LECTURES ON THE KATO SQUARE ROOT PROBLEM

PASCAL AUSCHER

Abstract. This is the text of a series of three lectures on the recent solution of the square root problem for divergence form elliptic operators, a long-standing conjecture posed by Kato in the early 60's. In this text, the motivations for this problem and its setting are given. The ideas from harmonic analysis on the T(1) theorem and T(b) theorem for square functions are described. In particular, an apparently new formulation of a local T(b) theorem for square functions is stated. The ideas of the full proof are presented.

Contents

These three lectures were given at the Centre for Mathematics and its Applications, Australian National University, during July and August, 2001. I want to thank the CMA at the Australian National University for inviting me to the special program on scattering theory and spectral problems and for the nice and stimulating atmosphere created by the mathematicians at the CMA.

1. Elliptic operators

Consider an open subset Ω of $\mathbb{R}^n$, $n \geq 1$. Let V be a closed subspace of $H^m(\Omega) = W^{m,2}(\Omega)$ which contains $H^m_0(\Omega)$, the closure of smooth functions supported in Ω in $H^m(\Omega)$.

Let N, m be positive integers and define a sesquilinear form on $V^N \times V^N$ by

$$Q(f,g) = \int_\Omega \sum_{\substack{|\alpha|,|\beta| \leq m \\ 1 \leq i,j \leq N}} a^{ij}_{\alpha\beta}(x) \partial^\beta f_j(x) \partial^\alpha \overline{g}_i(x)\, dx$$

Here $f = (f_1, \ldots, f_N)$ and $g = (g_1, \ldots, g_N)$ belongs to V^N, and the coefficients $a^{ij}_{\alpha\beta}$ are complex-valued L^∞ functions on Ω. We use the standard notations of differential calculus in $\mathbb{R}^n$: multiindices, partials, and so on.

One assumes that

$$|Q(f,g)| \leq \Lambda \|\nabla^m f\|_2 \|\nabla^m g\|_2 + \kappa' \|f\|_2 \|g\|_2 \tag{1}$$

and the Gårding inequality

$$\operatorname{Re} Q(f,f) \geq \lambda \|\nabla^m f\|_2^2 - \kappa \|f\|_2^2 \tag{2}$$

for some $\lambda > 0$, $\kappa, \kappa' \geq 0$ and $\Lambda < +\infty$ independent of $f, g \in V^N$. Here, $\nabla^k f = (\partial^\alpha f)_{|\alpha|=k}$ and $\|\nabla^k f\|_2 = \big(\sum_{|\alpha|=k} \int_\Omega |\partial^\alpha f|^2\big)^{1/2}$.

A well-known representation theorem of Kato asserts that one can represent the form by

$$Q(f,g) = \langle Lf, g \rangle, \quad f \in \mathcal{D}(L), g \in V^N$$

where $D(L)$ is the subspace of those $f \in V^N$ such that $g \mapsto Q(f,g)$, originally defined on V^N, extends to a bounded anti-linear form on $L^2(\Omega, \mathbb{C}^N)$. As usual, it is convenient to denote the operator (system) as

$$(Lf)_i = \sum_{\substack{|\alpha|,|\beta| \leq m \\ 1 \leq j \leq N}} (-1)^{|\alpha|} \partial^\alpha (a^{ij}_{\alpha\beta} \partial^\beta f_j), \quad 1 \leq i \leq N. \tag{3}$$

In fact, the operator L is defined from V^N into its dual and $\mathcal{D}(L)$ can be seen as the subspace of $f \in V^N$ such that $Lf \in L^2(\Omega, \mathbb{C}^N)$. The restriction of L to $\mathcal{D}(L)$ is a maximal-accretive operator and $\mathcal{D}(L)$ is dense in V^N [11]. By abuse of notation, we do not distinguish in the notation L from its restriction. We remark that L^*, the adjoint of L is similarly obtained from the coefficients $\overline{a^{ji}_{\beta\alpha}}$.

