiable sets are purely unrectifiable (or better, purely 1-unrectifiable), i.e., null on every rectifiable curve.

A mysterious vectorfield

Let us see where the directional derivative arguments lead to. For simplicity, consider just the planar case.

Let $E \subset \mathbb{R}^2$ be a Lebesgue null set contained in the set of points of non-differentiability of a Lipschitz $f : \mathbb{R}^2 \to \mathbb{R}^m$.

The key observation is:

Fact. For all $x \in E$ except for those belonging to a uniformly purely unrectifiable set $N \subset E$, there is a unique differentiability direction $\tau(x)$ of f.

Fact. The vector field τ is Borel measurable and, up to a uniformly purely unrectifiable set, is uniquely determined by E. (In other words, it is independent of f!)

Reason. If E is contained in the non-differentiability set of both $f : \mathbb{R}^2 \to \mathbb{R}^m$ and $g : \mathbb{R}^2 \to \mathbb{R}^k$, the τ defined by (f, g) must coincide with the ones defined by f or g whenever f, g and (f, g) have a unique direction of differentiability.

A very wrong argument

We have seen that if every Lebesgue null set $E \subset \mathbb{R}^2$ is a set of non-differentiability, then every such set carries a uniquely (modulo purely unrectifiable sets) determined tangent direction. This seems to be in immediate contradiction with several arguments:

- How can it be unique? We can rotate the set!
- By Davies's theorem, every null set E is contained in another null set F such that through every point of E there pass lines lying in F whose directions are dense. How can the original set E remember its unique directions?

Perhaps such tangent field doesn't exist and this is the reason why a particular null set in $\mathbb{R}^2$ contains a point of differentiability of every Lipschitz $f : \mathbb{R}^2 \to \mathbb{R}$? **NO!**

Theorem. *For every Lebesgue null set $E \subset \mathbb{R}^2$ there is a Lipschitz $f : \mathbb{R}^2 \to \mathbb{R}^2$ non-differentiable at any point of E.*

The tangent field argument

Definition. Let $\mathcal{N}_{n,k}$ be the σ-ideal of subsets of $\mathbb{R}^n$ generated by sets for which there is a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ differentiable in at most k linearly independent directions.

So $\mathcal{N}_{n,0}$ are exactly the uniformly purely unrectifiable sets while $\mathcal{N}_{n,n-1}$ are the non-differentiability sets we are mainly interested in.

Theorem. *Whenever $E \in \mathcal{N}_{n,k}$, there is $\tau : E \to G(n,k)$ such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable in the direction $\tau(x)$ for all $x \in E$ except those belonging to an $\mathcal{N}_{n,k-1}$ set. Moreover, this property determines τ uniquely up to a set from $\mathcal{N}_{n,k-1}$.*

We shall say that τ is the k-dimensional tangent field of E.

Smallness of sets of directional non-differentiability

The set of (directional) nondifferentiability of $f\colon \mathbb{R}^n \to \mathbb{R}$ can be written as a countable union of sets E, for each of which we may find direction u and numbers $a < b$ such that

$$\liminf_{t\to 0} \frac{f(x+tu)-f(x)}{t} < a < b < \limsup_{t\to 0} \frac{f(x+tu)-f(x)}{t}.$$

Since our f is Lipschitz, such set E is null not only on every line in direction u, but also on every curve $\gamma\colon \mathbb{R} \to \mathbb{R}^n$ provided that $\|\gamma' - u\|$ is small enough.

We can do slightly better: If $\delta > 0$ is small enough, for every $\varepsilon > 0$ there is an open set $G \supset E$ such that the length of $G \cap \gamma$ is $< \varepsilon$ for every curve $\gamma\colon \mathbb{R} \to \mathbb{R}^n$ with $\|\gamma' - u\| < \delta$.

We will use this observation as one of the motivations to introduce a new way of measuring size of sets in $\mathbb{R}^n$.

How construct 'bad' Lipschitz functions?

- Given $E \subset \mathbb{R}$ define $f(x) = \int_0^x \mathbf{1}_E(t)\,dt$.
- Try: Given $E \subset \mathbb{R}^2$ let $f(x,y) = \int_0^y \mathbf{1}_E(x,t)\,dt$. For Lipschitz strips this is the second coordinate of the cutting map; but usually this does not give a Lipschitz functions.
- Try: For $y \geq 0$, define

$$f(x,y) = \inf_g \int_0^y \mathbf{1}_E(g(t),t)\,dt$$

where $\mathrm{Lip}(g) \leq 1$ and $g(y) = x$. This almost works, but a modification is still needed.

- For $y \geq 0$, define

$$f(x,y) = \inf_g \int_0^Y \mathbf{1}_E(g(t),t)\,dt$$

where $\mathrm{Lip}(g) \leq 1$, $Y \geq y$ and $g(Y) = x$.

Width of a set

Given a proper convex cone C and a unit vector $e \in C$, define $M = M_{C,e} : \mathbb{R}^n \to \mathbb{R}$ by

$$M(x) = \sup\{\lambda \in \mathbb{R} : x - \lambda e \in C\}.$$

Exercise. For $C = \{x \in \mathbb{R}^n : \langle x, e\rangle > \sigma|x|\}$, the corresponding function $M(x)$ is $M(x) = c(\langle x, e\rangle - \sigma|x|)$.

Definition The width $w(G) = w_{C,e}(G)$ of an open set $G \subset \mathbb{R}^n$ is defined as the supremum of the numbers

$$\int_{\{t:\gamma(t)\in G\}} M(\gamma'(t))\,dt$$

over all Lipschitz curves $\gamma : \mathbb{R} \to \mathbb{R}^n$ with $\gamma' \in C$.

For an arbitrary set $E \subset \mathbb{R}^n$ we define $w(E)$ as the infimum of $w(G)$ over all open sets G containing E.

Basic Lipschitz functions

Suppose G is an open set with finite $w(G) = w_{C,e}(G)$. Define $\omega(x)$ as the supremum of the numbers

$$-\lambda + \int_{t\in[a,b],\gamma(t)\in G} M(\gamma'(t))\,dt$$

among all $a, b \in \mathbb{R}$, $\lambda \geq 0$ and $\gamma : [a, b] \to \mathbb{R}^n$ such that $\gamma(b) - x = \lambda e$ and $\gamma' \in C$. Then

1. $0 \leq \omega(y) - \omega(x)$ if $y - x \in C$;
2. $M(y - x) \leq \omega(y) - \omega(x)$ if $y - x \notin C$; in particular, $\omega(y + te) - \omega(y) \leq t$ for every $y \in \mathbb{R}^n$ and $t \geq 0$;
3. $\omega(x + te) = \omega(x) + t$ if the segment $[x, x + te]$ lies entirely in G;
4. $0 \leq \omega(x) \leq w(G)$ for all $x \in \mathbb{R}^n$.

Also let $u(x) = x - \omega(x)e$ and notice that $u(G)$ is $n - 1$ rectifiable.

Example of a decomposition result

Let $C_1 = \{(x, y) \in \mathbb{R}^2 : x > |y|\}$, $e_1 = (1, 0)$. and $C_2 = \{(x, y) \in \mathbb{R}^2 : y > |x|\}$, $e_2 = (0, 1)$.

Theorem. *If $E \subset \mathbb{R}^2$ has $|E| < ab$, then there is open $G \subset \mathbb{R}^2$ with $w_1(G) < a$ such that $w_2(u_1(E \setminus G)) < b$.*

Idea of proof. Assume E open, choose an open set $G_0 \supset \mathbb{Q} \times \mathbb{R}$ with small width and then take a maximal open $G \supset G_0$ such that $w_1(G) - w_1(G_0) \leq (|G \cap E| - |G_0 \cap E|)/b$.

Application. Take open $H \supset u_1(E \setminus G)$ and the corresponding function u_2. Then $u_2 \circ u_1$ maps E onto a rectifiable set.

Using this for $E = [0, 1]^2 \setminus K$, where $K \subset [0, 1]^2$ is a compact of measure close to one, we get a map of $\mathbb{R}^2$ onto a set with nonempty interior such that the image of $\mathbb{R}^2 \setminus K$ is 1-rectifiable.

First attempt at understanding non-differentiability sets

We do not know if the non-differentiability sets of Lipschitz functions are exactly described by the property that they can be covered by countably many sets, each of which has width zero with respect to some cone.

We can, however, show that the non-differentiability sets have a (formally) stronger property that for every $\tau > 0$ they can be covered by (a finite number of) sets each of which has width zero with respect to some cone that is only τ-far from a halfspace.

These ideas also lead to a full description of sets from $\mathcal{N}_{n,k}$ (the σ-ideal generated by sets at which f can be differentiable in at most k linearly independent directions). We will state these results in a slightly different language.

Back to tangent fields

Definition. If $E \subset \mathbb{R}^n$, the mapping $\tau : E \to G(n,k)$ is called a k-dimensional tangent field of E if for every proper closed convex cone C the set $\{x \in E : \tau(x) \cap C = \{0\}\}$ has C-width zero.

Theorem. *The σ-ideal $\mathcal{N}_{n,k}$ of subsets of $\mathbb{R}^n$ generated by sets for which there is a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ differentiable in at most k linearly independent directions coincides with the family of those subsets of $\mathbb{R}^n$ that admit a k-dimensional tangent field.*

Theorem. *The k-dimensional tangent field of E (as defined above) is determined uniquely up to an $\mathcal{N}_{n,k-1}$-set and coincides with the tangent field defined via Lipschitz functions. In particular, every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable in the direction $\tau(x)$, for $\mathcal{N}_{n,k-1}$ almost every $x \in E$.*

The key open problem

The key open problem. *Is the restriction to $n \le 2$ necessary?* In other words, is it true that for all n that

- Lebesgue null sets coincide with $\mathcal{N}_{n,n-1}$;
- every Lebesgue null set in $\mathbb{R}^n$ admits an $n-1$ dimensional tangent field?

Problem. Is it true that for every $m < n$ there is a set $E \in \mathcal{N}_{n,m}$ such that every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at some point of E?

The answer is "yes" for $m = 1$ and any n; this is (essentially) the result mentioned in the beginning.

Dore and Maleva (using methods developed by Lindenstrauss, Tišer and Preiss in their study of infinite dimensional differentiability) will probably show "yes" also for $m = 2$.
But the methods of LTP cannot work for $m \ge 3$!

Description of sets of non-differentiability

Theorem. *For every set $E \subset \mathbb{R}^n$, the following are equivalent.*

- *There is a Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^n$ that is non-differentiable at any point of E.*
- *There is a sequence (possibly infinite) of Lipschitz functions $f_j : \mathbb{R}^n \to \mathbb{R}$ such that at every point of E at least one of the f_j is non-differentiable.*
- *The set E admits an $n-1$-dimensional tangent field.*
- *If $n \le 2$: E has Lebesgue measure zero.*

Corollary. *The following statements about a (σ-finite Borel) measure in $\mathbb{R}^n$ are equivalent.*

- Every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable μ almost everywhere.
- Every set admitting an $n-1$-dimensional tangent field is μ null.

Measures with unique $n-1$-dimensional tangent field

For what measures is the $n-1$-dimensional tangent field uniquely defined a.e?

(a) The measure has to be concentrated on $\mathcal{N}_{n,n-1}$ and
(b) absolutely continuous with respect to $\mathcal{N}_{n,n-2}$.

Observe that sets from $\mathcal{N}_{n,n-2}$ are purely $n-1$-unrectifiable. So for (b) it suffices that the measure be absolutely continuous with respect to purely $n-1$-unrectifiable sets.

The requirement (a) would be equivalent to singularity if $\mathcal{N}_{n,n-1}$ coincided with Lebesgue null sets, which we do not know. But the methods used to prove it when $n = 2$ are powerful enough to show that singularity plus absolute continuity with respect to purely $n-1$-unrectifiable sets suffices:

Theorem. *Every Lebesgue null set in $\mathbb{R}^n$ is a union of a set from $\mathcal{N}_{n,n-1}$ and a purely $n-1$-unrectifiable set.*

Understanding Alberti's theorem

Alberti's theorem can be understood as saying that certain class of (positive) measures in $\mathbb{R}^n$, namely those that arise as singular parts of derivatives of BV functions, have a.e. uniquely defined normal directions.

Since it is known that they are absolutely continuous with respect to the purely $n-1$-unrectifiable sets, and since they are singular by definition, they have uniquely defined $n-1$-dimensional tangent fields in our sense.

One can now adapt Alberti's argument of "rectifiably representable measures" to show that the gradient directions of the derivative of a BV function are, almost everywhere with respect to its singular part, orthogonal to our tangent field.

The case of uniformly purely unrectifiable sets

The construction of the required Lipschitz functions is considerably simpler in two special cases: functions non-differentiable on $\mathcal{N}_{n,0}$ sets and functions non-differentiable μ almost everywhere.

If $E \in \mathcal{N}_{n,0}$, for any unit vector e we can choose an open set $G \supset E$ with C-width null where $C \ni e$ is as close to the halfspace $\{x : \langle x, e\rangle \geq 0\}$ as we wish.

