Real Sociedad Matemática Española

All that Math

Portraits of mathematicians as young readers

Celebrating the Centennial of
Real Sociedad Matemática Española

Revista Matemática Iberoamericana

All that Math

Portraits of mathematicians as young readers

Revista Matemática Iberoamericana

All that Math

Portraits of mathematicians as young readers

Celebrating the Centennial of
Real Sociedad Matemática Española

Antonio Córdoba, José L. Fernández, Pablo Fernández
(Editors)

ISBN: 978-84-615-2900-1

Diseño de la edición: PFG.

Diseño de cubierta: Pepe Herrera Taller Gráfico.
`http://www.tallergraficopepeherrera.com`.

Ilustración de la cubierta. "La quinta luz", Santiago de la Quintana, 2011.

Impresión: Aventura Gráfica. C/ Lepanto 12-13, 19004-Guadalajara.
`http://www.aventuragrafica.com`.

Libro impreso en papel estucado semimate de 115 gramos fabricado con pastas libres de cloro elemental.

La quinta luz
Santiago de la Quintana, 2011.

Preface

As mathematicians we work on our research program, write up and publish our results, deliver lectures here and there disseminating our theorems and, once in a while, we attend the lectures and read the papers of our fellow mathematicians.

Some of us read a lot and some just rarely. But all of us have had that special experience of reading a paper (attending a lecture or having a mathematical discussion) which has been a personal turning point, perhaps because its techniques and elegance have illuminated ongoing work or because it has led us to explore a whole new area of mathematics or because it has given us a new set of mathematical objectives. We all treasure those occasions and some of us, in the right context, even like to share them.

On the occasion of its centennial, the Real Sociedad Matemática Española asked the editors of the Revista Matemática Iberoamericana to contribute to the occasion by editing and publishing a commemorative issue of the journal. We decided to take the opportunity to ask distinguished mathematicians to write down their personal experiences as readers of research papers. We are convinced that such a collection of essays, written by a set of world-class researchers, will be interesting reading for mathematicians in general and that it will contribute to understanding the role of research journals in the development of mathematics.

We asked longtime friends of the Revista Matemática Iberoamericana, both authors and members of past and present editorial boards who have fostered and cared for Revista, to contribute with an essay about a paper –not necessarily the most important publication in the field– which, in one way or another, had a deep impact on their own mathematical careers, especially at its early stages, and giving the special reasons why "that particular publication" got their attention and affected their research.

We have been very pleased, even overwhelmed, with the response of the friends of the Revista to our call. The reader will find among the articles in this book a surprisingly wide and rich variety of points of view, styles and approaches to this topic, all of which we appreciate greatly and have fully respected. Some are personal recollections with interesting anecdotes and charming stories about some well-known mathematician, while others present the state of the art of some specific field. Some papers are exquisitely technical while some others are perfect bedtime stories.

We offer our deepest thanks to all of the authors for sharing these contributions with all of us.

The Editors

Authors

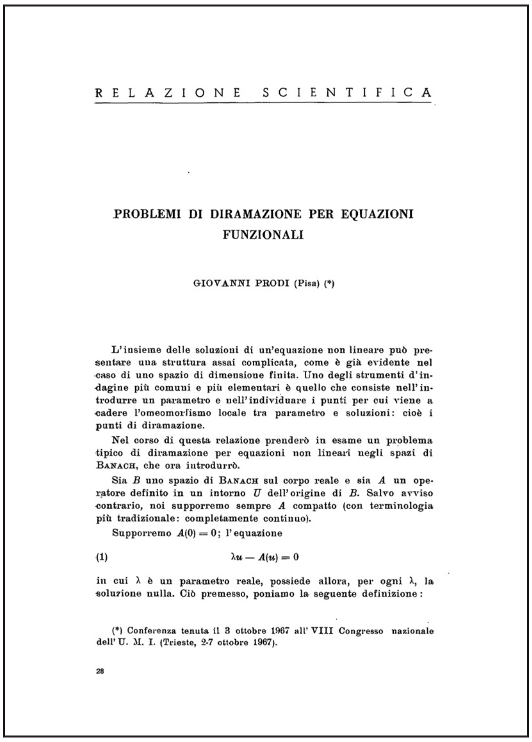
ANTONIO AMBROSETTI was born in Bari (Italy), in 1945. He is currently Full Professor at SISSA, Trieste. His research field is nonlinear analysis.

On a paper by Giovanni Prodi on bifurcation

ANTONIO AMBROSETTI*

1. Introduction

Many papers have influenced my work, but the very first one is a survey paper by G. Prodi [3] dealing with bifurcation theory. I was fascinated by the interplay between abstract methods which are motivated by, and hence applied to, concrete problems like fluid-dynamics and elasticity. Among other things, I learned that partial differential equations arising in mathematical physics are strongly linked to nonlinear functional analysis. The correct approach to this field must have its starting point in concrete problems modeled by differential equations. These applications motivate the right choice of the abstract tools and the advances to be carried out in order to obtain the expected results.



RELAZIONE SCIENTIFICA

PROBLEMI DI DIRAMAZIONE PER EQUAZIONI FUNZIONALI

GIOVANNI PRODI (Pisa) (*)

L'insieme delle soluzioni di un'equazione non lineare può presentare una struttura assai complicata, come è già evidente nel caso di uno spazio di dimensione finita. Uno degli strumenti d'indagine più comuni e più elementari è quello che consiste nell'introdurre un parametro e nell'individuare i punti per cui viene a cadere l'omeomorfismo locale tra parametro e soluzioni: cioè i punti di diramazione.

Nel corso di questa relazione prenderò in esame un problema tipico di diramazione per equazioni non lineari negli spazi di BANACH, che ora introdurrò.

Sia B uno spazio di BANACH sul corpo reale e sia A un operatore definito in un intorno U dell'origine di B. Salvo avviso contrario, noi supporremo sempre A compatto (con terminologia più tradizionale: completamente continuo).

Supporremo $A(0) = 0$; l'equazione

(1) $$\lambda u - A(u) = 0$$

in cui λ è un parametro reale, possiede allora, per ogni λ, la soluzione nulla. Ciò premesso, poniamo la seguente definizione:

(*) Conferenza tenuta il 3 ottobre 1967 all' VIII Congresso nazionale dell' U. M. I. (Trieste, 2-7 ottobre 1967).

28

Many physical problems can be formulated as an operator equation $F(\lambda, u) = 0$ where u belongs to an appropriate Banach function space X and λ is a real parameter with a specific physical meaning. Often there is a threshold λ^* such that new solutions arise when λ crosses λ^*, which is called a bifurcation value for the problem.

Prodi outlines three different approaches that can be used to find bifurcation points: (i) analytical methods, (ii) degree theoretical arguments and, in the case that the problem is variational, (iii) Morse theory. The abstract results are used to discuss two main applications:

(i) the motions of a viscous fluid between two cylinders and

(ii) the equilibria of an elastic clamped plate.

*ambr@sissa.it.

2. Analytical methods

Let X, Y be Banach spaces and let $F \in C^2(\mathbb{R} \times X, Y)$ (weaker regularity assumption could be given). We assume the equation $F(\lambda, u) = 0$ has the *trivial solution* $u = 0$ for all $\lambda \in \mathbb{R}$. Let us denote by S the set of nontrivial solutions of $F(\lambda, u) = 0$. A value λ^* is a bifurcation point for $F = 0$ if $(\lambda^*, 0) \in \overline{S}$, the closure of S in $\mathbb{R} \times X$. It is easy to see that if λ^* is a bifurcation point then $D_u F(\lambda^*, 0)$ cannot be invertible as a linear map from X to Y. One main purpose of bifurcation theory is to find sufficient conditions for the existence of bifurcation points. Suppose that

(F1) there exists $\phi^* \in X$ such that $\ker[D_u F(\lambda^*, 0)] = \mathbb{R}\phi^*$;

(F2) range$[D_u F(\lambda^*, 0)]$ is closed and has codimension 1.

Using the Lyapunov-Schmidt reduction one proves

Theorem 1 *Suppose that* (F1) *and* (F2) *hold. Moreover, letting* $M := D_{u,\lambda}F(\lambda^*, 0)$, *assume that* $M(\phi^*) \notin$ *Range*$[D_u F(\lambda^*, 0)]$. *Then* λ^* *is a bifurcation point for* $F = 0$.

Remark 1 If $X = Y$ and $F(\lambda, u) = \lambda u - G(u)$, the previous assumptions are verified provided λ^* is a *simple eigenvalue* of $G'(0)$.

The interest of theorem 1 relies on the fact that it deals with a very general equation $F(\lambda, u) = 0$.

Theorem 1 has many applications to Mathematical Physics. Among them, Prodi considers a problems arising in fluid-dynamics. Consider two coaxial rotating cylinders of ray $r_1 < r_2$ filled by a viscous fluid. Starting from the stationaly Navier–Stokes system, one is lead to look for solutions $(u(r,z), v(r,z))$, T-periodic in z, of the bounday value problem

$$\begin{cases} \nu\mathcal{L}^2 u + a(r)v_z + M(u,v) = 0, \\ \nu\mathcal{L}^2 v + bu_z + N(u,v) = 0, \\ u(r_1,z) = u(r_2,z) = u_r(r_1,z) = u_r(r_2,z) = 0, \\ v(r_1,z) = v(r_2,z) = 0. \end{cases} \tag{1}$$

Above $\nu > 0$ is the inverse of the Reynolds number, $a(r) \geq 0$, $b > 0$,

$$\mathcal{L} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2} - \frac{1}{r^2}.$$

and M, N are suitable homogeneous polynomials of degree 2 depending upon $u, u_r, u_z, u_{rr}, u_{zz}, v, v_r, v_z$.

Problem (1) has the trivial solution $u(r) = v(r) \equiv 0$ which corresponds to the *Couette flow*. It is possible to show that for suitable choices of T, Theorem 1 applies and (1) has a bifurcation point $\nu^* > 0$. When $\nu > \nu^*$ the fluid motion becomes turbulent.

3. Degree theoretical approach

We restrict here the attention to the case in which $X = Y$ and $F(\lambda, u) = \lambda u - G(u)$, where $G : X \to X$ is *compact*. The form of F allows us to employ the Leray-Schauder topological degree. Since the Leray-Schauder index of the trivial solution $u = 0$ changes when λ passes through an odd eigenvalue of $G'(0)$, it can be proved that

> *Every eigenvalue of $G'(0)$ with odd multiplicity is a bifurcation point for $G(u) = \lambda u$.*

This theorem goes back to a celebrated paper by M. A. Krasnoselski [1]. Moreover, in [4] a global version of the result is given.

4. The case of variational operators

If X is a Hilbert space and G is variational (namely there exists $\Phi \in C^2(X, \mathbb{R})$ such that $\nabla\Phi(u) = G(u)$) one can use critical point theory. Actually, solutions of $G(u) = \lambda u$ are the stationary points $u \in X$ of the functional

$$J_\lambda(u) = \frac{\lambda}{2}\|u\|^2 - \Phi(u).$$

For every $\lambda \in \mathbb{R}$, $u = 0$ is a critical point of J_λ. Moreover $u = 0$ is nondegenerate in the sense of Morse whenever λ is not an eigenvalue of $G'(0)$ and its nature changes when λ crosses the eigenvalues. To evaluate this change we set $U_\lambda^- = \{U \in H : J_\lambda(u) < 0\}$ and let m^j denote the Betti number of the j-th homology group

$$H_j(U_\lambda^- \cap B_\varepsilon, U_\lambda^- \cap B_\varepsilon \setminus 0),$$

where B_ε is the ball centered at 0 with radius ε.

Let us introduce the multi-index

$$\ell_\lambda = [m^0, m^1, \ldots, m^q, \ldots],$$

It is known that $m^q = \delta_q^s$ where δ_q^s is the Kronecker symbol and s is the number of the eigenvalues λ_i of $G'(0)$ greater than λ, counted with their multiplicity. Therefore, if $\alpha < \lambda_k < \beta$ it follows that $\ell_\alpha \neq \ell_\beta$. It is worth pointing out that ℓ_λ is sharper than the Leray–Schauder index. Actually, this latter equals $(-1)^s$ and hence does not change crossing an even eigenvalue of $G'(0)$, while ℓ_λ does.

By an elegant argument, it is possible to show that ℓ_λ remains constant if λ varies in an interval T which does not contain any eigenvalue λ_k. It follows:

Theorem 2 *If G is a compact potential operator then any eigenvalue of $G'(0)$ with finite multiplicity is a bifurcation point for $G(u) = \lambda u$.*

Remark 2 Theorem 2 was first proved by Krasnoselski by means of a different argument. The complete proof of Theorem 2 is carried out in [2].

As an important application of Theorem 2, Prodi discusses a bifurcation problem arising in nonlinear elasticity. According to the Van Karman theory, the buckling states of a clamped plate $\Omega \subset \mathbb{R}^2$, Ω bounded, are the nontrivial solutions (u, f) of

$$\begin{cases} \Delta^2 f & = -\mu\, \mathcal{B}(u,u), & \text{in } \Omega \\ \Delta^2 u & = \mu\, \mathcal{B}(F,u) + \mathcal{B}(f,u), & \text{in } \Omega \\ u = f & = u_\nu = f_\nu = 0, & \text{on } \partial\Omega \end{cases} \tag{2}$$

where $\mu > 0$ is the bifurcation parameter, F is given, u_ν, f_ν denote the outer normal derivatives at $\partial\Omega$ and

$$\mathcal{B}(f,g) = f_{xx}g_{yy} + f_{yy}g_{xx} - 2f_{xy}g_{xy}.$$

If K denotes the Green operator of Δ^2 on $X = H_0^2(\Omega)$, we have that

$$f = -\mu\, K\, \mathcal{B}(u,u).$$

Hence weak (and by regularity strong) solutions of (2) are the $u \in X$ such that $u = \mu G(u)$ where

$$G(u) = K\mathcal{B}(F,u) - K\mathcal{B}(K\mathcal{B}(u,u),u).$$

It is easy to check that G is a compact variational operator. If $\mu \neq 0$, we are in the preceding abstract setting, with $\lambda = \mu^{-1}$. The linearized problem $v = \mu G'(0)v$ is nothing but

$$\Delta^2 v = \mu\, \mathcal{B}(F,v), \quad v \in X. \tag{3}$$

Let us point out that the multiplicity of the eigenvalues, including the first one, of this linear problem can be an even integer. However, an application of Theorem 2 yields to show that any eigenvalue μ_k of (3) is a bifurcation point for (2).

I believe that any young researcher could have a great benefit reading this paper: he will have an idea of many of the most useful tools of nonlinear functional analysis, learning the correct way to carry out research in this beautiful field.

References

[1] Krasnoselski, M. A.: *Topological methods in the teory of nonlinear integral equations*. Pergamon Press, 1964.

[2] Marino, A. and Prodi, G.: La biforcazione nel caso variazionale. *Boll. Unione Mat. Ital.* **3** (1975), 1–32.

[3] Prodi, G.: Problemi di diramazione in analisi nonlineare. *Boll. Unione Mat. Ital.* **22** (1967), 413–433.

[4] Rabinowitz, P. H.: Some global results for nonlinear eigenvalue problems. *J. Funct. Anal.* **7** (1971), 485–513.

Kari Astala (Helsinki, Finland, 1953) obtained his PhD in mathematics at Helsingin Yliopisto in 1980 (advisor: Klaus Vala). He is Professor at the University of Helsinki. His main research areas are complex and harmonic analysis, PDEs and quasiconformal mappings with their applications.

Miracles of holomorphic motions

Kari Astala*

1. λ-lemma

What draws us to mathematics and maintains its spell upon us? We all have our own reasons, tastes and backgrounds, but I cannot think I would be alone in admiring the unexpected strikes of power our science presents every now and then: No matter how implausible or even counter-intuitive a mathematical fact might at first appear, once its proven and booked, there will be no objections towards the result –it is our own thinking and prejudices that we must change, according to the unquestionable facts of nature.

A beautiful example of such phenomena are no doubt the Holomorphic Motions, discovered by Mañé, Sad and Sullivan [10] and then developed by many others, that have hold their magic since their invention in the mid 80's. The notion can be explained in few lines to anyone interested in mathematics, and with an intuitive interpretation even beyond –a deformation of the space with a nonstandard concept of time; time is assumed to vary holomorphically. Nothing more is assumed and yet the conclusions are very strong and unexpected. The first properties of holomorphic motions can be proven with classical methods from the beginning of 20th century –and yet the method was found only recently, in the mid 80's within the modern study of complex dynamics. One cannot avoid an everlasting astonishment!

Computer animations revolutionized complex dynamics in the early 80's, and among phenomena observed was strong geometric –not only topological– stability in perturbations of hyperbolic systems; see Figure 1 for a typical illustration. Mañé, Sad and Sullivan [10] realized that these phenomena can be completely understood in terms of the following fundamental notion.

Definition 1. Let A be any subset of the complex plane $\mathbb{C}$. Then a *holomorphic motion* of A, parametrized by the unit disk $\mathbb{D}$, is a map

$$\Phi : \mathbb{D} \times A \to \mathbb{C}$$

such that

i) For any fixed $\lambda \in \mathbb{D}$, the map

$$a \to \Phi(\lambda, a) = \Phi_\lambda(a) \quad \text{is an injection.}$$

*kari.astala@helsinki.fi.

ii) For any fixed $a \in A$, the map

$$\lambda \to \Phi(\lambda, a) \text{ is holomorphic in } \mathbb{D}.$$

iii) The mapping Φ_0 is the identity on A,

$$\Phi(0, a) = a, \qquad \text{for every } a \in A.$$

Motions in the complex plane are enough for this presentation, but it does not take much effort to extend the discussion to the Riemann sphere $\overline{\mathbb{C}}$. Combining with a Möbius transform, we may and will always assume that $\Phi_\lambda(z)$ fixes the points $z = 0$ and $z = 1$.

In practical terms, we are discussing deformations of A with no assumptions at all on the "space" variable –continuity or even measurability– while the key property is the holomorphic dependence required from the "time" parameter λ. As innocent as this mere last assumption may appear, let us see how far it can take us –be ready for surprises!

That continuity occurs after all is a consequence of the remarkable λ-lemma of Mañé–Sad–Sullivan [10], the first step towards classifying and understanding the holomorphic motions. The result is best formulated in terms of a notion describing a weak form of scale invariance –the quasisymmetric mappings of Tukia and Väisälä [19]: A mapping $f : A \to \mathbb{C}$ is called quasisymmetric if the condition

$$\frac{|f(z) - f(x)|}{|f(z) - f(w)|} \leq \eta\left(\frac{|z - x|}{|z - w|}\right) \qquad \text{for all } \; x, z, w \in A,$$

holds for some modulus of continuity η (= increasing homeo $\eta : \mathbb{R}_+ \to \mathbb{R}_+$). Here recall that similarities, *i.e.* compositions of rotations, translations and scalings, are exactly the quasisymmetric maps with $\eta(t) \equiv t$. An arbitrary quasisymmetry preserves the relative sizes and geometric roundishness of sets, though not in general the rectifiability or even the Hausdorff dimension.

Theorem 1 (λ-lemma, Mañé–Sad–Sullivan) *If $\Phi : \mathbb{D} \times A \to \mathbb{C}$ is a holomorphic motion, then Φ has an extension to $\widehat{\Phi} : \mathbb{D} \times \overline{A} \to \mathbb{C}$ such that*

i) $\widehat{\Phi}$ is a holomorphic motion of the closure $\overline{A}$.

ii) Each $\widehat{\Phi}_\lambda(\cdot) : \overline{A} \to \mathbb{C}$ is quasisymmetric, in particular continuous.

iii) $\widehat{\Phi}$ is jointly continuous in (λ, a).

In view of the result, in a world where the (physical) time is holomorphic, according to *ii*) one can not even tear a piece of paper!

A typical illustration of the λ-lemma is given in Figure 1, where the polynomial $P_c(z) = z^2 + c$ with Julia set the “Douady rabbit” is holomorphically perturbed. It is easy to see that under a holomorphic deformation of parameters the repelling fixed points move holomorphically, and these are dense in the Julia set. By Theorem 1 the motion extends to $J(P_c)$, inducing an equivariant quasisymmetry between the respective Julia sets.

For the lemma, simply consider the holomorphic mapping

$$g(\lambda) := \frac{\Phi(\lambda, z) - \Phi(\lambda, x)}{\Phi(\lambda, z) - \Phi(\lambda, w)}, \quad \lambda \in \mathbb{D},$$

with values in $\Omega = \mathbb{C} \setminus \{0, 1\}$. Since holomorphic mappings contract the hyperbolic metric and this is complete, we see that $|g(\lambda)| \leq \eta_{|\lambda|}\big(|g(0)|\big)$, where $\eta_{|\lambda|}(t) \to 0$ as $t \to 0$. This establishes the quasisymmetry, condition *ii*), and the two other claims quickly follow.

The Mañé–Sad–Sullivan proof was so simple and the result so unexpected, that –to be honest– upon seeing the argument each of us complex analysts stood in bewilderment –how is this possible?

The authors themselves used the λ-lemma to prove an important result in complex dynamics, the density of structurally stable maps within the polynomials of any given degree. This just falls short of the celebrated Fatou conjecture of the density of hyperbolic maps.

Density of structural stability for all rational maps requires an improved version of the λ-lemma, provided by Sullivan and Thurston [18]. They showed that any motion ϕ of any set A extends to a motion of the whole space $\mathbb{C}$, but were able to prove this only for a small fraction of the parameters, for $|\lambda| \leq \varepsilon$; in an accompanying paper Bers and Royden [6] showed that one can take $\varepsilon = 1/3$. It is natural to enquire whether the extension works for the whole parameter disk $\mathbb{D}$. This important question was open for some years, with researchers in the area debating in both directions and looking for counterexamples –a complete extended version might be too good to be true? However, with an unexpected approach, using methods from several complex variables, Z. Słodkowski [12] hit the jackpot.

Theorem 2 (Słodkowski) *Every holomorphic motions of any set $A \subset \mathbb{C}$ is the restriction of a holomorphic motion of the entire complex plane $\mathbb{C}$, with the same parameter disk $\mathbb{D}$.*

There are now several approaches to Słodkowski’s theorem –it suffices to find a problem, typically a PDE, for which solutions depending holomorphically on a parameter are uniquely determined by their values at any single point. All these tended to be somewhat technical, until Chirka gave an elegant approach in terms of a non-linear Cauchy problem; see [7] or *e.g.* [3, Chapter 12].

Theorem 2 appears a complete and final word –yet, is there still room for further improvements? It is well known that for mappings of the entire plane $\mathbb{C}$, the quasisymmetry admits an equivalent analytic or infinitesimal formulation, leading us to first order PDE's and the well developed theory of quasiconformal mappings. Thus we have a detailed description of holomorphic motions of the entire plane –and through Słodkowski's theorem– of all holomorphic motions.

Put the story short, a mapping $f : \mathbb{C} \to \mathbb{C}$ is quasisymmetric if and only if it is a homeomorphic $W^{1,2}_{loc}(\mathbb{C})$-solution to the elliptic Beltrami equation

$$\frac{\partial f}{\partial \bar{z}}(z) = \mu(z)\frac{\partial f}{\partial z}(z), \qquad \|\mu\|_\infty \le k < 1, \tag{1}$$

where μ is bounded and measurable. Further, for compactly supported coefficients, a Neumann-series development shows that in (1) the derivatives depend holomorphically on μ, hence the same is true for the values $f(z) = f_\mu(z)$. For general coefficients we use a limiting argument combined with a suitable normalization of the mapping. As a consequence we see that holomorphic deformations of μ give rise to holomorphic motions.

We now arrive at the characterization of all global motions; for further details see [3, Section 12.3].

Theorem 3 *The following are equivalent:*

- $f_\lambda(z) = \Phi(\lambda, z)$, $\lambda \in \mathbb{D}$, *defines a holomorphic motion of* $\mathbb{C}$.

- $f_\lambda \in W^{1,2}_{loc}(\mathbb{C})$ *are homeomorphic solutions to the PDE*

$$\frac{\partial f}{\partial \bar{z}}(z) = \mu_\lambda(z)\frac{\partial f}{\partial z}(z), \qquad f(0) = 0,\ f(1) = 1,$$

where $\|\mu_\lambda\|_\infty \le |\lambda|$, $\lambda \in \mathbb{D}$, *and* $\lambda \mapsto \mu_\lambda(z)$ *is holomorphic (as an* L^∞*-valued function).*

For a given motion $\Phi(\lambda, z)$, the ellipticity bounds $\|\mu_\lambda\|_\infty \le |\lambda|$ in the second condition follow from Schwartz lemma (once holomorphicity [6] of μ_λ is shown), and the bounds are optimal. As we will see below, this means that the distortion under every holomorphic motion can be described by exact quantitative bounds.

At least for the author of this article, it is still –after all these years– almost breathtaking to see how far the simple notion of Definition 1 can take us!

2. Quasiconformal mappings

A more traditional way to state the condition (1) is the following equivalent form, requiring the operator norm

$$\|Df(z)\|^2 \le K J(z,f) \quad \text{for almost every } z, \qquad K := \frac{1+k}{1-k} \in [1,\infty). \quad (2)$$

The $W^{1,2}_{loc}$-mappings with this property are called K-quasiregular, or K-quasiconformal when homeomorphisms.

In the late 80's a small Revista-paper of mine [1] got me interested in optimal bounds for distortion of Hausdorff dimension under a K-quasiconformal mapping. Particularly necessary appeared the bounds for the dimension of a K-quasicircle, the image of a line or circle under a global K-quasiconformal mapping $f : \mathbb{C} \to \mathbb{C}$. At the time many of the distortion properties of quasiconformal mappings had already obtained their optimal form by classical methods, such as the Hölder continuity with exponent $1/K$ where *e.g.* symmetrization methods are available. However, the dimension bounds were a mystery, even though Iwaniec and Martin [9] had already found the natural conjectures.

One expects the worst dimension distortion under a quasiconformal mapping to happen within the family of Cantor sets –but how to get a global control of Cantor sets with either of the conditions (1) or (2)? Other standard methods in the quasiconformal theory, such as path families or capacity estimates, seemed equally useless.

Around these times came Słodkowski's theorem, the extended λ-lemma and all the consequences it brought with: As we see with Theorems 2–3, quasiconformal mappings and holomorphic motions are just different sides of the same coin! With the precise ellipticity bounds of the theorem, it became an equivalent problem to study distortion of dimension in holomorphic motions of Cantor sets. The Cantor sets one can identify as an attractor of a simple dynamical system, and this brings us the Thermodynamical formalism [11], [20], developed for understanding invariant measures for general dynamical systems. Within the formalism we have quantities such as the "topological pressure" determining the dimension of the attractor. It turns out that these quantities are not difficult to control under a holomorphic motion.

In the end, making the above intuition rigorous by using decompositions of quasiconformal mappings [2], using holomorphic motions and ideas from thermodynamical formalism led to even stronger results, the optimal Sobolev regularity and area distortion bounds for quasiconformal mappings conjectured by Gehring and Reich [8].

Theorem 4 ([2]) *Suppose $f : \Omega \to \Omega'$ is a K-quasiconformal mapping. Then $f \in W^{1,p}_{loc}(\Omega)$ for every exponent $p < \frac{2K}{K-1}$.*

In fact, it follows that locally $Df \in$ weak-L^{p_0}, $p_0 = \frac{2K}{K-1}$. Simple examples show that in general $f \notin W^{1,p_0}_{loc}(\Omega)$. As a consequence of the Theorem,

Corollary 1. *If $f : \Omega \to \Omega'$ is a K-quasiconformal mapping and $A \subset \Omega$ is compact, then*

$$\frac{1}{K}\Big(\frac{1}{\dim(A)} - \frac{1}{2}\Big) \leqslant \frac{1}{\dim(f(A))} - \frac{1}{2} \leqslant K\Big(\frac{1}{\dim(A)} - \frac{1}{2}\Big) \qquad (3)$$

Further, for both estimates we have sets A and maps f so that the equality holds.

Recalling that basically, quasiconformal mappings do not differ from holomorphic motions we have obtained the optimal bounds for distortion of Hausdorff dimension under an arbitrary Holomorphic motion.

What remained above was the original goal, the dimension of K-quasicircles. Given a holomorphic motion $\Phi(\lambda, z)$, Corollary 1 gives for any set A of Hausdorff dimension one the estimate $\dim(\Phi_\lambda A) \leq 1 + |\lambda|$. However, by the Mañé–Sad–Sullivan λ-lemma, continuums are preserved under the motion, and this extra constraint forces an improved dimension bound.

It was natural to conjecture [2] that $\dim(\Phi_\lambda \mathbb{S}^1) \leq 1 + |\lambda|^2$ for any holomorphic motion of the circle. And indeed, recently this was proved by Smirnov [16] who discovered a new class, the so called antisymmetric quasiconformal mappings, having extra symmetries, and then showed that any quasicircle is the image of $\mathbb{S}^1$ under such a mapping. Holomorphic perturbations of antisymmetric maps give motions with extra symmetries, yielding improved bounds in the thermodynamical arguments.

Theorem 5 (Smirnov) *Let Γ be a $K = \frac{1+k}{1-k}$-quasicircle. Then the Hausdorff dimension*

$$\dim(\Gamma) \leqslant 1 + k^2 \qquad (4)$$

The work of Smirnov leads also to new understanding of distortion properties quasisymmetric maps of the circle [14], [15], as well as suggests the use of holomorphic motions towards fundamental open questions in complex analysis, such as exact bounds for the multifractal spectrum of the harmonic measure [17]. These questions are also related to the sharpness of the bound (4), which remains open. We therefore pose the reader the following "homework":

Find a holomorphic motion of $\mathbb{S}^1$ with $\dim(\Phi_\lambda \mathbb{S}^1) = 1 + |\lambda|^2, \quad \lambda \in \mathbb{D}$!

For an experimental step towards this problem see [5].

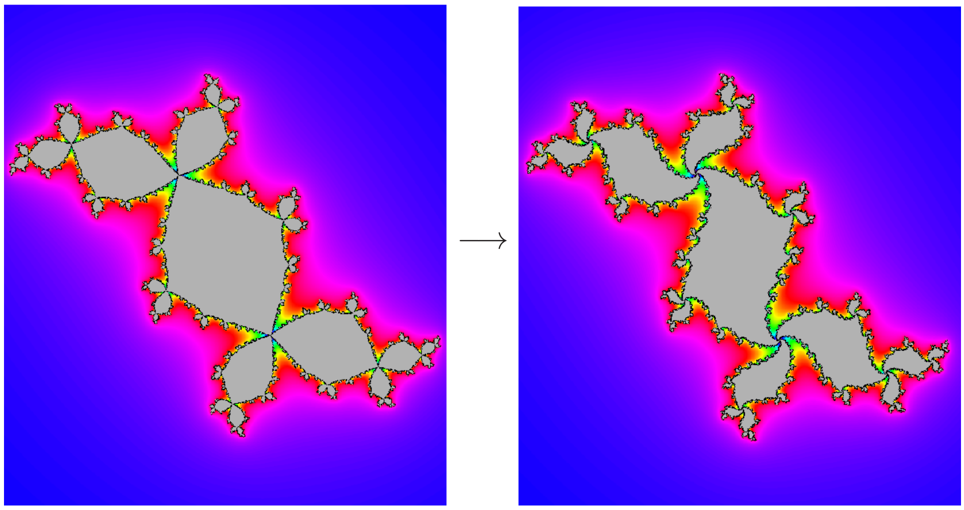


Figure 1: Holomorphic perturbations induce a quasisymmetry of Julia sets (Pictures drawn with a program of C. McMullen.)

What I have described above is, naturally, not the whole story –but I hope I have been able give some flavor of the developments the work [10] has led to. The holomorphic motions have of course become a basic tool in complex dynamics, but they remain a deep method also in analysis and PDE. At this moment I cannot resist the temptation and hint on the work [4] where the motions are used –towards lower semicontinuity properties in vector valued calculus of variations. I would expect this not to be the last area where holomorphic motions will appear indispensable.

References

[1] Astala, K.: Calderón's problem for Lipschitz classes and the dimension of quasicircles. *Rev. Mat. Iberoamericana* **4** (1988), 469–486.

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

[3] Astala, K., Iwaniec, T. and Martin, G.: *Elliptic partial differential equations and quasiconformal mappings in the plane.* Princeton University Press, 2009.

[4] Astala, K., Iwaniec, T., Prause, I. and Saksman, E.: Burkholder Integrals, Morrey's Problem and Quasiconformal Mappings. (Submitted)

[5] Astala, K., Rohde, S. and Schramm, O.: Self-similar Jordan curves. (Preprint)

[6] Bers, L. and Royden, H.: Holomorphic families of injections. *Acta Math.* **157** (1986), 259–286.

[7] Chirka, E. M.: On the propagation of holomorphic motions. *Dokl. Akad. Nauk* **397** (2004), 37–40.

[8] Gehring, F. W. and Reich, E.: Area distortion under quasiconformal mappings. *Ann. Acad. Sci. Fenn. Ser. A. I. Math.* **388** (1966), 3–15.

[9] Iwaniec, T. and Martin, G.: Quasiconformal mappings and capacity. *Indiana Univ. Math. J.* **40** (1991), 101–122.

[10] Mañé, R., Sad, P. and Sullivan, D.: On the dynamics of rational maps. *Ann. Sci. École Norm. Sup.* **16** (1983), 193–217.

[11] Ruelle, D.: *Thermodynamic formalism.* Encyclopedia of Mathematics and its Applications **5**. Addison-Wesley, Reading, Mass., 1978.

[12] Słodkowski, Z.: Holomorphic motions and polynomial hulls. *Proc. Amer. Math. Soc.* **111** (1991), 347–355.

[13] Słodkowski, Z.: Holomorphic motions commuting with semigroups. *Studia Math.* **119** (1996), 1–16.

[14] Prause, I.: A remark on quasiconformal dimension distortion on the line. *Ann. Acad. Sci. Fenn. Math.* **32** (2007), no. 2, 341–352.

[15] Prause, I. and Smirnov, S.: Quasisymmetric distortion spectrum. *Bull. Lond. Math. Soc.* **43** (2011), 267–277.

[16] Smirnov, S.: Dimension of quasicircles. *Acta Math.* **205** (2010), 189–197.

[17] Smirnov, S.: *Quasiconformal maps and harmonic measure.* Available at `http://www.unige.ch/s̃mirnov/slides/index.html`.

[18] Sullivan, D. and Thurston, W.: Extending holomorphic motions. *Acta Math.* **157** (1986), 243–257.

[19] Tukia, P. and Väisälä, J.: Quasisymmetric embeddings of metric spaces. *Ann. Acad. Sci. Fenn. Ser. A. I. Math.* **5** (1980), 97–114.

[20] Zinsmeister, M.: *Thermodynamic formalism and holomorphic dynamical systems.* SMF/AMS Texts and Monographs **2**, 2000.

Luis Barreira obtained his PhD in mathematics at Penn State University, USA, in 1996. He is currently Full Professor of the Instituto Superior Técnico, Lisboa. His research interests include differential equations, dynamical systems and ergodic theory.

Geometric constructions and thermodynamic formalism

Luis Barreira*

I discuss briefly the early contribution of Moran to dimension theory of dynamical systems in his paper "Additive functions of intervals and Hausdorff measure", *Proc. Cambridge Philos. Soc.* **42** (1946), 15–23.

Pat Moran (1917–1988) was an Australian mathematician who made significant contributions to probability theory and its applications. Today, the Moran Medal is awarded by the Australian Academy of Science, normally every two years, recognizing outstanding research by scientists under the age of forty in the fields of probability and statistics.

Pat Moran

Thus, it may come as a surprise that I want to emphasize a particular paper of Moran, namely [9], at the very early stage of his mathematical career –during which incidentally he never got a PhD. In my opinion, this paper should be considered one of the most significant early contributions to dimension theory of dynamical systems.

1. Dimension theory of geometric constructions

Let us first describe a particular geometric construction in $\mathbb{R}$. We consider constants $\lambda_1, \dots, \lambda_p \in (0,1)$ and disjoint closed intervals $\Delta_1, \dots, \Delta_p \subset \mathbb{R}$ of lengths $\lambda_1, \dots, \lambda_p$. For each $k = 1, \dots, p$, we choose again p disjoint closed intervals $\Delta_{k1}, \dots, \Delta_{kp} \subset \Delta_k$ of lengths $\lambda_k\lambda_1, \dots, \lambda_k\lambda_p$. Iterating this procedure, for each $n \in \mathbb{N}$ we obtain p^n disjoint closed intervals $\Delta_{i_1\cdots i_n}$ of lengths $\prod_{k=1}^n \lambda_{i_k}$, and we define the limit set

$$F = \bigcap_{n=1}^{\infty} \bigcup_{i_1\cdots i_n} \Delta_{i_1\cdots i_n}. \tag{1}$$

The following formula for the Hausdorff dimension $\dim_H F$ of the set F was obtained by Moran in [9].

*barreira@math.ist.utl.pt.

Theorem 1 *We have* $\dim_H F = s$*, where* s *is the unique real number such that*

$$\sum_{k=1}^{p} \lambda_k^s = 1. \tag{2}$$

It is remarkable that the Hausdorff dimension of the limit set F does not depend on the location of the intervals $\Delta_{i_1\cdots i_n}$ but only on their lengths. Even today, most probably this observation is not given the value that it rightfully deserves (in particular, the result in Theorem 1 has been often referred to in the special case when the intervals $\Delta_{i_1\cdots i_n}$ are obtained from a family of contraction maps).

Moran's result can be broadly generalized. Before describing briefly some of these generalizations, we recall the notion of geometric construction for an arbitrary symbolic dynamics. Given $p \in \mathbb{N}$, we consider the space of sequences $\Sigma_p = \{1, \dots, p\}^{\mathbb{N}}$ equipped with the distance

$$d(\omega, \omega') = \sum_{k=1}^{\infty} e^{-k} |i_k - i'_k|,$$

where $\omega = (i_1, i_2, \dots)$ and $\omega' = (i'_1, i'_2, \dots)$. We also consider the shift map

$$\sigma \colon \Sigma_p \to \Sigma_p$$

defined by $\sigma(i_1, i_2, \dots) = (i_2, i_3, \dots)$. A *geometric construction* in $\mathbb{R}^m$ consists of:

1. a compact σ-invariant set $\Sigma \subset \Sigma_p$, for some $p \in \mathbb{N}$;
2. a decreasing sequence of compact sets $\Delta_{i_1\dots i_n} \subset \mathbb{R}^m$ for each $\omega \in \Sigma$, with diameters $\operatorname{diam} \Delta_{i_1\dots i_n} \to 0$ when $n \to \infty$, such that

   $$\operatorname{int} \Delta_{i_1\dots i_n} \cap \operatorname{int} \Delta_{j_1\dots j_n} \neq \varnothing$$

   whenever $(i_1, \dots, i_n) \neq (j_1, \dots, j_n)$.

We also say that the geometric construction is *modeled* by Σ, and define its *limit set* by (1), with the union taken over all $i_1, \dots, i_n \in \{1, \dots, p\}$ such that $(i_1, \dots, i_n) = (j_1, \dots, j_n)$ for some sequence $(j_1, j_2, \dots) \in \Sigma$. For example, to model repellers and hyperbolic sets, one can consider geometric constructions modeled by topological Markov chains.

Let us point out that in order to determine or even to estimate the dimension of the limit set F, sometimes it is not sufficient to know the geometric shape of the sets $\Delta_{i_1\cdots i_n}$, in strong contrast to what happens in Theorem 1.

For example, the dimension can be affected by certain number-theoretical properties. To explain this phenomenon, let us consider a geometric construction in $\mathbb{R}^2$ modeled by Σ_2 such that the sets

$$\Delta_{i_1\cdots i_n} = (f_{i_1} \circ \cdots \circ f_{i_n})([0,1] \times [0,1])$$

are rectangles with sides of lengths a^n and b^n. More precisely, let us assume that they are obtained from the composition of the functions

$$f_1(x,y) = (ax, by) \quad \text{and} \quad f_2(x,y) = (ax - a + 1, by - b + 1),$$

for some fixed constants $a \in (0,1)$ and $b \in (0,1/2)$. The projection of the rectangle $\Delta_{i_1\ldots i_n}$ on the horizontal axis is an interval with right endpoint equal to

$$a^n + \sum_{k=0}^{n-1} j_k a^k, \tag{3}$$

where

$$j_k = \begin{cases} 0 & \text{if } i_k = 1, \\ 1-a & \text{if } i_k = 2. \end{cases}$$

For $a = (\sqrt{5}-1)/2$, we have $a^2 + a = 1$, and thus, for each $n > 2$ there is more than one vector $(i_1, \ldots, i_n)$ with the same given value in (3). This causes a larger concentration of the rectangles $\Delta_{i_1\ldots i_n}$ in certain regions of the limit set F. Therefore, in order to compute its Hausdorff dimension, it may be possible to replace several elements of a given cover by a single element. Thus, the set F may have a smaller Hausdorff dimension than expected, with respect to a certain generic value obtained by Falconer in [7]. An example was described by Pollicott and Weiss in [14], modifying a construction of Przytcki and Urbański in [15] that depends on delicate number-theoretical properties. The same phenomenon was observed by Neunhäuserer in [10].

2. Thermodynamic formalism and dimension theory

In [13], Pesin and Weiss extended the result of Moran in Theorem 1 to an arbitrary symbolic dynamics, with the help of the thermodynamic formalism (we refer to the books [5, 17, 20] for detailed presentations of the theory). The notion of topological pressure, which is the most basic notion of the thermodynamic formalism, was introduced by Ruelle in [16] for expansive transformations and by Walters in [19] in the general case.

In order to illustrate the relation between the dimension theory of dynamical systems and the thermodynamic formalism, we consider the numbers $\lambda_1, \ldots, \lambda_p \in (0,1)$ and we define a function $\varphi\colon \Sigma_p \to \mathbb{R}$ by

$$\varphi(i_1, i_2, \ldots) = \log \lambda_{i_1}. \tag{4}$$

Then, given $s \in \mathbb{R}$, the topological pressure of the function $s\varphi$ is given by

$$\begin{aligned}
P(s\varphi) &= \lim_{n\to\infty} \frac{1}{n} \log \sum_{i_1,\ldots,i_n} \exp\Big(s \sum_{k=1}^{n} \log \lambda_{i_k}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} \log \sum_{i_1,\ldots,i_n} \prod_{k=1}^{n} \lambda_{i_k}^s \\
&= \lim_{n\to\infty} \frac{1}{n} \log \Big(\sum_{i=1}^{p} \lambda_i^s\Big)^n \\
&= \log \sum_{i=1}^{p} \lambda_i^s.
\end{aligned}$$

That is, equation (2) is equivalent to the new equation

$$P(s\varphi) = 0, \tag{5}$$

which involves the topological pressure. Equation (5) was introduced by Bowen in [4] (in his study of quasi-circles) and is usually called *Bowen's equation*. It is also appropriate to call it *Bowen–Ruelle equation*, taking not only into account the fundamental role of the thermodynamic formalism substantially developed by Ruelle, but also his article [18]. Equation (5) establishes the connection between the thermodynamic formalism and the dimension theory of dynamical systems. Moreover, it has a rather universal character: virtually all known equations used to compute or to estimate the dimension of an invariant set of a dynamical system are particular cases of this equation or of an appropriate generalization (see [3]).

Now we present the result of Pesin and Weiss in [13] that extends Theorem 1 to an arbitrary symbolic dynamics. The value of the dimension is again given by Bowen's equation in (5).

Theorem 2 *For a geometric construction modeled by $\Sigma \subset \Sigma_p$ such that the sets $\Delta_{i_1\ldots i_n} \subset \mathbb{R}^m$ are balls of diameter $\prod_{k=1}^n \lambda_{i_k}$, for some numbers $\lambda_1, \ldots, \lambda_p \in (0,1)$, we have $\dim_H F = s$, where s is the unique real number satisfying $P(s\varphi) = 0$ with the function $\varphi\colon \Sigma \to \mathbb{R}$ given by (4).*

It turns out that the topological pressure is not adapted to all geometric constructions. Let us consider again a geometric construction such that all sets $\Delta_{i_1\ldots i_n}$ are balls. In Theorems 1 and 2 we have

$$\operatorname{diam} \Delta_{i_1\ldots i_n} = \prod_{k=1}^{n} \lambda_{i_k} \tag{6}$$

for every set $\Delta_{i_1\dots i_n}$. A geometric construction for which identity (6) fails is called a *nonstationary geometric construction.* The thermodynamic formalism is of no help in this situation. However, the nontrivial generalization given by the so-called *nonadditive thermodynamic formalism* can still be used with success. The main idea is to replace the sequence of functions

$$\varphi_n = \sum_{k=0}^{n-1} \varphi \circ \sigma^k \tag{7}$$

in the definition of topological pressure by an arbitrary sequence $\Psi = (\psi_n)_n$. While the functions φ_n in (7) satisfy the identity

$$\varphi_{n+m} = \varphi_n + \varphi_m \circ \sigma^n,$$

the functions ψ_n may have no similar property, giving rise to the expression of *nonadditive.* The notion of nonadditive topological pressure $P(\Psi)$ was introduced by Barreira in [1] using the theory of Carathéodory dimensions developed by Pesin (we refer to [11] for references and full details). This is a generalization of the notion of topological pressure, and it contains as a particular case the subadditive version earlier introduced by Falconer in [8]. In the additive case it recovers the notion of topological pressure introduced by Pesin and Pitskel' in [12].

The following result gives a formula for the dimension of the limit sets of a class of nonstationary geometric constructions, in terms of the nonadditive topological pressure. It was obtained by Barreira in [1].

Theorem 3 *Consider a geometric construction modeled by $\Sigma \subset \Sigma_p$ such that the sets $\Delta_{i_1\dots i_n}$ are balls of diameter $r_{i_1\dots i_n} < 1$. If there exists $\lambda \in (0,1)$ such that*

$$\lambda^{m+n} \leq r_{i_1\dots i_{n+m}} \leq r_{i_1\dots i_n} r_{i_{n+1}\dots i_m}$$

for every $(i_1, i_2, \dots) \in \Sigma$ and $n, m \in \mathbb{N}$, then $\dim_H F = s$, where s is the unique real number satisfying $P(s\Psi) = 0$, where Ψ is the sequence of functions $\psi_n \colon \Sigma \to \mathbb{R}$ defined by

$$\psi_n(i_1, i_2, \dots) = \log \operatorname{diam} \Delta_{i_1\dots i_n}.$$

We note that the proof in [1] requires that $r_{i_1\dots i_{n+1}} \geq \delta r_{i_1\dots i_n}$ for some constant δ. This assumption was recently removed by Cao, Feng and Huang in [6].

The equation $P(s\Psi) = 0$ is a nonadditive version of Bowen's equation in (5). We note that Theorem 3 contains both the result of Moran in Theorem 1 and the result of Pesin and Weiss in Theorem 2. It is not too much to consider that Theorem 3 is ultimately inspired in Moran's work, although certainly the theory required for the generalization is much more involved.

We refer the reader to the books [2, 3, 11] for related discussions and for a detailed exposition of selected topics of the dimension theory of dynamical systems.

References

[1] Barreira, L.: A non-additive thermodynamic formalism and applications to dimension theory of hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **16** (1996), 871–927.

[2] Barreira, L.: *Dimension and Recurrence in Hyperbolic Dynamics*. Progress in Mathematics **272**. Birkhäuser, 2008.

[3] Barreira, L.: *Thermodynamic Formalism and Applications to Dimension Theory*. Progress in Mathematics **294**. Birkhäuser, 2011.

[4] Bowen, R.: Hausdorff dimension of quasi-circles. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 259–273.

[5] Bowen, R.: *Equilibrium States and Ergodic Theory of Anosov Diffeomorphisms*. Lect. Notes in Math. **470**. Springer, 1975.

[6] Cao, Y.-L., Feng, D.-J. and Huang, W.: The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst.* **20** (2008), 639–657.

[7] Falconer, K.: The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 339–350.

[8] Falconer, K.: A subadditive thermodynamic formalism for mixing repellers. *J. Phys. A: Math. Gen.* **21** (1988), 1737–1742.

[9] Moran, P.: Additive functions of intervals and Hausdorff measure. *Proc. Cambridge Philos. Soc.* **42** (1946), 15–23.

[10] Neunhäuserer, J.: Number theoretical peculiarities in the dimension theory of dynamical systems. *Israel J. Math.* **128** (2002), 267–283.

[11] Pesin, Ya.: *Dimension Theory in Dynamical Systems: Contemporary Views and Applications*. Chicago Lectures in Mathematics. Chicago University Press, 1997.

[12] Pesin, Ya. and Pitskel', B.: Topological pressure and the variational principle for noncompact sets. *Functional Anal. Appl.* **18** (1984), 307–318.

[13] Pesin, Ya. and Weiss, H.: On the dimension of deterministic and random Cantor-like sets, symbolic dynamics, and the Eckmann–Ruelle conjecture. *Comm. Math. Phys.* **182** (1996), 105–153.

[14] Pollicott, M. and Weiss, H.: The dimensions of some self-affine limit sets in the plane and hyperbolic sets. *J. Statist. Phys.* **77** (1994), 841–866.

[15] Przytycki, F. and Urbański, M.: On Hausdorff dimension of some fractal sets. *Studia Math.* **93** (1989), 155–186.

[16] Ruelle, D.: Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification. *Trans. Amer. Math. Soc.* **185** (1973), 237–251.

[17] Ruelle, D.: *Thermodynamic Formalism.* Encyclopedia of Mathematics and its Applications **5**. Addison-Wesley, 1978.

[18] Ruelle, D.: Repellers for real analytic maps. *Ergodic Theory Dynam. Systems* **2** (1982), 99–107.

[19] Walters, P.: A variational principle for the pressure of continuous transformations. *Amer. J. Math.* **97** (1976), 937–971.

[20] Walters, P.: *An Introduction to Ergodic Theory.* Graduate Texts in Mathematics **79**. Springer, 1982.

Pilar Bayer (Barcelona, Spain, 1946) obtained her PhD in mathematics at the University of Barcelona in 1975. Since 1982, she has been Full Professor of Algebra at the University of Barcelona. Her research field is number theory.

(Photo by Marcel·lí Bayer)

Early experiences in Number Theory

Pilar Bayer*

I first encountered number theory in the autumn of 1967, during the last academic year of my undergraduate studies at the University of Barcelona. I had enrolled on a course entitled *Teoría de números*, which was to be given by Professor Enrique Linés.

Linés was an analyst who loved to teach undergraduate students, and whom I already knew from his earlier, highly stimulating lectures on Real Analysis. What is more, one of his assistants at the University was Ms. Griselda Pascual, my wonderful mathematics teacher at secondary school.

Linés began his lectures by saying that the purpose of the course was to carry out a joint study of a book that had just arrived in the Faculty Library. The book was, no less, the French translation of Borevich-Shafarevich's *Théorie des Nombres* [3].

As is well known, the late sixties was a tumultuous period everywhere. Unauthorized meetings, riots, fights between students and police, and strikes paralysed academic life at Spanish universities day after day.

Because of this, Linés' expectations for the course would remain unfulfilled, since we could only go through the first chapters of [3]; the local methods presented in chapter 4 and the analytical methods in chapter 5 remained completely untouched.

In 1970, after teaching mathematics for a year at a secondary school, I obtained a scholarship from the Spanish Ministry to take part in an innovative three-year program for training scientific researchers. At the same time, I began to teach as an assistant at the University of Barcelona and at the newly created Autonomous University of Barcelona (in Bellaterra, a small village 22 km from Barcelona).

To begin my research training, I studied the whole of Borevich–Shafarevich's book from scratch, trying to solve as many exercises as possible. This took me the whole of 1971. My next step was spelled out in a passage from Borevich–Shafarevich's book:

> *La théorie du corps de classes décrit la loi de décomposition des diviseurs premiers d'un corps quelconque k de nombres algébriques en facteurs dans une extension K/k si le groupe de Galois de cette extension est abélien* [...] *On connaît trés peu de résultats sur les lois de décomposition des nombres premiers dans les corps dont le groupe de Galois est non abélien.*

*bayer@ub.edu.

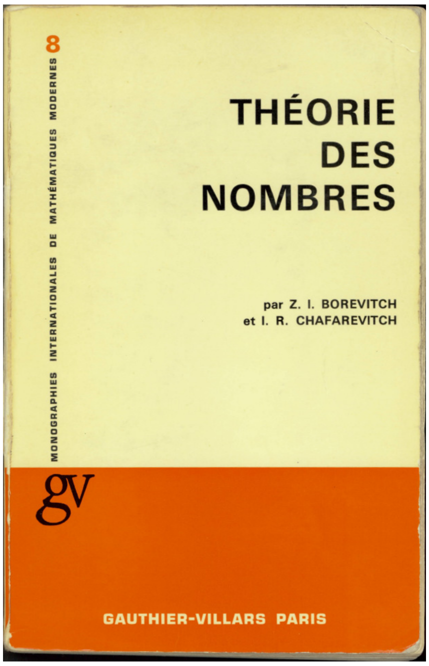


Figure 1: My copy of Borevich–Shafarevich's book *Théorie des Nombres*.

My journey through class field theory would last for the next two years. Luckily I was not alone, because Griselda (who was 20 years my senior) joined me on the project. It took us over a year to study the local theory, through Serre's book *Corps locaux* [6], and almost another year to familiarize ourselves with the global theory, through Artin–Tate's book *Class Field Theory* [1]. Both books were innovative, marvelously written, and presented an extremely profound theory in an accessible way. We studied their pages with great excitement, without even noticing the heat of the summer afternoons. It was not only a matter of understanding the contents of the books, but also of becoming familiar with the variety of tools used in them to obtain the main theorems. (We were blissfully unaware of the fact that the development of class field theory had taken more than 100 years.)

Our encounter with class field experience left Griselda and me exhausted, but we had the impression that we now had a good grounding in the modern developments in the field and might be ready to undertake some sort of personal research. The 100 bibliographical references at the end of Serre's *Corps locaux* were tantalizing, but where should we start?

I took advice from two mathematicians. The first was Jean Dieudonné. In a visit of him to Barcelona, he told me of Jacques Martinet's thesis, which had recently appeared in *Annals de l'Institut Fourier* [4]. Dieudonné was very kind and encouraging. The second was Francesc Tomàs, a Mexican mathematician with Catalan ancestors, who had just enrolled at the Autonomous University of Barcelona. Tomàs recommended that I read Jürgen Neukirch's papers.

Griselda studied Martinet's thesis carefully, in which he proved the existence of normal bases for the ring of integers of dihedral number fields of degree equal to $2p$, p denoting an odd prime. The analogous result in the abelian case had been proven by Hilbert long before.

My choice was Neukirch's paper [5], in *Inventiones Mathematicae*, dealing with the remarkable result that normal algebraic number fields are characterized by their absolute Galois group. (But first, I had to translate the paper from German to Catalan –my mother tongue– since at that time I could not read German fluently.)

The problems studied in those papers were an inspiration for us. Two years later, in April 1975, my former teacher and I were able to defend our theses at the University of Barcelona (becoming, on the same day, the second and third female Doctors in Mathematics at the University).

In May 1977, I moved to Regensburg University in Germany. There, in an ideal environment, I enjoyed the immense privilege of working in number theory with the Research Group conducted by Jürgen Neukirch, Günter Tamme and Manfred Knebusch, for more than three years.

Today, the Number Theory Research Group in Barcelona comprises some 40 people from the University of Barcelona, the Autonomous University of Barcelona, and the Polytechnic University of Catalonia. When I look back on these lines for the *Revista Matemática Iberoamericana*, the question I cannot help asking is this: what would have happened if the French translation of Borevich–Shafarevich's work had not been published in 1967?

References

[1] Artin, E.; Tate, J.: *Class field theory.* W. A. Benjamin, New York–Amsterdam, 1968.

[2] Bayer, P.; Neukirch, J.: On values of zeta functions and l-adic Euler characteristics. *Invent. Math.* **50** (1978/79), no. 1, 35–64.

[3] Borevitch, Z. I.; Shafarevitch, I. R.: *Théorie des nombres.* Traduit par Myriam et Jean-Luc Verley. Traduction faite d'après l'édition originale russe. Monographies Internationales de Mathématiques Modernes **8**. Gauthier-Villars, Paris, 1967.

[4] Martinet, J.: Sur l'arithmétique des extensions galoisiennes à groupe de Galois diédral d'ordre $2p$. *Ann. Inst. Fourier (Grenoble)* **19** (1969), fasc. 1, 1–80.

[5] Neukirch, J.: Kennzeichnung der p-adischen und der endlichen algebraischen Zahlkörper. *Invent. Math.* **6** (1969), 296–314.

[6] Serre, J.-P.: *Corps locaux.* Publications de l'Institut de Mathématique de l'Université de Nancago. VIII Actualités Sci. Indust. **1296**. Hermann, Paris, 1962.

Christopher J. Bishop (Syracuse, New York, 1959) obtained his PhD in mathematics at the University of Chicago in 1987 (advisor: Peter W. Jones). He is Professor at the Department of Mathematics of Stony Brook University. His main research interests are geometric function theory and its connections to other areas such as quasiconformal geometry, conformal dynamics, hyperbolic geometry, probability, computational geometry and numerical conformal mapping.

A random walk in analysis

Christopher J. Bishop*

1. Introduction

Which papers have had a big impact on my own work? When Antonio Córdoba and José Luis Fernández asked me to write about this, I started by making a list of some topics I've worked on and drew arrows to indicate when one idea led to another. The final version of my diagram is in Figure 11 and it includes three papers besides my own: Nick Makarov's paper on the dimension of harmonic measure, Peter Jones' traveling salesman paper and Dennis Sullivan's paper on hyperbolic convex hulls. Below I'll try to explain why each of these caught my attention and how it pushed my work in new directions.

2. Harmonic measure

I'm a Chicago Ph.D., but spent two years at Yale when my advisor, Peter Jones, moved there and I briefly shared an office with Stephen Semmes and Tim Steger who were Gibbs instructors. Stephen told me about his construction of a non-rectifiable closed curve such that the harmonic measures ω_1, ω_2 for opposite sides had a bounded ratio (*i.e.*, $\log d\omega_1/d\omega_2$ is bounded. If you don't know what ω is, just think of a random path running until it hits the curve and ω is the probability distribution of that first hitting point.) His paper [58] was hard for me to follow, but while trying to sort through it, I built a curve with dimension > 1 and the same bounded ratio property (giving me the first part of a thesis). This is a useful technique: fail to understand what some smart person has done and prove a different result with a simpler technique instead. (Applying this method to Tim Steger's description of his work resulted in our joint paper [30] about Fuchsian groups, representations and rigidity.)

I told Peter about the curve when he returned from a visit to UCLA and it prompted him to share his conversations with Lennart Carleson and John Garnett about a related problem: harmonic measures ω_1, ω_2 corresponding to opposite sides of closed curve are mutually absolutely continuous on the tangent points, but what happens on the set of non-tangent points? Must the measures be singular there? Luckily, Nick Makarov had already invented the right tool to solve this problem.

*bishop@math.sunysb.edu.

The idea of Makarov's paper [51] is that harmonic measure on the boundary of a simply connected domain acts like a random walk. More precisely, if we consider a disk $D(x, r)$ with $x \in \partial\Omega$ and $r \searrow 0$, then $\log \frac{\omega(D)}{r}$ behaves like a random walk on $\mathbb{R}$ whose step size is related to the "flatness" of the boundary near x at scale r. At a.e. tangent point the boundary is very flat and this quantity approaches a finite limit because the steps become small. At non-tangent points we expect $\log \frac{\omega}{r}$ to oscillate between $+\infty$ and $-\infty$. Christian Pommerenke [54] proved $\limsup = +\infty$ soon after Makarov's paper, although $\liminf = -\infty$ took another twenty years (see the beautiful paper of Sunhi Choi [38]). The Ahlfors distortion theorem implies

$$\omega_1(D)\omega_2(D) = O(r^2),$$

so for a disk where $\omega_1 \gg r$, we must have $\omega_2 \ll r$. Thus by Pommerenke's $\limsup = \infty$ result, ω_1 and ω_2 must be singular (written $\omega_1 \perp \omega_2$). on the non-tangent points. This gave me the second part of my thesis and a joint paper with Jones, Garnett and Carleson [26] (I still consider this paper a highlight of my career).

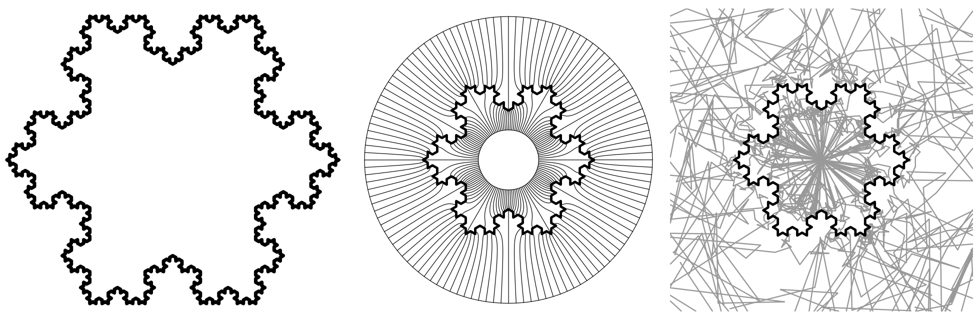


FIGURE 1: The von Koch snowflake (left) has singular harmonic measures which we visualize in two ways. In the center we plot the images of 120 radials lines under conformal maps to the inside and outside of a 196-sided approximation of the snowflake. On the right we simulate 100 Brownian paths per side by a discrete random walk that steps the distance to the boundary. Using 10000 such paths gives two 196-vectors whose normalized dot product is .0213 (so the vectors form an angle of 88.67°; almost perpendicular).

The final part of my thesis was an application of singular harmonic measures. If $f : \mathbb{C} \to \mathbb{D}$ is continuous and is holomorphic off a smooth curve γ, then it must be entire and hence constant (*i.e.*, smooth curves are removable). However, using an indirect duality argument, John Wermer and Andrew Browder [33, 34] had proven that if $\omega_1 \perp \omega_2$, then there are many such non-constant functions. Moreover, every non-trivial example is "space-filling", *i.e.*, it maps the curve to set that is the closure of its interior. Curious about what these functions looked like, I gave a new proof of the Browder–Wermer theorem using explicit constructions [9]; these methods later led to new results about function algebras [10, 16, 17], conformal welding [15, 21, 20] and Martin boundaries [12, 13].

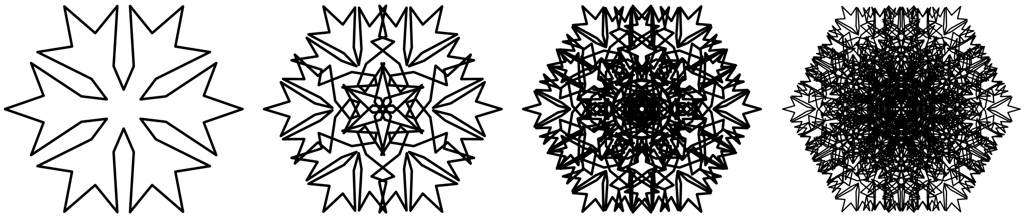


Figure 2: Polygons whose vertices are

$$v_k = 4^{-n} \sum_{j \neq k} (z_j - z_k)^{-1},$$

where $\{z_k\}$ are the vertices of the nth generation of von Koch snowflake (this approximates the convolution of $1/z$ with Hausdorff measure on the snowflake). These converge to a space-filling image of the snowflake, an example of the functions given by the Browder-Wermer theorem.

Next I looked for other existence proofs that lacked an explicit construction. Don Sarason [45] had indirectly proven there are infinite Blaschke products in the little Bloch space (*i.e.*, $|f'(z)| = o(1/1 - |z|)$ for $|z| < 1$) and had asked for an explicit example. I was able to build one [11, 14] using another idea from Makarov's paper: the radial behavior of harmonic functions on a disk is tied to the pointwise convergence of dyadic martingales on the boundary. To solve Sarason's problem, I constructed a martingale with certain smoothness properties on the unit circle and placed the zeros of the Blaschke product in the tops of Carleson boxes that corresponded to dyadic intervals where the martingale was zero.

As mentioned above, Makarov and Pommerenke proved that harmonic measure on the non-tangent points gives full mass to a set of zero length. What can we say about this set? Makarov proved it can't be too small (*i.e.*, dimension < 1 is impossible) and Bernt Øksendal conjectured that it must be big in the sense that it cannot be contained in any finite length curve. As a postdoc at MSRI and UCLA I thought a lot about this problem, but could only prove it in special cases (it's easy if Ω is a quasidisk). The difficulty is that most nice properties of a rectifiable curve γ only hold a.e.; how does a zero length subset of γ differ from a general zero length set? Fortunately, the answer became available right on schedule.

3. Traveling salesman and rectifiable sets

One summer I visited Peter Jones at Yale and he described his new "traveling salesman theorem" (TST) that estimates the length of the shortest path γ containing a given set E [47, 48]. For a disk $D = D(x,t)$, define "Jones' β-numbers"

$$\beta_E(x,t) = \inf_L \sup_{z \in E \cap D} \operatorname{dist}(z, L),$$

where the infimum is over all lines L hitting D. Peter proved that

$$\ell(\gamma) \simeq \operatorname{diam}(E) + \iint \beta(x,t)^2 \, \frac{dx\,dt}{t}.$$

His proof was simplified by Kate Okikiolu [53] who extended the result to $\mathbb{R}^d$ and was extended to Hilbert space by Raanan Schul [56, 57].

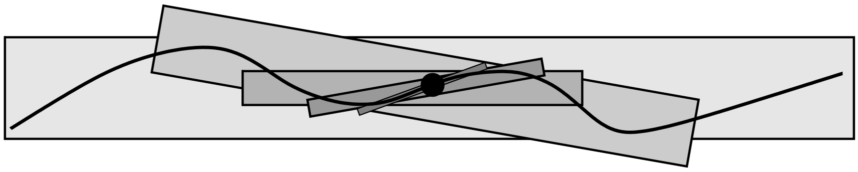


Figure 3: $\beta(x,t)$ measures the eccentricity of the narrowest rectangle covering $E \cap D(x,t)$. A curve is wiggly if $\beta > 0$ uniformly and x is a tangent point almost surely iff $\sum \beta(x, 2^{-n})^2 < \infty$.

If a set E lies on a rectifiable curve, Jones' TST gives concrete bounds for how "flat" E must be and we turned these into bounds on the Green's function for the complement of E, and eventually into a proof of Øksendal's conjecture [27] and a generalization of the Hayman-Wu theorem. We wrote a sequel [28] that simplified the proof, extended work of Astala and Zinsmeister [4, 5] on BMO domains and gave an a.e. characterization of tangent points of a curve in terms the β's. See the excellent discussion by Garnett and Marshall in [44].

The TST allowed us to use Littlewood–Paley type estimates, but in place of the usual second derivatives of a function, our estimates involved the β-numbers and Schwarzian derivatives (the usual second derivative measures deviation from a linear function, the β's measure deviation from a line and Schwarzians measure deviation from a linear fractional transformation). The basic idea in Jones' TST is that sets can be analyzed by quadratic sums just as functions can be, and this fact distinguishes Euclidean space in a way that I don't fully understand, but can illustrate with an example. At the 2005 Ahlfors–Bers colloquium in Ann Arbor, Juha Heinonen reminded me of the question of whether every A_1 weight on the plane is comparable to the Jacobian of some planar quasiconformal (QC) map. This problem is in his "33 Yes/No problems" paper [46] with Stephen Semmes, so I had seen it before, but I hadn't thought it was "up my alley". However, Juha's comments made me realize a counterexample would follow from a zero area set E with the property that every small-constant QC image of E contains a rectifiable curve. I constructed a Sierpinski carpet E where the holes are large enough to give zero area, but small enough (even after a QC mapping) so that we can construct rectifiable curves that avoid the holes [22] (the length is estimated using Jones' TST and the distribution of hole sizes). From this construction we can also obtain a quasisymmetric image of $\mathbb{R}^2$ in $\mathbb{R}^3$ that is not a biLipschitz image of $\mathbb{R}^2$. Hence characterizing Euclidean space up to biLipschitz equivalence is tied to understanding rectifiability and Jones' TST better.

Peter Jones and I also used his TST to prove "wiggly sets" have dimension > 1 [29] (a set is wiggly if it is connected and has β's uniformly bounded away from zero). This seems like an obvious result, but I still know no simpler proof than using the TST. Moreover, this basic result led to more subtle variations. A Brownian motion run for unit time defines a compact set in the plane and the complementary components are simply connected open sets, so their boundaries, called Brownian frontiers, are connected sets that look quite wiggly. Motivated by physical arguments, Benoit Mandelbrot had conjectured Brownian frontiers have dimension 4/3 and this was later proven using SLE type techniques by Lawler, Schramm and Werner [49, 50]. At the time, only infinite length was known, but Peter Jones, Robin Pemantle, Yuval Peres and I were able to prove Brownian frontiers have dimension > 1. The β's are not bounded away from zero, but they do have positive probability of being non-zero with enough independence between different locations and scales to prove the result (but not without a few tricks, *e.g.*, we used a fractal partition of the plane instead of the usual dyadic grid).

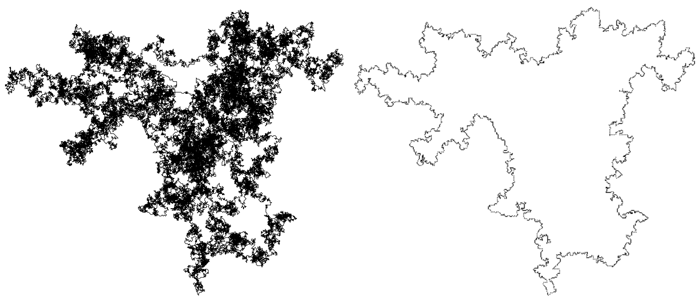


Figure 4: A Brownian path and its outer frontier. Jones' TST implied it has dimension > 1 and more recent work shows the exact dimension is 4/3, verifying a conjecture of Benoit Mandelbrot.

4. Kleinian groups and the convex hull theorem

By this time I had moved to Stony Brook and started to learn about Kleinian groups. Ed Taylor, a student of Bernie Maskit, asked me if the limit set of a finite generated, geometrically infinite Kleinian group has dimension > 1, a problem that seemed related to wiggly sets. Briefly, a Kleinian group G is a discrete group of Möbius transformations acting as isometries on hyperbolic 3-space $\mathbb{H}$ (identified with the upper half-space $\mathbb{R}^3_+$). The limit set $\Lambda \subset \partial\mathbb{H} = \mathbb{R}^2 \cup \{\infty\}$ is the accumulation set of any orbit and usually has a fractal structure. $\Omega = \partial\mathbb{H} \setminus \Lambda$ is open and we define the "dome" S_Ω of Ω as the upper envelope in $\mathbb{H}$ of all hemispheres with base disk in Ω. The region above the dome is the hyperbolic convex hull of Λ (assuming $\infty \in \Lambda$), and is denoted by $C(\Lambda)$. If G is finitely generated then the surface $S_\Omega = \partial C(\Lambda)$ has finite hyperbolic area mod G, but $C(\Lambda)/G$ itself may have either finite or infinite hyperbolic volume. These cases are called geometrically finite and infinite respectively. Ed Taylor's question was a weaker version of a well known conjecture that limit sets of geometrically infinite groups must have dimension 2. Like Brownian frontiers, Kleinian limit sets need not be uniformly wiggly, but in the finitely generated case there are only countably many points at which β tends to zero, so it was possible to prove $\dim(\Lambda) > 1$ using TST [29].

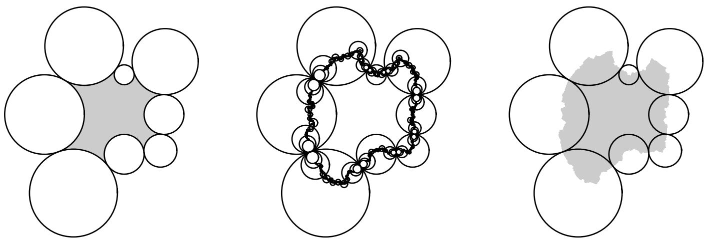


Figure 5: Here is a Kleinian limit set generated by circle reflections. The group is finitely generated and the limit set is connected but not a circle, so must have dimension > 1. However, the β's are zero where generating circles touch and along the orbits of such points.

Eventually, I was able to prove $\dim(\Lambda) = 2$ as well. The critical exponent δ of a Kleinian group measures the exponential rate growth of the G-orbits (there are at most $O(e^{\delta n})$ orbits points in any hyperbolic ball of radius n). Peter Jones and I showed that $\delta \le \dim(\Lambda)$ for any Kleinian group, so the conjecture reduces to the case when $\delta < 2$. Dennis Sullivan [60] had related δ to the base eigenvalues for the Laplacian on a hyperbolic manifold $M = \mathbb{R}^3_+/G$, and using this I showed that if $\delta < 2$ and G is geometrically infinite, then a Brownian motion started inside $C(\Lambda)$ has a positive probability of never crossing $\partial C(\Lambda)$. This implies Λ has positive area, hence $\dim(\Lambda) = 2$.

(In fact, area(Λ) > 0 is impossible for finitely generated groups by later work of Danny Calegari, David Gabai [35] and Ian Agol [1] proving the Ahlfors measure conjecture.)

This Brownian motion argument uses the fact mentioned earlier that S_Ω/G has finite hyperbolic area because we estimate the probability a Brownian motion crosses $S_\Omega = C(\Lambda)$ by integrating heat kernel bounds over S_Ω. Finite area is a consequence of two theorems: the Ahlfors finiteness theorem (hyperbolic area(Ω/G) $< \infty$) and Dennis Sullivan's convex hull theorem (CHT) [59]: Ω is biLipschitz equivalent to S_Ω with a universal bound K (so they have comparable areas). The CHT holds for any simply connected domain, as established by David Epstein and Al Marden [40, 41]. Peter Jones and I avoided quoting the CHT by using an alternate argument in the dim(Λ) = 2 paper, but it was not long before I needed to understand the CHT much better.

A Fuchsian group G is a Kleinian group that preserves the unit disk, $\mathbb{D}$. A deformation of G is a conformal map $f : \Omega \to \mathbb{D}$ that conjugates G to a Kleinian group $G' = f^{-1} \circ G \circ f$ acting on Ω. If the group is cocompact (*i.e.*, $R = \mathbb{D}/G$ is compact) then Rufus Bowen [32] proved $\partial\Omega$ is either a circle or has dimension > 1. This is "Bowen's dichotomy". Dennis Sullivan extended it to cofinite groups (R has finite area), and Kari Astala and Michel Zinsmeister [3, 6, 7] showed it fails whenever G is convergence type (R is a surface with a Green's function). This left open the case when R has infinite area but no Green's function (divergence type groups).

Thurston had observed that the hyperbolic path metric on the dome S_Ω is isometric to the hyperbolic unit disk (geometrically, the dome is just a hyperbolic disk that has been folded along certain geodesics). Composing Sullivan's map $\sigma : \Omega \to S_\Omega$ with this isometry gives a hyperbolically biLipschitz (hence QC) map from Ω to $\mathbb{D}$ with uniform constants. We call this the iota map. I observed (perhaps others had as well) that iota is locally Lipschitz $\Omega \to \mathbb{D}$ and deduced a factorization theorem: any conformal map $f : \Omega :\to \mathbb{D}$ is the composition of a locally Euclidean Lipschitz QC map $\varphi : \Omega \to \mathbb{D}$ and a hyperbolically biLipschitz map $\psi : \mathbb{D} \to \mathbb{D}$, both with uniform constants (assuming Ω has inradius ≥ 1).

Why does this help with Bowen's dichotomy? Suppose we have a non-circular deformation of a divergence type group G. We can show the β's for $\partial\Omega$ are large a.e. with respect to harmonic measure, but we need them large on positive length to get dim > 1. Since Makarov showed harmonic measure can be concentrated on a zero length set, the strategy seems to fail, but Sullivan's CHT saves the day. The factorization theorem implies that a conformal deformation of G via f is also a QC deformation of the divergence type group $G' = \psi \circ G \circ \psi^{-1}$ via the map φ (divergence type is a QC invariant by Pfluger, [55]). Moreover, $\varphi^{-1} : \mathbb{D} \to \Omega$ is locally expanding; this implies the β's are large on positive length, as desired [18].

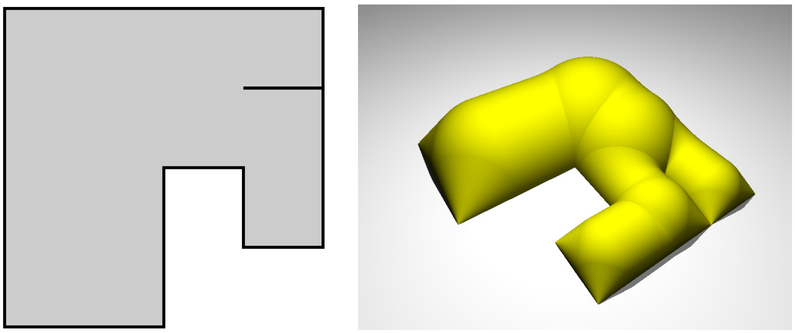


Figure 6: A polygon and its dome. The dome is the upper envelope of all hemispheres with base disk inside the polygon. The centers of hemispheres touching the dome form the medial axis of the polygon, a well studied object in computational geometry.

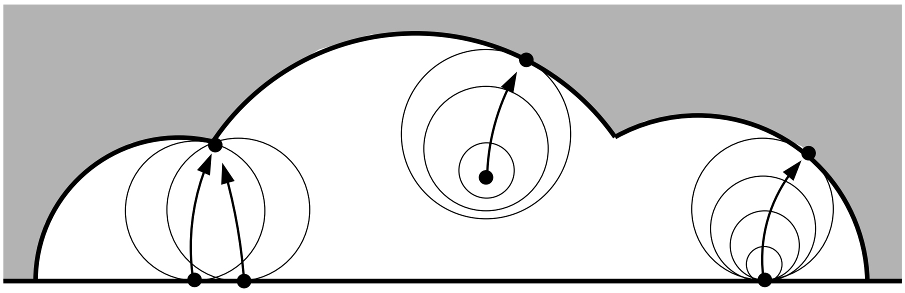


Figure 7: Here is a "side view" of a hyperbolic dome. The nearest point retraction is defined by expanding a hyperbolic ball until it hits the dome. For points in the base domain we expand a horoball instead. This defines a quasi-isometry, giving Sullivan's theorem. The map is not necessarily a homeomorphism since two distinct points can map to the same point.

The factorization theorem implies that if $f : \Omega \to \mathbb{D}$ is conformal, then $|f'| = |\varphi'| \cdot |\psi'|$ where $\varphi' \in L^\infty$ and ψ' is in weak-L^p for $p = 2K/(K-1)$ by a celebrated result of Kari Astala [2]. This is reminiscent of Brennan's conjecture that $f' \in L^{4-\epsilon}(dxdy, \Omega)$ for any $\epsilon > 0$. In fact, if Sullivan's theorem held with constant $K = 2$, Brennan's conjecture would follow. This motivated me to try to give the best explicit constant I could. In [19] I proved $K < 7.82$ by carefully examining the Epstein-Marden proof of CHT in [40]. Unfortunately, Epstein and Markovic found a logarithmic spiral domain for which $K > 2.1$, [42]. It is still possible that every simply connected domain has a 2-QC, locally Lipschitz map to the disk (this would imply Brennan's conjecture), but iota itself doesn't always work.

5. Computational conformal geometry

For polygons, the Riemann map is given by the Schwarz–Christoffel formula, but this involves unknown parameters, namely, the points on the circle that get mapped to the polygon's vertices. Solving for these can be quite difficult. On the other hand, the iota map can be applied to every vertex of an n-gon in time $O(n)$. This depends on the close relation between the dome of a domain and its medial axis. The medial axis is a term from computer science [31] that refers to the centers of subdisks of Ω whose boundaries hit $\partial\Omega$ in ≥ 2 points (Erdös [43] called the same set M_2 twenty years earlier). The medial axis of an n-gon can be computed in time $O(n)$ by a result of Chin, Snoeyink and Wang [36, 37] and iota can be computed in linear time from the medial axis. Thus iota gives a "fast" map to the disk that is uniformly close to conformal by the CHT. Dennis Sullivan told me he originally thought of the CHT as a "constructive version of the Riemann mapping theorem".