Such an operator has a holomorphic functional calculus. It satisfies resolvent estimates such as

$$\|(\zeta - (L+\kappa))^{-1}\|_{op} \leq \mathrm{dist}(\zeta, \Gamma)^{-1}, \zeta \notin \Gamma$$

where Γ is an open sector with vertex 0, and half angle ω from the positive x-axis, where $w \in [0, \pi/2)$ depends on $\lambda, \Lambda, N, m, n$. Such estimates allow one, by the Cauchy formula, to define $f(L)$ for some appropriate holomorphic functions f defined on conic neighborhoods of the spectrum of $L + \kappa$ (ie, defined on larger open sectors).

In particular, one can take $f(\zeta) = \zeta^\alpha$ for $\alpha \in [-1, 1]$ and obtain the fractional powers of $L + \kappa$. They are closed unbounded operators with the expected properties such as

$$(L+\kappa)^\alpha (L+\kappa)^\beta = (L+\kappa)^{\alpha+\beta}$$

when $\alpha + \beta \in [-1, 1]$. In particular, $(L+\kappa)^{1/2}$ is the unique maximal-accretive square root of $L + \kappa$.

Kato first studied this question: is it possible to identify the domains of the positive fractional powers of $L + \kappa$?

2. In what way are square roots critical?

Kato found the following answer by abstract functional analytic methods [12]. He proved that for $\alpha \in (0, 1/2)$ then $\mathcal{D}((L+\kappa)^\alpha) = \mathcal{D}(L^* + \kappa)^\alpha)$. This result was completed by J.L. Lions [13] which found other identifications by compex interpolation and one has $\mathcal{D}((L+\kappa)^\alpha) = [L^2(\Omega), V^N]_{2\alpha}$. Whenever such interpolation spaces are known then one gets a result.

Also Lions proved that for any $\alpha \in (0, 1)$, the domain $\mathcal{D}((L+\kappa)^\alpha)$ is given by $[L^2(\Omega), \mathcal{D}(L)]_\alpha$ but this result is in practice useless as we do not know the domain of L. This implies nevertheless that whenever $\mathcal{D}((L+\kappa)^{1/2})$ and $\mathcal{D}((L^*+\kappa)^{1/2})$ are both contained in V^N then the three spaces are the same.

But the methods break down at $\alpha = 1/2$ and the result cannot be true by purely abstract reasoning as we see in the next section. The remaining question is the following.

Conjecture 4 (Kato square root problem)**.** *Does $\mathcal{D}((L+\kappa)^{1/2})$ coincide with the domain of the form Q?*

One case is easy. When L is self-adjoint then

$$\|(L+\kappa)^{1/2} f\|_2^2 = \langle (L+\kappa)f, f\rangle = Q(f,f) + \kappa\|f\|_2^2 \geq \lambda \|\nabla^m f\|_2^2$$

for all $f \in \mathcal{D}(L)$. Thus, $\mathcal{D}((L+\kappa)^{1/2})$ is contained in V^N, hence the spaces coincide.

Let us see why $\alpha = 1/2$ is critical. Let $n=1$, $m=1$ and $N=1$. That is, consider $L = DaD$ with $D = -id/dx$ with domain $H^1(\mathbb{R})$ and a is the multiplication by a bounded **real**-valued function $a(x)$ on $\mathbb{R}$ such that $a \geq 1$. In such a case, L is self-adjoint and the domain of L is the space of $f \in H^1(\mathbb{R})$ such that $af' \in H^1(\mathbb{R})$. It is not too hard to construct functions in the space [Actually, this space can even be characterized by an adapted wavelet basis, see [4]].

By self-adjointness, we have $\mathcal{D}(L^{1/2}) = H^1(\mathbb{R}^n)$ [This holds for **complex** a with $\operatorname{Re} a \geq 1$, but this is much harder]. Using interpolation we find that

$$\mathcal{D}(L^\alpha) = \begin{cases} H^{2\alpha}(\mathbb{R}), & \text{if } \alpha \in (0,1/2), \\ \{f \in H^1(\mathbb{R}); af' \in H^{2\alpha-1}(\mathbb{R})\}, & \text{if } \alpha \in (1/2,1). \end{cases}$$

In one dimension, the surjectivity of $-id/dx$ and the injectivity of its adjoint make the understanding of the domain of L easier. In higher dimensions, these properties are lost.