The function $\langle x, e\rangle - \omega(x)$ sees, from every point of G, some points in the direction e with slope almost one, but has local Lipschitz constant close to zero on G. This allows us to iterate the construction locally. Moving also the vectors e through a dense subset of the unit sphere, we get a function which is non-differentiable at any point of E in any direction.

Problem. For Borel sets in $\mathbb{R}^n$, does uniform pure unrectifiability coincide with pure unrectifiability?

Non-existence of weak derivatives

Let μ be a measure such that $\mu(S) > 0$ for some $S \in \mathcal{N}_{n,n-1}$.

Find $E \subset S$ with $\mu(E) > 0$ of C-width zero for some cone C.

Pick e in the interior of C and construct ω_j for $w = 1/j$.

So we get a sequence of functions $\omega_j : \mathbb{R}^n \to \mathbb{R}$ with uniformly bounded Lipschitz constants, converging to $\omega := 0$ and such that $\omega_j'(x; e) \geq 0$ everywhere and $\omega_j'(x; e) = 1$ for $x \in E$.

A moment's reflection shows that ω_j' cannot converge to $0 = \omega'$ in any weak sense with respect to μ. And a straightforward argument can make them C^1.

Theorem. For a measure μ in $\mathbb{R}^n$, weak derivatives of Lipschitz functions may be defined iff μ is absolutely continuous with respect to $\mathcal{N}_{n,n-1}$, hence iff every Lipschitz function is differentiable μ a.e.

For $n = 2$ we know that the above holds iff μ is absolutely continuous with respect to the Lebesgue measure. This answers a problem due to G. Mokobodzki.

Some problems of geometric measure theory: Part II: A proof of differentiability in null sets

based on joint work with J. Lindenstrauss[1] and J. Tišer[2]

[1]Hebrew University, Jerusalem
[2]Technical University, Prague

Barcelona, June 2009

What we prove, plus basic notation

Theorem. *Let $E \subset \mathbb{R}^2$ be a G_δ set containing all rational lines. Then every Lipschitz $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable at a point of E.*

Recall that the directional (or partial) derivative of $f : \mathbb{R}^2 \to \mathbb{R}$ at $x \in \mathbb{R}^2$ in direction $u \in \mathbb{R}^2$ is

$$f'(x; u) := \lim_{t\to 0} \frac{f(x + tu) - f(x)}{t}$$

provided that the limit exists.

In the rest of this proof, E will always denote the above set, $f : \mathbb{R}^2 \to \mathbb{R}$ a bounded Lipschitz function and

$$M = \{(x, u) \in E \times \mathbb{R}^2 : f'(x; u) \text{ exists}\} \neq \emptyset$$

Pre-basic observation

Observation. *Suppose that $\Theta \in C^1(\mathbb{R}^2)$, $\Phi \in C(E)$, $(x_0, u_0) \in M$ and the function*

$$(x, u) \in M \to f'(x; u) + \Phi(x) + \Theta(u)$$

attains its minimum at (x_0, u_0). Then f is differentiable at x_0.

Note that $u \to f'(x_0, u) + \Theta(u)$ attains minimum at u_0. So, assuming $f'(x_0)$ exists, we get by differentiating with respect to u that $f'(x_0) + \Theta'(u_0) = 0$. In the direction u_0 we may differentiate, so we have that $f'(x_0; u_0) = -\Theta'(u_0; u_0)$.

Pre-idea. Take a minimizing sequence (x_k, u_k) for $f'(x; u) + |u|^2$; with some luck it may converge to some (x_∞, u_∞) and f may happen to be differentiable at x_∞.

A long distance mean value estimate

Key Lemma. Let $h\colon [a, b] \longrightarrow \mathbb{R}$ be a Lipschitz function, $h(b) = h(a) = 0$, and let $\kappa = \mathrm{Lip}(h)$. Then there is a set $S \subset [a, b]$ such that $3\kappa |S| \geq \|h\|_\infty$ and for every $\xi \in S$,

1. h is differentiable at ξ;
2. $h'(\xi) \geq \dfrac{\|h\|_\infty}{3(b - a)}$;
3. $|h(t) - h(\xi)| \leq 48\sqrt{\kappa h'(\xi)}|t - \xi|$ for every $t \in [a, b]$.

A better idea, but still not enough. Try to construct a "minimizing sequence" (x_k, u_k) for $f'(x, u) + |u|^2$ by choosing the next term by this lemma, so we can get differentiability in the limit point in the limit direction. A minor problem is that $\sqrt{\cdot}$ is not additive, but the main problem is that the sequence u_k may not converge. All existing proofs solve this problem by some kind of perturbation argument.

A useful modification of the key lemma

Corollary. Let $g, h\colon \mathbb{R} \longrightarrow \mathbb{R}$ be functions, $0 < a, b \leq \delta$, and let $h(t) = g(t)$ for $-\delta \leq t \leq -a$ and $b \leq t \leq \delta$. Suppose further that $\mathrm{Lip}(h) \leq \kappa$ and that the parameter $0 < \tau \leq \kappa$ is such that

$$|h(0) - g(0)| \geq 13\tau \max(a, b) \text{ and } |g(t) - g(0)| \leq \frac{1}{\kappa}\tau^2|t| \text{ for } |t| \leq \delta.$$

Then there is $\xi \in (-a, b) \setminus \{0\}$ with the following properties:

1. h is differentiable at ξ;
2. $h'(\xi) \geq |h(0) - g(0)|/9 \max(a, b) \geq \tau$;
3. $|h(t) - h(\xi)| \leq 99\sqrt{\kappa h'(\xi)}|t - \xi|$ for every $t \in [-\delta, \delta]$.

Proof. Use Key Lemma for $h -$ *affine* and choose $\xi \neq 0$ as far from the end points of $(-a, a)$ as you can.

A variational principle (of Ekeland's type)

Theorem. *Let (M, d) be a complete metric space and let $F_j\colon M \times M \to [0, \infty]$, $j \geq 0$, be lower semi-continuous functions such that $F_j(x, x) = 0$ for all $x \in M$ and*

$$\inf_{d(x,y) > r_j} F_j(x, y) > 0 \quad \text{for some sequence } r_j \searrow 0 .$$

Then for every lower semi-continuous $h_0 : M \to \mathbb{R}$ bounded from below and every $\varepsilon_j > 0$ there is a sequence $(x_j)_{j=0}^{\infty}$ in M converging to some $x_\infty \in M$ such that, denoting $\varepsilon_\infty = 0$ and

$$h_j(x) = h_0(x) + \sum_{i=0}^{j-1} F_i(x_i, x),$$

we have that for $0 \leq i \leq \infty$,

$$h_\infty(x_\infty) \leq \min\{h_i(x_i),\ \varepsilon_i + \inf_{x \in M} h_i(x)\}$$

Notice that for $i = \infty$ the last inequality is just a complicated way of saying that h_∞ attains its minimum on M at x_∞.

Remetrizing $M = \{(x, u) \in E \times \mathbb{R}^2 : f'(x; u) \text{ exists}\}$

With the Euclidean metric, M is not complete. Since E is G_δ, we can find a metric d_E inducing the Euclidean topology on E such that (E, d_E) is complete. But M is not G_δ, and so we cannot use the same trick for it. So we change the topology of M.

We map M into the normed function space

$$\mathcal{P} = \{g \in \mathrm{Lip}(\mathbb{R}, \mathbb{R}) : g(0) = 0\}, \qquad \|g\|_{\mathcal{P}} = \sup_{t \in \mathbb{R} \setminus \{0\}} \frac{|g(t)|}{|t|}$$

by assigning to (x, u) the function $f_{x,u} \in \mathcal{P}$ defined by

$$f_{x,u}(t) = f(x + tu) - f(x).$$

The wished for metric on M is

$$d((x, u), (y, v)) = \max\{d_E(x, y), |u - v|, \varrho(f_{x,u}, f_{y,v})\}.$$

Fact. (M, d) is complete and $h_0(x, u) = f'(x, u) + |u|^2$ is continuous and lower bounded on (M, d).

The perturbation scheme

We define $F_i\colon M \times M \longrightarrow [0, \infty)$, $i \geq 0$ by

$$F_i((x, u), (y, v)) = \lambda_i d_E(x, y) + \beta_i |u - v|^2 + \gamma_i |f'(x; u) - f'(y; v)|^2 + \sigma_i \Delta_i((x, u), (y, v)),$$

where

$$\Delta_i((x, u), (y, v)) = \max\{0, \|f_{x,u} - f_{y,u}\|_{\mathcal{P}}\} - s_i\}$$

and $\lambda_i, \beta_i, \gamma_i, \sigma_i, s_i > 0$, $\sum(\lambda_i + \beta_i + \gamma_i + \sigma_i) < \infty$ and $s_i \to 0$.

The peculiarity in the definition of Δ_i is not a misprint: Δ_i really does not depend on v.

Fact. The functions F_i satisfy the assumptions of the variation principle.

Use of the variation principle.

We choose $\varepsilon_j > 0$ and find a sequence $(x_j, u_j)_{j=0}^{\infty}$ in M converging to $(x_\infty, u_\infty) \in M$ such that, denoting $\varepsilon_\infty = 0$ and

$$h_j(x,u) = h_0(x,u) + \sum_{i=0}^{j-1} F_i((x_i,u_i),(x,u))$$

we have that for $0 \le i \le \infty$,

$$h_\infty(x_\infty, u_\infty) \le \min\{h_i(x_i,u_i),\ \varepsilon_i + \inf_{(x,u)\in M} h_i(x,u)\}$$

We let

$$\begin{aligned} h_\infty(x,u) &= f'(x;u) + \Phi(x) + \Psi(u) + Q(x,u) + \Delta(x) \\ \Phi(x) &= \textstyle\sum_{i=0}^{\infty} \lambda_i d_E(x_i, x), \\ \Psi(u) &= |u|^2 + \textstyle\sum_{i=0}^{\infty} \lambda_i |u_i - u|^2, \\ Q(x,u) &= \textstyle\sum_{i=0}^{\infty} \sigma_i \big(f'(x_i;u_i) - f'(x;u)\big)^2, \\ \Delta(x) &= \textstyle\sum_{i=0}^{\infty} \Delta_i((x_i,u_i),(x,u)). \end{aligned}$$

Guessing the derivative

Here it is important that Δ does not depend on u.

Assuming that $f'(x_\infty)$ exists, we differentiate $u \to h_\infty(x_\infty, u)$ at the point $u = u_\infty$ to get that

$$f'(x_\infty) + \Psi'(u_\infty) - \kappa f'(x_\infty) = 0$$

where

$$\kappa = 2\sum_{i=0}^{\infty} \sigma_i \big(f'(x_i;u_i) - f'(x_\infty;u_\infty)\big)$$

The choice of small σ_i gives $|\kappa| < 1$, so

$$f'(x_\infty) = -\Psi'(u_\infty)/(1+\kappa) =: L$$

provide the derivative exists.

In the direction u_∞ we may differentiate, so we get

$$f'(x_\infty; u_\infty) = L(u_\infty).$$

The final contradiction (1/2)

Assumption. There is $\eta > 0$ such that for every $\delta > 0$ there is $v \in X$ with $0 < |v| < \delta$ and

$$|f(x_\infty + v) - f(x_\infty) - L(v)| > 1000\eta|v|.$$

We choose small δ and the corresponding v and define the path γ as in the pre-basic observation, but this time choose the point ξ by using the corollary of the Key Lemma. We again define $y = \gamma(\xi)$ and $w = \gamma'(\xi)$. The inequality

$$\begin{aligned} &f'(y;w) + \Phi(y) + \Psi(w) + Q(y,w) \\ &\quad < f'(x_\infty;u_\infty) + \Phi(x_\infty) + \Psi(u_\infty) + Q(x_\infty,u_\infty) - 900\beta\eta \end{aligned}$$

is proved as in the pre-basic observation. (We could have included the quadratic term in the derivatives already there.)

The final contradiction (2/2)

It remains to estimate the Δ term. Recall that

$$\Delta_i((x_i,u_i),(x,u)) = \max\{0, \|f_{x_i,u_i} - f_{x,u_i}\|_{\mathcal{P}}\} - s_i\}.$$

This is done by a somewhat involved induction, whose key points are:

1. The contribution to $\|f_{x_i,u_i}, f_{y,u_i}\|_{\mathcal{P}} - \|f_{x_i,u_i}, f_{x_\infty,u_i}\|_{\mathcal{P}}$ coming from large values of the variable (say t) is tiny (by choosing δ small.
2. So big contribution may come only from values of t where we have control by the Key Lemma and near linearity of f at x_i in direction u_i ($0 \le i \le$ some large k).
3. The Q-term guarantees that the differences of derivatives are tiny, so all control for such t is by $\sqrt{\ }$ of the possible change of derivatives.
4. We follow this change by induction and show it stays below s_i. So these t do not contribute to Δ at all. $\boxed{\text{END}}$

Proof of the pre-basic observation (1/2)

We claim that $f'(x_0) = L := -\Theta'(x_0)$. We know $f'(x_0, u_0) = Lu_0$.