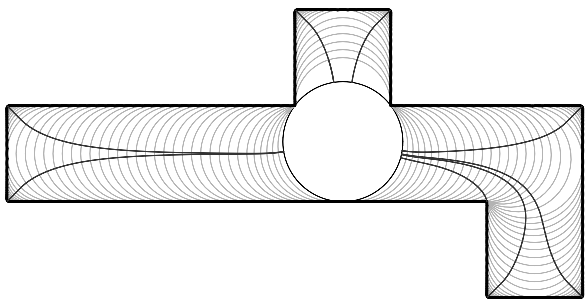


Figure 8: A polygon is foliated by arcs of medial axis disks; the orthogonal flow gives the iota map from the polygon to a circle (it can also computed algebraically in time $O(n)$ for any n-gon).

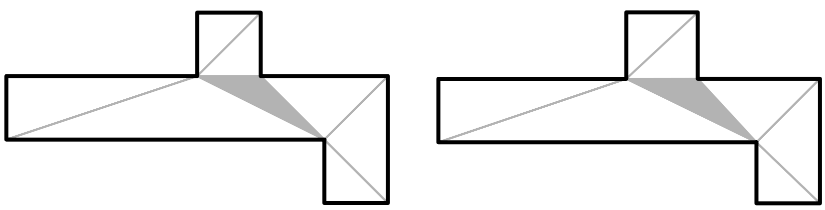


Figure 9: A polygon and the Schwarz–Christoffel image using the correct angles but pre-vertices guessed using iota. By the CHT there is a K-QC, vertex preserving map between them. We can get an upper bound for K by triangulating both polygons and computing the maximum dilatation of the corresponding piecewise linear map; $|\mu| \leq .108$ is this case. The most distorted triangle is shaded.

In fact, the iota map was also discovered in numerical analysis, but under a different name. While trying to numerically compute the best K in Sullivan's theorem, I came across the paper [39] by Toby Driscoll and Steve Vavasis. It describes their CRDT algorithm, a numerical conformal mapping method that uses a map from simple n-gons to n-tuples on the circle defined in terms of cross ratios and the Delaunay triangulation of the polygon (hence the name) and while reading this paper (for the fourth or fifth time), I realized the CRDT map was a version of iota and I was able to prove the same uniform QC bounds as for the "real" iota [23]. (Actually, Vavasis had sent me a preprint of the CRDT paper a few years earlier, but I hadn't appreciated it without knowing about CHT and iota, and had forgotten about it. My "rediscovery" of their paper led to a workshop, a joint grant with Vavasis and several results about domain decomposition and conformal maps.)

CRDT uses the Schwarz–Christoffel formula, so each iteration gives a true conformal map onto an approximate domain. The algorithm tries to improve this domain at each step, but the dependence on the parameters is so subtle that no proof of convergence is known (at least to me). Failing to prove CRDT converges, I tried the opposite approach: approximately conformal maps onto the true domain, i.e, consider QC maps $\mathbb{D} \to \Omega$ and solve a Beltrami problem to lower the QC constant at each step. I showed this method can compute a $(1+\epsilon)$-QC map from the disk onto any n-gon in time $O(n \log \frac{1}{\epsilon} \log\log \frac{1}{\epsilon})$ [24]. The maps are held in memory using $O(n)$ series, each of length $p = \log \frac{1}{\epsilon}$. The iteration has quadratic convergence, so using iota as a starting point, Sullivan's CHT implies only $O(\log\log \frac{1}{\epsilon})$ iterations are needed to reach accuracy ϵ, independent of the domain.

The CHT is used in other parts of this algorithm as well. A key ingredient is the idea of a thick/thin decomposition of a polygon analogous to the thick/thin decomposition of a Riemann surface. Thin parts of a polygon are certain generalized quadrilaterals with a pair of sides whose extremal distance inside the polygon is less than ϵ. Decomposing a polygon into its thick and thin parts makes various mapping and meshing problems easier to understand. The iota map and Sullivan's CHT allow us to compute extremal distances (up to a bounded factor) in linear time and this leads to a linear time algorithm to find all the thin parts.

Marshall Bern and David Eppstein (two "p"'s this time; not the same David Epstein mentioned before) had proven in [8] that any simply n-gon has a quadrilateral mesh with $O(n)$ elements and no angle bigger than $120°$ (and this is sharp). They asked if a lower angle bound was possible, and using thick/thin decompositions and the mapping theorem above, I showed [25] we could also take all new angles $\geq 60°$ (small angles in the original polygon must remain).

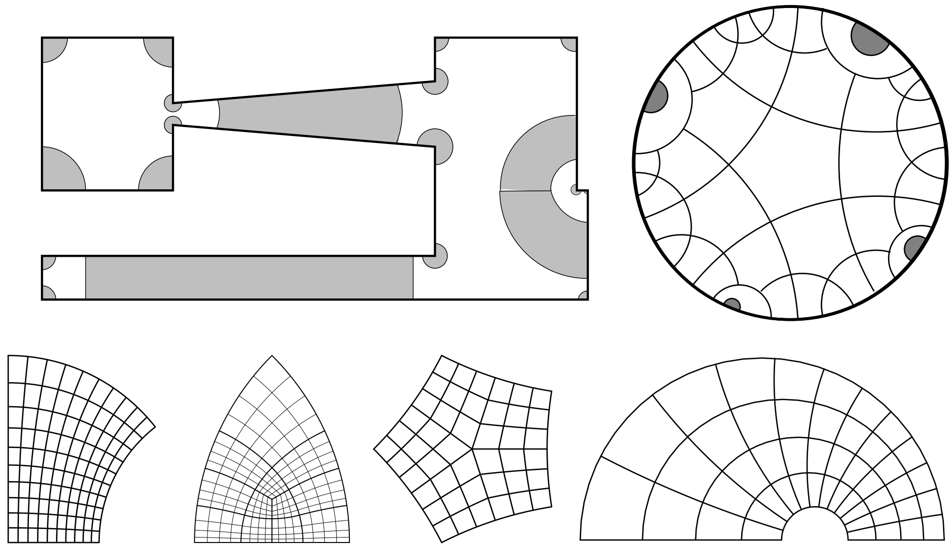


Figure 10: Meshing with optimal angle and complexity bounds. We decompose a polygon into thick and thin pieces (white and shaded). Thin parts are meshed "by hand" and the Riemann map of the polygon sends each thick part to a region in $\mathbb{D}$ that can be subdivided into four types of hyperbolic polygons as shown. Each type has a mesh with angles in $[60°, 120°]$ that we transfer back to the thick part by the conformal map.

This quickly leads to a more general problem. A planar straight line graph (PSLG) is any finite, disjoint collection of line segments and points (polygons are a special case where the edges meet end-to-end). A mesh of a PSLG is a mesh of its convex hull whose vertices and edges covers all the vertices and edges of the PSLG. I was able to show that any PSLG has a quadrilateral mesh with $O(n^2)$ elements and the same angle bounds as above (and n^2 is sharp). By adding diagonals to the quadrilaterals, we get a $O(n^2)$ triangulation of any PSLG with all angles $\leq 120°$, improving the bound 157.5° by Scott Mitchell [52] and 132° by Tiow-Seng Tan [61]. In fact, 120° can be replaced by any bound $> 90°$ (but the constant in $O(n^2)$ grows) and there is even a polynomial algorithm for nonobtusely triangulating a PSLG (all angles $\leq 90°$).

The proof uses thick/thin decompositions and a foliation of the thin parts similar to those used by Epstein and Marden in their proof of CHT. In each thin part, the leaves are just circular arcs, but when joined together the leaves can become quite complicated. If every path hits at most $O(n)$ thin parts, we get an $O(n^2)$ nonobtuse triangulation. In general, I show that by adding $O(n^{1.5})$ extra paths and bending the original paths slightly we can cause collisions which terminate every path after crossing at most $O(n)$ thin parts; this gives an $O(n^{2.5})$ triangulation. The best lower bound is $O(n^2)$, so a gap remains open (I either need to understand CHT a bit better or it is time for another serendipitous result to appear; Dennis Sullivan suggested looking at closing lemmas in dynamics).

That's the story so far: Nick Makarov's paper helped me write my thesis and led to various problems including Øksendal's conjecture; Peter Jones' traveling salesman theorem was the key to solving that conjecture and involved me with Brownian motion, geometric measure theory and Kleinian groups; the $\dim(\Lambda) = 2$ problem for Kleinian limit sets led me to Dennis Sullivan's convex hull theorem, which then solved Bowen's dichotomy and pushed me towards new results in numerical conformal mappings and computational geometry. Each paper was first useful because it contained a fact I needed, but their real value lay in the new problems they inspired.

Postscript

This essay is an edited version of an even more rambling previous attempt. Trying to compress it further, I projected into a lower dimension rhyming space:

Under logs are measures walking
along paths with betas stalking
over domes of bounded bending
questions answered and unending.

Projecting into the even lower dimensional haiku space gives

Flatness abandoned
deep origamic thunder
echos off my pen.

This may be useful if NSF proposal limits ever drop from 15 pages to 17 syllables.

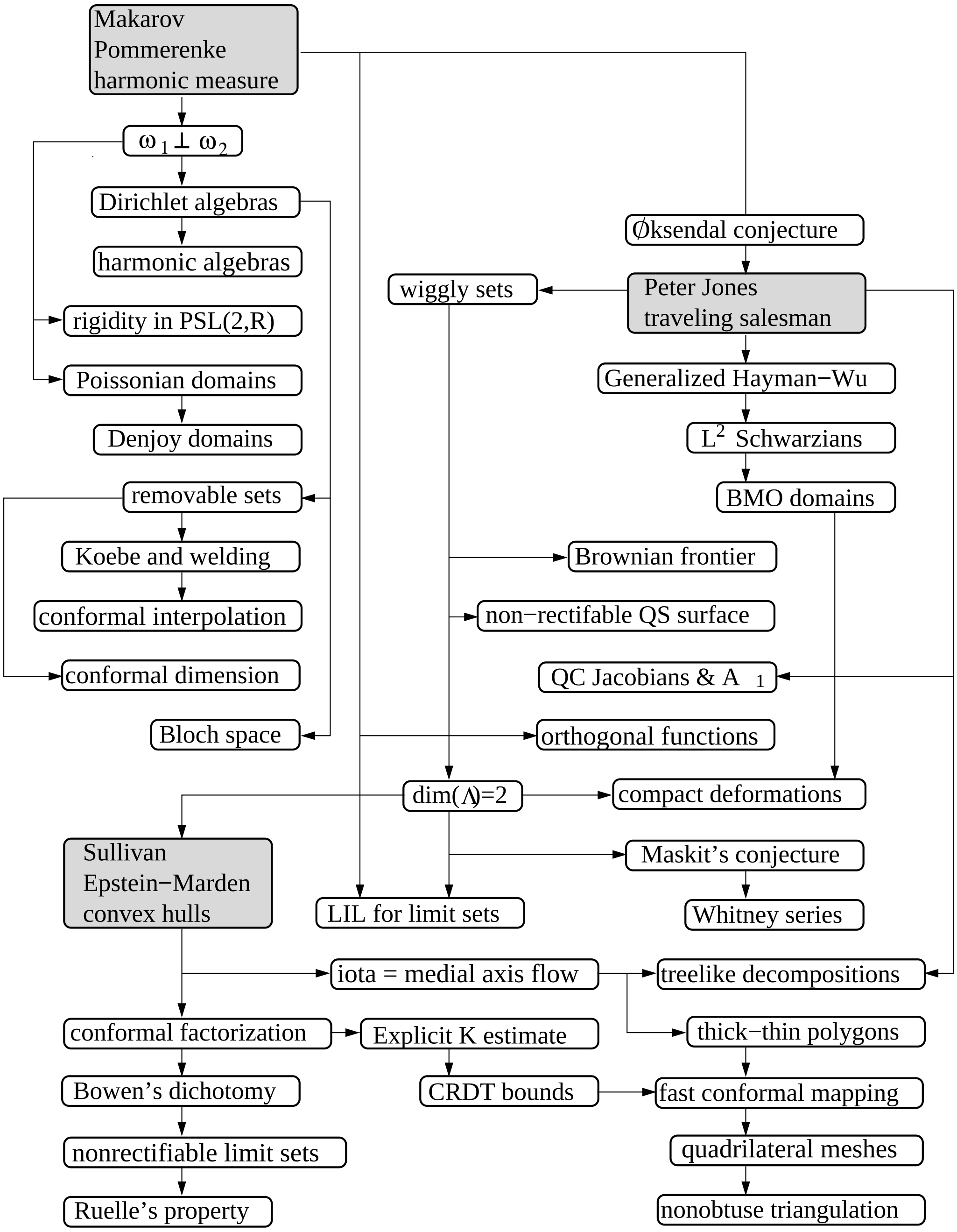


FIGURE 11: Some of my work as a directed graph.

References

[1] Agol, I.: Tameness of hyperbolic 3-manifolds. Preprint, 2004. Available at arXiv:math/0405568v1.

[2] Astala, K.: Area distortion of quasiconformal mappings. *Acta Math.* **173** (1994), no. 1, 37–60.

[3] Astala, K. and Zinsmeister, M.: Mostow rigidity and Fuchsian groups. *C. R. Acad. Sci. Paris Sér. I Math.* **311** (1990), no. 6, 301–306.

[4] Astala, K. and Zinsmeister, M.: Teichmüller spaces and BMOA. *Math. Ann.* **289** (1991), no. 4, 613–625.

[5] Astala, K. and Zinsmeister, M.: Rectifiability in Teichmüller theory. In *Topics in complex analysis (Warsaw, 1992)*, 45–52. Banach Center Publ. **31**. Polish Acad. Sci., Warsaw, 1995.

[6] Astala, K. and Zinsmeister, M.: Holomorphic families of quasi-Fuchsian groups. *Ergodic Theory Dynam. Systems* **14** (1994), no. 2, 207–212.

[7] Astala, K. and Zinsmeister, M.: Abelian coverings, Poincaré exponent of convergence and holomorphic deformations. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20** (1995), no. 1, 81–86.

[8] Bern, M. and Eppstein, D.: Quadrilateral meshing by circle packing. In "Selected papers from the Sixth International Meshing Roundtable, Part II (Park City, UT, 1997)". *Internat. J. Comput. Geom. Appl.* **10** (2000), no. 4, 347–360.

[9] Bishop, C. J.: Constructing continuous functions holomorphic off a curve. *J. Funct. Anal.* **82** (1989), no. 1, 113–137.

[10] Bishop, C. J.: Approximating continuous functions by holomorphic and harmonic functions. *Trans. Amer. Math. Soc.* **311** (1989), no. 2, 781–811.

[11] Bishop, C. J.: Bounded functions in the little Bloch space. *Pacific J. Math.* **142** (1990), no. 2, 209–225.

[12] Bishop, C. J.: A characterization of Poissonian domains. *Ark. Mat.* **29** (1991), no. 1, 1–24.

[13] Bishop, C. J.: Brownian motion in Denjoy domains. *Ann. Probab.* **20** (1992), no. 2, 631–651.

[14] Bishop, C. J.: An indestructible Blaschke product in the little Bloch space. *Publ. Mat.* **37** (1993), no. 1, 95–109.

[15] Bishop, C. J.: Some homeomorphisms of the sphere conformal off a curve. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **19** (1994), no. 2, 323–338.

[16] Bishop, C. J.: Some characterizations of $C(M)$. *Proc. Amer. Math. Soc.* **124** (1996), no. 9, 2695–2701.

[17] Bishop, C. J.: A distance formula for algebras on the disk. *Pacific J. Math.* **174** (1996), no. 1, 1–27.

[18] Bishop, C. J.: Divergence groups have the Bowen property. *Ann. of Math. (2)* **154** (2001), no. 1, 205–217.

[19] Bishop, C. J.: An explicit constant for Sullivan's convex hull theorem. In *In the tradition of Ahlfors-Bers. III*, 41–69. Contemp. Math. **355**. Amer. Math. Soc., Providence, RI, 2004.

[20] Bishop, C. J.: Boundary interpolation sets for conformal maps. *Bull. London Math. Soc.* **38** (2006), no. 4, 607–616.

[21] Bishop, C. J.: Conformal welding and Koebe's theorem. *Ann. of Math. (2)* **166** (2007), no. 3, 613–656.

[22] Bishop, C. J.: An A_1 weight not comparable with any quasiconformal Jacobian. In *In the tradition of Ahlfors-Bers. IV*, 7–18. Contemp. Math. **432**. Amer. Math. Soc., Providence, RI, 2007.

[23] Bishop, C. J.: Bounds for the CRDT conformal mapping algorithm. *Comput. Methods Funct. Theory* **10** (2010), no. 1, 325–366.

[24] Bishop, C. J.: Conformal mapping in linear time. *Discrete Comput. Geom.* **44** (2010), no. 2, 330–428.

[25] Bishop, C. J.: Optimal angle bounds for quadrilateral meshes. *Discrete Comput. Geom.* **44** (2010), no. 2, 308–329.

[26] Bishop, C. J., Carleson, L., Garnett, J. B. and Jones, P. W.: Harmonic measures supported on curves. *Pacific J. Math.* **138** (1989), no. 2, 233–236.

[27] Bishop, C. J. and Jones, P. W.: Harmonic measure and arclength. *Ann. of Math. (2)* **132** (1990), no. 3, 511–547.

[28] Bishop, C. J. and Jones, P. W.: Harmonic measure, L^2 estimates and the Schwarzian derivative. *J. Anal. Math.* **62** (1994), 77–113.

[29] Bishop, C. J. and Jones, P. W.: Wiggly sets and limit sets. *Ark. Mat.* **35** (1997), no. 2, 201–224.

[30] Bishop, C. and Steger, T.: Representation-theoretic rigidity in $\mathrm{PSL}(2,\mathbb{R})$. *Acta Math.* **170** (1993), no. 1, 121–149.

[31] Blum, H.: A transformation for extracting new descriptors of shape. In *Proc. Symp. Models for the perception of speech and visual form.* MIT Press, Cambridge, 1967.

[32] Bowen, R.: Hausdorff dimension of quasicircles. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 11–25.

[33] Browder, A. and Wermer, J.: Some algebras of functions on an arc. *J. Math. Mech.* **12** (1963), 119–130.

[34] Browder, A. and Wermer, J.: A method for constructing Dirichlet algebras. *Proc. Amer. Math. Soc.* **15** (1964), 546–552.

[35] Calegari, D. and Gabai, D.: Shrinkwrapping and the taming of hyperbolic 3-manifolds. *J. Amer. Math. Soc.* **19** (2006), no. 2, 385–446 (electronic).

[36] Chin, F., Snoeyink, J. and Wang, C. A.: Finding the medial axis of a simple polygon in linear time. In *Algorithms and computations (Cairns, 1995)*, 382–391. Lecture Notes in Comput. Sci. **1004**. Springer, Berlin, 1995.

[37] Chin, F., Snoeyink, J. and Wang, C. A.: Finding the medial axis of a simple polygon in linear time. *Discrete Comput. Geom.* **21** (1999), no. 3, 405–420.

[38] Choi, S.: The lower density conjecture for harmonic measure. *J. Anal. Math.* **93** (2004), 237–269.

[39] Driscoll, T. A. and Vavasis, S. A.: Numerical conformal mapping using cross-ratios and Delaunay triangulation. *SIAM J. Sci. Comput.* **19** (1998), no. 6, 1783–1803.

[40] Epstein, D. B. A. and Marden, A.: Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces. In *Analytical and geometric aspects of hyperbolic space (Coventry/Durham, 1984)*, 113–253. London Math. Soc. Lecture Note Ser. **111**. Cambridge University Press, Cambridge, 1987.

[41] Epstein, D. B. A. and Marden, A.: Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces. In *Fundamentals of hyperbolic geometry: selected expositions*, 117–266. London Math. Soc. Lecture Note Ser. **328**. Cambridge Univ. Press, Cambridge, 2006.

[42] Epstein, D. B. A. and Markovic, V.: The logarithmic spiral: a counterexample to the $K = 2$ conjecture. *Ann. of Math. (2)* **161** (2005), no. 2, 925–957.

[43] Erdös, P.: On the Hausdorff dimension of some sets in Euclidean space. *Bull. Amer. Math. Soc.* **52** (1946), 107–109.

[44] Garnett, J. B. and Marshall, D. E.: *Harmonic measure*. New Mathematical Monographs **2**. Cambridge University Press, Cambridge, 2008.

[45] Havin, V. P., Hruščëv, S. V. and Nikol'skiĭ, N. K. (editors): *Linear and complex analysis problem book*. Lecture Notes in Mathematics **1043**. Springer-Verlag, Berlin, 1984.

[46] Heinonen, J. and Semmes, S.: Thirty-three yes or no questions about mappings, measures, and metrics. *Conform. Geom. Dyn.* **1** (1997), 1–12 (electronic).

[47] Jones, P. W.: Rectifiable sets and the traveling salesman problem. *Invent. Math.* **102** (1990), no. 1, 1–15.

[48] Jones, P. W.: The traveling salesman problem and harmonic analysis. In "Conference on Mathematical Analysis (El Escorial, 1989)". *Publ. Mat.* **35** (1991), no. 1, 259–267.

[49] Lawler, G. F., Schram, O. and Werner, W.: Values of Brownian intersection exponents. I. Half-plane exponents. *Acta Math.* **187** (2001), no. 2, 237–273.

[50] Lawler, G. F., Schram, O. and Werner, W.: The dimension of the planar Brownian frontier is 4/3. *Math. Res. Lett.* **8** (2001), no. 4, 401–411.

[51] Makarov, N. G.: On the distortion of boundary sets under conformal mappings. *Proc. London Math. Soc. (3)* **51** (1985), no. 2, 369–384.

[52] Mitchell, S. A.: Refining a triangulation of a planar straight-line graph to eliminate large angles. In *34th Annual Symposium on Foundations of Computer Science (Palo Alto, CA, 1993)*, 583–591. IEEE Comput. Soc. Press, Los Alamitos, CA, 1993.

[53] Okikiolu, K.: Characterization of subsets of rectifiable curves in $\mathbb{R}^n$. *J. London Math. Soc. (2)* **46** (1992), no. 2, 336–348.

[54] Pommerenke, C.: On conformal mapping and linear measure. *J. Analyse Math.* **46** (1986), 231–238.

[55] Pfluger, A.: Sur une propriété de l'application quasi conforme d'une surface de Riemann ouverte. *C. R. Acad. Sci. Paris* **227** (1948), 25–26.

[56] Schul, R.: Analyst's traveling salesman theorems. A survey. In *In the tradition of Ahlfors-Bers. IV*, 209–220. Contemp. Math. **432**. Amer. Math. Soc., Providence, RI, 2007.

[57] Schul, R.: Subsets of rectifiable curves in Hilbert space —the analyst's TSP. *J. Anal. Math.* **103** (2007), 331–375.

[58] Semmes, S.: A counterexample in conformal welding concerning chord-arc curves. *Ark. Mat.* **24** (1986), no. 1, 141–158.

[59] Sullivan, D.: Travaux de Thurston sur les groupes quasi-fuchsiens et les variétés hyperboliques de dimension 3 fibrées sur S^1. In *Bourbaki Seminar, Vol. 1979/80*, 196–214. Lecture Notes in Math. **842**. Springer, Berlin, 1981.

[60] Sullivan, D.: Related aspects of positivity in Riemannian geometry. *J. Differential Geom.* **25** (1987), no. 3, 327–351.

[61] Tan, T.-S.: An optimal bound for high-quality conforming triangulations. *Discrete Comput. Geom.* **15** (1996), no. 2, 169–193.

ALINE BONAMI (Randan, France, 1944) obtained her PhD in mathematics at Orsay (advisor: Yves Meyer). She is Emeritus Professor at Orléans University. Her main research areas are real and complex analysis and applications of Fourier analysis to signal processing, probability and statistics.

The Uncertainty Principle for Fourier analysts

Aline Bonami*

Which paper? Why?

There are certainly many kinds of mathematicians. Some dig the same furrow to infinity, some others pick right or left the little flowers they find during an endless promenade, while probably most of them move from one of these attitudes to the other. The paper that I will speak of, *The Uncertainty Principle: A Mathematical Survey*, by Gerald B. Folland and Alladi Sitaram [9], definitely invites the reader to a kind of walk around the uncertainty principle, with the possibility of picking right and left armfuls of flowers.

The Journal of Fourier Analysis and Applications
Volume 3, Number 3, 1997

Research Tutorial

The Uncertainty Principle: A Mathematical Survey

Gerald B. Folland and Alladi Sitaram

ABSTRACT. We survey various mathematical aspects of the uncertainty principle, including Heisenberg's inequality and its variants, local uncertainty inequalities, logarithmic uncertainty inequalities, results relating to Wigner distributions, qualitative uncertainty principles, theorems on approximate concentration, and decompositions of phase space.

Introduction

The uncertainty principle is partly a description of a characteristic feature of quantum mechanical systems, partly a statement about the limitations of one's ability to perform measurements on a system without disturbing it, and partly a meta-theorem in harmonic analysis that can be summed up as follows.

A nonzero function and its Fourier transform cannot both be sharply localized. (0.1)

When translated into the language of quantum mechanics, as we shall do in §2, (0.1) says that the values of a pair of canonically conjugate observables such as position and momentum cannot both be precisely determined in any quantum state. Therefore, it leads to mathematical formulations of the physical ideas first developed in Heisenberg's seminal paper [51] of 1927 and widely promulgated thereafter.

However, the uncertainty principle also has a useful interpretation in classical physics. Namely, if $f(t)$ represents the amplitude of a signal (a sound wave or light wave, perhaps) at time t, the Fourier transform $\widehat{f}$ tells how f is built from sine waves of different frequencies and (0.1) expresses a limitation on the extent to which a signal can be both time-limited and band-limited. This aspect of the uncertainty principle was already expounded by Norbert Wiener in a lecture in Göttingen in 1925. Unfortunately, no written record of this lecture seems to have survived, apart from the nontechnical account in Wiener's autobiography [119, pp. 105–107], so one can only guess at what precise versions of (0.1) it might have contained. Whatever influence this lecture might have had on the physicists in the audience, however, the uncertainty principle did not really sink into the minds of signal analysts until Gabor's fundamental work [40] in 1946. Since then, it has become firmly embedded in the common culture.

On the mathematical side, there were sporadic developments relating to the uncertainty principle in the fifty years after the initial work in the 1920's, followed by a steady stream of results in

Math Subject Classifications. Primary 42-02, 42B10, 26D15; Secondary 43A25, 43A30, 81Q10, 81S30, 94A12, 94A17.
Keywords and Phrases. uncertainty principle, Fourier transform, Heisenberg's inequality, logarithmic inequalities, Wigner distribution, Landau–Pollak–Slepian theory, phase space.

©1997 CRC Press, Inc.
ISSN 1069-5869

This choice may be considered as not serious, and the serious (and slightly critical) reader will ask questions.

- Why did I choose a survey paper, not an original one? Such a paper is always somewhat superficial. In this particular case, for the reader who wants to study the Uncertainty Principle, the book of Havin and Jöricke *The Uncertainty Principle in Harmonic Analysis* [11] is certainly a much richer reference[1].
- The Uncertainty Principle is, in some ways, a fashion. It is certainly fundamental in physics, but this term, when used in Harmonic Analysis, not only refers to new concepts or new methods but also covers many well-known properties. Why not some deeper paper?

These are legitimate questions, and there are of course *deeply original* papers that influenced me. But this is the precise paper that came naturally to my mind as an answer to the question raised by the Editors of the Revista

*Aline.Bonami@univ-orleans.fr.

[1] I read later on the survey paper of Havin [10], which is also an excellent introduction to the subject.

Matemática Iberoamericana. Reasons of this are very subjective. I already gave one: I like promenades in mathematics as well as in other areas of activity. A second one is also related to my tastes in mathematics. Again one can observe two categories of mathematicians: those who spend their time reading papers and those who like most listening to talks. I belong definitely to this last category[2]. When I read the paper of Folland and Sitaram in 1999, I felt the same pleasure that an inspiring lecture usually gives me. There were fascinating open questions, some of which have agitated me during the last twelve years. I did recognize the kind of questions that were discussed in Orsay in the seventies but they were approached from a different point of view and I understood them better. The paper itself does not contain tricky proofs, but looking a little further into the subject there are plenty of them. It offers new perspectives and gives the desire to go further, which is the best compliment one can give to a survey paper.

I cannot help adding some general comment on mathematical journals. Even if there exists top level journals that contribute to select the best papers, one has to deal with two characteristics of the present time: the fact that there are an increasing number of institutions in the world where some good mathematics is done[3] on the one hand, and on the other, the huge amount of documentation available on the web, with the permanent launch of new journals and with arXiv as a daily instrument. In order to help young mathematicians to make good choices within the literature that is available, survey papers should play an increasing role.

The Uncertainty Principle as a fairy tale

The impossibility of measuring precisely both the position and the velocity of a sub-atomic particle sounds like a mystery to most mathematicians[4]. They rapidly need to go back to some clear mathematical setting. Folland and Sitaram start from Quantum Mechanics to explain the Uncertainty Principle but refer hastily to a lecture of Norbert Wiener in Göttingen in 1925, which may be taken as the first occurrence of the Uncertainty Principle in

[2] Having been a young mathematician in Orsay in the late sixties and early seventies has played a large role in this. My whole work has been primarily influenced by the course of E. Stein on *Singular Integrals* in 1966-67, which has been transformed into the book *Singular integrals and differentiability properties of functions*.

[3] It is no more possible to hope that most young mathematicians will benefit from word of mouth, as it was for me at Orsay in the sixties.

[4] To quote Heisenberg: *Any use of the words "position" and "velocity" with an accuracy exceeding that given by the uncertainty equation is just as meaningless as the use of words whose sense is not defined* (from Physical Principles of the Quantum Theory, 1930).

Signal Analysis. Recall that, as Folland and Sitaram write, the Uncertainty Principle may be seen as the fact that

> *a nonzero function and its Fourier transform cannot both be sharply localized,*

which is not a mathematical statement and deserves to be translated into rigorous statements. This is the moral of the story but we need characters and intrigue.

When speaking of this lecture in his autobiography *I am a Mathematician*, Norbert Wiener starts his explanation from the musical notation, which gives the false impression that one can independently fix the pitch of a sound (Do, Re, ..., or C, D, ...) and its length (whole note, half note, quarter note, ...). On the contrary if we quote Norbert Wiener, *Precision in time means a certain vagueness in pitch, just as precision in pitch an indifference to time. (...) You can't play a jig on the lowest register of the organ.*

Even if this looks more familiar than Quantum Mechanics, it holds some mystery. The importance of the Uncertainty Principle in Signal Analysis was only plainly felt twenty years later, even though Hardy says in 1933 that he was directly influenced by lectures of Wiener.

Historical aspects are sufficiently present in the paper of Folland and Sitaram to arouse the curiosity of the reader without any risk of boring him with long considerations. I would like to quote two stories that are told in a few words. One concerns Benedicks' Theorem [2] (and I quote Folland and Sitaram) *whose elegant proof, first circulated as a preprint in 1974 but not formally published for another decade.* The other one concerns Beurling's Theorem, or Beurling–Hörmander's Theorem depending on the authors, whose proof was lost even though it is stated in the collected works of Beurling, then published in 1991 by Hörmander [12]. These few words were sufficient for me to feel an irresistible urge to look at proofs and papers. Let me recall the two theorems.

Theorem (Benedicks) *If a function $f \in L^2(\mathbb{R}^n)$ and its Fourier transform are supported on measurable sets of finite measure, then f vanishes a.e.*

Theorem (Beurling) *If the function $f \in L^1(\mathbb{R})$ satisfies the inequality*

$$\int_{\mathbb{R}} |f(x)|\, |\widehat{f}(y)|\, e^{2\pi|x||y|}\, dx\, dy < \infty, \tag{1}$$

then f vanishes a. e.

Here the Fourier transform in $\mathbb{R}^n$ is defined by

$$\widehat{f}(y) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i\, x\cdot y}\, dx.$$

These are two elegant statements, with very nice proofs. On the other hand, they may seem incomplete and one does not have any definitive conjecture to describe the ultimate point up to which methods could be pushed forward. This may be why their authors hesitated to publish these results. In fact this is the charm of the Uncertainty Principle, to be broken down into many mathematical statements that cannot be easily compared between them. It is the great merit of the paper of Folland and Sitaram to transform separate results into a coherent whole.

More to read, more to do

It is very difficult to read a mathematical paper extensively. We all have our own way of picking this or that, then looking at the bibliography and jumping to another paper, then coming back to read again and again the same lines, as if there was some secret clue to find. The selection of the theorems of Benedicks and Beurling corresponds to my own selection and they are those that have piqued my curiosity the most.

For the first one, it is easy to see that the proof, which is given by Folland and Sitaram and relies on Poisson Formula, can be adapted to some sets of infinite measure, such as the interior of an hyperbola in $\mathbb{R}^2$. In general one does not know at all for which kind of sets this type of property is valid. In the case of two measurable sets of finite measure A and B, this implies easily (see [3] for these comments) the existence of some constant c such that, for all $f \in L^2(\mathbb{R}^n)$, one has the inequality

$$\int_{\mathbb{R}^n} |f|^2 dx \leq c \left(\int_{\mathbb{R}^n \setminus A} |f|^2 dx + \int_{\mathbb{R}^n \setminus B} |\widehat{f}\,|^2 dy \right). \tag{2}$$

This last inequality has also been proved through another method by Amrein and Berthier [1]. In one dimension, there is a sharp estimate of the constant given in a beautiful paper of Nazarov [16], which has not yet been entirely exploited (see [13] for a partial generalization in higher dimension). At this point, it is necessary to read carefully the book of Havin and Joricke.

As I said before, there is no clear conjecture that one can deduce from Benedicks' Theorem. Instead, there are a lot of possible directions, some of them having already partially explored, while others have not. One of the most fascinating recent approaches is due to T. Wolff (his former student Kovrijkine [14] speaks of Wolff's version of the Uncertainty Principle, even though the corresponding theorem may be found in a joint work [17], where it is used to prove the existence of a spectral gap). Wolff's Principle asserts that (2) holds for some constant c when ε is small enough, and A and B are

ε-thin in the following sense: A is called ε-thin if, for each $x_0 \in \mathbb{R}$ and for $\delta(x_0) = \max(1, |x_0|^{-1})$, one has

$$\int_{|x-x_0|<\delta(x_0)} \chi_A(x)\, dx \leq \varepsilon\, \delta(x_0)^n.$$

The typical example of such a situation is given by $A = \{|q| < \eta\}$, where q is a non degenerate quadratic form. Such inequalities give rise to Heisenberg type inequalities. They have been considered later on by Demange [7] and Kovrijkine [15]. It is striking that Wolff's methods come from real analysis and use a Littlewood–Paley decomposition of functions.

By contrast, the proof of Beurling's Theorem given by Hörmander relies on complex analysis and mainly on the theorem of Phragmén–Lindelöf as does the classical proof of Hardy's Theorem:

Theorem (Hardy) *If the function f satisfies the two inequalities*

$$|f(x)| \leq Ce^{-\pi|x|^2}, \qquad |\widehat{f}(y)| \leq Ce^{-\pi|y|^2}, \tag{3}$$

then f is, up to a constant, equal to the function $e^{-\pi|x|^2}$.

Complex methods have been pursued in [4], then by Demange in [6], where the use of Bargmann transforms allows to have Hardy type theorems, but with a non degenerate bilinear form replacing the Euclidean scalar product, under some supplementary assumptions.

Many other results concern Hardy type or Beurling type theorems in different contexts and especially in a non commutative setting, see in particular the book of Thangavelu [19].

The general feeling was that complex methods were unavoidable for Hardy type theorems. The development of real methods by Escauriaza, Kenig, Ponce and Vega appears miraculous (see [8]). It is a real *tour de force* to have achieved (jointly with Cowling) a new proof of the classical Hardy's Theorem [5] as a refinement of these methods. These may also be used for unique continuation properties of Schrödinger Equation with potential, while complex methods seem to be inadequate for this.

All this arises out of a partial and personal reading of the paper of Folland and Sitaram, which contains many other results and references. Now fifteen years have passed. Their overview is still valid, except that there are new avenues to add to the promenade that they propose. I have spoken of Wolff's Uncertainty Principle, as well as of the entirely new methods of Escauriaza, Kenig, Ponce and Vega. The Uncertainty Principle on finite Abelian groups should be added, with, following Tao [18], the use of Chebotarev's Lemma, which allows to prove that the sum of the cardinalities of the support of a function and its Fourier transform is bounded below by $p+1$ on the group $\mathbb{Z}/p\mathbb{Z}$, when p is prime. This, in turn, proves that sparse signals on $\mathbb{Z}/p\mathbb{Z}$ are determined by a small number of Fourier coefficients, a central issue in the mathematical theory of compressed sensing. No doubt it is a subject full of life!

References

[1] Amrein, W. O. and Berthier, A. M.: On support properties of L^p-functions and their Fourier transforms. *J. Funct. Anal.* **24** (1977), no. 3, 258–267.

[2] Benedicks, M.: On Fourier transforms of functions supported on sets of finite measure. *J. Math. Anal. Appl.* **106** (1985), 180–183.

[3] Bonami, A. and Demange, B.: A survey on uncertainty principles related to quadratic forms. *Collect. Math.* **2** (2006), Vol. Extra, 1–36.

[4] Bonami A., Demange, B. and Jaming, P.: Hermite functions and uncertainty principles for the Fourier and the windowed Fourier transforms. *Rev. Mat. Iberoamericana* **19** (2003), 23-55.

[5] Cowling M., Escauriaza, L., Kenig, C. E., Ponce, G. and Vega, L.: The Hardy Uncertainty principle revisited. To appear in *Indiana Univ. Math. J.*

[6] Demange, B.: Uncertainty principles associated to non-degenerate quadratic forms. *Mém. Soc. Math. Fr. (N.S.)* **119**, 98 pp. (2010).

[7] Demange, B.: Inégalités d'incertitude associées à des fonctions homogènes. *C. R. Math. Acad. Sci. Paris* **340** (2005), no. 10, 709–714.

[8] Escauriaza, L., Kenig, C. E., Ponce, G. and Vega, L.: The sharp Hardy uncertainty principle for Schrödinger evolutions. *Duke Math. J.* **155** (2010), no. 1, 163–187.

[9] Folland, G. B. and Sitaram, A.: The uncertainty principle –a mathematical survey. *J. Fourier Anal. Appl.* **3** (1997), 207–238.

[10] Havin, V. P.: On the uncertainty principle in harmonic analysis. In *Twentieth century harmonic analysis –a celebration (Il Ciocco, 2000)*, 3–29. NATO Sci. Ser. II Math. Phys. Chem. **33**. Kluwer Acad. Publ., Dordrecht, 2001.

[11] Havin, V. and Jöricke, B.: *The uncertainty principle in harmonic analysis.* Springer-Verlag, Berlin, 1994.

[12] Hörmander, L.: A uniqueness theorem of Beurling for Fourier transform pairs. *Ark. Mat.* **29** (1991), 237–240.

[13] Jaming, P.: Nazarov's Uncertainty Principles in Higher Dimension. *J. Approx. Theory* **149** (2007), 30–41.

[14] Kovrijkine, O.: Some results related to the Logvinenko–Sereda Theorem. *Proc. Amer. Math. Soc.* **129** (2001), 3037–3047.

[15] Kovrizhkin, O.: The uncertainty principle for certain densities. *Int. Math. Res. Not.* **17** (2003), 933–951.

[16] Nazarov, F. L.: Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type. *St. Petersburg Math. J.* **5** (1994), 663–717.

[17] Shubin C., Vakilian, R. and Wolff, T.: Some Harmonic Analysis questions suggested by Anderson–Bernoulli models. *Geom. Funct. Anal.* **8** (1998), 932–964.

[18] Tao, T.: An uncertainty principle for cyclic groups of prime order. *Math. Res. Lett.* **12** (2005), no. 1, 121–127.

[19] Thangavelu, S.: *An introduction to the uncertainty principle. Hardy's theorem on Lie groups*. Progress in Mathematics **217**. Birkhäuser, Boston, MA, 2004.

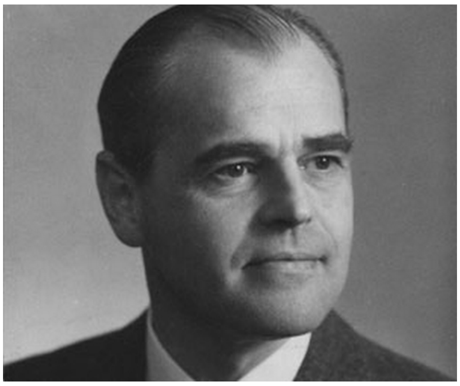
Joaquim Bruna (Barcelona, Spain, 1953) obtained his PhD in mathematics at the Universitat Autònoma de Barcelona in 1978 (advisor: Joan Lluís Cerdà). He is professor at the Departament de Matemàtiques of the Universitat Autònoma de Barcelona and Director of the Centre de Recerca Matemàtica. His main research areas are complex and harmonic analysis.

Some work by Beurling

Joaquim Bruna*

1. Introduction

When the editors of *La Revista* informed me about their excellent idea for this volume and asked for my contribution, it took me little time to decide that I would chose some articles of Arne Beurling[1].



Arne Beurling

I profess special admiration for Beurling, for his work, his influence and his taste in analysis. I did not have the chance to meet him so I know nothing first hand about his apparently strong character, or about how he viewed mathematical research. It's been said that he did not like to publish articles that were not polished, definite and complete, and consequently a substantial part of his research have never appeared as papers in scientific journals.

An immense gift to the mathematical community was delivered by Lennart Carleson, Paul Malliavin, John Neuberger and John Wermer when they edited and published *The Collected Works by Arne Beurling* in the eighties [6], containing plenty of previously unpublished research material. Yet, I have recently known from Misha Sodin that the Angstrom library in Uppsala still contains unpublished material of Beurling, which is supposed to be available online in a near future. Sodin and Michael Benedicks reported in September 2008 that they had gone through Beurling's archives and have selected 133 items "that might contain new results, observations and problems that should be of interest for many mathematicians working in different areas of analysis". All of this says much about how Beurling looked about publishing original research. I know about Beurling's work mostly through [6], as I guess it is the case for most analysts. In spite of the impact that of these two volumes have had, I think that Beurling's work is not as well known as it deserves.

*`bruna@mat.uab.cat`.

[1] He was born in Göteborg in 1905, and died in Princeton in 1986.

The work of Beurling, centered mostly in the interplay between harmonic and complex analysis, has been one of the most important and influential in analysis in the second half of the last century, but I have the feeling that maybe because of his personality his achievements have not gained the recognition they deserve, and that they have been overshadowed by those of a few no less important and influential analysts of his generation. Anyway, although I am not in a position to analyze all of this, I want to emphasize that I have decided to use this opportunity to pay homage to Beurling.

I have chosen two sets of papers by Beurling which have influenced my personal work and that of my students too. They are not his most famous results; probably his theorem on the description of shift invariant subspaces is better known. The first set consists of the two papers that Beurling published with P. Malliavin in the sixties providing the solution of the so called *radius of completeness problem* [2, 3], which had been already proposed by Paley and Wiener. As far as I know the proof was completed as early as 1961, yet the Acta paper appeared later in 1967. This problem had been studied by many other authors, who have contributed some partial results, and it has grown to become a central topic in classical harmonic analysis. My choice is also due to the fact that these papers perfectly explain the interplay between harmonic and complex analysis.

The second set consists of the two papers on *Balayage and Interpolation for Fourier–Stieltjes transforms*, which constitute the basis, together with the Shanon–Whittaker–Kotolnikov theorem, of modern sampling and interpolation theory. These two articles, as far as I know, were never published as research papers, though they constitute chapters IV and V, respectively, of the *Mittag–Leffler Lectures on Harmonic Analysis (1977–1978)*, and are included in [6]. Related to these is the paper of H. L. Landau [12], who undoubtedly influenced by Beurling carried over a similar study for multidimensional signals. This area of analysis has received renewed attention in recent years, mainly because of its connections with wavelets.

2. The radius of completeness problem

Let Λ be a discrete sequence of non-negative numbers and let $\mathcal{E}(\Lambda) = \{e^{i\lambda x}, \lambda \in \Lambda\}$ be the set of complex exponentials with frequencies in Λ. The linear space (topologically) spanned by $\mathcal{E}(\Lambda)$ is thus what we can get by superposition of sines and cosines with frequencies in Λ.

The question posed by Paley and Wiener was to describe, for a given Λ, how big this space is. To be more precise, let us define the *closure radius* or *radius of completeness* of Λ as

$$\rho(\Lambda) = \sup\left\{r \geq 0 : \mathcal{E}(\Lambda) \text{ spans } L^2(-r, r)\right\} \leq +\infty.$$

It is easy to show that this quantity does not change if one replaces the L^2 norm by any other reasonable norm. It does not change either if one adds or removes from Λ a finite number of points; this means that $\rho(\Lambda)$ is an asymptotic characteristic of Λ. Paley and Wiener wanted a description of $\rho(\Lambda)$ in metric or geometric terms, presumably using some appropriate notion of density.

By its very definition, one has that $\mathcal{E}(\Lambda)$ spans $L^2(-r,r)$ for $r < \rho(\Lambda)$ and does not span it if $r > \rho(\Lambda)$. For instance, Fourier's classical statement means simply that $\rho(\mathbb{Z}) = \pi$. A dual formulation of the problem invokes complex analysis. For $\mathcal{E}(\Lambda)$ does not span $L^2(-r,r)$ if and only if there exists *something* nonzero orthogonal to it, that is, a function $g \in L^2(-r,r)$ such that

$$\int_r^r g(x)e^{-i\lambda x}\,dx = 0, \qquad \text{for all } \lambda \in \Lambda.$$

This means that the Fourier–Laplace transform G of g,

$$G(z) = \int_r^r g(x)e^{-izx}\,dx, \quad z \in \mathbb{C},$$

vanishes on Λ. By the Paley–Wiener theorem, the Fourier–Laplace transform is one to one and isometric from $L^2(-r,r)$ onto the Paley–Wiener space of entire functions of exponential type $\leq r$, with L^2 values along the real line,

$$\mathrm{PW}_r = \{G : G\,\text{entire},\ |G(z)| = O(e^{r|z|}),\ G|_{\mathbb{R}} \in L^2(\mathbb{R})\}.$$

Hence it follows that

$$\rho(\Lambda) = \sup\{r \geq 0 : \Lambda\,\text{is a uniqueness set for}\,\mathrm{PW}_r\}.$$

Here we call Λ a uniqueness set for PW_r if whenever $G \in \mathrm{PW}_r$ and G vanishes on Λ, then $G \equiv 0$.

Thus the Paley–Wiener question is a problem about zero sequences of functions in Paley–Wiener spaces. It is intuitive then that the answer must be given by some sort of density $\mathcal{D}(\Lambda)$, say, defined for a certain class Σ of sequences; by dilation invariance, one should guess that $\rho(\Lambda)$ equals $c\mathcal{D}(\Lambda)$ for some constant c. Also note that if we agree that any reasonable notion of density should be 1 on $\mathbb{Z}$, then the constant c must be π. One has two things to prove:

1. $\pi\mathcal{D}(\Lambda) \leq \rho(\Lambda)$, *i.e.*, if $r < \pi\mathcal{D}(\Lambda)$ and $G \in \mathrm{PW}_r$ vanishes on Λ, then $G \equiv 0$.
2. $\pi\mathcal{D}(\Lambda) \geq \rho(\Lambda)$, *i.e.*, if $r > \pi\mathcal{D}(\Lambda)$, then there exists a nonzero function $G \in \mathrm{PW}_r$ vanishing on Λ.

The first one amounts to say that if a nonzero $G \in \mathrm{PW}_r$ vanishes on Λ then $\mathcal{D}(\Lambda) \leq \frac{r}{\pi}$. This is dealt usually with variants of Jensen's formula, which give information about the repartition of the *whole* sequence $Z(G)$ of zeroes of an entire function G. This means that it is rather the *outer density* $\mathcal{D}_e$ associated to $\mathcal{D}$ the one to be considered,

$$\mathcal{D}_e(\Lambda) = \inf \left\{ \mathcal{D}(\Gamma), \Lambda \subset \Gamma, \Gamma \in \Sigma \right\},$$

and that the result should read

$$\pi \mathcal{D}_e(\Lambda) = \rho(\Lambda).$$

Hence one must prove the above inequalities with $\mathcal{D}_e$ replacing $\mathcal{D}$.

Of course, the hard work is to find the right definition of $\mathcal{D}$. Every result providing information about the zero sequence of a function in the Paley–Wiener class in terms of some density $\mathcal{D}$ would lead to an estimate $\pi\mathcal{D}_e(\Lambda) \leq \rho(\Lambda)$. For instance, the ordinary density or Pólya density is defined as

$$D(\Lambda) = \lim_{t\to\infty} \frac{n(\Lambda, t)}{|t|},$$

whenever this makes sense, where $n(\Lambda, t)$ denotes the number of points of Λ between 0 and t. By Levinson's theorem [9], the zeroes of a function $G \in \mathrm{PW}_r$ satisfy

$$D(Z(G)) \leq \frac{r}{\pi}.$$

This leads to

$$\pi D_e(\Lambda) \leq \rho(\Lambda);$$

this much was known before Beurling and Malliavin.

The second type of inequality, $\pi\mathcal{D}_e(\Lambda) \geq \rho(\Lambda)$, is harder to prove, since it requires *constructing* an adequate function and estimating, in general, infinite products. At the time of Beurling and Malliavin it was known that

$$\pi D^u(\Lambda) \geq \rho(\Lambda),$$

with D^u some kind of uniform density.

Beurling and Malliavin found the right density $\mathcal{D}_e$, which they called *effective density*. It turns out to be the outer density associated to what Kahane [8] quotes as "densité maligne". A sequence Λ is said to have evil density $\mathcal{D}(\Lambda)$ if

$$\int_0^\infty \frac{|n(\Lambda, t) - \mathcal{D}(\Lambda) \cdot t|}{1 + t^2}\, dt < +\infty.$$

Theorem 1 *The radius of completeness is given by*

$$\rho(\Lambda) = \pi \mathcal{D}_e(\Lambda),$$

with

$$\mathcal{D}_e(\Lambda) = \inf\{\mathcal{D}(\Gamma), \Lambda \subset \Gamma\}.$$

As explained before, the proof of this result has two parts. The "easy" one consists in proving that if $G \in \mathrm{PW}_r$ vanishes on a sequence Λ with $\mathcal{D}(\Lambda) > \frac{r}{\pi}$, then $G \equiv 0$. A proof using Jensen's formula for ellipses can be found in [9].

For the harder part one must show that if $r > \pi \mathcal{D}_e(\Lambda)$ then there exists a nonzero $G \in \mathrm{PW}_r$ vanishing on Λ. A key role in this direction is played by the so called *Beurling–Malliavin multiplier theorem*, which is the object of their first Acta paper. Here the term multiplier is used in a sense different from the one used in Fourier Analysis. A weight $w(x) = e^{\omega(x)} \geq 1$ is said to *admit multipliers* if the translation operators are bounded in the weighted spaces $L^p(w)$ and if this latter space contains entire functions in PW_r for arbitrarily small $r > 0$. The characterization of the weights that they provide involves a Lipschitz type condition on ω and, more importantly, the integral condition

$$\int_{-\infty}^{+\infty} \frac{\omega(x)}{1+x^2}\, dx < +\infty.$$

What is relevant regarding the radius of completeness problem is that the same proof works for a weight $w(x) = |G(x)|$, G being an entire function of exponential type, leading to:

Theorem 2 *Let G be an entire function of exponential type such that*

$$\int_{-\infty}^{+\infty} \frac{\log^+ |G(x)|}{1+x^2}\, dx < +\infty.$$

Then for each $r > 0$ there exists $H \in \mathrm{PW}_r$ such that the product $G \cdot H$ is in some Paley–Wiener space.

These results allow to replace everywhere entire functions in Paley–Wiener spaces by entire functions of exponential type with finite logarithmic integral. So given $r > \pi \mathcal{D}_e(\Lambda)$ one must construct one such function G of exponential type $\leq r$ vanishing on Λ. As a first step, one considers a sequence Γ containing Λ with evil density $\mathcal{D}(\Gamma)$ bigger and arbitrarily close to $\mathcal{D}_e(\Lambda)$ and forms the Hadamard product with zeroes on Γ,

$$H(z) = \prod_{\sigma \in \Gamma} \left(1 - \frac{z}{\sigma}\right) e^{\frac{z}{\sigma}},$$

which has exponential type less than $\pi\mathcal{D}(\Gamma)$. The second step is to multiply this infinite product by another function of exponential type making the logarithmic integral finite. A careful analysis of this step (known as the little multiplier theorem) reveals that this is essentially a question about the behaviour of the Hilbert transform in L^1, hence a question related to H^1. It is noteworthy that in their proof, Beurling and Malliavin exhibit an explicit decomposition in what now are known as H^1-atoms.

Since Beurling–Malliavin original one, a number of alternative proofs have appeared. The papers by Beurling and Malliavin are very hard to read; a number of alternative expositions based on the same ideas have appeared since then, notably those by Kahane [8] and Koosis [9, 10] (see also [4] and [7]). A historical overview is to be found in [25]. Among the last ones, a broad generalization of the Beurling–Malliavin theory using the language of Toeplitz operators has appeared recently in [16].

I would like to use this opportunity to comment on another classical problem related to the Beurling–Malliavin theory, not yet completely solved. We are interested in the following situation: ϕ is a function in $L^p(\mathbb{R})$ and Λ is a discrete set in $\mathbb{R}$ such that the family of translates $\phi_\Lambda = \{\phi_\lambda(x) = \phi(x-\lambda), \lambda \in \Lambda\}$ span $L^p(\mathbb{R})$. We say that ϕ is a *generator* for $L^p(\mathbb{R})$ if for some discrete Λ, ϕ_Λ spans; and that Λ is a *spectral set* for $L^p(\mathbb{R})$ if for some ϕ, ϕ_Λ spans.

In [5] it is proved that Λ is a spectral set for $L^1(\mathbb{R})$ if and only if $\rho(\Lambda) = +\infty$. That this condition is necessary is rather evident. First note that obviously the Fourier transform $\hat{\phi}$ of the generator ϕ vanishes nowhere. The Fourier transform of the linear span of ϕ_Λ is $\hat{\phi}$ times the linear span of $\mathcal{E}(\Lambda)$, and it will be dense in the L^∞-norm in the Fourier image of $L^1(\mathbb{R})$. Since the latter is dense in $L^\infty(I)$ for any interval I and $\hat{\phi}$ is bounded above and below over I, it turns out that $\mathcal{E}(\Lambda)$ must be dense in $L^\infty(I)$ for all intervals I, whence $\rho(\Lambda) = +\infty$. The situation is quite different in $L^p, p \neq 2$, for there are slight perturbations of the integers which are spectral for $L^2(\mathbb{R})$, and $\mathbb{Z}$ is spectral for $L^p(\mathbb{R}), p > 2$ (see [1, 20, 21]).

It is not hard to see that ϕ is not a generator for $L^p(\mathbb{R})$ if and only if there exists $g \in L^q(\mathbb{R})$, q being the conjugate exponent of p, such that the convolution $\phi * g$ has compact support. It follows that ϕ is not a generator for $L^2(\mathbb{R})$ if and only if $\hat{\phi}$ is the quotient of some function in a Paley–Wiener space over an L^2-function, a situation similar to that studied by Beurling and Malliavin. This implies that

$$\int_{-\infty}^{+\infty} \frac{\log|\hat{\phi}(\xi)|}{1+|\xi|^2}\, d\xi > -\infty.$$

The same conclusion can be proved for $L^1(\mathbb{R})$. The converse is not true (private communication of A. Borichev). Hence together with $\hat{\phi}(\xi) \neq 0$, divergence of the above integral is a sufficient (but not necessary) condition for ϕ to be a generator for $L^1(\mathbb{R})$.

3. Sampling and interpolation of band–limited signals

In the papers [6, 341–365], Beurling sets the grounds for modern sampling and interpolating theory. Beurling uses an equivalent formulation borrowed from potential theory. Here I will describe it using the notation employed today.

The Bernstein space B_r is the L^∞ version of the Paley–Wiener space: it consists of the entire functions F of exponential type less than or equal than r, bounded on the real line. This space coincides with the space of bounded continuous functions on $\mathbb{R}$ with spectrum (band-limited) in $[-r, r]$. A discrete set Λ in $\mathbb{R}$ is called a *sampling set* for B_r if there exists a constant K such that

$$\sup_{x\in\mathbb{R}} |F(x)| \leq K \sup_{x\in\Lambda} |F(x)|, \quad \text{for each } F \in B_r.$$

This means that the samples of F on Λ allow to reconstruct the whole function F in a stable way. The accompanying notion of *interpolating set* means that every bounded sequence (c_λ) can be interpolated by some $F \in B_r$, *i.e.*, $F(\lambda) = c_\lambda$ for all $\lambda \in \Lambda$.

Beurling gave a complete description of both types of sequences in terms of uniform densities defined as follows. We denote by $n^+(t)$ the largest number of points of Λ in a closed interval of length t; similarly $n_-(t)$ denotes the smallest number. The upper and lower uniform densities are respectively defined by

$$D^+(\Lambda) = \lim_{t\to+\infty} \frac{n^+(t)}{t}, \qquad D^-(\Lambda) = \lim_{t\to+\infty} \frac{n^-(t)}{t}.$$

Theorem 3 *Λ is a sampling sequence for B_r if and only if contains a separated sequence Λ' such that $D^+(\Lambda') > \frac{r}{\pi}$, and is an interpolating sequence if and only if it separated and $D^-(\Lambda) < \frac{r}{\pi}$.*

We recall that a sequence is called *separated* if the distance between two points in the sequence is uniformly bounded from below.

Beurling proofs depends on his notion of weak limits of translates. Let $W(\Lambda)$ denote the collection of sets Γ for which there exists translation parameters τ_n such that the sets $\Lambda + \tau_n$ converges weakly to Γ. Here the weak convergence of sets means that for every compact K, the traces on K tend to $\Gamma \cap K$ in the Fréchet distance. It turns out that Λ is a sampling set if and only if all sets $\Gamma \in W(\Lambda)$ are uniqueness sets for B_r, meaning that there is no nonzero function in B_r vanishing on Γ. This shows the connection with the problems discussed in the previous section and with complex analysis in particular.

Of course, these notions can be defined if the interval $[-r, r]$ is replaced by an arbitrary set S in $\mathbb{R}^n$ of finite measure. The space of bounded continuous on $\mathbb{R}^n$ with spectrum in S can be described as a space of entire functions only if S is convex, so complex analysis techniques are not good anymore, in general. In a marvelous paper [12] that I recommend every analyst to read, Landau found a real analysis proof of the necessary conditions

$$D^+(\Lambda) > \frac{|S|}{2\pi}, \qquad D^-(\Lambda) < \frac{|S|}{2\pi},$$

valid for an arbitrary bounded set. The sufficient conditions cannot hold in general; no density can provide a complete solution of the problem.

The relation of all this with the previous work of Beurling and Malliavin is best seen in the L^2 context. The notion of sampling for the Paley–Wiener space is

$$\int_{-\infty}^{+\infty} |G(x)|^2\, dx \leq K \sum_{\lambda \in \Lambda} |G(\lambda)|^2, \quad \text{for each } G \in \mathrm{PW}_r,$$

while Λ is interpolating if every sequence (c_λ), $\sum_\lambda |c_\lambda|^2 < +\infty$, can be interpolated by some $F \in \mathrm{PW}_r$. For a set Λ being sampling for PW_r amounts to say that the family of exponentials $\mathcal{E}(\Lambda)$ is a frame of $L^2(-r, r)$, while being interpolating means that this family is free; in particular, sequences that are simultaneously sampling and interpolating in PW_r (which, by the way, do not exist in B_r) correspond to Riesz basis $\mathcal{E}(\Lambda)$ of exponentials. Beurling's results give, for separated sequences,

$$D^+(\Lambda) > \frac{r}{\pi} \implies \Lambda\,\text{sampling} \implies D^+(\Lambda') \geq \frac{r}{\pi},$$

and analogously for interpolating sequences. In particular,

$$\sup\left\{r \geq 0 \,:\, \mathcal{E}(\Lambda)\,\text{frame for}\, L^2(-r, r)\right\} = D^+(\Lambda).$$

The frames of exponentials in L^2 were fully described in [23]. Riesz basis occur at the critical density, and they were characterized in [24].

In my view, the two papers of Beurling and the one of Landau, both their results and their techniques, have been very influential in this area of analysis. I think that thanks to [6] many analysts discovered Beurling's work and somewhat this sparked new interest. Also, the development of the theory of wavelets, for which sampling is basic, attracted new interest. Beurling uniform densities have been generalized to a number of different contexts. For instance, by adapting Beurling ideas to the hyperbolic setting, Seip was able to describe sampling and interpolating sequences for various function spaces of holomorphic functions in the unit disc (see [26] and the references therein). They were described as well for the Fock space, an isometric copy of $L^2(\mathbb{R})$ via the Bargmann transform, in terms of the natural uniform densities; in this case there is a critical density above which one has sampling sequences and below which one has interpolating sequences. The non existence of sampling and interpolating sequences in the Fock space is equivalent to the non existence of Riesz basis of the so called gaborlets (time-frequency translates of the gaussian function), a particular case of the Balian–Low theorem. It is said that this result lead Yves Meyer to try to prove that wavelet bases well localized both in time and frequency did not exist either, a hint that fortunately turned out to be false and opened a whole new area. In the electronic engineering community, Landau paper is a milestone. Altogether, the impact of Beurling's work is very noticeable nowadays.

I use the opportunity to call the attention of the reader to a couple of basic problems that remain open in the area. For instance, it is not known yet whether in the case where S is the unit ball in $\mathbb{R}^n$ there are Riesz basis of complex exponentials. In this case, the corresponding Paley–Wiener space consists of entire functions F in $\mathbb{C}^n$ of exponential type at most one, with $F|_{\mathbb{R}^n} \in L^2(\mathbb{R}^n)$. The question is to decide whether there are sampling and interpolating sequences in this space. In the L^p- version of this space, $p \neq 2$, this turns out to be false, as a consequence of Fefferman's theorem on the ball multiplier (J. Ortega-Cerdà, private communication). On the other hand, in [14] it is shown that the answer is affirmative for a class of convex polygons. In the case where S is a finite union of intervals, it is not known in general whether Riesz basis of exponentials exist; this is known as the *multiband problem.* Partial results appear in [15] and [27]. A very interesting connection with quasicrystals is shown in [11], [13], [17] and [19]. Connections with compressed sensing through the notion of *universal sampling* have been undertaken in [18] and [22].

References

[1] Atzmon, A. and Olevskii, A.: Completeness of integer translates in function spaces. *J. Approx. Theory* **83** (1996), 291–327.

[2] Beurling, A. and Malliavin, P.: On Fourier transforms of measures with compact support. *Acta Math.* **107** (1962), 291–309.

[3] Beurling, A. and Malliavin, P.: On the closure of characters and the zeros of entire functions. *Acta Math.* **118** (1967), 79–93.

[4] De Branges, L.: *Hilbert spaces of entire functions.* Prentice–Hall, 1968.

[5] Bruna, J., Olevskii, A. and Ulanovskii, A.: Completeness in $L^1(\mathbb{R})$ of discrete translates. *Rev. Mat. Iberoamericana* **22** (2005), 1–16.

[6] Carleson, L., Malliavin, P., Neuberger, J. and Wermer, J. (editors): *The collected works of Arne Beurling, volumes I and II.* Birkhauser, 1989.

[7] Havin, V. and Joricke, J.: *The uncertainty principle in harmonic analysis.* Springer, 1994.

[8] Kahane, J. P.: *Travaux de Beurling et Malliavin.* Seminaire Bourbaki, exposés 223 a 228 (1968).

[9] Koosis, P.: *The logarithmic integral, vols I and II.* Cambridge University Press, 1968.

[10] Koosis, P.: *Leçons sur le theoreme de Beurling et Malliavin.* Les publications CRM, Montreal, 1996.

[11] Kozma, G. and Lev, N.: Exponential Riesz bases, discrepancy of irrational rotations and BMO. To appear in *J. Fourier Anal. Appl.*

[12] Landau, H. J.: Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.* **117** (1967), 37–52.

[13] Lev, N.: Riesz bases of exponentials on multiband spectra. To appear in *Proc. Amer. Math. Soc.*

[14] Lyubarskii, Y. and Rashkovskii, A.: Complete interpolating sequences for Fourier transforms supported by convex symmetric polygons. *Ark. Mat.* **38** (2000), 139–170.

[15] Lyubarskii, Y. and Seip, K.: Sampling and interpolating sequences for multiband-limited functions and exponential bases on disconnected sets. *J. Fourier Anal. Appl.* **3** (1997), 597–615.

[16] Makarov, N. and Poltoratski, A.: Beurling-Malliavin theory for Toeplitz kernels. *Invent. Math.* **180** (2010), 443–480.

[17] Matei, B. and Meyer, Y.: Quasicrystals are sets of stable sampling. *C. R. Acad. Sci. Paris, Ser. I* **346** (2008), 1235–1238.

[18] Matei, B. and Meyer, Y.: A variant of compressed sensing. *Rev. Mat. Iberoam.* **25** (2009), 669–692.

[19] Matei, B. and Meyer, Y.: Simple quasicrystals are sets of stable sampling. *Complex Var. Elliptic Equ.* **55** (2010), 947–964.

[20] Nikolskii, N.: *Selected problems of weighted approximation and spectral analysis.* Proceedings Steklov Institute of Mathematics **120** (1974); American Mathematical Society, Providence, RI, 1976.

[21] Olevskii, A. and Ulanovskii, A.: Almost integer translates; do nice generators exist? *J. Fourier Analysis Appl.* **10** (2004), 93–104.

[22] Olevskii, A. and Ulanovskii, A.: Universal sampling of band-limited signals. *C. R. Math. Acad. Sci. Paris* **342** (2006), 927–931.

[23] Ortega, J. and Seip, K.: Fourier frames. *Ann. of Math. (2)* **155** (2002), 789–806.

[24] Pavlov, B.: The basis property of a system of exponentials and the condition of Muckenhoupt. *Dokl. Akad. Nauk SSSR* **247** (1979), 37–40.

[25] Redheffer, R.: Completeness of sets of complex exponentials. *Adv. Math.* **24** (1977), 1–62.

[26] Seip, K.: *Interpolation and Sampling in Spaces of Analytic Functions.* American Mathematical Society, Providence, RI, 2004.

[27] Seip, K.: A simple construction of exponential bases in L^2 of the union of several intervals. *Proc. Edinburgh Math. Soc.* **38** (1995), 171–177.

Luis A. Caffarelli (Buenos Aires, Argentina, 1948) obtained his PhD in mathematics from the University of Buenos Aires (advisor: Calixto P. Calderón). He currently holds the Sid Richardson Chair at the University of Texas at Austin. His main research area is nonlinear analysis, in particular the study of nonlinear partial differential equations arising from geometry and mechanics.

A lemma of De Giorgi

Luis A. Caffarelli*

In 1956, at age 28, Ennio de Giorgi [9] solved the 19th Hilbert problem by proving the regularity and analyticity of variational (weak) solutions to nonlinear elliptic variational problems. More precisely, given a variational integral

Ennio de Giorgi

$$\mathcal{T}(u) = \int_\Omega F(\nabla u)\, dx \ ,$$

local minimizers, u_0, of $\mathcal{T}(u)$ satisfy the Euler–Lagrange equation

$$\text{div} F_j(\nabla u) = 0$$

(with $F_j = D_j F$) in the sense that $\nabla u \in L^2$ and

$$\int_\Omega \nabla\varphi \ F_j(\nabla u)\, dx = 0$$

for any φ compactly contained in Ω, and $\nabla\varphi \in L^2$.

The function F is supposed to be strictly convex and smooth, *i.e.*, $\lambda I \leq F_{ij}(\cdot) \leq \Lambda I$, and one observes easily that

a) $\text{div} F_j(\nabla u) = F_{ij}(\nabla u) D_{ij} u$

also that

b) if w is a directional derivative $D_\sigma u$, then $D_i F_{ij}(\nabla u) D_j w = 0$.

It follows from a) that if ∇u_0 were C^α, then u_0 would be $C^{2,\alpha}$ (from Schauder estimates), ∇u_0 would be $C^{1,\alpha}$ and so on. However, all we know at this point is that $\nabla u \in L^2$ and also the ellipticity estimate $\lambda I \leq F_{ij}(\cdot) \leq \Lambda I$.

So, De Giorgi set out to prove a linear theorem.

If in $B_1 \subset \mathbb{R}^n$ the function w satisfies

$$D_i a_{ij}(x) D_j w = 0 \qquad \text{with} \qquad \lambda I \leq a_{ij} \leq \Lambda I \ \forall x$$

and $\nabla w \in L^2$ (this is attained by replacing $w = D_\sigma u_0$ by an incremental quotient),

then in $B_{1/2}$, the function w is C^α.

*caffarel@math.utexas.edu.

Note that the theorem is linear, but it is in a higher invariance class.

If a_{ij} where smooth, under dilations of the space it becomes “constant” coefficients (the Laplacian) and this was the basis of the existing regularity theory (Schauder, Calderón–Zygmund, Cordes–Nirenberg).

But now, no matter how much we rescale the equation, the coefficients from $a_{ij}(x)$ to $a_{ij}(\varepsilon x)$ remain “bounded measurable” and far from constant.

The complete proof of Hilbert problem, of just 18 pages, is very elegant and geometric and the underlying ideas have been used extensively for regularity of solutions to integral equations [3, 6], the partial regularity of solutions to Navier–Stokes equation, [8, 10], homogenization, phase transitions [1], free boundary problems [2, 4].

The lemma referred to in the title is the first part of his proof, and consists in showing that if $w^+|_{B_1} \in L^2$, then $w^+|_{B_{1/2}} \in L^\infty$ and

$$\sup_{B_{1/2}} w^+ \lesssim \|w^+\|_{L^2(B_1)}.$$

The main ingredient in the proof is that for all truncations $w_\lambda = (w-\lambda)^+$ of w, we have competing inequalities between $\|w\|_{L^2}$ and $\|\nabla w\|_{L^2}$ and they have different homogeneities.

The first one is a general fact, Sobolev inequality: If φ is a cutoff function in B_1, *i.e.*, $\varphi \equiv 0$ near ∂B_1,

$$\|\varphi w_\lambda\|_{L^p} \le C\|\nabla(\varphi w_\lambda)\|_{L^2}$$

for some $p(n) > 2$.

The second, instead, the energy formula, holds for solutions w of an equation

$$D_i a_{ij} D_j w = 0.$$

It says that

$$\int_{B_1} \nabla(\varphi w_\lambda^2) \le \sup_{B_1} |\nabla\varphi| \cdot \int_{B_1} w_\lambda^2\ .$$

A very original interplay between these two inequalities implies the lemma.

The underlying geometric idea in De Giorgi's lemma is that if we have a hypersurface (the boundary of a set, or a graph) where two “elliptic quantities” of different homogeneity compete, this imposes some “scale invariant” control on the geometry of the surface.

Examples are:

1) The interaction between *“Sobolev and energy inequalities”* for problems in the calculus of variations: by Sobolev “(the L^2 norm of) derivatives control the (L^p norm of the) function” while, for energy minimizers, the (L^2 norm of the) function controls the (L^2 norm of) derivatives [9].

2) *"Volume and area":* By the isoperimetric inequality ("Sobolev") area controls volume (to the $\frac{n-1}{n}$), and for a minimal boundary, locally "volume of the set controls area of the boundary (from appropriate perturbations) [5].
3) *"Harmonic measure and area"*: For a free boundary problem where curvature and harmonic measure interact, boundary curvature controls harmonic measure (with different homogeneities) and from the free boundary condition, harmonic measure controls boundary area (and curvature) [2].

In this note, I would like to show how these ideas work in two completely different circumstances, to provide important information.

The first example, from a work in collaboration with A. Vasseur [6] (see also [3]), uses these ideas to prove that solutions to the surface quasi-geostrophic equations, with initial data in L^2, become instantaneously bounded. It is a simple case since being an "all space" theorem it does not need space truncations.

The second, in collaboration with Roquejoffre and Savin [5], concerns "nonlocal" minimal surfaces, and it exhibits the fact that the De Giorgi lemma has a geometrical underpinning beyond the apparent "functional character" of his lemma.

1. The parabolic lemma

We assume that $u(x,t)$ satisfies an integral heat equation

$$u_t(x,t) = \int_{\mathbb{R}^n} \Big[u(y,t) - u(x,t)\Big] K(x,y,t)\, dy$$

- where the kernel $K(x,y,t)$ is comparable to the s-fractional Laplacian kernel

  $$K^{(s)}(x,y) = \frac{(1-s)}{|x-y|^{n+2s}}$$

  in the sense that

  $$\lambda K^{(s)} \le K \le \Lambda K^{(s)} \quad \text{for all} \quad x,y,t$$

  (this replaces the uniform ellipticity hypothesis: $\lambda I \le a_{ij} \le \Lambda I$);
- and K is symmetric in x,y, *i.e.*, $K(x,y,t) = K(y,x,t)$.

We want to prove:

Theorem 1

$$u^+(0,1) \le C\|u^+(x,0)\|_{L^2}.$$

As we pointed out before, this is obtained by the interplay of two competing inequalities: Sobolev and energy.

We assume that the weak solution $u(x,t)$ is good enough so that we can multiply by a truncation, $u_\lambda(x,t) = (u(x,t)-\lambda)^+$ and integrate by parts:

$$\int_{t_1}^{t_2}\int_{\mathbb{R}^n} u_\lambda(z,t)\; u_t(z,t)\,dz\,dt =$$
$$= \int_{t_1}^{t_2}\int_{\mathbb{R}^n\times\mathbb{R}^n} u_\lambda(z,t)\Big[(u(y,t)-u(z,t)\Big]K(z,y,t)\,dz\,dy\,dt$$

We note that, from the symmetry of K we can interchange the roles of z and y:

$$= \int_{t_1}^{t_2}\int_{\mathbb{R}^n\times\mathbb{R}^n} u_\lambda(y,t)\Big[u(z,t)-u(y,t)\Big]K(z,y,t)\,dy\,dz\,dt$$

add and divide by 2:

$$= \int_{t_1}^{t_2}\int_{\mathbb{R}^n\times\mathbb{R}^n} \Big[u_\lambda(z,t)-u_\lambda(y,t)\Big]\cdot K(z,y,t)\Big[u(y,t)-u(z,t)\Big]\,dz\,dy\,dt$$
$$= -\int_{t_1}^{t_2} B_t(u_\lambda,u)\;dt,$$

where

$$B_t(f,g) = \int_{\mathbb{R}^n\times\mathbb{R}^n}\Big[f(z)-f(y)\Big]K(z,y,t)\Big[g(z)-g(y)\Big]\,dz\,dy.$$

That is, we have now the formula

$$\int_{\mathbb{R}^n} u_\lambda^2(z,t_2)\,dz + \int_{t_1}^{t_2} B_t(u_\lambda,u)\,dt = \int_{\mathbb{R}^n} u_\lambda^2(z,t_1)\,dz.$$

To complete the inequality, we note that $B_t(u_\lambda, u-u_\lambda)\geq 0$ by considering the four cases $x,y\in\{u_\lambda>0\}$ or $\{u_\lambda=0\}$, and we get the final energy inequality

$$E_{t_1,t_2}(u_\lambda) = \int_{\mathbb{R}^n} u_\lambda^2(z,t_2)\,dz + \int_{t_1}^{t_2} B_t(u_\lambda,u_\lambda)\,dt \leq \int_{\mathbb{R}^n} u_\lambda^2(z,t_1)\,dz$$

But, from the hypothesis

$$\lambda K^{(s)} \leq K \leq \Lambda K^{(s)}$$

the bilinear form verifies

$$B_t(u_\lambda,u_\lambda)\sim \|u_\lambda(\cdot,t)\|^2_{H_s}.$$

We want to use now, complementary to the energy inequality, the Sobolev inequality for some $p > 2$ (the important thing about p is $p > 2$):

$$\begin{aligned}\left[\int_{t_1}^{t_2}\int_{\mathbb{R}^n} u_\lambda(z,t)^p\,dz\,dt\right]^{2/p} &\le C\left[\sup_{t_1\le t\le t_2}\|u(\cdot,t)\|_{L^2} + \int_{t_1}^{t_2}\|u_\lambda(\cdot,t)\|_{H_2}^2\,dt\right]\\ &\le C\sup_{t_1\le t\le t_2} E_{t_1,t}(u_\lambda)\\ &\le C\left[\int_{\mathbb{R}^n} u_\lambda^2(z,t_1)\,dz\right].\end{aligned}$$

Finally, from Hölder inequality,

$$\int_{t_1}^{t_2}\int_{\mathbb{R}^n} u_\lambda(z,t)^2\,dz\,dt \le \left[\int_{t_1}^{t_2}\int_{\mathbb{R}^n} u_\lambda^p(z,t)\,dz\,dt\right]^{2/p}\left|\{u_\lambda > 0\}\right|^{(1-\frac{2}{p})}.$$

Now, we reproduce the De Giorgi lemma.

We will choose $t_2 = \infty$, for a sequence of truncations of u: $\lambda_k = 1 - 2^{-k}$, and for an appropriate sequence of times t_k:

$$t_k \in \mathbb{I}_k = \left(1 - 2^{-k}, 1 - 2^{-(k+1)}\right)$$

one has

$$\int_{\mathbb{R}^n} u_{\lambda_k}^2(z,t_k)\,dz = \inf_{\mathbb{I}_k}\left[\int_{\mathbb{R}^n} u_{\lambda_k}^2(z,t)\,dz\right] \le 2^k\int_{\mathbb{I}_k}\int_{\mathbb{R}^n} u_{\lambda_k}^2(z,t)\,dz\,dt$$

Finally, we will denote

$$A_k = \left[\int_{t_k}^{\infty}\int_{\mathbb{R}^n}\left[u_{\lambda_k}(z,t)\right]^p\,dz\,dt\right]^{2/p}$$

and obtain a recurrence relation for A_k ($t_0 = 0$, $\lambda_0 = 0$).

Then

$$A_0 \le C\int_{\mathbb{R}^n} u^2(z,0)\,dz$$

and

$$\begin{aligned}A_{k+1} &\le C_1\cdot E_{t_{k+1},\infty}(u_{\lambda_{k+1}})\\ &\le C_2\int_{\mathbb{R}^n} u_{\lambda_{k+1}}^2(z,t_{k+1})\;dz\\ &\le C_3\cdot 2^{k+1}\int_{t_k}^{\infty}\int_{\mathbb{R}^n} u_{\lambda_{k+1}}^2(z,t)\;dz\,dt \qquad (1)\\ &\le \left[\int_{t_k}^{\infty}\int_{\mathbb{R}^n} u_{\lambda_{k+1}}^p(z,t)dz\,dt\right]^{2/p}\left|\{u_{\lambda_{k+1}} > 0\}\right|^{\varepsilon}.\end{aligned}$$

But now, we jump from $k+1$ to k:

a) $u_{\lambda_k} > u_{\lambda_{k+1}}$

b) $\{u_{\lambda_{k+1}} > 0\} = \{u_{\lambda_k} > 2^{-k}\}$ and

c) $|\{u_{\lambda_{k+1}} > 0\}| = |\{u_{\lambda_k} > 2^{-k}\}| \leq 2^{pk} \int (u_{\lambda_k})^p$

So we get

$$A_{k+1} \leq C\ 2^{\varepsilon kp} \left[\int_{t_k}^{\infty} \int_{\mathbb{R}^n} u^p_{\lambda_k}(z,t)\ dz\, dt \right]^{\frac{2}{p}+\varepsilon}. \tag{2}$$

All together we get the *nonlinear* relation

$$A_{k+1} \leq C\ 2^{Mk} A_k^{(1+\bar{\varepsilon})}.$$

If A_0 is small enough ($\leq D$) (depending on $C, M, \bar{\varepsilon}$), $A_k \to 0$. That is,

$$\int_1^{\infty} \int_{\mathbb{R}^n} (u(z,t)-1)^+\ dz\, dt \equiv 0.$$

Since the operator is linear, it follows that

$$\|u\|_{L^\infty(\mathbb{R}^n, t>1)} \leq \frac{1}{D} \|u(\cdot,0)\|_{L^2(\mathbb{R}^n)}.$$

2. Nonlocal minimal surfaces

The second application concerns nonlocal minimal surfaces in the context of "minimal boundaries" [5].

The idea of minimal boundaries was also introduced by De Giorgi, following Cacciopoli:

Definition. A set Ω has *locally finite perimeter* if its indicator function has locally finite bounded variation.