3. Abstract methods are insufficient

We present an adaptation of an abstract counterexample by McIntosh [14]. On $H = \ell^2(\mathbb{Z})$, define an unbounded selfadjoint operator D by $De_j = 2^j e_j$ and a bounded operator B by $Be_j = \sum_{n\in\mathbb{Z}} b_n e_{j+n}$, where (e_j) is the natural orthonormal basis of H and (b_n) is a sequence of complex numbers such that $\hat{b}(\theta) = \sum b_n e^{in\theta}$ satisfies $\|\hat{b}\|_\infty = 1$. Clearly, the operator B has norm equal to $\|B\| = \|\hat{b}\|_\infty = 1$. For $z \in \mathbb{C}$ with $|z| < 1$, one can define the maximal-accretive operator $L_z = DA_zD$ with $A_z = Id + zB$ by the method of forms. Let $R_z = (L_z)^{1/2}$.

Assume that $\|R_z u\| \leq c\|Du\|$ for all $u \in \mathcal{D}(D)$ and uniformly for $|z| \leq r < 1$. As a function of z, R_z is an operator valued holomorphic function so that $R_0' D^{-1}$ is bounded on H. Differentiating at $z=0$ the equation $R_z R_z = L_z$, we find

$$R_0' D + DR_0' = DBD.$$

Solving for R_0' one finds that

$$R_0' e_j = 2^j \sum c_n e_{j+n}\ , \quad c_n = \frac{b_n 2^n}{1+2^n}.$$

Hence, $\|R_0' D^{-1}\| = \|\hat{c}\|_\infty$ with evident notation. Now take $b_n = \frac{i}{\pi n}$, then $\hat{b}(\theta) = -\frac{2}{\pi}\sum_{n>0} \frac{\sin(n\theta)}{n} = \frac{\theta}{\pi} - 1$, $0 < \theta < 2\pi$, so that $\|\hat{b}\|_\infty = 1$.

But $\hat{c}(\theta) \sim -\frac{i}{\pi} \ln|\sin(\theta/2)|$ near 0 so that $\hat{c}$ is not bounded. This is a contradiction, hence $\|R_z u\| \le c\|Du\|$ fails for some z.

We shall find out that the Kato conjecture for elliptic operators belongs to the realm of harmonic analysis.

4. Why complex coefficients?

Take two pure second order **self-adjoint** operators L_1 and L_2 on $\mathbb{R}^n$ defined as in Section 1 and denote by A_1 and A_2 the matrix of coefficients corresponding to L_1 and L_2. Is is true that

$$\|(L_1)^{1/2} f - (L_2)^{1/2} f\|_2 \le C\|A_1 - A_2\|_\infty \|\nabla f\|_2 \quad ? \tag{5}$$

This apparently simple question is equivalent to asking about the strong regularity of the (non-linear) mapping

$$\text{coefficients} \mapsto \text{square root}$$

from an open set in $L^\infty(\Omega, E)$ into the space of bounded operators from V^N to $L^2(\Omega, \mathbb{C}^N)$, where E is some finite dimensional space.

This question is highly non-trivial. The solution of the conjecture for all possible **complex** coefficients (or least those complex coefficients that are perturbations of self-adjoint coefficients) gives us boundedness of this mapping on complex balls, hence analyticity by the use of complex function theory.

Here is an application of (5). Consider the solutions $u_k(t,x)$, $t \in \mathbb{R}$, $x \in \mathbb{R}^n$, $k = 1, 2$, of the wave equations

$$\partial_t^2 u_k(t) + L_k u_k(t) = 0, \quad t \in \mathbb{R}, x \in \mathbb{R}^n,$$

with same Cauchy data $\partial_t u_{k_{|t=0}} = g \in L^2(\mathbb{R}^n)$ and $u_k(0) = f \in H^1(\mathbb{R}^n)$. Then, starting from the ansatz

$$u_k(t) = e^{it(L_k)^{1/2}} \tilde{g} + e^{-it(L_k)^{1/2}} \tilde{f}$$

and using (5) one obtains for $t > 0$

$$\|u_1(t) - u_2(t)\|_2 + \left\|\int_0^t \nabla(u_1(s) - u_2(s))\, ds\right\|_2 \le Ct\|A_1 - A_2\|_\infty (\|\nabla f\|_2 + \|g\|_2).$$

This estimate is sharp. It suffices to take $L_1 = -\Delta$ and $L_2 = -(1+b)\Delta$ with b small to show this.