Assumption. There is $\eta > 0$ such that for every $\delta > 0$ there is $v \in \mathbb{R}^2$ with $0 < |v| < \delta$ and

$$|f(x_0 + v) - f(x_0) - Lv| > 3\eta|v|.$$

Fix $\beta > 0$ so that for $|w - u_0| < \beta$,

$$|\Theta(w) - \Theta(u_0) - L(w - u_0)| \le \eta|w - u_0|$$

Choose small $\delta > 0$ (depending on η and β), find v as above; make $x_0 + v$ rational and find $|v|/\beta < a, b < 2|v|/\beta$ so that

$$[x_0 - au_0, x_0 + v] \cup [x_0 + v, x_0 + bu_0] \subset E.$$

Define $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ and $h : \mathbb{R} \to \mathbb{R}$ by

$$\gamma(t) = x_0 + tu_0 + \max\{0, \min\{1 + t/a, 1 - t/b\}\} v$$
$$h(t) = f(\gamma(t)) - f(x_0) - L(\gamma(t) - x_0)$$

Then h is Lipschitz on $[-a, b]$ and, if δ is small,

$$|h(0)| > 3\eta|v| > \eta|v| > \max\{|h(-a)|, |h(b)|\}$$

Proof of the pre-basic observation (2/2)

Reminder: $h(t) = f(\gamma(t)) - f(x_0) - L(\gamma(t) - x_0)$

Find $-a < \xi < b$, $\xi \ne 0$ so that
$h'(\xi) < -2\eta|v|/\max\{a, b\} < -2\eta\beta$.
Let $y = \gamma(\xi)$ and $w = \gamma'(\xi)$. So $f'(y; w) - Lw = h'(\xi) < -2\eta\beta$.
Since $|w - u_0| \le |v|/min(a, b) < \beta$, we may estimate $\Theta(w)$ expansion at u_0. Also y is close to x_0 so $\Phi(y)$ is close to $\Phi(x_0)$.
Together we get

$$\begin{aligned} & f'(y; w) + \Phi(y) + \Theta(w) \\ &\quad < f'(y; w) + \Phi(x_0) + \eta\beta + \Theta(u_0) + \Theta'(u_0; w - u_0) + \eta\beta \\ &\quad = (f'(y; w) + Lw) + \Phi(x_0) + f'(x_0, u_0) + \Theta(u_0) + 2\eta\beta \\ &\quad < f'(x_0, u_0) + \Phi(x_0) + \Theta(u_0). \end{aligned}$$

Recalling that $(y, w) \in M$, we have our contradiction.

Proof of the Key Lemma (1/3)

Lemma. Denote for $t \in [a, b]$

$$H(t) = \sup\Big\{\frac{|h(\beta) - h(\alpha)|}{\beta - \alpha} \;\Big|\; a \le \alpha \le t \le \beta \le b,\ \alpha < \beta\Big\}.$$

Then for every $\lambda > 0$,

$$|\{t \in [a, b] \mid H(t) > \lambda\}| \le \frac{8}{\lambda}\int_a^b \max\{0, h'(t)\}\, dt.$$

Proof. Let $\mathbf{M}$ be the Hardy-Littlewood maximal operator. Since $H(t) \le 2\,\mathbf{M}h'(t)$, we get that

$$|\{t \in [a, b] \mid H(t) > \lambda\}| \le \frac{4}{\lambda}\int_a^b |h'(t)|\, dt.$$

The condition $h(a) = h(b)$ implies that

$$\int_{\{t|h'(t)>0\}} h'(t)\, dt = \int_{\{t|h'(t)<0\}} -h'(t)\, dt, \text{ so}$$
$$\int_a^b |h'(t)|\, dt = 2\int_a^b \max\{0, h'(t)\}\, dt.$$

Proof of the Key Lemma (2/3)

Denote by $S \subset [a, b]$ the set of all points $\xi \in (a, b)$ for which the three statements hold and suppose to the contrary that the measure $|S| < \frac{1}{3\kappa}\|h\|_\infty$.
Then for a.e. $t \in [a, b] \setminus S$ either

$$h'(t) < \frac{\|h\|_\infty}{3(b - a)} \quad \text{or} \quad h'(t) \ge \frac{\|h\|_\infty}{3(b - a)} \text{ and } H(t) > 48\sqrt{\kappa h'(t)}.$$

Recalling also that $h' \le \kappa$, we see that for a.e. $t \in [a, b]$,

$$\max\{0, h'(t)\} < \max\{0, h'(t)\}\chi_S(t) + \frac{\|h\|_\infty}{3(b - a)} + \min\Big\{\kappa, \frac{1}{48^2\kappa}H^2(t)\Big\},$$

where χ_S denotes the characteristic function of S.

Proof of the Key Lemma (3/3)

Hence

$$
\begin{aligned}
\int_a^b \max\{0,h'(t)\}\,dt &< \int_S \max\{0,h'(t)\}\,dt + \int_a^b \frac{\|h\|_\infty}{3(b-a)}\,dt \\
&\qquad + \int_a^b \min\Big\{\kappa, \frac{1}{48^2\kappa}H^2(t)\Big\}\,dt \\
&\le \kappa\,|S| + \frac{1}{3}\|h\|_\infty + \int_0^\kappa \Big|\Big\{t\in[a,b] \;\Big|\; \frac{1}{48^2\kappa}H^2(t) > \lambda\Big\}\Big|\,d\lambda \\
&\le \frac{2}{3}\|h\|_\infty + \int_0^\kappa \frac{1}{6\sqrt{\kappa\lambda}}\,d\lambda \int_a^b \max\{0,h'(t)\}\,dt \\
&\le \frac{2}{3}\int_a^b \max\{0,h'(t)\}\,dt + \frac{2}{6}\int_a^b \max\{0,h'(t)\}\,dt \\
&= \int_a^b \max\{0,h'(t)\}\,dt.
\end{aligned}
$$

Proof of the variational principle (1/1)

Diminishing the ε_j if necessary, we assume that

$$\inf_{d(x,y)>r_j} F_j(x,y) > \varepsilon_j.$$

Choose x_0 so that $h_0(x_0) < \varepsilon_0 + \inf_{x\in M} h_0(x)$. Recursively define $h_{j+1}(x) = h_j(x) + F_j(x_j,x)$ and choose x_{j+1} so that

$$h_{j+1}(x_{j+1}) \le \min\big(h_j(x_j), \varepsilon_{j+1} + \inf_{x\in M} h_{j+1}(x)\big).$$

We show that x_n is Cauchy. For that notice that for $k>j$,

$$h_j(x_k) + F_j(x_j,x_k) \le h_k(x_k) \le \varepsilon_j + \inf_{x\in M} h_j(x) \le \varepsilon_j + h_j(x_k).$$

Hence $F_j(x_j,x_k) \le \varepsilon_j$ and so $d(x_j,x_k) \le r_j$.

Second Order Operators in Divergence Form and Quasiconformal Mappings

Zhong

Title of the course: *Second order operators in divergence form and QC mappings.*

Abstract: The series of lectures will focus on the well-known theory of De Giorgi-Nash-Moser for linear elliptic equations. We will prove the local boundedness, Hölder continuity and the Harnack inequality of solutions of the following second order, linear, elliptic equations with divergence structure

$$\operatorname{div}(\mathbb{A}(x)\nabla u(x)) = 0,$$

where $\mathbb{A}(x) = [a_{ij}(x)]_{i,j=1,2,\dots,n}$, defined in a domain $\Omega \subset \mathbb{R}^n, n \geq 2$, is a symmetric matrix with measurable coefficients satisfying for $0 < \lambda \leq \Lambda < \infty$

$$\lambda|\xi|^2 \leqslant \langle \mathbb{A}(x)\xi, \xi\rangle \leqslant \Lambda|\xi|^2$$

for all $\xi \in \mathbb{R}^n$ and for almost every $x \in \Omega$. We will also mention the connections between this type of equations and quasiconformal mappings.

Outline of the course:

- **Lecture 1:** Second order elliptic euqations with divergence structure: introduction.
- **Lecture 2:** Moser's iteration.
- **Lecture 3:** De Giorgi's method.
- **Lecture 4:** De Girogi's method (continued).
- **Lecture 5:** Connections with quasiconformal mappings; open problems.

Some references:

(1) D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order.* 2nd ed., Springer-Verlag, New York (1983).
(2) O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations.* Academic Press, New York (1968).
(3) Q. Han and F. H. Li, *Elliptic partial differential equations.* Courant Lecture Notes in Mathematics, 1 (1997).

Xiao Zhong, Department of Mathematics and Statistics, University of Jyväskylä,, P.O. Box 35 (MaD), Fin-40014, University of Jyväskylä, Finland.
E-mail address: `zhong@maths.jyu.fi`

DE GIORGI–NASH–MOSER THEORY

XIAO ZHONG

1. Introduction

1.1. **Equations.** We consider the second order, linear, elliptic equations with divergence structure

$$\operatorname{div}(\mathbb{A}(x)\nabla u(x)) = \sum_{i,j=1}^{n} \partial_{x_i}(a_{ij}(x)\partial_{x_j} u(x)) = 0. \tag{1}$$

Here $\mathbb{A}(x) = [a_{ij}(x)]_{i,j=1,2,\dots,n}$ is an $n \times n$ symmetric matrix with measurable entries $a_{ij}(x)$, defined in a domain $\Omega \subset \mathbb{R}^n, n \geq 2$. We assume the following ellipticity and boundedness conditions. That is, we assume that

$$\lambda|\xi|^2 \leqslant \langle \mathbb{A}(x)\xi, \xi\rangle = \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leqslant \Lambda|\xi|^2 \tag{2}$$

for all $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ and for almost every $x \in \Omega$. Here $0 < \lambda \leq \Lambda < \infty$ are constants.

Note that since $\mathbb{A}(x)$ is a symmetric matrix, it has n real eigenvalues $\lambda_i(x), i = 1, \dots, n$. The condition (2) is equivalent to

$$\lambda \leq \lambda_i(x) \leq \Lambda$$

for all $i = 1, \dots, n$, and almost every $x \in \Omega$ (Exercise).

Example 1.1. Let $\mathbb{A}(x)$ be the identity matrix. Then the condition (2) is true with $\lambda = \Lambda = 1$, and the equation (1) is reduced to the Laplace equation

$$\Delta u = \sum_{i=1}^{n} \partial_{x_i}\partial_{x_i} u = 0.$$

Example 1.2. Let α be a constant, $0 < \alpha < 1$. Define $\mathbb{A}(x)$ in $\mathbb{R}^2$ as

$$\mathbb{A}(x) = \begin{pmatrix} \frac{x_1^2+\alpha^2 x_2^2}{|x|^2} & (1-\alpha^2)\frac{x_1x_2}{|x|^2} \\ (1-\alpha^2)\frac{x_1x_2}{|x|^2} & \frac{\alpha^2x_1^2+x_2^2}{|x|^2} \end{pmatrix}, \quad x = (x_1, x_2).$$

Then we have (Exercise)

$$\alpha^2|\xi|^2 \leq \langle \mathbb{A}(x)\xi, \xi\rangle \leq |\xi|^2, \quad \forall\, x \in \mathbb{R}^2, \xi \in \mathbb{R}^2.$$

Define the function $u : B(0,1) = \{y \in \mathbb{R}^2 : |y| < 1\} \to \mathbb{R}$ as

$$u(x) = |x|^{\alpha-1}x_1, \quad \text{for} \;\; x = (x_1, x_2) \in B(0,1).$$

Date: May 2009.

2000 *Mathematics Subject Classification.* Primary 30J70; Secondary 35D10, 35J15.

Key words and phrases. Continuity, Weak solutions, Second order elliptic equations, Sobolev spaces, De Giorgi–Nash–Moser theory.

Supported by the Academy of Finland, project 127639.

Then it is a weak solution (see definition in Section 2) of equation (1) with the coefficients $\mathbb{A}(x)$ defined above (Exercise).

1.2. **Motivation: a variational problem.** We start with the problems, raised by Hilbert in the ICM in 1990.

- 20th problem: Has not every regular variational problem a solution, provided certain assumptions regarding the given boundary conditions are satisfied, and provided also if need be that the notions of a solution shall be suitably extended?
- 19th problem: Are the solutions of regular problems in the calculus of variations always necessarily analytic?

These problems are stated in a general way. We will consider the following specific variational problem to illustrate these problems and their solutions.

Let $\Omega \subset \mathbb{R}^n$ be a bounded smooth domain and $F : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. We assume that it satisfies for constants $0 < \lambda \leq \Lambda < \infty$

$$\lambda|\xi|^2 \leq \langle D^2F(\eta)\xi, \xi\rangle \leq \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \eta \in \mathbb{R}^n. \tag{3}$$

Now we consider the following functional

$$I(v) = \int_\Omega F(\nabla v)\, dx, \tag{4}$$

among the admissible class

$$K' = \{v : v \in C^1(\bar{\Omega}) \text{ and } v = \phi \text{ on } \partial\Omega\},$$

where $\phi \in C^1(\bar{\Omega})$ is a given function. We say that $u \in K$ is a minimizer of the functional I among the class K, if

$$I(u) \leq I(v), \quad \forall\, v \in K'.$$

Now the problems are the existence of minimizers (20th problem) and the regularity of minimizers (19th problem).