A weak definition can be given, for instance, by using formally Greens theorem:

$$\text{“} \int \nabla \chi_\Omega \varphi + \int \chi_\Omega \nabla \varphi = 0 \text{”}$$

The first term does not make "L^p sense" but the second does, provided that φ is C^1:

Definition. Ω has *finite perimeter* in the unit ball B_1 if, for any φ in $C_0(B_1)$,

$$\left| \int \chi_\Omega \nabla \varphi \right| \leq C_{B_1} \|\varphi\|_{L^\infty}.$$

The optimal C_{B_1} is the perimeter of Ω in B_1.

Then, we can define: The set Ω has "minimal perimeter" in B_1, if any other set $\widetilde{\Omega}$ that coincides with Ω outside B_1. ($\Omega \triangle \widetilde{\Omega} \subset B_1$) has larger perimeter than Ω.

Minimal surface theory of which De Giorgi was a main contributor then shows that $\partial\Omega \cap B_1$ is an analytic surface except for a set of $(n-8)$-Hausdorff measure 0. An important property in the theory of sets of minimal perimeter is the "uniform density property": For any ball $B_\rho(x_0)$ contained in the domain of minimality (say B_1), and centered on $\partial\Omega$,

$$|B_\rho \cap \Omega| \text{ and } |B_\rho \cap \Omega^C| \geq \lambda |B_\rho|$$

for some $\lambda > 0$ depending only on dimension. Further, $\exists\, \mu > 0$ also depending only on dimension such that for some z_1, z_2

$$B_\rho(x_0) \cap \Omega \supset B_{\mu\rho}(z_1) \text{ and } B_\rho(x_0) \cap \Omega^C \supset B_{\mu\rho}(z_2).$$

In particular, $H^{n-\varepsilon}(\partial\Omega) = 0$ for some small ε.

There is also a compacity result within this family of sets having locally finite perimeter, namely: given a sequence Ω_k that converges to Ω_0 in L^1, then it actually converges uniformily (*i.e.*, in Hausdorff distance).

Nonlocal minimal surfaces (or sets of *nonlocal minimal perimeter*) arise in geometry and in problems involving phase transition with long range interactions.

In this case, the set Ω, instead of minimizing perimeter, minimizes the H^s norm of its characteristic function:

$$E_s(\Omega) = \iint \frac{|\chi_\Omega(x) - \chi_\Omega(y)|^2}{|x-y|^{n+2s}} = \left(\iint \frac{\chi_\Omega(x)\chi_{\Omega^C}(y)}{|x-y|^{n+2s}} \right).$$

Some remarks

1) $E_s(\Omega)$ is infinite for $s \geq 1/2$, even for Ω a ball. In fact, if we renormalize E_s to $\widetilde{E}_s = (1-2s)E_s$ and Ω is a smooth and bounded set, then $\widetilde{E}_s(\Omega)$ converges to Perimeter(Ω) as $s \to 1/2$ (see [7]).

2) Let us define the bilinear form on sets

$$L(S,T) = \iint \frac{\chi_S(x)\ \chi_T(y)}{|x-y|^{n+2s}}.$$

Then, if Ω is a minimizer in B_1, $A \subset \Omega \cap B_1$ and we compare E_Ω with $E_{\Omega \setminus A}$, a simple computation gives us that

$$L(A, \Omega \setminus A) - L(A, \Omega^C) \geq 0,$$

while if $A \subset \Omega^C \cap B_1$,

$$L(A, \Omega) - L(A, (A \cup \Omega)^C) \leq 0.$$

3) If we dilate the sets by λ,

$$L(\lambda S, \lambda T) = \lambda^{n-2s} L(S, T).$$

4) $L(B_1, B_1^C) = C_0 = C_0(n, s) < \infty$ for $s < 1/2$.

5) If Ω is a minimizer in B_1 and $A = (\Omega \cap B_1)$, from 2) we get

$$L(\Omega \cap B_1, \Omega \setminus B_1) \geq L(\Omega \cap B_1, \Omega^C).$$

In particular, the "internal energy" verifies

$$L(\Omega \cap B_1, \Omega^C \cap B_1) \leq L(\Omega \cap B_1, \Omega \setminus B_1) \leq L(\Omega \cap B_1, \mathbb{R}^n \setminus B_1).$$

This inequality uses that Ω is a minimizer and replaces the "energy inequality".

The opposite one is Sobolev inequality: There exists a $p = p(n, s) > 2$ so that for any bounded set Σ,

$$\|\chi_\Sigma\|_{L^p} \leq C \|\chi_\Sigma\|_{H^s}.$$

We now show:

Density Lemma. *There exists a small constant δ_0, so that if Ω is a minimizer in B_1, and $|\Omega \cap B_1| \leq \delta_0$, then $|\Omega \cap B_{1/2}| = 0$.*

Proof. We will consider a sequence of truncations

$$\Omega_{r_k} = \Omega \cap B_{r_k}, \qquad \text{with } r_k = \tfrac{1}{2} + 2^{-k}$$

and

$$A_k = \|\chi_{\Omega_{r_k}}\|_{L^p}^2 = (\text{Vol } \Omega_{r_k})^{2/p}$$

In particular $A_1 \leq \delta_0$. Then, from Sobolev inequality for a generic r we have

$$A_r \leq L(\Omega_r, \Omega_r^C) = \|\chi_{\Omega_r}\|_{H_s}^2$$

But, from minimality (see 5) above)

$$L(\Omega_r, \Omega_r^C) \leq L(\Omega_r, \mathbb{R}^n \setminus B_r)$$

Next we estimate $L(\Omega_r, \mathbb{R}^n \setminus B_r)$ in terms of $a(t)$ the area of $\Omega_r \cap \partial B_t$.

Note that if $|x| = t < r$

$$\int_{B_r^C} \frac{1}{|x-y|^{n+2s}}\, dy = C(r-t)^{-2s}.$$

Therefore, for $r > 1$,

$$L(\Omega_r, \mathbb{R}^n \setminus B_r) \leq \int_0^r a(t)(r-t)^{-2s}\, dt.$$

Together with the Sobolev inequality this gives us

$$A_r \leq C \int_0^r a(t)(r-t)^{-2s}\, dt.$$

Integrating in r from 0 to r_k we get

$$\int_0^{r_k} V^{2/p}(\Omega_r) \leq C r_k^{2-s} V(\Omega_{r_k}).$$

Since $V(\Omega_r)$ is monotone in r we get, as before, the recurrence relation

$$\Big(V(\Omega_{r_{k+1}})\Big)^{2/p} \leq C\; 2^k V(\Omega_{r_k}).$$

As before, this completes the proof of the lemma. It follows from here, with appropriate covering lemmas that

a) There exists $\lambda > 0$ such that, if X_0 belongs to the "minimal surface", $\partial\Omega$, then for any r, there are balls

$$B_{\lambda r}(Y_1) \subset \Omega \cap B_r(X_0),$$
$$B_{\lambda r}(Y_2) \subset \Omega \cap B_r(X_0),$$

and

b) $H^{n-2s}(\partial\Omega \cap B_r(X_0)) \leq r^{n-2s}$.

References

[1] Athanasopoulos, I., Caffarelli, L. A. and Salsa, S.: The free boundary in an inverse conductivity problem. *J. Reine Angew Math.* **534** (2001), 1–31.

[2] Athanasopoulos, I., Caffarelli, L. A., Kenig, C. and Salsa, S.: An area-Dirichlet integral minimization problem. *Comm. Pure Appl. Math.* **54** (2001), no. 4, 479–499.

[3] Caffarelli, L. A., Chan, C. H. and Vasseur, A.: Regularity theory for nonlinear integral operators. `arXiv:1003.1699`. To appear in *J. Amer. Math. Soc.*

[4] Caffarelli, L. A. and Mellet, A.: Random homogenization of an obstacle problem. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **26** (2009), no. 2, 375–395.

[5] Caffarelli, L. A., Roquejoffre, J.-M. and Savin, O.: Nonlocal minimal surfaces. *Comm. Pure Appl. Math.* **63** (2010), no. 9, 1111–1144.

[6] Caffarelli, L. A. and Vasseur, A.: Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math. (2)* **171** (2010), no. 3, 190–1930.

[7] Caffarelli, L. A. and Valdinoci, E.: Uniform estimates and limiting arguments for nonlocal minimal surfaces. *Calculus of Variations* **41** (2011), no. 1-2, 203–240.

[8] Chen, C., Strain, R., Tsai, T. and Yau, H. T.: Lower bounds on the blow up rate of the axisymmetric Navier–Stokes equations. *Int. Math. Res. Notices*, no. 9 (2008).

[9] De Giorgi, E.: Sulla differenziabilitá e l'analiticitá delle estremali degli integrali multipli regolari. *Mem. Acad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)* **3** (1957), 25–43.

[10] Vasseur, A.: Regularity criterion for 3D Navier–Stokes equations in terms of the direction of the velocity. *Appl. Math.* **54** (2009), no. 1, 47–52.

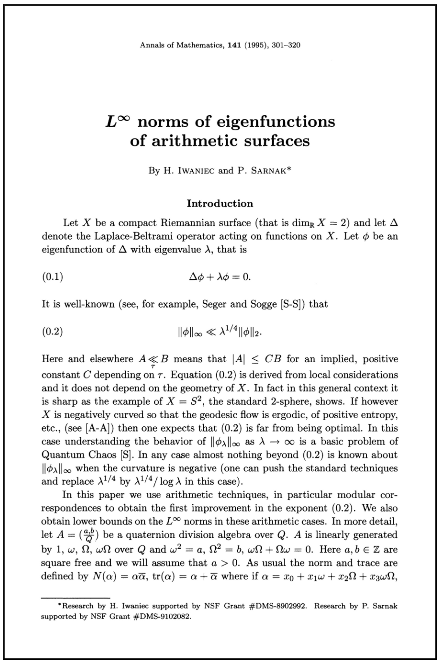
Fernando Chamizo (Madrid, Spain, 1967) obtained his PhD in mathematics at the University Autónoma of Madrid in 1994 (advisor: Antonio Córdoba), and currently is a professor at its Department of Mathematics. His main research area is analytic number theory.

Amplification and Quantum Chaos

Fernando Chamizo*

1. The paper and its importance to me

I am writing about the paper by H. Iwaniec and P. Sarnak entitled L^∞ *norms of eigenfunctions of arithmetic surfaces* published in Annals of Mathematics (see in [11] the full reference). It establishes an upper bound for the absolute value of the eigenfunctions of the Laplacian (more properly, of the Laplace-Beltrami operator) on some Riemann surfaces with arithmetic significance. This is a fundamental problem in the so-called *Arithmetic Quantum Chaos* (after [18]), a branch of Mathematics that employs number theoretical methods to answer some questions coming from Physics.



Annals of Mathematics, **141** (1995), 301–320

L^∞ norms of eigenfunctions of arithmetic surfaces

By H. Iwaniec and P. Sarnak*

Introduction

Let X be a compact Riemannian surface (that is $\dim_{\mathbb{R}} X = 2$) and let Δ denote the Laplace-Beltrami operator acting on functions on X. Let ϕ be an eigenfunction of Δ with eigenvalue λ, that is

$$\Delta\phi + \lambda\phi = 0. \tag{0.1}$$

It is well-known (see, for example, Seger and Sogge [S-S]) that

$$\|\phi\|_\infty \ll \lambda^{1/4}\|\phi\|_2. \tag{0.2}$$

Here and elsewhere $A \underset{\tau}{\ll} B$ means that $|A| \le CB$ for an implied, positive constant C depending on τ. Equation (0.2) is derived from local considerations and it does not depend on the geometry of X. In fact in this general context it is sharp as the example of $X = S^2$, the standard 2-sphere, shows. If however X is negatively curved so that the geodesic flow is ergodic, of positive entropy, etc., (see [A-A]) then one expects that (0.2) is far from being optimal. In this case understanding the behavior of $\|\phi_\lambda\|_\infty$ as $\lambda \to \infty$ is a basic problem of Quantum Chaos [S]. In any case almost nothing beyond (0.2) is known about $\|\phi_\lambda\|_\infty$ when the curvature is negative (one can push the standard techniques and replace $\lambda^{1/4}$ by $\lambda^{1/4}/\log\lambda$ in this case).

In this paper we use arithmetic techniques, in particular modular correspondences to obtain the first improvement in the exponent (0.2). We also obtain lower bounds on the L^∞ norms in these arithmetic cases. In more detail, let $A = (\frac{a,b}{Q})$ be a quaternion division algebra over Q. A is linearly generated by $1, \omega, \Omega, \omega\Omega$ over Q and $\omega^2 = a$, $\Omega^2 = b$, $\omega\Omega + \Omega\omega = 0$. Here $a, b \in \mathbb{Z}$ are square free and we will assume that $a > 0$. As usual the norm and trace are defined by $N(\alpha) = \alpha\overline{\alpha}$, $\mathrm{tr}(\alpha) = \alpha + \overline{\alpha}$ where if $\alpha = x_0 + x_1\omega + x_2\Omega + x_3\omega\Omega$,

*Research by H. Iwaniec supported by NSF Grant #DMS-8902992. Research by P. Sarnak supported by NSF Grant #DMS-9102082.

There are other papers more influential to me but, as far as I remember, this was the first time in my career that I understood entirely a contemporary paper in a top-level journal like Annals of Mathematics. Moreover, it dealt with a very young topic (I was lucky attending some of the first lectures on Arithmetic Quantum Chaos) that I considered highly technical. I felt self-confident but I would not have understood anything without the expertise guidance of Professor Iwaniec who, by the way, gave me the preprint and whose explanations in his book [9] published in the Biblioteca de la Revista Matemática Iberoamericana, were crystal clear.

Besides, I have chosen this paper because it gives me the opportunity of writing some lines about the beautiful *amplification method* and exposing some basic ideas about the spectral theory of automorphic forms escaping from the technical details.

*fernando.chamizo@uam.es.

2. The amplification method

Let me employ rigorlessly the term *harmonic* to mean something capable of being included in some kind of spectral resolution.

Spectral completeness is crucial in analysis and in its applications to number theory. Usually it turns out to be important and difficult to single out the contribution of only one harmonic. A natural approach resides in using peak functions in the spectral resolution to select a range of harmonics, dropping all except one by some instance of positivity to get an upper bound.

The amplification method is a simple and effective method (or a trick working several times, if we accept the definition in [17, p. 117]) that breaks the analytic barrier established by the uncertainty principle in some conspicuous situations. To my knowledge, it was originally introduced in [8]. The key idea is the definition of an *amplifier*, playing the role of the peak function, given by the square of a (usually short) linear form depending on the harmonics. Opening the square we expect interferences (quasi-orthogonality) but on the other hand, with a suitable choice of the variables of the form we can amplify the contribution of a single harmonic. This has some slight resemblances with Selberg's technique in sieve theory in which squares of linear forms are employed to imitate Möbius μ function minimizing upper bounds.

To specify the idea, I prefer to go through the note [1] that is the simplest application to my taste of the amplification method.

Dirichlet characters to modulus q are the multiplicative periodic functions $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$ such that $|\chi(n)| = 1$ if $\gcd(n,q) = 1$ and $\chi(n) = 0$ otherwise. They define homomorphisms from the multiplicative group of $\mathbb{Z}/q\mathbb{Z}$ onto S^1 and they are the commonly employed harmonics to analyze multiplicative periodic functions. An important problem in number theory is to give non-trivial upper bounds for $S(\chi,N) = \sum_{n=1}^{N} \chi(n)$ with N very small in comparison with q. In [1] it is considered this problem for $\chi = \chi_{r_0}\chi_{s_0}$ where the notation χ_k means a Dirichlet character to modulus k. Using the orthogonality relations and some known character sum bounds it is possible to prove

$$\sum_{\psi} \Big| \sum_{l \le L} c_l \psi(l) \Big|^2 \big| S(\psi\chi_{s_0}, N)\big|^2 \ll \|\vec{c}\|^2 F(N, L, r_0, s_0)$$

for certain F and with ψ running through all characters to the modulus r_0. We are not interested in an average bound of this kind but in selecting the case $\psi = \chi_{r_0}$. To this end one takes $c_l = \overline{\chi_{r_0}}(l)$, then $\big|\sum_{l\le L} c_l\psi(l)\big|^2$ amplifies the value of $\big|S(\chi_{r_0}\chi_{s_0}, N)\big|^2$ multiplying it by something close to L^2. Even more, at least conjecturally the orthogonality suggests that the amplifier is very small for $\psi \neq \chi_{r_0}$ and dropping these terms is not a big loss. In this way the method beats uncertainty principle and indeed beats the best known bounds for $S(\chi, N)$ in some ranges.

3. Arithmetic Quantum Chaos

In quantum mechanics the classical well-determined trajectories are substituted by something blur described by a wave function and the square of its absolute value gives, according to Copenhagen interpretation, a probability of detecting the particle. Using spectral analysis these wave functions are expressed in terms of eigenfunctions of a second order self-adjoint differential operator and the eigenvalues correspond to energy levels.

Quantum chaos asks about how some properties of classical dynamics are reflected in the quantum picture, specially those corresponding to chaotic features. It is a quite open subject and surprisingly arithmetic has provided a number of results, examples and counterexamples [2, 18].

To mention an interesting instance of Arithmetic Quantum Chaos and for the rest of the paper, we introduce the *modular surface*

$$X = \Gamma\backslash\mathbb{H} \qquad \text{where } \Gamma = \left\{ z \mapsto \frac{az+b}{cz+d} \quad \text{with} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \right\}$$

and $\mathbb{H}$ is the Poincaré upper half plane with the area element (invariant measure) $d\mu(x+iy) = y^{-2}\ dx\ dy$. The group Γ can also be described as generated by the translation $z \mapsto z+1$ and the inversion $z \mapsto -1/z$.

Laplace-Beltrami operator in X has a discrete spectrum $\{\lambda_j\}_{j=0}^{\infty}$ (and also a continuous spectrum). The corresponding L^2-normalized eigenfunctions are called *Maass wave forms* $\{u_j\}_{j=0}^{\infty}$. Conjecturally the multiplicity is one and they are uniquely determined. A recent major advance in Arithmetic Quantum Chaos is the proof of the QUE conjecture (see [14, 5, 21]) that reads

$$|u_j(z)|^2\ d\mu(z) \to d\mu(z) \qquad \text{as} \qquad \lambda_j \to \infty.$$

In this acronym Q stands for Quantum, because we are dealing with eigenfunctions; U stands for unique, because the limit was known for a subsequence and one conjectures that it is the same for any other subsequence; and E stands for ergodicity, because it shows equidistribution when the energy grows (large eigenvalues).

The proof of QUE depends heavily on the arithmetic symmetries of X through the Hecke operators and on some techniques from analytic number theory like sieve methods and subconvexity bounds for L-functions.

4. The result and its proof

The funny shape of atomic orbitals shows that eigenfunctions can have large peaks. Quantum chaos heuristic suggests that the negative curvature prevents hyperbolic surfaces from similar examples and we expect a kind of boundedness of the eigenfunctions when the energy grows as a manifestation of the chaotic geodesic flow. This boundedness opposes *scarring*, the situation of physical significance in which large values are concentrated along a curve [4].

We expect for z in a compact set $K \subset X$

$$|u_j(z)| \le C_{\epsilon,K}\lambda_j^{\epsilon} \qquad \text{for every } \epsilon > 0,$$

where $u_j(z)$ is a Maass wave form with eigenvalue λ_j (see p. 178 of [10], the second edition of [9], for the lack of uniformity in z). In the paper it is proved that this is true for every $\epsilon > 5/24$ and that it does not extend to $\epsilon = 0$, revealing the statistical nature of QUE. We focus on the proof of the first part.

The spectral analysis in $L^2(X)$ implies a kind of Poisson summation formula, sometimes called *pretrace formula*, that reads

$$\sum_{\gamma\in \mathrm{SL}_2(\mathbb{Z})} K(z,\gamma w) = \sum_j h(\lambda_j)u_j(z)\overline{u_j(w)} + \cdots$$

where the dots represent the contribution of the continuous spectrum and K and h are related by means of an integral transform. Note that the right hand side contains spectral information and the left hand side contains a certain geometric information (Selberg trace formula captures in a very elegant and appealing way the duality between analysis and geometry embodied in the pretrace formula [10]). A suitably chosen peak function h would give, taking $z = w$, a bound for $|u_j(z)|^2$. The idea is to employ the amplification method to beat it.

The Hecke operators T_n are some arithmetically defined operators such that $T_n u_j(z) = \eta_j(n)u_j(z)$ (we depart from the usual notation to avoid confusions). A fundamental point is that the numbers $\eta_j(n) \in \mathbb{R}$ satisfy the multiplicative relation

$$\eta_j(n)\eta_j(m) = \sum_{d|\gcd(n,m)} \eta_j\Big(\frac{nm}{d^2}\Big).$$

Consider the expression

$$S = \sum_j h(\lambda_j)A_j|u_j(z)|^2 \qquad \text{with} \qquad A_j = \Big|\sum_{n\le N} a_n\eta_j(n)\Big|^2.$$

Opening the square in the amplifier A_j, the multiplicative relation expresses S as a sum of Hecke operators applied to the pretrace formula, giving an upper bound for S (of course one needs to manipulate the exact definition of the Hecke operators that we do not include here). It is conjectured but not known that A_j is large for $a_n = \eta_j(n)$. To avoid this insidious difficulty the authors appeal again masterly to arithmetical properties: Choosing $a_n = -1$ if n is the square of a prime, $a_n = \eta_j(n)$ if n is a prime less than $\sqrt{N}$, and $a_n = 0$ otherwise. The multiplicative property assures $\eta_j(p)^2 - \eta_j(p^2) = 1$ for p prime and the choice of a_n gives readily by the prime number theorem $A_j \sim \sqrt{N}/2\log N$. In this way the value of $|u_j(z)|^2$ is amplified by this amount. For other eigenfunctions, the amplifier is expected to be small (due to some independence difficult to quantify) and they are anyway disregarded by positivity. Combining all the estimates the result is proved.

The beauty of the paper stems from the explotation of the arithmetical properties in an analytic context. We only know so strong bounds for eigenfunctions in hyperbolic surfaces using these techniques.

5. A glimpse on Kloostermania

Maass wave forms and their spectral theory play a role on the broader picture of Langlands program and it can be employed to justify the interest of number theorists on these topics but my little understanding of this point of view, forces me to focus on the part that boosted the subject and is still fruitful. I have recovered for the title the word Kloostermania that was in use 20 years ago to express the boom of this approach (see the comments in [3] and [16]).

Kloosterman sums are defined as

$$S(m,n;c) = \sum_{\substack{k=1\\(k,c)=1}}^{c} e^{2\pi i(mk+nk')/c} \qquad \text{with} \quad kk' \equiv 1 \pmod{c}.$$

They appear naturally when studying the distribution of the inverses modulo c. For instance, to localize simultaneously a class and its inverse one uses $\sum w_1(k)w_2(k')$ for some bump functions w_1, w_2 and their Fourier series expansions lead to Kloostermann sums. Historically they appeared in a different context, in a version of the circle method due to H. D. Kloostermann [12]. He gave a non-trivial bound for $S(m,n;c)$ and two decades later A. Weil [22] got $|S(m,n;c)| \le d(c)c^{1/2}$ for m, n and c coprimes and $d(c)$ meaning the number of divisors, which is optimal in a certain sense according to average results.

On the other hand H. Maass introduced in 1949 [15] non-holomorphic automorphic forms (in particular Maass wave forms) in connection with the study of real number fields. Automorphic forms (from the arithmetical point of view) are linked to the group $\mathrm{SL}_2(\mathbb{Z})$, defined by the determinant equation $ad - bc = 1$. Fixing c and d (the lower row of the matrix) a and b are determined by a choice of the inverse of d modulo c. With this idea in mind one can cook up special automorphic forms (Poincaré's series) having Kloosterman sums as Fourier coefficients. Their generating properties allowed A. Selberg in 1965 [20] to bound Fourier coefficients of "general" automorphic forms (in the narrow arithmetic sense) through Weil bound.

The breakthrough came with a paper by N. V. Kuznetsov [13]. The name *Kloostermania* was introduced by M. N. Huxley who wrote a nice early survey explaining the main ideas [6]. Instead of writing Kuznetsov formula fully I prefer just to outline the main terms

$$\sum_c \frac{S(m,n;c)}{c} f\left(\frac{4\pi\sqrt{|mn|}}{c}\right) = \sum_j \widetilde{f}(t_j)\bar{\rho}_j(m)\rho_j(n) + \dots$$

where $f \in C_0^\infty[0,\infty)$, $\widetilde{f}$ is a certain Bessel transform and $\{\rho_j(n)\}_{n\in\mathbb{Z}}$ are the Fourier coefficients of the Maass wave form $u_j(z)$. It can be employed in both directions. Our knowledge in spectral theory applied to the right hand side implies

$$\left|\sum_{c\le x} \frac{S(m,n;c)}{\sqrt{c}}\right| < Cx^{2/3}\log x.$$

This paradoxical improvement on Weil optimal bound comes from the variation of the sign of $S(m,n;c)$. In arithmetical terms, we can control the distribution of inverses for varying moduli and it has a profound impact on many problems [7, 16, 19].

Acknowledgements. I am grateful to Adrián Ubis for his valuable suggestions. I also appreciate the encouragement and patience of E. Valenti. Finally, I would like to take this opportunity to show my gratitude to Professor Iwaniec for his influence on my career beyond any of his papers.

References

[1] Friedlander, J. B. and Iwaniec, H.: A note on character sums. In *The Rademacher legacy to mathematics (University Park, PA, 1992)*, 295–299. Contemp. Math. **166**. Amer. Math. Soc., Providence, RI, 1994.

[2] Gutzwiller, M. C.: Physics and Selberg's trace formula. In *The Selberg trace formula and related topics (Brunswick, Maine, 1984)*, 215–251. Contemp. Math. **53**. Amer. Math. Soc., Providence, RI, 1986.

[3] Heath-Brown, D. R.: Arithmetic applications of Kloosterman sums. *Nieuw Arch. Wiskd. (5)* **4** (2000), no. 1, 380–384.

[4] Heller, E. J.: Bound-state eigenfunctions of classically chaotic Hamiltonian systems: scars of periodic orbits. *Phys. Rev. Lett.* **53** (1984), no. 16, 1515–1518.

[5] Holowinsky, R. and Soundararajan, K.: Mass equidistribution for Hecke eigenforms. *Ann. of Math. (2)* **172** (2010), no. 2, 1517–1528.

[6] Huxley, M. N.: Introduction to Kloostermania. In *Elementary and analytic theory of numbers (Warsaw, 1982)*, 217–306. Banach Center Publ. **17**. PWN, Warsaw, 1985.

[7] Iwaniec, H.: Promenade along modular forms and analytic number theory. In *Topics in analytic number theory (Austin, TX, 1982)*, 221–303. Univ. Texas Press, Austin, TX, 1985.

[8] Iwaniec, H.: The spectral growth of automorphic L-functions. *J. Reine Angew. Math.* **428** (1992), 139–159.

[9] Iwaniec, H.: *Introduction to the spectral theory of automorphic forms.* Biblioteca de la Revista Matemática Iberoamericana. Revista Matemática Iberoamericana, Madrid, 1995.

[10] Iwaniec, H.: *Spectral methods of automorphic forms.* Graduate Studies in Mathematics **53**. American Mathematical Society, Providence, RI, second edition, 2002.

[11] Iwaniec, H. and Sarnak, P.: L^∞ norms of eigenfunctions of arithmetic surfaces. *Ann. of Math. (2)* **141** (1995), no. 2, 301–320.

[12] Kloosterman, H. D.: On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$. *Acta Math.* **49** (1926), 407–464.

[13] Kuznetsov, N. V.: The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums. *Mat. Sb. (N.S.)* **111** (1980), 334–383.

[14] Lindenstrauss, E.: Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)* **163** (2006), no. 1, 165–219.

[15] Maass, H.: Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen. *Math. Ann.* **121** (1949), 141–183.

[16] Michel, P.: Some recent applications of Kloostermania. In *Physics and number theory*, 225–251. IRMA Lect. Math. Theor. Phys. **19**. Eur. Math. Soc., Zürich, 2006.

[17] Petkovšek, M., Wilf, H. S. and Zeilberger, D.: $A = B$. A.K. Peters Ltd., Wellesley, MA, 1996.

[18] Sarnak, P.: Arithmetic quantum chaos. In *The Schur lectures (Tel Aviv, 1992)*, 183–236. Israel Math. Conf. Proc. **10**. Bar-Ilan Univ., Ramat Gan, 1995.

[19] Sarnak, P.: Kloosterman, quadratic forms and modular forms. *Nieuw Arch. Wiskd. (5)* **1** (2000), no. 4, 385–389.

[20] Selberg, A.: On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, 1–15. Amer. Math. Soc., Providence, RI, 1965.

[21] Soundararajan, K.: Quantum unique ergodicity for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. *Ann. of Math. (2)* **172** (2010), no. 2, 1529–1538.

[22] Weil, A.: On some exponential sums. *Proc. Nat. Acad. Sci. USA* **34** (1948), 204–207.

Sun-Yung Alice Chang (Cian, China, 1948; she grew up in Taiwan) obtained her PhD in mathematics at the University of California, Berkeley, in 1974 (advisor: Donald E. Sarason). She is Professor at Princeton University. Her research interests include the study of geometric types of nonlinear partial differential equations and problems in isospectral geometry.

In the search of Q-curvature

Sun-Yung Alice Chang*

In my career, there have been a number of outstanding papers by distinguished mathematicians which have greatly influenced the way I think about mathematics. For example, during the years 1971–1974, when I was a graduate student studying under the supervision of Donald Sarason at the University of California, Berkeley, the whole field of analysis was energized by the result of Charlie Fefferman and Elias Stein on the duality of H^1 and BMO. In particular, the linking of their result to the work of Carleson on the Corona Problem have electrified a whole generation of mathematicians working in the field of real and complex analysis. This greatly influenced me as a young student studying the topic of bounded analytic functions in the complex plane.

Here I would like to share an experience I have had, that by sheer good luck, after many years working on a topic, when I learned a result which has since played a most important role in my work.

Starting from mid 1980s, I began to do some joint work with my husband Paul Yang –who was one of my fellow graduate students and who was trained as a geometer. We had started seriously discussing mathematics only after we had known each other for over ten years and eventually found some common interests in a problem. The problem we were studying is the problem known as the "Nirenberg problem" —when can a function K defined on the sphere be identified as the Gaussian curvature of a metric on the sphere? Denote the metric by $\hat{g} = e^{2w}g$, where g denotes the surface measure on the sphere. Nirenberg's problem can be formulated as a PDE problem, that is, when does the non-linear PDE

$$-\Delta_g w + K_g = K e^{2w} \tag{1}$$

allows a solution w on the sphere? (Here on the sphere $K_g \equiv 1$, $K = K_{\hat{g}}$).

One of main tools Paul and I have used to study the problem is a sharp borderline version of the Sobolev inequality which Jurgen Moser has introduced earlier to study this problem; inequality nowadays called Moser–Trudinger inequality. It is still one of the most important tools in the study of curvature variational problems in both Riemannian and Kähler geometry.

The importance of Gaussian curvature in the theory of surfaces is partly due to the role that it plays in the Gauss–Bonnet formula and in the Uniformization Theorem. In the late 1980s to mid 1990s, in the mathematical

*chang@math.princeton.edu.

community there was also an intensive study of the "Yamabe problem" on manifolds of dimension $n \geq 3$. The problem can also be thought as a generalization of the uniformization theorem —that in each conformal class of metrics, one can find a representative, whose scalar curvature is a constant function. There are the celebrated important works of the solution of the problem by Yamabe, Trudinger, Aubin and Schoen. On the other hand, the sign of the integral of scalar curvature alone does not classify manifolds. In the meantime, through the work of Tom Branson [1], Paul and I began to pay attention to a new type of curvature called the Q-curvature, which was initially defined on four manifolds. It is related to a fourth-order differential operator P discovered in 1983 by S. Paneitz [11]; the relation of the pair (P, Q) on four manifolds is very similar to that of the pair $(-\Delta, K)$ on 2-surfaces:

$$P_g w + Q_g = Q_{\hat{g}} \, e^{4w} \tag{2}$$

on four manifolds, where $\hat{g} = e^{2w} g$. The Q-curvature is defined as

$$Q \, = \, \frac{1}{6} \, (-\Delta R + R^2 - 3|Ric|^2), \tag{3}$$

where R denotes the scalar curvature and Ric the Ricci curvature on the Riemannian manifold.

By the Gauss-Bonnet formula, the integral of Q is a conformal invariant; and through the work of Matt Gursky [10] and others, the sign of this integral has a geometric and topological implication on the four manifold. It also turns out that on general n manifolds, there exists a pair of (P, Q) of n-th order operators and curvatures as discovered by Graham, Jenny, Mason and Sparling [8] with their construction based on an earlier work of C. Fefferman and R. Graham [6].

Many of the tools in the study of Gaussian curvature can be extended to study the Q-curvature and the effort still is going on today. There are questions in the subject which really puzzled me for a long time.

On compact manifolds without boundary, the divergence term ΔR in the expression of Q integrates to zero. Thus the term does not contribute to the integral of Q; while in terms of the number of differentiation of the metric, it is a 4-th order hence the highest order term in Q, what role does it really play?

Is there a geometric meaning in the curvature term Q? For example, the sign and the size of the scalar curvature compares the volume of the ball near the point to that of a Euclidean ball, can one ask what does the Q-curvature measure? One can also ask the same question for the quadratic polynomial $(R^2 - 3|Ric|^2)$ which appears in the expression of the Q-curvature.

In 1998–99, after having been a faculty at UCLA for over 16 years, I moved to Princeton. Almost immediately after my move, I happened to have a talk with Jeffrey Viaclovsky, a graduate student at Princeton, who at the time had just written his thesis. In this thesis, he pointed out, through the angle of conformal geometry and Cartan's computation, that the full Riemannian curvature R_m on a Riemannian manifold (M^n, g) can be decomposed into two parts:

$$R_m(g) \ = \ W_g \ \oplus \ A_g \owedge g,$$

where W_g denotes the Weyl tensor and A_g denotes the Schouten tensor of the metric g:

$$A_g = \frac{1}{n-2}\Big(Ric_g - \frac{R_g}{2(n-1)}\,g\Big).$$

Since Weyl curvature is a pointwise invariant under the conformal change of metric (g to $\hat{g}$), all the information of the conformal change of metric of the full curvature tensor of R_m is contained in the change from A_g to $A_{\hat{g}}$. This led Viaclovsky [12] to study in his thesis the functional F_k

$$F_k : \ g \to \int \sigma_k(A_g) dv_g$$

for $k < \frac{n}{2}$. Here $\sigma_k(A_g)$ denotes the k elementary symmetric function of the eigenvalues of A_g; *e.g.*, when $k = 1$, $\sigma_1(A_g) = \frac{1}{2(n-1)} R_g$, where R_g denotes the scalar curvature of the metric g (thus the study of F_1 is the study of the Yamabe functional); when $k = n$, $\sigma_n(A_g) = \text{determinant of}\, A_g$.

It struck me that this notion is exactly what I had been searching for! As when $n = 4$ and $k = \frac{n}{2} = 2$ we have exactly

$$\sigma_2(A_g) = \frac{1}{24}({R_g}^2 - 3|Ric_g|^2), \quad \text{thus } Q = -\frac{1}{6}\Delta R \ + \ 4\sigma_2(A_g).$$

It also turns out that in this case, with the additional assumption that the scalar curvature is either positive or negative, the sign of $\sigma_2(A_g)$ is the same as that of the curvature tensor Ric_g, thus has a strong geometric implication.

One can then use the techniques developed in the study of fully nonlinear PDE (in particular the Monge–Ampère equations) to study the equation $\sigma_k(A_{\hat{g}}) =$ constant. This study has since become an active branch of research in itself.

Built on this piece of pure luck —knowing a crucial fact at a crucial point— Matt Gursky, Paul Yang and I continued our program on the study of the Q-curvature; this eventually led to some classification results of four manifolds [2, 3]; thus confirmed our earlier belief that $\int Q_g dv_g$ is an important conformally invariant quantity on four manifolds, the sign and the size of which both carry important geometric information.

Later on we continued to learn more about the connection of this conformal invariant to other invariants in the conformal compact Einstein setting through the work of Fefferman–Graham [6], Graham–Zworski [9]. In particular, we finally understood more about the role played by the 4-th order curvature term ΔR in the expression of the Q-curvature; that in the setting of compact manifold with boundary (or in conformally compact manifold with conformal infinity) it links the behavior of the metric on the boundary to its interior. We are able to understand some global invariant quantity on the manifold in the more recent work of Chang–Qing–Yang [5], Graham–Juhl [7] and Chang–Fang [4]. The study of this very special geometric invariant has been quite a fruitful adventure for me!

References

[1] Branson, T.: Differential operators canonically associated to a conformal structure. *Math. Scand.* **57** (1985), 293–345.

[2] Chang, S.-Y. A., Gursky, M. J. and Yang, P. C.: An equation of Monge-Ampère type in conformal geometry, and four-manifolds of positive Ricci curvature. *Ann. of Math. (2)* **155** (2002), 709–787.

[3] Chang, S.-Y. A., Gursky, M. J. and Yang, P. C.: A conformally invariant sphere theorem in four dimensions. *Publ. Math. Inst. Hautes Études Sci.* **98** (2003). 105–143.

[4] Chang, S.-Y. A. and Fang, H.: On a new class of conformal variational functional. *Int. Math. Res. Notices* (2008), article ID rnm016, 16 pages.

[5] Chang, S.-Y. A., Qing, J. and Yang, P. C.: On the renormalized volumes for conformally compact Einstein manifolds. *J. Math. Science* **149** (2008), no. 6, 1755–1769 (English translation).

[6] Fefferman, C. and Graham, C. R.: Conformal invariants. Élie Cartan et les Mathématiques d'aujourd'hui. *Asterisque* (1985), 95–116.

[7] Graham, C. R. and Juhl, A.: Holographic formula for Q-curvature. *Adv. Math.* **216** (2006), 841–853.

[8] Graham, C. R., Jenne, R., Mason, L. and Sparling, G.: Conformally invariant powers of the Laplacian, I: existence. *J. London. Math. Soc. (2)* **46** (1992), 557–565.

[9] Graham, C. R. and Zworski, M.: Scattering matrix in conformal geometry. *Invent. Math.* **152** (2003), no. 1, 89–118.

[10] Gursky, M. J.: The Weyl functional, deRham cohomology, and Kähler–Einstein metrics. *Ann. of Math. (2)* **148** (1998), 315–337.

[11] Paneitz, S.: A quartic conformally covariant differential operator for arbitrary pseudo-Riemannian manifolds. Preprint, 1983.

[12] Viaclovsky, J. A.: Conformal geometry, Contact geometry and the Calculus of Variations. *Duke Math. J.* **101** (2000), no. 2, 283–316.

Antonio Córdoba (Murcia, Spain, 1949) obtained his PhD in mathematics at the University of Chicago in 1974 (advisor: Charles Fefferman). Since 1979, he has been Full Professor at the Universidad Autónoma of Madrid. His research fields are harmonic analysis, number theory, PDEs and mathematical physics. Antonio is Co-Director of *Revista Matemática Iberoamericana*.

All those rectangles

Antonio Córdoba*

Among the different publications that I could have considered to fill the goals of this issue, namely, to sustain that papers matters, there are three which have specially influenced my early mathematical career:

- "Note on the differentiability of multiple integrals", *Fund. Math.* **25** (1935), by B. Jessen, J. Marcinkiewicz and A. Zygmund;
- "Extremum problems with subsidiary conditions", by F. John (published in *Studies and Essays*, Interscience, NY, 1948);
- and "The multiplier problem for the ball", *Ann. of Math.* **94** (1971), by C. Fefferman.

Although one generally agrees with Simone Signoret who wrote that to be nostalgic is always a mistake, let me mention that during the academic year 71–72, I was a beginning graduate student at the Mathematics Department of the University of Chicago, but the years before, as an undergraduate at the Universidad Complutense of Madrid, was mostly fascinated by abstract algebra and Grothendick's approach to algebraic geometry. Nevertheless, I was fortunate to attend a series of lectures given by Antoni Zygmund in Madrid in 1971, and also by Alberto Calderón, the year before, where he presented his results on uniqueness for the Cauchy problem and his theory of pseudodifferential operators. Miguel de Guzmán encouraged me to follow those lectures which somehow changed my life, among other things because I was offered a fellowship to become a graduate student at Chicago. But before arriving at that university, I was exposed and grasped a certain understanding about the role of singular integrals, Fourier multipliers, maximal functions, covering lemmas and all that.

As most students in the late sixties, I also participated in the effervescence of ideas and cultural changes which spread around many university campuses during those years, but that, in the particular case of Spain, were entangled with the fight for democracy. Regarding science, there was a publication which produced a great impact in our vision of research, namely *The structure of scientific revolutions* by T. Kuhn. We liked the idea of "paradigm" and the description of "ordinary" science as the search for new results following the established ones, together with some "contradictions" which, eventually, will built up enough evidence helping us to discover new

*antonio.cordoba@uam.es.

paradigms in a kind of revolutionary process. Naturally we were so young and naïf to consider the idea of a scientific hero as the one who would lead us in that project. At the university of Chicago, while following the graduate program and sharpening my analytical tools, I wrote an essay about Lebesgue's integral being one of the best established paradigms of modern mathematics, as it was the Calderón–Zygmund theory of singular integrals in Fourier Analysis. It was then when I paid an special attention to the first of the three articles mentioned above. Let me give a brief description of its content:

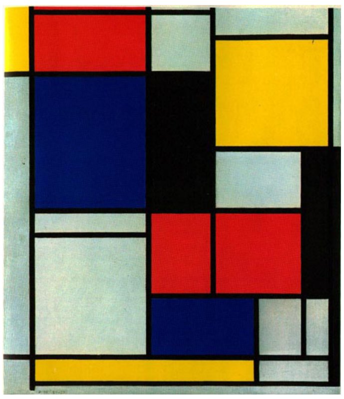
An important tool in Harmonic Analysis is the Hardy–Littlewood maximal function, defined over locally integrable functions f as the "sup" of its averages over all cubes containing x:

$$Mf(x) = \sup_{x \in Q} \frac{1}{\mu(Q)} \int_Q |f(y)|\, d\mu(y),$$

where μ denotes Lebesgue's measure in $\mathbb{R}^n$.

At the time when I read that paper it was well understood how quantitative estimates for Mf imply the fundamental theorem of calculus in the context of Lebesgue's measure theory. Furthermore, the properties of the maximal function were reduced to the understanding of the simple geometry of cubes in $\mathbb{R}^n$ throughout the so called Vitali's covering lemma. But also the Calderón–Zygmund decomposition of a function f at a level α yields a family of disjoint dyadic cubes $\{Q_j\}$ so that $|f(x)| \leq \alpha$ a.e. in the complement of their union $\cup Q_j$ and satisfies that

$$\alpha < \frac{1}{\mu(Q_j)} \int_{Q_j} |f(y)| \leq 2^n\, \alpha.$$



If we now paint each dyadic generation of those cubes with a different colour, we will obtain (in dimension 2) a Mondrian, and such collection of neoplasticist paintings will help us to understand the action of singular integrals upon the function f.

That is the case for classical (isotropic) singular integrals but not, for instance, for the double Hilbert transform. The paper of Jessen, Marcinkiewicz and Zygmund affords the needed extension introducing what they called strong maximal function

$$M_S f(x) = \sup_{x \in R \in \mathcal{B}_n} \frac{1}{\mu(R)} \int_R |f(y)|\, d\mu(y),$$

where the "sup" is now taken over all parallelepipeds containing the point x and whose sides are parallel to the coordinate axes ($\mathcal{B}_n$).

It was observed by E. Saks that Lebesgue's differentiation theorem failed for general locally integrable functions in $\mathbb{R}^n$, $n > 1$, when we substitute cubes by "rectangles" in $\mathcal{B}_n$. However, the work of J-M-Z showed us that the fundamental theorem of calculus remains true if we restrict our attention to functions which are locally in the space $L\log^+(L)^{n-1}(\mathbb{R}^n)$ and, furthermore, that this is the best possible result.

In their proof, J-M-Z used in a crucial manner the product structure of $\mathbb{R}^n$ but they left open the question about what kind of covering lemma was satisfied by the class $\mathcal{B}_n$.

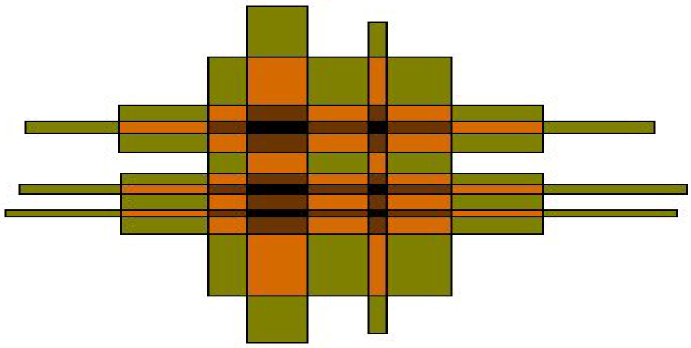
In the year 1975, in collaboration with Robert Fefferman, we were able to give a precise answer to that question showing that the family $\mathcal{B}_n$ satisfies an exponential type covering property:

Given a collection of parallelepipeds $\{R_\alpha\} \in \mathcal{B}_n$, one can select a countable family $\{R_j\}_{j=1,2\ldots}$ such that:

$$i) \quad \mu(\cup R_\alpha) \le C_n\, \mu(\cup R_j);$$

$$ii) \quad \int_{\cup R_j} \exp\Big(\sum\nolimits_k \chi_{R_k}(x)\Big)^{1/(n-1)} d\mu(x) \le C_n\ \mu(\cup R_j)$$

for some constant C_n.

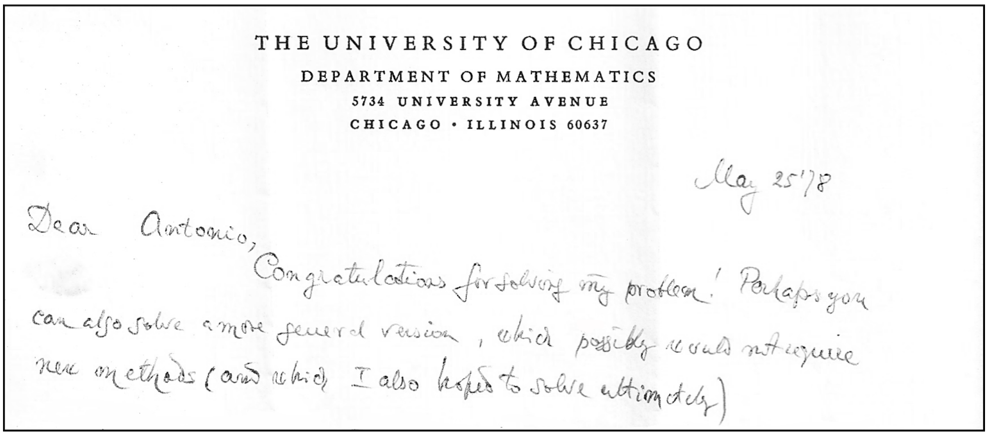


Exponential overlapping

As a continuation of J-M-Z, Antoni Zygmund proposed a problem which, for a long time, became an object of desire among harmonic analysts: Let us consider, in Euclidean space $\mathbb{R}^3$, the basis $\mathcal{B}_\phi$ of parallelepipeds whose sides are parallel to the coordinate axes but whose dimensions are given by $(s, t, \phi(s,t))$, $s, t \in \mathbb{R}^+$ and ϕ is positive and monotonic on each variable separately. The question asked by Zygmund was to decide if from the differentiation point of view $\mathcal{B}_\phi$ behaves like $\mathcal{B}_2$ or $\mathcal{B}_3$.

In 1978, I was able to solve Zygmund's problem showing that $\mathcal{B}_\phi$ differentiates integrals of functions which are locally in $L\log^+ L(\mathbb{R}^3)$. That year I was on leave from Princeton and had visited the universities of Paris (Orsay), Complutense (Madrid) and also the Mittag–Leffler Institute. It was at this last institution where I received a letter from Zygmund, dated at Chicago, which started with:

> Dear Antonio, May 25'78
>
> Congratulations for solving my problem! [. . .]



THE UNIVERSITY OF CHICAGO
DEPARTMENT OF MATHEMATICS
5734 UNIVERSITY AVENUE
CHICAGO • ILLINOIS 60637

May 25'78

Dear Antonio,
Congratulations for solving my problem! Perhaps you can also solve a more general version, which possibly would not require new methods (and which I also hope to solve ultimately)

and ended asking for the more general question about parallelepipeds in $\mathbb{R}^n$ whose dimensions are given by $(\phi_1(t_1, \dots, t_k), \dots \phi_n(t_1, \dots, t_k))$, where ϕ_j is monotonic in each variable separately.

In my answer, I told Zygmund that with that generality the conjecture was not true, but the important particular case $(t_1, \dots, t_{n-1}, \phi(t_1, \dots, t_{n-1}))$, with $n > 3$, was then, and still is, an intriguing open problem.

As I said before, my first meeting with Zygmund took place in Madrid, in June 1971. He was visiting the Complutense and I had recently been admitted to the graduate school at Chicago. I had studied parts of his monumental book on trigonometric series and considered him a living legend. Walking with Zygmund the streets of Madrid was for me an unforgettable experience and, among other advices, he gave me an excellent hint about how to get an appropriate apartment in the campus of the university of Chicago, very close to Eckhart Hall.

During the academic year 1973-74, Felix Browder, then chairman of the Math Department, asked me to be Zygmund's teaching assistant, which implied, among other tasks, that I was supposed to help him with the organization of the analysis seminar. It happened that Alberto Calderón was in Argentina and Charles Fefferman enjoyed a sabbatical, so I was the only available analyst. Besides the more rutinary tasks, Zygmund proposed me to fill the gaps for those Monday's afternoon (3:45 p.m.) when we did not have an invited speaker. To that end he provided me with a series of his favourite problems, and adequate references to be prepared for those occasions; he also told me many interesting anecdotes and stories about his colleagues, students and collaborators. We had many conversations, usually at his office, but also at his apartment where he offered me his excellent liquors.

One day, during the spring term of 1974, Zygmund said to me something like "Antonio, we have the same names, why you do not call me Antoni?" I was flattened because it was notorious that very few persons had that privilege and, frankly, was not easy for me at the beginning; nevertheless, with time, and a certain effort from my part, I learned to use it. But, of course, in the mathematical genealogy Zygmund is my great-grandfather, and the members of my generation, now on their sixties, have already experienced how one gets specially indulgent and tolerant with grandchildren.

The article of Fritz John about convex sets was the first serious mathematical work that I read directly from its original source. I did it following the suggestion made by Miguel de Guzmán in his lectures at the Universidad Complutense de Madrid, during the academic year 1969–70. Let me point out that, at that time, Miguel had just returned to Madrid, after completing his PhD at the University of Chicago under the advice of Alberto P. Calderón, and was beginning to display a very positive influence in the development of mathematics in Spain helping and encouraging the career of many Spanish mathematicians, like myself. He was then mainly interested in understanding the properties of different differentiation basis, and John's lemma was the adequate tool to reduce a general "convex basis" to the particular case of those consisting in parallelepipeds.

But the original proof given by F. John was rather complicated and Miguel asked for a more geometric and easier one. Antonio Gallego and I were among the students of that course that got interested in answering the question. First independently, and later joining our efforts, we read John's paper and figure out a more elementary proof which had the advantage of provide a better (sharp) constant. It was presented at the meeting of the Spanish Mathematical Society (RSME) celebrated in Murcia during the winter of 1970, and later published in the Springer *Lecture Notes about Differentiation of Integrals in* $\mathbb{R}^n$, written by Miguel in the year 1975. However, at that time I was totally unaware about the importance of mathematical journals and research articles. Reading textbooks was my main relationship with scientific literature and the idea of writing a paper and publishing that proof did not occurred to us.

Our proof of John's lemma, taken directly from the manuscript of 1970, is very simple. A kind of mathematical haiku written in elementary Euclidean geometry. But it was my first serious mathematical contribution and I can still remember the excitement and pleasure that it gave to me.

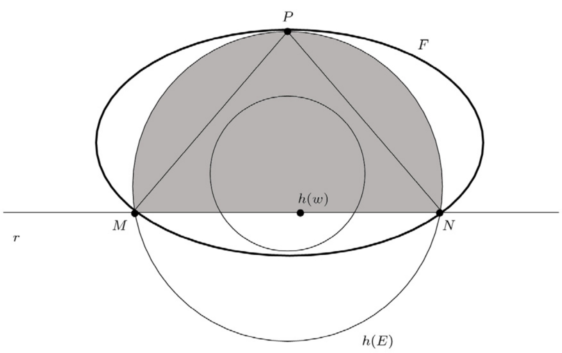
Lemma 1 *Let K be a convex, open and bounded set in the Euclidean space $\mathbb{R}^n$ and let E be an open ellipsoid of minimal volume containing K. Then the ellipsoid $\frac{1}{n}E$, obtained contracting E by the factor $1/n$ with respect to its centre, is contained in K.*

In the following I will sketch the proof in dimension 2.

First the existence of such an ellipse of minimal area containing K follows easily by a compactness argument. Next let us consider the following facts whose proof is immediate:

If C is an open circle in the plane and T is a triangle of maximal area inscribed in C, then, necessarily, must be equilateral. Conversely, if T is an open equilateral triangle and E is an ellipse of minimal area containing it, then E must be its circumscribed circle.

As a consequence we obtain that if T is the isosceles triangle XYZ ($XZ = YZ$) and h is an affine map fixing the vertices X and Y and mapping Z into W, so that the triangle XYW is equilateral, then the ellipse of minimal area containing XYZ is the inverse image by the mapping h of the circumscribed circle to the triangle XYW.



Suppose now that E is an ellipse of minimal area containing K and assume the hypothesis that $\frac{1}{2}E$ is not contained in K. There must be then a point w in the interior of $\frac{1}{2}E$ which is also placed at the boundary of K. Let h be an affine transformation such that the image $h(E)$ is a circle. Then $h(w)$ is located in the intersection of the interior of $\frac{1}{2}h(E)$ and the boundary of $h(K)$. Let us consider the supporting line r to the convex set $h(K)$ through the point $h(w)$ which intersect the circle $h(E)$ at the points M, N and consider also the isosceles triangle MNP (see figure).

But since $h(K)$ lies inside the shaded region and the isosceles triangle satisfies the conditions of our previous considerations, there must exists an ellipse F containing the shaded region (and therefore containing $h(K)$) which has strictly smaller area than $h(E)$. But then the inverse image of F by h

contains K and is an ellipse of smaller area than E in contradiction with our original hypothesis. Q.E.D.

As I said before, neither Antonio Gallego or myself were at that time aware of the importance of publications and the fundamental role played by John's lemma in different mathematical areas. We enjoyed reading John's article, getting involved in the search of a more geometrical proof and, of course, answering the question of Miguel de Guzmán, someone we admired and who was a constant reference for us. With the years I have learned that such proof is now the standard one for experts in convex sets theory, and that it belongs to the folklore literature. On a few occasions, however, we have received public credit for the authorship of the haiku, but that is by now totally irrelevant as it is very well expressed in a popular Spanish song:

Hasta que el pueblo las canta,
las coplas, coplas no son.
Y cuando el pueblo las canta,
las coplas del pueblo son.

Let me now turn to C. Fefferman 1972 paper. At that time it was a kind of bomb destroying the established conjecture which asserted that the ball multiplier in $\mathbb{R}^n$, $n \geq 2$, should be bounded on $L^p(\mathbb{R}^n)$, so long as

$$\frac{2n}{n+1} < p < \frac{2n}{n-1}.$$

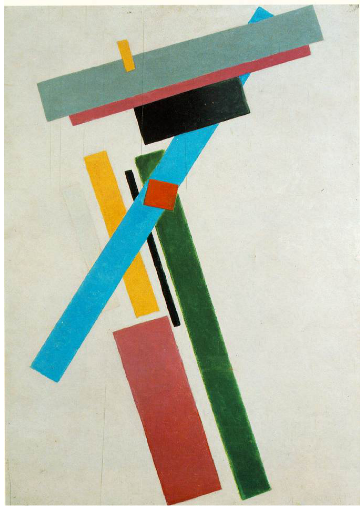
Fefferman showed that this was not the case and that, in fact, T is bounded only in $L^2(\mathbb{R}^n)$. The method was brilliant: based on a clever observation of Yves Meyer, Fefferman made use of the properties of Kakeya's set to disprove the conjecture. In the terminology of Littlewood that set was the real enemy for the ball multiplier to be bounded on L^p, $p > 2$.

Shortly after Fefferman's counterexample, L. Carleson and P. Sjölin proved that if we smooth out a little the characteristic function of the disc, namely if we consider the so-called Bochner–Riesz or spherical summation operators given by the formula

$$\widehat{T_\alpha f}(\xi) = (1 - |\xi|^2)_+^\alpha \; \widehat{f}(\xi)\,, \qquad \frac{1}{2} > \alpha > 0,$$

then, in dimension $n = 2$, T_α extends to a bounded operator on $L^p(\mathbb{R}^2)$ whenever

$$\frac{4}{3+2\alpha} < p < \frac{4}{1-2\alpha}.$$



Using a different approach, C. Fefferman obtained also the same result and published it as a short note in the *Israel Journal of Mathematics*. Those three papers had a direct and strong influence in my mathematical career because my thesis problem was, precisely, to understand their hidden connections. That task leaded me to study the properties of the Kakeya maximal function and to consider different kinds of "square functions" in order to extends the paradigm of C–Z theory to this more complicated setting, where cubes, or rectangles having sides parallel to the coordinate axes, are now substituted by rectangles with arbitrary directions, and the neoplasticists painting of Mondrian changes into examples of Malevich suprematism.

Williamstown, 1978

In September, 1978, in an important meeting in Harmonic Analysis organized by the AMS at Williamstown, I had the opportunity to deliver a plenary lecture on those topics. There I quoted the great J. E. Littlewood who, in 1964, had published an article on *The Scientific Speculates* developing the idea of the "enemy in mathematics". Following Littlewood, in order to obtain a good theorem, one must find first the most dangerous enemy of that property that one wants to be established, and then introduce the hypothesis needed to kill him. What remains is the theorem. Littlewood presents several examples of that principle, and applies it to the case of the "three lines theorem" in complex analysis. But he also speculates and use it to support that Lusin's conjecture should be false!

It is now a common place to say that Kakeya's set (*i.e.*, sets of measure zero containing a unit segment on each direction) are the enemies for the theory of Bochner–Riesz operators and for the understanding of the interaction of Hilbert transforms in different directions of the space. They are negligible from the point of view of Lebesgue measure theory, but not, in a certain sense, for the Fourier transform.

However, in dimension $n = 2$ we have achieved a rather complete theory where one of its main characters is the Kakeya maximal function:

$$\mathcal{M}_N f(x) = \sup_{x \in R} \frac{1}{\mu(R)} \int_R |f(y)| \, d\mu(y).$$

Here the "sup" is taken over all rectangles of eccentricity=bigger side/smaller side $= N$, but with arbitrary directions.

It happens that

$$\|\mathcal{M}_N f\|_2 \lesssim (\log N)^{1/2} \|f\|^2$$

and this estimate yields easily, among other facts, that a set containing a rectangle of dimensions $\varepsilon \times 1$ on every direction of the plane must have area greater than $C/|\log(\varepsilon)|$, implying that a Kakeya set in the plane has fractal dimension 2. But the case $n \geq 3$ remains as an outstanding open problem.

Understanding the ideas contained in those papers and trying to go further has been an important leitmotiv of my own research, and I believe that also of many others harmonic analysts. They have provided us with a suggestive plan to go beyond the classical C–Z theory of singular integrals, because either the singularity set of the kernels is more complicated, or they involve many components with symbols lying in different regions of phase space and whose L^p estimates need subtle cancellations. It is a very suggestive but difficult plan whose success seems to need new, and perhaps revolutionary, ideas.

To finish, let us point out that the paradigm of C–Z theory seems to indicate that a really interesting proof must always have a good stopping time. I believe that such a principle can be applied to many other aspects of life: Let me stop here.

David Drasin (Philadelphia, USA, 1940) obtained his PhD in mathematics at Cornell University in 1966 (advisors: W. H.J. Fuchs and C. J. Earle). He is currently Full Professor at Purdue University. His research fields are complex variables and potential theory.

Some mathematical influences

David Drasin*

1. Introduction

Many papers, lectures and contacts have inspired significant parts of my mathematical career, but two topics stand out, both having their roots in Picard's famous theorem:

If f is analytic in the complex plane $\mathbb{C}$ with $f \neq 0, 1$, then f is constant

(an alternate formulation is that if f is meromorphic nonconstant in the plane, it can have at most two 'Picard values').

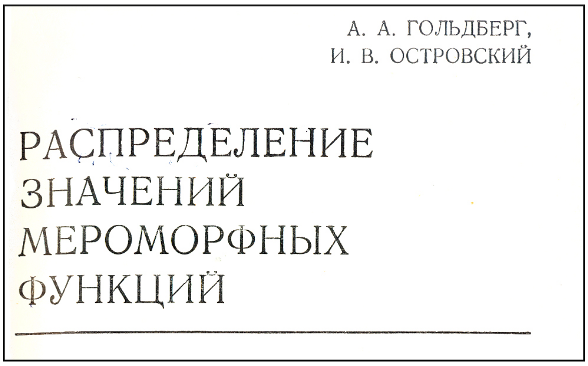
Figure 1: The Revista was not available at the time of this photograph.

The first theme (A) discusses the influence of [8], especially its Chapter 7, on my applications of quasiconformal mappings to value-distribution theory in $\mathbb{C}$. Rolf Nevanlinna developed his value-distribution theory with Picard's theorem in mind, as the title to his initial paper [11] makes clear. The second subject (B) focuses on the analogue of Picard's theorem in $\mathbb{R}^n$, $n > 2$. Here analytic/meromorphic functions must be replaced by K-quasiregular mappings, where $1 \leq K < \infty$ (and $K > 1$ in all nontrivial cases). I have already spent over a decade studying Rickman's surprising example [15] which, although valid only in three dimensions, indicates, in contrast to the situation when $n = 2$, that the number of Picard exceptional values depends on the distortion coefficient K, and can be large (if always finite).

These two themes illustrate different ways that a mathematical publication can be influential. That in (A) concerns material that had been well-integrated into the discipline for some decades, and I was lucky to study [8] with my own orientation and several possible applications in mind (only [4] is discussed here, but this technique is again a standard tool; [5] develops an alternative approach). In (B) the subject is less mature, and so the impact of [15] less obvious.

*drasin@math.purdue.edu. The author thanks O. Lehto, P. Pankka and P. Poggi-Corradini for helpful comments.

My generation learned classical Nevanlinna theory from [9] and [14]. Chapter 4 of [9] introduced us to some of A. A. Goldberg's recent work on Nevanlinna deficiencies. Thus it was natural around 1970 for Allen Weitsman and me to learn scientific Russian in order to study Goldberg's papers. This introduced us to exciting new ways to handle analytic functions. Questions that seemed to resist solution by constructing explicit analytic expressions (power series, infinite products) could be convincingly resolved by non-analytic exhaustions or Riemann surface arguments. What was especially exciting was his application of quasiconformal mappings to produce entire and meromorphic functions and uniformize Riemann surfaces. We were very fortunate that just at this time his carefully-prepared monograph [8] (written with I. V. Ostrovskii) appeared, and our first dividend was [3].


А. А. ГОЛЬДБЕРГ,
И. В. ОСТРОВСКИЙ

РАСПРЕДЕЛЕНИЕ
ЗНАЧЕНИЙ
МЕРОМОРФНЫХ
ФУНКЦИЙ

Figure 2: My 'secret weapon'.

While preparing this essay, I was reminded that quasiconformal mappings were applied to questions relating to Picard's theorem and value-distribution theory very soon after Grötzsch's paper (of 1928, considered the beginning of the subject), see §2. Teichmüller's celebrated habilitation thesis [18] gives the first formal presentation of the theory of quasiconformal mappings as a complete subject in itself, with starting point the application of quasiconformal mappings to the uniformization of some Riemann surfaces relevant to Nevanlinna theory (§2).

By the 1960s, however, interest in the West was centered on other applications of this theory, many of which Teichmüller himself introduced later (quadratic differentials, Riemann surfaces, extremal problems, Teichmüller spaces), and neither [18] nor this original motivation are mentioned in the survey [2]. Teichmüller had already published several papers in algebra and operator theory, and Nevanlinna's work was already well-known in Germany when Nevanlinna arrived from Finland to Göttingen for the 1936-7 academic year. Teichmüller spoke with Nevanlinna and attended his lectures [10], and then made important contributions to value-distribution theory (as well as [18]). Quasiconformal mappings with a variety of unanticipated applications became the thrust of many of his future papers: his interaction with quasiconformal mappings had a decisive impact on complex analysis.

In two dimensions a quasiregular mapping g, the theme of (B), may be factored as

$$g = f \circ \varphi,$$

where f is meromorphic in the plane φ a K-quasiconformal homeomorphism of $\mathbb{C}$ with $1 \leq K < \infty$. Hence Picard's theorem adapts immediately to planar quasiregular mappings.

In higher dimensions, this theory has its special character [16], and is distinct from that of several complex variables, although its (nonlinear) potential theory retains much of the flavor of that associated to the Laplacian in the plane. Rickman [16] proved that if $f : \mathbb{R}^n \to \mathbb{R}^n$ is K-quasiregular and non-constant, then $\mathbb{R}^n \setminus f(\mathbb{R}^n)$ has at most $q(K, n) < \infty$ points, and by now this is a nice application of nonlinear potential theory. More isolated at present is Rickman's stunning example [15]:

Let $q < \infty$ be given. Then there is $K < \infty$ and a (nonconstant) K-quasiregular mapping f on $\mathbb{R}^3$ with

$$f : \mathbb{R}^3 \to \mathbb{R}^3 \setminus \{a_1, a_2, \ldots a_q\}$$

(the $\{a_j\}_1^q$ may be assigned arbitrarily).

After nearly 30 years, it remains the only evidence that Rickman's form of Picard's theorem is precise, although at present is valid only in dimension 3 (an example of V. Zorich, based on the quasiregular analogue of the two-dimensional exponential function had led to the conjecture that $q(K, n) \equiv 1$ $(n \geq 2)$, and [16] begins with a discussion relating his example to that of Zorich). Rickman's paper is scrupulously written, in principle has almost no prerequisites (an indication of its originality) and uses surprisingly elementary techniques. It is the most impressive single work of mathematics I have read, one which requires mastering literally every line (see [7]). It remains a challenge to summarize its key ideas, to say nothing of using them in other situations.

Today's young mathematician has abundant opportunities to attend conferences and obtain materials (usually via the internet); when my career began, mail could be slow and contacts with mathematicians in much of the world could be very difficult or even impossible. I was very lucky to have the Purdue Mathematical Sciences Library as my mathematical portal. An unexpected catalyst to my first post-thesis work was encountering a paper on the New Journals shelf by the Chinese mathematicians Yang Lo and Zhang Guang Hou on normal families; many years had to pass before we could have personal contact. The book [8] also arrived unheralded to our library at a perfect time. Of course, all standard books in the subject ([19] is a rare exception) are now available in English.

2. Quasiconformal compositions in value-distribution theory

The 'modern' theory of value-distribution begins with Rolf Nevanlinna's fundamental papers [11, 12]. Nevanlinna obtained Picard's theorem as an immediate consequence of (in his words) "a general method to study the roots of the equation"

$$f(z) = a, \qquad a \in \hat{\mathbb{C}}. \tag{1}$$

Since the number of solutions to (1) will usually be infinite, Nevanlinna applies what is now considered standard potential theory to the δ-subharmonic function $\log|f|$, with f restricted to each disk $B(r) := \{|z| \le r\}$, and then takes appropriate limits as $r \to \infty$. Shelves of books and monographs are devoted to this subject, but his striking and easily-stated *defect relation* (2) (and its extension (3), a consequence of his *Second Fundamental Theorem*) readily confirms its power. The theory has many applications, but the focus here is only on (2) and (3).

Thus, let f be meromorphic (non-constant) in the z-plane. To each $a \in \hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ Nevanlinna associates a number $\delta(a)$, $0 \le \delta(a) \le 1$, such that $\delta(a) = 1$ if (1) has no solution (in which case the number of solutions to (1) is certainly 'maximally deficient'). Nevanlinna proved

$$\sum_{a \in \hat{\mathbb{C}}} \delta(a) \le 2. \tag{2}$$

This yields Picard's theorem, but offers infinitely many variants, since the possible values $\{\delta(a)\}$ form a continuum. Inequality (2) may be augmented to

$$\sum_{a} \delta(a) + \theta(a) \le 2, \tag{3}$$

where $0 \le \theta(a) \le 1$ measures the branching of f over a; M. Heins refers to (3) as the 'transcendental analogue' of the Riemann-Hurwitz relation.

After completing [3], I returned to Chapter 7 of [8]. Since I knew from [9] that (2) was sharp for entire functions (where $\delta(\infty) = 1$; Fuchs and Hayman did not consider the more general (3) and their methods failed in the meromorphic case), it was natural to suppose that restrictions (2) and (3) were best-possible in general. *Inverse problem:* given sequences $\{\delta_n\}, \{\theta_n\}, \{a_n\}, 0 \le \delta_n, \theta_n, 0 < \delta_n + \theta_n \le 1, a_n \in \hat{\mathbb{C}}$, find a meromorphic function f with

$$\delta(a_n) = \delta_n,\ \theta(a_n) = \theta_n; \qquad \delta(a) = \theta(a) = 0 \ \ (a \notin \{a_n\}). \tag{4}$$

The solution is in [4]; a short proof for (2) using different methods is in [5].

The treatment in [8], Chapter 7, presents a catalogue of examples from several authors which show (4) possible in many situations. It is written in a clear and leisurely style, with many examples and discussions that were helpful to us novices, unlike the very condensed outline in Chapter VIII of [19]. The examples begin with Nevanlinna [13] and lead to Goldberg's own contributions from the 1950s.

The solution to (4) in each of these cases has two steps. The first (α) was to create a Riemann surface $\mathfrak{R}$ which is presumed to be the image of the z-plane by the map f, the simplest example being the familiar spiraling surface associated to the exponential function, where $\delta(0) = \delta(\infty) = 1$. The surface would be constructed with appropriate branching over the $\{a_n\}$, but in enumerating these special cases, the procedure became increasingly rococo. The exposition is enriched by a superb account of Speiser graphs, for which [8] is still the best reference. The second step (β) is to find a conformal map $f : \mathbb{C} \to \mathfrak{R}$ such that when $\mathfrak{R}$ is exhausted by f-images of the $\{B(r)\}$, the data obtained solves (4).

Step (α) was present in all cases considered in [8]. Nevanlinna's surfaces [13] were generalizations of the exponential function, with equality in (2). Their uniformization relied on the theory of second-order differential equations for step (β), but this approach could not work in other situations. In retrospect, the paper of Ahlfors [1], in the same volume of *Acta* as [13], introduced the essential idea of constructing an explicit non-analytic exhaustion of the class of surfaces $\{\mathfrak{R}\}$ in [13]. These maps $g : \mathbb{C} \to \mathfrak{R}$ were no longer complex-analytic, but (in language not available at the time) were *quasiconformal with small average maximal dilatation.* It was easy to check that each g formally solved the inverse problem under consideration, and Ahlfors then applied his length-area technique to show that the analytic map $f : \mathbb{C} \to \mathfrak{R}$ imitated g sufficiently well that f also was a genuine meromorphic solution.

This procedure (which by 1938 had also been used by other authors) was formalized in Teichmüller's famous [18]. In his opening remarks, Teichmüller states as his goal to encompass these partial solutions in a unified orientation. This leads to the problem of transferring data from quasiconformal to analytic mappings (β). Using simplifications allowed by later authors, let $g : \mathbb{C} \to \mathfrak{R}$ be a $W^{1,1}$ function whose value-distribution formally satisfies (4) when data are computed with respect to the g-exhaustion. Set $\mu(z) = g_{\bar{z}}/g_z(z)$, and suppose that $\|\mu\|_\infty \leq k_0 < 1$ for some fixed k_0. (Note: $\mu \equiv 0 \iff g$ is analytic, and μ and K are related by $\|\mu\|_\infty = (K-1)/(K+1)$.) Obviously g satisfies (Beltrami equation)

$$g_{\bar{z}}(z) = \mu(z) g_z(z), \tag{5}$$

but there is also a homeomorphic solution to (5) on $\mathbb{C}$, say $\varphi(z)$, which may be taken to fix 0 and 1. If $\psi = \varphi^{-1}$, it is easy to check that $f := g \circ \psi$ is meromorphic, and the issue is to show that φ approximates the identity well enough that data for f transfer to g relative to $\{B(r)\}$-exhaustions.

Teichmüller encapsulates his principle on the first page of [18]: if

$$\int\int_{\{|z|>1\}} |\mu(z)| \frac{dxdy}{|z|^2} < \infty, \tag{6}$$

then $|\psi/\varphi| \to A,\ 0 < A < \infty$, as $z \to \infty$. This makes transferring data directly from g to f routine and applies to all surfaces considered in [8].

Unfortunately, (6) did not seem to hold for the surfaces studied in [3], but after the usual period of floundering, we noticed that our dilatations did satisfy

$$\int_0^{2\pi} |\mu(re^{i\theta})|\, d\theta \to 0 \qquad (r \to \infty) \tag{7}$$

An inspection of the proof of (6) showed (7) implied that for suitable functions $\sigma(r), \eta(r) = o(1)$ $(r \to \infty)$ that

$$B\Big((1-\eta(r))\int_1^r (1+\sigma(t))t^{-1}dt\Big) \subset \varphi(B(r)) \subset$$
$$\subset B\Big((1+\eta(r))\int_1^r (1+\sigma(t))t^{-1}dt\Big) :$$

again 'circles correspond to circles', and so the $\{\delta(a),\, \theta(a)\}$ transfer from g to f (however, in retrospect (7) had been analyzed earlier).

3. Quasiconformal modifications

Now let data

$$\{a_n\}, \{\delta_n\}, \{\theta_n\} \tag{8}$$

be given. From the repertoire in [8] (augmented by the classes in [3]), we can readily select a sequence of meromorphic functions $\{f_j\}$ whose 'Nevanlinna data' $\{\delta_{j,n}, \theta_{j,n}, a_{j,n}\}$ converge (as $j \to \infty$) to (8); in turn each f_j arises from a quasiconformal map $g_j : \mathbb{C} \to \mathfrak{R}_j$. This leads to another modification from the scheme of §2: instead of constructing one Riemann surface $\mathfrak{R}$ as some sort of limit of the $\{\mathfrak{R}_j\}$ and worrying about how $\mathfrak{R}$ is to be uniformized (as suggested in (α)), we directly produce $g : \mathbb{C} \to \hat{\mathbb{C}}$ which, as $r \to \infty$, mirrors g_j near $\{|z| = r\}$, so that $j \to \infty$ with r. In this situation (7) readily applies and the inverse problem is solved. Riemann surfaces never appear, but (7) is essential.

4. Quasiregular mappings

By 1986, most problems in classical Nevanlinna theory relating to (2) and (3) were either settled or considered hopelessly difficult (although important progress continued, albeit at a slower pace, *e.g.*, [17, 20, 21]; in addition, there was activity in other significant sub-areas). About this time the theory of quasiregular mappings in space had entered its adolescence. In many ways this subject was natural extension of the two-dimensional geometric and potential-theoretic aspects of complex function theory. We avoid formal definitions (see [16]), but note that in two dimensions Ahlfors in 1936 had already shown these mappings were the natural setting for his theory of covering surfaces.

This seemed an attractive area, since the relevant literature was modest, and its antecedent lay in the geometric aspects of complex-analytic functions in $\mathbb{C}$. Unlike the classical theory, it lacks a catalogue of functions already in the mathematical culture, although the Zorich function $f : \mathbb{R}^n \to \mathbb{R}^n \setminus \{0\}$ has been mentioned.

Let us discuss a special situation in two and three dimensions to see informally why the number of Picard values could well be different. With $x = (x_1, x_2)$ or $x = (x_1, x_2, x_3)$ the variable as appropriate, let $U_1 = \{|x| > 1\}$ and for $n = 2, 3$ take $U_2 = \{|x| < 1\} \cap \{x_n < 0\}$, $U_3 = \{|x| < 1\} \cap \{x_n > 0\}$. Then U_1, U_2, U_3 are neighborhoods of $y_1 = \infty, y_2 = -(1/2)e_n, y_3 = +(1/2)e_n$. Notice that whenever $x \in \hat{U} := \cap_j \partial U_j$, any neighborhood of x intersects all three U_j. Relative to $\mathbb{R}^2, \hat{U}$ consists of two points, but $\hat{U}$ is connected in space: $\hat{U} = S^2 \cap \{x_3 = 0\} = S^1$. Let's choose three points a, b, c equally spaced on S^1; these determine arcs $\{\gamma_i\}_1^3 \subset S^1$.

Now suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is K-quasiregular, nonconstant and f omits all three y_j. Let $W_j = f^{-1}(U_j)$, and note that each (component of each) W_j must be unbounded. In $\mathbb{R}^3$, Rickman creates one domain W_1 and W_2, lying respectively in the upper and lower half-spaces; these have obvious antecedents the upper and lower half-spaces in $\mathbb{R}^3$ which are preimages of $\{|x| > 1\}$ and $\{|x| < 1\}$ respectively of a Zorich 'exponential' function. Rickman then inserts six cone-like symmetric domains between W_1 and W_2, whose union is W_3. His function f is quasiregular in $\mathbb{R}^3$, with $f : W_j \to U_j \setminus \{y_j\},\ 0 \le j \le 3$.

The linking of the boundaries of these domains is given explicitly in §7 of [15], so that $f : \cap_j \partial W_j \to \hat{U}$, much as with the exponential function in $\mathbb{R}^2$, $\hat{e}(z) := e^{-iz}$, where $\hat{e} : \mathbb{R}(= \partial W_1 \cap \partial W_2) \to S^1 \subset \mathbb{R}^2$. Consider the domain W_1, similar observations hold for W_2 and each of the $\{W_3\}$. The f-preimage of $\{|y| = r\},\ r > 1$ is a surface inside W_1 which can be triangulated. As $r \downarrow 1$, each triangle $T \subset \partial W_1$ maps to $U_1 \cap U_3$ or $U_1 \cap U_2$, so that the images of triangles T and T° with a side in common alternate. Thus those of the first class have image in $S^2_+ := S^2 \cap \{x_3 \geq 0\}$, the others are mapped to $S^2_- := S^2 \setminus S^2_+$, and the arrangement of the $\{U_1\}_1^3$ forces similar correspondences of the images of triangles in $\partial W_2 \cup \partial W_3$.

Suppose $f : T \to U_1 \cap U_3$. It follows that as x crosses ∂T in ∂W_1, $f(x)$ must be able to access both domains U_2 and U_3. Let the sides of T be labeled Γ_i, $1 \leq i \leq 3$, with $f : \Gamma_i \to \gamma_i \subset S^1$. Rickman exploits $\mathbb{R}^3$ by viewing each image γ_i as a 'hinge', so that triangles

$$T' \subset \partial W_1 \cap \partial W_2, \quad T'' \subset \partial W_2 \cap \partial W_3$$

share Γ_i with T as a common boundary arc. Note that on other arcs $\Gamma_j \subset \partial T$, the particular T', T'' may change (although not their classes).

While this final step is very hands-on, Rickman also needs a beautiful deformation theory of two-dimensional surfaces to construct $f : W_j \to U_j$ away from $\cap_j W$. Modifying this to higher dimensions seems very difficult.

5. Concluding remarks

The two-dimensional theory in §2 has existed for nearly a century, and its importance has multiple confirmations. But there remain at least two classical problems relating to problems discussed here for which present technology seems inadequate, both arising from the fact that it is hard to obtain stronger conclusions than already mentioned when $\mu = \mu_g$ satisfies (6) or (7).

(A) *Examples with preassigned 'moving targets'.* The classical Nevanlinna relations (2), (3) are sharp, but they also hold with the $\{a_n\}$ replaced by 'small functions' $\{a_n(z)\}$ ([17, 20, 21]). To produce examples showing (2) and (3) sharp in this setting at present seems out of reach (it is likely that the method of Fuchs–Hayman [9] gives a partial solution for entire functions). If the target $a = a(z)$ is constant, the meromorphic function $f = g \circ \psi$ approximates $a \circ \psi \equiv a$, but this is no longer valid if $a(z)$ is nonconstant.