5. The known results

Here we state the positive answers to conjecture (5).

Theorem 6. *Let $n \ge 1$ and $L = -\operatorname{div} A\nabla$ be a pure second order operator on $\mathbb{R}^n$. Then $\mathcal{D}(L^{1/2}) = H^1(\mathbb{R}^n)$ with the estimate $\|L^{1/2} f\|_2 \sim \|\nabla f\|_2$.*

This is the result we shall explain in the following sections.

The case $n = 1$ was due to Coifman, McIntosh and Meyer in 1981 [9]. The general case is due to Hofmann, Lacey, McIntosh, Tchamitchian and the author [1]. It came after a successful attempt in 2 dimensions by Hofmann and McIntosh (unpublished manuscript). See the introduction [1] for references to earlier partial results.

Theorem 7. *Let $n \geq 1$ and L be an homogeneous elliptic $N\times N$-system of arbitrary order m on $\mathbb{R}^n$. Then $\mathcal{D}((L+\kappa)^{1/2}) = H^m(\mathbb{R}^n, \mathbb{C}^N)$ with the estimate $\|(L+\kappa)^{1/2} f\|_2 \leq C(\|\nabla^m f\|_2^2 + \kappa\|f\|_2^2)^{1/2}$.*

This result is due to Hofmann, McIntosh, Tchamitchian and the author [2].

Theorem 8. *Let $n \geq 1$ and $L = -\operatorname{div} A\nabla$ be a pure second order operator on a proper open set Ω of $\mathbb{R}$. Then one has $\mathcal{D}(L^{1/2}) = V$ with the estimate $\|L^{1/2} f\|_2 \leq C(\|\nabla f\|_2 + \|f\|_2)$ in the following cases*

(i) *$n = 1$ and all possible choices of Ω and V.*

(ii) *$n \geq 2$, Ω is a strongly Lipschitz domain and $V = H_0^1(\Omega)$ (Dirichlet boundary condition) or $V = H^1(\Omega)$ (Neumann boundary condition).*

This theorem is due to Tchamitchian and the author. In one dimension, this is achieved by constructing an adapted wavelet basis [4]. We mention the approach by interpolation methods and the result on $\mathbb{R}$ by McIntosh, Nahmod and the author [3]. In higher dimensions, this goes by transferring the result from $\mathbb{R}^n$ [6]. It is likely that the method applies to second order systems with Dirichlet or Neumann boundary conditions.

Proposition 9. *Assume that L is as in one of the previous theorems. Then one can perturb L by lower order terms (ie, obtain an inhomogeneous operator) and still answer positively the square root problem for the perturbed operator.*

We have separated this result from the others because it is an "abstract" statement proved in [5], Chapter 0, Proposition 11. Basically, any positive result for the square root of a given homogeneous operator is "stable" under perturbations by lower order terms.

6. Open problems

We list some problems ranked by level of difficulty, the first being most likely more tractable.

Problem 10. *Find a direct proof of Theorem 8 following the ideas of [1].*

Problem 11. *This problem was already posed by Lions. Prove the Kato conjecture for second order operators with mixed boundary conditions on strongly Lipschitz domains.*

Problem 12. *More generally, prove the Kato conjecture for second order operators under general boundary conditions on strongly Lipschitz domains.*

Problem 13. *Prove the Kato conjecture for higher order operators or systems with Dirichlet or Neumann boundary conditions on smooth domains, then on strongly Lipschitz domains. Study other types of boundary conditions.*

7. Harmonic analysis

Our goal is to understand when a square function estimate (SFE) of the form

$$\left(\int_0^\infty \|U_t f\|_2^2 \frac{dt}{t}\right)^{1/2} \leq C\|f\|_2, \tag{14}$$

can hold, where $(U_t)_{t>0}$ is a family of operators acting boundedly and uniformly on $L^2(\mathbb{R}^n)$.