We can not prove the existence of the minimizers directly in the class K', due to the lack of compactness of the space $C^1(\bar{\Omega})$. We need to extend the space $C^1(\bar{\Omega})$ to a bigger space. The classical derivatives are extended to the weak ones, and the classical solutions to the weak ones, as suggested by Hilbert. A natural function space for this variational problem is the Sobolev space $W^{1,2}(\Omega)$. We will give a brief introduction to the Sobolev spaces in Section 2. Now let

$$K = \{v : v \in W^{1,2}(\Omega) \text{ and } v - \phi \in W_0^{1,2}(\Omega)\}.$$

We can easily prove the existence of minimizers of the functional I in K, by the directly method in the calculus of variations.

THEOREM 1.1. *Suppose that $F \in C^\infty(\mathbb{R}^n)$ satisfies (3) and $\phi \in W^{1,2}(\Omega)$ is given. Then there is a unique $u_0 \in K$ such that*

$$I(u_0) = \inf_{u\in K} I(u).$$

Proof. Read Section 2 first. It is easy to show from (3) that (Exercise)

$$c_1|\eta|^2 - c_2 \leq F(\eta) \leq c_3|\eta|^2 + c_4, \tag{5}$$

where $c_1 > 0, c_3 > 0, c_2, c_4$ are constants depending only on $\lambda, \Lambda, F(0), \nabla F(0)$. Now let $m = \inf_{u\in K} I(u)$. It is easy to see that $-\infty < m < \infty$. Let $\{u_i\}_{i=1}^\infty \subset K$ be a minimizing sequence, that is, $I(u_i) \to m$ as $i \to \infty$. It is easy to prove that $\{u_i\}_{i=1}^{\infty}$ is a bounded sequence in $W^{1,2}(\Omega)$. Then by the weak compactness theorem, Theorem 2.5, there is a subsequence, still denoted by itself, such that it converges weakly in $W^{1,2}(\Omega)$ to a function u_0. Since $u_i \in K$, we also have $u_0 \in K$. Now we claim that

$$\liminf_{i\to\infty} \int_\Omega F(\nabla u_i)\,dx \geq \int_\Omega F(\nabla u_0)\,dx,$$

which shows that $I(u_0) = m$ and hence u_0 is a minimizer of the functional I among the class K. Indeed, we have

$$F(\eta) \geq F(\eta_0) + \langle \nabla F(\eta_0), \eta - \eta_0 \rangle + \frac{\lambda}{2}|\eta - \eta_0|^2, \quad \forall\, \eta, \eta_0 \in \mathbb{R}^n. \tag{6}$$

The proof of (6) is as follows. We have that

$$\begin{aligned} F(\eta) - F(\eta_0) - \langle \nabla F(\eta_0), \eta - \eta_0 \rangle &= \int_0^1 \frac{d}{dt} F(t\eta + (1-t)\eta_0)\,dt - \langle \nabla F(\eta_0), \eta - \eta_0 \rangle \\ &= \int_0^1 \langle \nabla F(t\eta + (1-t)\eta_0) - \nabla F(\eta_0), \eta - \eta_0 \rangle\,dt, \end{aligned}$$

and that

$$\begin{aligned} \nabla F(t\eta + (1-t)\eta_0) - \nabla F(\eta_0) &= \int_0^1 \frac{d}{ds} \nabla F(st\eta - st\eta_0 + \eta_0)\,ds \\ &= t\int_0^1 D^2 F(st\eta - st\eta_0 + \eta_0)(\eta - \eta_0)\,ds. \end{aligned}$$

Thus (6) follows from (3). Now by (6), we have

$$\int_\Omega F(\nabla u_i) - F(\nabla u_0)\,dx \geq \int_\Omega \langle \nabla F(\nabla u_0), \nabla u_i - \nabla u_0 \rangle\,dx.$$

The last integral goes to zero as i goes to infinity, since u_i converges to u_0 weakly in $W^{1,2}(\Omega)$. Thus the claim is true. It remains to prove the uniqueness of minimizer. Let $u_0, \bar{u}$, be minimizers and let $u = (u_0 + \bar{u})/2$. Note that $u \in K$. Then by (6), we have that

$$F(\nabla u_0) \geq F(\nabla u) + \langle \nabla F(\nabla u), \nabla u_0 - \nabla u \rangle + \frac{\lambda}{2}|\nabla u - \nabla u_0|^2,$$

and that

$$F(\nabla \bar{u}) \geq F(\nabla u) + \langle \nabla F(\nabla u), \nabla \bar{u} - \nabla u \rangle + \frac{\lambda}{2}|\nabla u - \nabla \bar{u}|^2.$$

Adding this two inequalities together, we obtain that

$$F(\nabla u_0) + F(\nabla \bar{u}) \geq 2F(\nabla u) + \frac{\lambda}{4}|\nabla u_0 - \nabla \bar{u}|^2.$$

Integrate both sides over Ω. We arrive at

$$2m \geq 2\int_\Omega F(\nabla u)\,dx + \frac{\lambda}{4}\int_\Omega |\nabla u_0 - \nabla \bar{u}|^2\,dx \geq 2m + \frac{\lambda}{4}\int_\Omega |\nabla u_0 - \nabla \bar{u}|^2\,dx,$$

from which, we deduce that $u_0 = \bar{u}$. This proves the uniqueness. The theorem is proved. $\square$

The above theorem gives a positive solution to the 20th problem. The good point to study the weak minimizers (solutions) is that it is easy to prove the existence, but the price we have to pay is the regularity. Now we turn to the 19th problem: the regularity of minimizers. Our goal is to show that the minimizer u_0 is smooth. In order to do this, first, we study the Euler-Lagrange equation corresponding to the functional I. It is easy to show that the minimizer u_0 that we obtained in Theorem 1.1 is a weak solution of the Euler-Lagrange equation. Note that at this moment, we only know that u_0 is from the Sobolev space $W^{1,2}(\Omega)$.

THEOREM 1.2. *Let u_0 be the minimizer as in Theorem 1.1. Then u_0 is a weak solution of the Euler-Lagrange equation*

$$\operatorname{div}(\nabla F(\nabla u_0)) = 0, \qquad \textit{in } \Omega, \tag{7}$$

that is,

$$\int_\Omega \langle \nabla F(\nabla u_0), \nabla \varphi \rangle \, dx = 0, \quad \forall \, \varphi \in C_0^\infty(\Omega). \tag{8}$$

Proof. Fix $\varphi \in C_0^\infty(\Omega)$. Define a function $g : \mathbb{R} \to \mathbb{R}$ as

$$g(t) = I(u_0 + t\varphi) = \int_\Omega F(\nabla u_0 + t\nabla\varphi) \, dx.$$

Then $g \in C^1(\mathbb{R})$ (Prove it), and we have

$$g'(t) = \int_\Omega \langle \nabla F(\nabla u_0 + t\nabla\varphi), \nabla\varphi \rangle \, dx.$$

Since u_0 is a minimizer, the function g reaches its minimum at 0. Thus $g'(0) = 0$, which gives (8). $\square$

Next, we will show that $u_0 \in W^{2,2}_{\mathrm{loc}}(\Omega)$ and that its weak derivatives are weak solutions of equation (1) with suitable coefficients $\mathbb{A}(x)$. In this way, we build up the connection between this variational problem and equation (1).

THEOREM 1.3. *Let u_0 be the minimizer as in Theorem 1.1. Then $u_0 \in W^{2,2}_{\mathrm{loc}}(\Omega)$ and $v_i = \partial_{x_i} u_0, i = 1, 2, \ldots, n$, is a weak solution of the equation*

$$\operatorname{div}(\mathbb{A}(x)\nabla v_i) = 0, \qquad \textit{in } \Omega, \tag{9}$$

where $\mathbb{A}(x) = D^2F(\nabla u_0(x))$.

The formal proof involves the difference quotients. Informally, we may just differentiate equation (7) with respect to $x_i, i = 1, 2, \ldots, n$, to obtain equation (9). The essential point is the Caccioppoli inequality, which will be studied extensively in this course. Roughly speaking, the Caccioppoli inequality says that the L^2-norm of the derivatives of the solutions is controlled by that of the solutions. We refer to Section 2 for the notations in the proof.

Proof. Fix a cut-off function $\eta \in C_0^\infty(\Omega)$. Let Ω' be an open set such that $spt(\eta) \subset \Omega' \Subset \Omega$. Let $\varphi = \Delta_i^{-h}((\Delta_i^h u_0)\eta^2)$, where $i = 1, 2, \ldots, n$ and h is so small that $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)/8$, and

$$\Delta_i^h v(x) = \frac{v(x + he_i) - v(x)}{h}$$

is the i^{th} difference quotient of size h. Since $u_0 \in W^{1,2}(\Omega)$, we have $\varphi \in W_0^{1,2}(\Omega)$ (Exercise). By Theorem 1.2, u_0 is a weak solution of equation (7). Thus,

$$
\begin{aligned}
0 &= \int_\Omega \langle \nabla F(\nabla u_0), \nabla\varphi\rangle \, dx \\
&= \int_\Omega \langle \nabla F(\nabla u_0), \Delta_i^{-h}(\nabla((\Delta_i^h u_0)\eta^2))\rangle \, dx \qquad (10)\\
&= -\int_\Omega \langle \Delta_i^h \nabla F(\nabla u_0), \nabla((\Delta_i^h u_0)\eta^2)\rangle \, dx,
\end{aligned}
$$

where in the second equality, we used the fact that ∇ and Δ_i^h are commutative, and in the last equality, the so-called integration by parts for difference quotients. Now we write

$$
\begin{aligned}
\Delta_i^h \nabla F(\nabla u_0(x)) =& \frac{\nabla F(\nabla u_0(x+he_i)) - \nabla F(\nabla u_0(x))}{h} \\
=& \frac{1}{h}\int_0^1 \frac{d}{dt} \nabla F(t\nabla u_0(x+he_i) + (1-t)\nabla u_0(x))\, dt \\
=& \int_0^1 D^2F(t\nabla u_0(x+he_i) + (1-t)\nabla u_0(x))\, dt \Delta_i^h \nabla u_0(x) \\
=& \mathbb{B}(x)\Delta_i^h \nabla u_0(x),
\end{aligned}
$$

and

$$\nabla((\Delta_i^h u_0)\eta^2) = \eta^2 \Delta_i^h \nabla u_0 + 2\eta\nabla\eta\Delta_i^h u_0.$$

Thus (10) becomes

$$\int_\Omega \langle \mathbb{B}(x)\Delta_h^i \nabla u_0, \Delta_i^h \nabla u_0\rangle \eta^2 \, dx = -2\int_\Omega \langle \mathbb{B}(x)\Delta_i^h \nabla u_0, \nabla\eta\rangle \eta\Delta_i^h u_0 \, dx.$$

Note that the matrix $\mathbb{B}(x)$ also satisfies (2). By Cauchy-Schwarz inequality and Hölder's inequality, we can easily deduce the following Caccioppoli type inequality

$$\int_\Omega |\Delta_i^h \nabla u_0|^2 \eta^2 \, dx \le \frac{4\Lambda}{\lambda}\int_\Omega |\Delta_i^h u_0|^2 |\nabla\eta|^2 \, dx,$$

from which, together with Theorem 2.6, proves the theorem. □

One question here: can we repeat the above argument to prove that $u_0 \in W^{3,2}(\Omega)$? The answer is no. One may informally differentiate equation (9) and see if we can obtain the Caccioppoli type inequality to control the L^2-norm of the third order derivatives of u_0.

Finally, the main goal of this course is to prove that the weak solutions of equation (1) are Hölder continuous.

THEOREM 1.4. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (1). Then $u \in C^{0,\alpha}(\Omega)$, where $0 < \alpha = \alpha(n, \lambda, \Lambda) \le 1$.*

In the planar case, the study goes back to the work of Morrey [11, 12], see [21], [10] and [17] for the study of the best Hölder continuity exponent. In higher dimensions ($\mathbb{R}^n, n \geqslant 3$), Hölder continuity of solutions was settled in the late 1950's by De Giorgi [1] and Nash [15]. Hölder continuity also follows from the Harnack inequality, due to Moser [13, 14].

Now, we go back to our variational problem. Combining Theorem 1.2 and Theorem 1.4, we obtain that the minimizer $u_0 \in C^{1,\alpha}(\Omega)$ for some $\alpha > 0$. Then we can show that actually $u_0 \in C^\infty(\Omega)$, by the Schauder estimates.

The above is the line to deal with Hilbert's 20th and 19th problems. As we can see, the essential and difficult point is the De Giorgi-Nash-Moser theory: the Hölder continuity of solutions of equation (1) in Theorem 1.4. In this course, we will first learn Moser's method and then De Giorgi's method. The students are expected to compare these two methods. While these two methods are further exploited and applied to many other problems, the argument of Nash is rather difficult to penetrate and consequently his work has not been extensively used in the literature. We refer to [6] for the applications of Nash's ideas.