(B) *Examples with fast growth.* The general solution to (4) must have 'infinite order', but the solution in [4] grows very slowly given this necessary restriction. It seems very likely that given a growth function $\Psi(r) \uparrow \infty$, there should be a solution to the inverse problem having

$$T(r, f) > \Psi(r)$$

for all large r. The requirement of slow growth is to ensure that the deficiencies and indices of ramification for g transfer to f.

References

[1] AHLFORS, L. V.: Über eine in der neueren Wertverteilungstheorie betrachtete Klasse transzendenter Funktionen. *Acta Math.* **58** (1932), no. 1, 375–406.

[2] BERS, L.: Quasiconformal mappings, with applications to differential equations, function theory and topology. *Bull. Amer. Math. Soc.* **83** (1977), no. 6, 1083–1100.

[3] DRASIN, D. AND WEITSMAN, A.: Meromorphic functions with large sums of deficiencies. *Adv. Math.* **15** (1975), 93–126.

[4] DRASIN, D.: The inverse problem of the Nevanlinna theory. *Acta Math.* **138** (1976), 83–151.

[5] DRASIN, D.: On Nevanlinna's inverse problem. *Complex Variables Theory Appl.* **37** (1998), 123–143.

[6] DRASIN, D.: On the Teichmüller-Wittich-Belinskiı theorem. *Results Math.* **10** (1986) no. 1-2, 54–65.

[7] DRASIN, D.: Picard's theorem and the Rickman construction. *Sci. China Math.* **53** (2010), no. 3, 523–532.

[8] GOLDBERG, A. A. AND OSTROVSKII, I. V.: *Distribution of values of meromorphic functions.* Nauka, Moscow, 1970.

[9] HAYMAN, W. K.: *Meromorphic Functions.* Oxford Math. Monographs. Clarendon Press, Oxford, 1964.

[10] LEHTO, O.: *Korkeat Maalimat: Rolf Nevanlinnan elämä* (in Finnish). Otava, Helsinki, 2001.

[11] NEVANLINNA, R.: Untersuchungen über den Picard'schen Satz. *Acta Soc. Sci. Fenn.* **50** (1924), no. 6, 42 pp.

[12] NEVANLINNA, R.: Zur Theorie der meromorphen Funktionen. *Acta Math.* **46** (1925), 1–99.

[13] NEVANLINNA, R.: Uber Riemannsche Flächen mit endlich vielen Windungspunkten. *Acta Math.* **58** (1932), 292–373.

[14] NEVANLINNA, R.: *Eindeutige analytische Funktionen.* Springer, 1953.

[15] RICKMAN, S.: The analogue for Picard's theorem for quasiregular mappings in dimension 3. *Acta Math.* **154** (1985), 195–242.

[16] RICKMAN, S.: *Quasiregular Mappings.* Results in Mathematics and Related Areas **26**. Springer-Verlag, Berlin, 1993.

[17] STEINMETZ, N.: Eine Verallgemeinerung des zweiten Nevanlinnaschen Hauptsatzes. *J. Reine Angew. Math.* **368** (1986), 134–141.

[18] TEICHMÜLLER, O.: Untersuchungen über konforme und quasikonforme Abbilduing. *Deutsche Math.* **3** (1938), 621–678; included in his *Collected Works.*

[19] WITTICH, H.: *Neuere Untersuchungen Über eindeutige Analytische Funktionen.* Springer, 1968.

[20] Yamanoi, K.: Defect relation for rational functions as targets. *Forum Math.* **17** (2005), 169–189.

[21] Yamanoi, K.: The second main theorem for small functions and related problems. *Acta Math.* **192** (2004), 225–299.

Charles Fefferman (Maryland, USA, 1949) obtained his PhD in mathematics at Princeton University in 1966 (advisor: Elias Stein), and currently is Herbert Jones Professor at Princeton University. His main research areas are mathematical physics, harmonic analysis, fluid dynamics, neural networks, geometry, mathematical finance and spectral analysis, amongst others.

A reminiscence on BMO

Charles Fefferman*

As a grad student, I read a paper entitled "Singular integrals and differentiability properties of functions" by Eli Stein. From that paper, I learned that a singular integral operator carries L^∞ into BMO. The proof was very simple, and the result struck me as charming.

In the fall of 1970, I arrived at University of Chicago as a new assistant professor, and I met Antoni Zygmund (known to me for the next several years as "Professor Zygmund"). Zygmund immediately offered me an interesting problem to work on: Find a non-trivial characterization of functions $f \in \mathrm{BMO}(\mathbb{R})$ in terms of the Poisson integral of f. ("Non-trivial" means that we don't merely say that the Poisson integral $u(x,t)$ has a bounded BMO norm as a function of x uniformly in t.)

The author, *circa* 1985

I thought the problem was fascinating, and I couldn't stop thinking about it. My first thought was that since BMO is close to L^∞, maybe the right condition is the boundedness of the Lusin area integral

$$S(u)(x) = \Big(\iint_{|y-x|<t} |\nabla u(y,t)|^2 dydt \Big)^{1/2}.$$

That guess wasn't very smart. Every student of Fourier analysis knows $S(u)$ needn't be bounded, even when f is bounded. I quickly realized my mistake.

I remembered Eli's paper: The Hilbert transform of a bounded function belongs to BMO, and of course every bounded function belongs to BMO. I wondered: Are there any more examples? Are there any functions in BMO that could not be written in the form $f + Hg$, with $f, g \in L^\infty$ and $H =$ Hilbert transform?

I tried to find a function in BMO that could not be expressed as $f + Hg$, but at the end of a frustrating day, I couldn't find any such example.

The next morning I thought: What if every function in BMO could be expressed in the form $f + Hg$? What could that mean? I quickly saw that

*cf@math.princeton.edu.

it would then follows easily that BMO is the dual of H^1, thanks to the Hahn–Banach theorem. That seemed exciting, and I sat out to prove it.

A natural strategy to relate H^1 to BMO is to use the identity

$$\int_{\mathbb{R}} f\, g\, dx = \iint_{y\in\mathbb{R},\ t>0} t\ \nabla u(y,t)\cdot\nabla v(y,t)\, dy\, dt \tag{1}$$

(with u, v the Poisson Integrals of f, g), followed by the Cauchy–Schwartz inequality

$$\begin{aligned}
\Big|\iint_{y\in\mathbb{R},\ t>0} t\ \nabla u(y,t)\cdot\nabla v(y,t)\, dy\, dt\Big| &= \qquad (2)\\
&= \Big|c\int_{x\in\mathbb{R}}\int_{|y-x|<t}\nabla u(y,t)\cdot\nabla v(y,t) dy\, dt\, dx\Big|\\
&\le c\int_{x\in\mathbb{R}}\Big(\int_{|y-x|<t}|\nabla u(y,t)|^2 dy\, dt\Big)^{1/2}\Big(\int_{|y-x|<t}|\nabla v(y,t)|^2 dy\, dt\Big)^{1/2} dx\\
&= c\int_{x\in\mathbb{R}} S(u)(x)\, S(v)(x)\, dx.
\end{aligned}$$

I had learned that trick from Eli Stein's wonderful grad course, that subsequently formed the first half of his book *Singular integrals and differentiability properties of functions*.

Alberto Calderón had recently proven that $S(u) \in L^1$ for H^1-function f; that was a big step in his great work on the commutator integral.

If only $S(v) \in L^\infty$ for $g \in$BMO, then the identity (1) and the inequalities (2) would imply that

$$\Big|\int fg\, dx\Big| \le C\|f\|_{H^1}\|g\|_{BMO}$$

and thus establish the duality of H^1 and BMO.

Damn it! $S(v)$ needn't be bounded for $g \in$ BMO.

Duality of H^1 and BMO hung on finding a good answer to Zygmund's question: How can we characterize $g \in$ BMO in terms of the Poisson integral v?

After a few days without further progress, I hit on the right idea: Instead of using the full area integral $S(v)$, I should bring in a "partial area integral"

$$S_{\text{partial}}(v)(x) = \Big(\iint_{|y-x|<t<h(x)}|\nabla v(y,t)|^2 dy\, dt\Big)^{1/2},$$

where the height $h(x)$ is defined to be as large as possible subject to the constraint that

$$S_{\text{partial}}(v) \le C_1$$

for a large constant C_1.

Thus,

$$S_{\text{partial}}(v)(x) = \min\left\{S(v)(x),\, C_1\right\} \leq C_1$$

automatically and $h(x)$ may be a positive real number or $+\infty$.

By replaying the argument (1), (2) using S_{partial} instead of S, one finds easily that

$$\left|\int f\,g\,dx\right| \leq C\|f\|_{H^1}$$

provided the following holds:

> Let $(y,t) \in \mathbb{R} \times (0,\infty)$, and let $h(x)$ be the height function defined above. Then the set $\{x \in (y-t, y+t) : h(x) > t\}$ has measure al least ct. (3)

Thus, a function g may be paired with H^1, provided its height function $h(x)$ satisfies (3).

A little further experimentation led to the condition

$$\int_{|y-x|<h,\ 0<t<h} t|\nabla\, v(y,t)|^2 dy\, dt \leq Ch \quad \text{for every } x \in \mathbb{R}, h > 0. \qquad (4)$$

It is easy to see that (4) implies (3), and that any BMO function with norm 1 satisfies (4).

Consequently $g \in \text{BMO} \Rightarrow v$ satisfies (4) $\Rightarrow g \in (H^1)^* \Rightarrow g = f_1 + Hf_2$ (with $f_i \in L^\infty$) $\Rightarrow g \in \text{BMO}$, answering Zygmund's question and proving the duality of H^1 and BMO.

The whole thing took less than two weeks. It struck me that, for the first time, I had proven an interesting result without hard work. I had simply asked (and received) questions, and made conjectures until the conjectures proved one another.

For many years, I've worked very hard to prove theorems. With luck, I've found complicated proofs after much suffering. With extraordinary luck, I found simple proofs after even more suffering. To find a simple proof, without suffering for it, is a very rare success. I will always be grateful for my incredible luck in reading Eli's paper and hearing Zygmund's question.

José Luis Fernández (Santa Cruz de Tenerife, Spain, 1956) obtained his PhD in mathematics at Washington University in St. Louis, in 1983 (advisor: Albert Baernstein II). He is professor at the Department of Mathematics of the Universidad Autónoma of Madrid. His main research area is complex analysis. Josechu is Co-Director of *Revista Matemática Iberoamericana*.

¡Olé!

José L. Fernández*

To Albert Baernstein II, ¡maestro!, *with appreciation and love*

In the course of a general lecture, in Madison, certainly in 1984, Marc Kac was recounting his training as a young mathematician in Poland. It was a large lecture room in Van Vleck Hall, one of those devoted to massive calculus classes.

Kac was a marvelous lecturer who dominated the scenario (no power-point, mind you!) with elegant demeanor and dynamic rhythm sprinkled with the appropriate theatrical stops here and there. I am quite certain that he was grateful that so many young faces were present and attentive. His hosts for that visit to Madison were the Rudin, Walter and Mary Ellen, and quite probably they had spread the news among students that the great man was around and that nobody could afford to miss the chance of being enlightened by him.

For starters, he recalled that the very very first lecture he attended upon entering college to study Mathematics was on . . . Dedekind cuts. Of course!, what else!, if you plan to study Mathematical Analysis, you must come equipped with a good functioning real line and not just with a näive intuition[1].

At that time, Kac was finishing writing his splendid *Enigmas of Chance: An Autobiography*[2] and the lecture used liberally some of its mathematical anecdotes. One of them, the one I recall, described his amazement when Steinhaus, who was his thesis supervisor[3], explained to him how to understand –following Borel– the probabilistic notion of *infinite* independent throws of a regular coin by choosing a number at random (uniformly) in the unit interval. Probability was not what it is nowadays, but a kind of Cinderella in the realm of serious and respected Mathematics. To material-

*a.k.a Josechu Fernández, joseluis.fernandez@uam.es.

[1] By the way, that was also the customary way to start the math curriculum in Spain in the early 70's; the pendulum have shifted to the opposite extreme. Kac intention with that opening, besides making an impression on the younger part of the audience, was to emphasize that to study Mathematics at the University (which meant almost only Math, with some courses in Physics) was reserved then to a few selected and very well trained students.

[2] Published posthumously; Marc Kac died shortly after, in 1985.

[3] Whatever *thesis supervisor* meant then.

ize that virtual coin model in such a precise and clear cut[4] analytical way seemed almost miraculous to him.

The coefficients $\xi_j(x)$ in the dyadic representation of numbers $x \in [0,1]$,

$$x = \sum_{j=1}^{\infty} \xi_j(x) 2^{-j}$$

form an *infinite* sequence of independent identically distributed *functions* in $[0,1]$, with Lebesgue measure as the underlying probability; Borel's (strong) law of large numbers and consequently the fact that almost every number in $[0,1]$ is normal can be proved directly and *analytically* within that model. This simple and beautiful fact is the starting point of Kac's delightful MAA Carus Monograph *Statistical Independence in Probability, Analysis and Number Theory*[5].

Before I continue any further, a DISCLAIMER seems appropriate. This is not a historical article, but purely and simply truthful storytelling, nothing else. It contains, subjectively ordered and subjectively chosen and subjectively interpreted, some bits and pieces (or even, traces) of mathematical history, precisely those that my story requires. It is just how I saw it or, as Walter Rudin would put it, *The way I remember it*.

Random walk

Random walk sums ± 1's instead of 0's and 1's. For independent ± 1's, one could use the Rademacher functions ϱ_j obtained by declaring $\varrho_j(x) = +1$, if $\xi_j(x) = 0$, and $\varrho_j(x) = -1$, if $\xi_j(x) = 1$. The sequence of partial sums of Rademacher functions:

$$\left\{ \sum_{j=1}^{n} \varrho_j(x) \right\}_{n=1}^{\infty}$$

is an arithmetic model for the symmetric random walk, encoded as a sequence of step functions in $[0,1]$.

Sine functions with exponentially (or lacunarily) growing periods

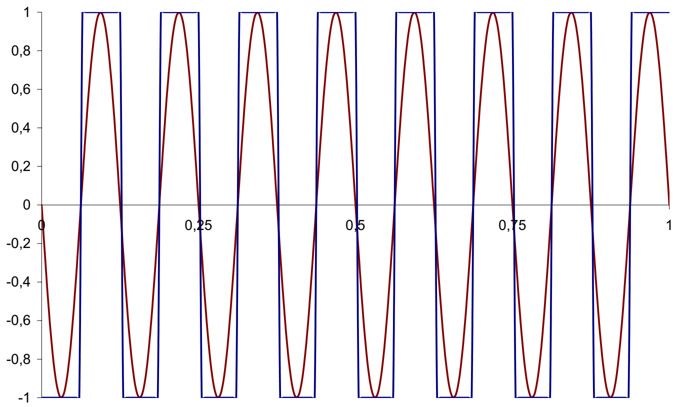
$$\sin\bigl(2^{j-1}\, 2\pi x\bigr)$$

resemble Rademacher functions –actually the latter are just squared versions of the former: $\varrho_j(x) = \operatorname{sign}\bigl(\sin(2^j \pi x)\bigr)$– and are very close to being statistically independent.

---

[4] Oops!, Dedekind.

[5] One of the few books I am planning to take with me, for occasional consolation, to that fabled isolated desert island, no kidding!



RADEMACHER AND SINE

This similarity naturally incites to look with probabilistic eyes at lacunary trigonometric series; for instance, they satisfy a Central Limit Theorem: the sequence of functions

$$S_N(x) = \sum_{j=1}^{N} \Big(a_j \sin(2^j \pi x) + b_j \cos(2^j \pi x) \Big) ,$$

appropriately normalized with

$$\sigma_N = \Big(\sum_{j=1}^{N} \frac{1}{2} (a_j^2 + b_j^2) \Big)^{1/2}$$

converges in distribution to a standard normal variable as $N \to \infty$, assuming the Lindeberg type conditions:

$$\lim_{N \to \infty} \frac{\max_{j \le N} (a_j^2 + b_j^2)}{\sigma_N^2} = 0 \quad \text{and} \quad \lim_{N \to \infty} \sigma_N = +\infty .$$

You may find a detailed exposition of this result of Raphaël Salem and Antoni Zygmund in the latter's *Trigonometric series*, or, as well, you may not find it[6,7]. For lacunary series there is even a law of iterated logarithm

[6] Zygmund's *Trigonometric series*, that amazing book of cosmological proportions, is so huge, intricate and russelian that it is almost unable to contain itself. I, or better, a former version of myself, *knew* that this Central Limit Theorem was discussed at length in it. Well, it is in there, but hidden: in the notes of the first volume, chapter V, page 380, it is claimed that it will be proved in the second volume, chapter XV, §4; but no, it is actually in chapter XVI. Now you know. Please, go to footnote 7.

[7] Try the experience of describing what a trigonometric series is to anyone who is cultivated enough, and show the book, or even better, ask him to hold it with one hand, while you explain that the whole book is devoted to that single specific subject. Wait for a while. Do not spoil the experience by talking about Joseph Fourier and the generality of the concept.

due to Mary Weiss[8]: assuming now that

$$\lim_{N\to\infty} \frac{\max_{j\le N}(a_j^2+b_j^2)}{\sigma_N^2 \big/ \ln\ln\sigma_N} = 0 \quad \text{and} \quad \lim_{N\to\infty} \sigma_N = +\infty\,,$$

then

$$\limsup_{N\to\infty} \frac{S_N(x)}{\sqrt{2\sigma_N^2 \ln\ln\sigma_N}} = 1\,,$$

for *a.e.* $x \in [0,1]$.

Martingales

Lacunary series are, of course, very special: they correspond to sums of independent identically distributed variables. So, what about general power series? From the probabilistic side, martingales furnish the appropriate general language.

The k^{th} generation of dyadic intervals consists of the intervals

$$I_{k,j} = \Big[\frac{j-1}{2^k}, \frac{j}{2^k}\Big), \qquad j = 1, 2, \ldots, 2^k$$

and generates an algebra of sets $\mathcal{F}_k$. The sequence $\{\mathcal{F}_k\}$ –the dyadic filtration– is the simplest model of information flow or of unveiling of uncertainty. If $I_k(x)$ denotes the k^{th} generation interval that contains x, unveiling $I_{k+1}(x)$ given $I_k(x)$ reduces the uncertainty of the location of x by half.

Martingales with respect to the dyadic filtration, although very concrete example of martingales, are a natural test ground for general results, abstract notions and fruitful connections. For instance, for a function f in $L^1[0,1]$, the sequence of averages

$$\frac{1}{|I_k(x)|} \int_{I_k(x)} f(u)du$$

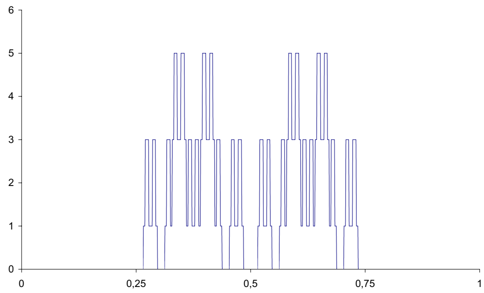
coincides with the sequence of conditional expectations $\mathbf{E}(f|\mathcal{F}_k)(x)$ and thus Lebesgue (dyadic) differentiation corresponds to martingale convergence[9].

[8] In an obituary –M. Weiss died at age 35– Zygmund has this to say which pertains to the *story* of this article: *[...] she attended my class in trigonometric series, our mathematical contact originating when she began working on her Ph.D. thesis on a topic I suggested, lacunary trigonometric series; the behavior of these series resembles in many ways that of series of independent random variables in the calculus of probability, a subject which she had mastered while working for the Advisory Board on Simulation, a project of the University of Chicago.*

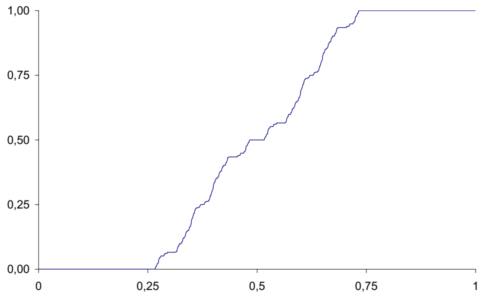
[9] ,*Martingales à temps discret*, the book of Jacques Neveu, is simply elegant.

A further, and relevant, illustration of this connection (or point of view, if you prefer) is Kahane's[10] construction of a probability measure μ in the interval $[0,1]$ which is singular with respect to Lebesgue measure but nonetheless has certain smoothness, namely that $|\mu(I) - \mu(I')| \leq C|I|$, for every pair of contiguous intervals with $|I| = |I'|$. Any primitive f of this μ belongs to the Zygmund class[11].

Kahane's delightfully simple construction begins with the arithmetic model for random walk starting from level 1, which he[12] stops as soon as it reaches 0. This procedure generates a positive martingale whose almost everywhere limit is 0 and defines the sought after singular measure[13].



4^{TH} STEP OF KAHANE'S MARTINGALE



KAHANE'S FUNCTION

One can represent a general dyadic martingale $\{X_k\}_{k\geq 0}$, or better its successive differences, $\Delta X_k = X_k - X_{k-1}$, in terms of Haar functions. The Haar functions $\chi_{k,i}(x)$, $i = 1, 2, \ldots, 2^{k-1}$, are given by

$$\chi_{k,i}(x) = \begin{cases} -1, & \text{if} \quad x \in I_{k,2i-1}, \\ +1, & \text{if} \quad x \in I_{k,2i}, \\ 0, & \text{otherwise.} \end{cases}$$

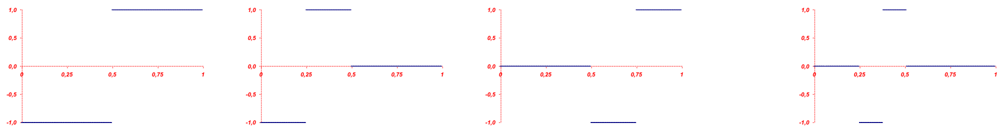
[10] In his own words: *Nous développons ici un exemple qui avait été mentionné dans [a paper of G. Piranian] sous le nom de Kahane's example...* what a lovely exercise in self-reference!

[11] The significance of the example of Kahane's is that for appropriately small $\delta > 0$, Jochen Becker's criterion of univalence shows that the holomorphic function given by $\ln f'(z) = -\delta \int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} d\theta$ is conformal, and, since μ is singular, the image under f of the unit circle is not a Smirnov domain. Kahane's was not the first example of such a conformal mapping, but its simplicity is striking.

[12] He, Kahane, himself, with enough power to personally stop a random walk at will.

[13] It fulfills the required condition $|\mu(I) - \mu(I')| \leq C|I|$, but only for contiguous dyadic intervals of the same length whose union is also a dyadic interval. To get the general result Kahane replaces, in an elegant twist, dyadic expansions with 4-adic expansions: $x = \sum_{i=0}^{\infty} \frac{\tau_i(x)}{4^i}$, defines $\omega_i(x) = +1$, if $\tau_i(x) = 0$ or 3, and $\omega_i(x) = -1$, if $\tau_i(x) = 1$ or 2, and considers instead the random walk given by $\{\sum_{i=1}^{n} \omega_i(x)\}_{n\geq 1}$.

These are the first cases:



$\chi_{1,1}$ $\chi_{2,1}$ $\chi_{2,2}$ $\cdots$ $\chi_{3,2}$

The representation is then

$$\Delta X_k = \sum_{i=1}^{2^{k-1}} a_{k,i}\, \chi_{k,i}\,, \qquad \text{or} \qquad X_n = X_0 + \sum_{k=1}^{n} \underbrace{\Bigg(\sum_{i=1}^{2^{k-1}} a_{k,i}\, \chi_{k,i} \Bigg)}_{=\Delta X_k}$$

for suitable *constants* $a_{k,i}$. If all the a's coincide, one has random walk, while if, for each level k, the $a_{k,\cdot}$'s coincide, the differences ΔX_k are independent, and one has a weighted sum of Rademacher functions. Martingales appear here as stochastic sums with respect to random walk or, if you wish, as martingale transforms.

Now, let's go back to power series to describe how the language of martingales and the probabilistic lense[14] illuminates them. Consider a function f holomorphic in the unit disk with power series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Let us split[15] the whole power series into a series of dyadic blocks:

$$f(z) = \sum_{n=0}^{\infty} a_n\, z^n = a_0 + \sum_{k=1}^{\infty} \underbrace{\Bigg(\sum_{2^{k-1} \leq n < 2^k} a_n\, z^n \Bigg)}_{=\Delta_k(f;z)}.$$

Here is an excerpt from the renowned Dictionary for aficionados from power series to martingales: consider the sequence of radii

$$r_k = 1 - \frac{1}{2^k}$$

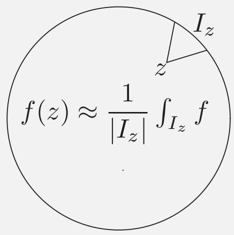
tending to 1, and assume that the function f at $|z| = 1$ is properly defined.

[14] As in *The probabilistic method* of Noga Alon and Joel Spencer.
[15] Purely classical Littlewood–Paley approach.

→ **function f restricted to $\{|z| = r_n\}$**

1. as conditional expectation

The function f restricted to $\{|z| = r_n\}$ is very much like the conditional expectation of $f(e^{2\pi ix})$ with respect to $\mathcal{F}_k$. The value of f at a point z, $|z| = r_n$ is a weighted average (Poisson, Cauchy) of its boundary values, but those in the boundary interval, I_z, just in front of z and of length $1 - r_n$ have the most weight.



2. as dyadic martingale

Thus one may think of the sequence of functions

$$x \in [0, 1] \mapsto f(r_k e^{2\pi ix})$$

as sort of a dyadic martingale. Moreover, when restricted to $|z| = r_n$, the function f and the partial sum $\sum_{k \le n} \Delta_k(f; z)$ are very close and behave similarly; so that $\Delta_k(f; r_n e^{2\pi ix})$ plays the role of the n^{th} martingale difference of the martingale generated by $f(e^{2\pi ix})$.

3. as hyperbolic derivative estimate

Finally, to complete this round of intuitions, an estimate:

$$|\Delta_k(f; r_n e^{2\pi ix})| \approx |f'(z)|\,(1 - |z|^2); \quad z = r_n e^{2\pi ix}.$$

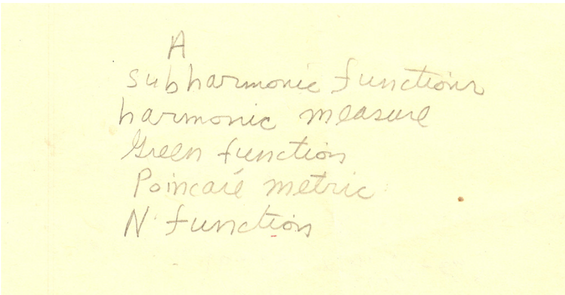
The term $f'(z)(1 - |z|^2)$ is the hyperbolic derivative, the distortion factor from hyperbolic metric in the unit disk to Euclidean metric in the complex plane.

So there you have it; a sort of one way dictionary, shared by a number of devotees and aficionados. At this point, this dictionary is just an analogy; a fruitful, alternative and suggestive way of looking at power series and holomorphic functions.

Geometric function theory

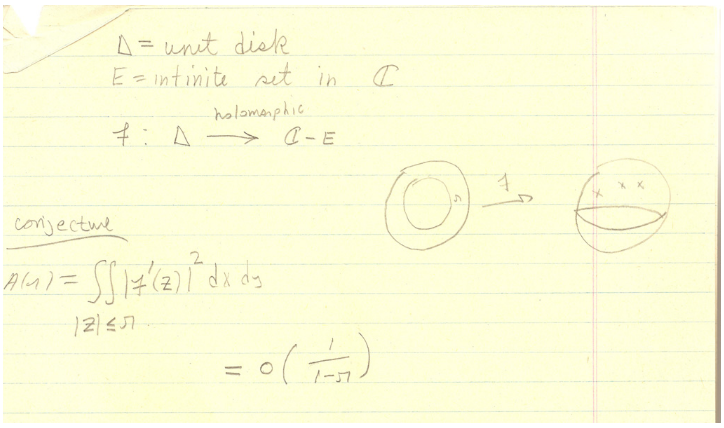
Loosely speaking, in good old classical Geometric Function Theory one tries to deduce analytical properties of holomorphic functions from geometric information on their range, or on their Riemann surfaces, or on their value distribution, or... De Branges theorem is one glorious example: the Taylor coefficients $\{a_n\}$ of conformal mappings defined in the unit disk verify that $|a_n| \leq n$, for any $n \geq 2$.

In my second[16] visit to Al Baernstein's office, he first suggested that over the incoming summer vacation I should read Rolf Nevanlinna's *Analytic functions*[17] and Lars Ahlfors' *Conformal invariants: topics in geometric function theory* to get acquainted with the appropriate tools[18] of the trade,



TOOLS OF THE TRADE

and then afterwards, he took his pad (of paper, of course) and proposed to me a question:



SUMMER VACATION

In a further side discussion in the blackboard, Al rewrote the statement of the question and preceded it with the phrase "TO PROVE:", or maybe

[16] The purpose of my first visit, just in case you care to know, was to ask Al if he would not mind being my thesis advisor.

[17] Also known at that time, and among the selected circle of disciples, and a few converted aficionados like myself, as the New Testament; the Old Testament being Nevanlinna's *Le théorème de Picard-Borel et la théorie des fonctions meromorphes*.

[18] By the way, the A stands for the area function which appears in the problem of the second picture.

just “TP:”, so I took it as sort of training problem, part of the initiation that I was just embarking upon, and that just a week of my vacation devoted to it should be enough; so I reserved it for the second week, to fully enjoy it. How touchingly naive!; no, it took a bit longer.

That problem was closely connected[19] with a question about Taylor coefficients and range of holomorphic functions. Assume that we know that the range of a holomorphic function f lies in a certain domain $\Omega \subset \mathbb{C}$. If the distance to the boundary of points in Ω is bounded, the f has bounded Taylor coefficients, and the smaller the bound on the distance, the smaller the bound on the coefficients; that much was known. But, assume further that the distance to the boundary tends to 0 when $\infty_{\mathbb{C}}$ is approached from within Ω; in that case, is it true that the Taylor coefficients tend to 0?[20]

A simpler looking –actually, more ambitious– challenge would be to prove that holomorphic functions f in the unit disk whose hyperbolic derivative $f'(z)(1-|z|^2)$ vanishes asymptotically as $f(z)$ tends to ∞ have vanishing Taylor coefficients. Now, no restriction on the range of the function is imposed. A positive answer to this question would mean a positive answer to the original one.

For this, Al suggested that I should consider *first* a convenient dyadic martingale analogue, to wit: is it true that $\mathbf{E}(|X_n|) \to 0$ for a dyadic martingale X_n whose sequence of differences ΔX_n tends to 0 whenever X_n tends to infinity?

The answer turned out to be no: the key[21] is to run random walk endowed with stopping times at an appropriate sequence of barriers of, alternatively, very large and very small levels[22].

[19] Only in Al’s mind, at that moment, and he was right.

[20] It would be natural to expect so, since in a certain loose sense, Ω may be split into a bounded part and a part with very small distance to the boundary, and the Taylor coefficients of bounded functions tend to 0. Besides, the property was known to hold for the universal covering map of Ω.

[21] Or, at least, one key since this is just general construction of examples.

[22] I must confess that this idea of iterating Kahane’s construction occurred to me while attending a seminar lecture. I was really paying attention to the lecture. It was a good lecture whose main thrust, completely unrelated to my interests at the time (or everafter) I still remember –well, vaguely. I have to confess too that another good idea had occurred to me during a lecture, but this time ... I was the lecturer and again the subject was removed from my most immediate research interest at the time. I am referring here to this definite flash that occurs when you perceive a connection between two mathematical ideas, the final step of an argument, the trick that will show you the way to conquer a so far impenetrable calculation, clearly and distinctively, and that you *know* that it is right. They, ideas, are temperamental by nature and may come to you at any time, like Poincaré’s *au moment où je mettais le pied sur le marchepied, l’idée me vint*, or in the shower, peripatetically, dozing away in a sofa, even... well, you know... and, still, attending or delivering lectures, too.

Starting from there, one could construct a holomorphic function which mimicked the behavior of the martingale, thus resolving in the negative the original question[23]. There was no operator transferring the martingale into the holomorphic function, no canonical construction, only just building of examples by sheer imitation. More care was required if the holomorphic function should have a given range, but again the answer is no, quite generally.

Conformal mapping

It was[24] a sunny, limpid, crispy morning day in Madison[25,26]. A few weeks before, the news that a young Russian has done good work on boundary distortion under conformal mappings have reached me. Nothing was said about what kind of results, whether positive, partial or not, or counterexamples, just that it was a very good paper[27]. And then the reprint arrived in the mail, with a polite note from the author hoping that I would find the paper interesting. In those times, one wrote *letters* and got *reprints* by *regular* mail.

The metric behavior of conformal mappings is clearly an interesting matter, basic and substantial[28]. It was not clear, at least not to me, if a result valid for all conformal mappings was to be expected, maybe a scale of distortion results was possible. I had given some thought for a while to the subject[29]. As a dutifully professional mathematician, I was awaiting the paper with some expectation and was ready to devote time and effort to read it and to understand and apprehend some new interesting techniques to be added to my bag of tricks. I was expecting –with delectation, I must confess– long chains of gruesome lemmas with careful and intricate estimates of the angular derivative near prime ends[30] of different colors and species and convoluted fractal and tree like constructions of the appropriate simply

[23] In the paper of mine pertaining this (counter-)example, martingales are not mentioned, while precisely that power series-martingale connection was the most interesting part.

[24] Well, maybe not, maybe it was a gloomy and rainy fall afternoon. But that would be a simple and certain fact, and I, today, just want the truth. So it was sunny and winter. QED.

[25] A senior colleague in Madison angrily refused to praise those few gorgeous winter days, he assured that the university claimed them as part of the salary and that each of them costed him a thousand dollars per year.

[26] 1984, once again, shut up Orwell!

[27] The news by letter probably came from Al, since I seem to remember that the *suggestion* was interpreted by me as the command *Josechu, read and learn.*

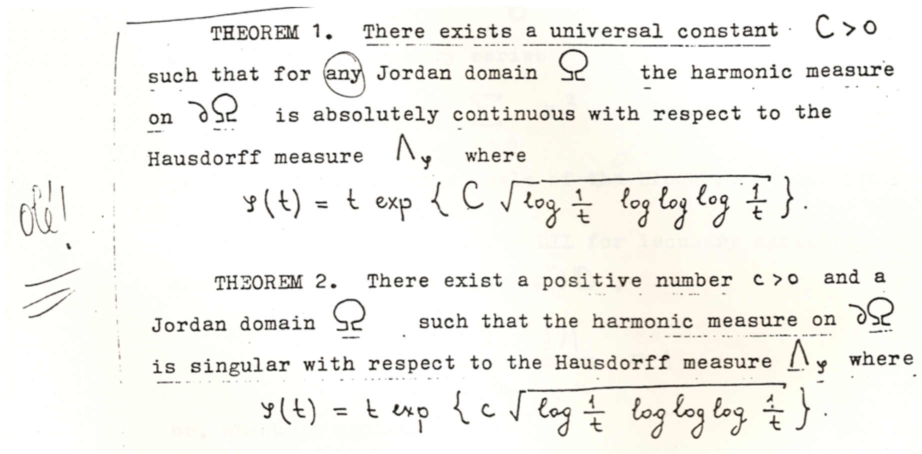
[28] It is not the kind of thing I would claim aloud in the bar next door, but *you* do understand me.

[29] Or lots of thought for quite a while.

[30] That's my sole aim in life: to become a *prime end.*

connected domains which could serve, just in case the need arises, as counterexamples. You know, as Littlewood claimed, pioneering work is (almost) always clumsy.

I was stunned: the paper was a masterpiece, so elegant and graceful. The main result was absolutely general, valid for any Jordan domain, no scale whatsoever. I could read it almost effortlessly. I was well trained and reasonably knowledgeable of the subject, I was ready for it, that helped a bit. The arguments progressed smoothly, for long stretches I knew what was coming next and could jump ahead but then once in a while there was this sudden unforeseeable (for me) turn of ideas which shortly after seemed unavoidable and inevitable and proper and so obviously right.[31] I read the paper in a continuos rupture, a few hours went by, but when I was through, with just a few details to be assimilated *¡mañana, mañana!*, I recounted the whole argument just for myself, and decided enthusiastically that it was worthy of the highest marks:


THEOREM 1. There exists a universal constant $C>0$ such that for any Jordan domain Ω the harmonic measure on $\partial\Omega$ is absolutely continuous with respect to the Hausdorff measure Λ_γ where

$$\gamma(t) = t \exp\left\{ C \sqrt{\log \tfrac{1}{t} \log\log\log \tfrac{1}{t}} \right\}.$$

THEOREM 2. There exist a positive number $c>0$ and a Jordan domain Ω such that the harmonic measure on $\partial\Omega$ is singular with respect to the Hausdorff measure Λ_γ where

$$\gamma(t) = t \exp\left\{ c \sqrt{\log \tfrac{1}{t} \log\log\log \tfrac{1}{t}} \right\}.$$

Olé!

AND ¡VUELTA AL RUEDO! TOO[32,33]

[31] In the 1984 movie *Amadeus*, in a memorable scene near the end, a dying Mozart wrapped in a blanket dictates a music score to Antonio Salieri, while he distractedly keeps playing around with the cue ball over a pool table. The whole score is complete in Mozart's head; Salieri –played superbly by a marvelous Murray Abraham– is simply amazed, since there is no hesitation, no turning back, no trial and error, no need for that. Once in a while Mozart stops, waves his hand and mumbles *and so on and so forth*, expecting Salieri to fill in the blanks, but Salieri simply cannot, and Mozart impatiently and tiredly does it for him, and every time this happens Salieri is shocked to realize how unavoidably right Mozart solution was.

[32] Please, notice the enthusiastic triply underlined handwritten *Olé!* on the left margin. One left exclamation mark (¡) is missing, maybe just the price of interculturality, or intertextuality, or text deconstruction, or what have you.

[33] Notice, please, the typesetting technology of the time. First you typewrited the text leaving empty and appropriate spaces for all the mathematical symbols which are afterwards meticulously handwritten; amazing craftsmanship.

Stop! I almost forget: the author and the paper:

Nikolai G. Makarov
On the distortion of boundary sets under conformal mappings
Proc. London Math. Soc. **51** (1985), 369–384.

So, what's in the paper? Bloch functions[34] are those holomorphic functions g in the unit disk whose hyperbolic derivative

$$g'(z)(1-|z|)^2$$

is bounded. Recalling the dictionary for aficionados, Bloch functions would *correspond* to martingales with bounded differences. First, Makarov establishes a law[35] of iterated logarithm for Bloch functions:

$$\limsup_{r\to 1^-} \frac{|g(re^{2\pi ix})|}{\sqrt{\ln\frac{1}{1-r}\ln\ln\ln\frac{1}{1-r}}} \leq C\cdot \sup_{z:|z|<1} |g'(z)|(1-|z|^2)\,;$$

C is an absolute constant. An unexpected and general result, with a direct, clean, economical, almost functorial proof.

And? Well, if f is a conformal mapping in the unit disk, then $\ln(f')$ is a Bloch function; that is the basic distortion theorem. The converse (with a factor added for good measure) is also true; that is exactly the content of Becker's univalence criteria (see footnote 11). Bloch functions conform a linear space; it is easy to work with them, but, amazingly, they capture within the far more complex and nonlinear notion of univalence and conformality.

And thus, Makarov's law of iterated logarithm for Bloch functions furnishes good (well, even sharp) general estimates of the derivative of conformal mappings near the boundary; you see: conformal mapping results obtained via martingale thinking.

Next step is to combine, in purely geometric function theory style, those distortion estimates with the extremal length techniques of Lars Ahlfors and Arne Beurling to produce harmonic measure estimates. The final outcome is the absolutely general (and quite sharp):

[34] The class of Bloch functions was termed (because of its connection with André Bloch's theorem) and introduced in an elegant and pathbreaking paper of Milne Anderson, Jim Clunie and Christian Pommerenke entitled *On Bloch functions and normal functions.*

[35] Weaker, as it correspondes to martingales and not to i.i.d. sequences.

THEOREM 1. There exists a universal constant $C > 0$ such that for any Jordan domain Ω the harmonic measure on $\partial\Omega$ is absolutely continuous with respect to the Hausdorff measure Λ_φ where

$$\varphi(t) = t \exp\left\{ C \sqrt{\log \frac{1}{t} \log\log\log \frac{1}{t}} \right\}.$$

Lennart Carleson, in a characteristic *tour de force*, had obtained some ten years earlier that there exists some number $\beta > \frac{1}{2}$, so that harmonic measure on any Jordan domain is absolutely continuous with respect to Hausdorff measure Λ_{t^β}; the case $\beta = \frac{1}{2}$ was a trivial consequence of the basic Beurling's projection lemma.

Beautiful; no question. It fulfills to perfection Hardy's canon for serious (significance, generality, depth) and beautiful (unexpectedness, inevitability, economy) Mathematics.

Coda

So now, I hope, you may appreciate how well founded was my enthusiastic and irrepressible *¡Olé!* Of course, there have been many, well, maybe not so many, but at least a few, other occasions of such enthusiastic appreciation of mathematical beauty in a new discovery of others, and, yes, if you ask, my own work has also been a source of deep excitement[36].

Appreciation of beauty is a cumulative cultural affair continuously evolving: you have to be prepared and ready to discriminate, to appreciate and to *share* it. But to *create* beauty, well, that is altogether a different matter. *¡Olé, Makarov!*

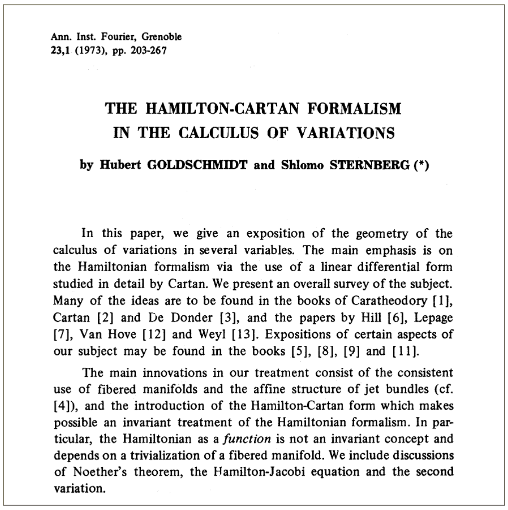
[36] Actually, as I believe it occurs to any of us, any minor discovery –if one may use that word, discovery– of oneself has been a much more profound source of joy, but that is, at least in my case, a pretty partial perception.

Pedro Luis García (Cartagena, Spain, 1938) obtained his PhD in mathematics at the University of Barcelona in 1967 (advisor: Juan Sancho Guimerá). He currently holds the position of Professor Emeritus at the University of Salamanca. His main research areas are differential geometry, variational calculus and geometric mechanics.

The Hamilton-Cartan formalism in the calculus of variations

Pedro Luis García*

In 1973, Hubert Goldschmidt and Shlomo Sternberg published with this title an important survey paper, [10], on the geometrical foundations of the calculus of variations on fibred manifolds. This article was much celebrated among researchers who were working at that time on that subject, that emerged in the process of renewal of the foundations of differential geometry dating back to the 1950s after the rigorous definition of the concept of differentiable manifold. I was among them [3, 4, 5]. Thus, I believe that this monograph and its influence on my subsequent research, increased by the strong friendship that I shared with Sternberg, fits in perfectly with the purposes of this special issue of the Revista Matemática Iberoamericana, on the occasion of the first centenary of the Real Sociedad Matemática Española.


Ann. Inst. Fourier, Grenoble
23,1 (1973), pp. 203-267

THE HAMILTON-CARTAN FORMALISM
IN THE CALCULUS OF VARIATIONS

by Hubert GOLDSCHMIDT and Shlomo STERNBERG (*)

In this paper, we give an exposition of the geometry of the calculus of variations in several variables. The main emphasis is on the Hamiltonian formalism via the use of a linear differential form studied in detail by Cartan. We present an overall survey of the subject. Many of the ideas are to be found in the books of Caratheodory [1], Cartan [2] and De Donder [3], and the papers by Hill [6], Lepage [7], Van Hove [12] and Weyl [13]. Expositions of certain aspects of our subject may be found in the books [5], [8], [9] and [11].

The main innovations in our treatment consist of the consistent use of fibered manifolds and the affine structure of jet bundles (cf. [4]), and the introduction of the Hamilton-Cartan form which makes possible an invariant treatment of the Hamiltonian formalism. In particular, the Hamiltonian as a *function* is not an invariant concept and depends on a trivialization of a fibered manifold. We include discussions of Noether's theorem, the Hamilton-Jacobi equation and the second variation.

1. The formalism of Hamilton–Cartan in the calculus of variations, following Goldschmidt and Sternberg

Let $\pi\colon Y \to X$ be a fibred manifold over an n-dimensional manifold X, oriented by a volume element ω. Let $j^1\pi\colon J^1Y \to X$ be its 1-jet extension and let $\mathcal{L}$ be a differentiable function (the Lagrangian) defined on J^1Y.

If A is a compact n-dimensional submanifold of X with boundary a $(n-1)$-dimensional submanifold ∂A, we can define a functional on the set of sections $s \in \Gamma(X, Y)$ by the rule:

$$\mathbb{L}_A(s) = \int_A (\mathcal{L} \circ j^1 s)\omega$$

where $j^1 s$ is the 1-jet extension of the section s.

*pgarcia@usal.es.

In these conditions, we say a section $s \in \Gamma(X,Y)$ is stationary (for fixed boundary values on ∂A) when:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{L}_A(s_t)\bigg|_{t=0} = 0$$

for every 1-parameter smooth family of sections $s_t \in \Gamma(X,Y)$, $t \in (-\epsilon,\epsilon)$, $\epsilon \in \mathbb{R}^+$, such that $s_0 = s$ and such that for each t, $s_t = s$ on ∂A.

The idea of Goldschmidt and Sternberg to characterise the stationarity was to use the theory of linear differential operators, as follows:

If $\delta s = \frac{\mathrm{d}s_t}{\mathrm{d}t}\big|_{t=0} \in \Gamma(X, s^*VY)$ and $j^1\delta s \in \Gamma(X, J^1(s^*VY))$ is its 1-jet extension (s^*VY being the bundle induced through s by the bundle VY of π-vertical vector fields on Y), there holds:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{L}_A(s_t)\bigg|_{t=0} = \int_A \mathcal{P}_s(\delta s)\omega$$

where $\mathcal{P}_s \colon \Gamma(X, s^*VY) \to \mathcal{C}^\infty(X)$ is the first linear differential operator:

$$\mathcal{P}_s(\delta y) = (\mathrm{d}\mathcal{L} \circ j^1 s)(j^1\delta s)$$

From here it follows that a section $s \in \Gamma(X,Y)$ is stationary if and only if:

$$\int_A \mathcal{P}_s(\delta s)\omega = 0 \tag{1}$$

for every section $\delta s \in \Gamma(X, s^*VY)$ whose support is interior to A.

Let us now consider the adjoint operator $\mathcal{E}_s = \mathcal{P}_s^*$, that is, the only first order linear differential operator:

$$\mathcal{E}_s \colon \Gamma(X, \Lambda^n T^*X) \to \Gamma(X, s^*VY \otimes \Lambda^n T^*X)$$

such that

$$\int_X \langle \mathcal{P}_s\delta s, \alpha\rangle = \int_X \langle \delta s, \mathcal{E}_s\alpha\rangle$$

for every $\alpha \in \Gamma(X, \Lambda^n T^*X)$ and every $\delta s \in \Gamma(X, s^*VY)$ with compact support (where $\langle\ ,\ \rangle$ is the bilinear duality product).

In these conditions, the "integration by parts" formula:

$$(\mathcal{P}_s\delta s)\omega = \langle \delta s, \mathcal{E}_s\omega\rangle + \mathrm{d}i[\operatorname{symb}\mathcal{P}_s(\ ,\delta s)]\omega \tag{2}$$

where $\operatorname{symb}\mathcal{P}_s$ is the symbol of $\mathcal{P}_s$, and the fact that δs vanishes on ∂A allow us to write (1) as:

$$\int_A \langle \delta s, \mathcal{E}_s\omega\rangle = 0$$

for every δs whose support is interior to A.

This allows us to conclude that

$$\mathcal{E}_s\omega = 0 \quad \text{on the interior of } A \tag{3}$$

which are Euler–Lagrange equations that characterize the stationarity of the section $s \in \Gamma(X, Y)$.

Formula (2) is thus the key element to obtain Euler–Lagrange equations of the variational problem. In this formalism, it expresses the **first variation formula** of the Lagrangian density $\mathcal{L}\omega$. From this formula we also obtain the so-called Noether's Theorem of the calculus of variations: If D is a vector field on Y, π-projectable on a vector field $\bar{D}$ on X and such that $L_{j^1 D}\mathcal{L}\omega = 0$ (infinitessimal symmetry), then for each stationary section s:

$$\mathrm{d}i[\mathcal{L}(j^1 s)\bar{D} + (\operatorname{symb}\mathcal{P}_s)(\ , D_s^v)]\omega = 0\,,$$

where $D_s^v = D \circ s - s_*\bar{D}$ is the π-vertical component of D along s.

All this is explained in Sections 1 and 2 of the monograph. Section 3 constitutes the heart of the work, where the formalism of Hamilton–Cartan is introduced by means of the concepts of momentum form, Poincaré–Cartan form and Legendre transformation of a variational problem.

The **momentum form** is the $(n-1)$-form on J^1Y with values on the bundle p^*V^*Y ($p\colon J^1Y \to Y$ the natural projection):

$$\Omega_{j_x^1 s} = (i[\operatorname{symb}\mathcal{P}_s]\omega)_x\ . \tag{4}$$

The **Poincaré–Cartan form** is the ordinary n-form on J^1Y:

$$\Theta = \theta\bar{\wedge}\Omega + \mathcal{L}\omega\,, \tag{5}$$

where θ is the 1-form on J^1Y with values on p^*VY introduced in Proposition 1.1 of the monograph and where $\bar{\wedge}$ is the exterior product with respect to the duality between p^*VY and p^*V^*Y.

Finally, the **Legendre transformation** is the bundle morphism Leg: $J^1Y \to \pi^*TX \otimes V^*Y$ (of bundles over Y) defined by:

$$i[\operatorname{Leg}(j_x^1 s)]\omega = \Omega_{j_x^1 s}\,. \tag{6}$$

With these new objects, the main observation in the Hamilton–Cartan formalism of the calculus of variations is the following:

Theorem 1 *If the mapping* Leg: $J^1Y \to \pi^*TX \otimes V^*Y$ *is an immersion (regular problems) and* $u \in \Gamma(X, J^1Y)$ *is a section such that:*

$$u^*(i_\xi \mathrm{d}\Theta) = 0,\ \textit{for every } \xi \in \mathfrak{X}(J^1Y)$$

then, if $s = p \circ u \in \Gamma(X, Y)$, *we have:*

$$u = j^1 s, \quad \mathcal{E}_s\omega = 0\,,$$

that is, s *is a stationary section of the variational problem.*

From here the route is clear to complete the theory:

So, in Section 4 (the Poisson bracket), if $\mathcal{V} = \{s \in \Gamma(X, Y) : \mathcal{E}_s\omega = 0\}$ is the set of stationary sections and if for each $s \in \mathcal{V}$ we define the tangent space $T_s\mathcal{V}$ to be the vector space of solutions $\delta s \in \Gamma(X, s^*VY)$ of the linearized Euler–Lagrange equations, then the $(n+1)$-form $\mathbf{\Omega} = \mathrm{d}\Theta$ defines, when restricted to those solutions, a hemi-symmetric metric on $\mathcal{V}$ with values on the set of functions from $\mathcal{V}$ to closed $(n-1)$-forms of X, which represents a generalization for the calculus of variations of the symplectic manifold structure of analytical mechanics. This is the "multi-symplectic manifold of solutions", where the theory of Noether symmetries can be interpreted in a dynamic way. More precisely: infinitesimal symmetries induce vector fields on the multi-symplectic manifold of solutions, which are the multisymplectic gradients of its corresponding conserved currents.

Section 5 is devoted to the theory of Hamilton–Jacobi. It is the first formulation in the language of modern differential geometry of the theory of geodesic fields of De Donder and Weyl [1, 13], in which criteria are given for the minimality of a stationary section immersed in a geodesic field, generalizing to several independent variables results of Hilbert and Weierstrass of the late 19th century. It must be pointed out that this formulation has been the only one to be found in the modern literature until very recently.

Finally Section 6 deals with Morse theory of the second variation with particular emphasis, as could be expected, on its Hamiltonian formulation.

2. Influence on my subsequent research

Shortly after its publication, I read with great interest the article by Goldschmidt and Sternberg, whose subject and results agreed with previous research of my own while trying to formulate a symplectic quantum field theory [4, 5]. I first noticed the interest of the subject itself in the framework of the geometrical theory of the calculus of variations; a matter that I had also studied in a work with the title "The Poincaré–Cartan invariant in the calculus of variations", which I presented at the Convegno di Geometria Simplettica e Fisica Matematica, that took place in January, 1973, at the Instituto Nazionale di Alta Matematica in Rome [3].

Unlike the approach of Goldschmidt and Sternberg, the starting point of [3] is the geometry of the 1-jet bundle, J^1Y, defined by its structure 1-form

$$\theta\colon j^1_x s \mapsto \theta_{j^1_x s},$$

where:

$$\theta_{j^1_x s}(\xi_{j^1_x s}) = (\mathrm{d}^v s)_x(p_* \xi_{j^1_x s}) = p_* \xi_{j^1_x s} - (s \circ \pi)_* p_* \xi_{j^1_x s}\,, \quad \xi_{j^1_x s} \in T_{j^1_x s}(J^1Y)\,.$$

This is a 1-form on J^1Y with values on the vector bundle p^*VY, which justifies as methodology the use of differential calculus on J^1Y with values on the vector bundle p^*VY and its various derived bundles (p^*V^*Y, $\mathrm{End}(p^*VY)$, etc.).

In particular, in terms of this geometric structure we can characterize the 1-jet extension of sections and of vector fields on Y: a section $u \in \Gamma(X, J^1Y)$ is the 1-jet extension of a section $s \in \Gamma(X, Y)$ (that is, $u = j^1 s$) if and only if $u^*\theta = 0$, and a vector field ξ on J^1Y is the 1-jet extension of a vector field D on Y (that is, $\xi = j^1 D$) if and only if there exists a section $f \in \Gamma(J^1Y, \mathrm{End}(p^*VY))$ such that $L_\xi \theta = f \circ \theta$, where the Lie derivative is taken with respect to any linear connection on p^*VY and the product $\circ$ is given by the natural duality product.

The idea now is to define the stationarity of a section $s \in \Gamma(X, Y)$ by:

$$\int_X (j^1 s)^* L_{j^1 D} \mathcal{L}\omega = 0 \tag{7}$$

for every vector field D on Y whose support projects onto a compact subset of X, and to characterize this notion proceeding as follows:

We can define the momentum form associated to the Lagrangian density $\mathcal{L}\omega$ as the unique $(n-1)$-form Ω on J^1Y with values on p^*V^*Y such that $\Omega = i_F\omega$, where F is any vector field on J^1Y with values on p^*V^*Y solution of the equation:

$$i_F \mathrm{d}_\nabla \theta = \mathrm{d}\mathcal{L}$$

on p-vertical vector fields of J^1Y, where the exterior derivative is taken with respect to the connection on p^*VY induced by any linear connection on VY and where the bilinear products are the obvious ones.

Using this notion, we can introduce the Poincaré–Cartan form and the Legendre transformation using formulas (4) and (5), respectively.

The key point of the formalism is now the following:

Lemma 2 *There exists a unique $j^1\pi$-horizontal n-form $\mathcal{F}(\nabla)$ on J^1Y with values on p^*V^*Y such that:*

$$\mathrm{d}\Theta = \theta \bar{\wedge} (\mathcal{F}(\nabla) - \mathrm{d}_\nabla \Omega)\,, \tag{8}$$

where the exterior derivative d_∇ *is taken with respect to the connection on* p^*V^*Y *induced by a linear connection* ∇ *on* VY *with vanishing vertical torsion (that is,* $\nabla_{D_1}D_2 - \nabla_{D_2}D_1 - [D_1, D_2] = 0$ *for any pair of* π*-vertical vector fields* D_1, D_2 *on* Y*).*

The n-form $\mathbb{E}(\nabla) = \mathcal{F}(\nabla) - d_\nabla\Omega$ (Euler–Lagrange form of $\mathcal{L}\omega$ with respect to the connection ∇) allows us to obtain the Euler–Lagrange operator $\mathcal{E}_s\omega$ as:

$$\mathcal{E}_s\omega = (j^1s)^*\mathbb{E}(\nabla)\,. \tag{9}$$

From here, the following results are readily obtained:

Theorem 2 (First variation formula) *For each vector field* D *on* Y *there exists an* $(n-1)$*-form* α *with values on* p^*V^*Y *such that:*

$$L_{j^1D}(\mathcal{L}\omega) = \theta(j^1D)\circ\mathbb{E} + d(i(j^1D)\Theta) + \theta\bar{\wedge}\alpha\,,$$

or, if we restrict to the 1*-jet extension* j^1s *of any section* $s\in\Gamma(X,Y)$*:*

$$(j^1s)^*L_{j^1D}(\mathcal{L}\omega) = \langle D^v_s, \mathcal{E}_s\omega\rangle + d(j^1s)^*(i(j^1D)\Theta)\,.$$

Corollary 1 (Euler–Lagrange equation) *A section* $s\in\Gamma(X,Y)$ *is stationary if and only if:*

$$\mathcal{E}_s\omega = 0\,.$$

Corollary 2 (Cartan equation) *A section* $s\in\Gamma(X,Y)$ *is stationary if and only if for every vector field* ξ *on* J^1Y *we have:*

$$(j^1s)^*i_\xi d\Theta = 0\,.$$

Corollary 3 (Noether theorem) *If* $L_{j^1D}\mathcal{L}\omega = 0$ *and* $s\in\Gamma(X,Y)$ *is stationary then:*

$$d(j^1s)^*(i(j^1D)\Theta) = 0$$

Also, as we might expect, there follows Theorem 1 under the same regularity assumption.

In 1974 I met Sternberg personally and since then we kept a good scientific relation regarding the above described research and other related subjects (geometric quantization, hamiltonian reduction, super-symmetry, etc.). In those years he proposed me to generalize the Hamilton–Cartan formalism to higher order variational calculus. He had done it for $X=\mathbb{R}$ (higher order mechanics) with an axiomatic characterization of a Poincaré–Cartan form that allowed him to prove the equivalence of the Euler–Lagrange and Hamilton equations under a certain regularity condition [12].

At the beginning of the 80's, J. Muñoz-Masqué and myself solved the problem by applying the methodology described above for the structure 1-form θ^k of the k-jet bundle J^kY ($k \geq 1$), which is now a 1-form on J^kY with values on the induced vector bundle $(VJ^{k-1}Y)_{J^kY}$ (where $VJ^{k-1}Y$ is the bundle of vector fields on $J^{k-1}Y$ that are vertical over X) [6, 11]. In those works we obtained a family $\{\Theta(\nabla)\}$ of Poincaré–Cartan forms on $J^{2k-1}Y$, parameterized by the set $\{\nabla\}$ of linear connections on the base manifold X, such that for each section $s \in \Gamma(X, Y)$ and each vector field D on Y the following first variation formula holds:

$$(j^k s)^* L_{j^k D} \mathcal{L}\omega = \langle D_s^v, \mathcal{E}_s \omega \rangle + \mathrm{d}\left((j^{2k-1}s)^* i_{j^{2k-1}D} \Theta(\nabla) \right),$$

where $\mathcal{E}_s \colon \Gamma(X, \Lambda^n T^* X) \to \Gamma(X, s^* V^* Y \otimes \Lambda^n T^* X)$ is the Euler–Lagrange operator of the variational problem.

From this point on, it is easy to generalize Corollaries 1, 2 and 3 of first order problems. In particular, for $k \leq 2$ and arbitrary n, or for $X = \mathbb{R}$ and arbitrary k, the family $\{\Theta(\nabla)\}$ consists of a single Poincaré–Cartan form, that coincides with the ones that were already known for those cases. More problematic was the regularity issue, since for $k > 1$, $n > 1$ there is no equivalence between the Euler–Lagrange equations and the Hamilton equations. This drove us into a closer study of this concept in [7, 8]. In particular, in certain situations of "reducibility" it is possible to derive a complete equivalence of the reduced Lagrange and Hamiltonian formalisms.

Let us finally say that this formalism, when appropriately combined with the original one by Goldschmidt and Sternberg has more recently allowed us to obtain new results on Stress-Energy-Impulse tensors of natural variational problems [2] and on the problem of Lagrange on fibred manifolds [9], two subjects of current interest in the non-holonomic and vakonomic theory of fields of geometric mechanics.

This is, in outline, my personal experience with this magnificent monograph by Goldschmidt and Sternberg, whose account I wish to dedicate to the other researchers of those years (R. Herman, D. Krupka, J. Kijowski, I. Kolar, M. Ferraris, M. Gotay, J. Marsden…) for whom this work was also a motive of inspiration.

References

[1] De Donder, Th.: *Théorie invariantive du calcul des variations.* Nouvelle edit. Paris: Gauthier-Villars & Cie. XXI, 1935.

[2] Fernández, A., García, P. L. and Rodrigo, C.: Stress-energy-momentum tensors in higher order variational calculus. *J. Geom. Phys.* **34** (2000), no. 1, 41–72.

[3] García, P. L.: The Poincaré–Cartan invariant in the calculus of variations. *Symposia Math.* **14** (1974), 219–246.

[4] García, P. L. and Pérez-Rendon, A.: Symplectic approach to the theory of quantized fields I. *Comm. Math. Phys.* **13** (1969), 24–44.

[5] García, P. L. and Pérez-Rendon, A.: Symplectic approach to the theory of quantized fields II. *Arch. Rat. Mech. Anal.* **43** (1971), no. 2, 101–124.

[6] García, P. L. and Muñoz-Masqué, J.: On the geometrical structure of higher order variational calculus. *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* **117** (1983), no. 1, 127–147.

[7] García, P. L. and Muñoz-Masqué, J.: Le problème de la régularité dans le calcul des variations du second ordre. *C. R. Acad. Sci. Paris Sér. I Math.* **301** (1985), no. 12, 639–642.

[8] García, P. L. and Muñoz-Masqué, J.: Higher order regular variational problems. In *Symplectic geometry and mathematical physics (Aix-en-Provence, 1990)*, 136–159. Progr. Math. **99**. Birkhäuser, Boston, 1991.

[9] García, P. L., García, A. and Rodrigo, C.: Cartan forms for first order constrained variational problems. *J. Geom. Phys.* **56** (2006), 571–610.

[10] Goldschmidt, H. and Sternberg, S.: The Hamilton–Cartan formalism in the calculus of variations. *Ann. Inst. Fourier (Grenoble)* **23** (1973), no. 1, 203–267.

[11] Muñoz Masqué, J.: Poincaré–Cartan forms in higher order variational calculus on fibred manifolds. *Rev. Mat. Iberoamericana* **1** (1985), no. 4, 85–126.

[12] Sternberg, S.: Some preliminary remarks on the formal variational calculus of Gelfand and Dikii. Lecture Notes in Math. **676**, 399–407. Springer Verlag, 1978.

[13] Weyl, H.: Geodesic fields in the calculus of variations for multiple integrals. *Ann. of Math.* **36** (1935), 607–629.

EVARIST GINÉ (Falset, Catalonia, 1944) obtained his PhD in mathematics in 1973 at MIT (adviser: R. M. Dudley). He has been a professor at the University of Connecticut since 1990 and his research focuses on: probability in Banach spaces, empirical processes, U-processes and U-statistics, probabilistic limit theorems and non-parametric statistics.

Some of the articles that influenced my work

Evarist Giné*

Mathematics research very seldom happens out of a social/scientific environment, and the scientific part of this environment is mostly defined by one's peers and colleagues, seminars, talks, conversations at meetings, and, last but not least, the mathematics one reads.

I'll try to make a short account of some aspects of my mathematics experience, as it pertains to reading mathematics. Before starting however, I should mention that oral, colloquial communication is at least as important as reading. Interchange of ideas does often –or at least some times– generate new ones, and conversation about mathematics very often provides insights that are impossible to obtain by reading alone, or even from formal talks.

Seldom in my career have I read entire books in order to learn a whole theory, except when I was a student. More often, I have read new papers (new at the time) or preprints that were dealing with whatever subject I was researching at the time, or I have read parts of many old and new papers in order to become knowledgeable about what I was working on, or in search of ideas or looking for clues to solve a specific problem.

My PhD thesis was on a very concrete and in a sense isolated problem: use Sobolev norms to construct invariant tests of uniformity on the sphere that can be effectively computed (say $O(n)$ operations from n data). After my thesis and some other work were completed in 1973 I did not have a clue about what to do next, and remember naively asking my adviser R. M. Dudley how should I go about finding problems on my own: I would have preferred to be part of a large team trying to complete a central and ambitious program of research, but instead I was going back to Venezuela, where there was only one more probabilist at the time, and wanted to know what to do a few months from then. Of course, his answer was: READ the journals. I read some of his work in search of open problems to solve, which was enough to keep me going for a few months, until I went to Berkeley in 1974 and was invited to Oberwolfach for a conference on Probability in Banach spaces. Thanks to this conference and the preprint of a 1977 paper by Joel Zinn (later my most frequent coauthor and a very good friend), where he was showing how the relatively abstract methods of Pisier, Hoffmann–Jørgensen and others would solve a problem I had worked on, about the central limit theorem for processes with continuous trajectories, I ended up in the middle

*gine@math.uconn.edu.

of a subject that at the time was undergoing vigorous development: Probability in Banach spaces was a relatively large and unexplored field, and this conference and Joel's paper directed me to it. So, reading journals did not help me acquire my first extensive research program, it was more a question of the right conference at the right time, and a short but inspiring preprint. I worked on this subject mainly with Joel Zinn and Aloisio Araujo. The works that influenced us more in this area, besides Zinn's, are Dudley and Strassen (1969), Hoffmann–Jørgensen (1974), Pisier's thesis (1976), his paper with Hoffmann–Jørgensen (1976) and Jain and Marcus (1975).

The next large subject I worked on was empirical processes, and in this case, yes, reading a published paper was determinant. Dudley (1979) had written a landmark paper on the subject, and when in 1982 we moved to Texas A&M with J. Zinn and M. Marcus, it happened that Zinn and I, who had already written several papers together on Probability on Banach spaces, both had the project of studying Dudley's paper. We started reading it together and never really finished since soon after we started, in the middle of the proof of one of the main results, we got ideas on how the subject should develop and started working on it, which we did for about ten years! I was lucky to have the right paper to inspire me and I was even luckier to have the stimulating company of my coauthor, which made 'creative reading' possible and fun. Zinn and I ended up writing twenty five articles on Probability in Banach spaces, empirical processes, bootstrap (a statitical method, whereby one uses data to simultaneously make predictions and assess their quality, in a very nice way), U-statistics, and other subjects, but mostly empirical processes. While working on Probability in Banach spaces and later on empirical processes we were part of a relatively large group of people, that ranged from mathematical statisticians to pure mathematicians working on Gaussian processes and more abstract probability. We were exchanging papers and reading them before they would get published, and this was a sort of an ideal situation to be in. One of the best results in this area (to be commented upon below) is due to Michel Talagrand, and his first contact with the subject were conversations with Zinn and I at a meeting on Probability in Banach spaces held in New England: he asked both Joel and I, separately, to explain to him our first main contribution to the subject, an invited paper in Annals of Probability, including open problems. Very soon we saw these problems solved, and more.

We moved again, in 1988, to New York and, in 2000, to Connecticut, and I brought with me a very good student, Miguel Arcones (who passed away two years ago). At least for me, subjects always end up losing their appeal: it is like you end up tired and having nothing else to say about them. When I thought that empirical processes and the bootstrap were exhausted for me, it was time to move to some other area. Two papers directed me to U-processes and U-statistics, one by Jean Bretagnolle (on the bootstrap of U-statistics, 1983) and two by David Pollard and Deborah Nolan (on U-processes, 1987, 1988). Each of them had left things for others to do, so to speak, and it also became clear, after looking at the field, that some important basic questions on U-statistics had not been totally resolved –law of large numbers, law of the iterated logarithm, for example–. I knew from Zinn about a technique in martingale theory called decoupling and asked Víctor de la Peña, who attended our seminar, whether it would be possible to 'decouple U-statistics'. He came up with a very nice solution to the problem, that later was strengthened by himself and Montgomery–Smith, and this gave great impetus to the theory of U-statistics, U-processes and multinomial forms in independent random variables. I was once more in a situation where I had no need to look for problems –the problems were staring at my face– and could work with several people on different aspects of the theory of U-statistics and U-processes for several years (Arcones, Latała, Kwapién, Zhang, Zinn, Koltchinskii, and a book with De la Peña).

I like applications, or perhaps more exactly, I like for my work to have applications, whatever this means –in my case, it mainly means applications in mathematical statistics, itself still at some degrees of separation from 'real world' applications–. In this respect, empirical process theory is quite rewarding as it is one of the main foundations of non-parametric statistics (asymptotic or not). U-statistics are also applicable but a little less; they appear mainly as higher order terms in Taylor developments of statistical functions on a space of probability measures, and they are also related to multilinear forms in independent random variables, like *e.g.*, Gaussian chaos. Even while I was working on U-statistics, I was looking for subjects where they could be of use. I found more than one, but specially I could develop, or several of us could develop, one of them, namely, density estimation (some aspects of this large subject). Winfried Stute (1993) had been applying U-processes to study the empirical distribution function for censored data. I asked a graduate student to look at his papers and see if he could do better and/or attack related problems using the new developments on the theory of U-statistics. But he left for a well paying job in information technologies. I took the subject with A. Guillou, a short term visitor from Paris. We found a very good paper on the same subject (S. Csorgo, 1996), which was inspiration for our first results. Then, after the distribution function comes the density, and while obtaining a law of the logarithm for the density in the censored data case, I discovered that a similar result did not exist in the regular uncensored case, and was easier to prove given Talagrand's inequality! This was the start a new research topic that has lasted for more than ten years, density estimation, and which has generated several excellent collaborations. When working in density estimation several papers have been particularly useful. The main one is Talagrand (1996): he proved an exponential inequality for tail probabilities of suprema of collections of sums of independent random variables: they concentrate about their mean at the same rate as the worst individual sum in the collection. I believe this result to be one of the best in Probability in at least the last 25 years. I had used this inequality in a paper about the Wasserstein distance between the empirical and the true distributions with del Barrio and Matrán in 1999, but I have used it in at least ten articles on density estimation since then. U. Einmahl and D. Mason (2000) actually introduced the use of Talagrand's inequality in that subject. Their paper had an influence in may main article with Guillou. I had my program, that overlapped with other people's program and this led to very nice and fruitful collaborations with Zinn and Koltchinskii once more, and particularly, on this subject, with D. Mason and with R. Nickl. In my work with Nickl on adaptive estimation of densities (the rate of uniform or L_p approximation of a density estimator to the true density depends on the smoothness of the true density, and an estimator is adaptive if it estimates the density at the best rate even if you do not know its smoothness –then you don't know the rate, you only know that you are doing best possible–) there were two articles that had a great influence on our work, namely those of Donoho, Johnstone, Kerkyacharian and Picard (1996) and Lepski (1991) together with Lepski and Spokoiny (1997) where the idea of adaptation is put forward in different ways. Recently, the work of van der Vaart and collaborators on frequentist properties of Bayesian estimation in infinite dimensions has also influenced the work of Nickl and I.

I have only described the influence of other people's work on the main directions of research I have had over the years. Besides those, I have worked on occasion on other problems either because a paper I refereed or reviewed inspired me, or as a result of conversations with or questions from colleagues. Among this type of collaborations, besides one with Koltchinskii about estimating the Laplacian of a compact manifold by sampling from the manifold according to the normalized volume element, perhaps the most notable is one with F. Götze and D. Mason, where we proved that the Student t-statistic is asymptotically standard normal if and only if the independent data come from a distribution in the domain of attraction of the normal law. The origin of this problem were papers by Logan, Mallows, Rice and Shepp (1973), where the problem we solved is more or less stated, and Griffin and Mason (1991), where it is solved in the symmetric case, and the interesting part of this collaboration is that it was specifically designed by F. Götze for the solution of the problem.

I conclude by listing the mentioned references, but not my work.

References

[1] Bretagnolle, J.: Lois limites du bootstrap de certaines fonctionelles. *Ann. Inst. H. Poincaré Sect B* **19** (1983), 281–296.

[2] Csörgő, S.: Universal Gaussian approximations under random censorship. *Ann. Statist.* **24** (1996), 2744–2778.

[3] Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D.: Density estimation by wavelet thresholding. *Ann. Statist.* **24** (1996), 508–539.

[4] Dudley, R. M.: Central limit theorems for empirical measures. *Ann. Probab.* **6** (1978), 899–929.

[5] Einmahl, U. and Mason, D.: An empirical process approach to the uniform consistency of kernel type function estimators. *J. Theoret. Probab.* **13** (2000), 1–37.

[6] Griffin, P. and Mason, D.: On the asymptotic normality of self-normalized sums. *Math. Proc. Cambridge Phil. Soc.* **109** (1991), 597–610.

[7] Jain, N. and Marcus, M.: Central limit theorems for $C(S)$-valued random variables. *J. Functional Analysis* **19** (1975), 216–231.

[8] Hoffmann-Jørgensen, J.: Sums of independent Banach space valued random variables. *Studia Math.* **52** (1974), 259–266.

[9] Hoffmann-Jørgensen, J. and Pisier, G.: The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.* **4** (1976), 587–599.

[10] Lepski, O.: Asymptotic minimax adaptive estimation I. Upper bounds. *Theorey Probab. Appl.* **36** (1991), 645–659.

[11] Lepski, O. and Spokoiny, V.: Optimal pointwise adaptive methods in nonparametric estimation. *Ann. Statist.* **25** (1997), 2512–2546.

[12] Logan, B. F., Mallows, C. L., Rice, S. O. and Shepp, L. A.: Limit distributions of self-normalized sums. *Ann. Probab.* **1** (1973), 788–809.

[13] Pisier, G.: Le théorème limite centrale et la loi du logarithme iteré dans les espaces de Banach. I et II. *Séminaire Maurey-Schwartz (1975-1976): Espaces L^p, applications radeonifiantes et géométrie des espaces de Banach* Exp. 3 et 4. Centre Math. École Polytechnique, Palaiseau, 1976.

[14] Pollard, D. and Nolan, D.: U-processes, rates of convergence. *Ann. Statist.* **15** (1987), 780–799.

[15] Pollard, D. and Nolan, D.: Functional limit theorems for U-processes. *Ann. Probab.* **16** (1988), 1291–1298.

[16] Strassen, V. and Dudley, R. M.: The central limit theorem and ε-entropy. In *1969 Probability and Information Theory*, 224–231. Springer, Berlin.

[17] Stute, W.: Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21** (1993), 146–156.

[18] Talagrand, M.: New concentration inequalities in product spaces. *Invent. Math.* **126** (1996), 505–563.

[19] Zinn, J.: A note on the central limit theorem in Banach spaces. *Ann. Probab.* **5** (1977), 283–286.

Tadeusz Iwaniec (Elbląg, Poland, 1947) obtained his PhD in mathematics at the University of Warsaw in 1975 (advisor: Bogdan Bojarski). Currently, he is John Raymond French Professor of Mathematics at Syracuse University and Finland Distinguished Professor (FiDiPro) at the University of Helsinki. His main research areas are: multidimensional geometric function theory, quasiconformal mappings, nonlinear partial differential equations, harmonic analysis, calculus of variations and mathematical theory of nonlinear hyperelasticity.

Let the beauty of Harmonic Analysis be revealed through nonlinear PDEs

A work of art in three sketches

Tadeusz Iwaniec*

Prologue

Mathematicians, like me, have the privilege to enjoy the ingenious ideas and splendid theories imagined and brilliantly developed by previous mathematicians. Their beauty inspires us to ask questions to create our own little theory with grace and prospective applications. And there is never an end to new questions, which in effect is the key to advances in mathematics. But advances come after hours and hours of intense work, trapping and holding our attention for years. This can be a dream, sometimes immense pleasure, sometimes a breathtaking moment when we spot the underlying ideas that are actually relevant to our aspirations. Genuine mathematics does not abide in complexity but, contrary to what one might think, somewhere in the unlimited beauty of applications of sophisticated ideas. Paraphrasing Luciano Pavarotti on music, let me say:

> *"Learning mathematics by only reading about it is like making love by e-mail".*

From the very beginning of my mathematical life I fell in love with logic and later as a young scholar with geometry and harmonic analysis. There are so many captivating topics in geometric analysis. I was especially fascinated by the foundation of Geometric Function Theory (GFT, quasiconformal mappings), the mysteries in the Calculus of Variations (nonconvex energy integrals) and Nonlinear Partial Differential Equations (PDEs, elliptic type). Nowadays, these fields are essential in material science and nonlinear elasticity, which are critical in modern technology and many engineering problems. Myriad practical problems of nonlinear elasticity and numerous elegant conjectures are very appealing to me. But I cannot fully treat these topics here. I will only indicate briefly a few adventurous moments of my studies on these topics by means of applications of maximal functions due to *Hardy* and *Littlewood*, *Fefferman* and *Stein*, as well as nonlinear commutators which originated with *Coifman*, *Rochberg* and *Weiss*.

*tiwaniec@syr.edu.

The results presented here are not my best, though there is some element of aesthetic beauty in them. It is for these reasons that:

Fefferman *and* ***Stein*** *(the architects of maximal inequalities),*
Coifman, ***Rochberg*** *and* ***Weiss*** *(the founders of singular commutators)*
became my mathematical luminaries.

FIGURE 1: During the Sierpiński Medal lecture (Warsaw, 2009). Title of the lecture: *An Invitation to Quasiconformal Hyperelasticity.*

Some further results (on quasiconformal mappings in even dimensions) are discussed in this issue by my amazingly imaginative colleague *Gaven Martin*[1]. We have presented GFT in all dimensions in our book [12]. Well, we did not make a fortune with this book, nor did we become famous. But I have heard someone say, “Hey, I have read your book”. How satisfying!

As I share the beauty and joy of mathematics with you I also remember Polish mathematicians whose glorious scientific careers came to a cruel end during Nazi-Soviet occupation. *Józef Marcinkiewicz*, *Stanisław Saks* and *Juliusz Paweł Schauder* were inspirations to me. I am mindful of them not only as mathematicians [10]. Marcinkiewicz, along with 22 thousand Polish patriots who dared to exhibit a love and pride of an independent Poland, were executed by the order of J. Stalin, and buried secretly in mass graves in gloomy forested sites near Starobielsk, Ostashkovo and the most documented Katyń.

Someday maybe a great musician will rise up,
will transform speechless rows of gravestones into a keyboard,
a great Polish song writer will compose a frightening ballad with blood and tears.
[...]
And there will emerge untold stories,
strange hearts, bodies bathed in light...
And the Truth again will embody
The Spirit
with living words-of the sand of Katyń.

Katyń Carol. KAZIMIERA IŁŁAKOWICZÓWNA
(translated by the author of this article)

[1] Gaven speaks of me as “After all, he has quite a good memory even if it is a bit short”. Yes indeed, I have a good memory for masterpieces, but a short one for trivia.

Professor Antoni Zygmund remarked once about Marcinkiewicz:

> "[...] his early death may be seen as a great blow to Polish Mathematics, and probably its heaviest individual loss during the Second World War."

I have had the privilege of growing up in the environment these mathematicians left for us.

1. The natural domain of definition

While singular integrals are naturally defined in $\mathscr{L}^2(\mathbb{R}^n)$, there is also such a thing as the natural space in which we look for the solutions of a differential equation; just to mention a few of those readily seen as being natural:

- The Sobolev space $\mathscr{W}^{1,2}(\Omega)$ for the Laplacian.
- The Sobolev space $\mathscr{W}^{1,p}(\Omega)$ for the p-harmonic operator

$$\operatorname{div}\left(|\nabla|^{p-2}\nabla\right) = 0\,.$$

- The space $\mathscr{W}^{1,n}(\mathbb{X},\mathbb{Y})$ for quasiconformal mappings $f : \mathbb{X} \xrightarrow{\text{onto}} \mathbb{Y}$ between n-manifolds, in which the major player is the Jacobian determinant $J(x,f)\,\mathrm{d}x = f^\sharp(\mathrm{d}y)$-pullback of the volume form in $\mathbb{Y}$.