We shall present the ideas in a model case and say how to generalize them. Proofs will not be given and the reader is referred to [5] and [8] for the $T(1)$ theorem. The version of the $T(b)$ theorem given here is new. Related ideas are in [7].

7.1. **The T(1) theorem.** The first part of the program is to find a simple statement equivalent to SFE.

We consider a model case in which one can compute $U_t f(x)$ as

$$\int U_t(x,y) f(y)\,dy$$

where

$$\text{the kernel } U_t(x,y) \text{ is supported in } |x-y| \leq t \tag{15}$$

and satisfies

$$|U_t(x,y)| \leq t^{-n} \qquad \text{and} \qquad |\nabla_y U_t(x,y)| \leq t^{-n-1}. \tag{16}$$

Notice that only regularity in the second variable is imposed.

Let Q be a cube with side parallel to the axes. We denote by $|Q|$ its volume in $\mathbb{R}^n$ and by $\ell(Q)$ its sidelength. Also cQ denotes the cube obtained by dilating c times Q from the centre of Q. If we apply (14) to $f = \mathbf{1}_{3Q}$ (the indicator function of $3Q$) and observe from (16) that

$$(U_t 1)(x) = U_t(\mathbf{1}_{3Q})(x)$$

whenever $x \in Q$ and $0 < t \le \ell(Q)$, then we obtain

$$\int_Q \int_0^{\ell(Q)} |(U_t 1)(x)|^2 \frac{dxdt}{t} \le C|3Q| = C3^n|Q|.$$

Such an estimate means that $|(U_t 1)(x)|^2 \frac{dxdt}{t}$ is a Carleson measure, that is a (positive Borel regular) measure μ on $\mathbb{R}^n \times (0, +\infty)$ such that

$$\sup \frac{\mu(\mathcal{R}_Q)}{|Q|} < +\infty$$

where the supremum is taken over all cubes Q. We have set $\mathcal{R}_Q = Q \times (0, \ell(Q)]$. We denote this supremum by $\|\mu\|_c$ and call it the Carleson norm of μ.

There is a converse to this which begins with the celebrated Carleson inequality.

Lemma 17. *Assume that P_t is an operator with kernel satisfying* (15) *and* (16) *(only the size estimate is used at this point). Then for any Carleson measure μ,*

$$\int_0^\infty \int_{\mathbb{R}^n} |P_t f(x)|^2 d\mu(x,t) \le C\|\mu\|_c \int_{\mathbb{R}^n} |f|^2$$

Assuming now that $|(U_t 1)(x)|^2 \frac{dxdt}{t}$ is a Carleson measure, this tells us that the operator

$$f \mapsto (U_t 1) \cdot (P_t f)$$

satisfies SFE. Hence, the SFE for U_t is the same as the SFE for V_t with

$$V_t = U_t - (U_t 1) \cdot P_t.$$

The latter operator has a kernel satisfying (16) (the regularity for $P_t(x, y)$ in the second variable is used here). If, in addition, we impose

$$P_t 1 = 1$$

then we have

$$V_t 1 = 0,$$

that is

$$\int_{\mathbb{R}^n} V_t(x,y)\, dy = 0. \tag{18}$$

This cancellation condition permits almost-orthogonality arguments in a second step.

Let us begin with the Schur Lemma.

Lemma 19. *Let $(\Delta_s)_{s>0}$ be a family of self-adjoint (this is just to make life easy) operators on $L^2(\mathbb{R}^n)$ such that*

$$f = \int_0^\infty \Delta_s^2 f \, \frac{ds}{s} \tag{20}$$

in the L^2-sense. Assume also the almost-orthogonality $L^2 - L^2$ bound

$$\|V_t\Delta_s\|_{op} \leq C\left(\inf\left(\frac{t}{s}, \frac{s}{t}\right)\right)^\alpha. \tag{21}$$

for some $\alpha > 0$. Then V_t satisfies SFE.