One comment: To my knowledge, De Giorgi's, Nash's and Moser's methods are the only existing approaches to prove the Hölder continuity of weak solutions of equation (1). I do not know any other way to prove Theorem 1.4.

2. Sobolev spaces

2.1. **A brief introduction to Sobolev spaces.** We refer to [7], [4] and [9] for the proofs of Theorems in this subsection. Let Ω be an open set in $\mathbb{R}^n$.

2.1.1. *Hölder space.*

Definition 2.1. We say a function $u \in C^{0,\alpha}(\Omega)$, if for every open $\Omega' \Subset \Omega$,

$$||u||_{C^{0,\alpha}(\bar{\Omega}')} = \sup_{x\in\Omega'} |u(x)| + \sup_{x,y\in\Omega', x\neq y} \frac{|u(x)-u(y)|}{|x-y|} < \infty.$$

2.1.2. *Weak derivatives.* Let $C_0^\infty(\Omega)$ denote the space of infinitely differentiable functions with compact support in Ω.

Definition 2.2. Suppose that $u, v \in L^1_{\mathrm{loc}}(\Omega)$. We say that v is the i^{th} weak partial derivative of u, written

$$\partial_{x_i} u = v,$$

provided that

$$\int_\Omega u \partial_{x_i}\phi \, dx = -\int_\Omega v\phi \, dx$$

for all test functions $\phi \in C_0^\infty(\Omega)$.

A weak partial derivative of a function u, if it exists, is uniquely defined up to a set of measure zero. We denote by $\nabla u = (\partial_{x_1} u, \partial_{x_2} u, \ldots, \partial_{x_n} u)$ the weak gradient of u.

2.1.3. *Definition of Sobolev spaces.*

Definition 2.3. The Sobolev space

$$W^{1,p}(\Omega), \quad 1 \le p \le \infty$$

consists of all locally integrable functions $u : \Omega \to \mathbb{R}$ such that $u \in L^p(\Omega)$ and the weak derivatives $\partial_{x_i} u \in L^p(\Omega)$ for all $i = 1, 2, \ldots, n$. We define its norm to be

$$||u||_{W^{1,p}(\Omega)} = \begin{cases} \left(\int_\Omega |u|^p \, dx + \sum_{i=1}^n \int_\Omega |\partial_{x_i} u|^p \, dx\right)^{1/p}, & (1 \le p < \infty); \\ \sup_\Omega |u| + \sum_{i=1}^n \sup_\Omega |\partial_{x_i} u|, & (p = \infty). \end{cases}$$

We denote by $W_0^{1,p}(\Omega)$ the closure of $C_0^\infty(\Omega)$ in $W^{1,p}(\Omega)$ with respect to the norm defined above. Thus $u \in W_0^{1,p}(\Omega)$ if and only if there exist functions $u_k \in C_0^\infty(\Omega)$ such that

$$u_k \to u \quad \text{in } W^{1,p}(\Omega),$$

that is,

$$\lim_{k\to\infty} ||u_k - u||_{W^{1,p}(\Omega)} = 0.$$

Exercise: $W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$ are Banach spaces.

2.1.4. *Inequalities.*

THEOREM 2.1 (Gagliardo-Nirenberg-Sobolev inequality)**.** *Assume $1 \le p < n$. There is a constant $c = c(n,p) > 0$ such that*

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

for all $u \in W_0^{1,p}(\Omega)$, where $p^* = np/(n-p)$.

We also have the following version of Sobolev inequality.

THEOREM 2.2. *Assume that $1 \le p < n$. Suppose that $u \in W^{1,p}(B(y,r))$ and $|\{x \in B(y,r) : u(x) = 0\}| \ge \delta |B(y,r)|$ for some $\delta > 0$. Then*

$$\left(\int_{B(y,r)} |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} \le c\left(\int_{B(y,r)} |\nabla u|^p\,dx\right)^{\frac{1}{p}},$$

where $c = c(n,p,\delta) > 0$.

The following theorem is an easy consequence of Theorem 2.2.

THEOREM 2.3. *Suppose that $u \in W^{1,1}(B(y,r))$ and that $|\{x \in B(y,r) : u(x) \le k\}| \ge \delta|B(y,r)|$ for some $\delta > 0$ and $k \in \mathbb{R}$. Then*

$$(l-k)|\{x \in B(y,r) : u(x) > l\}|^{1-\frac{1}{n}} \le c\int_{k<u<l} |\nabla u|\,dx,$$

for any $l > k$, where $c = c(n,\delta) > 0$.

THEOREM 2.4 (Poincaré inequality)**.** *Assume $1 \le p < \infty$. Then there is a constant $c = c(n,p) > 0$ such that*

$$\int_{B(y,r)} |u - u_{B(y,r)}|^p\,dx \le cr^p \int_{B(y,r)} |\nabla u|^p\,dx$$

for every $u \in W^{1,p}(\Omega)$ and every ball $B(y,r) \subset \Omega$, where $u_{B(y,r)} = ⨍_{B(y,r)} u\,dx$ is the average of u over the ball $B(y,r)$.

2.1.5. *Weak compactness theorem.*

THEOREM 2.5 (Weak compactness theorem)**.** *Assume $1 < p < \infty$. Suppose that $\{u_k\}_{k=1}^{\infty}$ is a bounded sequence in $W^{1,p}(\Omega)$. Then there are a function $u \in W^{1,p}(\Omega)$ and a subsequence of $\{u_k\}$, still denoted by itself, such that u_k converges weakly in $W^{1,p}(\Omega)$ to u, that is, we have that*

$$\lim_{k\to\infty} \int_\Omega u_k \phi \, dx = \int_\Omega u\phi \, dx, \quad \forall \, \phi \in L^{\frac{p}{p-1}}(\Omega),$$

and that

$$\lim_{k\to\infty} \int_\Omega \partial_{x_i} u_k \phi \, dx = \int_\Omega \partial_{x_i} u \phi \, dx, \quad \forall \, \phi \in L^{\frac{p}{p-1}}(\Omega), i = 1, 2, \ldots, n.$$

2.1.6. *Difference quotients.* Let $v : \Omega \to \mathbb{R}$ be a locally integrable function, and $\Omega' \Subset \Omega$.

Definition 2.4. The i^{th} difference quotient of size h is

$$\Delta_i^h v(x) = \frac{v(x + he_i) - v(x)}{h}, \quad i = 1, 2, \ldots, n$$

for $x \in \Omega'$ and $h \in \mathbb{R}$, $0 < |h| < \mathrm{dist}(\Omega', \partial\Omega)$. We write $\Delta^h v = (\Delta_1^h v, \Delta_2^h v, \ldots, \Delta_n^h v)$.

THEOREM 2.6 (Difference quotients and weak derivatives)**.** *(i) Assume that $1 \le p < \infty$ and $v \in W^{1,p}(\Omega)$. Then for every $\Omega' \Subset \Omega$,*

$$\int_{\Omega'} |\Delta^h v|^p \, dx \le \int_\Omega |\nabla v|^p$$

for all $0 < |h| < dist(\Omega', \partial\Omega)/2$.

(ii) Assume that $1 < p < \infty$ and $u \in L^p(\Omega')$. Suppose that there is a constant c such that

$$\int_{\Omega'} |\Delta^h v|^p \, dx \le c$$

for all $0 < |h| < dist(\Omega', \partial\Omega)/2$. Then

$$v \in W^{1,p}(\Omega')$$

and

$$\int_{\Omega'} |\nabla v|^p \, dx \le c.$$

2.2. **Definition of weak solutions.**

Definition 2.5. We say that a function $u \in W^{1,2}(\Omega)$ is a *weak solution* of equation (1), if

$$\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\varphi\rangle \, dx = \int_\Omega \sum_{i,j=1}^n a_{ij}(x)\partial_{x_j} u \partial_{x_i}\varphi \, dx = 0 \tag{11}$$

for every $\varphi \in C_0^\infty(\Omega)$. We say that a function $u \in W^{1,2}(\Omega)$ is a *weak subsolution* of equation (1), if

$$\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\varphi\rangle \, dx \le 0$$

for every non-negative $\varphi \in C_0^\infty(\Omega)$. Similarly, we define the *weak supersolutions.*

Remark 2.1. i) If u is a classical solution of equation 1, then by integration by parts, it is a weak solution.
ii) In the definition, we require that the weak solutions are from the Sobolev space $W^{1,2}(\Omega)$. This is the natural Sobolev space for the definition of weak solutions of equation (1). Actually, the formula (11) makes sense if we only require that $u \in W^{1,p}(\Omega)$ for some $p \geq 1$. We call this kind of solutions *very weak solutions*. See [8] for the study of very weak solutions.
iii) Since in the definition we require that $u \in W^{1,2}(\Omega)$, it is easy to prove by an approximation argument that (11) holds for all $\varphi \in W_0^{1,2}(\Omega)$ (Exercise).

3. Moser's iteration

In this lecture, we will prove Theorem 1.4. In the following, $u \in W^{1,2}(\Omega)$ is a weak solution of equation

$$\operatorname{div}(\mathbb{A}(x)\nabla u(x)) = 0, \tag{12}$$

where $\mathbb{A}(x)$ is a symmetric matrix satisfying for $0 < \lambda \leq \Lambda < \infty$

$$\lambda|\xi|^2 \leqslant \langle \mathbb{A}(x)\xi, \xi\rangle \leqslant \Lambda|\xi|^2, \quad \forall\, \xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n, a.e.\, x \in \Omega. \tag{13}$$

3.1. **Harnack's inequality.** We will prove the following Harnack's inequality.

THEOREM 3.1. *Let $u \in W^{1,2}(\Omega), u \geq 0$ in Ω, be a weak solution of equation (12). Then there is a constant $c = c(n, \lambda, \Lambda) > 0$ such that for every ball $B(y, r) \subset \Omega$, we have*

$$\sup_{B(y,r/2)} u \leq c \inf_{B(y,r/2)} u.$$

The proof of Harnack's inequality is divided into two parts in subsection 3.2 and subsection 3.3. The Hölder continuity of solutions is an easy consequence of Harnack's inequality. We leave the proof as an exercise. We use the notation $\operatorname{osc}_{B(y,t)} u = \sup_{B(y,t)} u - \inf_{B(y,t)} u$.

THEOREM 3.2. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Then there is $\alpha = \alpha(n, \lambda, \Lambda), 0 < \alpha \leq 1$, such that $u \in C^{0,\alpha}(\Omega)$. Moreover, for every ball $B(y, R) \subset \Omega$ and every $0 < r \leq R < \infty$, we have*

$$\operatorname{osc}_{B(y,r)} u \leq 2^\alpha \Big(\frac{r}{R}\Big)^\alpha \operatorname{osc}_{B(y,R)} u.$$

3.2. **Weak Harnack's inequality: sup.** We will prove the local boundedness of the weak solutions.

THEOREM 3.3. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Then $u \in L^\infty_{\mathrm{loc}}(\Omega)$. Moreover, for every ball $B(y, r) \subset \Omega$, we have*

$$\sup_{B(y,r/2)} |u| \leq c \left(\fint_{B(y,r)} |u|^2 \, dx \right)^{\frac{1}{2}}, \tag{14}$$

where $c = c(n, \lambda, \Lambda) > 0$.

The essential ingredients of the proof are the Caccioppoli type inequality and the Sobolev inequality. An iteration argument is involved. The starting point is the following Caccioppoli inequality.

Lemma 3.1. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Then for any $\alpha \geq 0$, any $\eta \in C_0^\infty(\Omega)$, we have*

$$\int_\Omega |u|^\alpha |\nabla u|^2 \eta^2 \, dx \leq c \int_\Omega |u|^{\alpha+2} |\nabla \eta|^2 \, dx, \tag{15}$$

where $c = c(\lambda, \Lambda) > 0$, provided that $u \in L^{\alpha+2}_{\mathrm{loc}}(\Omega)$.