There is no genuine distinction between linear and nonlinear differential operators. Indeed, once we depart from their natural domain of definition the application of singular integrals in the extended settings such as: $\mathscr{L}^p(\mathbb{R}^n) \to \mathscr{L}^p(\mathbb{R}^n)$, $\mathscr{L}^1(\mathbb{R}^n) \to \mathscr{L}^1_{\text{weak}}(\mathbb{R}^n)$, $\mathscr{H}^1(\mathbb{R}^n) \to \mathscr{L}^1(\mathbb{R}^n)$ and $\mathscr{L}^\infty(\mathbb{R}^n) \to \mathrm{BMO}(\mathbb{R}^n)$, becomes equally pressing in both cases. But first we need some definitions.

2. Maximal operators

Maximal inequalities, traditionally discussed in the entire space $\mathbb{R}^n$, can actually be considered for functions $f \in \mathscr{L}^1_{\text{loc}}(\Omega)$ on any open set $\Omega \subset \mathbb{R}^n$:

- *Hardy–Littlewood maximal function* [8] (1930),

$$\mathbf{M}f(x) = \sup\left\{ \fint_B |f(y)|\,\mathrm{d}y\,;\;\; B = B(x,r)\subset\Omega \right\}, \qquad \fint_B = \frac{1}{|B|}\int_B .$$

- *Fefferman's sharp operator* [4] (1971),

$$\mathbf{M}^\sharp f(x) = \sup_B \left\{ \fint_B \left|f(y) - f_B\right|\mathrm{d}y\,;\;\; B = B(x,r)\subset\Omega \right\}.$$

- *Spherical operator of E. Stein* [21],

$$\mathbf{S}f(x) = \sup_{\partial B} \Big\{ \fint_{\partial B} |f(y)|\, \mathrm{d}y\,; \quad B = B(x,r) \subset \Omega \Big\}.$$

Theorem 1 (Three Fundamental Maximal Inequalities) *For every* $f \in \mathscr{L}^q(\mathbb{R}^n)$ *with* $1 < q < \infty$ *, we have*[2]

$$\| f \|_{\mathscr{L}^q(\mathbb{R}^n)} \preccurlyeq (q-1) \cdot \| \mathbf{M}f \|_{\mathscr{L}^q(\mathbb{R}^n)} \preccurlyeq \| f \|_{\mathscr{L}^q(\mathbb{R}^n)} . \tag{1}$$

If $f \in \mathscr{L}^1(\mathbb{R}^n) \cap \mathscr{L}^q(\mathbb{R}^n)$ *and* $1 < q < \infty$ *, then*

$$\| \mathbf{M}f \|_{\mathscr{L}^q(\mathbb{R}^n)} \preccurlyeq \| \mathbf{M}^\sharp f \|_{\mathscr{L}^q(\mathbb{R}^n)} \preccurlyeq \| \mathbf{M}f \|_{\mathscr{L}^q(\mathbb{R}^n)} \tag{2}$$

In both inequalities (1) *and* (2) *the implied constants stay bounded as* q *approaches 1.*

If $f \in \mathscr{L}^s(\mathbb{R}^n)$ *with* $s > \frac{n}{n-1}$ *, then*

$$\| \mathbf{S}f \|_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq \| f \|_{\mathscr{L}^s(\mathbb{R}^n)} \tag{3}$$

This time the implied constant blows up as s *approaches* $\frac{n}{n-1}$.

Our discussion of the Hardy space $\mathscr{H}^1(\Omega)$ becomes somewhat simpler if we confine ourselves to a rotationally invariant approximation to the identity. Thus we choose and fix a function $\Phi \in \mathscr{C}_\circ^\infty[0,1)$ such that $\int_{\mathbb{R}^n} \Phi(|x|)\, \mathrm{d}x = 1$, and set $\Phi_t(x) = t^{-n}\Phi\big(\frac{|x|}{t}\big)$. Given any $F \in \mathscr{L}^1_{\mathrm{loc}}(\Omega)$, we smooth it as

$$\big(F * \Phi_t\big)(x) = \int_\Omega \Phi_t(x-y)\, F(y)\, \mathrm{d}y\ , \qquad 0 < t < \mathrm{dist}(x, \partial\Omega).$$

A maximal operator that accounts for cancellations of positive and negative terms is now given by

- $\mathcal{M}F(x) = \sup \big\{ \, \big|F * \Phi_t\big|(x) \ : \ 0 < t < \mathrm{dist}(x, \partial\Omega) \, \big\}.$

- *The Hardy Space* $\mathscr{H}^1(\Omega)$ consists of functions $F \in \mathscr{L}^1(\Omega)$ such that

$$\big\|F\big\|_{\mathscr{H}^1(\Omega)} \stackrel{\mathrm{def}}{=\!=} \big\|\mathcal{M}F\big\|_{\mathscr{L}^1(\Omega)} \ < \ \infty\,.$$

[2] Hereafter the notation $\preccurlyeq$ refers to inequalities with the so-called *implied constants* in the right hand side, which vary from line to line. Their precise values are readily perceived from the context. We shall indulge in this harmless convention for aesthetic reasons.

The above maximal operators, brilliantly developed by C. Fefferman and E. Stein, not only gave birth to a new discipline to effectively handle singular integrals: they also provided Geometric Analysts, like me, with the means of solving demanding problems in Geometric Function Theory (GFT) and nonlinear PDEs. Maximal inequalities saved us from laborious computation once used in a clever, sometimes artistic, way. Let us take on stage, as the first sketch, the Jacobian determinant $F = J(x, f) = \det Df(x)$ of the differential matrix $Df(x) \in \mathbb{R}^{n\times n}$ of a mapping $f : \Omega \to \mathbb{R}^n$ [13].

Sketch I: Let Hardy and Littlewood meet Fefferman and Stein

It is quite easily seen that for the rotationally invariant approximation of unity, we have

$$\big|F * \Phi_t\big|(x) \leqslant \frac{C_\Phi}{t^{n+1}} \int_0^t \Big| \int_{B(x,r)} F(y)\,\mathrm{d}y \Big|\,\mathrm{d}r \,, \qquad C_\Phi = \big\|\Phi'\big\|_{\mathscr{L}^\infty[0,1)} \tag{4}$$

It should be noted that the absolute value is administered only upon integrating F over the ball $B(x, r)$. Such an observation, though elementary, is vital when dealing with *null-Lagrangians.* Null-Lagrangians, like Jacobians, are the nonlinear differential expressions whose integral mean over any subdomain reduces to the boundary integral, basically due to cancellation of second order partial derivatives when integrating by parts.

Theorem 2 [13] *Suppose* $f \in \mathscr{W}^{1,n-1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ *and the matrix of cofactors* $D^\sharp f \in \mathscr{L}^{\frac{n}{n-1}}(\Omega, \mathbb{R}^{n\times n})$. *Then the Jacobian determinant of* f *lies in the Hardy space* $\mathscr{H}^1(\Omega)$. *Furthermore,*

$$\big\| \det Df \,\big\|_{\mathscr{H}^1(\Omega)} \preccurlyeq \int_\Omega |\, D^\sharp f(x)\,|^{\frac{n}{n-1}}\,\mathrm{d}x \tag{5}$$

Proof. We sketch the proof with emphasis on the spherical maximal function that comes into play. Let us commence with an isoperimetric type inequality

$$\Big| \fint_{B(x,r)} F(y)\,\mathrm{d}y \,\Big| \preccurlyeq \Big(\fint_{S(x,r)} \big|D^\sharp f(y)\big|\mathrm{d}y \Big)^{\frac{n}{n-1}} \,, \quad F(y) = J(y, f) \tag{6}$$

Thus, in particular, $|\,F\,| \leqslant |D^\sharp f\,|^{\frac{n}{n-1}} \in \mathscr{L}^1(\Omega)$. A skillful reader with patience can find it in Federer's Book, see also [19] for a legible proof.

Trivially, we have a pointwise inequality $\mathbf{M}F \leqslant \mathbf{M}\big(|D^\sharp f|^{\frac{n}{n-1}}\big)$, but it does not yield the desired $\mathscr{L}^1$-integrability of $\mathbf{M}F$. Combining (4) and (6) yields a better inequality $\mathcal{M}F \preccurlyeq [\mathbf{S}(D^\sharp f)]^{\frac{n}{n-1}}$. Unfortunately, we still find ourselves in a borderline of the $\mathscr{L}^p$-theory of the spherical maximal operator [1, 21]. Neither Hardy–Littlewood nor Fefferman–Stein would ensure us that $\mathcal{M}F \in \mathscr{L}^1(\Omega)$. But together they actually come to the rescue. To this effect we introduce a one-parameter family $\{\mathfrak{M}_s\}_{1\leqslant s\leqslant\infty}$ of maximal operators,

$$\mathfrak{M}_s F(x) = \sup_{0<t<\operatorname{dist}(x,\partial\Omega)} \Big[\frac{n}{t^n} \int_0^t r^{n-1} \Big(⨍_{S(x,r)} |F(y)| \, \mathrm{d}y \Big)^s \mathrm{d}r \Big]^{\frac{1}{s}}. \tag{7}$$

The observant reader may wish to note that this family interpolates between $\mathbf{M} = \mathfrak{M}_1$ and $\mathbf{S} = \mathfrak{M}_\infty$.

Theorem 3 *The sublinear operator* $\mathfrak{M}_s : \mathscr{L}^p(\Omega) \to \mathscr{L}^p(\Omega)$ *is bounded for all exponents* $p > \frac{n}{n-1+\frac{1}{s}}$ *; thus for* $p = \frac{n}{n-1}$ *when* $1 \leq s < \infty$.

Having this result in mind we now complete the proof of (5), by using another straightforward consequence of (4) and (6),

$$\big| F * \Phi_t \big|(x) \preccurlyeq \Big(\mathfrak{M}_s |D^\sharp f| \Big)^{\frac{n}{n-1}}, \quad \text{where, incidentally or not, } s = \frac{n}{n-1}.$$

Thus $\mathcal{M}F \preccurlyeq \big(\mathfrak{M}_s |D^\sharp f| \big)^{\frac{n}{n-1}}$. In conclusion,

$$\| F \|_{\mathscr{H}^1(\Omega)} = \| \mathcal{M}F \|_{\mathscr{L}^1(\Omega)} \preccurlyeq \int_\Omega |D^\sharp f|^{\frac{n}{n-1}}.$$

■

Exploring the BMO–$\mathscr{H}^1$ *duality.* Two bonus results can readily be deduced from Theorem 2. First, since $\mathrm{BMO}(\mathbb{R}^n)$ is the dual space to $\mathscr{H}^1(\mathbb{R}^n)$ (see [4]), we obtain for every $\varphi \in \mathrm{BMO}(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} \varphi(x) \, J(x,f) \, \mathrm{d}x \preccurlyeq \|\varphi\|_{\mathrm{BMO}} \int_{\mathbb{R}^n} \big| D^\sharp f(x) \big|^{\frac{n}{n-1}} \mathrm{d}x, \tag{8}$$

provided $f \in \mathscr{W}^{1,n-1}(\mathbb{R}^n, \mathbb{R}^n)$ and $|D^\sharp f| \in \mathscr{L}^{\frac{n}{n-1}}(\mathbb{R}^n)$.

Second, since $\mathscr{H}^1(\mathbb{R}^n)$ is the dual of $\mathrm{VMO}(\mathbb{R}^n)$, we infer compactness of the Jacobian determinants in the weak star topology of $\mathscr{H}^1(\mathbb{R}^n)$,

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi(x) \, J(x,f_k) \, \mathrm{d}x = \int_{\mathbb{R}^n} \varphi(x) \, J(x,f) \, \mathrm{d}x, \qquad \varphi \in \mathrm{VMO}(\mathbb{R}^n),$$

whenever $f_k \rightharpoonup f$ weakly in $\mathscr{W}^{1,n-1}(\mathbb{R}^n,\mathbb{R}^n)$ and $|D^\sharp f_k|$ stay bounded in $\mathscr{L}^{\frac{n}{n-1}}(\mathbb{R}^n)$. Be cautious, this fails for $\varphi(x) = \log|x|$.

Other implications. Assuming that $f \in \mathscr{W}^{1,n}(\mathbb{R}^n, \mathbb{R}^n)$ we see that Theorem 2 covers the popular result of [2] on $\mathscr{H}^1$-regularity of the Jacobians, because $|D^\sharp f|^{\frac{n}{n-1}} \leqslant |Df|^n \in \mathscr{L}^1(\mathbb{R}^n)$. It also covers a very useful result by S. Müller [18] on local $\mathscr{L}\log\mathscr{L}$-integrability of a nonnegative Jacobian. In fact we have, for every pair of concentric balls $B \subset 2B \subset \mathbb{R}^n$,

$$\int_B F \cdot \log\Big(e + \frac{F}{F_B}\Big) \preccurlyeq \|F\|_{\mathscr{H}^1(2B)} \preccurlyeq \int_{2B} |Df|^n\,, \quad F = J(x,f) \geqslant 0\,. \quad (9)$$

Out of curiosity, the left hand side represents a norm in the Zygmund space $\mathscr{L}\log\mathscr{L}(B)$; the triangle inequality holds. Many further inequalities are to be found in [12]. One of the central problems in GFT is to determine minimal regularity of a Sobolev map under which the nonnegative Jacobian is locally integrable. In fact [7, 14], we have somewhat dual estimates below the natural domain of definition of the Jacobian function,

$$\int_B J(x,f)\mathrm{d}x \preccurlyeq \int_{2B} |D^\sharp f|^{\frac{n}{n-1}} \log^{-1}\Big(\mathrm{e} + \frac{|D^\sharp f|}{|D^\sharp f|_{2B}}\Big) \preccurlyeq \int_{2B} \frac{|Df(x)|^n\,\mathrm{d}x}{\log\big(\mathrm{e} + \frac{|Df(x)|}{|Df|_{2B}}\big)}$$

The true value of these estimates goes beyond theoretical interest; they play a significant role in establishing the existence of energy-minimal deformations in the theory of n-dimensional hyperelasticity.

Sketch II: The p-harmonic transform, a play with the sharp maximal operator

We consider the nonhomogeneous p-harmonic equation, a prototype of many nonlinear PDEs,

$$\operatorname{div}\big(|\nabla u|^{p-2}\nabla u\big) = \operatorname{div}|\mathfrak{f}|^{p-2}\mathfrak{f}\,, \quad 1 < p < \infty \quad (10)$$

The operator that carries a given vector field $\mathfrak{f} \in \mathscr{L}^p(\mathbb{R}^n, \mathbb{R}^n)$ into the gradient of the (unique) solution $u \in \mathscr{W}^{1,p}(\mathbb{R}^n)$ will be called p-*Harmonic Transform*, denoted by

$$\mathbf{R}_p : \mathscr{L}^p(\mathbb{R}^n, \mathbb{R}^n) \to \mathscr{L}^p(\mathbb{R}^n, \mathbb{R}^n)\,, \qquad \mathbf{R}_p\,\mathfrak{f} \overset{\text{def}}{=\!=} \nabla u$$

The linear operator $\mathbf{R}_2 = -\big[R_{ij}\big]_{1\leqslant i,j\leqslant n} : \mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n) \to \mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n)$ is a matrix of second order Riesz transforms, $R_{ij} = R_i \circ R_j$. This is a device for the $\mathscr{L}^2$-projection of a vector field onto the gradient and divergence-free components, known as *Hodge decomposition*:

$$\mathfrak{f} = \nabla\varphi + \mathfrak{h} = \mathbf{R}_2\mathfrak{f} + \mathbf{T}\mathfrak{f}\,, \quad \mathbf{T} = \mathbf{I} - \mathbf{R}_2 : \mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n) \to \mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n)$$

Note that $\mathbf{T}$ vanishes on gradient fields. The idea in the sequel is to use φ as a test function for a divergence type nonlinear differential expressions.

While the Sobolev space $\mathscr{W}^{1,p}(\mathbb{R}^n)$, together with the given vector field $\mathfrak{f} \in \mathscr{L}^p(\mathbb{R}^n, \mathbb{R}^n)$, is considered the natural setting for the equation (10), we shall depart from it and move into the realm of exponents $s \geqslant p$. We shall show that the p-harmonic transform is bounded in $\mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n)$ in the sense of the following

Theorem 4 *If $\mathfrak{f}$ belongs to $\mathscr{L}^p(\mathbb{R}^n, \mathbb{R}^n) \cap \mathscr{L}^s(\mathbb{R}^n, \mathbb{R}^n)$ then so does $\boldsymbol{R}_p \mathfrak{f}$. Moreover, we have the uniform bound,*

$$\| \boldsymbol{R}_p \mathfrak{f} \|_s \preccurlyeq \| \mathfrak{f} \|_s , \qquad s \geqslant p . \tag{11}$$

Proof. We only sketch the proof of the uniform estimate (11), and stick to the case $p \geqslant 2$ for simplicity. A laborious proof of $\mathscr{L}^s$-integrability of $\mathbf{R}_p\mathfrak{f}$, based on Gehring's Lemma on reverse Hölder inequalities [6], can be found in [9]. Choose and fix a ball $B \subset \mathbb{R}^n$. The weak form of equation (10) reads as

$$\int_B \langle\, |\nabla u|^{p-2} \nabla u \;\big|\; \nabla \varphi \,\rangle \;=\; \int_B \langle\, |\mathfrak{f}|^{p-2} \mathfrak{f} \;\big|\; \nabla \varphi \,\rangle \,, \quad \text{for} \quad \varphi \in \mathscr{W}^{1,p}_{\circ}(B) . \tag{12}$$

Let $v \in u + \mathscr{W}^{1,p}_{\circ}(B)$ be a (unique) function that agrees with u on ∂B and has smallest p-harmonic energy $\mathscr{E}_p[v] = \int_B |\nabla v|^p$. Thus v is a p-harmonic function, meaning that

$$\int_B \langle\, |\nabla v|^{p-2} \nabla v \;\big|\; \nabla \varphi \,\rangle \;= 0 \quad \text{and} \quad ⨍_B |\nabla v|^p \leqslant ⨍_B |\nabla u|^p . \tag{13}$$

We subtract this integral from the left hand side of (12) and test the resulting equation with $\varphi = u - v$. From this it is straightforward to derive basic local estimates

$$⨍_B |\nabla u - \nabla v|^p \;\preccurlyeq\; ⨍_B |\,\mathfrak{f}|^p . \tag{14}$$

We aim to replace ∇v by a constant. For this we recall the $\mathscr{C}^{1,\alpha}$-regularity of p-harmonic functions, $0 < \alpha = \alpha(n,p) \leqslant 1$. Precisely, for every $0 < \tau \leqslant 1$,

$$⨍_{\tau B} \big|\, \nabla v \,-\, (\nabla v)_{\tau B} \,\big|^p \;\preccurlyeq\; \tau^{\alpha p} ⨍_B |\nabla v|^p \;\leqslant\; \tau^{\alpha p} ⨍_B |\nabla u|^p , \tag{15}$$

where the implied constants in the inequalities $\preccurlyeq$ depend only on n and p. On the other hand, (14) yields

$$⨍_{\tau B} \big|\, \nabla u \,-\, \nabla v \,\big|^p \;\preccurlyeq\; \tau^{-n} ⨍_B |\,\mathfrak{f}\,|^p , \tag{16}$$

whence it is readily inferred that

$$\fint_{\tau B} \left| \nabla u - (\nabla u)_{\tau B} \right|^p \preccurlyeq \tau^{-n} \fint_B |\mathfrak{f}|^p + \tau^{\alpha p} \fint_B |\nabla u|^p . \tag{17}$$

From now on there are two ways to obtain the estimate (11), both via a pointwise inequality between maximal functions. In the first approach, we apply Hölder's inequality to the left hand side of (17) and then take supremum over all balls centered at a given point,

$$\left| \mathbf{M}^\sharp \nabla u \right|^p \preccurlyeq \tau^{-n} \mathbf{M} |\mathfrak{f}|^p + \tau^{\alpha p} \mathbf{M} |\nabla u|^p , \quad \text{pointwise in } \mathbb{R}^n .$$

We raise this to the power $\frac{s}{p} > 1$ and, with the aid of maximal inequalities, obtain

$$\| \nabla u \|^p_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq \tau^{-n} \| \mathfrak{f} \|^p_{\mathscr{L}^s(\mathbb{R}^n)} + \tau^{\alpha p} \| \nabla u \|^p_{\mathscr{L}^s(\mathbb{R}^n)} ,$$

where the implied constant depends on n, p and s, but not on the parameter τ. The observant reader may be concerned that this constant blows up as s approaches p. But still we can chose τ small enough so that the last term in the right hand side will be absorbed by the left hand side, establishing the desired estimate (11).

In the second approach, to avoid the undue anomaly near the natural Sobolev space $\mathscr{W}^{1,p}(\mathbb{R}^n)$, we relinquish the idea of using the sharp maximal inequality (2). Instead, we appeal to the full force of Hardy–Littlewood maximal inequalities near $\mathscr{L}^1(\mathbb{R}^n)$. For this purpose, we rewrite (17) as

$$\fint_{\tau B} | \nabla u |^p \preccurlyeq \tau^{-n} \fint_B |\mathfrak{f}|^p + \left| (\nabla u)_{\tau B} \right|^p + \tau^{\alpha p} \fint_B |\nabla u|^p , \tag{18}$$

Taking supremum over the balls centered at a given point we capture a pointwise inequality for maximal functions

$$\mathbf{M} |\nabla u|^p \preccurlyeq \tau^{-n} \mathbf{M} |\mathfrak{f}|^p + |\mathbf{M} \nabla u|^p + \tau^{\alpha p} \mathbf{M} |\nabla u|^p . \tag{19}$$

We now eliminate the operator $\mathbf{M}$ by computing the $\mathscr{L}^q(\mathbb{R}^n)$-norm, $q = \frac{s}{p} \approx 1$, of both sides. Maximal inequalities (1) yield

$$\| \nabla u \|^p_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq \tau^{-n} \| \mathfrak{f} \|^p_{\mathscr{L}^s(\mathbb{R}^n)} + (s-p) \| \nabla u \|^p_{\mathscr{L}^s(\mathbb{R}^n)} + \tau^{\alpha p} \| \nabla u \|^p_{\mathscr{L}^s(\mathbb{R}^n)} .$$

This time we are not troubled with the exponent s approaching p; the implied constant remains bounded. Choose $s = s(n,p) > p$ close enough to p and τ sufficiently small so that the last two terms will be absorbed by the left hand side. We obtain uniform bounds (11) for $\mathbf{R}_p$ near its natural domain of definition. ∎

Remark. The unplanned bonus coming from the asymptotically precise Hardy–Littlewood maximal inequalities (1) gives the above proof its beauty, doesn't it?

Sketch III: The splendor of commutators

In any preliminary analysis of a differential equation, linear or nonlinear, one often encounters undesirable higher order terms which eventually cancel out. The instruments for rigorous performance of such an analysis are the commutators of a singular integral $\mathbf{T} : \mathscr{L}^s(\mathbb{R}^n) \to \mathscr{L}^s(\mathbb{R}^n)$, $1 < s < \infty$, with suitable nonlinear algebraic operations on the gradient of the solution. Let us look briefly at three commutators, together with their underlying estimates:

- The linear commutator of Coifman-Rochberg-Weiss [3]

$$\| \mathbf{T}(\lambda f) - \lambda(\mathbf{T}f) \|_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq \| \lambda \|_{\mathrm{BMO}(\mathbb{R}^n)} \| f \|_{\mathscr{L}^s(\mathbb{R}^n)} .$$

- The Rochberg-Weiss commutator [20]

$$\| \mathbf{T}(f \log|f|) - (\mathbf{T}f) \log |\mathbf{T}f| \|_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq \| f \|_{\mathscr{L}^s(\mathbb{R}^n)} .$$

- The commutator of $\mathbf{T}$ and a power type operation [15]

$$\| \mathbf{T}(|f|^{\pm\varepsilon} f) - |\mathbf{T}f|^{\pm\varepsilon}(\mathbf{T}f) \|_{\mathscr{L}^s(\mathbb{R}^n)} \preccurlyeq |\varepsilon| \cdot \| \, |f|^{1\pm\varepsilon} \|_{\mathscr{L}^s(\mathbb{R}^n)} , \; 0 \leqslant \varepsilon < 1 - \frac{1}{s}.$$

The proof of this latter estimate captures the ideas of the complex method of interpolation originated in the celebrated work by G. O. Thorin [22]. Actually, it yields the estimate of the Rochberg–Weiss commutator through the L'Hôpital's rule, which in turn gives us Müller's $\mathscr{L} \log \mathscr{L}$-integrability of nonnegative Jacobians in a stylish way [12].

Although the linear commutator of Coifman–Rochberg–Weiss has been known for a long time, and numerous deep studies have been devoted to it, its usefulness in solving PDEs still remains magical. For example, good estimates of the p-norms of the tensor products of the Riesz transforms combined with the Fredholm index theory (via compactness of the Coifman–Rochberg–Weiss commutators) are elegant tools in elliptic PDEs with VMO coefficients [16]; there is no need to go again and again through the foundational details of singular integrals.

I have saved the best for last:

Very weak solutions of nonlinear PDEs. These are the solutions weaker than those in the natural domain of definition. The chief difficulty is to launch some estimates in order to take the very weak solution off the ground. Let us return to the weak formulation of the p-harmonic equation (12) in which $\mathfrak{f} \in \mathscr{L}^{p-\varepsilon}(\mathbb{R}^n, \mathbb{R}^n)$, so we must look, naturally, for the solution $u \in \mathscr{W}^{1,p-\varepsilon}(\mathbb{R}^n, \mathbb{R}^n)$. The legitimate test function φ must lay in $\mathscr{W}^{1,\frac{p-\varepsilon}{1-\varepsilon}}(\mathbb{R}^n)$. We steal it from the Hodge decomposition

$$|\nabla u|^{-\varepsilon}\nabla u = \nabla\varphi + \mathfrak{h}\,, \quad \text{where } \mathfrak{h} = \mathbf{T}(|\nabla u|^{-\varepsilon}\nabla u)$$

to obtain

$$\int_{\mathbb{R}^n} |\nabla u|^{p-\varepsilon} \preccurlyeq \int_{\mathbb{R}^n} |\mathfrak{f}|^{p-\varepsilon} + \int_{\mathbb{R}^n} |\mathfrak{h}|^{\frac{p-\varepsilon}{1-\varepsilon}}\,. \tag{20}$$

Since the operator $\mathbf{T}$ vanishes on gradient fields we can write $\mathfrak{h}$ as a commutator of $\mathbf{T}$ and the power function

$$\mathfrak{h} = \mathbf{T}(|\nabla u|^{-\varepsilon}\nabla u) - |\mathbf{T}\nabla u|^{-\varepsilon}(\mathbf{T}\nabla u)\,.$$

Our estimate for the power type commutator shows that

$$\int_{\mathbb{R}^n} |\mathfrak{h}|^{\frac{p-\varepsilon}{1-\varepsilon}} \preccurlyeq \varepsilon \cdot \int_{\mathbb{R}^n} |\nabla u|^{p-\varepsilon}\,.$$

Consequently, this term can be sucked up by the left hand side, which results in the desired estimate of the p-harmonic transform slightly below its natural domain of definition:

$$\big\|\mathbf{R}_p\,\mathfrak{f}\big\|_{\mathscr{L}^{p-\varepsilon}(\mathbb{R}^n)} \preccurlyeq \big\|\mathfrak{f}\big\|_{\mathscr{L}^{p-\varepsilon}(\mathbb{R}^n)}\,. \tag{21}$$

A study of very weak solutions of nonlinear PDEs is largely motivated by removability of singularities [11].

3. Moral of the story

If one dark rainy night you find yourself in the midst of Whitney cubes, covering lemmas, Calderón–Zygmund decomposition, etc., then you should remind yourself that instead you might cleverly apply singular integral stuff and thereby see the light.

> *“Every block of stone has a statue inside it and it is the task of the sculptor to discover it.”*

Michelangelo di Lodovico Buonarroti Simoni

> *“No profit grows where is no pleasure taken; in brief, sir, study what you most affect”.*

William Shakespeare[3]

[3] An English teacher assigned a student to read some Shakespeare, and a week later he asked: “How did you like it?” The student answered: “Well, nothing special; just a collection of quotations”.

However,

> *In order for something to remain beautiful, it must stay "long enough to be noticed and enjoyed, never so long as to outstay its welcome."*

"The Rainbow and Cartesian Wonder"
PHILIP FISHER

Finally, as Thomas Edison said,

> *"No sooner does a fellow succeed in making a good thing, than some other fellows pop up and tell you they did it years ago".*[4]

References

[1] BOURGAIN, J.: Estimations de certaines fonctions maximales. *C. R. Acad. Sci. Paris St. I Math.* **301** (1985), no. 10, 499–502.

[2] COIFMAN, R., LIONS, P.-L., MEYER, Y. AND SEMMES, S.: Compensated Compactness and Hardy Spaces. *J. Math. Pures Appl.* **72** (1993), no. 9, 247–286. *Arch. Ration. Mech. Anal.* **195** (2010), no. 3, 899–921.

[3] COIFMAN, R., ROCHBERG, R. AND WEISS, G.: Factorization theorems for Hardy spaces in several variables. *Ann. of Math. (2)* **103** (1976), 611–635.

[4] FEFFERMAN, C.: Characterizations of bounded mean oscillation. *Bull. Amer. Math. Soc.* **77** (1971), 587–588.

[5] FEFFERMAN, C. AND STEIN, E. M.: H^p spaces of several variables. *Acta Math.* **129** (1972), 137-193.

[6] GEHRING, F. W.: The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* **130** (1973), 265-277.

[7] GIANNETTI, F., IWANIEC, T., ONNINEN, J. AND VERDE, A.: Estimates of Jacobians by Subdeterminants. *J. Geometric Analysis* **12** (2002), no. 2, 223–254.

[8] HARDY, G. H. AND LITTLEWOOD, J. E.: A maximal theorem with function-theoretic applications. *Acta Math.* **54** (1930) 81–116; also in *Collected Papers of G. H. Hardy*, vol. 2, 509–545. Oxford Univ. Press, 1967.

[9] IWANIEC, T.: Projections onto gradient fields and $\mathscr{L}^p$-estimates for degenerate elliptic operators. *Studia Mathematica* **75** (1983), 293–312.

[4] Once a scholar loudly informed the speaker of a plenary lecture "Sir, your result follows from my theorem, I proved it years ago". The speaker answered: "Yes it does follow, obviously. However, I am not sure if your theorem follows from my result".

[10] IWANIEC, T.: *An Essay on the Interpolation Theorem of Józef Marcinkiewicz —Polish Patriot.* arXiv:1102.0167v1.

[11] IWANIEC, T.: p-harmonic tensors and quasiregular mappings. *Ann. of Math. (2)* **136** (1992), no. 3, 589–624.

[12] IWANIEC, T. AND MARTIN, G.: *Geometric function theory and non-linear analysis.* Oxford University Press, New York, 2001.

[13] IWANIEC, T. AND ONNINEN, J.: $\mathscr{H}^1$-estimates of Jacobians by subdeterminants. *Mathematische Annalen* **324** (2002), 341–358.

[14] IWANIEC, T. AND SBORDONE, C.: On the inegrability of the Jacobian under minimal hypothesis. *Arch. Rat. Mech. Anal.* **119** (1992), 129–143.

[15] IWANIEC, T. AND SBORDONE, C.: Weak minima of variational integrals. *J. Reine Angew. Math.* **454** (1994), 143–161.

[16] IWANIEC, T. AND SBORDONE, C.. Riesz transforms and elliptic PDEs with VMO coefficients. *J. Anal. Math.* **74** (1998), 183–212.

[17] MARCINKIEWICZ, J.: Sur l'interpolation d'opérations *C. R. des Séances de l'Académie des Sciences. Série A e B* **208** (1939), 1272–1273.

[18] MÜLLER, S.: Higher integrability of determinants and weak convergence in $\mathscr{L}^1$. *J. Reine Angew. Math.* **412** (1990), 20–34.

[19] MÜLLER, S., QI, T. AND YAN, B. S.: On a new class of elastic deformations not allowing for cavitation. *Ann. Inst. H. Poincaré Anal. Non Lineaire* **11** (1994), no. 2, 217–243.

[20] ROCHBERG, R. AND WEISS, G.: Derivatives of analytic families of Banach spaces. *Ann. of Math.* **118** (1983), 315–347.

[21] STEIN, E. M.: Maximal functions: spherical means.*Proc. Nat. Acad. Sci. USA* **73** (1976), 2174-2175.

[22] THORIN, G. O.: An extension of a convexity theorem due to M. Riesz. *Kungl. Fysiogr. Sällsk. i Lund Förh.* **8** (1938), 166–170.

Joseph J. Kohn (Prague, Czech Republic, 1932) obtained his PhD in mathematics in 1956 at Princeton University (adviser: Donald C. Spencer). He is Professor Emeritus of Mathematics at Princeton University. His work focuses, among other things, on the use of partial differential operators in the theory of functions of several complex variables and microlocal analysis.

Mathematical encounters

Joseph J. Kohn*

This article, in celebration of the 100th anniversary of the Real Sociedad Matemática Española, deals with my first encounters with mathematics and mathematicians. I became seriously interested in mathematics in high school inspired by dedicated teachers and competitive fellow students. I began my college education in MIT in 1950 and ever since I have had the privilege of meeting many remarkable mathematicians. Many were in the forefront of their profession and I was fortunate in learning a great deal from them. They were diverse individuals. However, as a group they had in common intense dedication to and enthusiasm for mathematics. Some of them had reputations as eccentrics and many amusing ("se non son vere son ben trovate") anecdotes circulated about them in the mathematical community. Here I will recall some of these individuals, their influence on me and some of the amusing incidents which I witnessed.

In 1950 the most famous person at MIT was **Norbert Wiener**. He was enormously erudite and very knowledgeable in vast areas of mathematics, science, and engineering. He was also deeply interested in the impact of science on society. He predicted the tremendous technological revolution that would be brought by the development of computers.

During my last year at MIT I took Wiener's course on Fourier Series and Integrals. In many ways it was inspiring; it taught me a lot about analysis. At the same time the course was disorganized. For example, one lecture was a highly technical and advanced exposition of his Tauberian theorem. In the next lecture he proclaimed that he will now apply this to number theory. "And what is number theory?", he asked. "Essentially number theory is the study of prime numbers". He then proceeded to give a lengthy definition of a prime number and went on: "two is a prime number, let us now consider the number three and determine whether or not it is prime". He continued in this vein systematically until he passed one hundred, then he said: "Now that we have a feeling for prime numbers let us consider the infinite sum

$$\sum \frac{1}{p}$$

summed over all primes." He wrote this on the blackboard, then he stepped back looked at the sum and raising his voice dramatically said: "this series clearly diverges because of its arithmetical character." He followed this up

*kohn@princeton.edu.

by a speedy introduction to the Riemann zeta function. I was puzzled; I couldn't imagine what he meant by the "arithmetical character"; now, over half a century later, I am still puzzled. One day Wiener started his class saying that this time he will not talk about mathematics, instead, he will tell us about his latest detective story. In a confidential tone he told us that he writes detective stories under the name "W. Norbert". The story was a murder mystery in which all clues were mathematical, such as the notation for a line integral, convolution, time derivative, etc. The detective had to be well versed in the mathematical literature to decipher this and then, by elimination, he could determine the nationality and mathematical specialty of the murderer. This analysis led to the identification, capture, trial, and conviction of the culprit.


Norbert Wiener

I first met Wiener when I was a freshman. A number of us were playing a game that involved drawing circles and lines on a blackboard. Suddenly Wiener appeared. It was an impressive sight. He was heavy, smoked a cigar and wore extremely thick glasses. It seemed that he would not be able to see anything through those glasses and cigar smoke but he gave the impression that he saw everything that mattered in the world without paying much attention to his immediate surroundings. He looked around for a moment and then in a booming voice he asked: "who here is working in binary arithmetic?" A few days later I was working on a chess problem when I suddenly was startled to hear that booming voice again: "young man, do you play the game of chess?" I was a mediocre player and intimidated by the challenge of this famous person. "I am not a good player", I answered in a trembling voice. "No matter, let us have a game", he shouted. We played and right from the beginning I was puzzled, Wiener was making some very strange moves. I assumed that he was playing at some deep level which was beyond my understanding. Suddenly, in the initial stages of the game, Wiener moved his queen so that I could take it with my knight. This was incomprehensible to me, I could not fanthom why he would sacrifice his queen. I felt like a fool, I started to sweat, I didn't know what to do. Wiener was enjoying my distress and chuckled. I concentrated very hard and tried to anticipate what possible advantage could Wiener have if I captured the queen, I could not figure it out. Finally I said: "Professor Wiener, I am sorry this is way above my head, I do not understand your queen sacrifice”. Wiener looked at the board and seemed bewildered: “Oops, I will have to take that move back”. Then the game deteriorated Wiener was making one bad move after another. Soon he decided to end the game, it was not a good day and he had to go back to his office, he said.

Several other Wiener stories come to mind but to tell them would take up a disproportionate part of this account.

Perhaps the best undergraduate course, the course in which I learned the most, was the junior full year course in real analysis. The teacher was **John F. Nash**. He was brilliant, arrogant, and eccentric. At this time he was in the midst of his spectacular work on embedding theorems, nevertheless, his course was meticulously prepared and beautifully presented. The course started with an introduction to mathematical logic and set theory and covered, with great originality, the central topics of analysis culminating in the study of differential and integral equations. Nash is now known by the public at large because of the movie “A Beautiful Mind” which is loosely based on the book with the same name by Sylvia Nasar (Simon & Shuster, 1998). The book recounts many incidents in Nash’s life including some of the interactions that he had with me.

The most carefully organized course I took at MIT was in differential geometry given by **Warren Ambrose**. Ambrose was a dynamic lecturer and he followed the Bourbaki idea of keeping things invariant and very general. It is in this course that I first learned about the Hodge theorem and became fascinated by the applications of the theory of partial differential equations to the study of manifolds. In particular it led me to study the result known as the Weyl’s lemma (Weyl, H., The method of orthogonal projection in potential theory, *Duke Math. J.* **7** (1940), 414–444). For me this result opened a new and exciting vision of mathematics: the intimate connections between partial differential equations and several complex variables. This vision was reinforced when, as a graduate student, I went to a series of lectures by **Hermann Weyl** at the Institute for Advanced Study on his work on Riemann surfaces.

Returning to Ambrose, he was a very colorful character with strong opinions on mathematics, religion, and politics. In the case of mathematics he had been an expert in functional analysis but later turned against this subject and talked about it with disdain, using expressions such as “trivial”, “irrelevant”, and “nonsense” to describe it. At this time the Unitarian Church in Boston presented a weekly series of talks to promote religious tolerance and understanding. The talks had titles such as: “Why I am a Catholic”, “Why I am a Protestant”, “Why I am a Jew”, etc. Ambrose gave a talk in this series entitled “Why I am an atheist”. Naturally many students in his class came to the talk and the church was packed. The talk was not well received, a low point was reached when Ambrose elaborated on his reasons for believing that all religions impede human progress. His thesis was that religion inhibits scientific experimentation, especially in the social sciences. For instance, he argued, all religions would object to the type of research that Kinsey was doing.

In June of 1966 I went to give a series of lectures in Buenos Aires. Just then there had been a military coup but the universities had not yet been taken over. The atmosphere was tense, the military was not allowed to enter the campus and outside there were confrontations between the police and the students. Everyone expected that the military would take over the university and one day it did. Many students and faculty members were terribly upset and planning how to react. The university had a massive building in downtown Buenos Aires and a group of students decided to "liberate" it. They scheduled a protest meeting in that building for one night. That evening I had dinner with Ambrose and Alberto González Domínguez, who was an important figure in mathematics and in the academic community. Ambrose said that he plans to join the students in the protest meeting. González Domínguez and I tried to talk him out of it. We argued that he does not know Spanish, that he doesn't understand the political intricacies and that this could be dangerous. We could not convince him, he went and we did not. As soon as the meeting hall was full the organizers locked the massive doors of the building and put up signs proclaiming that the building was liberated. Soon the police came and demanded, with loudspeakers, that the building be vacated. This building has very thick walls and the police demands were not heard inside. The police broke in and the audience was lined up to go out of the building through the exit where they were beaten, one by one, with rifle butts. Late that night I got a phone call that Ambrose had disappeared. I immediately called the American Embassy. The next morning Ambrose came to the hotel. He told me that this was a fantastic experience; he was thrilled to have been part in it. When he was in line waiting his turn to be beaten he realized that there has to be a rhythm to beat so many people as they passed. Ambrose had been a jazz musician so he was trained to move rhythmically. Now he put the training to good use so that by going along with the rhythm of the beating the blows he received were only glancing. He said that he was happy not to have missed such a great adventure.

After the universities in Argentina were taken over by the military many mathematicians fled to other countries. Ambrose was extremely helpful in settling a number of them in the United States.

In 1953 I started graduate work in Princeton. The chairman of the Mathematics Department was the renowned mathematician **Solomon Lefschetz**. He had been one of the foremost developers of modern topology and algebraic geometry. After the war most of his work was in ordinary differential equations. He was a very effective chairman, instrumental in building up the department to its preeminent status. He could be very blunt. He gave the entering class the following words of encouragement: "Congratulations, you have been admitted to Princeton after a very stiff competition. There are thirteen of you and you have a right to be proud of this accomplishment. I have no doubt that if you work at the same level that enabled you to get in here you will be able to satisfactorily complete the requirements for a doctorate. Soon most of you will get their PhD in mathematics. But how many of you will become real mathematicians? Maybe one, perhaps even two, but certainly not more than three." Two other instances of the Lefschetz style are the following. 1. One day Kodaira and Spencer told Lefschetz that they finally were able to complete the proof of one of his major theorems. His response was: "This is Princeton; we do not do baby mathematics here." 2. When Zariski asked Lefschetz for advice on how to categorize an algebraic geometry paper submitted to the Annals of Mathematics – should it be classified as algebra or geometry. Lefschetz answered: "If the paper consists of a lot of symbol manipulation, then it is algebra. But if there is an idea, then it is geometry." To be fair, more often than not Lefschetz was very generous and supportive. When Steenrod died, at the age of 62, Lefschetz was distraught; he said that the world was not just and that he should have died before Steenrod.

In my time, graduate students did not have offices; we were assigned carrels in the library. Lefschetz had a huge office, he liked to have people around so he invited a few graduate students to use it; there were several spare desks and an elegant blackboard. The rule was that whenever he had business there we had to vacate and whenever he held seminars or mathematical discussions we were welcome to stay. One day I was studying a beautifully written book by **Norman Steenrod**: "The Topology of Fibre Bundles" (Princeton University Press). When Lefschetz saw this he said that this is a great book and that there is a lot I could learn from it. However, he said, it does not confront the main issue. The book is based on the assumption of constant rank and the really interesting and relevant problems arise when the rank degenerates. This view was reenforced by **René Thom**. On a visit to Princeton, Thom had lunch with Spencer and me. He told us that the trouble with current mathematics is that it is dedicated mainly to describing phenomena that are smooth and regular whereas the exciting phenomena happen when such regularity breaks down in the spirit of his catastrophe theory. The remarks by Lefschetz and Thom had a profound effect on my work. I was studying the $\bar{\partial}$-Neumann problem on strongly pseudoconvex domains. This is a boundary value problem in which the ellipticity breaks down. Nevertheless, the break down is controlled by the fact that the Levi form has constant rank. A much more radical departure from ellipticity is when the domains are weakly pseudoconvex. For such problems in general, the solutions cannot be given by any explicit formulas and have to be investigated by means of estimates. The study of these estimates often leads to problems in algebraic geometry of the sort that Lefschetz referred to.

Lefschetz was devoted to the development of mathematics in México. He turned his organizational talent to building up Mexican mathematics. He spent a lot of time lecturing there and brought a series of Mexican students to do graduate work in Princeton. His efforts were spectacularly successful, Mexican mathematics is flourishing in large part because of the impulse from Lefschetz.

When Lefschetz found out that I speak Spanish he urged me to come to México. Before I went, the well known Mexican topologist, José Adem (who was one of the mathematicians that Lefschetz brought to study in Princeton), asked me to give some lectures on sheaves since I had been following a seminar on the subject in which the latest developments were exposed. Just before my first lecture Lefschetz told me that whatever I do I should not lecture on sheaves, he claimed that the audience was not prepared for it. I stood in front of the audience not knowing what to do. Fortunately, when the bell rang Adem walked in. I took him to a side and explained my predicament. He told me not to worry, that I should start with a few general comments and that within a few minutes Lefschetz would be sound asleep after which I could start exposing sheaves. I followed his advise. It was right on target.

Spencer (Bombay, 1963)

At Princeton, my thesis advisor was **Donald C. Spencer**. He was enormously enthusiastic about mathematics and also about his students. He always made us feel that our ideas were really important. He was truly inspiring. When he lectured he wanted to get to the important point as quickly as possible and he did not mind if the foundations that he presented were somewhat sketchy. For example, in the opening lecture of the above mentioned seminar on sheaf theory, Spencer got very frustrated with all the conditions in the definition of a sheaf and after trying unsuccessfully to summarize them quickly, he said:

"It is really very simple, a sheaf is just algebra this way (waving his hands vertically) and topology this way (waving his hands horizontally)." The intensity of Spencer's enthusiasm was contagious. He often talked about the great project of deformation of complex structures that he and **Kunihiko Kodaira** were working on. The collaboration with Kodaira was extremely fruitful. It had and continues to have enormous influence in a number of mathematical fields. In the introduction to his remarkable book "Complex Manifolds and Deformations" (Springer Verlag, 1981), Kodaira conveys the nature of their work: "Spencer and I developed the theory of deformation of compact complex manifolds. The process of the development was the most interesting experience in my whole mathematical life. It was similar to an experimental science developed by the interaction between experiments (examination of examples) and theory. In this book I have tried to convey this interesting experience; however I could not fully convey it. Such an experience may be a passing phenomenon which cannot be reproduced."

A great attraction to all those around Spencer was his weekly seminar, called the "Nothing Seminar". It was very informal and very stimulating. A speaker presented some current work and the audience, led by Spencer, interrupted with comments and questions. Spencer would usually take the speaker to dinner to an out of town restaurant. Very often Kodaira and I were in the back seat of the car and the guest in front. Invariably Spencer had an urgent mathematical idea which he had to discuss with Kodaira while driving. He would turn around to face Kodaira during this conversation while driving full speed ahead. This made the guest and me very nervous while Kodaira and Spencer did not seem to feel any danger. Fortunately it always came out well, evidently Spencer's peripheral vision was extraordinary.

Physically, Spencer was very strong, he looked like of a western movie hero. He could also hold his liquor. Spencer told me that the most serious drinking challenge he faced was when Lefschetz invited him for a drink. When he got there Mrs. Lefschetz told Spencer that her husband would be late but that he had instructed her to pour Spencer a drink. She opened a bottle of scotch and filled up a tall glass with it (Mrs. Lefschetz was a midwesterner and not familiar with liquor). Spencer managed to drink it all without hesitation. Once

Spencer in Durango,
after retirement

when I had dinner with him, we started with double martinis. He was on his fourth drink while I was still slowly sipping my first. I said: “Don, I don’t know how you do it, I am still working on the first drink and I am already dizzy.” He answered: “Joe, you don’t know how lucky you are, it takes me at least six to get to that point.” Spencer retired at 65, three years before the then mandatory age of 68. He said that he did not want to grow old pontificating about mathematics, like many of his colleagues. He moved to Durango Colorado and spent most of his time hiking and working on his garden. He soon became known to everyone in town, he had a charismatic personality. One year after his move to Durango his new friends threw a birthday party for him with over 150 guests. In 2008, several years after his death, one of his favorite lookout points near Durango was officially named **Spencer’s Peak** in his honor.

Apart of the above mentioned paper by H. Weyl my early research was greatly influenced by two papers. First, the ground breaking paper by **Hans Lewy**: “An example of a smooth linear partial differential equation without solution” (*Annals of Math.* **66** (1957), no. 1, 155–158). This paper motivated my research in two directions. In several complex variables where it leads to analysis on CR manifolds and in PDE where it has led me to study hypoelliptic operators that lose derivatives.

The other famous paper that had a major influence on me was by **Alberto Calderón** “Uniqueness of he Cauchy problem for partial differential equations” (*Amer. J. Math.* **80** (1958), 16–36). This paper motivates the study of pseudo differential operators and especially their use in reducing estimates in PDE to problems in algebra and algebraic geometry.

Of course I greatly benefited by studying many other seminal papers. The three mentioned here come to mind as being particularly striking and inspiring in my work.

I spent the summer of 1957 at Stanford as assistant to **Stefan Bergman**. He was a very original mathematician. His introduction of the Bergman projection and of the Bergman kernel function became a major tools in the study of conformal mappings (see “The kernel function and conformal mappings” by Stefan Bergman, Amer. Math. Soc., 1950). He then went on to use these methods in the study of functions in several complex variables and these ideas gave rise to many important developments. He was a very generous man dedicated to his students and colleagues. He also was very effective in helping Polish mathematicians during and after the war. At the same time he was quite a character. When asked how many languages he speaks he replied: “I shpeak sieven langvidges und Eanglish is de bestest”. His wife told me that when they first arrived in Stanford, Bergman went to introduce himself to the local butcher. He told him that he was a well known mathematician, one of the greatest experts in the theory of several complex variables and that he had just been appointed to a prestigious professorship at Stanford. The butcher was not impressed and said: "OK, but I bet that I make at least twice as much money as you do."

One day Bergman told me that the Soviet launching of Sputnik awakened a great curiosity about space. He said that now the American public would be eager to learn about the geometry of four dimensional space. Further he said that the best way to understand this geometry is through the theory of two complex variables. He felt that he was the ideal person to explain this and he wanted me to go to San Francisco to set up a series of television appearances for him in which he would lecture on this. I told him that, as a recent Ph.D., I was unknown and that it would be very difficult for me to see the right people. I suggested that, instead of sending me, he should write a letter on official Stanford stationary making the suggestion. As a world renowned scientist he could not be ignored. After thinking about it Bergman agreed to this and asked me to draft such a letter. I did, Bergman signed it and sent it. No reply came and the matter was forgotten.

This account concerns some of my early encounters with mathematicians. The list is by no means complete, among the other inspiring mathematicians that were my teachers are: **Witold Hurewicz, Isidore Singer, William Feller, Emil Artin, and Solomon Bochner**. I could write a great deal about each one of them but the limitations on the length of this note preclude that.

Pierre Gilles Lemarié-Rieusset (Douai, France, 1960) obtained his PhD in mathematics in 1984 at Parix XI-Orsay (advisor: Yves Meyer). He has been Professor of the Université d'Evry since 1995 and his research focuses on real harmonic analysis with applications to functional analysis and non-linear PDEs (mainly Navier–Stokes equations).

Good or not so good papers I have read or have not read

Pierre Gilles Lemarié-Rieusset*

However simple the question raised by the Revista Matemática Iberoamericana may look, I have met some trouble to find an adequate answer. I was asked to write down a couple of pages on "the particular publications that got my attention and affected my own personal research". The aim was to "place in a proper light the role that research journals play in the development of Mathematics".

There are many papers that have had a decisive role in my research. However, when I look more closely at the question, there are very few ones which play that role as publications in scientific journals *per se*. First of all, the influence they have had on my research is due to the scientific content: results and methods, and not the printed medium. The most influential ones were so influential that I did not have to read them: I learned their content through books, conferences, seminars, lectures, or derived papers, and thus I actually did not have to read them, I only had to quote them in the references at the end of my own papers (thus propagating their influence a bit further)...

Another problem was the fact that I most usually did not read them as journal papers. I remember being thrilled as a student as I read an old paper of Lebesgue in his *Oeuvres complètes*, but to know in which journal it was originally released is no longer meaningful. For modern papers, this question is even more meaningless: the most striking ones are widely disseminated as soon as they are posted on the Internet, on `arXiv` for instance, and when they are released on a specific copyrighted journal it is important to know which journal for bibliographic or bibliometric reasons, but the scientific impulse often has been exerted before the official publication. Even when the scientific publisher is important in providing the paper, most often, it is mediated through a data base of hundreds of journals (such as *Science Direct* or *Springerlink*) more than through the production of a specific journal.

Even in the case of direct confrontation with the results (precluding any use of the huge electronic data bases offered on the Internet, or a second-hand presentation of the paper), most of the time, the actual journal is not involved, as the direct confrontation is with the author, not with the journal: either the author presents his results at a seminar or a conference, or he sends

*pierre-gilles.lemarie@univ-evry.fr.

me his preprint, or his work is relayed to me by an enthusiastical co-worker. In all cases, the journal does not seem to play a very important role.

For instance, the celebrated $T(1)$ theorem of David and Journé had a long (everlasting) impact on my research; first of all, the first draft of my Ph.D. in 1983 thrown in the dustbin, then every paper of mine, or almost every one, thereafter used this theorem. However, I never read the paper which was published in 1984 in the *Annals of Mathematics*, since I had the wonderful opportunity to meet Yves Meyer, Guy David and Jean-Lin Journé on a regular basis throughout the years 1982 and 1983.

Another of my favorite papers which I did not read was the block spin construction of ondelettes by Guy Battle, a paper which was published in 1987 in *Communication in Mathematical Physics*. I was drastically concerned with this paper, as Guy Battle constructed his spline wavelet basis in the heart of Texas in just the same week as I constructed my own basis on the sun-bathed shore of Sidi Bou Saïd. Yves Meyer received my Tunisian letter a short time after he received Battle's Texan letter; Battle was kind enough to retitle his paper as "Lemarié's functions", as I was a younger student in a precarious status. The Battle–Lemarié wavelet basis was born. I tried to read Battle's paper, but it sounded so mysterious to me (and to every co-worker of mine) that I cannot say I really read it. It was only three years later that I understood Battle's paper, not by reading it once more, but when listening to a talk given by a young student at the École Polytechnique, whose end of year's work was to expose Battle's paper: the poor guy had great difficulties in explaining the paper, the audience in the room was totally lost in perplexity, while I had the satisfaction to feel that a three year long ripening of understanding eventually allowed me to understand a physicist's point of view. From time to time, I now find the answer to some open questions I have trouble with just by trying to think "as Battle would do"... But again, human mediation was stronger than printed rough material.

Definitely, the top paper I have not read is Hedberg's paper which, in only six pages published in 1972 in the *Proceedings of the American Mathematical Society*, established a pointwise inequality for Riesz potentials, allowing a direct and simple proof of Sobolev embeddings. I find it a seminal paper; I can track its influence in many important works in functional analysis and partial differential equations. When I heard of this paper (I was very lucky: randomly turning the pages of Adams and Hedberg's book on potential theory at my University's library, I had the surprise to see the book open just by itself to the very page where this inequality is recalled), I entered a new dimension: now, there are distinctly two categories of analysts in my opinion, those who know this inequality and those who don't. In a recent paper of Adams about his "love affair" with Sobolev inequalities, I could find another warm tribute to this short paper of Hedberg. But, one more time, the actual journal did not play a special role, as even Adams heard very late about Hedberg's paper.

Thus far, I was at a total loss about which paper and which journal I could cite. Journals do play a central role in the dissemination of science and knowledge, peer review play a crucial role in the validation of results. But which paper should I cite?

I finally chose two papers, which had a prominent role in my research and which I encountered truly as papers published in a material journal.

The first one is a short paper by Yves Meyer: "Remarques sur un théorème de J. M. Bony" (*Supplemento ai Rendiconti del Circolo Matematico di Palermo, Serie II* **1** (1981), 1–20). This paper was the basis for the $T(1)$ theorem of David and Journé, introducing the paradifferential calculus of J. M. Bony in the formalism of Littlewood–Paley decomposition. This formalism has been a faithful companion to my own work for thirty years now. But most of all, I was deeply attached to the reprint Meyer gave me, a nice well-printed and perfectly binded bunch of pages that traveled with me for years and that I read from time to time with an ever renewed pleasure: it was the first reprint I had ever had, and I loved the object. Nowadays, most "reprints" are electronic files, or hastily printed and stapled pages, and this material pleasure of holding such a little treasure is vanishing. One of my greatest pleasures in publishing in the *Revista Matemática Iberoamericana* has always been the neatly designed reprints they send to the authors. When you do some hard work, you definitely appreciate to see it enclosed in a nice environment.

The second one is a paper published some six years ago in the *Revista Matemática Iberoamericana*: Wang, W.-K. and Xu, C.-J., "The Cauchy problem for viscous shallow water equations" (*Rev. Mat. Iberoamericana* **21** (2005), no. 1, 1–24). It is not "the most important" in the field, but as a matter of fact we are not asked to quote the most important papers in our field, but those which are special to us. So, why is that paper so special to me? When I read the paper, I did not feel it as very striking: the methods were not very new (they follow the formalism introduced by J. Y. Chemin for studying the equations in fluid mechanics with help of the Littlewood–Paley decomposition) and the results were far from optimal (the optimal result was given in 2008 by Chen, Miao and Zhang in the *SIAM Journal of Mathematical Analysis*). Thus, it could seem strange that I retain this paper as my special paper. But, indeed, it is special paper for me, published in a special journal. I work in a young University (it was founded only twenty years ago), in a small mathematics department. Our library is very modest. Now, we may access to many papers online, but at the beginning we had only few papers that our library received: in analysis, there was only *Annales de l'Institut Fourier*, *Potential analysis*, *SIAM Journal of Mathematical Analysis*... and *Revista Matemática Iberoamericana*.

With such a restricted choice, I had time to have a random look to the new arrivals in the library. Thus, I found that paper on shallow water, and I read it. It was a mixture of Navier–Stokes equations (a theme I had been working on for ten years) and Euler equations (something which remained mysterious to me) Then, I read it with growing interest and dissatisfaction: clearly, the method was a method that I could understand and the result was not optimal. I could try and get a better result. But I had no time and when I tried to find some time, the optimal result had already been published. However, it was the first time I was confident in my ability to handle transport equations, and thereafter I developed my own approach of Euler equations, as a remote answer to my first dissatisfaction. By now, Euler equations have turned into a thrilling field of investigation for me.

I don't think that the "googleized" situation we all know nowadays allows such bifurcations. Finding some unexpected sources of inspiration (what is coined as "serendipity" in the science of information, following a term introduced by H. Walpole at the end of the 18th century) is not only a matter of key words and data bases. The actual meeting with the material support of ideas, such as books and journals, remains important for opening new horizons, and I still hope that libraries will not wholly migrate into the digital world.

Rafael de la Llave (Madrid, Spain, 1957) obtained his PhD in mathematics at Princeton University in 1983 (advisor: Arthur S. Wightman), and currently is Full Professor in Mathematics at the University of Texas at Austin (starting fall 2011 at Georgia Tech.). His main research areas are dynamical systems and mathematical physics.

Looking back at “A Regularity lemma for functions of several variables”

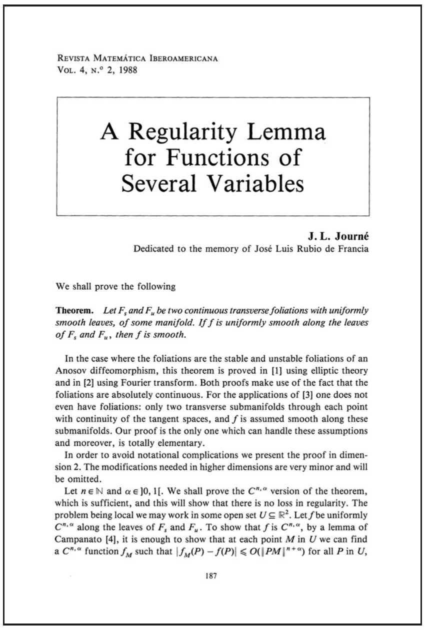
RAFAEL DE LA LLAVE*

We look back at the regularity result established in the paper quoted in the title, presenting its context, as well as further developments and related open questions. The presentation is rather informal.

1. Introduction

It is an honor to be invited by the editors of *“La Revista”* to make a presentation of an article that has been influential for me. Since there were so many choices, I decided to take a paper published in Revista, [19], in which I was somewhat involved.

Following the editors’ instructions, I will try to explain why I was interested in this paper and which are its repercussions. I will propose several questions that, to the best of my knowledge, still remain open. I cannot give a comprehensive view of the problem, but I hope I will be able to show some of the connections to other areas of mathematics.



REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

A Regularity Lemma for Functions of Several Variables

J. L. Journé
Dedicated to the memory of José Luis Rubio de Francia

We shall prove the following

Theorem. *Let F_s and F_u be two continuous transverse foliations with uniformly smooth leaves, of some manifold. If f is uniformly smooth along the leaves of F_s and F_u, then f is smooth.*

In the case where the foliations are the stable and unstable foliations of an Anosov diffeomorphism, this theorem is proved in [1] using elliptic theory and in [2] using Fourier transform. Both proofs make use of the fact that the foliations are absolutely continuous. For the applications of [3] one does not even have foliations: only two transverse submanifolds through each point with continuity of the tangent spaces, and f is assumed smooth along these submanifolds. Our proof is the only one which can handle these assumptions and moreover, is totally elementary.

In order to avoid notational complications we present the proof in dimension 2. The modifications needed in higher dimensions are very minor and will be omitted.

Let $n \in \mathbb{N}$ and $\alpha \in]0, 1[$. We shall prove the $C^{n,\alpha}$ version of the theorem, which is sufficient, and this will show that there is no loss in regularity. The problem being local we may work in some open set $U \subseteq \mathbb{R}^2$. Let f be uniformly $C^{n,\alpha}$ along the leaves of F_s and F_u. To show that f is $C^{n,\alpha}$, by a lemma of Campanato [4], it is enough to show that at each point M in U we can find a $C^{n,\alpha}$ function f_M such that $|f_M(P) - f(P)| \leqslant O(\|PM\|^{n+\alpha})$ for all P in U,

187

As an modest personal note, let me mention that the article [19] was written while Jean Lin Journé and I shared the exciting atmosphere of the Mathematics Department at Princeton and had a good personal relation outside the office. So, reflecting on these 20+ year old questions brings me a bit of melancholy. I can also look back a few years earlier and remember when I first heard about the project of having a first class journal in Spain. It sounded exciting but somehow impossible. I think that one of the greatest achievement of the creators of *La Revista* is that the new generations cannot even imagine what an achievement it was to carry out at that time projects that then seemed impossible, but that now look so normal.

*llave@math.utexas.edu.

2. Description of the result in [19]

The problem considered in [19] is a very basic regularity question of functions of several variables.

Roughly, we assume that a function is several times differentiable along the leaves of two transversal foliations. Then we want to conclude that the function is smooth. We assume that the leaves of each of the foliations are uniformly smooth, *but we do not assume* any regularity of the leaves when we change the point.

We note that the problem is local, hence, we can set it up in $\mathbb{R}^d$ and consider only a neighborhood. As a matter of fact, as we will explain later, we can reduce it to the case $d = 2$. See the end of Section 3.2.

More precisely. We assume:

H1) We are given two foliations in $\mathbb{R}^2$, that is, two collections $\mathcal{W}^1, \mathcal{W}^2$ of smooth curves indexed by the base points $x \in \mathbb{R}^2$, we assume that $\mathcal{W}^i_x$ are uniformly $C^{r+\alpha}$. $r \in \mathbb{N}$, $r \geq 1$, $\alpha \in (0, \mathrm{Lip})$.

1) $x \in \mathcal{W}^i_x$
2) $\mathcal{W}^1_x \cap \mathcal{W}^i_y \neq \emptyset \quad \Longrightarrow \mathcal{W}^i_x = W^i_y$
3) When parameterized by arc length the manifold are uniformly $C^{r+\alpha}$.
4) The jets $J^r \mathcal{W}_x$ depend continuously on the point x.
5) The foliations are transversal. $T_x\mathcal{W}^1_x \otimes T_x\mathcal{W}^2_x = \mathbb{R}^2$.

Assumptions 1), 2) are the standard definition of foliation. Assumption 3) formalizes that the leaves are uniformly differentiable. We denote by J^r the r-jet of the manifolds. In this simple context, one could also consider the $r + \alpha$ jet (say that the mapping $t \to \mathcal{W}^i_x$ is continuous when the curves are given the $C^{r+\alpha}$ topology), but it is not geometrically natural. Also note the case $\alpha = 0$ will give difficulties later when $r > 1$.

An example to keep in mind about the foliations is as follows. Consider increasing continuous real functions f_1, f_2 such that $f_i(x) \geq x + 0.1$, $f_i(0) = 0.1$ and $f_i(1) = 1.1$. The foliation $\mathcal{W}^1$ will consist of the straight lines joining $(0, x)$ to $(1, f_1(x))$ and the foliation $\mathcal{W}^2$ will consist of the straight lines joining $(0, f_2(x))$ to $(1, x)$. They satisfy the assumptions in a open neighborhood around $(1/2, 1/2)$.

H2) We have a function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ such that

- $\|\varphi|_{W^i_x}\|_{C^{r+\alpha}} \leq K$.

Theorem 1 *Under assumptions* **H1** *and* **H2** *we have that* $\varphi \in C^{r+\alpha}(\mathbb{R}^2)$.

3. Some preliminary remarks

3.1. The case $r = 1$

The case $r = 1$ is classical. In this case, we can also allow $\alpha = 0, \mathrm{Lip}$. It is classical that a function is $C^{1+\alpha}$ if and only if it has C^α partial derivatives. The proof, based on integrating the mean value theorem, can be very easily adapted to the situation when, rather then having partial derivatives, one has derivatives along a foliation.

For derivatives of higher order, the problem is that one has only derivatives along several directions and has to reconstruct the mixed derivatives.

It should be remarked that this is false for C^2 and that there are classic examples (see, for instance, [22]) which show that there are functions whose derivatives along coordinate axis are smooth but that lack mixed derivatives (this is clearly related to the fact that the inverse Laplacian is not bounded from C^0 to C^2).

3.2. Coordinate foliations

When the foliations are the coordinate lines, the result is classical. One can find proofs in [35, 22].

Of course, when the foliations are C^∞ (that is, the leaves are C^∞ and the jet depends continuously on the base point), we can reduce the problem to the coordinate case.

The proofs of [35, 22] are based on the theory of singular integrals (Riesz potentials), which is a natural way to reconstruct the mixed derivatives from the derivatives along certain directions.

It is well known by specialists in harmonic analysis that the theory of Riesz potentials is not simple in $C^{r+\alpha}$ spaces when $\alpha = 0, \mathrm{Lip}$. This is related to the reasons why we eliminated $\alpha = 0$ in our statements.

Note also that this allows us to reduce to the case $d = 2$. If the foliations are in $\mathbb{R}^d$, we just take intersections with two dimensional planes where they are local. Once we conclude the regularity for arbitrary two dimensional planes, we can use the coordinate argument.

4. Some idea of the proof in [19]

The proof in [19] is based on some elementary observations. For simplicity, we will just describe it in the case of coordinate foliations. Then, we will indicate the changes needed. Fix $0 < \lambda < 1$.

The goal is to show that we can approximate the function by a Taylor polynomial in a neighborhood of a point, that we will assume, to simplify the notation, it is at the origin.

We consider grids of points in $\mathbb{R}^2$ of the form

$$\mathcal{G}^{N,M} = \{(\lambda^n, \lambda^m)|\ 0 \le n \le N + D, 0 \le m \le M + D\} \tag{1}$$

Because the grid $\mathcal{G}^{N,M}$ is a product, it is possible to find a polynomial of degree $D \times D$, $P^{N,M}(x, y) = \sum_{0 \le n,m \le D} c_{n,m} x^n y^m$, that agrees with φ on the grid.

The polynomial $P^{N,M}$ is found by using Lagrange interpolation on each vertical line using polynomials in y. Therefore, we have a function from x taking values in y-polynomials. Then, we can use again Lagrange interpolation and find a x-polynomial whose coefficients are y polynomials. Of course, one could do the interpolation in x first and then the interpolation in y. The result is the same. Indeed, there is a uniqueness lemma that states that the interpolating polynomial is unique.

The key property to keep in mind is that, because the exponential separation of the points, the coefficients of the Lagrange interpolation have good stability properties (these stability properties are absent, *e.g.*, if we take equally spaced points, as it is well known).

Noting that φ is smooth along the x direction, we can compare $P^{N+1,M}$ to $P^{N,M}$ because the interpolating polynomials along the horizontal lines do not change too much. Then, the second interpolation in y does not change much either. A similar argument, using the regularity along the y direction, allows to compare $PN, M + 1$ to $P^{N,M}$ so that, by doing the process twice, we can compare $P^{N+1,M+1}$ to $P^{N,M}$.

Of course, we have omitted the precise statement of the comparisons, but in [19] it is shown that we can get enough control to show that as $K \to \infty$, the coefficients of degree less or equal than D of $P^{N+K,M+K}$ converge.

An extra argument shows that this polynomial (which is independent of the sequence of grid used and of λ) satisfies the same bounds as a Taylor approximation.

Since the point is arbitrary, we have shown that in a neighborhood there is a "Taylor approximation" around every point. It is not hard to show that, because it is uniformly bounded and continuous (A compactness argument and the uniqueness of the Taylor approximation gives that the graph of the mapping $(x, y) \to P^D_{x,y}$ is closed, hence it is compact).

The argument is robust enough so that it can incorporate the fact that the foliations are not exactly the product, but, of course, are close to the product in a small enough neighborhood. The stability of the interpolation when we change the interpolation grid slightly is closely related to the fact –that we mentioned without detail– that in exponential grid the coefficients of the interpolation polynomial are controlled by the values of the function.

Once one has a Taylor expansion at every point, it is an standard result to show that the function is smooth. This is done in [1, 23], where it is called *Converse Taylor theorem*. Note that there are more general results in [6], where the approximation of Taylor's theorem is just required to be in the sense of L^p of a small ball.

4.1. The origins of the problem and some previous developments

The problem is, of course, a natural one and could have been asked hundreds of years ago. But as a matter of fact it was originated in Dynamical Systems. Since this origin also motivated several of the later developments, we will try to discuss it.

When one considers a diffeomorphism $f : M \to M$, M a manifold, and $\eta : M \to \mathbb{R}$, the simplest natural and nontrivial linear equation one can form is

$$\varphi \circ f - \varphi = \eta \tag{2}$$

Indeed, these equations appear very frequently in dynamical systems, representation theory, etc., and are called *cohomology equations*. Their regularity theory opens the gate to many questions in stability, rigidity, inverse spectral theory, etc.

If f is induced by algebraic operations, *e.g.*, f is a rotation on a torus or an automorphisms of a homogeneous space, one can consider using representation theory (Fourier analysis in the case of rotations). See [26] for a rather general use of representation theory methods in the study of (2) and [33] for the case of rotations. The case of rotations is the basis of KAM theory. For applications to rigidity theory, see [41].

One interesting case for dynamicists is when f is an Anosov system. That is, there is a splitting $T_xM = E^s_x \oplus E^u_x$ invariant under f, $Df(x)E^{s,u}_x = E^{s,u}_{f(x)}$. and, for some N large enough $\|Df^N(x)|_{E^s_x}\| \le 1/2$, $\|Df^{-N}(x)|_{E^u_x}\| \le 1/2$. That is, one can find directions that contract exponentially either in the future or in the past.

It was known very early in the theory of Anosov systems (certainly it is reviewed in [2]) that one can find foliations $\mathcal{W}^s$ and $\mathcal{W}^u$ whose leaves are tangent to E^s and E^u respectively. The leaves are as smooth as the map, nevertheless, the dependence from point to point of the jets could be only Hölder, even if the map f is analytic. The integrability of these stable and unstable distributions is a very surprising fact, which depends a lot on dynamic arguments. It does not follow from regularity properties of the jets in the case that they are one-dimensional or –as far as we know– from Frobenius theorem in higher dimension. Indeed, there are some other dynamically defined distributions with very similar properties, which are not integrable [18].

Equations of the form (2) over Anosov systems were studied in [24, 15] for very different motivations. The paper [24] developed a theory of existence of C^α $0 < \alpha <$ Lip solutions that was extended to $C^{1+\alpha}$ $(0 \leq \alpha \leq \text{Lip})$ solutions in [15] for general systems and for C^∞ solutions for geodesic flows on surfaces of negative curvature.

It was observed in [7] that the C^0 solutions of (2) are differentiable along the leaves of the stable and unstable foliations. Formally, if (2) holds, we have

$$\varphi(x) = \eta(x) + \eta \circ f(x) + \cdots + \eta \circ f^n(x) + \cdots \tag{3}$$

Even if the formula (3) is quite formal, one can observe that the formula that one obtains taking derivatives along the stable directions is very convergent. Denoting by D_x the derivative along a vector field tangent to E^s,

$$\begin{aligned} D_S\varphi(x) &= D_s\eta(x) + D_s\eta \circ f(x) + \cdots + D_s\eta \circ f^n(x) + \cdots \\ &= D\eta(x)v_s + D\eta \circ f\, Df(x)v_s + \cdots + D\eta \circ f^n(x)Df^n(x)v_s + \cdots \end{aligned} \tag{4}$$

Since the factors $Df^n(x)v_s$ decrease exponentially fast with n, the above one is a very convergent series.

Once one can justify (4) by taming the application of (3) (one justification using approximation by periodic orbits can be found in [7], another one using Cesàro sums in [8]), one can proceed to take further derivatives. Of course, an analogue argument works in the inverse direction.

Hence we are in the situation of Theorem 1. So the Theorem 1 leads to a regularity theory of solutions of (2), which as mentioned before leads to quite a number of results.

4.2. Some previous results

It is interesting to compare the proof of Theorem 1 with other previous proofs of similar results.

4.2.1. The elliptic regularity approach

The first similar result appears in [7]. The main idea is that, because we can take derivatives along the stable and unstable directions, we can compute and find that given $a \gg 1$, φ satisfies

$$(-1)^N [D_s^{2N} + D_u^{2N}]\varphi + a\varphi = \nu \tag{5}$$

where $\nu \in C^0(M)$.

It is amusing to note that the properties that in dynamical systems are called *hyperbolicity* correspond closely to the fact that the operator above is *elliptic*.

So, one can hope to use the regularity theory of the elliptic equations to find the φ has to be smooth.

One has to be careful for several reasons: The first is that the coefficients are not very smooth (only Hölder). The second one is that the equation is satisfied in the sense of the old fashioned derivatives along certain leaves.

The fact that the coefficients are not smooth is not a problem for the method of freezing of coefficients and it is possible to find a function Ψ such that $(-1)^N[D_s^{2N} + D_u^{2N}]\Psi + a\Psi = \nu$. Now, the question is to prove uniqueness to conclude that $\Psi = \varphi$.

There are several ways to prove uniqueness of elliptic PDE's of high order. All of them somehow require integrating by parts and that $D_s^* = -D_s + \text{loworder}$.

To justify the integration by parts, the method of [7] was to improve on a result of [2] called *absolute continuity of foliations*. This property shows that there is mapping that sends the stable (resp. unstable) foliation into the standard foliation and that this transformation has an absolutely continuous Jacobian. This absolute continuity is crucial to prove the ergodicity of Anosov systems, in particular, geodesic flows of manifolds of negative curvature following a classic argument of Hopf.

The method of [7] was to prove that the Jacobian of this transformation can be chosen to be differentiable along the stable directions. So that using the transformation Γ of Jacobian J, we can transform

$$\int \Psi D_s \varphi = \int \Psi \circ \Gamma \frac{\partial}{\partial x} \varphi \circ \Gamma\, J$$

So that the integration by parts is possible.

The proof of the differentiability of the Jacobian in [7] follows by examining carefully the proof of absolute continuity in [2]. (The better known proof in [29] does not seem to yield so well to generalizations). Another method was remarked in [8]. One observes that by the invariance of the parameterization, one can get that the Jacobian satisfies in patches $J \circ f \det D_s f = J$ and taking logarithms $\log J \circ f - \log J = -\det D_s f$. In other words (ignoring issues about patches, which are irrelevant to computing derivatives), we obtain that $\log J$ satisfies an equation of the form (2) so that the argument for existence of derivatives can be adapted.

Another proof of a similar regularity lemma appeared in [17]. The proof is based on studying $\hat{\varphi}_k \equiv \int \varphi(x,y) e^{i(kx+ly)}\,dxdy$ (of course, one has to use cutoff functions etc.).

We can transform variables using the stable foliation and the smoothness of the Jacobian (already established in [7]) and obtain that one can bound $|\hat{\varphi}_k| \le C|\Pi_s(k,l)|^{-r}$ and a similar argument in the unstable direction gives a similar decay in the other direction. One, of course, is reminded of the use of Fourier analysis to conclude regularity of elliptic equations.

Of course, using Fourier series and estimates on the coefficients rather than integral representations leads to a loss of regularity which grows with the dimension because the size of the Fourier coefficients of a function does not capture very well its regularity. Even if one tries to obtain estimates of the identity operator one gets a loss.

The elliptic approach was extended significantly in [8, 30]. Of course, the use of the smoothness of the Jacobian remains.

4.2.2. Results for Hölder regular foliations

Before [19], there was another argument in [20], again based on proving by induction in the degree the existence of a Taylor approximation. This, however, required that the foliation was Hölder. The idea of the proof was to consider the Taylor expansion at neighboring points. One shows that if they were not differentiable with respect to the point, it would be impossible that the original function has the assumed regularity.

5. Some further developments

5.1. Analytic regularity

The method using Fourier coefficients was revisited in [9] to deal with analytic regularity. This needs basically to be very careful because many of the tools in elliptic regularity (cut-offs, partitions of unity) do not work.

5.2. Results for more complicated sets

In [10] it was noted that the proof above can be extended to situations when the assumption of foliations is somehow weakened.

It suffices that for points in *thick sets*, one has smooth curves that pass through the point. The fact that these leaves intersect (and hence fail to be a foliation) does not have a big importance.

A motivation for this generalization was the examples in [18], which show that several distributions were not integrable by smooth leaves. Nevertheless, it was shown in [14] that one can find families of smooth leaves that map into each other and have a dynamical meaning. They typically cross and fail to satisfy part 2) of **H1**.

One of the thickness properties we require is that if p, p' are points in the set, then $W^s_p \cap W^u_{p'}$ intersect (so that we can define the grid). This property is well known to dynamicists in many situations and is called the *"local product structure"*.

Another much more subtle property is that in the set, if we have approximation by a Taylor polynomial, we can use the arguments in the converse Taylor theorem and conclude that the function is smooth.

Note that for general closed sets, the fact that Taylor approximations suffice to conclude regularity is false. One has to use the much more sophisticated Whitney extension theorem [38, 39, 35, 13]. The hypotheses of the Whitney extension theorem require not only the existence of a Taylor approximation, but also subsequent conditions that say that these Taylor expansions are differentiable with respect to the basis.

Nevertheless, if the sets have sufficient densities, the Taylor polynomials are unique and one can imitate the proof of the converse Taylor theorem. Informally, the sets are so close to being the whole space that the Whitney embedding theorem reduces to the Converse Taylor Theorem.

The motivation of [10] was the study of non-uniformly hyperbolic maps. This is a very deep theory (also known as Pesin theory) [28, 32, 5]. It shows that, in many circumstances, one can find the stable and unstable manifolds of Anosov systems, but not defined everywhere, only on sets of large measure (for any invariant measure). Furthermore, the foliations studied in Pesin theory are only measurable, not continuous in general. Fortunately, thanks to Lusin theorem, one can get large sets where they are continuous (and, using dynamical arguments, bootstrap continuity to Hölder). Hence, the arguments of [10] apply in closed sets of large measure, contained in sets of density points of the Pesin sets, which are of measure $1 - \epsilon$.

A remarkable improvement of these circle of ideas happened in [27], where the analogue questions for hyperbolic sets in two dimensions were studied (one example is the famous Smale horseshoe). These sets are rather thin (they have Hausdorff dimension smaller than 2). Nevertheless, [27] show that they have quite a number of very interesting geometric properties which allow us to extend the argument.

In [27], the somewhat crude density arguments of [10] get replaced by very remarkable delicate geometric properties that are then verified for hyperbolic sets in two-dimensional systems.

5.2.1. Partially hyperbolic systems

A very active area in dynamical systems is partially hyperbolic systems. Two recent surveys are [16, 31], but the field is growing very fast. These partially hyperbolic systems have stable and unstable directions, satisfying the same properties as those in Anosov systems, but they do not span the whole tangent space. These tangent spaces integrate to foliations.

One can also study the cohomology equations over these systems. In particular, they appear naturally in the study of actions [21].

Studying the regularity of cohomology equations (2), one can adapt the argument leading to (4) (the approximation by periodic orbits is iffy, but the Cesàro means works) and obtain that the solution is indeed differentiable. Now, one is left with the problem of concluding regularity. If one tries to use the method of [7], one finds that the operator (5) is not elliptic.