In practice, one takes $\Delta_t^*(= \Delta_t)$ satisfying properties (15), (16) and (18). Very often, Δ_t is an operator of convolution type and (20) is checked by use of the Fourier transform. Now to see that the almost-orthogonality bound holds we compute the kernel of $V_t\Delta_s$ as

$$\int_{\mathbb{R}^n} V_t(x, z)\Delta_s(z, y)\, dz.$$

When $|x - y| \geq 2\sup(t, s)$, then the support condition gives us 0, which is to say that the two functions of z are orthogonal. When $|x-y| \leq 2\sup(t, s)$ then, we see that the function with smaller support oscillates while the other is regular on that support. Thus one can perform an "integration by parts" by writing, if say $s \leq t$,

$$\int_{\mathbb{R}^n} V_t(x, z)\Delta_s(z, y)\, dz = \int_{\mathbb{R}^n} (V_t(x, z) - V_t(x, y))\Delta_s(z, y)\, dz.$$

Using the mean value inequality, we get the bound

$$C\frac{s}{t}t^{-n}\mathbf{1}_{|x-y|\leq 2t}$$

from which we obtain one of the almost-orthogonality bound. The other one is exactly symmetric since we have the cancellation condition (18).

Hence, the SFE for V_t is always valid. Let us summarize the results.

Theorem 22 (T(1) theorem)**.** *Assume U_t and P_t as above with $P_t1 = 1$. Then, the following are equivalent*

(i) *U_t satisfies SFE.*
(ii) *$(U_t1) \cdot P_t$ satisfies SFE.*
(iii) *$|(U_t1)(x)|^2 \frac{dxdt}{t}$ is a Carleson measure.*

Moreover, one has

$$\int_0^\infty \int_{\mathbb{R}^n} |U_tf(x) - (U_t1)(x) \cdot (P_tf)(x)|^2 \frac{dxdt}{t} \leq C \int_{\mathbb{R}^n} |f|^2.$$

The idea of comparing $U_t f$ to $(U_t 1)(P_t f)$ is natural in probability theory where U_t would be a positive linear operator. It was brought into the realm of square function estimates and Carleson measures by Coifman and Meyer [10].

Remark. 1) By handling tails, one can assume that $U_t(x,y)$ has some integrable decay at infinity such as

$$|U_t(x,y)| \le t^{\varepsilon}(t+|x-y|)^{-n-\varepsilon}, \quad \varepsilon > 0.$$

One can also replace the Lipschitz regularity by a Hölder type regularity.

2) One can take for P_t a dyadic averaging operator: Given a family of dyadic cube Q of $\mathbb{R}^n$, define

$$S_t f(x) = \frac{1}{|Q|}\int_Q f, \quad \text{when } x \in Q \text{ and } \ell(Q)/2 < t \le \ell(Q).$$

The difference is that the kernel of S_t is not Hölder smooth in its second variable. However, it is Sobolev smooth, in the sense that it belongs to $H^s(\mathbb{R}^n)$ when $s \in (0, 1/2)$. This is enough.

In our applications, U_t will neither have a nice kernel, nor regularity in the second variable. Here is the statement which applies.

Lemma 23. *Let $U_t : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, $t > 0$, be a measurable family of bounded operators with $||U_t||_{op} \le 1$. Assume that*

(i) *U_t has a kernel, $U_t(x,y)$, that is a measurable function on $\mathbb{R}^{2n}$ such that for some $m > n$ and for all $y \in \mathbb{R}^n$ and $t > 0$,*

$$\int_{\mathbb{R}^n}\left(1+\frac{|x-y|}{t}\right)^{2m}|U_t(x,y)|^2\,dx \le t^{-n}.$$

(ii) *For any ball $B(y,t)$ with center at y and radius t, U_t has a bounded extension from $L^\infty(\mathbb{R}^n)$ to $L^2(B(y,t))$ with*

$$\frac{1}{t^n}\int_{B(y,t)}|U_t f(x)|^2\,dx \le \|f\|_\infty^2.$$

and $U_t(f\mathcal{X}_R)$ converges to $U_t f$ in $L^2(B(y,t))$ as $R \to \infty$ where $\mathcal{X}_R$ stands for the indicator function of the ball $B(0,R)$.