Proof. Fix $\eta \in C_0^\infty(\Omega)$. Let $t \geq 0$ and define $v = (u-t)^+ = \max(u-t, 0)$. Using $\varphi = v\eta^2 \in W_0^{1,2}(\Omega)$ (prove it) as a test-function in equation (12), we obtain that

$$\begin{aligned} 0 &= \int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla \varphi \rangle \, dx \\ &= \int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla (u-t)^+ \rangle \eta^2 \, dx + 2\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla \eta \rangle (u-t)^+ \eta \, dx. \end{aligned} \tag{16}$$

We use Cauchy-Schwarz inequality

$$|\langle \mathbb{A}(x)\nabla u, \nabla \eta \rangle| \leq \langle \mathbb{A}(x)\nabla u, \nabla u \rangle^{\frac{1}{2}} \langle \mathbb{A}(x)\nabla \eta, \nabla \eta \rangle^{\frac{1}{2}}$$

and Hölder's inequality to estimate the last integral. Then (16) gives us

$$\int_{u>t} \langle \mathbb{A}(x)\nabla u, \nabla u \rangle \eta^2 \, dx \leq 4 \int_{u>t} \langle \mathbb{A}(x)\nabla \eta, \nabla \eta \rangle |(u-t)^+|^2 \, dx,$$

from which, together with (13), yields

$$\int_{u>t} |\nabla u|^2 \eta^2 \, dx \leq \frac{4\Lambda}{\lambda} \int_{u>t} |(u-t)^+|^2 |\nabla \eta|^2 \, dx \leq \frac{4\Lambda}{\lambda} \int_{u>t} |u^+|^2 |\nabla \eta|^2 \, dx. \tag{17}$$

Now the above inequality holds for all $t \geq 0$. We multiply both sides by $\alpha t^{\alpha-1}$ and integrate with respect to t over $(0, \infty)$. A direct calculation gives

$$\int_\Omega |u|^\alpha |\nabla u^+|^2 \eta^2 \, dx \leq \frac{4\Lambda}{\lambda} \int_\Omega |u^+|^{\alpha+2} |\nabla \eta|^2 \, dx.$$

Similarly, the above inequality is also true for u^-. Then we obtain (15) with $c = 4\Lambda/\lambda$. This proves the lemma. □

Now by the Sobolev inequality, we obtain the following reverse inequality.

Lemma 3.2. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Then $u \in L^{(\alpha+2)\chi}_{\mathrm{loc}}(\Omega)$, if $u \in L^{\alpha+2}_{\mathrm{loc}}(\Omega)$ for any $\alpha \geq 0$, where $\chi = n/(n-2)$ when $n \geq 3$. Moreover, we have for any $\eta \in C_0^\infty(\Omega)$,*

$$\left(\int_\Omega |u|^{(\alpha+2)\chi} \eta^{2\chi} \, dx \right)^{\frac{1}{\chi}} \leq c(\alpha+2)^2 \int_\Omega |u|^{\alpha+2} |\nabla \eta|^2 \, dx, \tag{18}$$

where $c = c(n, \lambda, \Lambda) > 0$.

Proof. Let $v = |u|^{\alpha/2} u \eta$. Then

$$\nabla v = \left(\frac{\alpha}{2} + 1\right) |u|^{\frac{\alpha}{2}} \eta \nabla u + |u|^{\frac{\alpha}{2}} u \nabla \eta.$$

Thus (15) in Lemma 3.1 gives us

$$\int_\Omega |\nabla v|^2 \, dx \leq c(\alpha+2)^2 \int_\Omega |u|^{\alpha+2} |\nabla \eta|^2 \, dx, \tag{19}$$

where $c = c(\lambda, \Lambda) > 0$. Now we use the Sobolev inequality in Theorem 2.1 (when $n \geq 3$)

$$\left(\int_\Omega |v|^{2\chi}\, dx\right)^{\frac{1}{\chi}} \leq c(n) \int_\Omega |\nabla v|^2\, dx. \tag{20}$$

Combining (19) and (20) yields (18). We finish the proof. □

The following corollary is an easy consequence of Lemma 3.2.

Corollary 3.1. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Then $u \in L^q_{\mathrm{loc}}(\Omega)$ for every $q \geq 1$. Moreover, for every $\alpha \geq 0$, for every ball $B(y, r) \subset \Omega$ and for every $0 < r' < r$, we have the following reverse type inequality*

$$\left(\int_{B(y,r')} |u|^{(\alpha+2)\chi}\, dx\right)^{\frac{1}{(\alpha+2)\chi}} \leq \frac{c^{\frac{1}{\alpha+2}}(\alpha+2)^{\frac{2}{\alpha+2}}}{(r-r')^{\frac{2}{\alpha+2}}} \left(\int_{B(y,r)} |u|^{\alpha+2}\, dx\right)^{\frac{1}{\alpha+2}}, \tag{21}$$

where $c = c(n, \lambda, \Lambda) > 0$.

Now we iterate (21) to prove Theorem 3.3.

Proof of Theorem 3.3. Fix a ball $B(y, r) \subset \Omega$. Define $\alpha_0 = 0, \alpha_i = 2\chi^i - 2, i = 1, 2, \ldots$. Let $r_0 = r$ and

$$r_i = \frac{r}{2} + \frac{r}{2^{i+1}}, \quad i = 1, 2, \ldots.$$

Apply (21) with $r = r_i, r' = r_{i+1}$ and $\alpha = \alpha_i$ for $i = 0, 1, \ldots$. We obtain that

$$M_{i+1} \leq c^{\frac{1}{\beta_i}} \beta_i^{\frac{2}{\beta_i}} \left(\frac{r}{2^{i+2}}\right)^{-\frac{2}{\beta_i}} M_i, \tag{22}$$

where $\beta_i = 2\chi^i$ and

$$M_i = \left(\int_{B(y,r_i)} |u|^{\beta_i}\, dx\right)^{\frac{1}{\beta_i}}.$$

An iteration of (22) gives us

$$M_{i+1} \leq c_i M_0,$$

from which we obtain (14) by letting $i \to \infty$. We leave the details as an exercise. We finish the proof of Theorem 3.3. □

Slightly modifying the above argument, we can prove the following version of Theorem 3.3 for the weak subsolutions (see Definition 2.5) when $n \geq 3$. Write down the details of the proof as an exercise.

THEOREM 3.4. *Let $u \in W^{1,2}(\Omega)$, $u \geq 0$ in Ω, be a weak subsolution of equation (12). Then $u \in L^\infty_{\mathrm{loc}}(\Omega)$. Moreover, for every ball $B(y, r) \subset \Omega$ and every $0 < \sigma < 1$, we have*

$$\sup_{B(y,\sigma r)} u \leq \frac{c}{(1-\sigma)^{n/2}} \left(\fint_{B(y,r)} u^2\, dx\right)^{\frac{1}{2}}, \tag{23}$$

where $c = c(n, \lambda, \Lambda) > 0$.

Finally, by another iteration argument, we will prove the following boundedness estimate for the subsolutions.

THEOREM 3.5. *Let $u \in W^{1,2}(\Omega)$, $u \geq 0$ in Ω, be a weak subsolution of equation (12). Then $u \in L^{\infty}_{\mathrm{loc}}(\Omega)$. Moreover, for every ball $B(y,r) \subset \Omega$ and every $0 < \sigma < 1, 0 < q \leq 2$, we have*

$$\sup_{B(y,\sigma r)} u \leq \frac{c}{(1-\sigma)^{n/q}} \left(\fint_{B(y,r)} u^q \, dx \right)^{\frac{1}{q}}, \tag{24}$$

where $c = c(q, n, \lambda, \Lambda) > 0$.

Proof. Fix $B(y,r) \subset \Omega$ and $0 < \sigma < 1$. Let $\sigma_0 = \sigma$ and

$$\sigma_i = 1 - \frac{1-\sigma}{2^i}, \quad i = 1, 2, \ldots.$$

By Theorem 3.4, (23) with $r = \sigma_{i+1} r$ and $\sigma = \sigma_i / \sigma_{i+1}$ gives us

$$\begin{aligned} \sup_{B(y,\sigma_i r)} u \leq & \frac{c}{\left(1 - \frac{\sigma_i}{\sigma_{i+1}}\right)^{n/2}} \left(\fint_{B(y,\sigma_{i+1} r)} u^2 \, dx \right)^{\frac{1}{2}} \\ \leq & \frac{c}{\left(1 - \frac{\sigma_i}{\sigma_{i+1}}\right)^{n/2}} \left(\fint_{B(y,\sigma_{i+1} r)} u^q \, dx \right)^{\frac{1}{2}} \left(\sup_{B(y,\sigma_{i+1})} u \right)^{\frac{2-q}{2}}, \end{aligned}$$

that is,

$$M_i \leq \frac{c}{\left(1 - \frac{\sigma_i}{\sigma_{i+1}}\right)^{n/2}} \left(\fint_{B(y,r)} u^q \, dx \right)^{\frac{1}{2}} M_{i+1}^{\frac{2-q}{2}}, \tag{25}$$

where $M_i = \sup_{B(y,\sigma_i r)} u$. An iteration of (25) gives (14). We leave the details as an exercise. We finish the proof. □

Exercise: figure out a version of Theorem 3.5 in the case $n = 2$, and write down the proof.

3.3. **Weak Harnack's inequality: inf.** In this subsection, we will prove

THEOREM 3.6. *Let $u \in W^{1,2}(\Omega)$, $u \geq 0$ in Ω, be a weak solution of equation (12). Then there are $q = q(n, \lambda, \Lambda) > 0$ and $c = c(n, \lambda, \Lambda) > 0$ such that for every ball $B(y, 2r) \subset \Omega$, we have*

$$\inf_{B(y,r/2)} u \geq c \left(\fint_{B(y,r)} u^q \, dx \right)^{\frac{1}{q}}. \tag{26}$$

By replacing u by $u + \varepsilon$ for $\varepsilon > 0$, we may assume that $u \geq \varepsilon$ in Ω. The essential point of the proof of Theorem 3.6 is that $\log u$ is a function of bounded mean oscillation (BMO). First, Theorem 3.5 yields the following estimate.

Lemma 3.3. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Suppose that $u \geq \varepsilon$ in Ω for some $\varepsilon > 0$. Then for any $q > 0$ there is $c = c(q, n, \lambda, \Lambda) > 0$ such that for every ball $B(y,r) \subset \Omega$, we have*

$$\inf_{B(y,r/2)} u \geq c \left(\fint_{B(y,r)} u^{-q} \, dx \right)^{-\frac{1}{q}}. \tag{27}$$

Proof. We claim that $v = 1/u$ is a subsolution of equation (12). Indeed, first, it is easy to show that $1/u \in W^{1,2}_{\mathrm{loc}}(\Omega)$. Second, for any $\eta \in C_0^\infty(\Omega), \eta \geq 0$, let $\varphi = \eta/u^2$. We use φ as a test-function in equation (12) to obtain that

$$\begin{aligned} 0 &= \int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\varphi\rangle\, dx \\ &= \int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\eta\rangle \frac{1}{u^2}\, dx - 2\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla u\rangle \frac{\eta}{u^3}\, dx. \end{aligned}$$

The last integral is non-negative. Thus we have

$$\int_\Omega \langle \mathbb{A}(x)\nabla v, \nabla\eta\rangle\, dx = -\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\eta\rangle \frac{1}{u^2}\, dx \leq 0,$$

which shows that $1/u$ is a subsolution. Now the lemma follows from Theorem 3.5. $\square$

Second, we show that $\log u$ is of BMO.

Lemma 3.4. *Let $u \in W^{1,2}(\Omega)$ be a weak solution of equation (12). Suppose that $u \geq \varepsilon$ in Ω for some $\varepsilon > 0$. Then for every ball $B(y, 2r) \subset \Omega$, we have*

$$\int_{B(y,r)} |\nabla v|^2\, dx \leq cr^{n-2}, \tag{28}$$

where $v = \log u$ and $c = c(n, \lambda, \Lambda) > 0$.

Proof. Fix $\eta \in C_0^\infty(\omega)$ and let $\varphi = \eta^2/u$. We use φ as a test-function in equation (12) to obtain that

$$\begin{aligned} 0 &= \int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\varphi\rangle\, dx \\ &= -\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla u\rangle \frac{\eta^2}{u^2}\, dx + 2\int_\Omega \langle \mathbb{A}(x)\nabla u, \nabla\eta\rangle \frac{\eta}{u}\, dx, \end{aligned}$$

from which we deduce that

$$\int_\Omega |\nabla v|^2\eta^2 \leq \frac{4\Lambda}{\lambda}\int_\Omega |\nabla\eta|^2\, dx. \tag{29}$$

Then (28) follows by choosing $\eta \in C_0^\infty(B(y, 2r))$ such that $\eta = 1$ on $B(y, r)$ and $|\nabla\eta| \leq 2/r$ in $B(y, 2r)$. $\square$

Definition 3.1. A function $v \in L^1_{\mathrm{loc}}(\Omega)$ is said to be of bounded mean oscillation, denoted by $v \in BMO(\Omega)$, if

$$[v]_{BMO} = \sup \fint_{B(y,r)} |v - v_{B(y,r)}|\, dx < \infty, \tag{30}$$

where $v_{B(y,r)} = \fint_{B(y,r)} v\, dx$ is the average of v over the ball $B(y, r)$ and the supremum in (30) is taken for all balls $B(y, r)$ such that $B(y, 2r) \subset \Omega$.

A fundamental property of BMO functions is the exponential integrability.

Lemma 3.5 (John-Nirenberg Lemma)**.** *Suppose that $v \in BMO(\Omega)$. Then there are positive constants c_1 and c_2, depending only on n and $[v]_{BMO}$, such that for every $B(y, 2r) \subset \Omega$, we have*

$$\fint_{B(y,r)} \exp(c_1|v - v_{B(y,r)}|)\, dx \le c_2. \tag{31}$$

Finally, we give the proof of Theorem 3.6.