The paper [21] considers systems of algebraic origin, so that the foliations are very smooth. In that case, one can define the commutators of derivatives along leaves and conclude that the operator (5) satisfies the assumptions of the regular hypoelliptic theory.

Unfortunately, for general partially hyperbolic systems, the coefficients are not smooth and one cannot study commutators. Nevertheless, there is a natural property that has been widely studied by dynamicists. We say that a system is *locally accessible in order L* when, starting at any point, we can reach a whole neighborhood by walking along stable and unstable directions (in such a way that to reach distance ρ one needs to travel $C\rho^{1/L}$). It is easy to see that this property is equivalent to the commutator property in the case of smooth foliations, but it makes sense in general.

It would be very natural to try to prove the result of regularity when one has regularity along foliations that allow one to move from one place to another.

The question of regularity of solutions of cohomology equations for partially hyperbolic stystems has been considered also in [40], but not following the above route. It introduces some extra hypothesis (bunching) that show that one can get some invariant objects along the missing directions and presents a remarkable argument that shows that one can get regularity along these directions, so that one can resort to the already established papers.

6. Some open questions

In this section, I will describe some open questions that are suggested by all the above developments. Some of the problems are more routine than others.

The problems are indeed problems in analysis, but the motivations of some of them come from Dynamical Systems. The area of Dynamical Systems is really not a subject but a problem (humans have for a long time tried to understand how to throw rocks efficiently or how the planets move), and it feeds on and stimulates many other areas of Mathematics.

So, here it is a very small sample of the problems which seem to require tools from analysis and that are closely related to [19].

6.1. The borderline cases

One straightforward problem is to study what happens in the borderline cases of the regularity. In the above discussion, sometimes the cases $\alpha = 0$ or $\alpha = \mathrm{Lip}$ were left out.

Clearly, in the borderline case, the analogue is to use the Λ_r spaces. As shown in [35], these are the right spaces for the theory of Riesz potentials. One can also reconstruct the regularity in $\mathbb{R}^2$ from the regularity along coordinate foliations. As it is well known, when $r \notin \mathbb{N}$, $\Lambda_r = C^{[r]+\{r\}}$, but when $r = \mathbb{N}$ they are genuinely new spaces.

Sometimes, the study of the last cases is just an academic endeavor. Nevertheless, there are some interesting cases of Dynamical Systems in which the foliations have naturally Λ_1 or Λ_2 regularity. See [17, 12].

I should mention that one of the really frustrating issues in Dynamical Systems is that there is a very large branch of knowledge establishing properties of C^1 generic systems [25]. Unfortunately, all the work on smooth ergodic theory (including the absolute continuity of foliations of [2] and the theory of invariant manifolds in Pesin theory) requires $C^{1+\alpha}$. So that the two programs do not have much common ground. It is clear that some of the results of smooth ergodic theory are obtained with a modulus of continuity for the derivative weaker than Hölder (which satisfies some summability conditions), but there are counterexamples of may results of smooth ergodic theory for C^1 maps.

For the analysts working in dynamical systems this gap is a source of frustration. Giving a proof of the result of [25] –or any similar result– in which C^1 is replaced by any modulus of continuity acceptable to smooth ergodic theory has been recognized as a great problem by many people.

6.2. Analytic regularity

A shortcoming of the proof in [9] mentioned in Section 5.1 is that the domain of analyticity established is roughly half of the domain of analyticity of the original function. This is unfortunate since one would like to use this arguments in KAM theorems. The proof of [9] also requires that the Jacobian of the foliations is analytic.

I think that it would be interesting to know if one can prove the analogue of Theorem 1 for analytic regularities (without using the regularity of the Jacobian) and obtain that φ has a radius of analyticity as large as that of the restrictions.

6.3. Reducing the Whitney extension to the converse Taylor theorem

We have seen in Section 5.2 that, in some sets, one can use only the Taylor approximation to conclude regularity rather than the most complicated conditions of Whitney extension theorem, that require that the Taylor approximations are also differentiable.

Remarkably, similar questions appear in other areas of dynamical systems such as studying the Whitney regularity of the KAM tori as a function of the frequency (the frequency is required to range in a complicated set of vectors satisfying Diophantine properties). A detailed study of this application can be found in [37].

It would be interesting to study more systematically which sets have the property that on them the converse Taylor theorem applies. In other words, to determine the geometric properties of sets which make all the hypothesis of Whitney's extension theorem implied by just the first.

As suggested in [27] it would be interesting to study whether some classes of hyperbolic sets in 3 or more dimensions satisfy these properties. It is possible that the "solenoid" [34] and its perturbations are among the sets for which the converse Taylor theorem applies.

6.4. Using several foliations

The argument in [19] is very tied up to the fact that we are considering only two foliations. On the other hand, the elliptic regularity methods of [7, 17, 9] can accommodate several foliations (provided, of course, that they satisfy the regularity property of the Jacobian).

I think that it would be very interesting to extend the method of [19] to three or more foliations. The main problem is that one cannot really construct a product grid unless the foliations integrate pairs. Maybe the mismatch is so small in small scales that one can adapt the argument.

Extending the method of [20] to three or more foliations (which are Hölder) seems more doable.

Either of them would have consequences for dynamical systems.

6.5. A hypoelliptic version

As mentioned in Section 5.2.1, the theory of partially hyperbolic systems leads to the question of whether there is an analogue of [19] for foliations in which one can reach all the neighborhood by moving along leaves.

We assume that a neighborhood of size ρ can be reached by segments contained in leaves of total length no more than $C\rho^{1/L}$

One suggestion that something like this should be true comes by comparing [26] and [11].

6.6. Anisotropic spaces

A very interesting recent development has been the systematic introduction of spaces that take into account some dynamic properties [4, 3, 36].

These spaces allow one to establish properties of dynamical systems in rather analytic ways. For example, the paper [36] allows to prove ergodicity of Anosov systems without taking the step of proving explicitly absolute continuity of the stable and unstable foliations.

The regularity lemmas could be a family of other results obtained some regularity in the direction of the foliations and some transversal (smaller) regularity.

Acknowledgment. The preparation of this paper has been supported by NSF grant DMS 0901389 and by Texas Coordinating board NHRP 0223.

References

[1] ABRAHAM, R. AND ROBBIN, J.: *Transversal mappings and flows.* W. A. Benjamin, New York–Amsterdam, 1967.

[2] ANOSOV, D. V.: Geodesic flows on closed Riemannian manifolds of negative curvature. *Trudy Mat. Inst. Steklov.* **90** (1967),

[3] BALADI, V. AND GOUËZEL, S.: Good Banach spaces for piecewise hyperbolic maps via interpolation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** (2009), no. 4, 1453–1481.

[4] Baladi, V. and Tsujii, M.: Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms. *Ann. Inst. Fourier (Grenoble)* **57** (2007), no. 1, 127–154.

[5] Barreira, L. and Pesin, Y.: *Nonuniform hyperbolicity.* Encyclopedia of Mathematics and its Applications **115**. Cambridge University Press, Cambridge, 2007.

[6] Campanato, S.: Proprietà di una famiglia di spazi funzionali. *Ann. Scuola Norm. Sup. Pisa (3)* **18** (1964), 137–160.

[7] De la Llave, R., Marco, J. M. and Moriyón, R.: Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math. (2)* **123** (1986), no. 3, 537–611.

[8] De la Llave, R.: Remarks on Sobolev regularity in Anosov systems. *Ergodic Theory Dynam. Systems* **21**, (2001), no. 4, 1139–1180.

[9] De la Llave, R., Analytic regularity of solutions of Livsic's cohomology equation and some applications to analytic conjugacy of hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **17** (1997), no. 3, 649–662.

[10] De la Llave, R., Smooth conjugacy and S-R-B measures for uniformly and non-uniformly hyperbolic systems. *Comm. Math. Phys.* **150** (1992), no. 2, 289–320.

[11] Dolgopyat, D.: On decay of correlations in Anosov flows. *Ann. of Math. (2)* **147** (1998), no. 2, 357–390.

[12] Fang, Y., Foulon, P. and Hasselblatt, B.: Zygmund strong foliations in higher dimension. *J. Mod. Dyn.* **4** (2010), no. 3, 549–569.

[13] Fefferman, C.: Whitney's extension problems and interpolation of data. *Bull. Amer. Math. Soc. (N.S.)* **46** (2009), no. 2, 207–220.

[14] Fontich, E., de la Llave, R. and Martín, P.: Invariant pre-foliations for non-resonant non-uniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **358** (2006), no. 3, 1317–1345.

[15] Guillemin, V. and Kazhdan, D.: On the cohomology of certain dynamical systems. *Topology* **19** (1980), no. 3, 291–299.

[16] Hasselblatt, B. and Pesin, Y.: Partially hyperbolic dynamical systems. In*Handbook of dynamical systems. Vol. 1B*, 1–55. Elsevier, Amsterdam, 2006.

[17] Hurder, S. and Katok, A.: Differentiability, rigidity and Godbillon-Vey classes for Anosov flows. *Inst. Hautes Études Sci. Publ. Math.* no. 72 (1990), 5–61.

[18] Jiang, M., Pesin, Ya. B. and De la Llave, R.: On the integrability of intermediate distributions for Anosov diffe omorphisms. *Ergodic Theory Dynam. Systems* **15** (1995), no. 2, 317–331.

[19] Journé, J.-L.: A regularity lemma for functions of several variables. *Rev. Mat. Iberoamericana* **4** (1998), no. 2, 187–193.

[20] Journé, J.-L.: On a regularity problem occurring in connection with Anosov diffeomorphisms. *Comm. Math. Phys.* **106** (1986), no. 2, 345–351.

[21] Katok, A. and Spatzier, R. J.: Subelliptic estimates of polynomial differential operators and applications to rigidity of abelian actions. *Math. Res. Lett.* **1** (1994), no. 2, 193–202.

[22] Krantz, S. G.: Lipschitz spaces, smoothness of functions, and approximation theory. *Exposition. Math.* **1** (1983), no. 3, 193–260.

[23] Krantz, S. G. and Parks, H. R.: *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. Birkhäuser Boston, Boston, MA, 2002.

[24] Livšic, A. N.: Certain properties of the homology of Y-systems. *Mat. Zametki* **10** (1971), 555–564.

[25] Mañé, R.: A proof of the C^1 stability conjecture. *Inst. Hautes Études Sci. Publ. Math.* no. 66 (1988), 161–210.

[26] Moore, C. C.: Exponential decay of correlation coefficients for geodesic flows. In *Group representations, ergodic theory, operator algebras, and mathematical physics (Berkeley, Calif., 1984)*, 163–181. Math. Sci. Res. Inst. Publ. **6**. Springer, New York, 1987.

[27] Nicol, M. and Török, A.: Whitney regularity for solutions to the coboundary equation on Cantor sets. *Math. Phys. Electron. J.* **13** (2007), Paper 6, 20.

[28] Pesin, Ja. B.: Characteristic Ljapunov exponents, and smooth ergodic theory. *Uspehi Mat. Nauk* **32** (1977), no. 4, 55–112.

[29] Pugh, C. and Shub, M.: Ergodicity of Anosov actions. *Invent. Math.* **15** (1972), 1–23.

[30] Rauch, J. and Taylor, M.: Regularity of functions smooth along foliations, and elliptic regularity. *J. Funct. Anal.* **225** (2005), no. 1, 74–93.

[31] Rodríguez Hertz, F., Rodríguez Hertz, M. A. and Ures, R.: A survey of partially hyperbolic dynamics. In *Partially hyperbolic dynamics, laminations, and Teichmüller flow*, 35–87. Fields Inst. Commun. **51**. Amer. Math. Soc., Providence, RI, 2007.

[32] Ruelle, D.: Ergodic theory of differentiable dynamical systems. *Inst. Hautes Études Sci. Publ. Math.* no. 50 (1979), 27–58.

[33] Rüssmann, H.: On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, 598–624. Lecture Notes in Phys. **38**. Springer, Berlin, 1975.

[34] Shub, M.: *Global stability of dynamical systems*. Springer-Verlag, New York, 1987.

[35] Stein, E. M.: *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series **30**. Princeton University Press, Princeton, N.J., 1970.

[36] Tsujii, M.: Quasi-compactness of transfer operators for contact Anosov flows. *Nonlinearity* **23** (2010), no. 7, 1495–1545.

[37] Vano, J.: *A Whitney-Zehnder implicit function theorem.* University of Texas at Austin, 2002.

[38] Whitney, H.: Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36** (1934), no. 1, 63–89.

[39] Whitney, H.: On the extension of differentiable functions. *Bull. Amer. Math. Soc.* **50** (1944), 76–81.

[40] Wilkinson, A.: The cohomological equation for partially hyperbolic diffeomorphisms, 2008. `arXiv:0809.4862`.

[41] Zimmer, R. J.: *Ergodic theory and semisimple groups.* Monographs in Mathematics **81**. Birkhäuser Verlag, Basel, 1984.

Gaven Martin (Rotorua, New Zealand, 1958) obtained his PhD in mathematics at the University of Michigan in 1985 (supervisor: Fred Gehring). He is Distinguished Professor and Director of the New Zealand Institute of Advanced Study, Massey University. His research focuses on geometric function theory, quasiconformal mappings and Kleinian groups.

A paper by S. K. Donaldson and D. P. Sullivan

Gaven Martin*

The paper I've chosen to acknowledge in this volume is "Quasiconformal 4-manifolds." I've also linked it to a wonderful period in my own career when I first started working with Tadeusz Iwaniec.

This essay is not only about the paper [4] which has hugely influenced the direction of my research, but also about a place and time. It's also about a collaboration starting twenty years ago and which continues to this day.

Figure 1: The Mittag–Leffler Institute

The place is Mittag–Leffler Institute in Djursholm, Sweden, and the time, around March 1990. Those of you who have been to Mittag–Leffler will know what a wonderful place it is, and those who haven't should take any opportunity they can to get there. I was visiting for six months with a good number of the quasi-world, our term for that group interested in geometric analysis, function theory, quasiconformal mappings and the like. I'd arrived in January and was working at that time on a project in the geometry of Kleinian groups with Fred Gehring, previously my PhD supervisor at Michigan. That project was ultimately completed a few years ago, solving a problem of Siegel in three dimensions –finding the minimal co-volume hyperbolic lattice [5]. But this essay is not about that.

So, sometime in March 1990 this (at that time) skinny young chain-smoking Polish mathematician Tadeusz Iwaniec shows up at Mittag–Leffler waving about in his hands a recent paper of Simon Donaldson and Dennis Sullivan [4] entitled "Quasiconformal 4-manifolds" published in Acta Mathematica (also run out of Mittag–Leffler). I'd seen the paper before and read it through nodding my head –as you do. Ostensibly the paper generalises some of Donaldson's remarkable earlier work on smooth four manifolds, for which he was awarded the Fields medal, to quasiconformal manifolds. These are manifolds for which the coordinate charts can be chosen to be quasiconformal. A few years earlier Sullivan had proved that every topological

*G.J.Martin@massey.ac.nz.

n-manifold, $n \neq 4$ admits such a quasiconformal structure allowing one to do analysis on a topological manifold. A remarkable fact in and of itself. He'd also proved the quasiconformal hauptvermutung and lots of other interesting things generalising the "furling" technique of Edwards and Kirby using hyperbolic geometry. The main theorems of Donaldson and Sullivan's papers were that there is a topological 4-manifold which does not admit a qc-structure, and there are two homeomorphic qc 4-manifolds which are not quasiconformally equivalent. Basically they showed that using qc maps and Sobolev theory you could set up the Yang-Mills equations on a qc-manifold and extract the same invariants from it that Donaldson had previously done in the smooth case.

I'd met Tadeusz before, he arrived with his family in Michigan the year I was finishing up there and lectured on quasiconformal mappings from the analytical point of view –I was far more interested in the geometric stuff so we didn't really talk too much about math (his many stories of life in Poland was another matter). I did give several research seminars at Michigan at that time on Pekka Tukia and Jussi Väisälä's "tidying up" of Sullivan's work mentioned above [9]. So it was sort of natural that we came together to talk about the Donalson-Sullivan paper at Mittag.

At this point, as an aside, I have to say that Dennis Sullivan is a bit of a personal hero of mine. So many of the things he has done have profoundly influenced the fields I work in and around that it's quite amazing. There is a quote (attributed to Bernie Maskit) along the following lines. Talking to Dennis, or listening to one of his lectures, is like getting a message from Mars. You'd better listen carefully, because although you won't understand a word he says, you know its important and when you finally decode what he means it is likely to change your view of the world.

So why was Tadeusz so excited about a paper that had been out for a while and that a few of us had seen? Well what he had seen (and was not telling anybody yet) was that embedded in the results of this paper was a proof of Fred Gehring's higher integrability result in four dimensions –a fact needed for technical reasons. Donaldson and Sullivan, perhaps not appreciating how hard this result is supposed to be, had more or less replicated in four dimensions an argument, in two dimensions, due to Bogdan Bojarski (Tadeusz' thesis advisor) and given in Lars Ahlfors' classic book "Lectures on quasiconformal mappings" [1]. Here, to help with the discussion a bit later, is a sketch of that idea –all the details (and much much more) can be found in [3]. The Beurling transform $\mathcal{S} : L^2(\mathbb{C}) \to L^2(\mathbb{C})$ is a singular integral operator defined by

$$(\mathcal{S}\varphi)(z) = \frac{-1}{\pi} \iint_{\mathbb{C}} \frac{\varphi(\zeta)\, d\zeta}{(z-\zeta)^2}\,. \tag{1}$$

It has the remarkable property that it is an isometry in L^2 and intertwines the z and $\overline{z}$ complex derivatives,

$$\mathcal{S} \circ \frac{\partial}{\partial \overline{z}} = \frac{\partial}{\partial z} \, . \tag{2}$$

Now, basically, the defining equation for a quasiconformal mapping f is the Beltrami equation,

$$\frac{\partial f}{\partial \overline{z}} = \mu(z) \frac{\partial f}{\partial z} \, , \tag{3}$$

where μ is a measurable function (the Beltrami coefficient or complex dilatation) and $\|\mu\|_\infty < 1$. Here is how improved regularity can be obtained. We put together (2) and (3) to get

$$\frac{\partial f}{\partial \overline{z}} = \mu \, \mathcal{S} \frac{\partial f}{\partial \overline{z}} \quad \text{or} \quad (I - \mu \mathcal{S}) \frac{\partial f}{\partial \overline{z}} = 0 \, .$$

Of course this is too good since the *qc*-map, apriori assumed to have locally square integrable first derivatives, does not have derivatives in $L^2(\mathbb{C})$ and so we can't apply $\mathcal{S}$. To overcome this, we multiply f by a cutoff function and rearrange terms leading to an equation of the form

$$(I - \mu \mathcal{S}) \frac{\partial f}{\partial \overline{z}} = \text{nice} \, . \tag{4}$$

where "nice" means about as good as f is as all the derivatives have fallen on the smooth cutoff function. Then of course $\frac{\partial f}{\partial \overline{z}} = (I - \mu \mathcal{S})^{-1}(\text{nice})$ and the same for $\frac{\partial f}{\partial \overline{z}}$ from (3). Thus the degree of integrability of the differential of a quasiconformal map depends on the invertibility properties of the operator $I - \mu \mathcal{S}$. Set $k = \|\mu\|_\infty$ and $\mathbf{S}_p = \|\mathcal{S}\|_{L^p(\mathbb{C}) \to L^p(\mathbb{C})}$. Then, as a consequence of the Neumann series expansion,

$$\|(I - \mu \mathcal{S})^{-1}\|_{L^p(\mathbb{C}) \to L^p(\mathbb{C})} \leq \frac{1}{1 - k \mathbf{S}_p} \, , \tag{5}$$

so the crucial thing are the p-norms $\mathbf{S}_p$, but note that $\mathbf{S}_2 = 1$ and $k < 1$. Bojarski realised that a weak-type estimate, along with the Calderón–Zygmund theory of singular integral operators, shows $\mathbf{S}_p < \infty$ and interpolation shows $\mathbf{S}_p \to 1$ as $p \to 2$. As $k < 1$ the equation (5) implies that Df is locally in a better class than L^2, this is improved regularity. The improvement depends on the number $k = \|\mu\|_\infty$ which is regarded as an ellipticity constant.

For many decades complex analysts have been struggling to prove $\mathbf{S}_p = p - 1$, for $p \geq 2$, so as to achieve the sharpest possible form of higher integrability –the Gehring–Reich conjecture. Fortunately, Kari Astala decided this was not the way to prove the optimal regularity and in another great Acta paper proved the sharp result $Df \in L^p_{loc}$ for all $p < 1 + 1/k$ (he proved much more, [2]).

Why these regularity estimates are important will become clear when we discuss what Tadeusz and I proved after reading Donaldson and Sullivan's paper.

So, back to Mittag. What Donaldson and Sullivan had done was identify many of the two-dimension structures in four dimensions. All the "quasi-world" were trying to make this stuff work in three dimensions first –bad luck there. Donaldson and Sullivan knew that in four dimensions the Hodge star $*$ acts as an isometry on 2-forms and $** = 1$ and so there are plus and minus eigenspaces giving the orthogonal decomposition $\Lambda^2 = \Lambda^+ \oplus \Lambda^-$. Then there are natural first order operators $d^+, d^- : \Lambda^1 \to \Lambda^2$ obtained by projecting the exterior derivative into each eigenspace. Using this, Donaldson and Sullivan wrote down the Beltrami equation for a mapping f as

$$d^+ f = \mu \, d^- f \, ,$$

where $\mu : \Lambda^- \to \Lambda^+$ is defined in terms of the differential of f acting on 2-forms (we later gave the explicit formula for μ). Now, an argument based around the Hodge decomposition and the conformal invariance of Hodge $*$ allowed them to follow the two-dimensional plan. The operators concerned are going to be expressed in terms of Riesz transforms –in one way or another– and so invertible near L^2.

Well any-one who knows Tadeusz knows that he hates these sorts of arguments. He wants (and loves) all the gory details. So we started going through them. Everything was more or less OK until we got to the "conformal invariance" of Hodge star bit. This is obvious geometrically when you are talking about multiplying a metric by a scalar function, but what does it really mean when you are talking about putative higher dimensional conformal mappings? It was the second or third night we'd been working on this when around 2 or 3 am we figured out that when written in terms of the differential of a Sobolev mapping the equations lead to the following observation:

Let O be an even dimensional, $n = 2\ell$, orthgonal transformation, $O^tO = Id$. Partition O up into the four $\ell \times \ell$ submatrices

$$O = \begin{pmatrix} A & B \\ C & D \end{pmatrix} .$$

Then

$$\det(A) = \det(D), \qquad \det(B) = (-1)^\ell \det(C) \, . \tag{6}$$

Of course this is obvious in two dimensions since

$$O = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

But to us, having grown up (mathematically speaking) with conformal mappings, this was obviously false in higher dimensions. First, we'd never seen anything like this before (and still haven't) and second there's the obvious consequence. Interchange as many rows and columns as you like, this mixes up A, B, C and D, yet leaves O orthogonal so the identity must still hold (up to sign). One literally gets lots of *linear* equations between minors of an orthogonal transformation.

Fred Gehring was the senior mathematician in residence at Mittag–Leffler then, and I remember discussing with Tadeusz for a bit about talking to Fred to ask advice about what we'd say to Donaldson and Sullivan to tell them they made a mistake (and it was in an Acta paper!). Fortunately caution prevailed and Tadeusz and I spent hours writing down even dimensional orthogonal matrices in four and six dimensions and calculating and finding the formulas working (miraculously). Then heading down the hill to bed about 4 or 5 am shaking our heads (Mittag is basically a castle on a hill, with the accommodation spread out below). The next day we were able to work out a proof. We asked around and looked everywhere, but no-one knew this crazy fact about orthogonal transformations.

At any rate, two things were clear. We had a lot of new identities for (the differentials of) conformal mappings, and that one could control the regularity of the differential if one could control the regularity of the minors. But these minors were polynomials of degree $\ell = n/2$ in the derivatives of f so we only needed to assume the derivatives of f were locally in $L^{n/2}$ is order to start working with them –as opposed to the usual assumption that they should be in L^n. Tadeusz knew these were key facts missing from the theory. From them we were able to proved the best possible form of Liouville's rigidity theorem in higher dimensions –roughly proving the higher dimensional equivalent of the Looman–Menchoff theorem: a $W^{1,1}$ solution to the Cauchy–Riemann equations is conformal.

We were then able to significantly refine and improve what Donaldson and Sullivan had done. We then set off in new directions and developed many new techniques. We found and gave explicit formulas for the operators used, identified the Beurling transform and showed they arose from singular integrals and identified the kernels –actually these are really nice tying in perfectly with the two-dimensional case. We jazzed up the arguments to the Dirac operators and more general mappings giving new and totally different proofs –that came with good estimates– on higher integrability. Tadeusz also knew that estimates below the ambient dimension were key to the analysis of singularities, and we had found them. Thus with a bit more work we were able to prove sharp versions (in even dimensions) of Painlevé's Theorem on removable singularities in higher dimensions as well. All this was quite novel and unexpected at the time –and certainly not foreshadowed in Donalson and Sullivan's paper.

There's one more thing I'd like to talk about that Tadeusz and I did during that time and after reading Donaldson and Sullivan's paper –it's a real "Eureka" moment– and was another thing quite independent of what they had done. It's one of those things that happens when you start a journey without a great idea of where you'll wind up.

As I said at the beginning, Tadeusz was a chain smoker. At Mittag at that time you were allowed to smoke in your office –provided you had one of the special offices at the very top of the building– basically in one of the Castle towers. Fortunately Tadeusz had such an office, and we could only work there since my office was nonsmoking. Anyway, you really could see the smoke coming out the window from the ground –so my wife knew where I was– at all hours (well, at least where Tadeusz was).

Figure 2: Quasi-dinner at Mittag–Leffler apartments. Clockwise: Fred Gehring (holding wine bottle); Dianne Brunton (my wife); Bruce Palka; Shanshuang Yang; Kari Hag; David Herron; Lois Gehring; me; Tadeusz Iwaniec (holding book).

By this time we had identified the higher dimensional Beurling transform and shown how to use it for the regularity theory of the PDEs associated with the geometric theory of mappings. We showed it's p-norms (the operator norm when it acts from L^p to L^p) precisely controlled the regularity of solutions to Beltrami systems both above and below the natural dimensions. We had two conjectures about these norms. First, that the p-norm of the higher dimensional operator was the same as that of the two dimensional operator (so dimension free) and that these norms were equal to $p-1$ for $p \geq 2$. So of course we hoped to solve the famous two-dimensional problem by solving the apparently harder higher dimensional problem. We'd written the Beurling transform as a product of Riesz transforms. But no-one even know what the p-norms of the Riesz transforms were. Finding the precise norms of integral operators is a tough business. Only the Hilbert transform in one real variable by S. K. Pichorides and the Fourier transform by Bill Beckner were known, even though there are many big thick books on the subject of singular integral operators. We found there was some work of Elias Stein in the 70's which showed the p-norms of the Riesz transforms were bounded independently of the dimension (but not necessarily dimension free of course).

We were up in Tadeusz' office trying to get an idea. He talked about the classical method of rotations and I suggested we could generalise it and get a good estimate. He of course challenged me with a piece of chalk and so I started writing on the blackboard. The first sketch was OK, but then Tadeusz homed in on the details. So we went through it correcting and improving until about 2am when we had what we thought was a pretty good estimate. We had a general theorem about lower bounds for the norms of these operators already worked out. In this case we had the lower bound $\cot(\pi/p)$, $p \geq 2$. The bound we just has was a multiple $\lambda \cot(\pi/p)$. Unfortunately λ came as a pretty complicated integral over a high dimensional sphere (well this attempt wound up that way). Of course such integrals are a real challenge for Tadeusz –probably from his math olympiad days– so he started working on it. An hour later he still hadn't figured it out, which meant to me it couldn't be done –only estimated. However, there is a great old library at Mittag–Leffler, so out went the cigarette and we went down to the Library, turned on all the lights (it was about 3am) and started looking through all the old math books from the nineteenth century –since obviously no new math book would have any integral formulas! About half and hour later we stumbled across an old book from about 1870 and low and behold there it was. Not only the calculation of λ but it said, in black and white, this complicated integral was equal to one. We looked at each other, both of our eyes went wide in surprise, a shiver went down my spine and an incredible feeling of elation. Because $\lambda = 1$ and because of our lower bound, we had identified the norms of the Riesz transforms, solving Stein's problem. It was a moment I will remember forever. That corner of the library at Mittag, and Tadeusz surprised face is etched on my mind.

All I've described above happened literally within the space of two and a half months –even writing and submitting the papers. It was lucky that there was a lot to do and see in Stockholm so my young family, which I simply didn't see over those few months, had things to do –I'd get up around 10 am and head straight up the hill, then come down about 4 am. I've been back to Mittag–Leffler a few times since (even to write a book with Kari Astala and Tadeusz) and it still is a magical place in my mind.

Figure 3: With Tadeusz Iwaniec, in New Zealand, around 2005.

Also, these are only my memories. Tadeusz might have a different recollection. After all, he has quite a good memory even if it is a bit short. He's also given up smoking and gained significant prestige from doing so. Finally, all I've talked about works in even dimensions [6]. The jury is out on the odd dimensional case, though there are plenty of partial results [8] and wonderful connections. But that's another story.

References

[1] Ahlfors, L.: *Lectures on quasiconformal mappings.* Van Nostrand, Princeton NJ. 1966.

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

[3] Astala, K., Iwaniec, T. and Martin, G. J.: *Elliptic partial differential equations and quasiconformal mappings in the plane.* Princeton Mathematical Series **48**. Princeton University Press, Princeton, NJ, 2009.

[4] Donaldson, S. K. and Sullivan, D. P.: Quasiconformal 4-manifolds. *Acta Math.* **163** (1989), 181–252.

[5] Gehring, F. W. and Martin, G. J.: Minimal co-volume hyperbolic lattices, I: The spherical points of a Kleinian group. *Ann. of Math. (2)* **170** (2009), 123–161.

[6] Iwaniec, T. and Martin, G. J.: Quasiregular mappings in even dimensions. *Acta Math.* **170** (1993), 29–81.

[7] Iwaniec, T. and Martin, G. J.: Riesz transforms and related singular integrals. *J. Reine Angew. Math.* **473** (1996), 25–57.

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

[9] Tukia, P. and Väisälä, J.: Lipschitz and quasiconformal approximation and extension. *Ann. Acad. Sci. Fenn. Ser. AI. Math.* **6** (1981), 303–342.

PERTTI MATTILA (Kuusankoski, Finland, 1948) obtained his PhD in mathematics at the University of Helsinki in 1973 (advisor: Jussi Väisälä). He is currently Professor at the University of Helsinki. His main research areas are geometric measure theory and harmonic analysis.

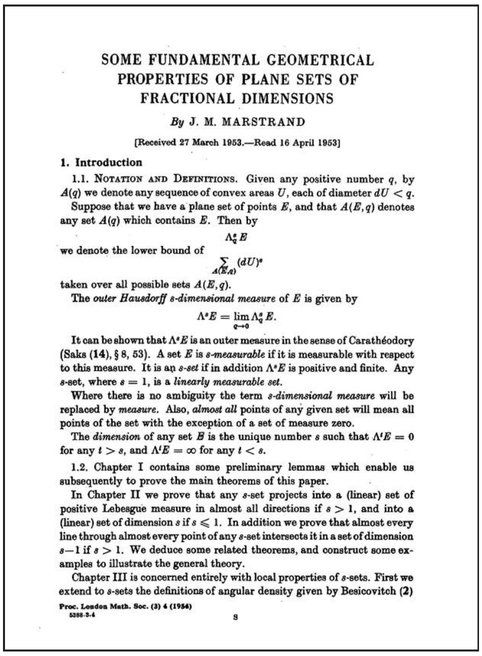
Marstrand's theorems

Pertti Mattila*

1. Introduction

I would like to thank the editors for inviting me to write a short essay describing a paper (or a couple of them) which, in one way or another, made a deep impact on my own mathematical career. I have chosen for this the following two papers of John Marstrand:

- J. M. Marstrand: Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc. (3)* **4** (1954), 257–302.
- J. M. Marstrand: The (ϕ, s) regular subsets of n space. *Trans. Amer. Math. Soc.* **113** (1964), 369–392.



SOME FUNDAMENTAL GEOMETRICAL PROPERTIES OF PLANE SETS OF FRACTIONAL DIMENSIONS

By J. M. MARSTRAND

[Received 27 March 1953.—Read 16 April 1953]

1. Introduction

1.1. Notation and Definitions. Given any positive number q, by $A(q)$ we denote any sequence of convex areas U, each of diameter $dU < q$.

Suppose that we have a plane set of points E, and that $A(E,q)$ denotes any set $A(q)$ which contains E. Then by

$$\Lambda_q^s E$$

we denote the lower bound of

$$\sum_{A(E,q)} (dU)^s$$

taken over all possible sets $A(E,q)$.

The *outer Hausdorff s-dimensional measure* of E is given by

$$\Lambda^s E = \lim_{q \to 0} \Lambda_q^s E.$$

It can be shown that $\Lambda^s E$ is an outer measure in the sense of Carathéodory (Saks (**14**), § 8, 53). A set E is *s-measurable* if it is measurable with respect to this measure. It is an *s-set* if in addition $\Lambda^s E$ is positive and finite. Any s-set, where $s = 1$, is a *linearly measurable set*.

Where there is no ambiguity the term *s-dimensional measure* will be replaced by *measure*. Also, *almost all* points of any given set will mean all points of the set with the exception of a set of measure zero.

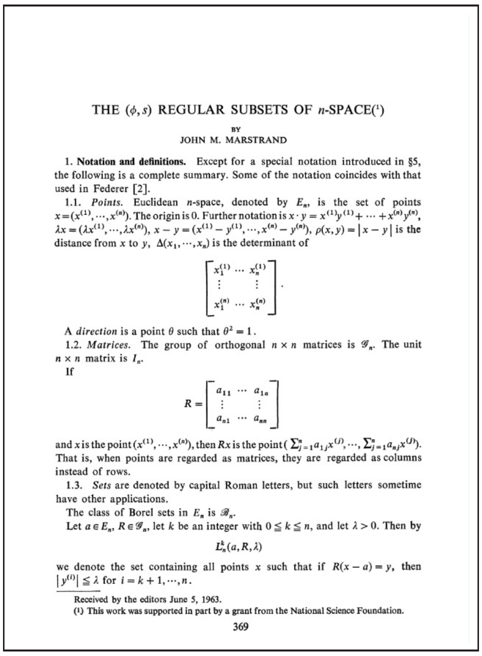
The *dimension* of any set E is the unique number s such that $\Lambda^t E = 0$ for any $t > s$, and $\Lambda^t E = \infty$ for any $t < s$.

1.2. Chapter I contains some preliminary lemmas which enable us subsequently to prove the main theorems of this paper.

In Chapter II we prove that any s-set projects into a (linear) set of positive Lebesgue measure in almost all directions if $s > 1$, and into a (linear) set of dimension s if $s \leqslant 1$. In addition we prove that almost every line through almost every point of any s-set intersects it in a set of dimension $s-1$ if $s > 1$. We deduce some related theorems, and construct some examples to illustrate the general theory.

Chapter III is concerned entirely with local properties of s-sets. First we extend to s-sets the definitions of angular density given by Besicovitch (**2**)

Proc. London Math. Soc. (3) 4 (1954)
5388.3.4 S



THE (ϕ, s) REGULAR SUBSETS OF n-SPACE(1)

BY
JOHN M. MARSTRAND

1. **Notation and definitions.** Except for a special notation introduced in §5, the following is a complete summary. Some of the notation coincides with that used in Federer [2].

1.1. *Points.* Euclidean n-space, denoted by E_n, is the set of points $x = (x^{(1)}, \cdots, x^{(n)})$. The origin is 0. Further notation is $x \cdot y = x^{(1)}y^{(1)} + \cdots + x^{(n)}y^{(n)}$, $\lambda x = (\lambda x^{(1)}, \cdots, \lambda x^{(n)})$, $x - y = (x^{(1)} - y^{(1)}, \cdots, x^{(n)} - y^{(n)})$, $\rho(x,y) = |x - y|$ is the distance from x to y, $\Delta(x_1, \cdots, x_n)$ is the determinant of

$$\begin{bmatrix} x_1^{(1)} & \cdots & x_n^{(1)} \\ \vdots & & \vdots \\ x_1^{(n)} & \cdots & x_n^{(n)} \end{bmatrix}.$$

A *direction* is a point θ such that $\theta^2 = 1$.

1.2. *Matrices.* The group of orthogonal $n \times n$ matrices is $\mathscr{G}_n$. The unit $n \times n$ matrix is I_n.

If

$$R = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

and x is the point $(x^{(1)}, \cdots, x^{(n)})$, then Rx is the point $(\sum_{j=1}^n a_{1j}x^{(j)}, \cdots, \sum_{j=1}^n a_{nj}x^{(j)})$. That is, when points are regarded as matrices, they are regarded as columns instead of rows.

1.3. *Sets* are denoted by capital Roman letters, but such letters sometime have other applications.

The class of Borel sets in E_n is $\mathscr{B}_n$.

Let $a \in E_n$, $R \in \mathscr{G}_n$, let k be an integer with $0 \leqq k \leqq n$, and let $\lambda > 0$. Then by

$$L_n^k(a, R, \lambda)$$

we denote the set containing all points x such that if $R(x - a) = y$, then $|y^{(i)}| \leqq \lambda$ for $i = k+1, \cdots, n$.

Received by the editors June 5, 1963.

(1) This work was supported in part by a grant from the National Science Foundation.

369

These papers have greatly influenced my work since the beginning of my career and continue to do so. They have also had a great impact on many other mathematicians. In the following, rather than describing Marstrand's papers in detail, I try to explain briefly why and how they have done that.

*pemattil@mappi.helsinki.fi.

There are many results that could be called Marstrand's theorem but the following two are probably most widely referred to as such:

Theorem 1 *Let A be a Borel subset of the plane $\mathbb{R}^2$ with Hausdorff dimension* $\dim A = s$. *Denote by p_θ the orthogonal projection onto the line* $\{t(\cos(\theta), \sin(\theta)) : t \in \mathbb{R}\}$.

1. *If $s \leq 1$, then* $\dim p_\theta(A) = s$ *for almost all $\theta \in [0, \pi)$.*
2. *If $s > 1$, then the Lebesgue measure $\mathcal{L}^1(p_\theta(A)) > 0$ for almost all $\theta \in [0, \pi)$.*

Theorem 2 *Let s be a positive number. If there is a non-trivial Borel measure μ on some $\mathbb{R}^n$ such that the positive and finite limit*

$$0 < \lim_{r \to 0} \frac{\mu(B(x, r))}{r^s} < \infty$$

exists for μ almost all $x \in \mathbb{R}^n$, then s must be an integer.

Theorem 1 was proved in the paper [4] and Theorem 2 in the paper [6]. Much more was done in these papers, but I shall mainly restrict to discuss these two theorems and their relatives.

In order not to make the reference list as long as the rest of the paper, I only give a few of them. Many others can be found in the books [1] and [7] and in the survey papers [8] and [9].

2. Marstrand's projection theorem

The paper [4] was the first work where geometric structure of general fractals in the plane was explored, about 20 years before Mandelbrot coined the term fractal and a wider interest in them started to develop. A general fractal here simply means a subset A of $\mathbb{R}^2$ which is measurable with respect to the s-dimensional Hausdorff measure $\mathcal{H}^s$ and has positive and finite Hausdorff s-measure: $0 < \mathcal{H}^s(A) < \infty$. Here s is any number with $0 < s < 2$. In the 1920's and 30's Besicovitch had in three papers laid the foundations of geometric measure theory by investigating in great detail the case $s = 1$, that is, subsets of the plane with positive and finite length. Marstrand did the same for general s. In addition to the projection properties, he derived fundamental results on line intersections and circular and conical density properties of such sets.

Follow-ups on Marstrand's paper came slowly; in those days Hausdorff dimension mainly appeared as a size estimate for various exceptional sets and it did not yet have such an independent role as it does today for example in dynamical systems. In 1968 Kaufman gave a new proof for Theorem 1 using a potential theoretic method for the part (1) and a Fourier analytic method for the part (2). Later on both of these techniques have played central roles in various generalizations and analogs of Marstrand's projection theorem, some of which we shall discuss later. Kaufman also showed that the exceptional set in (1) has Hausdorff dimension at most s. In 1975 I generalized Theorem 1 and Kaufman's exceptional set estimate to higher dimensions, for orthogonal projections into m-planes in $\mathbb{R}^n$. The proof was rather straightforward using Marstrand's ideas with Kaufman's potential theoretic method. In the same year we proved with Kaufman that the exceptional set estimates are sharp. In 1982 Falconer proved in general dimensions an exceptional set estimate corresponding to the case (2), in the plane the upper bound is $2-s$. Falconer used Fourier transform, no proof without it is known.

Here is a brief sketch how to prove (1) of Theorem 1. If $0 < t < s = \dim A \leq 1$, one can put a Borel probability measure μ on A such that the energy-integral

$$I_t(\mu) := \int\int |x-y|^{-t} d\mu x d\mu y < \infty$$

by classical results of Frostman. Letting μ_θ be the push-forward of μ under the projection p_θ; $\mu_\theta(B) = \mu(p_\theta^{-1}(B))$, one finds by Fubini's theorem that

$$\int_0^\pi I_t(\mu_\theta) d\theta = c(t) I_t(\mu),$$

where

$$c(t) = \int_0^\pi |p_\theta(0,1)|^{-t} d\theta < \infty,$$

since $t < 1$. Thus $I_t(\mu_\theta) < \infty$ for almost all θ. Another reference to Frostman yields $\dim p_\theta(A) \geq t$ for almost all θ, whence also $\dim p_\theta(A) \geq s$ for almost all θ. The opposite inequality is trivial.

Soon it started to become clear that Marstrand's projection theorem is not only a single theorem but a basic case of a general phenomenom: many parametrized families of mappings transform Hausdorff dimension in a similar way as orthogonal projections, as one might anticipate from the above proof sketch. This principle can be applied in various instances. One such instant is that of Bernoulli convolutions, that is, the random sums

$$\sum_j \pm\lambda^j,$$

where $0 < \lambda < 1$ is fixed. In 1995 Solomyak solved an old problem of Erdös by showing that the probability distribution of these sums is absolutely continuous for almost all $\lambda > 1/2$. A little later Peres and Solomyak gave a simpler proof interpreting this as a Marstrand type projection problem: defining

$$\Pi_\lambda : \{-1,1\}^{\mathbb{N}} \to \mathbb{R},\ \Pi_\lambda(\omega_j) = \sum_j \omega_j \lambda^j,$$

the probability distribution in question is the push-forward, that is, a kind of projection, of the uniformly distributed measure on $\{-1,1\}^{\mathbb{N}}$ and projection methods can be used, although highly non-trivially. Seemingly inspired by this Peres and Schlag developed in 2000 in [10] a setting of generalized projections and analyzed them with deep Fourier analytic methods. In particular, they estimated dimensions of various exceptional sets, sharpening the results on Bernoulli convolutions, orthogonal projections and many others.

Another case concerns invariant measures under geodesic flows. Ledrappier and Lindenstrauss proved in 2003 a Marstrand type result for such measures on tangent bundles of two-dimensional surfaces when projected onto the surface. Although there is only one projection, the problem can still be interpreted as a projection problem for a family of mappings. Later on this work has been continued by Ledrappier, E. and M. Järvenpää, Leikas and Hovila, for references, see [2].

Sum sets $A+B$ play important role in dynamical systems in many ways. For $A, B \subset \mathbb{R}$ the question whether $\mathcal{L}^1(A + tB) > 0$ for almost all $t \in \mathbb{R}$ is equivalent to $\mathcal{L}^1(p_\theta(A \times B)) > 0$ for almost all $\theta \in [0, \pi)$. Hence it is not surprising that Marstrand's projection theorem and its analogs have had many applications in dynamical systems by Palis, Yoccoz, Moreira, Lima, Peres, Shmerkin, Hochman and others, some references can be found in [3].

Also Falconer's distance set problem has a projection type flavour. It asks: for which $0 < s < n, n \geq 2$, is it true that for any Borel set $A \subset \mathbb{R}^n$, with $\dim A > s$ the distance set $D(A) = \{|x-y| : x, y \in A\}$ has positive Lebesgue measure? The projection type flavour becomes clearer when one looks at the pinned distance sets $D_x(A) = \{|x-y| : y \in A\}, x \in \mathbb{R}^n$. Then $D_x(A)$ is the image of A under the mapping $y \mapsto |x-y|$ and x serves as a parameter. Because of its connection to modern Fourier analysis the distance set problem has been studied by many people including Bourgain, Wolff, Erdogan, Katz, Tao, Iosevich, Sjölin and myself. The best known value of s guaranteeing that $\dim A > s$ implies $\mathcal{L}^1(D(A)) > 0$ is $n/2 + 1/3$ due to Wolff for $n = 2$ and to Erdogan for $n > 2$. The conjecture is that $s = n/2$ should suffice. The problem for pinned distance sets was studied by Peres and Schlag in their generalized projections setting in [10].

Other Marstrand type results and their applications include those of E. and M. Järvenpää on projections of SRB-measures for coupled map lattices, of Sauer and Yorke on dimension change under typical smooth mappings, of Hunt and Kaloshin on projections in infinite-dimensional spaces, of Olsen on multifractals, of Falconer, Howroyd, M. Järvenpää and myself on behaviour of box counting and packing dimensions under projections. Recently I have studied analogs of Marstrand's projection theorem in Heisenberg groups with Balogh, Durand Cartagena, Fässler and Tyson.

3. Marstrand's density theorem

The result of Theorem 2 was already proved by Besicovitch in the 1930's for subsets of the line and by Marstrand in [4] for subsets of the plane. The higher dimensional case was much trickier. It was preceded by the paper [5]. This is a third paper of Marstrand which had deep impact on my career, and it was the first of these three which I became acquainted with. In 1947 Federer had generalized Besicovitch's structure theory of 1-dimensional sets to general m-dimensional subsets of $\mathbb{R}^n$, m is now an integer, $0 < m < n$. The core of this theory is that any $\mathcal{H}^m$-measurable subset A of $\mathbb{R}^n$ with $0 < \mathcal{H}^m(A) < \infty$ can be decomposed as a union of a rectifiable and a purely unrectifiable part, and both of these parts can be characterized in four different ways:

1) existence (or non-existence for purely unrectifiable parts) of Lipschitz parametrizations from $\mathbb{R}^m$,
2) the density property
$$\lim_{r\to 0} \frac{\mathcal{H}^m(A \cap B(x,r))}{\mathcal{H}^m(B^m(0,1))r^m} = 1 \text{ for } \mathcal{H}^m \text{ almost all } x \in A,$$
3) almost everywhere existence of approximate tangent planes, and
4) integralgeometric properties.

Besicovitch had proved in the plane the equivalence of these four properties, Federer was able to generalize all others to higher dimensions except that 2) implies the other properties. Marstrand succeeded in doing this in [5] for the case $m = 2, n = 3$.

I pause here for a brief personal account. Federer's book *Geometric Measure Theory* had come out in 1969 and when I as a graduate student wanted go to US for a year, my supervisor Jussi Väisälä was far-sighted enough to see that this was an area for future. By his suggestion I spent the

Academic Year 1972-73 in Indiana University. There Bill Ziemer proposed to me the problem of generalizing Marstrand's result to general dimensions. This was a good problem for a beginning graduate student in the area. Marstrand's proof contained ingenuous ideas and it seemed quite possible that these with the addition of some technicalities could settle the general case. This turned out to be so and I wrote my second paper (my thesis was the first and on a different topic) on this. When I came back to Finland I stayed attracted to this type of geometric measure theory and continued to work on it. Later in the 80's and 90's I was also very fortunate to learn to know Marstrand personally during my two visits to Bristol and his two visits to Finland.

After proving that density one implies rectifiability the natural thing was to try to show that already the existence of the positive and finite limit

$$0 < \lim_{r \to 0} \frac{\mathcal{H}^m(A \cap B(x,r))}{r^m} < \infty \text{ for } \mathcal{H}^m \text{ almost all } x \in A$$

is enough to conclude rectifiability, a fact that Besicovitch had earlier established in the plane. In [6] Marstrand had studied this question. It can immediately be expressed as a question for general Borel measures μ: does the existence of

$$0 < \lim_{r \to 0} \frac{\mu(B(x,r))}{r^m} < \infty \text{ for } \mu \text{ almost all } x \in \mathbb{R}^n \tag{1}$$

imply that μ is m-rectifiable, which means that it lives on countably many m-dimensional Lipschitz surfaces? So Marstrand proved that if we start with an arbitrary positive number m, then (1) implies that m must be an integer. For integral m he also proved some partial results towards rectifiability but was not able to solve the problem. Neither was I. It turned out to be very difficult and was finally solved by Preiss in his remarkable paper [11] in 1987.

Preiss's method was based on tangent measures. They were already implicitly present in Marstrand's paper. The idea is to take normalized blow-up limits ν of a measure μ satisfying (1). They satisfy the much stronger restriction:

$$\nu(B(x,r)) = cr^m \text{ for } x \in \operatorname{spt} \nu. \tag{2}$$

Theorem 2 follows once one shows that such non-trivial uniformly distributed measures can exist only if m is an integer. This is essentially what Marstrand proved. Then the next task would be to show that for integral m they are just Lebesgue measures on m-planes, which would imply the rectifiability of μ by the results of Marstrand in [6]. In fact, this is true, as proven by Preiss, only for $m = 1, 2$ and for other values of m a more delicate analysis of uniformly distributed measures was performed by Preiss in order to complete the rectifiability proof, and by Kowalski and Preiss to characterize such uniformly distributed measures when $m = n-1$. For positive integers other than $1, 2, n-1$ no characterization is known. Another interesting problem concerns characterization of the larger class of measures ν for which

$$0 < \nu(B(x,r)) = \nu(B(y,r)) < \infty \text{ for } x, y \in spt\,\mu.$$

Kirchheim and Preiss solved this in the plane in 2002 and some partial results were proven earlier by Christensen in 1970.

A key how (2) is employed is that it implies identities like

$$\int g(|x_1 - y|) d\nu y = \int g(|x_2 - y|) d\nu y \text{ for } x_1, x_2 \in \operatorname{spt} \nu.$$

For Marstrand's density theorem 2 it is enough to use functions $g(y) = (r^2 - |x_j - y|^2)\chi_{B(x_j,r)}(y)$, that is, identities

$$\int_{B(x_1,r)} (r^2 - |x_1 - y|^2) d\nu y = \int_{B(x_2,r)} (r^2 - |x_2 - y|^2) d\nu y \quad \text{for } x_1, x_2 \in \operatorname{spt} \nu,$$

and the geometric information derived from them. For the rectifiability problem one needs also higher powers, such as

$$\int_{B(x_j,r)} (r^2 - |x_j - y|^2)^2 d\nu y.$$

To get the full Preiss theorem one has to use moments of all orders which are included in the expansions of exponential integrals

$$\int e^{-t|x-y|^2} d\nu y, \qquad t > 0.$$

It is natural to ask whether Theorem 2 holds for other norms in $\mathbb{R}^n$. In general this is open, but good partial results were proven by Lorent in 2007. In 2003 he also proved partial results for the rectifiability problem with the norm $\|x\| = \max_i |x_i|$, for which the balls are cubes, but the general result is open even for 2-dimensional measures in $\mathbb{R}^3$.

There are also intriguing questions in general metric spaces. Some easy examples show that both Marstrand's density theorem 2 and Preiss's theorem are false in general metric spaces, but in which metric spaces do they hold? No examples of metric spaces are known where 'density equals 1 implies rectifiability' fails. Neither are there examples of metric spaces where

Besicovitch's 1/2-problem for 1-dimensional sets would have a negative solution, this problem is open also for subsets of the plane. Good partial results on it in metric spaces were obtained by Preiss and Tiser in 1992.

Inspired by the above type of results I investigated the corresponding rectifiability problem for the principal values of the Cauchy singular integral with a hope that it could be applied to problems on analytic functions. I proved in 1990 that if μ is a finite Borel measure in the complex plane that satisfies for μ almost all $z \in \mathbb{C}$,

$$0 < \liminf_{r\to 0} \frac{\mu(B(z,r))}{r} \le \limsup_{r\to 0} \frac{\mu(B(z,r))}{r} < \infty$$

and

$$\exists \lim_{\epsilon\to 0} \int_{\mathbb{C}\setminus B(z,\epsilon)} \frac{1}{w-z} d\mu w \in \mathbb{C},$$

then μ is rectifiable. Later we proved with Preiss that if μ is a finite Borel measure in $\mathbb{R}^n$ that satisfies for μ almost all $x \in \mathbb{R}^n$,

$$0 < \liminf_{r\to 0} \frac{\mu(B(x,r))}{r^m} \le \limsup_{r\to 0} \frac{\mu(B(x,r))}{r^m} < \infty$$

and

$$\exists \lim_{\epsilon\to 0} \int_{\mathbb{R}^n\setminus B(x,\epsilon)} \frac{y-x}{|y-x|^{m+1}} d\mu y \in \mathbb{R}^n,$$

then m must be an integer and μ is rectifiable. The density condition can be weakened to

$$0 < \limsup_{r\to 0} \frac{\mu(B(x,r))}{r^m} < \infty$$

for μ almost all $x \in \mathbb{R}^n$ by the results of Tolsa in 2008 and Ruiz de Villa and Tolsa in 2010. Vihtilä proved in 1996 for Ahlfors m-regular measures the necessity of m being an integer starting from the L^2-boundedness of the related singular integral operator, the corresponding rectifiability problem is open except for $m = 1$. Recently I proved analogous results with Chousionis in Heisenberg groups.

The hope of applications to analytic functions was never realized and the problems have been solved by other methods by Melnikov, Verdera, David, Tolsa and others. But there is still a chance that the higher dimensional results could be applied to harmonic functions. A survey on this topic is given in [9].

Applications of Theorem 2 to Gauss-Weierstrass and Poisson integrals were given by Watson in 1994 and to Ginzburg–Landau type equations by Wang in 2002.

References

[1] FALCONER, K. J.: *Geometry of Fractal Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1985.

[2] HOVILA, R., JÄRVENPÄÄ, E., JÄRVENPÄÄ, M. AND LEDRAPPIER, L.: Besicovitch-Federer projection theorem and geodesic flows on Riemann surfaces. Available at `arXiv:1104.3453`.

[3] LIMA, Y. AND MOREIRA, C. G.: Yet another proof of Marstrand's Theorem, `arXiv:1101.3252`. To appear in *Bull. Brazilian Math. Society.*

[4] MARSTRAND, J. M.: Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc. (3)* **4** (1954), 257–302.

[5] MARSTRAND, J. M.: Hausdorff two-dimensional measure in 3 space. *Proc. London Math. Soc. (3)* **11** (1961), 91–108.

[6] MARSTRAND, J. M.: The (ϕ, s) regular subsets of n space. *Trans. Amer. Math. Soc.* **113** (1964), 369–392.

[7] MATTILA, P.: *Geometry of Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1995.

[8] MATTILA, P.: Hausdorff dimension, projections, and the Fourier transform. *Publ. Mat.* **48** (2004), 3–48.

[9] MATTILA, P.: Removability, singular integrals and rectifiability. *Rev. Roumaine Math. Pures Appl.* **54** (2009), 483–491.

[10] PERES, Y. AND SCHLAG, W.: Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions. *Duke Math. J.* **102** (2000), 193–251.

[11] PREISS, D.: Geometry of measures in $\mathbb{R}^n$: distribution, rectifiability, and densities. *Ann. of Math.* **125** (1987), 537–643.

Yves Meyer (Paris, France, 1939) obtained his PhD in mathematics at Université de Strasbourg in 1966 (advisor: Jean-Pierre Kahane). He currently holds the position of Professor Emeritus at the École Normale Supérieure de Cachan. His main research areas are harmonic analysis, number theory, singular integral operators and signal and image processing.

Two papers by Alberto P. Calderón

Yves Meyer*

Two papers by Alberto P. Calderón have been pivotal to my research work. They have been published in the Proceedings of the National Academy of Sciences which is unusual in mathematics. These two papers are extremely short notes (eight pages for [1] and four pages for [2]). The most outstanding features in Calderón's achievements are elegance, concision, profoundness and a vision of the future of mathematics. As much as real analysis, complex analysis, and operator theory are concerned, these two papers changed everything. Twelve elegant pages by Calderón gave rise to an intellectual revolution.

In the early sixties mathematicians were using a clumsy pseudo-differential calculus in which it was forbidden to multiply by functions which are not $\mathcal{C}^\infty$. Calderón wanted to overcome this limitation. His work was motivated by three problems:

(a) In 1956 E. de Giorgi proved that every solution of a scalar elliptic equation of second order in divergence form with bounded coefficients is Hölder continuous. De Giorgi's theorem is the crucial step to solve Hilbert's nineteenth problem, which consists in showing that a function which minimizes a convex integral functional of the calculus of variations is analytic if the functional is analytic. Calderón wanted to recover De Giorgi's theorem by using an improved pseudo-differential calculus in which the smoothness assumptions on the coefficients are minimal.

(b) The second motivation of Calderón came from nonlinear partial differential equations. Calderón was prophetic when he wrote:

> The aim of this greater generality is to obtain stronger estimates and to prepare the ground for applications to the theory of quasilinear and nonlinear differential equations.

(c) The third issue Calderón had in mind is the solution of Dirichlet or Neumann problem in Lipschitz domains by the method of the double layer potential. This approach leads to operators defined as singular integrals [3, 4]. Such operators are not amenable to the standard pseudo-differential calculus.

This is why in 1965 Calderón elaborated a new algebra of pseudo-differential operators containing pointwise multiplications by Lipschitz functions.

*Yves.Meyer@cmla.ens-cachan.fr.

The key tool in this new calculus is a deep theorem which is proved in [1] and will be discussed now:

Theorem 1 *Let T be a pseudo-differential operator commuting with translations. Let us assume that the symbol $\tau(\xi)$ of T is homogeneous of degree 1 and is $\mathcal{C}^\infty$ on the unit sphere. Let A be the operator of pointwise multiplication with a Lipschitz function $A(x)$. Then the commutator $[T, A]$ is bounded on $L^2(\mathbb{R}^n)$.*

Calderón was entirely right when he predicted that Theorem 1 will be pivotal to nonlinear partial differential equations. In a brilliant series of papers Tosio Kato and Gustavo Ponce have been using similar commutator estimates in nonlinear partial differential equations. One should also mention the para-differential operators by Jean-Michel Bony and the refined version of the div-curl lemma by Pierre-Louis Lions [4]. Let us quote Michael Taylor [6, 7]:

> The work of Kato and G. Ponce in 1988 on the Navier–Stokes equations produced the Kato–Ponce estimate, a commutator estimate that can be viewed as a microlocalized Moser estimate. This result can be analyzed from the point of view of paradifferential operators, introduced by J.-M. Bony as a tool for nonlinear analysis. This is a connection I found particularly intriguing, and Kato and I corresponded about related issues in paradifferential operator calculus as recently as 1996.

Theorem 1 is an obvious consequence of the $T(1)$ theorem by David and Journé [3], [4]. Today everything reduces to checking that $[T, A](1) = T(A) \in \mathrm{BMO}$ which is obviously true since pseudo-differential operators of order 0 map L^∞ to BMO. But the beauty of the proof given by Calderón remains intact. Instead of trying to prove Theorem 1 in its full generality, Calderón focused on a toy example which looked much simpler but happened to be the magic key opening all doors. The toy example is the "first commutator" $[\Lambda, A]$. Here we are in one dimension, $A(x)$ is a Lipschitz function of the real variable x. Therefore $a = a(x) = \frac{d}{dx}A(x)$ belongs to $L^\infty(\mathbb{R})$. The operator of pointwise multiplication by $A(x)$ is denoted by A. We set $D = \frac{d}{dx}$, $\mathcal{H}$ is the Hilbert transform and finally $\Lambda = D\mathcal{H}$ is the simplest pseudo-differential operator of order 1. The symbol of Λ is $\tau(\xi) = |\xi|$ which is homogeneous of degree 1. The commutator $[\Lambda, A]$ is defined by $[\Lambda, A](f) = \Lambda(Af) - A\Lambda(f) = C(a, f)$ where f is any testing function. This commutator can also be computed as a singular integral operator

$$C(a, f) = \frac{1}{\pi} p.v. \int \frac{A(x) - A(y)}{(x-y)^2} f(y)\, dy\,. \tag{1}$$

Then $C(a,f)$ can be studied as a bilinear operator, or as an operator depending on a parameter a and acting on a function f, or as a singular integral operator. This observation applies as well to the commutator $[T,A]$ in Theorem 1 and explains why the $T(1)$ theorem can be applied. These three options have been used by Calderón [5].

Let us write $\tilde{a} = \mathcal{H}(a)$. If Leibniz rule applied to Λ we would have $\|C(a,f)\|_2 \leq \|\tilde{a}f\|_2$ which is not true. Instead Calderón proved the following estimate in [1]

$$\|C(a,f)\|_2 \leq C\|a\|_\infty \|f\|_2 \,. \tag{2}$$

The operator norm of $[\Lambda, A]$ is larger than $\|a\|_\infty$ and (2) is sharp.

The proof of (2) given by Calderón is beautiful and unexpected. It relies on complex analysis. The complex Hardy space $H^p(\mathbb{R})$, $1 \leq p < \infty$, consists of all holomorphic functions $F(z)$ in the upper half-plane such that $\sup_{y>0} \int |F(x+iy)|^p \, dx = \|F\|_p^p < \infty$. Such a function has a trace on the real axis given by $F(x) = \lim_{y\to 0} F(x+iy)$. This limit exists a.e. and in L^p and the L^p norm of this trace is the H^p norm of F. The Lusin area integral $S(F)$ of F is the function of the real variable x defined by $S(F)(x) = \left(\int\int_{v\geq|u-x|} |F'(z)|^2 \, du\, dv\right)^{1/2}$ where $z = u+iv$ and F' is the derivative of F. The equivalence between the the L^p norms of $S(F)$ and of F is standard when $1 < p < \infty$. In [1] Calderón proved this equivalence when $p = 1$:

Theorem 2 *There exist two constants C_1 and C_2 such that for $F \in H^1(\mathbb{R})$*

$$C_1\|F\|_1 \leq \|S(F)\|_1 \leq C_2\|F\|_1 \,. \tag{3}$$

One way in this theorem is almost obvious. If $F \in H^1(\mathbb{R})$, then $F = GH$ where G, H belong to the Hardy space $H^2(\mathbb{R})$. Then $F' = G'H + GH'$ and the proof of the second inequality in (3) reduces to well known facts on $H^2(\mathbb{R})$. But the proof of the first inequality in (3) is much deeper as it can be seen in [1].

For proving (2) Calderón defined a "left paraproduct" between two functions G and H in $H^2(\mathbb{R})$ by $F' = G'H$ and Theorem 2 implies $\|F\|_1 \leq C\|G\|_2\|H\|_2$. One also defines a "right paraproduct" by $F' = GH'$ and the product GH is the sum between these two paraproducts as in Bony's work. Finally Calderón showed that the bilinear operator $C(a,f)$ is a transposed version of these paraproducts.

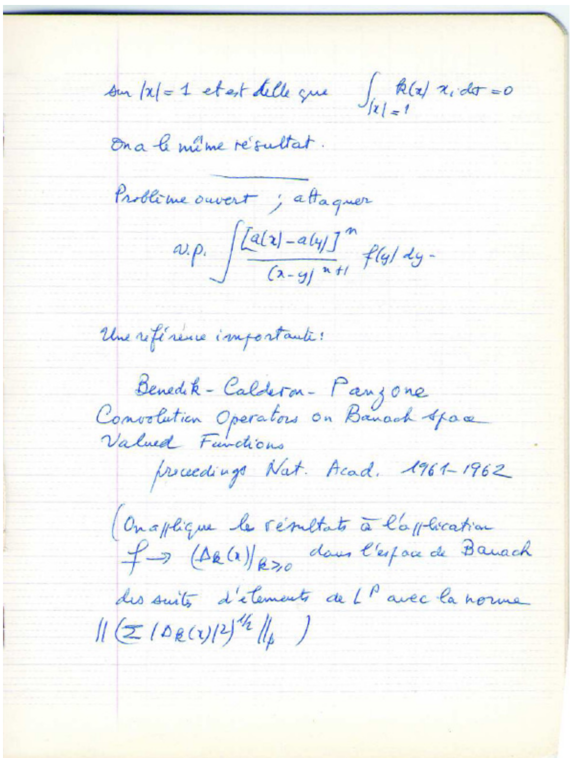
Using the "method of rotations" he created with Antoni Zygmund, Calderón showed that the boundedness of the first commutator implies Theorem 1 in full generality.

Calderón's achievements paved the way to the modern theory of Hardy spaces. The Hardy space $\mathcal{H}^1(\mathbb{R}^n)$ was defined by Elias Stein and Guido

Weiss as the Banach space consisting of all functions $f \in L^1(\mathbb{R}^n)$ such that the n Riesz transforms $R_j(f)$, $1 \leq j \leq n$, belong to $L^1(\mathbb{R}^n)$. Then Stein extended Theorem 2 to $\mathcal{H}^1(\mathbb{R}^n)$ and proved that Marcinkiewicz multipliers map $\mathcal{H}^1(\mathbb{R}^n)$ into itself. Charles Fefferman proved that BMO is the dual space of $\mathcal{H}^1(\mathbb{R}^n)$. These were the two problems I addressed without much success in my Ph.D. The modern theory of Hardy spaces culminated in the discovery of atomic decompositions by Coifman and Weiss [3].

My first encounter with Calderón took place forty-five years ago. At that time I was writing my Ph.D. On January 26th, 1966, I attended a lecture given by Calderón at Institut Henri Poincaré in Paris. Calderón lectured on Theorem 1. I do not remember if Calderón unveiled his research program at the end of his talk or if he offered me this precious gift during a discussion we had afterwards.

But I am certain that Calderón raised the issue of the continuity of the higher order commutators. I took notes of everything on a small copybook. These higher order commutators were much too deep, much too difficult to me. They were out of reach since I was unable to improve on the spectacular complex variable methods used by Calderón to prove the boundedness of the first commutator. I had to do something else. I fled. I moved to number theory. From 1967 to 1972 I worked on Pisot and Salem numbers. This detour was a very pleasant journey. But my fate was to become Calderón's disciple. The day came when I had to answer Calderón's call. This happened in 1974 when Raphy Coifman convinced me to attack Calderón's higher commutators.


sur $|x| = 1$ et est telle que $\int_{|x|=1} k(x)\, x_i\, d\sigma = 0$

On a le même résultat.

Problème ouvert ; attaquer

$$\text{v.p.} \int \frac{[a(x) - a(y)]^n}{(x-y)^{n+1}} f(y)\, dy.$$

Une référence importante :

Benedek - Calderon - Panzone
Convolution Operators on Banach space Valued Functions
proceedings Nat. Acad. 1961-1962

(On applique le résultat à l'application $f \to (\Delta_k(x))_{k \geq 0}$ dans l'espace de Banach des suites d'éléments de L^p avec la norme $\|(\Sigma |\Delta_k(x)|^2)^{1/2}\|_p$)

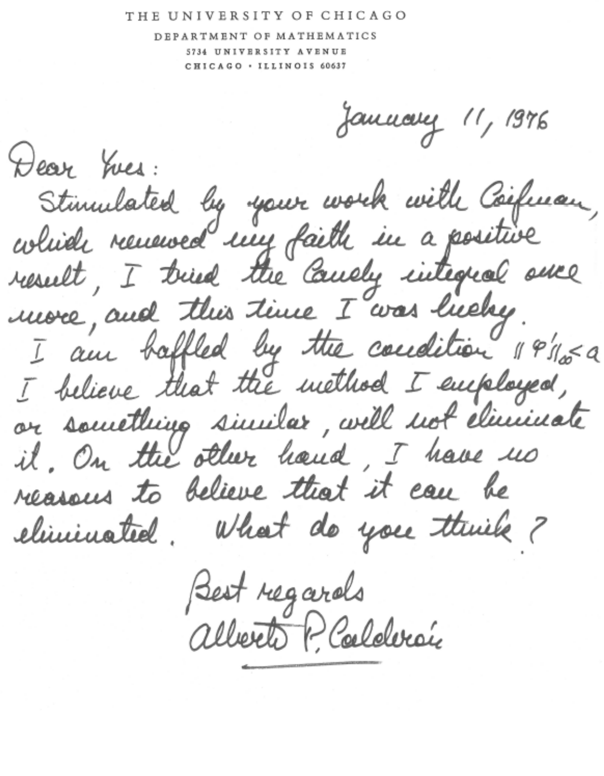
I have kept my precious copybook with an extreme care. Calderón is still alive there.

Higher order commutators are defined as follows. We start again with a Lipschitz function $A(x)$ of the real variable x and we denote by A the operator of pointwise multiplication by $A(x)$ as we did above. The simplest pseudo-differential operator of order m is $\Lambda_m = D^m\mathcal{H}$. The higher order commutators are then defined as $m!\,\Gamma_m = [A, [A, \ldots, [A, \Lambda_m]$. The kernel of Γ_m is

$$K_m(x, y) = \frac{(A(x) - A(y))^m}{(x - y)^{m+1}}.$$

I began to work on Calderón's program during my first visit to Washington University at Saint Louis, Missouri, in 1974. I was invited by Guido Weiss. The first day I was there, Raphy Coifman entered my office and said he was expecting my visit to attack Calderón's commutators. I accepted and we proved that the second commutator Γ_2 is bounded on L^2. Our proof did not rely on complex analysis. Instead we considered all trilinear operators $T(a, b, f)$ commuting with simultaneous translations and dilations on a, b, and f. We wanted to prove trilinear estimates. For that purpose we defined the trilinear symbol of the operator $T(a, b, f)$ and our strategy was to break this symbol into a series of building blocks. For each piece we could prove a basic estimate and the problem was to pile up these bounds. Moving from Γ_2 to Γ_3 took us a year so that we thought we would never finish. We were so awkward. Today the $T(1)$ theorem by David and Journé provides for free the required estimates for all higher order commutators. What took us years takes a second of thought. Raphy told me that these years of intense work on Calderón's program have been the most exciting in his mathematical life.

A short letter by Calderón came as a thunderbolt.


THE UNIVERSITY OF CHICAGO
DEPARTMENT OF MATHEMATICS
5734 UNIVERSITY AVENUE
CHICAGO · ILLINOIS 60637

January 11, 1976

Dear Yves:

Stimulated by your work with Coifman, which renewed my faith in a positive result, I tried the Cauchy integral once more, and this time I was lucky.

I am baffled by the condition $\|\varphi'\|_\infty < a$. I believe that the method I employed, or something similar, will not eliminate it. On the other hand, I have no reasons to believe that it can be eliminated. What do you think?

Best regards

Alberto P. Calderón

January 11, 1976

Dear Yves,

Stimulated by your work with Coifman, which renewed my faith in a positive result, I tried the Cauchy integral once more, and this time I was lucky. I am baffled by the condition $\|\phi'\|_\infty < a$. I believe that the method I employed, or something similar, will not eliminate it. On the other hand, I have no reasons to believe that it can be eliminated. What do you think?

Best regards, Alberto P. Calderón.

This letter announced [2]. What Coifman and I had been doing for years was no longer needed.

Let me say a few words about the Cauchy integral in Calderón's letter and explain its connection with the commutators. Here is the story. Let Γ be a closed Jordan curve in the complex plane $\mathbb{C}$ and let Ω_1, Ω_2 be the two components of the complement of Γ, Ω_1 being bounded. Let us assume that Γ is rectifiable and let s denote the corresponding arc length. The Hilbert space $L^2(\Gamma, ds)$ consists of functions f which are defined on Γ and are square integrable with respect to the arc length s. The Hardy space $H^2(\Omega_1)$ is the closure in $L^2(\Gamma, ds)$ of the vector space of polynomials $P(z)$ in the complex variable z, while $H^2(\Omega_2)$ is the closure in $L^2(\Gamma, ds)$ of the polynomial in $1/z$ which vanish at infinity.

Calderón wanted to know whether

$$L^2(\Gamma, ds) = H^2(\Omega_1) + H^2(\Omega_2)\,, \tag{4}$$

where the sum is direct, but not orthogonal in general.

A similar problem can be asked when Γ is an open Jordan curve but the Hardy spaces have a slightly distinct definition. The Hardy space $H^2(\Omega_1)$ is the closure in $L^2(\Gamma, ds)$ of all rational functions $P(z)/Q(z)$ which vanish at infinity and have no poles on the closure of Ω_1. The space $H^2(\Omega_2)$ is defined similarly.

When Γ is a straight line, (4) is obviously true. The sum is orthogonal and the projection operators are related to the Hilbert transform by

$$P_\pm = \frac{I \pm i\mathcal{H}}{2}\,.$$

This identity extends to the general case, the Hilbert transform $\mathcal{H}$ being replaced by the Cauchy integral $\mathcal{C}_\Gamma$. The Cauchy integral is defined by the singular kernel

$$p.v.\ \frac{1}{\pi}\,\frac{1}{z(s) - z(t)}\,,$$

where $z(s)$ denotes the arc length parametrization of Γ. Finally (4) is equivalent to the boundedness of $\mathcal{C}_\Gamma$ on $L^2(\Gamma, ds)$. This boundedness is the problem Calderón was mentioning in his letter.

Let us now assume that $A(\cdot)$ is a real valued Lipschitz function. Let Γ be the graph of A. We then write $\mathcal{C}_A$ instead of $\mathcal{C}_\Gamma$. If the harmless factor $1 + ia'(y)$ is ignored, we have

$$\mathcal{C}_A f(x) = p.v. \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(y)}{x - y + i(A(x) - A(y))} \, dy \, . \tag{5}$$

The boundedness of the Cauchy integral for Lipschitz curves reads

$$\|\mathcal{C}_A f\|_2 \leq C(A) \|f\|_2 \, , \tag{6}$$

where $C(A)$ only depends on the Lipschitz norm $\|\frac{d}{dx} A\|_\infty$.

We now arrive at a crossroads and several attacks to (6) have been proposed. The one Coifman and I used was mocked by Lennart Carleson as being "the pedestrian way". It relies on the Taylor expansion of $\mathcal{C}_A$ into a series of Calderón's commutators

$$\mathcal{C}_A = \sum_0^\infty (-i)^m \Gamma_m \tag{7}$$

Calderón did not use this "pedestrian way" when he proved the following theorem in [2]

Theorem 3 *There exists a constant $\eta_0 > 0$ such that for every Lipschitz function A satisfying $\|\frac{d}{dx} A\|_\infty < \eta_0$ the operator $\mathcal{C}_A$ is continuous on L^2.*

In Calderón's letter, this η_0 is denoted by a and the Lipschitz function $A(x)$ is denoted by ϕ. Calderón was wondering whether $\|\frac{d}{dx} A\|_\infty < \eta_0$ is actually needed. Theorem 3 immediately implies that a constant C exists such that

$$\|\Gamma_m(f)\|_2 \leq C^m \, \|a\|_\infty^m \, \|f\|_2 \, , \tag{8}$$

where

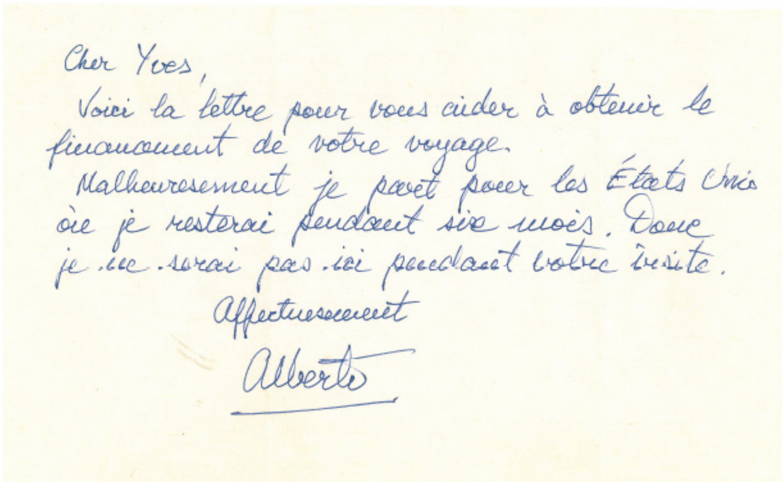
$$a(x) = \frac{d}{dx} A.$$

The operator norm of Γ_m cannot be smaller than $\|a\|_\infty^m$ which suggests that the optimal C might be 1. It is indeed the case, up to a polynomial factor [3].

Theorem 3 implies (4) whenever Γ is a closed Jordan curve of class $\mathcal{C}^1$ or is a Lipschitz graph with a small Lipschitz constant. Calderón's extraordinary proof relies on a perturbation argument and stresses the analytic dependence on A of the Cauchy operator $\mathcal{C}_A$. Both Carlos Kenig's work on generalized Hardy spaces and some subtle estimates on the conformal mapping are playing an important role in the argument. Once more Calderón used complex methods with an incredible virtuosity and once more I knew I could not compete. But I did not give up. Instead of complex methods I used a new magic trick provided by Alan McIntosh and coming from a world which had been mostly ignored by harmonic analysis people. This new world was familiar to those mathematical physicists who were opening new avenues in operator theory and quantum mechanics. Alan discovered that a very natural conjecture raised by Tosio Kato implies Theorem 3. This conjecture says that the domain of the square root of a maximal accretive operator coincides with the domain of the sesquilinear form defining this operator [4]. How a conjecture which seems so abstract could be connected with the Cauchy kernel on Lipschitz curves? This is the magic of Calderón's program. This new perspective reshaped everything. The commutators Γ_m were given a new decomposition into building blocks. I could prove (8) with C^m replaced by $(1+m)^6$ which implies "the full theorem", *i.e.*, the boundedness of the Cauchy kernel on every Lipschitz curve. But this detour by mathematical physics was not needed. Guy David built new "real variable methods" and deduced "the full theorem" from Theorem 3. If the fundamental result proved by Calderón in [2] is nurtured with Guy David's version of the "good lambda inequalities" the boundedness of the Cauchy kernel for all Lipschitz curves is obtained for free [3].

The story of the boundedness of the Cauchy integral did not stop there. Indeed in 1995, M. Melnikov and J. Verdera found an extraordinary proof. The starting point is a geometric identity due to Karl Menger and rediscovered by M. Melnikov. Karl Menger (1902–1985) was living in Chicago in these times but his work was not given the attention it deserved. M. Melnikov and J. Verdera cleverly used the Menger curvature and gave us the simplest and the most beautiful proof of the L^2 boundedness of the Cauchy integral on Lipschitz curves. Combining this new approach with some subtle variations on the $T(b)$-theorem, Guy David proved the Vitushkin conjecture which is a special case of Painlevé's problem on analytic capacity and Xavier Tolsa solved Painlevé's conjecture [4].

I visited Eckhart Hall every year from 1975 to 1987. I enjoyed discussing with Calderón. He treated me with irony and tenderness. Then a much deeper relation took place and I wished I could visit Buenos Aires with him. This never happened. In July 1989 Calderón sent me the following letter (in French) together with a formal letter of invitation to the Instituto Argentino de Matemática:


Cher Yves,
Voici la lettre pour vous aider à obtenir le financement de votre voyage.
Malheureusement je part pour les États Unis où je resterai pendant six mois. Donc je ne serai pas ici pendant votre visite.
Affectueusement
Alberto

Cher Yves,

Voici la lettre pour vous aider à obtenir le financement de votre voyage. Malheureusement je pars pour les États-Unis où je resterai pendant six mois. Donc je ne serai pas ici pendant votre visite.

Affectueusement, Alberto

References

[1] Calderón, A. P.: Commutators of singular integral operators. *Proc. Nat. Acad. Sci.* **53** (1965), 1092–1099.

[2] Calderón, A. P.: Cauchy integrals on Lipschitz curves and related operators. *Proc. Nat. Acad. Sci.* **74** (1977), 1324–1327.

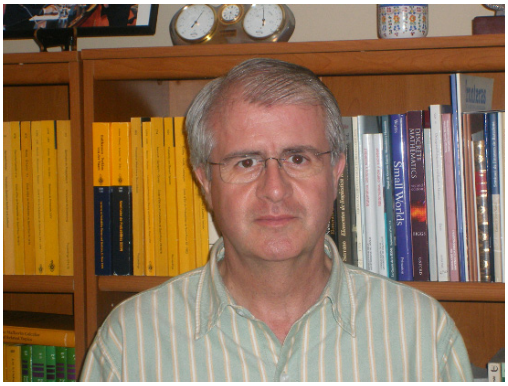
[3] Coifman, R. R. and Meyer, Y.: *Wavelets and Operators*. Cambridge Studies in Advanced Mathematics **48**, 1997.