Proof of Theorem 3.6. By Lemma 3.4 and the Poincaré inequality, we have for every ball $B(y, 2r) \subset \Omega$

$$\fint_{B(y,r)} |v - v_{B(y,r)}|^2\, dx \le c(n)r^2 \int_{B(y,r)} |\nabla v|^2\, dx \le c(n, \lambda, \Lambda).$$

Thus $u \in BMO(\Omega)$. Then the John-Nirenberg lemma yields

$$\fint_{B(y,r)} \exp(c_1|v - v_{B(y,r)}|)\, dx \le c_2$$

for $c_1 = c_1(n, \lambda, \Lambda) > 0$ and $c_2 = c_2(n, \lambda, \Lambda) > 0$. Then we have

$$\begin{aligned}
&\fint_{B(y,r)} u^{c_1}\, dx \times \int_{B(y,r)} u^{-c_1}\, dx \\
&= \fint_{B(y,r)} \exp(c_1(v - v_{B(y,r)}))\, dx \times \int_{B(y,r)} \exp(c_1(v_{B(y,r)} - v))\, dx \\
&\le \left(\fint_{B(y,r)} \exp(c_1|v - v_{B(y,r)}|)\, dx \right)^2 \le c_2^2,
\end{aligned}$$

from which, together with Lemma 3.3, proves (26) with $q = c_1$. This finishes the proof of Theorem 3.6. □

4. De Giorgi's method

4.1. **De Giorgi's class of functions.** In his fundamental work on linear elliptic equations, De Giorgi [1] established the local boundedness and the Hölder continuity for functions satisfying certain integral inequalities, known as the De Giorgi class of functions.

Let Ω be an open set in $\mathbb{R}^n$ and γ be a constant. The De Giorgi class $DG^+(\Omega, \gamma)$ consists of functions $u \in W^{1,2}(\Omega)$, which satisfy for every ball $B(y, r) \subset \Omega$, for every $0 < r' < r$ and for every $k \in \mathbb{R}$, the following Caccioppoli type inequality

$$\int_{B(y,r')} |\nabla(u - k)^+|^2\, dx \le \frac{\gamma}{(r - r')^2} \int_{B(y,r)} |(u - k)^+|^2\, dx, \tag{32}$$

where $(u - k)^+ = \max(u - k, 0)$. Similarly, we may define the class $DG^-(\Omega, \gamma)$ by replacing $(u - k)^+$ by $(u - k)^- = \min(u - k, 0)$. Thus $u \in DG^+(\Omega, \gamma)$ if and only if $-u \in DG^-(\Omega, \gamma)$. We denote $DG(\Omega, \gamma) = DG^+(\Omega, \gamma) \cap DG^-(\Omega, \gamma)$.

All weak solutions of equation (12) are in the De Giorgi class. We already proved the following lemma, see the proof of Lemma 3.1.

Lemma 4.1. *Let $u \in W^{1,2}(\Omega)$ be a weak subsolution of equation (12). Then $u \in DG^+(\Omega, \gamma)$ for some $\gamma = \gamma(\lambda, \Lambda) > 0$.*

4.2. **Boundedness of functions in** $DG(\Omega, \gamma)$**.**

THEOREM 4.1. *Let $\gamma > 0$ be a constant and $u \in DG^+(\Omega, \gamma)$. Then $u \in L^\infty_{\text{loc}}(\Omega)$. Moreover, for every ball $B(y, r) \subset \Omega$, we have*

$$\sup_{B(y,r/2)} u^+ \le c \left(⨍_{B(y,r)} |u^+|^2 \, dx \right)^{\frac{1}{2}}, \tag{33}$$

where $c = c(n, \gamma) > 0$.

Proof. Fix $B(y, r) \subset \Omega$. Let $M > 0$ be a number to be chosen later. We set

$$k_i = M - \frac{M}{2^i}, \quad i = 0, 1, 2, ...$$

and consider the sequence of radii

$$r_i = \frac{r}{2} + \frac{r}{2^{i+1}}, \quad \bar{r}_i = \frac{1}{2}(r_i + r_{i+1}) = \frac{r}{2} + \frac{3}{4} \cdot \frac{r}{2^{i+1}}, \quad i = 0, 1, 2, ...$$

Let $\eta_i \in C_0^\infty(B(y, \bar{r}_i))$ be a cut-off function such that $\eta_i = 1$ in $B(y, r_{i+1})$ and $|\nabla \eta_i| \le 2^{i+8}/r$. We only prove the case $n \ge 3$. By Hölder's inequality, we have

$$\begin{aligned} \int_{B(y,r_{i+1})} |(u - k_{i+1})^+|^2 \, dx &\le \int_{B(y,\bar{r}_i)} |(u - k_{i+1})^+ \eta_i|^2 \, dx \\ &\le \left(\int_{B(y,\bar{r}_i)} |(u - k_{i+1})^+ \eta_i|^{\frac{2n}{n-2}} \, dx \right)^{\frac{n-2}{n}} |A_i|^{\frac{2}{n}}, \end{aligned} \tag{34}$$

where $A_i = \{x \in B(y, \bar{r}_i) : u(x) > k_{i+1}\}$. We continue to estimate the first integral on the right hand side of (34) by Sobolev inequality

$$\begin{aligned} I = &\left(\int_{B(y,\bar{r}_i)} |(u - k_{i+1})^+ \eta_i|^{\frac{2n}{n-2}} \, dx \right)^{\frac{n-2}{n}} \\ \le &c(n) \int_{B(y,\bar{r}_i)} |\nabla((u - k_{i+1})^+ \eta_i)|^2 \, dx \\ \le &c \int_{B(y,\bar{r}_i)} |\nabla(u - k_i)^+|^2 \, dx + c \int_{B(y,\bar{r}_i)} |(u - k_{i+1})^+|^2 |\nabla \eta_i|^2 \, dx. \end{aligned}$$

Since $u \in DG^+(\Omega, \gamma)$, we have

$$\int_{B(y,\bar{r}_i)} |\nabla(u - k_i)^+|^2 \, dx \le \frac{\gamma}{(r_i - \bar{r}_i)^2} \int_{B(y,r_i)} |(u - k_i)^+|^2 \, dx.$$

Thus, we get

$$I \le \frac{c 2^{2i}}{r^2} \int_{B(y,r_i)} |(u - k_i)^+|^2 \, dx, \tag{35}$$

where $c = c(n, \gamma) > 0$. Now let

$$Y_i = \frac{1}{M^2 r^n} \int_{B(y,r_i)} |(u - k_i)^+|^2 \, dx$$

and observe that

$$Y_i \ge \frac{(k_{i+1} - k_i)^2}{M^2 r^n} |\{x \in B(y, r_i) : u(x) > k_{i+1}\}| \ge \frac{1}{2^{2i+2} r^n} |A_i|. \tag{36}$$

Thus combining (34), (35) and (36) yields

$$Y_{i+1} \le cb^i Y_i^{1+\delta}, \quad b = 2^{2+\frac{4}{n}}, \delta = 1 + \frac{2}{n}.$$

It is easy to prove (Exercise) that there is $\varepsilon_0 = \varepsilon_0(c, b, \delta) > 0$ such that if $Y_0 \le \varepsilon_0$, then $Y_i \to 0$ as $i \to \infty$. This means that

$$\sup_{B(y,r/2)} u \le M$$

if we choose M such that

$$Y_0 = \frac{1}{M^2 r^n} \int_{B(y,r)} |u^+|^2 \, dx = \varepsilon_0.$$

This proves the theorem. □

4.3. **Hölder continuity of functions in** $DG(\Omega, \gamma)$**.** In this subsection, we will prove the Hölder continuity for functions in the De Giorgi class.

THEOREM 4.2. *Let $\gamma > 0$ be a constant. There is an exponent $\alpha = \alpha(n, \gamma), 0 < \alpha \le 1$, such that for every $u \in DG(\Omega, \gamma)$, we have $u \in C^{0,\alpha}(\Omega)$. Moreover, there is $\delta = \delta(n, \gamma), 0 < \delta < 1$, such that for every ball $B(y, r) \subset \Omega$, we have*

$$\mathrm{osc}_{B(y,r/4)} u \le \delta \, \mathrm{osc}_{B(y,r)} u.$$

Theorem 4.2 follows from the following two lemmas.

Lemma 4.2. *For any $\theta > 0$, there is $s = s(\theta, \gamma, n) \ge 1$ such that the following holds: for every $u \in DG^+(\Omega, \gamma)$ and for every ball $B(y, r) \subset \Omega$ if*

$$|\{x \in B(y, r/2) : u(x) \le k_0\}| \ge \frac{1}{2} |B(y, r/2)| \tag{37}$$

holds for some $k_0 \in \mathbb{R}$, then we have

$$|\{x \in B(y, r/2) : u(x) > M_r - \frac{1}{2^s}(M_r - k_0)\}| \le \theta |B(y, r/2)|, \tag{38}$$

where $M_r = \sup_{B(y,r)} u$.

Proof. Fix $B(y, r) \subset \Omega$ and $k_0 \in \mathbb{R}$ such that (37) holds. We may assume that $k_0 < M_r$. Otherwise, there is nothing to prove. Since $u \in DG^+(\Omega, \gamma)$, then for any k,

$$\begin{aligned}\int_{B(y,r/2)} |\nabla (u-k)^+|^2 \, dx \le & \frac{4\gamma}{r^2} \int_{B(y,r)} |(u-k)^+|^2 \, dx \\ \le & c(n, \gamma) r^{n-2} (M_r - k)^2.\end{aligned}$$

For any $k_0 \le k < l < M_r$, we have by Hölder's inequality

$$\begin{aligned}\left(\int_{A_{k,l}} |\nabla u| \, dx \right)^2 & \le \int_{A_{k,l}} |\nabla u|^2 \, dx |A_{k,l}| \\ & \le \int_{B(y,r/2)} |\nabla (u-k)^+|^2 \, dx |A_{k,l}|,\end{aligned}$$

where $A_{k,l} = \{x \in B(y, r/2) : k < u(x) \le l\}$. Thus,

$$\left(\int_{A_{k,l}} |\nabla u| \, dx \right)^2 \le c(M_r - k)^2 r^{n-2} |A_{k,l}|. \tag{39}$$

Now note that

$$|\{x \in B(y, r/2) : u(x) \leq k\}| \geq \frac{1}{2}|B(y, r/2)|$$

for all $k \geq k_0$, due to (37). We apply Theorem 2.3 to obtain that

$$(l-k)|A_l|^{1-\frac{1}{n}} \leq c(n) \int_{A_{k,l}} |\nabla u| \, dx, \tag{40}$$

where

$$A_l = \{x \in B(y, r/2) : u(x) > l\}.$$

It follows from (39) and (40) that

$$(l-k)^2 |A_l|^{2-\frac{2}{n}} \leq c(M_r - k)^2 r^{n-2} |A_{k,l}| \tag{41}$$

for all $k_0 \leq k < l < M_r$. Now set

$$k_i = M_r - \frac{1}{2^i}(M_r - k_0), \quad i = 0, 1, 2, \ldots.$$

Then (41) with $k = k_i, l = k_{i+1}$ gives

$$|A_{k_{i+1}}|^{2-\frac{2}{n}} \leq c r^{n-2} |A_{k_i, k_{i+1}}|.$$

Let $i_0 \in \mathbb{N}$, to be chosen soon. Sum the above inequality from $i = 0$ to $i_0 - 1$. Note that $|A_{k_{i+1}}| \geq |A_{k_{i_0}}|$ for all $0 \leq i \leq i_0 - 1$. Thus we arrive at

$$i_0 |A_{i_0}|^{2-\frac{2}{n}} \leq c r^{n-2} |A_{k_0}| \leq c r^{2n-2}.$$

By choosing i_0 big enough, we proved the lemma. □

Lemma 4.3. *There is $\theta = \theta(n, \gamma) > 0$ such that the following holds: for every $u \in DG^+(\Omega, \gamma)$, for every ball $B(y, r) \subset \Omega$ and for any $k_0 \in \mathbb{R}$, if*

$$|\{x \in B(y, r) : u(x) > k_0\}| \leq \theta |B(y, r)|, \tag{42}$$

then we have

$$\sup_{B(y, r/2)} u \leq \frac{1}{2}k_0 + \frac{1}{2} \sup_{B(y,r)} u. \tag{43}$$