[4] Meyer, Y.: *Complex analysis and operator theory in Alberto Calderón's work*. Selected Papers of Alberto P. Calderón with commentary edited by Alexandra Bellow, Carlos E. Kenig and Paul Malliavin. American Mathematical Society, 2008.

[5] Stein, E. M.: Singular Integrals: The Roles of Calderón and Zygmund. *Notices Amer. Math. Soc.* **45** (1998), 1130–1140.

[6] Taylor, M.: Tosio Kato. *Notices Amer. Math. Soc.* **47** (2000), 650–657.

[7] Taylor, M.: Commutator estimates. *Proc. Amer. Math. Soc.* **131** (2002), 1501–1507.

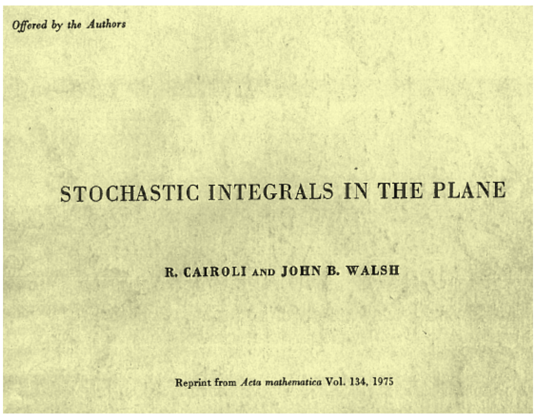


David Nualart (Barcelona, Spain, 1951) obtained his PhD in mathematics at the University of Barcelona in 1975 (advisor: Francesc Sales). He currently is Full Professor at Kansas University. His main research areas are stochastic processes and PDEs, fractional brownian motion and mathematical finance.

Stochastic integrals in the plane

David Nualart*

One of the papers that has very much influenced the first steps of my mathematical career in stochastic analysis was the article on "Stochastic integrals in the plane" by Renzo Cairoli and John B. Walsh, published in *Acta Mathematica* in 1975 (see [1]). This is a very long paper, and I still keep the original reprint offered by the authors. I came across this paper during my postdoctoral stay at the "Laboratoire d'Automatique et Analyse des Systèmes", in Toulouse, in 1976, in occasion of a seminar talk given by Eugene Wong from Berkeley on multiparameter processes. At that time, being at the beginning of my career, I was interested in stochastic analysis, but I still had not found a suitable research direction. Reading this paper was a discovery for me, and I found many sources of interesting open problems and new leads to follow.

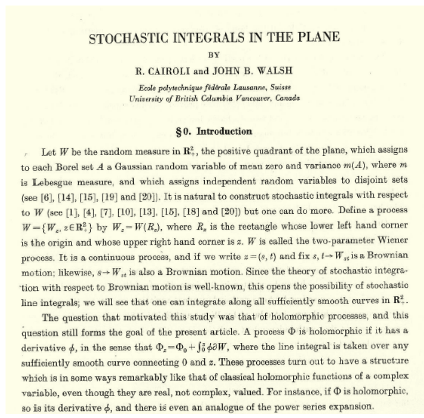


Offered by the Authors

STOCHASTIC INTEGRALS IN THE PLANE

R. CAIROLI AND JOHN B. WALSH

Reprint from *Acta mathematica* Vol. 134, 1975



STOCHASTIC INTEGRALS IN THE PLANE

BY

R. CAIROLI and JOHN B. WALSH

Ecole polytechnique fédérale Lausanne, Suisse
University of British Columbia Vancouver, Canada

§0. Introduction

Let W be the random measure in $\mathbf{R}^2_+$, the positive quadrant of the plane, which assigns to each Borel set A a Gaussian random variable of mean zero and variance $m(A)$, where m is Lebesgue measure, and which assigns independent random variables to disjoint sets (see [6], [14], [15], [19] and [20]). It is natural to construct stochastic integrals with respect to W (see [1], [4], [7], [10], [13], [15], [18] and [20]) but one can do more. Define a process $W=\{W_z, z\in\mathbf{R}^2_+\}$ by $W_z=W(R_z)$, where R_z is the rectangle whose lower left hand corner is the origin and whose upper right hand corner is z. W is called the two-parameter Wiener process. It is a continuous process, and if we write $z=(s,t)$ and fix s, $t\to W_{st}$ is a Brownian motion; likewise, $s\to W_{st}$ is also a Brownian motion. Since the theory of stochastic integration with respect to Brownian motion is well-known, this opens the possibility of stochastic line integrals; we will see that one can integrate along all sufficiently smooth curves in $\mathbf{R}^2_+$.

The question that motivated this study was that of holomorphic processes, and this question still forms the goal of the present article. A process Φ is holomorphic if it has a derivative ϕ, in the sense that $\Phi_z=\Phi_0+\int_0^z\phi\,\partial W$, where the line integral is taken over any sufficiently smooth curve connecting 0 and z. These processes turn out to have a structure which is in some ways remarkably like that of classical holomorphic functions of a complex variable, even though they are real, not complex, valued. For instance, if Φ is holomorphic, so is its derivative ϕ, and there is even an analogue of the power series expansion.

Figure 1: First pages of the old reprint offered by the authors with its yellowing pages that I still keep in my files.

The paper [1] is considered a fundamental work on the theory of two-parameter processes. This theory deals with stochastic processes

$$\{X_{s,t}, (s,t)\in\mathbb{R}^2_+\}$$

which depend on two parameters, instead of the usual time parameter. During the 70's, and starting from the pioneering work by Cairoli and Walsh, this field was developed and got the attention of leading probabilists like Paul

*nualart@math.ku.edu.

André Meyer. An important landmark was the conference on two-parameter processes that took place in Paris in 1980, whose proceedings were published in the volume 863 of the Lecture Notes in Mathematics.

A basic ingredient in this theory is the two-parameter Brownian motion also called Brownian sheet. This is a two-parameter process

$$\{W_{s,t}, (s,t) \in \mathbb{R}_+^2\}$$

defined in a probability space $(\Omega, \mathcal{F}, P)$, which is Gaussian, with zero mean and covariance function given by

$$E(W_{s_1,t_1} W_{s_2,t_2}) = \min(s_1, s_2) \min(t_1, t_2) \,.$$

The trajectories of this process, that is, the mappings $(s,t) \mapsto W_{s,t}(\omega)$ are continuous surfaces, and for any fixed s, $t \mapsto W_{s,t}$ is a Brownian motion, and likewise, $s \mapsto W_{s,t}$ is also a Brownian motion. The purpose of the article [1] is to construct a stochastic calculus for the two-parameter Brownian motion similar to the classical Itô calculus developed by Kyoshi Itô in the 40's for the standard Brownian motion. New ingredients appear here, for instance, one can define surface integrals and also curvilinear integrals. On other hand, this calculus should be related to the theory of two-parameter martingales.

Let us introduce some basic notation of the theory of two-parameter processes. For any point $z = (s,t) \in \mathbb{R}_+^2$ we denote by R_z the rectangle $[0,s] \times [0,t]$. Also, for any $z \in \mathbb{R}_+^2$ we denote by $\mathcal{F}_z$ the σ-field generated by the random variables $\{W_\xi, \xi \in R_z\}$. We say that a process $\{\phi(z), z \in \mathbb{R}_+^2\}$ is adapted if $\phi(z)$ is $\mathcal{F}_z$-measurable for each z. The notion of predictability is stronger than adaptability and is required to define stochastic integrals. The predictable σ-field $\mathcal{P}$ of subsets of $\mathbb{R}_+^2 \times \Omega$ is generated by the sets of the form $(s,t] \times (s',t'] \times \Lambda$, where $\Lambda \in \mathcal{F}_{s,s'}$. A two-parameter process X is called predictable if the mapping $(z,\omega) \mapsto X_z(\omega)$ is measurable with respect to the predictable σ-field. These notions are similar to the one-parameter case. The main difference is the fact that the parameter space $\mathbb{R}_+^2$ is partially ordered and this creates new difficulties. In addition to the σ-fields $\mathcal{F}_z$, one can consider also the bigger σ-fields $\mathcal{F}_z^1$ and $\mathcal{F}_z^2$, generated by the random variables $\{W_{s',t}, s' \le s\}$ and $\{W_{s,t'}, t' \le t\}$, respectively, where $z = (s,t)$.

Given $z = (s,t) \in \mathbb{R}_+^2$, the surface stochastic integral with respect to the Brownian sheet W on the rectangle R_z

$$\int_{R_z} \phi(\xi) dW_\xi \,,$$

is defined for processes $\{\phi(\xi), \xi \in \mathbb{R}_+^2\}$ which are predictable and square integrable, that is,

$$E\left(\int_{R_z} \phi(\xi)^2 d\xi\right) < \infty \,,$$

for each $z \in \mathbb{R}_+^2$. This is the counterpart of the Itô integral. In the case of a process continuous in $L^2(\Omega)$, this integral is the limit in $L^2(\Omega)$ of the Riemann sums:

$$\int_{R_z} \phi(\xi) dW_\xi = \lim_{n\to\infty} \sum_{i,j=0}^{n-1} \phi(z_{i,j}) W(\Delta_{i,j}) ,$$

where $z_{i,j} = (is/n, jt/n)$, $\Delta_{i,j} = (z_{i,j}, z_{i+1,j+1}]$, and $W(\Delta_{i,j})$ denotes the increment of the process W on the rectangle $\Delta_{i,j}$ defined by

$$W(\Delta_{i,j}) = W_{z_{i+1,j+1}} - W_{z_{i,j+1}} - W_{z_{i+1,j}} + W_{z_{i,j}} .$$

This integral has zero expectation and satisfies the classical Itô isometry property:

$$E\Big(\Big| \int_{R_z} \phi(\xi) dW_\xi \Big|^2 \Big) = E\Big(\int_{R_z} \phi(\xi)^2 d\xi \Big) .$$

This is a consequence of the fact that the process W has independent increments in disjoint rectangles, and we have considered Riemann sums based on the value of the process in the lower left corner of the rectangle.

A fundamental result in Itô calculus is the Martingale Representation Theorem that asserts that any square-integrable martingale relative to the natural fields of the Brownian motion can be written as a constant plus a stochastic integral. In order to extend this result to the framework of the two-parameter Brownian motion, we need first to introduce the notion of martingale for two-parameter processes. The simplest way to do this is to use the partial ordering on the plane: $z' = (s', t') \le z = (s, t)$ if and only if $s' \le s$ and $t' \le t$. An adapted stochastic process $M = \{M_z, z \in \mathbb{R}_+^2\}$, is called a martingale if $E(|M_z|) < \infty$ for each z, and

$$E(M_z | \mathcal{F}_{z'}) = M_{z'}$$

for each $z' \le z$. It turns out that the Martingale Representation Theorem is no longer true in the framework of the two-parameter Brownian motion. More precisely, Wong and Zakai [5] proved the following result: If $M = \{M_z, z \in \mathbb{R}_+^2\}$ is a square integrable martingale, then for each $z \in \mathbb{R}_+^2$,

$$M_z = M_0 + \int_{R_z} \phi(\xi) dW_\xi + \int_{R_z \times R_z} \psi(\xi, \xi') dW_\xi dW_{\xi'} ,$$

where the second integral is a double stochastic integral, and the process $\psi(\xi, \xi')$ vanishes except if $\xi = (s, t)$ and $\xi' = (s', t')$ satisfy $s < s'$ and $t > t'$, is square integrable and it satisfies a suitable predictability condition.

The stochastic integrals

$$\Big\{\int_{R_z} \phi(\xi)dW_\xi, z\in \mathbb{R}^2_+\Big\}$$

constitute a special type of martingales, called *strong martingales.* This means that, for any $z \le z'$, the process $M_z = \int_{R_z} \phi(\xi)dW_\xi$ satisfies

$$E(M((z,z'])|\mathcal{F}^1_z \vee \mathcal{F}^2_z) = 0,$$

where $M((z,z'])$ denotes the rectangular increment of M, and $\mathcal{F}^1_z \vee \mathcal{F}^2_z$ denotes the σ field generated by $\mathcal{F}^1_z$ and $\mathcal{F}^2_z$. The strong martingale property of these integrals is a consequence of the fact that the two-parameter Brownian motion W has independent increments on disjoint rectangles. Furthermore, all strong square integrable martingales vanishing on the axes are stochastic integrals of the form $M_z = \int_{R_z} \phi(\xi)dW_\xi$.

The notion of *quadratic variation* plays a basic role in Itô calculus, and it is the source of the complementary terms appearing in the classical Itô formula. The quadratic variation of a one-parameter continuous process $\{X_t, t\ge 0\}$ is defined, if it exists, as the limit in probability

$$\langle X\rangle_t = \lim_{n\to\infty}\sum_{i=0}^{n-1}(X_{(i+1)t/n} - X_{it/n})^2.$$

For example, if B_t is a Brownian motion, $\langle B\rangle_t = t$. It turns out that any continuous martingale $M = \{M_t, t\ge 0\}$ has an increasing and continuous quadratic variation $\langle M\rangle_t$. Now, the restriction of a two-parameter martingale $M = \{M_z, z\in\mathbb{R}^2_+\}$ to a continuous increasing path in the plane

$$\gamma = \{\gamma(t), 0\le t\le 1\},$$

starting at the origin, defines a one-parameter martingale

$$M^\gamma = \{M_{\gamma(t)}, 0\le t\le 1\}$$

and we can compute its quadratic variation $\langle M^\gamma\rangle_t$. We say that a two-parameter martingale $M=\{M_z, z\in\mathbb{R}^2_+\}$ has *path-independent variation* if

$$\langle M^\gamma\rangle_1 = \langle M^{\gamma'}\rangle_1\,,$$

for any two paths γ and γ' such that $\gamma(1) = \gamma'(1)$. This notion was introduced by Moshe Zakai. Then, strong martingales have path-independent variation, and Cairoli and Walsh said in their paper that "We have not succeeded in proving that, in general, the converse is true, that is, that each martingale with path-independent variation is a strong martingale. However, several indications let us believe that path-independence is a second characterization of the strong martingales".

This statement le me to be interested in this challenging open problem. After working for a while, I was able to prove the surprising fact that this converse result is not true, and there are path-independent variation martingales which are not strong. The construction of such martingales is very delicate and it is obtained by an approximation procedure. I was very proud of this result which I consider my first important contribution to stochastic analysis. It was published in the proceedings of the conference in Paris devoted to two-parameter processes (see [4]).

The paper by Cairoli and Walsh [1] was actually motivated by the study of holomorphic processes in the plane. A process Φ is holomorphic if it has a derivative ϕ, in the sense that

$$\Phi_z = \Phi_0 + \int_0^z \phi \partial W \,,$$

where $\int_0^z \phi \partial W$ is a line integral taken over any sufficiently smooth curve connecting $(0,0)$ and z. These processes turn out to have a structure which is in some ways remarkably like that of classical holomorphic functions of a complex variable, even though they are real. For instance, if Φ is holomorphic, so is its derivative ϕ, and there is even an analogue of the power series expansion. Some years later, in collaboration with Ely Merzbach (see [2]) using techniques of Malliavin calculus, and more precisely, the Clark-Ocone formula to represent Φ_z as a stochastic integral, I was able to obtain a condition on the Malliavin derivative of Φ that characterizes holomorphicity. This leads to a simple proof of the power series expansion of holomorphic processes.

The line integrals together with the surface integrals allowed Cairoli and Walsh to derive in [1] a Green formula for rectangles, and, as an application, to show a two-parameter version of the classical Itô formula. An immediate application of this formula was the existence and continuity of the local time for W by means of a suitable version of Tanaka's formula.

I continued working on two-parameter processes for a while, especially on regularity properties of martingales and their two-parameter quadratic variation. For instance, in [4], I was able to prove the continuity of the quadratic variation of a square-integrable two-parameter continuous martingale, which was also an open problem. As other researchers in the field, at the beginning of the eighties I shifted my research interests to other topics like stochastic partial differential equations which are also connected with multiparametric processes.

References

[1] Cairoli, R. and Walsh, J. B.: Stochastic integrals in the plane. *Acta Math.* **134** (1975), 111–183.

[2] Merzbach, E. and Nualart, D.: Generalized holomorphic processes and differentiability. *J. Theoret. Probab.* **2** (1989), 419–432.

[3] Nualart, D.: Martingales à variation indépendante du chemin. In *Two-index random processes (Paris, 1980)*, 128–148. Lecture Notes in Math. **863**. Springer, Berlin, 1981.

[4] Nualart, D.: On the quadratic variation of two-parameter continuous martingales. *Ann. Probab.* **12** (1984), 445–457.

[5] Wong, E. and Zakai, M.: Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrschienlichkeitstheorie und verw. Gebiete* **29** (1974), 109–122.

ANTONIO ROS (Frontignan, France, 1957; raised in Lorca, Spain) obtained his PhD in mathematics at the University of Granada in 1984 (advisor: Luis Esteban). He is Professor at the University of Granada and his research focuses on geometric analysis, and in particular on minimal surfaces and isoperimetric problems.

Stable minimal surfaces

Antonio Ros*

Minimal surfaces in Euclidean space $\mathbb{R}^3$ have played a key role in the development of differential geometry. They are linked to the *Plateau Problem* (determine the shape of least area surfaces whose boundary is a given Jordan curve) and are the mathematical model for soap films.

A surface S is minimal if any compact region $R \subset S$ has zero *first variation formula* (of the area), $A'(0) = 0$. This is equivalent to saying that its mean curvature is zero. In the particular case where the surface is the graph of a function $f(x, y)$, it is characterized by the *minimal surface equation*

$$(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy} = 0. \qquad (1)$$

Stable minimal surfaces are the local minima of the area functional. Analytically, if the surface is two-sided and K denotes its Gaussian curvature, stability means that for any regular function with compact support $u \in C_0^2(S)$ we have

$$A''(0) = \int_S |\nabla u|^2 + 2Ku^2 \geq 0, \qquad (2)$$

or equivalently, the operator $-(\Delta - 2K)$ is non-negative. At the same time, fundamental results [17] give that a minimal surface S is stable if and only if every compact region $R \subset S$ has less area than any other with the same boundary and close enough to R (in the C^0-sense).

The global theory of stable minimal surfaces began in 1915 with the *Bernstein's theorem,* which states that the only entire solutions of equation (1) are affine functions. As a generalization, the theorem leads us to study area minimizing properly embedded surfaces (since the graph of an entire solution of (1) has these properties), and in the next step, we arrive to the question we are interested in this note:

> ($\star$) *The only complete stable minimal surface, i.e., with nonnegative* second variation formula $A''(0) \geq 0$*, is the plane.*

The papers by Do Carmo and Peng [1], Fischer-Colbrie and Schoen [5] and Pogorelov [13] solved the above question in the affirmative for two-sided surfaces around the year 1980. The proof of [5] also applies to ambient spaces with curvature greater than or equal to zero. It is not unusual

*aros@ugr.es.

that different authors reach the same result independently, but in this case three different proofs appeared simultaneously figured out by some first class mathematicians.

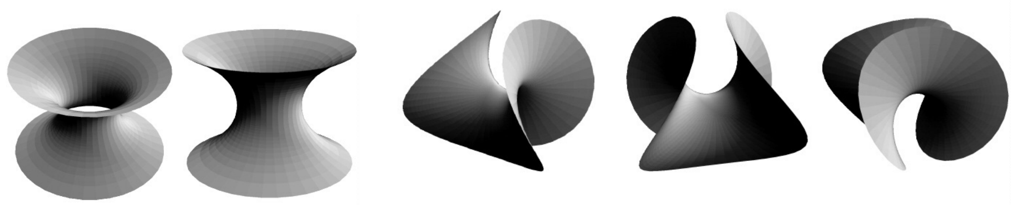
When we started to study these issues, the conjecture was already solved (in the orientable case) and the previous works have exerted a great influence on the geometry group of Granada and, in particular, in our training. In fact its role has been crucial throughout the geometry of surfaces and it has given rise to research lines which remain active nowadays, see for example the recent survey [11].

The condition $(\star)$ is equivalent to the fact that there exists a positive solution on S of the equation

$$\Delta u - 2Ku = 0.$$

Passing to the universal covering, we can assume that S is simply connected and therefore, the surface is conformally equivalent either to the plane or the disc (as a minimal surface cannot have the topology of the sphere). Since the Gaussian curvature of a minimal surface is less than or equal to zero, the Liouville theorem for subharmonic functions implies that, in the first case, $K = 0$ and so the surface is a plane in $\mathbb{R}^3$. The case of the disc is more delicate.

We started in the field of *Geometric Analysis*, the combination of geometry and partial differential equations, by studying these papers and, following in his wake, joint with Francisco López we showed in [9] that an orientable complete minimal surface has index one (*i.e.*, it has nonnegative second variation but only for deformations which are L^2-orthogonal to a certain direction on the surface), if and only if it is the Catenoid or the Enneper surface.



Catenoid Enneper Surface

It is worth noticing an unexpected connection between stability and another classic question in this area: The study of the image of the Gauss map. Xavier [18] had shown that the Gauss map of a complete nonflat minimal surface omits at most 6 points on the sphere. This is a geometric variant of the Theorem of Picard on the values omitted by entire holomorphic functions in one variable. The study of the above operators allowed us, joint with

Paco, to combine these techniques with the one of [18] to improve the exceptional 6 values of Xavier's theorem to 5. Our result was never published (although it appeared in Paco's thesis [8]) because in the meantime we knew that Fujimoto [6] had found the complete solution of the problem of the Gauss map, by showing that it omits at most 4 points (as the example of the surface of Scherk), see also Ros [14]. Our approach depended on the fact that if the Gauss map omits 3 values, then the operator $-(\Delta - a\,K)$ is nonnegative for some $a > 0$. This is a generalization of the concept of stability, whose consequences in the theory of surfaces are of great interest, although not sufficiently understood at the moment. A basic problem in this direction proposes the study of the values of $a > 0$ for which any complete Riemannian surface such that the operator $-(\Delta - aK)$ is nonnegative is necessarily parabolic. The work [5] contains this result for $a > 1$ and it is known that for the hyperbolic plane of curvature -1, the operator $-(\Delta - aK)$ is nonnegative just for the range $a \geq 1/4$. The complete solution to this problem was obtained by Castillon [2], who showed that if a complete Riemannian surface admits a positive solution of $\Delta u - aKu = 0$ with $a > 1/4$, then it is conformally equivalent to the plane or a quotient of the plane. The way in which Castillon dealt with this problem depends on techniques introduced by Colding and Minicozzi [3] that, in part, were a generalization of those of Pogorelov [13] and Kawai [7] who had previously considered the case $K \leq 0$, a restriction that permits a more comfortable treatment as on these surfaces the distance function to a given point is regular. This direction of research is still active, Espinar and Rosenberg [4], and in particular the study of stability in other ambient spaces, Manzano, Pérez and Rodríguez [10].

For nonorientable surfaces the conjecture about complete stable minimal surfaces remained open until some years ago and, although several partial results were known, we were lucky to find an argument that closed the problem in the general case, Ros [15]: there are not one-sided complete stable minimal surfaces in $\mathbb{R}^3$. This result is specific of $\mathbb{R}^3$ and does not extend to spaces of positive curvature. In fact, there are one-sided compact stable minimal surfaces with arbitrarily complicated topology in suitably chosen quotients of the three-dimensional sphere. Another fundamental difference with the orientable situation is that now, stability is not preserved when we pass to a covering space. The deformations we use in [15] are related to square integrable harmonic 1-forms over the surface.

There are complete one-sided stable nonflat minimal surfaces in quotients of $\mathbb{R}^3$. Ross [16] showed that the periodic examples P and D of Schwarz are stable and it would be interesting, both in geometric analysis and in mathematical crystallography, to classify global stable surfaces in these quotients. Such surfaces are necessarily one-sided and Ros [15] proved that the unique nonflat complete stable minimal surfaces in the quotient of $\mathbb{R}^3$ by one or two linearly independent translations are certain nonorientable quotients of the Helicoid and the Scherk surface. If we consider quotients involving more complicated rigid motions, then the question remains still open.

Of course it is natural to consider stable minimal hypersurfaces in $\mathbb{R}^n$. The study of Bernstein's Theorem in higher dimensions and the theory of absolute area minimizing hypersurfaces of $\mathbb{R}^n$, [12], have largely influenced both geometry and partial differential equations. However, the consequences of stability on complete stable minimal hypersurfaces of $\mathbb{R}^4$ are still unknown. Besides the importance of this problem in geometry, it represents a challenge to develop the techniques that will allow us to understand the implications of the nonnegativity of the operator $\Delta - aR$ on a complete Riemannian manifold, where a is a positive constant and R is the scalar curvature.

References

[1] Carmo, M. do and Peng, C. K.: Stable complete minimal surfaces in $\mathbb{R}^3$ are planes. *Bull. Amer. Math. Soc. (N.S.)* **1** (1979), 903–906.

[2] Castillon, P.: An inverse spectral problem on surfaces. *Comment. Math. Helv.* **81**(2) (2006), 271–286.

[3] Colding, T. and Minicozzi, W.: Estimates for parametric elliptic integrands. *Internat. Math. Res. Notices* **6** (2002), 291–297.

[4] Espinar, J. M. and Rosenberg, H.: A Colding-Minicozzi Stability inequality and its applications. *Trans. Amer. Math. Soc.* **363** (2011), 2447–2465.

[5] Fischer-Colbrie, D. and Schoen, R.: The structure of complete stable minimal surfaces in 3-manifolds of nonnegative scalar curvature. *Comm. Pure Appl. Math.* **33** (1980), 199–211.

[6] Fujimoto, H.: On the number of exceptional values of the Gauss map of minimal surfaces. *J. Math. Soc. Japan* **40** (1988), 235-247.

[7] Kawai, S.: Operator $\Delta - aK$ on surfaces. *Hokkaido Math. J.* **17** (1988), 147–150.

[8] López, F. J.: *Superficies minimales. Superficies de curvatura media constante.* Tésis de la Universidad de Granada, 1989.

[9] López, F. J. and Ros, A.: Complete minimal surfaces with index one and stable constant mean curvature surfaces. *Comment. Math. Helvet.* **64** (1989) 34–43.

[10] Manzano, J. M., Pérez, J. and Rodríguez, M.: Parabolic stable surfaces with constant mean curvature. To appear in *Calculus of Variations and PDEs*.

[11] Meeks III, W. H., Pérez, J. and Ros, A.: Stable constant mean curvature surfaces. In *Handbook of Geometric Analysis no. 1*, 301–381. Editors: Lizhen Ji, Peter Li, Richard Schoen, Leon Simon. International Press, 2008.

[12] Morgan, F.: *Geometric Measure Theory: a Beginner's Guide*. Academic Press, 4th ed., 2008.

[13] Pogorelov, A. V.: On the stability of minimal surfaces. *Dokl. Akad. Nauk SSSR* **260** (1981), 293–295.

[14] Ros, A.: The Gauss map of minimal surfaces. In *Differential Geometry (Valencia, 2001), Proceedings of the conference in honour of Antonio Naveira*, 235–252 (O. Gil-Medrano and V. Miquel, editors). World Scientific, 2002.

[15] Ros, A.: One-sided complete stable minimal surfaces. *J. Diff. Geom.* **74** (2006), 69–92.

[16] Ross, M.: Schwarz' P and D surfaces are stable. *Differential Geom. Appl.* **2** (1992), 179–195.

[17] White, B.: A strong minimax property of nondegenerate minimal submanifolds. *J. Reine. Angew. Math.* **457** (1994), 203–218.

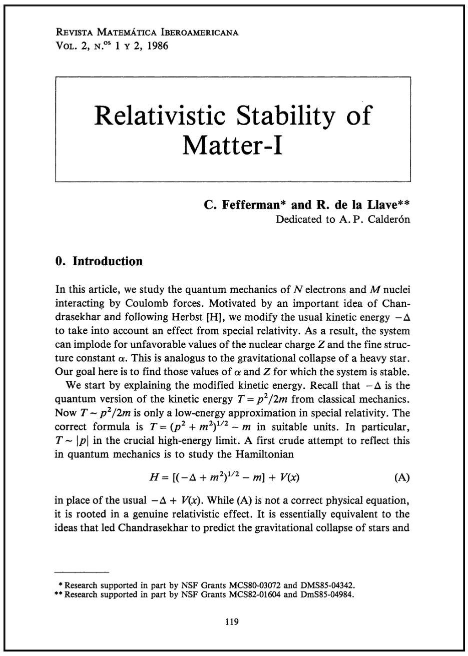
[18] Xavier, F.: The Gauss map of a complete non-flat minimal surface cannot omits 7 points of the sphere. *Ann. of Math.* **113** (1981), 211–214.

Luis Ángel Seco (Bilbao, Spain, 1962) obtained his PhD in mathematics from Princeton University in 1989 (advisor: Charles Fefferman). Since 1992 he has been at the University of Toronto, where he is presently Professor in the Department of Mathematics and Director of the RiskLab. He is also the President and CEO of Sigma Analysis and Management, a portfolio management firm that manages investments for pension plans. His areas of research include a past history in mathematical physics and present activity in financial risk management and portfolio theory.

Relativistic stability of matter

Luis Ángel Seco*

I arrived at Princeton University as a graduate student just as the paper "Relativistic Stability of Matter–I" (C. Fefferman and R. de la Llave, *Rev. Mat. Iberoamericana* **2** (1986)) was being finished; I had no knowledge of quantum mechanics, but listening to Charlie Fefferman's lecture about what would become this paper make it clear to me that this was a result unlike anything I had ever read; this is not a strong statement coming from a rookie graduate student but, many decades later, I can say that this paper is unlike any other mathematics paper I have ever read. And this includes a lot of papers now.



Revista Matemática Iberoamericana
Vol. 2, n.os 1 y 2, 1986

Relativistic Stability of Matter-I

C. Fefferman* and R. de la Llave**
Dedicated to A. P. Calderón

0. Introduction

In this article, we study the quantum mechanics of N electrons and M nuclei interacting by Coulomb forces. Motivated by an important idea of Chandrasekhar and following Herbst [H], we modify the usual kinetic energy $-\Delta$ to take into account an effect from special relativity. As a result, the system can implode for unfavorable values of the nuclear charge Z and the fine structure constant α. This is analogus to the gravitational collapse of a heavy star. Our goal here is to find those values of α and Z for which the system is stable.

We start by explaining the modified kinetic energy. Recall that $-\Delta$ is the quantum version of the kinetic energy $T = p^2/2m$ from classical mechanics. Now $T \sim p^2/2m$ is only a low-energy approximation in special relativity. The correct formula is $T = (p^2 + m^2)^{1/2} - m$ in suitable units. In particular, $T \sim |p|$ in the crucial high-energy limit. A first crude attempt to reflect this in quantum mechanics is to study the Hamiltonian

$$H = [(-\Delta + m^2)^{1/2} - m] + V(x) \tag{A}$$

in place of the usual $-\Delta + V(x)$. While (A) is not a correct physical equation, it is rooted in a genuine relativistic effect. It is essentially equivalent to the ideas that led Chandrasekhar to predict the gravitational collapse of stars and

* Research supported in part by NSF Grants MCS80-03072 and DMS85-04342.
** Research supported in part by NSF Grants MCS82-01604 and DmS85-04984.

119

For me, this is a ground-breaking article not for any of the usual reasons; of course, the result is important, the proof intelligent, the exposition is instructive but what is most striking about this paper is that it created an entirely new way of thinking about theorem proving in quantum mechanics, and for me it introduced a new way of thinking overall. In fact, reflecting on my own experience, I have never personally worked on relativistic quantum mechanics, but I did spend ten years working on classical quantum mechanics. Nevertheless, the concepts of this paper provided with enough inspiration to deal with ten years worth of mathematical challenges in quantum many body problems. This paper is rich with ideas, but I can say that some of these ideas actually transcend the realm of quantum mechanics and offer deep insights about how to attach seemingly impossibly hard problems, in mathematics and beyond. I confess that this paper has had a critical impact, not just on my research in quantum mechanics, but in my research in mathematical finance later in my life and in my activities as an entrepreneur. It is my objective in this short introspective to show why this is the case.

*seco@math.toronto.edu.

The paper starts with an excellent summary of quantum mechanics, the relativistic theory of matter, the concept of stability, but also introduces the main philosophy of the paper, which represented a new approach to dealing with kinetic energy and Coulomb potentials under a unique new common roof: in this regard, the key sentence in the paper appears in page 4:

> "More generally, we hope the ideas in our proofs will be useful tools in understanding many body problems."

and this really is the main achievement of this paper.

In a certain sense, this paper can be summarized as follows:

- Section 1 explains why is this problem is interesting and so hard.
- Sections 2 and 3 talks about what can we do about it.
- The rest (75% of the paper) are details.

The paper does not get obsessed with the problem, nor with the solution; instead, it gives us hints, here and there, about where the difficulties lie, and how to best approach them. That's why, in my opinion, this paper ends with section 3, after just 12 pages of a 40 page paper. I know the authors will disagree, because I know how hard the rest of the article is, but, as a consumer of this paper, my perspective is different and the first twelve pages of this paper are not just brilliant to the point of overshadowing the rest of the paper, but they are the ones that provide true insight into the problem and constitute a point of reference to the universal topic to tackle problem solving.

The problem is concerned with the relativistic stability of matter; from a mathematical viewpoint, the authors express a relativistic Hamiltonian of M nuclei of charge Z_i at points y_i and N electrons at points x_j, where $j = 1, \ldots, M$, and $i = 1, \ldots, N$ is given by

$$H_{Z,M,N} = \sum_{k=1}^{N} (-\Delta_{x_k})^{1/2} + \alpha \cdot V_{Z,M,N}\,, \tag{1}$$

where the Coulomb potential is given by

$$V_{Z,M,N} = \sum_{j<k} \frac{1}{|x_j - x_k|} + \sum_{j<k} \frac{Z_j\, Z_k}{|y_j - y_k|} - \sum_{j,k} \frac{1}{|x_j - x_k|}\,. \tag{2}$$

This Hamiltonian acts on the Hilbert space H of antisymmetric functions in $L^2(\mathbb{R}^{3N})$, antisymmetry reflecting the fact that our particles are fermions. This model assumes the nuclei infinitely heavy so they do not move, sets the mass of electrons equal to $1/2$, ignores spin, α is the fine structure constant (whose role is to emerge later) and a set of other mathematical licenses to turn the original physical problem into this form. But the main assumption is the statement that the relativistic kinetic energy is given by the square root of the Laplacian. The paper actually starts arguing in favour of this representation, citing other authors who took similar approaches, and essentially telling us that in relativity, kinetic energy becomes proportional to momentum, not its square, in the high energy limit, which is the domain of interest of this paper, which will focus on stability of matter and is a property shared by the square root of the Laplacian.

After these beautiful arguments in favour of the justification of their perspective, they present the objective of the paper: to prove the stability of a large system of interacting nuclei and electrons, expressed as the inequality

$$\langle H_{Z,M,N}\psi \,,\, \psi\rangle \geq 0\,. \tag{3}$$

Here, the fine structure constant plays a critical role, which the authors address pointing out the critical difficulty of the problem: both terms, the Coulomb potential and the relativistic kinetic energy, scale in the same way (they both have units of 1/inches), therefore (3) will depend critically on the values of the fine structure constant α.

Since scale invariance will make the problem hard, they will turn to use scale invariance to their advantage by rewriting the kinetic energy and the Coulomb potential in a convenient form, in Section 2 of the paper, which is one of the most brilliant gambits I have ever seen:

This is best explained thinking about the following expression:

$$\frac{1}{|x|} = \frac{1}{\pi}\int\int_{R>0,\, z\in\mathbb{R}^3} \left\{\begin{matrix} 1 & \text{if both } x, 0 \in B(z,R) \\ 0 & \text{otherwise.}\end{matrix}\right\} \frac{dz\, dR}{R^5}\,. \tag{4}$$

First, we must learn to think from the authors that in this expression, the left hand side is the hard one (yes, the seemingly simple function function $1/|x|$), and that the easy one is the right hand side, which simply attaches the number 0 or 1 to a ball, depending on whether a given point x and the origin are both inside it. Second, we must realize that this expression is practically trivial, since both sides are translation invariant, and they scale under x in the same way (as inches^{-1}); therefore, they must be equal up to a multiplicative constant; therefore, the only difficulty in this expression is the calculation of the proportionality constant $1/\pi$.

This thought process leads to equations (1) and (2) in the paper, which appear in page 6, use the translation and dilation invariants of the Laplacian and the Coulomb potential (remember, the source of our main obstacle) to rewrite the entire hamiltonian 1 as an integral over all balls in $\mathbb{R}^3$, parametrized by all centers $z \in \mathbb{R}^3$ and all radii $R > 0$, of two extremely simple ball-dependent quantities. In the case of expression (2) in the paper, they show us how the Coulomb potential is an integral over all balls of a function that most kindergarten graduates understand: counting of numbers of pairs of electrons and nuclei inside the ball. The contribution of the kinetic energy is not as brutally simple, but it is still remarkably easy: it is a type of electron density variance, as expressed in equation (1) in the paper. The net result is that the ultimate objective, the stability inequality 3 will be reduced to a single ball-based inequality. But in order to do this, the paper needs to make another two critical steps: first, they will reduce the stability problem over a single ball to one that involves just one nuclei: this is important, since in hierarchy of difficulty in many body quantum problems, single nuclei and many electrons is easier than a problem involving several nuclei; second, they will break the problem further by considering a single nucleus and a single electron, with a kinetic-type and a potential-type Hamiltonians specific to a single ball; this double-tiered decoupling, from many nuclei and electrons to a single nucleus and a single electron is a simple consequence of their ultra-simple energy expressions derived for each ball in $\mathbb{R}^3$.

This crakes the nut open; what remains is very hard technically, and it requires fine analysis, but at the end of the day is a problem of the simplest type in the hierarchy of difficulty of many-body quantum systems: a single nucleus and a single electron.

However, there is more; I was not completely accurate when I said the paper ended after the first 12 pages (of course). There is a small hidden detail, which actually makes up for the last 50 pages or so which are technically outside the paper proper. If the reader checks the article, about half of its 95 pages are devoted to a computer printout. What happened in the last section of the paper, the one I claimed were mere "details", is that the authors reduce the proof of their inequality 3 to a long but elementary set of properties that, if checked by a human will take a very, very long time, but if checked by a conveniently trained computer, they can be done in a reasonable timeframe. Hence, the paper ends with a computer assisted proof (emphasis on the word "proof"); a few years later, and with the help of Rafael de la Llave, I was able to also do my own computer-assisted proof; I have proved a lot of theorems in my life, but only two with the help of a computer; however, those two took a disproportionately high amount of my own human time; but more importantly, my computer-assisted activities emerged from my observation of the progress of this paper and, as with this paper, were unavoidable at the time if one was to pursue the original mathematical problem, and not a conveniently simplified version of itself.

In summary, this paper taught me several lessons, which have lasted many decades and proved to be useful many times, but can be summarized by saying that problem solving is not an end in itself, it is the school of life.

Stephen Semmes (Savannah, Georgia, USA, 1962) obtained his PhD in mathematics at Washington University in 1983 under the direction of Richard Rochberg. He is Professor of Mathematics at Rice University in Houston, Texas. His research focuses on harmonic analysis, geometry and related topics.

Some reflections on papers by Heinonen and Rubio de Francia

Stephen Semmes*

Let D be a simply-connected domain in the complex plane. By the Riemann mapping theorem, there is a conformal mapping f from the unit disk U onto D. If D is bounded by a Jordan curve, then a famous theorem of Carathéodory states that f extends to a homeomorphism from the closed unit disk onto the closure of D. If in addition the boundary of D is rectifiable, then it is well-known that the boundary values of f determine an absolutely continuous mapping from the unit circle to the boundary of D. Basically, this is because the derivative f' of f is an element of the Hardy space H^1, as a consequence of the F. and M. Riesz theorem.

However, the analogous absolute continuity property on the boundary does not work for quasiconformal mappings in the plane. More precisely, there are examples of Beurling and Ahlfors [2] of quasiconformal mappings from the unit disk onto itself whose boundary values on the unit circle are not absolutely continuous. In a series of papers, Juha Heinonen [10, 11, 12] considered similar questions for quasiconformal mappings in higher dimensions. In particular, in the original version of [11], Juha asked about a certain type of absolute continuity property of quasisymmetric mappings defined on subsets of $\mathbb{R}^n$.

Without getting too technical, a mapping between arbitrary metric spaces is said to be *quasisymmetric* [23] if it does not distort relative distances too much at any scale. This is a relative of the notion of a quasiconformal mapping on $\mathbb{R}^n$, which is defined classically in terms of the boundedness of the ratio of the maximal and minimal stretching of the differential of a mapping. Under suitable conditions, well-known distortion theorems allow infinitesimal quasiconformality conditions to be "integrated" to get quasisymmetry conditions at definite scales. However, quasisymmetric mappings can also be considered when distortion theorems like these are not available.

For example, it does not really make sense to talk about quasiconformal mappings on the unit circle, because the maximal and minimal stretching of the differential would be the same automatically. It does make sense to talk about quasisymmetric mappings on the unit circle, and indeed the main result of [2] says that an orientation-preserving homeomorphism on the unit circle corresponds to the boundary values of a quasiconformal mapping from the unit disk onto itself if and only if it is quasisymmetric. Thus the ex-

*semmes@rice.edu.

amples in [2] mentioned previously are actually examples of quasisymmetric mappings on the circle that are not absolutely continuous.

FIGURE 1: Looking to the future of quasiconformal mappings (Djursholm, Sweden, 1983). Although the classical theory was familiar to many of us studying analysis at that time, we could not have anticipated the developments to come, in which Juha Heinonen played a very important role.

By contrast, quasiconformal mappings on $\mathbb{R}^n$ are absolutely continuous when $n \geq 2$. In the original version of [11], Juha asked about absolute continuity of quasisymmetric mappings defined on arbitrary subsets of $\mathbb{R}^n$ when $n \geq 2$. I was very interested in Juha's question, in part because I was already thinking about some related matters, in connection with the BPI (or "big pieces of itself) theory of self-similar fractals in [5]. Of course, a key point in Juha's question is that quasisymmetric mappings on subsets of $\mathbb{R}^n$ cannot normally be extended to quasiconformal mappings on $\mathbb{R}^n$. The main idea in my approach was to work with measures instead of mappings, corresponding to the Jacobians of quasiconformal mappings. It is much easier to make various constructions with measures instead of mappings, avoiding topological obstructions in particular. At the same time, one has to keep track of enough of the geometric information in the measures to get absolute continuity.

Fortunately for us, the editors at the *Revista* were quite sympathetic to our situation. Thus we were able to publish our papers in the same issue in [11, 12, 18]. It was also very nice in the way that so many aspects of analysis and geometry came together.

Actually, Juha and I have been discussing questions related to absolute continuity from the beginning. I think that our first contact was an email that he sent me about a statement on pp. 102–103 in [4], to the effect that the Jacobian of a quasisymmetric mapping of $\mathbb{R}^n$ onto an n-dimensional Ahlfors-regular subset of $\mathbb{R}^m$ for some $m > n$ is a strong A_∞ weight. This extends the case of a quasiconformal mapping of $\mathbb{R}^n$ onto itself, and Juha was asking if we were inadvertently using results about absolute continuity of such mappings without mentioning it. I remember reading his email and thinking to myself "uh-oh", wondering if we had overlooked something. Of course, the A_∞

condition for Jacobians of quasiconformal mappings comes from Gehring's well-known paper [9]. Gehring's approach can also be used to show that a doubling measure is absolutely continuous under certain conditions, with a density that is an A_∞ weight, by approximating the measure by weights with bounded A_∞ constants. This version of Gehring's argument can be applied to measures obtained from quasisymmetric mappings of $\mathbb{R}^n$ onto n-dimensional Ahlfors-regular sets, including quasiconformal mappings of $\mathbb{R}^n$ onto itself. At any rate, this was not explained very well in [4], and more details can be found in in [17].

I would also like to say a few words about Rubio de Francia's famous paper [16] on square functions associated to arbitrary collections of pairwise-disjoint intervals in the real line. Let $\{I_j\}$ be a countable family of pairwise-disjoint intervals in the real line. Alternatively, one can take the I_j's to be closed intervals in $\mathbb{R}$ with pairwise-disjoint interiors, so that two such intervals may have a common endpoint. If f is a reasonable function on $\mathbb{R}$, then let $S_j(f)$ be the function defined on $\mathbb{R}$ by

$$(S_j(f))\widehat{\ }(\xi) = \mathbf{1}_{I_j}(\xi)\, \widehat{f}(\xi). \tag{1}$$

Here $\widehat{f}(\xi)$ denotes the Fourier transform of f, and $\mathbf{1}_{I_j}(\xi)$ is the characteristic or indicator function on $\mathbb{R}$ associated to I_j, equal to 1 when $\xi \in I_j$ and to 0 when $\xi \in \mathbb{R}\backslash I_j$. Under these conditions, Rubio shows in [16] that for each real number $p \geq 2$ there is a nonnegative real number C_p such that

$$\Big\| \Big(\sum_j |S_j(f)|^2 \Big)^{1/2} \Big\|_p \leq C_p \, \|f\|_p \tag{2}$$

for every $f \in L^p(\mathbb{R})$. This follows from classical results of Littlewood and Paley when the intervals are of the form $[2^k, 2^{k+1}]$ or $[-2^{k+1}, -2^k]$, where k is an integer. In this case, there is a similar estimate for $1 < p < 2$, but this does not work for arbitrary families of intervals with disjoint interiors. A very nice overview of Rubio's theorem and related work can be found on [6, pp. 185-187].

Around that time, Rubio visted Yale University for a semester, while I was a post-doc there. It was natural to try to find a project of common interest, and certainly a wonderful opportunity for me. In particular, I asked about the potential consequences of Rubio's theorem for boundedness of linear operators, since we all know that square functions are closely related to that. Coifman, Rubio and I managed to obtain some results along these lines in [3]. Rubio's paper caused quite a stir when it came out, and hopefully other contributors to this volume will comment on it as well.

References

[1] Ahlfors, L.: *Lectures on Quasiconformal Mappings*, 2nd edition. American Mathematical Society, 2006.

[2] Beurling, A. and Ahlfors, L.: The boundary correspondence under quasiconformal mappings. *Acta Math.* **96** (1956), 125–142.

[3] Coifman, R., Rubio de Francia, J. L. and Semmes, S.: Multiplicateurs de Fourier de $L^p(\mathbb{R})$ et estimations quadratiques. *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988), 351–354.

[4] David, G. and Semmes, S.: Strong A_∞ weights, Sobolev Inequalities and quasiconformal mappings. In *Analysis and Partial Differential Equations*, 101–111. Dekker, 1990.

[5] David, G. and Semmes, S.: *Fractured fractals and broken dreams: Self-similar Geometry through metric and measure*. Oxford Univ. Press, 1997.

[6] Duoandikoetxea, J.: *Fourier analysis*, translated and revised from the 1995 Spanish original by D. Cruz-Uribe, SFO. American Mathematical Society, 2001.

[7] Duren, P.: *Theory of H^p spaces*. Academic Press, 1970.

[8] García-Cuerva, J. and Rubio de Francia, J. L.: *Weighted Norm Inequalities and Related Topics*. North-Holland, 1985.

[9] Gehring, F.: The L^p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.* **130** (1973), 265–277.

[10] Heinonen, J.: The boundary absolute continuity of quasiconformal mappings. *Amer. J. Math.* **116** (1994), 1545–1567.

[11] Heinonen, J.: The boundary absolute continuity of quasiconformal mappings, II. *Rev. Mat. Iberoamericana* **12** (1996), 697–725.

[12] Heinonen, J.: A theorem of Semmes and the boundary absolute continuity in all dimensions. *Rev. Mat. Iberoamericana* **12** (1996), 783–789.

[13] Heinonen, J.: *Lectures on analysis on metric spaces*. Springer-Verlag, 2001.

[14] Journé, J.-L.: *Calderón–Zygmund Operators, Pseudodifferential Operators, and the Cauchy Integral of Calderón*. Lecture Notes in Mathematics **994**. Springer-Verlag, 1983.

[15] Lehto, O. and Virtanen, K.: *Quasiconformal Mappings in the Plane*, translated from the German by K. Lucas. Springer-Verlag, 1973.

[16] Rubio de Francia, J. L.: A Littlewood–Paley inequality for arbitrary intervals. *Rev. Mat. Iberoamericana* **1** (1985), 1–14.

[17] Semmes, S.: Bi-Lipschitz mappings and strong A_∞ weights. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **18** (1993), 211–248.

[18] Semmes, S.: Quasisymmetry, measure, and a question of Heinonen. *Rev. Mat. Iberoamericana* **12** (1996), 727–781.

[19] Stein, E.: *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.

[20] Stein, E.: *Topics In Harmonic Analysis Related to the Littlewood–Paley Theory*. Annals of Mathematics Studies **63**. Princeton University Press, 1970.

[21] Stein, E.: *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

[22] Stein, E. and Weiss, G.: *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, 1971.

[23] Tukia, P. and Väisälä, J.: Quasisymmetric embeddings of metric spaces. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **5** (1980), 97–114.

[24] Väisälä, J.: *Lectures on n-dimensional quasiconformal mappings*. Lecture Notes in Mathematics **229**. Springer-Verlag, 1971.

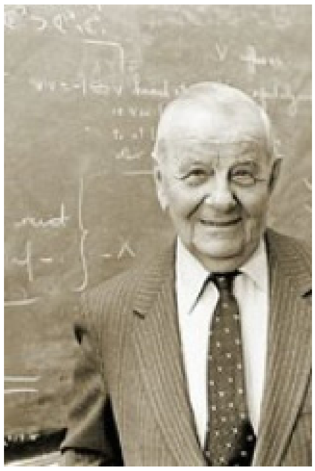
[25] Zygmund, A.: *Trigonometric Series, Volumes I and II,* 3rd edition, with a foreword by R. Fefferman. Cambridge University Press, 2002.

Elias M. Stein (Antwerp, Belgium, 1931) obtained his PhD in mathematics at the University of Chicago in 1955 (advisor: Antoni Zygmund). He is the Albert Baldwin Dod Professor of Mathematics at Princeton University. His fields of interest are: harmonic analysis, several complex variables, and analysis on symmetric spaces.

Some early mathematical reminiscences

Elias M. Stein*

Among my fondest mathematical remembrances are recollections of my teacher Antoni Zygmund, both his personality and the effect on me of his mathematical view-point. I can relate this best in terms of two papers he coauthored and a third paper, whose influence on me was much more indirect.



Antoni Zygmund

I want to begin with the famous paper "On the existence of certain singular integrals" with Calderón, which appeared in 1952 in Acta Mathematica. I know, both by what he said to me and by what I understood later, that his major mathematical goal at that time was to develop "real-variable" methods to allow extension to to higher dimensions of basic one-dimensional Fourier analysis. He stressed that these methods would be needed to supplant the "complex-methods" –involving analytic functions, Blaschke products, conformal mappings, etc.– which had proved to be so powerful in the one-dimensional theory, but that by their very nature were restricted to that setting. In this paper, he and Calderón had achieved a crucial break-through: the n-dimensional generalization of the L^1 and L^p theory of the Hilbert transform. Like all great works it of course owed a debt to earlier ideas of others. In this case to the work of Giraud and Mihlin in the formulation of the notion of singular integrals involved, and to Besicovitch's "real-variable" treatment of the Hilbert transform.

As we now know this paper had a major impact in the development of analysis. In the main, its early thrust was its applications for partial differential equations, both for elliptic and non-elliptic. For me it was in a different direction. I had become intrigued by what was known as the "Littlewood–Paley theory", the treatment of one-dimensional Fourier series by decomposing these series in dyadic blocks and obtaining L^p estimates for related square functions. At first it was only the statements of the results in Zygmund's "Trigonometrical Series"(the first edition of 1935 that was available at that time) that fascinated me. There, in Chapter 10, section 10.33, I read: "We shall state here without proof the most important of

*stein@math.princeton.edu.

Littlewood–Paley results" (and in a footnote it is said that detailed proofs had not yet been published). There followed statements of two of their theorems. The totality of these results seemed like a high and majestic peak whose top was hidden in the clouds. When I later found the papers giving the detailed proofs (written by Littlewood after Paley's death) I marvelled at the difficulty and ingenuity of the arguments, which relied heavily on complex methods. I decided to try to redo the theory by real-variable methods, so as to extend it to higher dimensions. Here I was able to follow the path laid out by the Acta paper, and use the Calderón–Zygmund decomposition to prove a weak-type L^1 estimate and L^p inequalities for various of the square functions. My initial efforts here were later recast more elegantly and systematically by others.

Another work of Zygmund that greatly influenced me was his paper with Calderón "On the theorem of Haussdorff-Young and its extensions" (1950). It gives a pithy version of Thorin's proof of the Riesz interpolation theorem, together with a multi-linear analogue that is applied to interpolation of Hardy spaces H^p. At that time (*circa* 1953), Zygmund had the practice of assigning his students the task of speaking about some recently published paper, and he suggested me that I present the interpolation paper. I was fascinated by the use of the maximum modulus principle (alias "three-lines lemma"), and spoke about the paper with what I thought was great enthusiasm. I knew Zygmund liked my presentation because he smiled at me during it; yet afterwards he asked me why I had paced back and forth so much! This paper left a deep impression on me which I was only to realize several years later.

During that time I had also become interested in the work of S. Bochner, who among other things was a pioneer in the study of the n-dimensional Fourier transform. I was particularly struck by his 1936 paper "Summation of multiple Fourier series by spherical means". In it, he pointed out the existence of the "critical index" $\frac{n-1}{2}$ (for what we now call "Bochner–Riesz summability") for Fourier series and Fourier transforms in n dimensions. He stressed that (at least for L^1) that order of summability was the analogue of ordinary convergence in one dimension. For example, summability of order $> \frac{n-1}{2}$ is like Cesàro summability of positive order for one dimension, in that it holds almost everywhere and in the norm for L^p, $1 \leq p$. However for the order $\frac{n-1}{2}$ this fails; in fact, by an ingenious argument involving Kronecker's theorem he showed that the analogue of Riemann's localization fails for L^1.

However, upon further reflection it became clear that known arguments for orthogonal expansions would prove that for L^2, dominated convergence (and almost everywhere results) would be valid for any strictly positive order of summability.

The question that occurred naturally was whether anything could be said about summability *below* the critical index for functions in L^p, when $p > 1$. I presented this as an interesting problem to Zygmund, asking for his opinion. He seemed indifferent. Very likely it was the multitude of uninteresting papers on summability of one-dimensional Fourier series that had been published during the past twenty years that soured him on anything involving "summability". Nevertheless, I continued to be intrigued by this question.

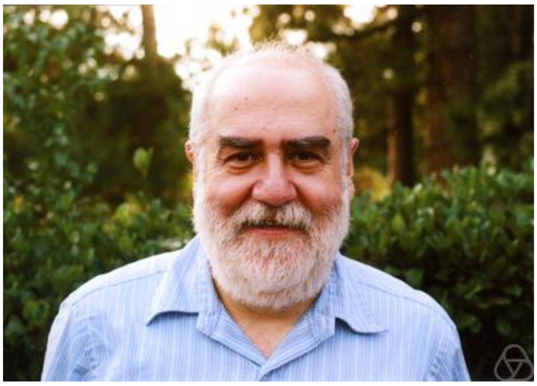
About a year later, in the summer of 1955, Zygmund went to Cornell to visit his friend Marc Kac. He suggested that I also go to Cornell. I rented a room near the campus, living a lonely life and spending much time thinking about mathematics. One day, while at the library, I realized that the ideas of the "Hausdorff-Young" paper would allow me to interpolate operators (of part of a suitable analytic family) as well as L^p spaces. This led quickly to results below the critical index for L^p. Thus two papers which had greatly influenced me, the Calderón–Zygmund work on interpolation and Bochner's paper on summability, were joined.

References

[1] BOCHNER, S.: Summation of multiple Fourier series by spherical means. *Trans. Amer. Math. Soc.* **40** (1936), 175–207.

[2] CALDERÓN, A. P. AND ZYGMUND, A.: On the theorem of Hausdorff-Young and its extensions. In *Contributions to Fourier Analysis*, 166–188. Ann. Math. Studies **25**. Princeton University Press, Princeton, NJ, 1950.

[3] CALDERÓN, A. P. AND ZYGMUND, A.: On the existence of certain singular integrals. *Acta Math.* **88** (1952), 85–139.



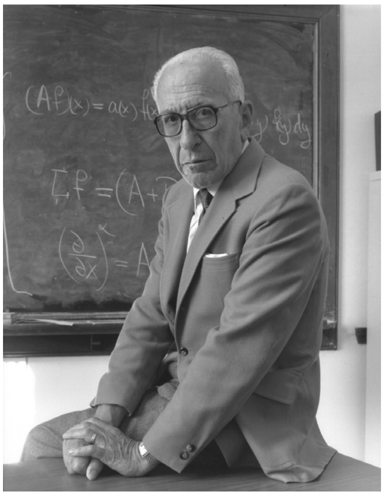
Gunther Uhlmann (Quillota, Chile, 1952) obtained his PhD in mathematics at MIT in 1976 (advisor: Victor W. Guillemin). Currently, he is Walker Family Endowed Professor of Mathematics at the University of Washington. His main research areas are inverse problems, partial differential equations, microlocal analysis and scattering theory.

Calderón's paper "On an inverse boundary value problem"

Gunther Uhlmann*

1. Introduction

An article that made a profound impact in my career is the short paper by Alberto P. Calderón published in 1980 entitled "On an inverse boundary value problem" [1]. In this article Calderón started the mathematical study of the following inverse problem: can one determine the electrical conductivity of a medium by making voltage and current measurements at the boundary? It is known in the community of mathematicians working on inverse problems as "Calderón's inverse problem" or in short "Calderón's problem". This seminal article has led to the development of *complex geometrical optics* that have had many applications [4].



Alberto P. Calderón

It may come as a surprise that Calderón, one of the most distinguished analysts of the 20th century, famous for his work on harmonic analysis and partial differential equations wrote a paper on an inverse problem. Calderón had a degree in engineering in Argentina and he worked in the late 40's in the geophysical research laboratory of the oil company of Argentina "Yacimientos Petrolíferos Fiscales" (YPF). Calderón in his speech accepting the Dr. *Honoris Causa* from the Universidad Autónoma de Madrid in 1997, "Reminiscencias de mi vida matemática" [2], mentioned his appreciation for the applications of mathematics:

> [...] *Estoy de acuerdo con el dicho de que las matemáticas son la reina de las ciencias, y además creía que toda buena reina debe servir a sus súbditos.*

A literal translation is:

> [...] *I agree with the saying that mathematics is the queen of science, and further I believed that every good queen must serve her subjects.*

*gunther@math.washington.edu.

He apparently thought of the inverse problem when he worked at YPF but did not publish his result until many years later. The motivation of Calderón was oil prospection. For different uses of electrical methods in geophysics see [5], for a recent study of the conductivity distribution underneath Yellowstone see

`http://www.sciencedaily.com/releases/2011/04/110411083533.htm`

In an interesting remark in "Reminiscencias de mi vida matemática", he said, referring to his work at YPF:

> *Como dije, el trabajo era muy interesante, pero me trataron mal. Pero, como veremos, esto fue para mi bien. Si me hubiesen tratado de otro modo, es casi seguro que me habría quedado allí el resto de mi vida activa. En cambio, renuncié.*

A literal translation is:

> *As I said my work was very interesting but I was not well treated. However, as we shall see this was good for me. If I would have been treated better is almost certain that I would have remained there for the rest of my active life. Instead, I resigned.*

And then his brilliant mathematical career started.

Calderón's problem is a form of tomography also known as *Electrical Impedance Tomography* (EIT). In this method one is attempting to determine the conductivity, a different property of a medium than other forms of tomography. X-ray tomography for instance attempts to determine the density of tissue by probing it with X-rays.

EIT also arises in medical imaging given that human organs and tissues have quite different conductivities. One potential application is the early diagnosis of breast cancer [6]. The conductivity of a malignant breast tumor is typically 0.2 mho which is significantly higher than normal tissue which has been typically measured at 0.03 mho. See the book [3] for other medical imaging applications of EIT.

We now describe more precisely the mathematical problem proposed by Calderón. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain with smooth boundary. The electrical conductivity of Ω is represented by a bounded and positive function $\gamma(x)$. In the absence of sinks or sources of current the equation for the potential is given by

$$\nabla \cdot (\gamma \nabla u) = 0 \text{ in } \Omega \tag{1}$$

since, by Ohm's law, $\gamma \nabla u$ represents the current flux.

> By A.P. Calderon
>
> In this note we shall diseuss the following problem. Let D be a bounded domain in $\mathbb{R}^n$, $n \geq 2$, with Lipschitzian boundary dD, and γ be a real bounded measurable function in D with a positive lower bound. Consider the differential operator
>
> $$L_\gamma(W) = \nabla.(\gamma\nabla W)$$
>
> acting on functions of $H^1(D)$ and the quadratic form $Q_\gamma(\phi)$, where the functions ϕ are restrictions to dD of functions in $H^1(\mathbb{R}^n)$, defined by
>
> $$Q_\gamma(\phi) = \int_D \gamma(\nabla W)^2 dx \;, \; W \in H^1(\mathbb{R}^n), \; W\big|_{dD} = \phi$$
>
> $$L_\gamma W = \nabla.(\gamma\nabla W) = 0 \text{ in } D.$$
>
> The problem is then to decide whether γ is uniquely determined by Q_γ and to calculate γ in terms Q_γ, if γ is indeed determined by Q_γ.

Given a potential $f \in H^{\frac{1}{2}}(\partial\Omega)$ on the boundary the induced potential $u \in H^1(\Omega)$ solves the Dirichlet problem

$$\begin{aligned} \nabla \cdot (\gamma \nabla u) &= 0 \text{ in } \Omega, \\ u\big|_{\partial\Omega} &= f. \end{aligned} \tag{2}$$

The Dirichlet to Neumann map, or voltage to current map, is given by

$$\Lambda_\gamma(f) = \left(\gamma \frac{\partial u}{\partial \nu}\right)\Big|_{\partial\Omega}, \tag{3}$$

where ν denotes the unit outer normal to $\partial\Omega$.

The inverse problem is to determine γ knowing Λ_γ. It is difficult to find a systematic way of prescribing voltage measurements at the boundary to be able to find the conductivity. Calderón took instead a different route.

Using the divergence theorem we have

$$Q_\gamma(f) := \int_\Omega \gamma |\nabla u|^2 dx = \int_{\partial\Omega} \Lambda_\gamma(f) f \, dS\,, \tag{4}$$

where dS denotes surface measure and u is the solution of (2). In other words $Q_\gamma(f)$ is the quadratic form associated to the linear map $\Lambda_\gamma(f)$, and to know $\Lambda_\gamma(f)$ or $Q_\gamma(f)$ for all $f \in H^{\frac{1}{2}}(\partial\Omega)$ is equivalent. $Q_\gamma(f)$ measures the energy needed to maintain the potential f at the boundary. Calderón's point of view is that if one looks at $Q_\gamma(f)$ the problem is changed to finding enough solutions $u \in H^1(\Omega)$ of the equation (1) in order to find γ in the interior. We will explain this approach further in the next section where we describe Calderón's paper on the linearization of the map

$$\gamma \xrightarrow{Q} Q_\gamma. \tag{5}$$

Here we consider Q_γ as the bilinear form associated to the quadratic form (4).

We now describe Calderón's paper and how he used complex exponentials to prove that the linearization of (5) is injective at constant conductivities. He also gave an approximation formula to reconstruct a conductivity which is, a priori, close to a constant conductivity.

2. Calderón's paper

Calderón proved in [1] that the map Q is analytic. The Fréchet derivative of Q at $\gamma = \gamma_0$ in the direction h is given by

$$dQ|_{\gamma=\gamma_0}(h)(f,g) = \int_\Omega h\nabla u \cdot \nabla v \, dx \tag{6}$$

where $u, v \in H^1(\Omega)$ solve

$$\begin{cases} \nabla \cdot (\gamma_0 \nabla u) = \nabla \cdot (\gamma_0 \nabla v) = 0 \text{ in } \Omega \\ u\Big|_{\partial\Omega} = f \in H^{\frac{1}{2}}(\partial\Omega), \quad v\Big|_{\partial\Omega} = g \in H^{\frac{1}{2}}(\partial\Omega). \end{cases} \tag{7}$$

So the linearized map is injective if the products of $H^1(\Omega)$ solutions of $\nabla \cdot (\gamma_0 \nabla u) = 0$ is dense in, say, $L^2(\Omega)$.

Calderón proved injectivity of the linearized map in the case γ_0 = constant, which we assume for simplicity to be the constant function 1. The question is reduced to whether the product of gradients of harmonic functions is dense in, say, $L^2(\Omega)$.

Calderón took the following harmonic functions

$$u = e^{x\cdot\rho}, \quad v = e^{-x\cdot\overline{\rho}} \tag{8}$$

where $\rho \in \mathbb{C}^n$ with

$$\rho \cdot \rho = 0. \tag{9}$$

We remark that the condition (9) is equivalent to the following

$$\begin{aligned} &\rho = \frac{\eta + ik}{2}, \eta, k \in \mathbb{R}^n, \\ &|\eta| = |k|, \eta \cdot k = 0. \end{aligned} \tag{10}$$

Then plugging the solutions (8) into (6) we obtain if $dQ|_{\gamma_0=1}(h) = 0$

$$|k|^2(\chi_\Omega h)^\wedge(k) = 0 \qquad \forall\, k \in \mathbb{R}^n\,,$$

where χ_Ω denotes the characteristic function of Ω and $^\wedge$ denotes Fourier transform. Then we conclude by the Fourier inversion formula that $h = 0$ on Ω. However, one cannot apply the implicit function theorem to conclude that γ is invertible near a constant since conditions on the range of Q that would allow use of the implicit function theorem are either false or not known.

Calderón also observed that using the solutions (8) one can find an approximation for the conductivity γ if

$$\gamma = 1 + h \tag{11}$$

and h small enough in the L^∞ norm.

We are given

$$G_\gamma = Q_\gamma\left(e^{x\cdot\rho}\Big|_{\partial\Omega}, e^{-x\cdot\overline{\rho}}\Big|_{\partial\Omega}\right)$$

with $\rho \in \mathbb{C}^n$ as in (2.4). Now

$$\begin{aligned} G_\gamma = &\int_\Omega (1+h)\nabla u\cdot\nabla v\,dx \\ &+\int_\Omega h(\nabla\delta u\cdot\nabla v + \nabla u\cdot\nabla\delta v)\,dx + \int_\Omega (1+h)\nabla\delta u\cdot\nabla\delta v\,dx \end{aligned} \tag{12}$$

with u, v as in (8) and

$$\begin{aligned} \nabla\cdot(\gamma\nabla(u+\delta u)) &= \nabla\cdot(\gamma\nabla(v+\delta v)) = 0 \text{ in } \Omega \\ \delta u\Big|_{\partial\Omega} &= \delta v\Big|_{\partial\Omega} = 0. \end{aligned} \tag{13}$$

Now standard elliptic estimates applied to (13) show that

$$\|\nabla\delta u\|_{L^2(\Omega)}, \quad \|\nabla\delta v\|_{L^2(\Omega)} \le C\|h\|_{L^\infty(\Omega)}|k|e^{\frac{1}{2}r|k|} \tag{14}$$

for some $C > 0$ where r denotes the radius of the smallest ball containing Ω.

Plugging u, v into (2.7) we obtain

$$\widehat{\chi_\Omega\gamma}(k) = -2\frac{G_\gamma}{|k|^2} + R(k) = \widehat{F}(k) + R(k) \tag{15}$$

where F is determined by G_γ and therefore known. Using (14), we can show that $R(k)$ satisfies the estimate

$$|R(k)| \le C\|h\|^2_{L^\infty(\Omega)}e^{r|k|}. \tag{16}$$

In other words we know $\widehat{\chi_\Omega\gamma}(k)$ up to a term that is small for k small enough. More precisely, let $1 < \alpha < 2$. Then for

$$|k| \leq \frac{2-\alpha}{r} \log \frac{1}{\|h\|_{L^\infty}} =: \sigma \tag{17}$$

we have

$$|R(k)| \leq C\|h\|^\alpha_{L^\infty(\Omega)} \tag{18}$$

for some $C > 0$.

We take $\widehat{\eta}$ a C^∞ cut-off so that $\widehat{\eta}(0) = 1$, $\text{supp}\widehat{\eta}(k) \subset \{k \in \mathbb{R}^n, |k| \leq 1\}$ and $\eta_\sigma(x) = \sigma^n \eta(\sigma x)$. Then we obtain

$$\widehat{\chi_\Omega\gamma}(k)\,\widehat{\eta}\left(\frac{k}{\sigma}\right) = \frac{-2G_\gamma\gamma}{|k|^2}\,\widehat{\eta}\left(\frac{k}{\sigma}\right) + R(k)\widehat{\eta}\left(\frac{k}{\sigma}\right).$$

Using this we get the following estimate

$$|l(x)| \leq C\|h\|^\alpha_{L^\infty(\Omega)} \left[\log \frac{1}{\|h\|_{L^\infty(\Omega)}}\right]^n \tag{19}$$

where $l(x) = (\chi_\Omega\gamma * \eta_\sigma)(x) - (F * \eta_\sigma)(x)$. Formula (19) gives then an approximation to the smoothed out conductivity, $\chi_\Omega\gamma * \eta_\sigma$, for h sufficiently small.

References

[1] CALDERÓN, A. P.: On an inverse boundary value problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physics (Río de Janeiro, 1980)*, 65–73. Soc. Brasil. Mat., Río de Janeiro, 1980.

[2] CALDERÓN, A. P.: *Reminiscencias de mi vida matemática*. Discurso de investidura de "Doctor Honoris Causa", Universidad Autónoma de Madrid, 117–125. Publicaciones UAM, 1997.

[3] HOLDER, D.: *Electrical Impedance Tomography*. Institute of Physics Publishing, Bristol and Philadelphia, 2005.

[4] UHLMANN, G.: Calderón's problem and electrical impedance tomography. *Inverse Problems*, 25th Anniversary Volume, **25** (2009), 123011 (39pp.)

[5] ZHDANOV, M. S. AND KELLER, G. V.: The geoelectrical methods in geophysical exploration. *Methods in Geochemistry and Geophysics* **31** (1994), Elsevier.

[6] ZOU, Y. AND GUO, Z.: A review of electrical impedance techniques for breast cancer detection. *Med. Eng. Phys.* **25** (2003), 79–90.

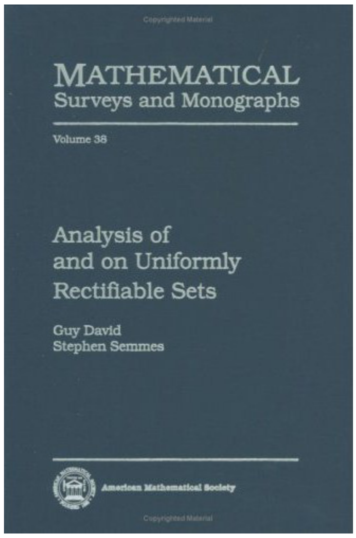
Alexander Volberg (Leningrad, Russia, 1956) obtained his PhD in mathematics at the V. A. Steklov Math. Institute in 1982 (advisor: Nikolai K. Nikolskii). He is currently Distinguished University Professor at Michigan State University. His research focuses on operator theory, complex analysis and harmonic analysis.

Singular integrals survival in bad neighborhoods

Alexander Volberg*

The book of Guy David and Steven Semmes "Analysis of and on uniformly rectifiable sets" is a pioneering text on the behavior of singular integrals on quite bad sets. In fact, it claims that analytic information, expressed in terms of the boundedness of a wide class of singular integrals of Calderón–Zygmund type, can be transferred into geometric information. From such analytic information one could infer very distinctive geometric consequences: it turns out that sets which a priori seem to be very bad cannot be hopelessly bad if such operators are bounded on them; such sets must have a special structure. These "analysis-to-geometry" structural theorems have become now the mainstream of modern "low regularity" Harmonic Analysis. We want to reflect here on certain ideas and problems from this book.

1. Introduction. David–Semmes' problem, its variants



I have selected the book of Guy David and Steven Semmes because it made a big impact on me and because I consider it to be the most important in the area of non-homogeneous harmonic analysis. It made a deep impact on my career because thinking about it has taken me 10 years at least. I am coming back to that stuff repeatedly after those many years. The interesting thing is that the book itself has nothing to do with "non-homogeneous" harmonic analysis. It is completely within the realm of homogeneous Calderón–Zygmund theory as extended by Michael Christ in [1]. In other words, the underlying measure μ, with respect to which the singular integrals are considered in [5], is always doubling. In fact, $\mu = H^s|E$, where E is assumed to be regular in the sense of Ahlfors:

$$c\, r^s \leq H^s(B(x,r)\cap E) \leq C\, r^s\,,\ x \in E, r \leq \operatorname{diam} E$$

*volberg@math.msu.edu.

uniformly. However, the feeling is that this regularity is not really needed and that it can be assumed “without loss of generality”. And this is indeed the case. It is not a simple fact, but it is true that non-homogeneous $T1$ theorems as in [17], [29] –see also the discussion in [6] and below– (often) allow us to reduce the general case to the case of Ahlfors regularity.

In fact, it is interesting how ubiquitous [5] was in my research and the research of my friends. I remember that during Lars Hedberg's conference in 1996, Mark Melnikov asked me whether I knew the description of measures μ for which Cauchy integral is bounded in $L^2(\mu)$. Immediately [5] came to my mind, where this is exactly the question, except that the Cauchy integral is replaced by its direct multi-dimensional analog called Riesz integral operators of singularity s, and μ is $H^s|E$, where $0 < H^s(E) < \infty$. This question of Mark, and the analogy with [5], brought very fruitful outcomes. Two or three years later, such measures were described by Nazarov–Treil–Volberg and Tolsa in [18] and [23], respectively.

These theorems opened the road to subsequent non-homogeneous non-accretive Tb theorems [3, 4, 20, 29], and turned out to be indispensable for Tolsa in his crown achievement [25] of solving Painlevé's problem.

Let me finish this introduction with a joke, which I heard from Peter Jones. He said that “there are exactly four people who have read [5]. These are: Guy David, Steven Semmes, him (Peter Jones)…and somebody else”. As the reader knows now, [5] had a deep impression on “somebody else”.

Let E be a compact set in $\mathbb{R}^d$ such that $0 < H^s(E) < \infty$, $0 < s \le d$. Let $R^s = (R^{1,s}, \dots, R^{d,s})$ be the vector Riesz kernel of singularity s: namely, $R^{i,s} = x_i/|x|^{1+s}$, $i = 1, \dots, d$. The question that David–Semmes posed in the book [5] is the following:

> let E be Ahlfors regular, that is,
>
> $$c\, r^s \le H^s(B(x,r) \cap E) \le C\, r^s\,,\ \forall x \in E, 0 < r < \operatorname{diam} E\,. \tag{1}$$
>
> Let $R^s : L^2(E, H^s|E) \to L^2(E, H^s|E)$ be bounded, with $s = d - 1$. Is is true that E is uniformly rectifiable?

Uniformly rectifiable means here that

> for all $x \in E$ and $0 < r < \operatorname{diam} E$, there exists a Lipschitz image $\Gamma_{x,r}$ of $\mathbb{R}^{d-1}$ into $\mathbb{R}^d$ with Lipschitz constant independent of x, r such that $H^{d-1}(B(x,r) \cap E \cap \Gamma_{x,r}) \ge c\, r^{d-1}$.

David–Semmes proved this “analysis-to-geometry” result under a stricter assumption, namely, boundedness is required for **all** Calderón–Zygmund operators –not only for Riesz transforms.

Mattila, Melnikov and Verdera [11] answered this question in the affirmative for $d = 2$. But it is known that in the case of the plane there is a miracle: Melnikov's formula, which introduces Menger's curvature as a tool into analysis, see [15, 16, 24]. The higher dimensional version seems to be one of the leading questions now. Let us now elaborate on David–Semmes question and consider several variants and reformulations.

Acknowledgements. 1) I am grateful to Nestor Guillén and Ryan Hynd for very valuable conversations, especially those concerning non-local free boundary problems with gradient constraints. I am also grateful to MSRI program on Free Boundary for the hospitality.

2) All purely mathematical content of this note is joint with Vladimir Eiderman and Fedja Nazarov. However, all mistakes are mine.

2. Variants of David–Semmes question

2.1. Integer $s = n$

The first natural thing to do is to consider $s = n \neq d - 1$, where $n \in \mathbb{Z}_+$ is an integer, $0 < n < d$. Then the question is exactly the same as before, only that uniform rectifiability becomes n-uniform rectifiability, namely, now $\Gamma_{x,r}$ is a Lipschitz image of $\mathbb{R}^n$ into $\mathbb{R}^d$ with Lipschitz constant independent of x, r such that $H^n(B(x,r) \cap E \cap \Gamma_{x,r}) \geq c\, r^n$. It is widely believed that the answer is correct: if the set E, with $0 < H^n(E) < \infty$ is n-Ahlfors regular (see (1) with $s = n$) and if all Riesz transforms of singularity n are bounded in $L^2(E, H^n|E)$, then E is n-uniformly rectifiable. Again, if one assumes that **all** Calderón–Zygmund operators of singularity n are bounded, then the conclusion follows ([5]).

2.2. Integer $s = n$, but Ahlfors regularity (1) is dropped

Let $s = n \leq d - 1$ be integer, but let us drop the assumption of Ahlfors regularity. The conclusion must be obviously altered. Instead of the existence of big pieces of Lipschitz images in all scales, one should hope for just one such Lipschitz image. So the assumption of the boundedness of all Riesz transforms in $L^2(E, H^n|E)$ remains, but the conclusion must be changed to

$$\text{there exists a Lipschitz image } \Gamma \text{ of } \mathbb{R}^n \text{ into } \mathbb{R}^d \\ \text{such that } H^n(E \cap \Gamma) > 0. \tag{2}$$

This follows from [24], combined with [9], for $d = 2, n = 1$. The proofs are quite difficult and they use the ubiquitous Melnikov's formula and Menger's curvature, the tool which, in the the words of Guy David, "is cruelly missing" in $d > 2, s \geq 1; d = 2, s > 1$.

2.3. What if s is not integer?

For non-integer s, $0 < s < d$, there is no "Lipschitz images of $\mathbb{R}^s$ into $\mathbb{R}^d$", because there is no $\mathbb{R}^s$, and there is no good way to express the structural condition on E saying that E has good "Lipschitz smooth" pieces. Therefore, it is natural to think that the Riesz transforms of singularity $s \notin \mathbb{Z}_+$ are not bounded in $L^2(H^s|E)$ if E is such that $0 < H^s(E) < \infty$.

This is actually proved in the case of s-Ahlfors regularity (1) of E by Vihtila [28]. In fact, more is proved in [28]. Instead of imposing a strong estimate from below as in (1), Vihtila in [28] requires only that for H^s-a.e. point $x \in E$ the lower density is strictly positive:

$$\liminf_{r\to 0} \frac{H^s(B(x,r)\cap E)}{r^s} > 0\,. \tag{3}$$

The technique of tangent measures then allows her to prove the non-existence of such sets having bounded Riesz transforms on them. However, dropping (1) and (3) completely seems to represent huge difficulties. We will further elaborate on this. Even the case $d = 2, 1 < s < 2$ is difficult and was open till very recently, see [6]. On the other hand, we already mentioned that the case $d = 2, s = 1$ was solved by Tolsa [24] (see algo Léger's [9]); again, these are very difficult papers. For $d = 2, s < 1$, one can use Prat's paper [21], and again the problem is solved: no such E exists. Here one uses the same Melnikov's approach but for Riesz kernels of singularity $s < 1$. A small miracle –a miracle known to the experts– happens: the symmetrization trick works and gives a positive kernel. As we already mentioned, this is "cruelly" false for $s > 1, d = 2$ and $s \geq 1, d > 2$.

Remark. This note is devoted much more to non-integer $s > 1$ than to the initial question of David–Semmes with $s = d - 1$. But notice that the main difficulty lies in both cases in the "cruel" lack of positivity of kernels after symmetrization, see [7]. Therefore, from now on we want to focus on the following implication of analysis-to-geometry type:

$$\begin{gathered} E \text{ is a compact in } \mathbb{R}^d,\ 0 < H^s(E) < \infty, \text{ such that} \\ R^s : L^2(E, H^s|E) \to L^2(E, H^s|E) \text{ is bounded} \Rightarrow s \text{ is integer}\,. \end{gathered} \tag{4}$$

For $s \in (d-1, d)$ this is proved in [6]. Namely, if s is strictly between $d-1$ and d, then the assumption of the above claim leads to a contradiction independently of the fact that the bounds in (1) hold (then it is Vihtila's case, the technique of tangent measures) or not. We never have boundedness of R^s, $s \in (d-1, d)$, on $L^2(E, H^s|E)$ if $H^s(E) < \infty$!

However, looking at our approach in [6], one can see that it leads to the following claim, which includes integer $s = d-1$ as well:

$$\begin{aligned}&E \text{ is a compact in } \mathbb{R}^d,\ 0 < H^{d-1}(E) < \infty,\ \text{such that}\\ &R^{d-1} : L^2(E, H^{d-1}|E) \to L^2(E, H^{d-1}|E) \text{ is bounded, and}\\ &\liminf_{r\to 0} \frac{H^{d-1}(B(x,r)\cap E)}{r^{d-1}} = 0\ H^{d-1} - a.e. \text{ on } E \Rightarrow \text{ contradiction.}\end{aligned} \tag{5}$$

3. Using $T1$ theorem to reformulate the problem

Let us consider in what sense the boundedness of operators with kernel R^s on $L^2(E,\mu)$ is understood, where μ stands for $\mu := H^s|E$, and where $0 < H^s(E) < \infty$. After all, the kernel R^s is singular with respect to such a measure, so the integral cannot be understood as absolutely convergent. There are two (non-trivially) equivalent senses in which this boundedness is dealt with. In the first sense, boundedness means that for two real smooth functions with compact support ϕ, ψ the form below is bounded by the right hand side with constant independent on ϕ, ψ:

$$\left|\frac{1}{2}\iint R^s(x-y)(\phi(x)\psi(y) - \phi(y)\psi(x))\, d\mu(x)\, d\mu(y)\right| \le C\, \|\phi\|_{L^2(\mu)} \|\psi\|_{L^2(\mu)}\,.$$

This form actually would be equal to the form $(R^s\phi, \psi)$ –just by anti-symmetry of the kernel– if $(R^s\phi,\psi)$ would make sense. But the above form by itself makes perfect sense, because it is absolutely convergent.

Another way to understand the boundedness of the operator with kernel R^s on $L^2(\mu)$ is to regularize it by, *e.g.*, considering

$$R^s_\epsilon \phi(x) := \int_{y:|y-x|\ge\epsilon} R^s(x-y)\phi(y)\, d\mu(y)$$

and requiring that the operators R^s_ϵ are *uniformly* bounded in $L^2(\mu)$.

It is well-known (see, e. g. [18], [29] and many other places) that both versions of boundedness imply that

$$\mu(B(x,r)) \le C\, r^s\,. \tag{6}$$

We denote by $\Sigma_s(E)$ the class of positive measures satisfying (6).

Let us introduce the *maximal singular* operator R^s_*: let σ be a measure on our compact set E, for example (but not necessarily) $\sigma = f\, d\mu$. Then

$$R^s_*\sigma(x) := \left|\sup_{\epsilon>0}\int_{y\in E:|y-x|\ge\epsilon} R^s(x-y)\, d\sigma(y)\right|.$$

We use the notation $R^s_* f(x)$ if $\sigma = f\, d\mu$.

We want to present the following result.

Theorem 1 *Let $E \subset \mathbb{R}^d$ be a compact set and $0 < H^s(E) < \infty$, $0 < s < d$, s being integer or not. As always, let μ denote $H^s|E$. Let $R^s : L^2(\mu) \to L^2(\mu)$ be bounded.*

Then there exists a strictly positive measure $\sigma \in \Sigma_s(E)$ such that $R^s\sigma$ is uniformly bounded in $\mathbb{R}^d \setminus E$ (or we can say, essentially uniformly bounded in $\mathbb{R}^d$). Moreover, $\sigma = h\,d\mu$, where $h \in L^\infty(\mu), h \geq 0$ and $h > 0$ on a subset E' such that $\mu(E') > 0$.

Interestingly, the inverse claim is also true to a large extent. Namely,

Theorem 2 *Let $E \subset \mathbb{R}^d$ be a compact set $1 < s < d$ and $0 < H^s(E) < \infty$, s being integer or not, $\mu = H^s|E$ as always. Let $\sigma \in M_+(E)$ be a strictly positive measure such that $R^s\sigma$ is bounded in $\mathbb{R}^d \setminus E$.*

Then $\sigma \in \Sigma_s(E)$, $R^s : L^2(E,\sigma) \to L^2(E,\sigma)$ is bounded, and there exists a piece $\tilde{E} \subset E$ such that $\mu(\tilde{E}) > 0$ and such that $R^s : L^2(\mu|\tilde{E}) \to L^2(\mu|\tilde{E})$ is bounded.

Theorem 2 is proved in [20, 29], see also [17]. It is a very difficult theorem. A slightly more sophisticated variant (when σ is allowed to be complex-valued) led Nazarov–Treil–Volberg to a solution of Denjoy's problem (which was different from the solution in David–Mattila's [4] and David's [3]) and led Tolsa to the final solution of Painlevé's problem [25, 27].

On the other hand, the proof of Theorem 1 is not that difficult.

Proof of Theorem 1. We will restrict ourself to the case $s \geq d-1$, as the main goal is to show the use of the maximal principle. The first step is non-homogeneous Harmonic Analysis. The Calderón–Zygmund operator R^s on the metric space with measure (E,μ) is bounded in $L^2(E,\mu)$. If the measure is doubling we can invoke Coifman–Weiss argument [2] to deduce that

$$R^s : L^1(\mu) \to L^{1,\infty}(\mu)\,. \tag{7}$$

For non-doubling measures (but satisfying $\mu \in \Sigma_s(E)$) the work [19] claims that (7) holds as well. Let us imagine for a second that in the right hand side above we also have $L^1(\mu)$. Then by duality (and the fact that the adjoint verifies $(R^s)' = -R^s$) we would have that *for any function $h \in L^\infty(\mu)$* one would have that $R^s h := (R^s)(h\,d\mu)$ is in $L^\infty(\mu)$. An amazing fact, see Christ's paper [1], is that if we, however, have $L^{1,\infty}(\mu)$, as it is the case in the relationship above, then *there exists at least one non-negative, non-zero* function $h \in L^\infty(\mu)$ such that $R^s h \in L^\infty(\mu)$! Therefore, for this h,

$$|R^s h(x)| \leq C_1 \qquad \text{for } \mu - \text{a. e. point } x \in E\,. \tag{8}$$

This inequality is not enough to use a maximal principle we are heading to. So we invoke [19], where it is proved that (8) implies

$$|R^s_* h(x)| \le C_2 \qquad \forall x \in E\,. \tag{9}$$

Now we use Lemma 3 of [28] to conclude that $R^s h(x)$ is bounded on $\mathbb{R}^d$. ∎

4. Maximum principle for fractional Laplacian

Let us consider $\alpha \in (\frac{1}{2}, 1]$, $s = d - 2\alpha + 1 \in [d-1, d)$, and a compact set E such that $H^{d-2\alpha+1}(E) \in (0, \infty)$. For $u \in S(\mathbb{R}^d)$ (Schwartz class) we have

$$L_\alpha u := (-\Delta)^\alpha u = (|\xi|^{2\alpha}\hat{u}(\xi))\check{}\,.$$

There is a nice formula for $L_\alpha u$ in [8]:

$$L_\alpha u(x) = c(d,\alpha) \int \frac{u(x+y) + u(x-y) - 2u(x)}{|y|^{d+2\alpha}}\, dy\,. \tag{10}$$

Here $c(d,\alpha) \asymp 1-\alpha$ as $\alpha \to 1$.

This holds not only if u is a Schwartz function; it holds, for example, if u is a potential of a measure lying on our compact E (or even the derivative of such a potential):

$$v(x) = P_\alpha \sigma := c \int_E \frac{1}{|x-y|^{d-2\alpha}}\, d\,\sigma(x)\,,$$

which is the solution of the distributional equation

$$L_\alpha v = \sigma\,.$$

Definition. If $L_\alpha v = 0$ in an open set Ω, we say that v is α-harmonic in Ω.

Formula (10) is a corollary of an amazing equality to be found in [8]: For $\alpha \in (\frac{1}{2}, 1)$,

$$\int \left(\frac{1}{|x+y|^{d-2\alpha}} + \frac{1}{|x-y|^{d-2\alpha}} - \frac{2}{|x|^{d-2\alpha}} \right) \frac{1}{|y|^{d+2\alpha}}\, dy = 0, \qquad x \neq 0\,.$$

So for a potential v as above,

$$\int \frac{v(x+y) + v(x-y) - 2v(x)}{|y|^{d+2\alpha}}\, dy = 0 \qquad \forall x \in \mathbb{R}^d \setminus E\,. \tag{11}$$

Let σ be a measure absolutely continuous with respect to Lebesgue measure m_d and with bounded density. Let $u := (\nabla v)_i, i = 1, \ldots, d$, where v is as above. Then we conclude that

$$\int \frac{u(x+y) + u(x-y) - 2u(x)}{|y|^{d+2\alpha}}\, dy = 0 \quad \forall x \in \mathbb{R}^d \setminus E\,. \tag{12}$$

We want to deduce from this that

$$\int \frac{u^2(x+y) + u^2(x-y) - 2u^2(x)}{|y|^{d+2\alpha}}\, dy \geq 0 \quad \forall x \in \mathbb{R}^d \setminus E\,. \tag{13}$$

If (13) were proved then assuming that $u = \partial_i v$ has a maximum at x_0 outside of the support of σ we would get

$$0 \leq \int \frac{u(x_0+y) + u(x_0-y) - 2u(x_0)}{|y|^{d+2\alpha}}\, dy \leq 0\,,$$

meaning that u is constant m_d-almost everywhere. But it goes to zero at infinity. So we would conclude that u is identically equal to zero.

Now let σ be a finite positive measure with compactly supported C^∞ density with respect to m_d. Applying the above to $u = R_i^s \sigma = \partial_i P_\alpha \sigma$ we would come to a contradiction, as it is not at all identically zero near infinity.

Therefore, we have established that, if $R := R^s$ and $s \in (d-1, d)$,

$$\max_{\mathbb{R}^d} |R\sigma| = \max_{\operatorname{supp}\sigma} |R\sigma|\,. \tag{14}$$

Also, clearly $\max_{\mathbb{R}^d} R_i^s \sigma \geq 0$ (since it is zero at infinity). So if this maximum is strictly positive it should be attained at some point $x \in \mathbb{R}^d$. Using formula (12) we get that

$$\max_{\mathbb{R}^d} R_i \sigma = \max_{\operatorname{supp}\sigma} R_i \sigma\,, \tag{15}$$

if the left hand side is strictly positive (again $R := R^s$, $s \in (d-1, d)$).

To prove (13) notice that

$$\begin{aligned}
\int &\frac{u^2(x+y) + u^2(x-y) - 2u^2(x)}{|y|^{d+2\alpha}}\, dy = \\
&= 2u(x) \int \frac{u(x+y) + u(x-y) - 2u(x)}{|y|^{d+2\alpha}}\, dy \\
&\quad + \int \frac{(u(x+y) - u(x))^2}{|y|^{d+2\alpha}}\, dy + \int \frac{(u(x-y) - u(x))^2}{|y|^{d+2\alpha}}\, dy\,,
\end{aligned}$$

So at a point x such that

$$\int \frac{u(x+y) + u(x-y) - 2u(x)}{|y|^{d+2\alpha}}\, dy = 0,$$

we must have

$$\int \frac{u^2(x+y)+u^2(x-y)-2u^2(x)}{|y|^{d+2\alpha}}\,dy = \\ = \int \frac{(u(x+y)-u(x))^2}{|y|^{d+2\alpha}}\,dy + \int \frac{(u(x-y)-u(x))^2}{|y|^{d+2\alpha}}\,dy \geq 0\,,$$

which is exactly (13).

Let R' denote the formal adjoint of $R = R^s$ (namely, $R'\eta := \sum_{i=1}^d R_i\eta_i$ for any d-tuple of measures), then in [6] the following variant of the maximal principle (14) was essentially used: Let σ and s be as above. Then,

$$\max_{\mathbb{R}^d} \left[|R\sigma| + R'((R\sigma)\,\sigma)\right] = \max_{\operatorname{supp}\sigma} \left[|R\sigma| + R'((R\sigma)\,\sigma)\right], \tag{16}$$

provided that the left hand side is positive.

We can now say –alas only very schematically and not at all accurately– how in [6] the relationship (16) was applied. It was applied to σ essentially minimizing the functional

$$\int |R\Sigma|^2\, d\Sigma \to \text{minimize}$$

under conditions like $\frac{1}{2}\mu \leq \Sigma \leq 2\mu$, $\Sigma(\mathbb{R}^d) = \mu(\mathbb{R}^d)$, where μ was a nice measure for which we wish to obtain the estimate from below for $\int |R\mu|^2\, d\mu$ (see the true statement in [6]). In fact, if the minimum is not too small (in terms of parameters depending on μ), then we are done. If it is small, then the extremal measure σ is proved to have the property that $[|R\sigma| + R'((R\sigma)\,\sigma)]$ is sufficiently small uniformly on the support of σ (extremal measures have certain structure, see [24] or [29, pp. 33–36]). Then the maximum over $\mathbb{R}^d$ is also estimated by (16). But for $R\Sigma$ with Σ in our class of measures, there is a weighted estimate from below of $\int |R\Sigma|^2\, \Psi\, dm_d$ in terms of parameters of μ (Ψ depends on μ itself, see its precise description in [6]). This brings the contradiction.

5. A theorem related to (4) in terms of fractional harmonic functions

Let $\alpha \in (\frac{1}{2}, 1)$. A function u, Lipschitz in $\mathbb{R}^d$ and zero at infinity, is α-harmonic in $\mathbb{R}^d \setminus E$ if it is a potential

$$u = c\frac{1}{|x|^{d-2\alpha}} \star S,$$

where S is a distribution with compact support in E.

After playing a little bit with the theory of Nazarov–Treil–Volberg [17, 20, 29], the claim (4) for $s = d - 2\alpha + 1 \in (d-1, d)$ becomes exactly

$$\begin{array}{l}\text{Let } E \text{ be a compact set in } \mathbb{R}^d \text{ such that } H^{d-2\alpha+1}(E) < \infty. \\ \text{Then there exists no function } u, \text{ Lipschitz in } \mathbb{R}^d \text{ and} \\ \alpha\text{-harmonic in } \mathbb{R}^d \setminus E \text{ which is a potential } u = c\frac{1}{|x|^{d-2\alpha}} \star \nu \\ \text{of a positive measure } \nu \text{ on } E.\end{array} \tag{17}$$

So the function u (which we want to show can be only a zero function) has the properties:

1) u is Lipschitz.
2) u is α-harmonic in $\mathbb{R}^d \setminus E$.
3) u is α-superharmonic: $(-\Delta)^\alpha u = \mu \geq 0$.

Notice that in the previous sections we have reduced problem (4) to the existence of such an α-harmonic, Lipschitz function

$$u = \frac{1}{|x|^{d-2\alpha}} \star (h\, dH^s|E),$$

which, because of its α-harmonicity in $\mathbb{R}^d \setminus E$ and lipschitzness in $\mathbb{R}^d$, it is also a potential of a positive measure $\nu = h\, dH^s|E$. On the other hand, the existence of non-zero positive ν (and so of its potential u) in (17) would give a measure from (4).

For $\alpha = \frac{1}{2}, s = d - 1$, such functions do exist, but **only if** E has a Lipschitz piece –this is yet another form of David–Semmes' conjecture. The **only if** part is open.

Requirement 3) on u appears naturally, but equally natural is the question whether a non-zero function u exists satisfying only requirements 1) and 2).

If such a function exists, and $H^s(E) < \infty$, then one can show that

$$u = \frac{1}{|x|^{d-2\alpha}} \star \nu$$

with ν a *complex* satisfying

$$d\nu = b\, dH^s|E, \qquad \|b\|_{L^\infty(E, H^s|E)} < \infty.$$

Then Nazarov–Treil–Volberg's "restricted" Tb-theorem [20, 29] for non-homogenous measures shows that there is a piece $E' \subset E$, $H^s(E') > c\, H^s(E)$, such that

$$R^s : L^2(E', H^s|E') \to L^2(E', H^s|E')$$

is a bounded operator. We are back to (4), which, as we just saw, can be reduced to the existence of u with 1), 2), and 3)! Hence, we have reduced the question of the existence of non-trivial Lipschitz, α-harmonic outside of E function, to the question of the existence of such a function with the extra property of being α-subharmonic. This is (17). We know that there is no such non-trivial function if $s \in (d-1, d)$, see [6]. One should prove that there is no such non-trivial function for all non-integer $s \in (0, d)$.

6. Related conjectures in terms of singular maximal function $R^s_* \mu$

Let $E \subset \mathbb{R}^d$ be a compact set and $0 < H^s(E) < \infty$, $0 < s < d$, s being integer or not, $\mu = H^s|E$ as always.

We want to mention that Nazarov–Treil–Volberg theory [17, 20, 29] implies that the problem of David–Semmes –even in a non-homogeneous situation (no (1))– and for any $s \in (0, d)$ has an

Equivalent formulation: Let E, μ be as above, $\mu \in \Sigma_s(E)$, let $h \in L^\infty(\mu), h \geq 0, h \neq 0$, and let $R^s_*(h\, d\mu) < \infty$ μ-a. e. Then s must be an integer and E must have a piece of Lipschitz image of $\mathbb{R}^s$ into $\mathbb{R}^d$.

Even for $h = 1$ this is a problem, and even for E's having the extra property (1). And even for one special $s = d - 1$. However, as we already mentioned above, this has been proved for $s \in (d-1, d)$, see [6].

On the other hand, there is a great progress if the assumption

$$\sup_\epsilon |R^s_\epsilon \mu(x)| =: R^s_*(d\mu)(x) < \infty \text{ for } \mu - \text{a.e } x \in E \tag{18}$$

is replaced by the seemingly only slightly stronger assumption (in fact it turns out to be *much* stronger):

$$\exists \lim_{\epsilon \to 0} R^s_\epsilon \mu \text{ for } \mu - \text{a.e } x \in E\,. \tag{19}$$

Probably the first such results were those of Mattila [15] and Mattila–Preiss [12]. In the latter paper, given the regularity (3) and the existence of principal values, the authors proved that s must be an integer and E must be s-rectifiable (the former paper is devoted to the case $d = 2, s = 1$, which is equivalent to the existence of principal value of the Cauchy integral; (3) is assumed). Getting rid of regularity condition is tough. Only the existence of principal values allowed Ruiz de Villa and Tolsa [22] and Tolsa [26] to prove that s is integer and to prove the existence of s-Lipschitz piece in E (s-rectifiability).

There are subsequent papers, [13, 14], where the existence of principal values is replaced by the finiteness of square function –with the same conclusions. But this is still far from the finiteness of maximal singular function as stated in (19), which remains the heart of the matter in all cases where we do not have the reduction to Menger's curvature, namely $d > 2, s \geq 1$ and $d = 2, s > 1$. In [6] the case $s \in (d-1, d)$ has been settled.

References

[1] Christ, M.: A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.* **60/61** (1990), no. 2, 601-628.

[2] Coifman, R. and Weiss, G.: *Analysis on homogeneous metric spaces.*

[3] David, G.: Unrectifiable 1-sets have vanishing analytic capacity. *Rev. Mat. Iberoamericana* **14** (1998), no. 2, 369–479.

[4] David, G. and Mattila, P.: Removable sets for Lipschitz harmonic functions in the plane. *Rev. Mat. Iberoamericana* **16** (2000), no. 1, 137–215.

[5] David, G. and Semmes, S.: *Analysis of and on uniformly rectifiable sets.* Mathematical surveys and monographs **38**. AMS, Providence, RI, 1993.

[6] Eiderman, V., Nazarov, F. and Volberg, A.: The non-existence of totally lower irregular s-dimensional measures with bounded Riesz transforms on the plane for $1 < s < 2$. Preprint, April 2011, 1–20.

[7] Farag, H. M.: The Riesz kernels do not give rise to higher-dimensional analogues of the Menger-Melnikov curvature. *Publ. Mat.* **43** (1999), no. 1, 251–260.

[8] Landkof, N. S.: *Foundations of modern potential theory.* Translated from the Russian by A. P. Dohovskoy. Die Grundlehren der mathematischen Wissenschaften **180**. Springer-Verlag, New York-Heidelberg, 1972. x+424 pp.

[9] Léger, J.-C.: Menger curvature and rectifiability. *Ann. of Math. (2)* **149** (1999), no. 3, 831–869.

[10] Mattila, P.: Cauchy singular integrals and rectifiability in measures of the plane. *Adv. Math.* **115** (1995), no. 1, 1–34.

[11] Mattila, P., Melnikov, M. and Verdera, J.: The Cauchy integral, analytic capacity, and uniform rectifiability. *Ann. of Math. (2)* **144** (1996), no. 1, 127–136.

[12] Mattila, P. and Preiss, D.: Rectifiable measures in Rn and existence of principal values for singular integrals. *J. London Math. Soc. (2)* **52** (1995), no. 3, 482–496.

[13] Mayboroda, S. and Volberg, A.: Finite square function implies integer dimension. *C. R. Math. Acad. Sci. Paris* **347** (2009), no. 21-22, 1271–1276.

[14] Mayboroda, S. and Volberg, A.: Boundedness of the square function and rectifiability. *C. R. Math. Acad. Sci. Paris* **347** (2009), no. 17-18, 1051–1056.

[15] Melnikov, M.: Analytic capacity: a discrete approach and the curvature of measure. (Russian) *Mat. Sb.* **186** (1995), no. 6, 57–76; translation in *Sb. Math.* **186** (1995), no. 6, 827–846 .

[16] Melnikov, M. and Verdera, J.: A geometric proof of the L^2 boundedness of the Cauchy integral on Lipschitz graphs. *Internat. Math. Res. Notices* (1995), no. 7, 325–331.

[17] Nazarov, F., Treil, S. and Volberg, A.: The *Tb*-theorem on nonhomogeneous spaces. *Acta Math.* **190** (2003), 151–239.

[18] Nazarov, F., Treil, S. and Volberg, A.: Cauchy integral and Calderón–Zygmund operators on nonhomogeneous spaces. *Int. Math. Res. Notices* (1997), no. 15, 703–726.

[19] Nazarov, F., Treil, S. and Volberg, A.: Weak type estimates and Cotlar inequalities for Calderón–Zygmund operators on nonhomogeneous spaces. *Internat. Math. Res. Notices* (1998), no. 9, 463–487.

[20] Nazarov, F., Treil, S. and Volberg, A.: The Tb-theorem on non-homogeneous spaces that proves a conjecture of Vitushkin (How to pull ourselves up by the hair). CRM preprints, No. 519, pp. 1–84. `http://www.crm.es/Publications/02/Pr519.pdf`.

[21] Prat, L.: Potential theory of signed Riesz kernels: capacity and Hausdorff measure. *Int. Math. Res. Notices* (2004), no. 19, 937–981.

[22] Ruiz de Villa, A. and Tolsa, X.: Non existence of principal values of signed Riesz transforms of non integer dimension. *Indiana Univ. Math. J.* **59** (2010), no. 1, 115–130.

[23] Tolsa, X.: L^2-boundedness of the Cauchy integral operator for continuous measures. *Duke Math. J.* **98** (1999), no. 2, 269–304.

[24] Tolsa, X.: Growth estimates for Cauchy integrals of measures and rectifiability. *Geom. Funct. Anal.* **17** (2007), no. 2, 605–643.

[25] Tolsa, X.: Painlevé's problem and the semiadditivity of analytic capacity. *Acta Math.* **190** (2003), no. 1, 105–149.

[26] Tolsa, X.: Principal values for Riesz transforms and rectifiability. *J. Funct. Anal.* **254** (2008), no. 7, 1811– 1863.

[27] Tolsa, X.: Analytic capacity, rectifiability, and the Cauchy integral. In *International Congress of Mathematicians. Vol. II*, 1505–1527. Eur. Math. Soc., Zürich, 2006.

[28] Vihtila, M.: The boundedness of Riesz s-transforms of measures in $\mathbb{R}^n$. *Proc. Amer. Math. Soc.* **124** (1996), no. 12, 3797–3804.

[29] Volberg, A.: *Calderón–Zygmund capacities and operators on non-homogeneous spaces.* CBMS **100**, 2003, 165p.

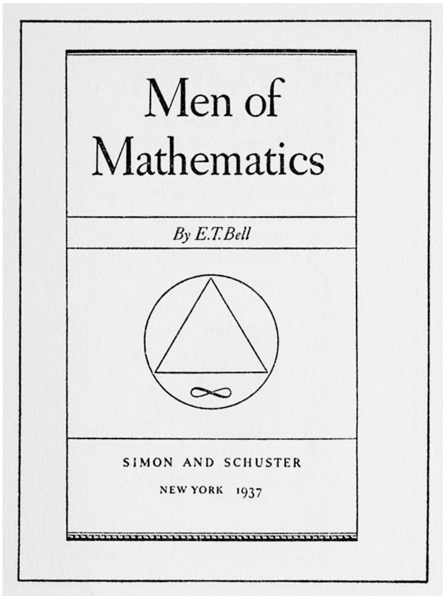
Stephen Wainger (Schenectady, New York, 1936) obtained his PhD in mathematics at the University of Chicago in 1961 (advisor: Elias M. Stein), and he is presently Professor Emeritus at the University of Wisconsin-Madison. His main research area is harmonic analysis.

Some recollections

Stephen Wainger*

When I was a teenager, I came across several books written by Eric Temple Bell. While I was already interested in mathematics, I became fascinated with these treatises, especially "Men of Mathematics".

I found engrossing both the biographies of these great mathematicians and the little of the mathematics that I could understand at the time.



As for the historical information, I have heard that "Men of Mathematics" contains what Mark Twain called "stretchers" (stretchings of the truth). But it certainly made good reading. Perhaps two of the most interesting biographies were those of Abel and Galois. Gauss threw away Abel's proof that the general fifth degree equation can not be solved by radicals. Also Cauchy seems to have neglected the work of both Abel and Galois on this subject. Abel was the luckier of the two. He was befriended by a Norwegian mathematician called Bernt Michael Holmboë and later by Crelle. Basically Galois' genious was neglected. He failed the entrance examinations for the École Polytechnique twice. Basically his talents were neglected although eventually Galois met a man Louis Paul Emile Richard, who recognize Galois's genius.

On the mathematical side there was much that I could not understand at all –such as the beginnings of group theory or the theory of elliptic integrals. On the other hand there were items that I found truly amazing. I had studied Euclidean geometry, and I knew how to construct a tangent line to a circle through a point on the circle. I remember being thunderstruck by the procedure of finding the tangent line to an "arbitrary" curve through a point on the curve. In fact I learned the basic idea of both differential and integral calculus from "Men of Mathematics" –though I could actually do only the simplest– if any of the problems that appearing a standard text. I have from time to time wondered if I had learned calculus the traditional

*wainger@math.wisc.edu.

way –learning the “rules”– whether I would have appreciated the subject as much as I have.

I also learned about non-Euclidean Geometry, the contributions of Cantor, and the impossibility of solving the general fifth degree equation by radicals from this book. Of course, Fermat’s conjecture is also there.

My recollection is that I also learned there the statement of the prime number theorem, but I couldn’t find it in the book later; although in the article on Riemann, Bell mentions the zeta function and the fact that it was used to study the number of primes less than x. Perhaps I read the statement of the prime number theorem in another of Bell’s books. For example, I found a statement of this theorem in his “Mathematics, Queen and Servant of Sciences”.

Looking back at “Men of Mathematics” I discover a few other surprises besides not finding a statement of the prime number theorem. I notice now that the statement that every integer is the sum of four squares is in that book. I don’t remember seeing it there when I was young. In “Men of Mathematics”, the statement of this result appears in the article on Jacobi, and I missed it. I might have skipped over much of the article on Jacobi because of the discussion of elliptic integrals –which of course I could not understand at all. (I first learned of this theorem in a year long course in complex variables which I took during my junior year in college. I remember being hardly able to believe it.)

On the historical side, I was surprised to see the modern beginnings of differential calculus was attributed to Fermat rather than to Newton or Leibniz. Another surprise was about the beginning of group theory. According to Bell, group theory had its origins in the work of Cauchy and Cayley.

One final surprise. Bell asserts that when Abraham Lincoln was a young trial lawyer in Illinois, he spent many nights studying Euclid in order to improve his skill in giving logical arguments. And it might be argued that the logic in Abraham Lincoln’s speech at The Cooper Union in New York City, April 27, 1860, propelled him from a little known midwesterner to the Republican candidate for president in 1860, and further that the logic of his arguments in the Lincoln Douglas debates played an important role in his winning of the presidency in 1860.

Twice I found distant cousins who showed an interest in mathematics, and I bought copies of “Men of Mathematics” for each of them. One works now with computers. I think that the other is still in high school, but her grandfather told me that he loved the book.

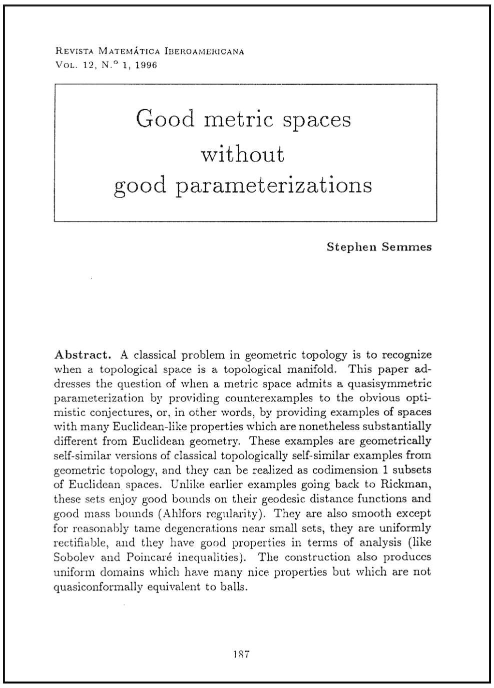
Jang-Mei Wu (Hangchow, China, 1948) obtained her BS at the National Taiwan University in 1970, and her PhD in mathematics at the University of Illinois in 1974 under the direction of Maurice H. Heins. She is Professor at the University of Illinois, Urbana-Champaign. Her research focuses on geometric analysis and potential theory.

Semmes spaces

Jang-Mei Wu*

To understand the underlying structure of a metric space, one seeks a parametrization of a special type. For example, every Riemannian manifold homeomorphic to the 2-sphere is conformally equivalent to $\mathbb{S}^2$.

In his 1996 Revista Matemática Iberoamericana papers [12, 13], Stephen Semmes gave unexpected counterexamples to several natural conjectures on the bilipschitz and quasisymmetric parametrizations of metric n-spheres. His examples are geometrically self-similar manifolds modeled on the decomposition spaces associated with the Whitehead continuum, Bing's dogbone, or Bing's double; these spaces admit metrics that are smooth Riemannian outside a totally disconnected closed set and, in some sense, indistinguishable from the standard metric on $\mathbb{S}^3$ geometrically and measure theoretically, and yet are not quasisymmetrically equivalent to $\mathbb{S}^3$.

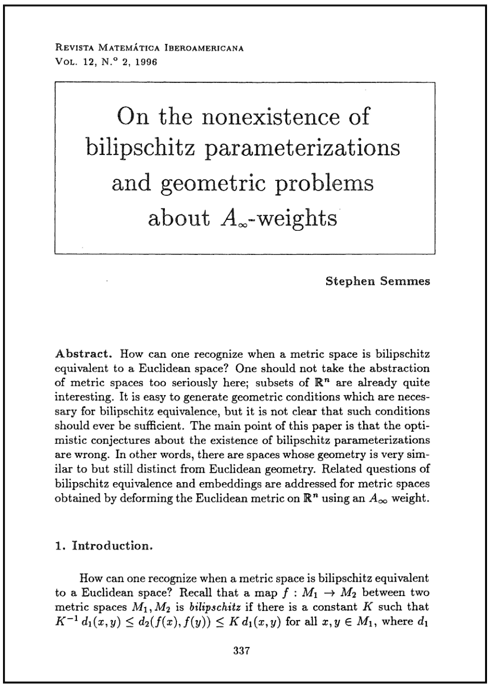

Revista Matemática Iberoamericana
Vol. 12, N.° 1, 1996

Good metric spaces without good parameterizations

Stephen Semmes

Abstract. A classical problem in geometric topology is to recognize when a topological space is a topological manifold. This paper addresses the question of when a metric space admits a quasisymmetric parameterization by providing counterexamples to the obvious optimistic conjectures, or, in other words, by providing examples of spaces with many Euclidean-like properties which are nonetheless substantially different from Euclidean geometry. These examples are geometrically self-similar versions of classical topologically self-similar examples from geometric topology, and they can be realized as codimension 1 subsets of Euclidean spaces. Unlike earlier examples going back to Rickman, these sets enjoy good bounds on their geodesic distance functions and good mass bounds (Ahlfors regularity). They are also smooth except for reasonably tame degenerations near small sets, they are uniformly rectifiable, and they have good properties in terms of analysis (like Sobolev and Poincaré inequalities). The construction also produces uniform domains which have many nice properties but which are not quasiconformally equivalent to balls.

187


Revista Matemática Iberoamericana
Vol. 12, N.° 2, 1996

On the nonexistence of bilipschitz parameterizations and geometric problems about A_∞-weights

Stephen Semmes

Abstract. How can one recognize when a metric space is bilipschitz equivalent to a Euclidean space? One should not take the abstraction of metric spaces too seriously here; subsets of $\mathbb{R}^n$ are already quite interesting. It is easy to generate geometric conditions which are necessary for bilipschitz equivalence, but it is not clear that such conditions should ever be sufficient. The main point of this paper is that the optimistic conjectures about the existence of bilipschitz parameterizations are wrong. In other words, there are spaces whose geometry is very similar to but still distinct from Euclidean geometry. Related questions of bilipschitz equivalence and embeddings are addressed for metric spaces obtained by deforming the Euclidean metric on $\mathbb{R}^n$ using an A_∞ weight.

1. Introduction.

How can one recognize when a metric space is bilipschitz equivalent to a Euclidean space? Recall that a map $f : M_1 \to M_2$ between two metric spaces M_1, M_2 is *bilipschitz* if there is a constant K such that $K^{-1}\, d_1(x,y) \le d_2(f(x), f(y)) \le K\, d_1(x,y)$ for all $x, y \in M_1$, where d_1

337

Through these examples, he addresses the roles of wildness, shrinkability and linking in questions of parametrization, and expresses his philosophical views on mappings and spaces. The usefulness of Semmes' construction is not limited to the problems at hand. Work on Semmes-type spaces there-

*wu@math.uiuc.edu.

after by Heinonen and Rickman [5] and by Pankka and Rajala [8] further highlights the topological properties of the spaces that admit or receive geometrically controlled branched covering maps.

In these papers of Semmes, ideas linking analysis and topology unroll in a story-telling style –more like a novel than a textbook. One can relax and watch Semmes effortlessly connecting the dots and unfolding the facts. In the end, the reader is rewarded with a mystery solved.

We now give a sample of Semmes' ideas and their implications.

Semmes metrics. Despite considerable attention in recent years, the problem of characterizing metric n-spheres that are bilipschitz or quasisymmetrically equivalent to the standard $\mathbb{S}^n$ is still far from understood. There exist finite 5-dimensional polyhedra (double suspension of homology 3-spheres with nontrivial fundamental groups) homeomorphic to the standard sphere $\mathbb{S}^5$ but not bilipschitz equivalent to $\mathbb{S}^5$ –an observation of Siebenmann and Sullivan (1979) based on deep work of Cannon (1978) and Edwards (1980). It is unknown whether these polyhedra are quasisymmetrically equivalent to $\mathbb{S}^5$.

When $\mathbb{R}^n$ is equipped with a path metric $D_\omega(x, y)$ associated with a continuous strong A_∞ weight ω ($\omega\, dx$ a doubling measure and $D_\omega(x, y)^n$ comparable to the ω-measure, $\int_{B_{x,y}} \omega$, of the smallest Euclidean ball $B_{x,y}$ containing x and y), the geometry of the space $(\mathbb{R}^n, D_\omega)$ is in many ways indistinguishable from $\mathbb{R}^n$ (David and Semmes 1990, Semmes 1993). For example, the metric D_ω is quasi-equivalent to the Euclidean metric in the sense that every D_ω-ball B contains a Euclidean ball and is contained in a Euclidean ball of comparable radii (in general very different from the radius of B); the ω-measure of any D_ω-ball of radius r is comparable to r^n (Ahlfors n-regular); and every D_ω-ball contains a definite portion in ω-measure that is uniformly bilipschitz equivalent to a subset of $\mathbb{R}^n$ in the Euclidean metric. Moreover $(\mathbb{R}^n, D_\omega)$ supports Sobolev and weak (1,1)-Poincaré inequalities which are crucial for differential calculus. Must $(\mathbb{R}^n, D_\omega)$ be bilipschitz equivalent to $\mathbb{R}^n$?

In [13] *Semmes found a strong A_∞ weight ω so that the associated space $(\mathbb{R}^3, D_\omega)$ is not bilipschitz equivalent to $\mathbb{R}^3$.*

Semmes' idea is to create a metric in $\mathbb{R}^3$ in terms of the distance function to a geometrically nice but wild Cantor set. Under Semmes' metric this Cantor set has a small Hausdorff dimension. Precisely, let $\mathcal{N}$ be a geometrically self-similar Antoine's necklace constructed in such a way that all tori used are similar and all tori in the same generation are congruent as illustrated in [10, p. 73]. The complement $\mathbb{R}^3 \setminus \mathcal{N}$ is non-simply connected. Then, for a fixed $s > 0$,

$$\omega(x) = \min(1, \operatorname{dist}(x, \mathcal{N})^s)$$

is a strong A_∞ weight in $\mathbb{R}^3$. The Hausdorff dimension of the necklace $\mathcal{N}$ in $(\mathbb{R}^n, D_\omega)$ is at most $(1+s/6)^{-1}$. On the other hand, every homeomorphism h of $\mathbb{R}^3$ maps $\mathcal{N}$ to a set with non-simply connected complement; this implies that the Hausdorff dimension of $h(\mathcal{N})$ is at least 1 for the Euclidean metric. The spaces $(\mathbb{R}^3, D_\omega)$ and $\mathbb{R}^3$ therefore are not bilipschitz equivalent.

Jacobians of quasiconformal mappings are classical A_∞-weights. The problem of characterizing weights in $\mathbb{R}^n$ that are comparable to a quasiconformal Jacobian is related to the bilipschitz parametrization problem. Semmes' example can be rephrased to give a counterexample the conjecture that every strong A_∞-weight is a quasiconformal Jacobian.

The topological obstructions above are restricted to dimension 3 or higher. Good metrics on 2-spheres that do not admit bilipschitz parametrization by $\mathbb{S}^2$ have been constructed later by Laakso (2002) and Bishop (2007).

The idea of constructing metrics based on distance is versatile. It can be adapted to create new metrics on subsets of $\mathbb{R}^3$ exploiting their topological characteristics in such a way that the obstruction to a particular problem caused by the Euclidean metric disappears and the solvability is determined by the topological nature of the sets. We call any metric constructed with this goal a Semmes-type metric.

Non-Euclidean Picard Theorem. Non-constant quasiregular mappings (higher dimension analogues of analytic functions or multivalent analogues of quasiconformal maps) from $\mathbb{R}^n$ to $\mathbb{R}^n$ can omit only finitely many values; and for any finite set of points in $\mathbb{R}^3$ there exists a quasiregular mapping from $\mathbb{R}^3$ into $\mathbb{R}^3$ omitting exactly those points –striking theorems of Rickman in 1980 and 1985.

Equipped with a Semmes-type metric, subsets of $\mathbb{S}^3$ become more amenable to receiving quasiregular maps. A sharp non-Euclidean Picard-type theorem in dimension 3 of Pankka and Rajala [8] inspired by Semmes' construction says that *if L is either an unknot (flat circle) or a Hopf link (two flat circles linked once) in $\mathbb{S}^3$, then there exists a Riemannian metric g in $\mathbb{S}^3 \setminus L$ so that $(\mathbb{S}^3 \setminus L, g)$ receives non-constant quasiregular maps from $\mathbb{R}^3$, i.e., is quasiregularly elliptic; on the other hand, if L is a link in $\mathbb{R}^3$ and there exists a Riemannian metric g in $\mathbb{S}^3 \setminus L$ so that $(\mathbb{S}^3 \setminus L, g)$ is quasiregularly elliptic, then L must be an unknot or a Hopf link.*

In the case of the classical Picard theorem, the non-existence of (non-constant) analytic functions into a twice punctured plane is due to the fact that the fundamental group $\pi_1(\mathbb{C} \setminus \{0, 1\})$ is a free group on two generators. The same topological obstruction occurs in the non-Euclidean theorem: the fundamental group of $\pi_1(\mathbb{S}^3 \setminus L)$ contains a free group of rank 2 as a subgroup if L is any link except the unknot or the Hopf link.

It is unknown however whether $\mathbb{S}^3 \setminus Wh$, the complement of a Whitehead continuum, can be equipped with a Semmes-type metric so that it is quasiregularly elliptic [8].

Semmes spaces. When is a metric n-sphere (X, d) quasisymmetrically equivalent to $\mathbb{S}^n$? A complete characterization is known only for dimensions 1 and 2 ([14], [2]). Conditions of Semmes [11] and of Bonk and Kleiner [2] imply that *if a metric 2-sphere is linearly locally contractible (every ball of radius r is contractible in the ball of radius Cr with the same center) and Ahlfors 2-regular (there exists a measure μ on the space so that the μ-measure of every ball of radius r is comparable to r^2 uniformly) then it is quasisymmetrically equivalent to $\mathbb{S}^2$.*

Could a metric n-sphere which resembles $\mathbb{S}^n$ geometrically (linearly locally contractible) and measure-theoretically (Ahlfors n-regular) fail to be quasisymmetrically equivalent to $\mathbb{S}^n$?

Semmes' negative example in dimension 3 is a geometrically self-similar metric space $(\mathbb{R}^3/Bd, d)$ modeled on the decomposition of $\mathbb{R}^3$ with respect to Bing's double [12].

The classical construction of R. H. Bing in geometric topology gives an involution in $\mathbb{S}^3$ whose fixed point set is a double horned sphere. Bing's double is a set constructed following Bing's procedure topologically, not necessarily geometrically. One construction of Bing's double Bd starts with a solid smooth torus T standardly embedded in $\mathbb{R}^3$ and two smooth tori T_1 and T_2 linked and embedded in the interior of T as illustrated in [1, Fig. 3, p. 357], or in [3, Fig 9-1, p. 63]. Let $\phi_j \colon T \to T_j$, $j = 1, 2$ be diffeomorphisms, $S_l = \{1, 2\}^l$, $\alpha = (\alpha_1, \ldots, \alpha_l) \in S_l$ and $\phi_\alpha = \phi_{\alpha_l} \circ \cdots \circ \phi_{\alpha_2} \circ \phi_{\alpha_1}$. Bing's double is

$$Bd = \bigcap_{l=0}^{\infty} \bigcup_{\alpha \in S_l} \phi_{\alpha T}.$$

The complement $\mathbb{R}^3 \setminus Bd$ is not simply connected, and as a topological space $\mathbb{R}^3/Bd$ is $\mathbb{R}^3$ for nontrivial reasons.

Semmes' idea is to embed $\mathbb{R}^3/Bd$ into $\mathbb{R}^4$ by unknotting and resizing the tori in the construction geometrically. As a first step, $\mathbb{R}^3 \setminus T$ is embedded in $\mathbb{R}^3 \times \{0\}$ by inclusion, and the linked tori $T_1 \cup T_2$ are mapped diffeomorphically onto tori, similar to T and of size λ times that of T, which are contained in two mutually disjoint Euclidean balls in $(\operatorname{int} T) \times \{0\}$. The embedding is then extended to a diffeomorphism θ from $\mathbb{R}^3$ into $\mathbb{R}^4$ by an unknotting argument. Careful construction of θ allows $\theta(T_1)$ and $\theta(T_2)$ to assume the role of T and the unknotting and resizing procedure to be iterated geometrically. At the limit we obtain a map that descends to a homeomorphism from $\mathbb{R}^3/Bd$ into $\mathbb{R}^4$. Semmes' geometrical realization of $\mathbb{R}^3/Bd$ is a 3-dimensional submanifold of $\mathbb{R}^4$ smooth outside a Cantor set Bd^*.

The space $(\mathbb{R}^3/Bd, d_\lambda)$ with the metric induced by the ambient Euclidean metric in $\mathbb{R}^4$ through the embedding is quasiconvex, linearly locally contractible and Ahlfors 3-regular and smooth except for well-controlled degeneracies near Bd^*. Moreover it satisfies the Sobolev and Poincaré inequalities needed for analysis.

However *the Semmes space* $(\mathbb{R}^3/Bd, d_\lambda)$ *is not quasisymmetrically equivalent to* $\mathbb{R}^3$.

Semmes' elegant explanation of this fact goes as follows. Suppose h is a homeomorphism from $\mathbb{R}^3/Bd$ onto $\mathbb{R}^3$. All l-th generation tori in $\mathbb{R}^3/Bd$ are similar to T and have diameter λ^l. Their images in $\mathbb{R}^3$ circulate around $h(\theta(T))$ at least 2^l times. Therefore at least one of 2^l image tori, call it τ_l, must have a longitude of length at least $c_0 > 0$, for every $l \geq 1$. Since $\operatorname{diam} \tau_l \to 0$ as $l \to \infty$, the tori τ_l can not be uniformly well-shaped, therefore h can not be quasisymmetric. This heuristic argument can be made precise by a lemma of Freedman and Skora (1987) using relative homology.

Any number of linked tori may be used in the first step of defining Bing's double. In case one torus T_1 is self-linked in T in such a way that a meridian of T and a longitude of T_1 form a Whitehead link as in [3, Fig. 9-7, p. 68] or in [10, p. 72], the resulting intersection is called a Whitehead continuum Wh. In case $k\,(\geq 3)$ tori $\bigcup_{1\leq j\leq k} T_j$ are linked in T as in [3, Fig. 9-9, p. 71] or in [10, p. 73], the resulting set is an Antoine's necklace.

Semmes' geometrization extends to all decomposition spaces of $\mathbb{R}^3$ defined by an initial package with a topological self-similarity. With an additional contractibility condition, the resulting spaces are generalized manifolds possessing all metric properties mentioned above for the space $(\mathbb{R}^3/Bd, d_\lambda)$. We call these spaces and other non-self similar ones constructed in this spirit Semmes-type spaces.

Branched covering maps. It seems that the existence of a bilipschitz parametrization for metric spheres is a rarity and that a concrete geometrical characterization is difficult. Heinonen and Rickman [5] however showed that *all spaces constructed in a geometrically self-similar manner from initial packages of Semmes on* $\mathbb{S}^3$ *with an additional contractibility condition, admit BLD-maps (maps of bounded length distortion –multivalent analogues of bilipschitz maps) onto* $\mathbb{S}^3$.

The example arising from the Bing's double decomposition space leads to a space $(\mathbb{S}^3/Bd, d_\lambda)$ that is homeomorphic to $\mathbb{S}^3$, although quasisymmetrically inequivalent to $\mathbb{S}^3$, but can be mapped onto $\mathbb{S}^3$ by a BLD-map whose branch set contains a wild Cantor set. This particular example shows a sharp contrast between the finite-to-one and injective cases and the power of Semmes-type metrics.

It follows from the above that there exists a branched cover $F : \mathbb{S}^3 \to \mathbb{S}^3$ (discrete open map) so that for no homeomorphism $h : \mathbb{S}^3 \to \mathbb{S}^3$ is $F \circ h : \mathbb{S}^3 \to \mathbb{S}^3$ quasiregular. An important question remains open: whether every such branch covering map F is topologically conjugate to a quasiregular map, *i.e.*, there exist homeomorphisms $g, h : \mathbb{S}^3 \to \mathbb{S}^3$ so that $g \circ F \circ h$ is quasiregular [5].

Quasisymmetric parametrization. At a meeting in 2005, Juha Heinonen suggested that we work on the question of quasisymmetric parametrization of the double suspension of homology 3-spheres $\Sigma^2 H^3$ [6, Question 12]. With no idea whether the answer would be yes or no, we set out in Fall 2006 to read Edwards' explicit construction of a homeomorphism between $\mathbb{S}^5$ and a particular $\Sigma^2 H^3$ (work of 1980, arXiv 2006). Our hope was that Edwards' map could be modified to be quasisymmetric; this task turned out to be more ambitious than originally expected. On the other hand, there is a subtle connection between the double suspension problem and the decomposition theory at the topological level [3, p.103].

As we struggled to make progress in quasisymmetric parametrization, experimenting with Semmes-type spaces built from classical examples in decomposition theory was fascinating. With Heinonen [7] we showed that the natural conditions mentioned earlier for good parametrization are also insufficient in dimension 4 or higher. More specifically, *the decomposition space $\mathbb{R}^3/Wh$ associated with the Whitehead continuum admits a Semmes-type metric that is linearly locally contractible and Ahlfors 3-regular but $(\mathbb{R}^3/Wh)\times\mathbb{R}^m$ is not quasisymmetrically equivalent to $\mathbb{R}^{3+m}$, for any $m \geq 1$.*

The complement of the Whitehead continuum Wh in $\mathbb{S}^3$ is a contractible non-compact 3-manifold that is not homeomorphic to $\mathbb{R}^3$. The decomposition space $\mathbb{R}^3/Wh$ is not $\mathbb{R}^3$, but $(\mathbb{R}^3/Wh)\times\mathbb{R}$ is homeomorphic to $\mathbb{R}^4$. The nonexistence of the quasisymmetric parametrization is due to the different roles of the meridians in the homotopy and the homology in the Whitehead construction and their roles in estimating moduli of surface families.

$\mathbb{R}^3/Wh$ is only one case of a non-trivial manifold factor of $\mathbb{R}^4$. By a theorem of Edwards and Miller [4], cell-like closed 0-dimensional upper semicontinuous decomposition spaces $\mathbb{R}^3/G$ are manifold factors of $\mathbb{R}^4$, $\mathbb{R}^3/G \times \mathbb{R} \approx \mathbb{R}^4$. Decomposition spaces satisfying Edwards and Miller's conditions are definable by defining sequences consisting of unions of cubes-with-handles (handlebodies), see Lambert and Sher (1968) and Sher and Alford (1968). This class provides a natural environment for testing quasisymmetric parametrization.

With Pankka [9] we consider a subclass of decomposition spaces $\mathbb{R}^3/G$ that are manifold factors and admit defining sequences (X_k) consisting of handlebodies of controlled topological complexity. As self-similar spaces these spaces may be equipped with Semmes-type metrics with controlled geometry that are linearly locally contractible and Ahlfors 3-regular.

We have noted that the existence of a quasisymmetric parametrization of $\mathbb{R}^3/G\times\mathbb{R}^m$ by $\mathbb{R}^{3+m}$ for any $m \geq 0$ imposed a necessary constraint on the geometry (growth of the handlebodies and the scaling factor of the metric) in terms of the topology (genus, welding and circulation of the handlebodies), which is needed for the quasi-invariance of the modulus. On the other hand, a strong self-similar welding structure on the decomposition suffices to guarantee the existence of a quasisymmetric parametrization of $(\mathbb{R}^3/G, d)$ for a properly chosen Semmes-type metric. Here, the growth defines how fast the handlebodies propagate; the welding describes an embedding relation between handlebodies of two consecutive generations; and the circulation, in some sense, sums up the (unsigned) winding numbers of handlebodies of one particular generation inside the previous one. Even for this subclass the gap between the known necessary and the sufficient conditions remains wide.

Semmes spaces combine a new kind of metrization with classical topology in a subtle and mysterious manner. In the field of quasiconformal analysis, if one searches, Semmes-type spaces exist everywhere.

References

[1] Bing, R. H.: A homeomorphism between the 3-sphere and the sum of two solid horned spheres. *Ann. of Math. (2)* **56** (1952), 354–362.

[2] Bonk, M. and Kleiner, B.: Quasisymmetric parametrizations of two-dimensional metric spheres. *Invent. Math.* **150** (2002), 127–183.

[3] Daverman, R. J.: *Decompositions of manifolds.* Pure and Applied Mathematics **124**. Academic Press, Orlando, FL, 1986.

[4] Edwards, R. D. and Miller, R. T.: Cell-like closed-0-dimensional decompositions of R^3 are R^4 factors. *Trans. Amer. Math. Soc.* **215** (1976), 191–203.

[5] Heinonen, J. and Rickman, S.: Geometric branched covers between generalized manifolds. *Duke Math. J.* **113** (2002), no. 3, 465–529.

[6] Heinonen, J. and Semmes, S.: Thirty-three yes or no questions about mappings, measures, and metrics. *Geom. Dyn.* **1** (1997), 1–12.

[7] Heinonen, J. and Wu, J.-M.: Quasisymmetric nonparametrization and spaces associated with the Whitehead continuum. *Geom. Topol.* **14** (2010), no. 2, 773–798.

[8] Pankka, P. and Rajala, K.: Quasiregularly elliptic link complements. *Geometriae Dedicata* (to appear).

[9] Pankka, P. and Wu, J.-M.: Quasisymmetric parametrization of Semmes spaces, preprint.

[10] Rolfsen, D.: *Knots and links.* Mathematics Lecture Series **7**. Publish or Perish, Berkeley, CA, 1976.

[11] Semmes, S.: Chord-arc surfaces with small constant. II. Good parameterizations. *Adv. Math.* **88** (1991), no. 2, 170–199.

[12] Semmes, S.: Good metric spaces without good parameterizations. *Rev. Mat. Iberoamericana* **12** (1996), no. 1, 187–275.

[13] Semmes, S.: On the nonexistence of bi-Lipschitz parametrizations and geometric problems about A_∞-weights. *Rev. Mat. Iberoamericana* **12** (1996), no. 2, 337–410.

[14] Tukia, P. and Väisälä, J.: Quasisymmetric embeddings of metric spaces. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **5** (1980), 97–114.

PAUL YANG (Changhua, Republic of China, 1947) obtained his PhD in mathematics at the University of California, Berkeley in 1973 (advisor: Hung-Hsi Wu). He is Professor at Princeton University. His main research areas are conformal and CR geometry.

CR Geometry

Paul Yang*

1. Introduction

Cauchy–Riemann geometry is concerned with the geometry of a smooth hypersurface inherited from the geometry of its ambient space. The study of CR geometry began with the work of E. Cartan who determined a complete set of local invariants associated with the geometry of a hypersurface in $\mathbb{C}^2$. In the mid 1970s, there followed the publication of several important papers which laid the geometric as well as the analytic foundation of CR geometry. In [6], Chern and Moser extended the work of E. Cartan to general dimensions and at the same time determined a normal form of a hypersurface in $\mathbb{C}^n$ and identified the coefficients with the curvature invariants. In the remarkable series of papers [7, 8, 9] C. Fefferman laid the ground work for future development in CR geometry as well as conformal geometry for many years to come. In [7], he proved the regularity of biholomorphic maps at the boundary of strictly pseudo-convex domains and thus making the previous work relevant for the study of geometry of strictly pseudo-convex domains. In [8], C. Fefferman introduced a complex Monge–Ampère equation:

$$J[u] = (-1)^{n+1} det \begin{pmatrix} u & u_{\bar{j}} \\ u_i & u_{i\bar{j}} \end{pmatrix} = 1 \,. \tag{1}$$

There soon followed the solution of this equation by Cheng–Yau ([5]) which provides the existence of a complete Kahler–Einstein metric in the interior of a spc domain. Around the same time the existence of global embedding of abstract CR structures was given by [1] Boutet de Monvel in dimensions greater than three leaving open the question in dimension three. In 1984 ([9]) Fefferman and Graham gave an outline of the ambient metric construction for conformal invariants which preceded the idea of ADS-CFT correspondence by some twenty years, and set the stage for much of subsequent work in conformal geometry. I had been fascinated by these developments for many years. In this short note, I discuss some aspects of the ensuing work.

I wish to thank the editors for the kind invitation to contribute to this volume.

*yang@math.princeton.edu.

2. Approximate solutions of (1) and the Kahler–Einstein metric

In [8], an iterative computational procedure is given to find approximate solutions of the equation (1) that is accurate to order $n+1$ near the boundary: let ψ be any smooth defining function, that is $\psi = 0$ on $\partial\Omega$ and $d\psi \neq 0$ on $\partial\Omega$; set

$$u_1 = \psi/J(\psi)^{1/3},$$
$$u_s = u_{s-1}\Big(1 + \frac{1 - J[u_{s-1}]}{(n+1-s)s}\Big), \quad \text{for } 2 \leq s \leq n+1.$$

This approximate solution gives the correct asymptotics of an actual solution to this nonlinear partial differential equation. In fact, the proof of [5] used this approximate solution as the initial step in a continuity argument to find the actual solutions.

An example of an explicit solution is found for the tube domain given by $x_1^2 + x_2^2 - 1 < 0$ where the complex coordinates are $z_1 = x_1 + \sqrt{-1}y_1,\ z_2 = x_2 + \sqrt{-1}y_2$. The exact solution is given by $u = x_1^2 + x_2^2 - 1$.

3. *Q* curvature and the Szegő kernel

Given any defining function u there is an associated contact form $\theta_u = d^c u = \sqrt{-1}(\bar\partial u - \partial u)$. All other contact form giving rise to the same contact structure are obtainable from this particular one by scaling: $\theta = e^{2v}\theta_u$, in exact analogy with conformal geometry. In [15] and [14], a connection compatible with the underlying CR structure is determined and it gives associated curvature invariants. In dimension three, Hirachi [11] determined a fourth order curvature invariant which we shall call Q-curvature: $Q = -\Delta_b R - 2ImA^{11}_{11}$.

The analytic significance of Q curvature lies in its relation to the Szegő kernel. Recall the holomorphic functions which belong to L^2 have their trace on the boundary, and when a contact form is given, one may form L^2 inner product on the boundary. The Szegő kernel gives the projection relative to this L^2 to the boundary. In [2] Boutet-de-Monvel and Sjöstrand proved the expansion of Szegő kernel in terms of the defining function u near the boundary:

$$S(z,z) = \phi(z)u(z)^{-2} + \psi(z)\log u(z),$$

where ϕ, ψ are smooth functions on $\bar\Omega$. Hirachi identified ([11]) the value of ψ at the boundary with the Q-curvature.

Remarkably, the Q curvature behaves in a completely analogous way as the Q curvature in conformal geometry in dimension four. It is related to the following fourth order operator:

$$Pu = div_b P_3 u, \tag{2}$$

where P_3 is the third order operator used by Lee [13] to characterise pluriharmonic functions. The operator P enjoys conformal covariance property under conformal change of contact form $\theta' = e^{2w}\theta$:

$$P' = e^{-4w}P.$$

Hirachi had also verified the following transformation rule for Q curvature:

$$Pw + Q = Q'e^{4w}.$$

In case when the CR structure in 3D is the boundary of a strictly pseudoconvex domain, for the contact form given by $\theta = d^c u$ where u is a third order approximate solution of the equation (1), the associated Q curvature vanishes. According to the transformation rule above, making the conformal change of contact form $\theta' = e^{2w}\theta$ where w is pluriharmonic, the new contact form also has vanishing Q-curvature. As a consequence, there is a large number of solutions of the equation $Q = 0$.

4. Embeddability and the CR-Paneitz operator

Recently [3] we had found the following CR-invariant condition for embeddability of CR structure in 3D. If the CR-Paneitz operator P is nonnegative, and the CR Yamabe operator is strictly positive, then the non-zero eigenvalues of Kohn's operator $\Box_b$ is strictly bounded below by a positive constant. As a consequence of Kohn's work [12], such structures are realizable as boundaries of strictly pseudoconvex manifolds. The condition $P \geq 0$ is a subtle condition since the kernel of the operator P contains the pluriharmonic functions which is often infinite dimensional.

This embedding criteria also has nice consequences for the underlying CR structure: in a forthcoming article, [4], we formulate a positive mass theorem for CR structures in 3-D, a rigidity result characterizing the standard CR structure on the three sphere. We reduce this result to a solution of the $\Box_b$ equation, which is solvable due to the embeddability of the CR structure. This is being carried out by P. Yung and H. Hsiao.

This extra condition $P \geq 0$ is in a strict sense necessary, since there exists many CR structures near the standard S^3 for which the mass is negative. As a consequence, the CR Yamabe problem will have a minimizing solution in case these two sign conditions are satisfied. This completes the previous work of Gamara [10], who proved existence of solutions via a topological method.

References

[1] Boutet de Monvel, L.: *Intégration des équations de Cauchy–Riemann induites formelles.* (French) Séminaire Goulaouic-Lions-Schwartz 1974–1975; Équations aux derivées partielles linéaires et non linéaires, pp. Exp. No. 9, 14 pp. Centre Math., École Polytech., Paris, 1975.

[2] Boutet de Monvel, L. and Sjöstrand, J.: Sur la singularité des noyaux de Bergman et de Szegö. (French) In *Journées: Équations aux Dérivées Partielles de Rennes (1975)*, 123–164. Asterisque **34-35**. Soc. Math. France, Paris, 1976.

[3] Chanillo, S., Chiu, H. and Yang, P.: Embeddability for the three dimensional CR manifolds and CR Yamabe invariants. Preprint 2010.

[4] Cheng, J-H., Malchiodi, A. and Yang, P.: A positive mass theorem in three dimensional CR geometry, preprint.

[5] Cheng, S. Y.; Yau, S. T.: On the existence of a complete Kähler metric on noncompact complex manifolds and the regularity of Fefferman's equation. *Comm. Pure Appl. Math.* **33** (1980), no. 4, 507–544.

[6] Chern, S. S. and Moser, J. K.: Real hypersurfaces in complex manifolds. *Acta Math.* **133** (1974), 219–271.

[7] Fefferman, C.: The Bergman kernel and biholomorphic mappings of pseudoconvex domains. *Invent. Math.* **26** (1974), 1–65.

[8] Fefferman, C.: Monge-Ampére equations, the Bergman kernel, and geometry of pseudoconvex domains. *Ann. of Math. (2)* **103** (1976), no. 2, 395–416.

[9] Fefferman, C. and Graham, R.: Conformal invariants. In *The mathematical heritage of Élie Cartan (Lyon, 1984)*, 95–116. Astérisque **1985**, Numero Hors Serie.

[10] Gamara, N.: The CR Yamabe conjecture in the case $n = 1$. *J. Eur. Math. Soc.* **3** (2001) 105–137.

[11] Hirachi, K.: Scalar pseudo-hermitian invariants and the Szegő kernel on three-dimensional CR manifolds. *Lect. Notes Pure Appl. Math.* **143** (1992), 67–76.

[12] Kohn, J. J.: The range of the tangential Cauchy–Riemann operator. *Duke. Math. J.* **53** (1986), 525–545.

[13] Lee, J.: Pseudo-Einstein structures on CR manifolds. *Amer. J. Math* **110** (1988), 157–178.

[14] Tanaka, N.: *A differential geometric study on strongly pseudo-convex manifolds.* Lectures in Mathematics, Department of Mathematics, Kyoto University **9**. Kinokuniya Book-Store, Tokyo, 1975.

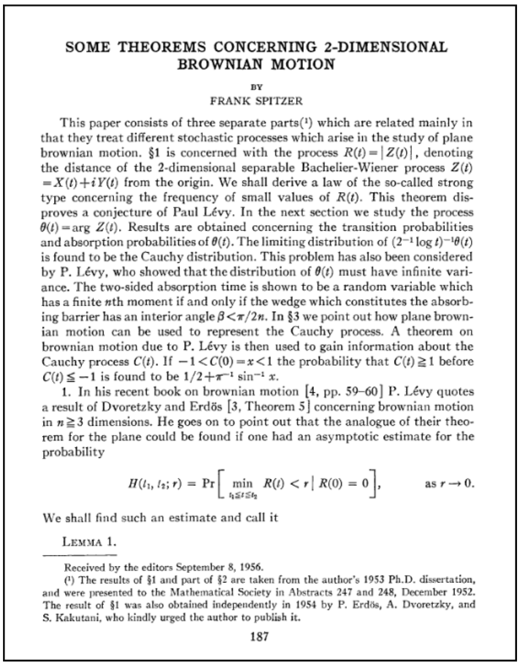
[15] Webster, S. M. Pseudo-Hermitian structures on a real hypersurface. *J. Differential Geom.* **13** (1978), no. 1, 25–41.

Marc Yor (Brétigny-sur-Orge France, 1949) obtained his Thèse d'Etat at the Université Paris VI-Pierre et Marie Curie in 1976 (advisor: Pierre Priouret). He has been Professor at the Pierre et Marie Curie Université (Paris 6) since 1981. His research focuses on Brownian motion and more generally on stochastic processes. He is also interested in connections with number theory.

Why I became especially interested to work from F. Spitzer's paper about the asymptotics of planar Brownian windings

Marc Yor*

1. First of all, I find the idea of writing a short note about "the very personal reasons why 'that particular publication', i.e: Spitzer's 1958 paper, got my attention and affected my personal research" very appealing. Indeed, I have written, quite independently from this project, notes about ten research themas in Probability [22], which are close –in spirit– to the aim of the present volume.



SOME THEOREMS CONCERNING 2-DIMENSIONAL BROWNIAN MOTION

BY

FRANK SPITZER

This paper consists of three separate parts(1) which are related mainly in that they treat different stochastic processes which arise in the study of plane brownian motion. §1 is concerned with the process $R(t)=|Z(t)|$, denoting the distance of the 2-dimensional separable Bachelier-Wiener process $Z(t)=X(t)+iY(t)$ from the origin. We shall derive a law of the so-called strong type concerning the frequency of small values of $R(t)$. This theorem disproves a conjecture of Paul Lévy. In the next section we study the process $\theta(t)=\arg Z(t)$. Results are obtained concerning the transition probabilities and absorption probabilities of $\theta(t)$. The limiting distribution of $(2^{-1}\log t)^{-1}\theta(t)$ is found to be the Cauchy distribution. This problem has also been considered by P. Lévy, who showed that the distribution of $\theta(t)$ must have infinite variance. The two-sided absorption time is shown to be a random variable which has a finite nth moment if and only if the wedge which constitutes the absorbing barrier has an interior angle $\beta<\pi/2n$. In §3 we point out how plane brownian motion can be used to represent the Cauchy process. A theorem on brownian motion due to P. Lévy is then used to gain information about the Cauchy process $C(t)$. If $-1<C(0)=x<1$ the probability that $C(t)\geqq 1$ before $C(t)\leqq -1$ is found to be $1/2+\pi^{-1}\sin^{-1}x$.

1. In his recent book on brownian motion [4, pp. 59–60] P. Lévy quotes a result of Dvoretzky and Erdös [3, Theorem 5] concerning brownian motion in $n\geqq 3$ dimensions. He goes on to point out that the analogue of their theorem for the plane could be found if one had an asymptotic estimate for the probability

$$H(t_1, t_2; r) = \Pr\Big[\min_{t_1\leqq t\leqq t_2} R(t) < r \,\Big|\, R(0) = 0\Big], \qquad \text{as } r \to 0.$$

We shall find such an estimate and call it

Lemma 1.

Received by the editors September 8, 1956.

(1) The results of §1 and part of §2 are taken from the author's 1953 Ph.D. dissertation, and were presented to the Mathematical Society in Abstracts 247 and 248, December 1952. The result of §1 was also obtained independently in 1954 by P. Erdös, A. Dvoretzky, and S. Kakutani, who kindly urged the author to publish it.

187

My attention got attracted to Spitzer's paper as I was writing my "Mémoire de DEA" (this would now be called: Master Thesis) in the summer of 1972, under the guidance of J. Azéma.

The title of this Memoir was: "Brownian Motion and Newtonian Potential Theory"[1]. For this purpose, I read L. Helms' classical treatise on Potential Theory [6] as well as about 40 pages I managed to understand from the "Bible": Itô–McKean [9]. At least, I thought I understood most of the results around p. 270 of [9], among which computations about the distribution (or the asymptotics) of the winding number of planar Brownian motion; in that "famous" p. 270, Itô and McKean present, in their own way, some of Spitzer's computations. Thoughts about this page literally haunted me for years, and indeed this page has been my starting point for years of investigations connected with planar Brownian windings. See, *e.g.*, Yor [18, 19], and 20 years later, [20]. Throughout the 80's, J. Pitman, J.F. Le Gall, P. Messulam and I kept ploughing this field. See below.

*deaproba@proba.jussieu.fr.

[1] This theme was very much in the air at that time; a few years later, the now well-known books by Port and Stone, and M. Rao –independently– appeared on this subject.

2. Spitzer's paper contains a few gems, among which:

a) A standard Cauchy process $(C_t, t \geq 0)$ may be constructed by subordination of a Brownian motion γ with the first passage process of another independent Brownian motion β, namely:

$$(C_t, t \geq 0) \stackrel{\text{(law)}}{=} (\gamma_{\sigma_t}, t \geq 0)\,, \tag{1}$$

where, on the right hand side, $(\gamma_u, u \geq 0)$ denotes a real-valued Brownian motion, independent from another real-valued one $(\beta_s, s \geq 0)$ say, and

$$\sigma_t = \inf\{s; \beta_s \geq t\}$$

(In fact, it may be that this representation goes back to Bochner (1955), but anyhow (1) plays an important role in Spitzer's paper).

Moreover, the representation (1) also plays an essential role in the study of level crossings of a Cauchy process, done in [14].

b) If $(Z_t, t \geq 0)$ denotes a planar (i.e: $\mathbb{C}$-valued) Brownian motion, starting from $Z_0 = 1 + i0$, then almost surely the path $(Z_t(\omega), t \geq 0)$ does not visit 0. Hence, there exists a continuous determination $(\theta_t(\omega), t \geq 0)$ of the windings of $\{(Z_u(\omega), u \leq t); t \geq 0\}$ around 0. One of the gems of Spitzer's paper is the following convergence in law result:

$$\frac{2\theta_t}{\log t} \xrightarrow[(t\to\infty)]{\text{(law)}} C_1\,, \tag{2}$$

where on the right hand-side C_1 denotes a standard Cauchy variable.

In fact, Spitzer [16] computes the characteristic function of θ_t, and deduces (2) from it.

3. I was very impressed by Spitzer's paper because, apart from the results (1) and (2) above, that paper was telling the reader that, although the planar Brownian motion trajectory is extremely complicated, it is possible to get a hand on the distribution (exact, or asymptotic) of an equally complicated functional of this trajectory, namely the winding number process. As an undergraduate student, I had learnt that the winding number around a point of a closed continuous curve avoiding that point could be computed from an integral –which was not a Riemann–Stieltjes integral– of (dz/z) along that curve. I started wondering whether, in case the curve is the Brownian one $(Z_s(\omega), s \leq t)$, that integral might be related to the Itô integral $\int_0^t dZ_s/Z_s$. It turned out that this was indeed true (see my derivation in [19]; at the same time, Ikeda–Manabe [8] discussed this kind of identities between stochastic line integrals and Itô integrals in a much more general geometric framework; finally, see also P. A. Meyer [11]). This opened –for

me, at least– the possibility of using stochastic calculus to recover Spitzer's result (2), say. In fact, the stochastic Itô integral

$$\Big(\int_0^t dZ_s/Z_s, t \geq 0\Big)$$

also allowed to consider quantities such as:

$$\int_0^t H_s \frac{dZ_s}{Z_s},$$

for some interesting adapted processes $(H_s, s \geq 0)$, showing once again that stochastic integration is more flexible than "ordinary" integration. Other famous examples are provided by the proofs via Brownian motion and stochastic calculus of the classical inequalities for harmonic functions, as well as the beautiful derivation by K. Carne [4] of the two main theorems of Nevanlinna theory about meromorphic functions f, taking advantage of the strict (martingale) locality of $(f(Z_t); t \geq 0)$. Nowadays, SLE theory (see [1] for an up to date exposition) is yet another masterful example, pushing forward vastly B. Davis' exploitation of the conformal invariance of planar Brownian motion to recover Picard's big theorem [5].

Following the above remark about stochastic integrals w.r.t dZ_s/Z_s, P. Messulam and I [10] realized that Spitzer's theorem (2) could be refined as follows: for any pair $0 < r < R < \infty$ of radii,

$$\frac{2}{\log t}\Big(\int_0^t 1_{(|Z_s| \leq r)} d\theta_s\,,\ \int_0^t 1_{(|Z_s| \geq R)} d\theta_s\Big) \tag{3}$$

converges in law towards:

$$\Big(\int_0^{\sigma_1} 1_{(\beta_s \leq 0)} d\gamma_s\,,\ \int_0^{\sigma_1} 1_{(\beta_s > 0)} d\gamma_s\Big) \tag{4}$$

(note that, in (4), r and R have disappeared!), where β and γ are two independent real valued Brownian motions, and $\sigma_1 = \inf\{t : \beta_t = 1\}$. Of course, taking the sum of the two components in (4), one recovers: γ_{σ_1}, which, by (1), is Cauchy distributed, thus recovering (2).

The above (simple!) remark could even be further exploited by considering the possibility of extending Spitzer's theorem (2) to the asymptotic study of Brownian windings around a finite number of points: $z_1, \dots, z_k$. However, this was not so simple to do!

Nonetheless, it turned out that developing the small and large windings strategy as in (3), but now for the k windings simultaneously, it was possible to show that the vector:

$$\frac{2}{\log t}\,(\theta_t^{z_1}, \dots, \theta_t^{z_k})$$

converges in law towards a vector of linked Cauchy variables. This was obtained jointly with J. Pitman ([13], [15]), along with many other limiting results. The short story of our joint work is told in [12].

4. To summarize, my interest in Spitzer's paper was startled by the fact that tools from stochastic analysis might be well adapted when dealing with (geometric) quantities related to the planar Brownian trajectory, viz: winding numbers of planar Brownian motion and were the counterparts of tools I had learnt in classical analysis.

It was the same feeling which guided me when I became interested in Cauchy's principal value:

$$H_t \stackrel{\text{def}}{=} \lim_{\varepsilon \to 0} \left(\frac{1}{\pi}\right) \int_0^t \frac{ds}{B_s} 1_{|B_s| \geq \varepsilon} \tag{5}$$

where, this time, $(B_s, s \geq 0)$ denotes real-valued Brownian motion.

Again, I found the definition (and existence!) of H_t in Problem 72 of Itô–McKean [9]; see also Yamada's paper [17], in which the author marvels about the rich offspring generated by this problem!

In the study of (5), the role of Brownian local times is essential. Using both stochastic calculus and excursion theory, P. Biane and I ([3]) were able to obtain the law of H_t, and a number of exact distributions for related quantities.

There is a striking connection/relationship with Spitzer's representation of the Cauchy process (1), namely: if $(\tau_t, t \geq 0)$ denotes the inverse of the local time at 0 of B, then

$$(H_{\tau_t}, t \geq 0) \text{ is a standard Cauchy process.} \tag{6}$$

However, an essential difference between (1) and (6) is that in (1), the process $(\sigma_t, t \geq 0) \stackrel{\text{(law)}}{=} (\tau_t, t \geq 0))$ is independent from the Brownian motion $(\gamma_u, u \geq 0)$ whereas in (6), the processes $(H_u, u \geq 0)$ and $(\tau_t, t \geq 0)$ depend on each other.

Despite this difference, I could not prevent asking myself whether, given $\tau_t = u$, H_{τ_t} might be distributed as γ_u, which would indeed explain (6).

It turned out that this "hope" was much too naive, and indeed completely wrong, but it motivated the authors in [3] to compute the joint Fourier–Laplace transform of (H_{τ_t}, τ_t), which then revealed some close connections with Lévy's stochastic area formula... Several authors (Bertoin, Fitzsimmons, Getoor, etc.) generalized this result by replacing Brownian motion by a general Lévy process. See, *e.g.*, Bertoin–Caballero [2] for the most advanced results to date and a number of references.

5. A fascinating feature of scientific research, and in particular mathematical research, is that "the symphony never ends", to quote (roughly) D. Williams.

The two topics which I very briefly outlined in Sections 3 and 4 above follow the rule:

- the asymptotic study of windings to which I participated is of homological nature, but homotopical studies were also developed by H. McKean, T. Lyons, J. Gruet, T. Mountford...
- principal values such as (5) were studied for symmetric Lévy processes by R. Getoor and P. Fitzsimmons, then in complete generality for Lévy processes by J. Bertoin; they also play a key role in some asymptotic studies of Brownian motion in random environments [7].

But the symphony echoes further: when, at the end of 1988, I was asked, by both H. Geman and M. Chesney independently to find the "price of Asian options", I realized that the knowledge accumulated about Bessel processes for the asymptotic study of planar Brownian windings could be used as a key tool; in other words, the Brownian windings problem and the Asian options problem are in a kind of duality. See, *e.g.*, the monograph [21], and especially, paper ♯7 there. The exploration of this duality keeps me busy to this very day.

Finally, I apologize in advance for related works (and authors) which (whom) I have not (unintentionally) mentioned.

References

[1] Alberts, T. and Sheffield, S.: The covariant measure of SLE on the boundary. *Prob. Th. Rel. Fields* **149** (2011), 331–371.

[2] Bertoin, J. and Caballero, M. E.: Regularity of the Cauchy principal value of the local times of some Lévy processes. *Bull. Sci. Math.* **123** (1999), no. 1, 47–58.

[3] Biane, P. and Yor, M.: Valeurs principales associées aux temps locaux browniens. *Bull. Sci. Math.* **111** (1987), no. 1, 23–101.

[4] Carne, K.: Brownian motion and Nevanlinna theory. *Proc. London Math. Soc.* **3-52** (1986), no. 2, 349–368.

[5] Davis, B.: Picard's theorem and Brownian motion. *Trans. Amer. Math. Soc.* **213** (1975), 353–362.

[6] Helms, L.: *Potential theory.* Universitext. Springer, London, 2009.

[7] Hu, Y., Shi, Z. and Yor, M.: Rates of convergence of diffusions with drifted Brownian potentials. *Trans. Amer. Math. Soc.* **351** (1999), no. 10, 3915–3934.

[8] Ikeda, N. and Manabe, S.: Stochastic integral of differential forms and its applications. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, 175–185. Academic Press, New York-London, 1978.

[9] Itô, K. and McKean, H. P.: *Diffusion processes and their sample paths.* Die Grundlehren der Mathematischen Wissenschaften **125**. Academic Press, Inc., Publishers, New York; Springer, 1965.

[10] Messulam, P. and Yor, M.: On D. Williams' "pinching method" and some applications. *J. London Math. Soc. (2)* **26** (1982), no. 2, 348–364.

[11] Meyer, P. A.: Géometrie différentielle stochastique. II. In *Seminar on Probability, XVI, Supplement*, 165–207. Lecture Notes in Mathematics **921**. Springer, Berlin-New York, 1982.

[12] Pitman, J. and Yor, M.: The asymptotic joint distribution of windings of planar Brownian motion. *Bull. Amer. Math. Soc.* (1984), 109–111.

[13] Pitman, J. and Yor, M.: Asymptotic laws of planar Brownian motion. *Ann. Probab.* **14** (1986), no. 3, 733–779.

[14] Pitman, J. and Yor, M.: Level crossings of a Cauchy process. *Ann. Prob.* **14** (1986), no. 3, 780–792.

[15] Pitman, J. and Yor, M.: Further asymptotic laws of planar Brownian motion. *Ann. Probab.* **17** (1989), no. 3, 965–1011.

[16] Spitzer, F.: Some theorems concerning 2-dimensional Brownian motion. *Trans. Amer. Math. Soc.* **87** (1958), 187–197.

[17] Yamada, T.: Principal values of Brownian local times and their related topics. In: *Itô's stochastic calculus and probability theory*, 413–422. Springer, Tokyo, 1996.

[18] Yor, M.: Formule de Cauchy relative à certains lacets browniens. *Bull. Soc. Math. France* **105** (1977), no. 1, 3–31.

[19] Yor, M.: Loi de l'indice du lacet brownien, et distribution de Hartman-Watson. *Z. Wahrsch. Verw. Gebiete* **53** (1980), no. 1, 71–95.

[20] Yor, M.: Generalized meanders as limits of weighted Bessel processes, and an elementary proof of Spitzer's asymptotic result on Brownian windings. *Studia Sci. Math. Hungar.* **33** (1997), no. 1-3, 339–343.

[21] Yor, M.: *Exponential Functionals of Brownian Motion and related Processes.* Springer Finance, 2001.

[22] Yor, M.: Dix thèmes de recherche en probabilité. In preparation for *Rev. Mat. Iberoam.*, June 2011.

Acabose de imprimir este libro en los talleres de Aventura Gráfica (Guadalajara) en el mes de septiembre de MMXI.