Proof. Fix $B(y, r) \subset \Omega$ and fix k_0 such that (42) holds for some θ, which will be chosen later. Denote $M_r = \sup_{B(y,r)} u$. We may assume that $k_0 < M_r$. Otherwise, the conclusion in the lemma is trivial. Now we consider a sequence

$$k_i = \frac{k_0}{2} + \frac{M_r}{2} - \frac{1}{2^{i+1}}(M_r - k_0), \quad i = 0, 1, 2, \ldots$$

and a sequence of radii

$$r_i = \frac{r}{2} + \frac{r}{2^{i+1}}, \quad i = 0, 1, 2, \ldots.$$

Since $u \in DG^+(\Omega, \gamma)$, (32) with $k = k_i, r = r_i, r' = r_{i+1}$ gives

$$\begin{aligned} \int_{A_{k_i, r_{i+1}}} |\nabla u|^2 \, dx \leq & \frac{\gamma}{(r_i - r_{i+1})^2} \int_{A_{k_i, r_i}} |(u - k_i)^+|^2 \, dx \\ \leq & \frac{c 2^{2i}}{r^2} (M_r - k_0)^2 |A_{k_i, r_i}|, \end{aligned} \tag{44}$$

where we denote

$$A_{k,\rho} = \{x \in B(y, \rho) : u(x) > k\}.$$

On the other hand, due to (42), we have

$$|A_{k_i,r_{i+1}}| \le |A_{k_0,r}| \le \theta |B(y,r)| \le \frac{1}{2}|B(y,r_{i+1})|,$$

if we assume that $\theta \le 1/2^{n+1}$. Then by Theorem 2.3 with $k = k_i, l = k_{i+1}$, we have

$$\begin{aligned}(k_{i+1}-k_i)|A_{k_{i+1},r_{i+1}}|^{1-\frac{1}{n}} \le & c(n)\int_{A_{k_i,r_{i+1}}} |\nabla u|\,dx \\ \le & c(n)\left(\int_{A_{k_i,r_{i+1}}} |\nabla u|^2\,dx\right)^{\frac{1}{2}} |A_{k_i,r_{i+1}}|^{\frac{1}{2}}.\end{aligned} \tag{45}$$

Combining (44) and (45) yields

$$(k_{i+1}-k_i)|A_{k_{i+1},r_{i+1}}|^{1-\frac{1}{n}} \le \frac{c2^i}{r}(M_r - k_0)|A_{k_i,r_i}|^{\frac{1}{2}}|A_{k_i,r_{i+1}}|^{\frac{1}{2}} \le \frac{c2^i}{r}|A_{k_i,r_i}|,$$

that is,

$$Y_{i+1} \le c b^i Y_i^{\frac{n}{n-1}}, \quad b = 2^{\frac{2n}{n-1}}, c = c(n,\gamma) > 0, \tag{46}$$

where we denote

$$Y_i = \frac{|A_{k_i,r_i}|}{|B(y,r)|}.$$

Now we can iterate (46) and prove (Exercise) that there is $\theta = \theta(b,c) > 0$ such that $Y_i \to 0$ as $i \to \infty$, provided that $Y_0 \le \theta$. This proves the lemma. □

Now we are in the position to prove Theorem 4.2.

Proof of Theorem 4.2. Let $u \in DG(\Omega,\gamma)$ and fix a ball $B(y,r) \subset \Omega$. Set

$$M_s = \sup_{B(y,s)} u, \quad m_s = \inf_{B(y,s)} u.$$

Then we let $k_0 = (M_r + m_r)/2$. We may assume that

$$|\{x \in B(y,r/2) : u(x) \le k_0\}| \ge \frac{1}{2}|B(y,r/2)|.$$

Otherwise, we consider $-u$, instead of u. Let $\theta = \theta(n,\gamma) > 0$ be the number determined as in Lemma 4.3. Now we apply Lemma 4.2 to obtain that

$$|\{x \in B(y,r/2) : u(x) > M_r - \frac{1}{2^s}(M_r - k_0)\}| \le \theta |B(y,r/2)|$$

where $s = s(n,\gamma) \ge 1$. Then we apply Lemma 4.3 to obtain that

$$M_{r/4} \le \frac{1}{2}[M_r - \frac{1}{2^s}(M_r-k_0)] + \frac{1}{2}M_{r/2} \le M_r - \frac{1}{2^{s+1}}(M_r - k_0).$$

Therefore, we have

$$M_{r/4} - m_{r/4} \le M_r - \frac{1}{2^{s+1}}(M_r-k_0) - m_r = (1 - \frac{1}{2^{s+2}})(M_r - m_r),$$

which proves the theorem. □

5. Further discussions

5.1. **Degenerate elliptic equations.** In equation (1), we assume the following for the coefficients $\mathbb{A}(x)$

$$\lambda(x)|\xi|^2 \leqslant \langle \mathbb{A}(x)\xi, \xi\rangle = \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leqslant \Lambda(x)|\xi|^2 \tag{47}$$

for all $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ and for almost every $x \in \Omega$. Here $0 \leq \lambda(x) \leq \Lambda(x) \leq \infty$ are given functions.

In 1995, De Giorgi gave a talk in Lecce, Italy, and discussed the natural question: are size assumptions on $\lambda(x)^{-1}$ and $\Lambda(x)$ sufficient to guarantee the continuity of weak solutions? He raised the following conjectures on the continuity of weak solutions of equation (1). The first one concerns the singular case in higher dimensions.

Conjecture 1. [De Giorgi [2]] Let $n \geq 3$. Suppose that $\mathbb{A}(x)$ satisfies (47) with $\lambda(x) = 1$ and with $\Lambda(x)$ satisfying

$$\int_\Omega \exp(\Lambda(x))\, dx < \infty. \tag{48}$$

Then all weak solutions of equation (1) are continuous in Ω.

The second one concerns the degenerate case in higher dimensions.

Conjecture 2. [De Giorgi [2]] Let $n \geq 3$. Suppose that $\mathbb{A}(x)$ satisfies (47) with $\Lambda(x) = 1$ and with $\lambda(x)$ satisfying

$$\int_\Omega \exp(\lambda(x)^{-1})\, dx < \infty. \tag{49}$$

Then all weak solutions of equation (1) are continuous in Ω.

The third one concerns the case of singular and degenerate equations in higher dimensions.

Conjecture 3. [De Giorgi [2]] Let $n \geq 3$. Suppose that $\mathbb{A}(x)$ satisfies (47) with $\Lambda(x) = \lambda(x)^{-1}$ satisfying

$$\int_\Omega \exp(\Lambda(x)^2)\, dx < \infty. \tag{50}$$

Then all weak solutions of equation (1) are continuous in Ω.

The fourth one concerns the planar case, $n = 2$, which is different from the higher dimensional cases.

Conjecture 4. [De Giorgi [2]] Let $n = 2$. Suppose that $\mathbb{A}(x)$ satisfies (47) with $\Lambda(x) = 1$ and with $\lambda(x)$ satisfying

$$\int_\Omega \exp(\sqrt{\lambda(x)^{-1}})\, dx < \infty. \tag{51}$$

Then all weak solutions of equation (1) are continuous in Ω.

Conjecture 1, Conjecture 2, and Conjecture 3 are still open. As far as we know, the best known result is due to Trudinger [20], which is far from the conjectures. It seems one needs new ideas to deal with these challenging problems. Concerning

Conjecture 4, Onninen and the author [16] recently proved that all weak solutions of equation (1) are continuous under the assumption that

$$\int_\Omega \exp(\alpha\sqrt{\lambda(x)^{-1}})\, dx < \infty$$

for some constant $\alpha > 1$.

Another interesting part of [2] is that De Giorgi also conjectured that his conjectures above are sharp; the integrability conditions (48), (49) and (51) are optimal to guarantee the continuity of weak solutions. For example, in Conjecture 1, one can not replace (48) by the following weaker one

$$\int_\Omega \exp(\alpha\Lambda(x)^{1-\delta})\, dx < \infty \tag{52}$$

for some $\delta > 0$ and any $\alpha > 0$. De Giorgi conjectured that one can construct a function $\Lambda(x)$ satisfying the integrability condition (52) such that equation (1) satisfying (47) with $\lambda(x) = 1$ and with this $\Lambda(x)$ has discontinuous weak solutions.

In [2], De Giorgi even gave hints how to construct such counter examples to show the sharpness of the above conjectures. He made the following precise conjectures. Let $\Omega = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : |x| < 1/e\}$ and A, B be subsets of Ω

$$A = \{x \in \Omega : 2|x_n| > |x|\}, \quad B = \{x \in \Omega : 2|x_n| < |x|\}.$$

The first conjecture would yield the sharpness of Conjecture 1.

Conjecture 5. [De Giorgi [2]] Let $n \geq 3$. For any $\varepsilon > 0$, define a function τ_1 in Ω as follows

$$\tau_1(x) = \begin{cases} |\log|x||^{1+\varepsilon} & \text{if } x \in A; \\ 1 & \text{if } x \in B. \end{cases}$$

Then equation (1) with $\mathbb{A}(x) = \tau_1(x)I$ has a weak solution, discontinuous at the origin.

The second one would yield the sharpness of Conjecture 2.

Conjecture 6. [De Giorgi [2]] Let $n \geq 3$. For any $\varepsilon > 0$, define a function τ_2 in Ω as follows

$$\tau_2(x) = \begin{cases} 1 & \text{if } x \in A; \\ |\log|x||^{-(1+\varepsilon)} & \text{if } x \in B. \end{cases}$$

Then equation (1) with $\mathbb{A}(x) = \tau_2(x)I$ has a weak solution, discontinuous at the origin.

The third one concerns the planar case, and would yield the sharpness of Conjecture 4.

Conjecture 7. [De Giorgi [2]] Let $n = 2$. For any $\varepsilon > 0$, define a function τ_3 in Ω as follows

$$\tau_3(x) = \begin{cases} 1 & \text{if } x \in A; \\ |\log|x||^{-(2+\varepsilon)} & \text{if } x \in B. \end{cases}$$

Then equation (1) with $\mathbb{A}(x) = \tau_3(x)I$ has a weak solution, discontinuous at the origin.

In [22], it was proved that Conjecture 5, Conjecture 6 and Conjecture 7 are true.

THEOREM 5.1. *Conjecture 5 is true.*

THEOREM 5.2. *Conjecture 6 is true.*

THEOREM 5.3. *Conjecture 7 is true.*

References

[1] 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.

[2] De Giorgi, E. *Congetture sulla continuità delle soluzioni di equazioni lineari ellittiche autoaggiunte a coefficienti illimitati.* Unpublished (1995).

[3] DiBenedetto, E. and Trudinger, N. S. *Harnack inequalities for quasiminima of variational integrals.* Ann. Inst. H. Poincaré Anal. Non Linéaire **1** (1984), no. 4, 295–308.

[4] Evans, L. C. *Partial differential equations.* Graduate Studies in Mathematics, **19**. American Mathematical Society, Providence, RI, 1998.

[5] Fabes, E. B., Kenig, C. E. and Serapioni, R. *The local regularity of solutions of degenerate elliptic equations.* Comm. Partial Diff. Equations **7** (1982), 77–116.

[6] Fabes, E. B. and Stroock, D. W. *A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash.* Arch. Rational Mech. Anal. **96** (1986), no. 4, 327–338.

[7] Gilbarg, D., and Trudinger, N. S. *Elliptic partial differential equations of second order.* 2nd ed., Springer-Verlag, New York (1983).

[8] Iwaniec, T. and Sbordone, C. *Weak minima of variational integrals.* J. Reine Angew. Math. **454** (1994), 143–161.

[9] Ladyzhenskaya, O. A., and Ural'tseva, N. N. *Linear and quasilinear elliptic equations.* Academic Press, New York (1968).

[10] Meyers, N. G. *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations.* Ann. Sc. Norm. Sup. Pisa **17** (1963), 189–206.

[11] Morrey, C. B. *On the solutions of quasi-linear elliptic partial differential equations.* Trans. Amer. Math. Soc. **43** (1938), 126–166.

[12] Morrey, C. B. *Multiple integral problems in the calculus of variations and related topics.* Univ. California Publ. Math.(N. S.) **1** (1943), 1–130.

[13] 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.

[14] Moser, J. *On Harnack's theorem for elliptic differential equations.* Comm. Pure Appl. Math. **14** (1961), 577–591.

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

[16] Onninen, J., and Zhong, X. *Continuity of solutions of linear, degenerate elliptic equations.* Ann. Sc. Norm. Super. Pisa Cl. Sc. (5) **6** (2007), 103–116.

[17] Piccinini, L. C., and Spagnolo, S. *On the Hölder continuity of solutions of second order elliptic equations in two variables.* Ann. Scuola Norm. Sup. Pisa (3) **26** (1972), 391–402.

[18] Serrin, J. *On the strong maximum principle for quasilinear second order differential inequalities.* J. Functional Analysis **5** (1970), 184–193.

[19] Trudinger, N. S. *On Harnack type inequalities and their application to quasilinear, elliptic equations.* Comm. Pure Appl. Math. **20** (1967), 721–747.

[20] Trudinger, N. S. *On the regularity of generalized solutions of linear, non-uniformly elliptic equations.* Arch. Rational Mech. Anal. **42** (1971), 51–62.

[21] Widman, K. O. *On the Hölder continuity of solutions of elliptic partial differential equations in two variables with coefficients in L^∞.* Comm. Pure Appl. Math. **22** (1969), 669–682.

[22] Zhong, X. *Discontinuous solutions of linear, degenerate elliptic equations.* J. Math. Pures Appl. (9) **90** (2008), 31–41.

Department of Mathematics and Statistics, University of Jyväskylä,, P.O. Box 35 (MaD), Fin-40014, University of Jyväskylä, Finland.

E-mail address: zhong@maths.jyu.fi