Let P_t be as above. Then U_tP_t satisfies SFE if and only if $|(U_t 1)(x)|^2\,\frac{dxdt}{t}$ is a Carleson measure. Moreover, one has

$$\int_0^\infty\int_{\mathbb{R}^n}|U_tP_tf(x) - (U_t1)(x)\cdot(P_tf)(x)|^2\frac{dxdt}{t} \le C\int_{\mathbb{R}^n}|f|^2.$$

The idea of proof is to go back to the previous theorem by using the operator $U_t^*U_tP_t$.

The same conclusion holds if one replaces P_t by S_t.

7.2. **The T(b) theorem.** The next part of the program is to be able to obtain the Carleson measure estimate involving $U_t 1$. The ideas here grew out from Semmes' work [16].

In practice, either $U_t 1 = 0$ and there is nothing to do or $U_t 1 \neq 0$ and it is usually impossible to compute. T(b) theorems are useful tools designed to overcome such problems.

Let us go back to a model operator U_t as in the previous section. Assume that for each cube Q, there are functions $b_Q : 3Q \to \mathbb{C}$ with the following properties

$$\int_{3Q} |b_Q|^2 \leq C|Q|, \tag{24}$$

$$|(S_t b_Q)(x)| \geq \delta, \quad \text{for } (x,t) \in \mathcal{R}_Q, \tag{25}$$

$$(U_t b_Q)(x) = 0 \quad \text{for } (x,t) \in \mathcal{R}_Q. \tag{26}$$

The constant C and δ are of course independent of Q. Here the dyadic cubes have been chosen so that Q is one of them. Then

$$\begin{aligned}
\int_Q \int_0^{\ell(Q)} |(U_t 1)(x)|^2 \frac{dxdt}{t} &\leq \frac{1}{\delta^2} \int_Q \int_0^{\ell(Q)} |(U_t 1)(x) \cdot (S_t b_Q)(x)|^2 \frac{dxdt}{t} \\
&\leq \frac{2}{\delta^2} \int_Q \int_0^{\ell(Q)} |(U_t b_Q)(x)|^2 \frac{dxdt}{t} \\
&\quad + \frac{2}{\delta^2} \int_Q \int_0^{\ell(Q)} |(V_t b_Q)(x)|^2 \frac{dxdt}{t} \\
&= \frac{2}{\delta^2} \int_Q \int_0^{\ell(Q)} |(V_t b_Q)(x)|^2 \frac{dxdt}{t} \\
&\leq C|Q|.
\end{aligned}$$

The first inequality comes from (25), the second from the definition of V_t, then one uses (26) and the last inequality comes from SFE for V_t combined with (24).

Let us see how to relax the hypotheses. First, (24) is OK as is. Secondly, (26) can clearly be replaced by

$$\int_Q \int_0^{\ell(Q)} |(U_t b_Q)(x)|^2 \frac{dxdt}{t} \leq C|Q|. \tag{27}$$

Next, (25) implies in particular that $|b_Q(x)| \geq \delta$ for $x \in Q$, which is often too strong. We shall need this lower bound only on a subset of $\mathcal{R}_Q$.

Lemma 28. *Let μ be a measure on $\mathbb{R}^n \times (0,\infty)$. Assume there are two constants $C>0$ and $\eta \in (0,1)$ such that for each cube Q one can find disjoint subcubes Q_i of Q with*

$$\sum |Q_i| \le (1-\eta)|Q| \tag{29}$$

and

$$\mu(\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}) \le C|Q|$$

Then μ is a Carleson measure and $\|\mu\|_c \le C/\eta$.

The proof is so simple that we give it. Suppose a priori that μ is a Carleson measure. We wish to obtain the bound above. Write

$$\begin{aligned}\mu(\mathcal{R}_Q) &= \mu(\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}) + \sum \mu(\mathcal{R}_{Q_i})\\ &\le C|Q| + \|\mu\|_c \sum |Q_i|\\ &\le C|Q| + (1-\eta)\|\mu\|_c |Q|.\end{aligned}$$

It remains to divide by $|Q|$, to take the supremum over Q and to solve for $\|\mu\|_c$.

Thus one can replace (25) by

$$|S_t b_Q(x)| \ge \delta \quad \text{for } (x,t) \in \mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i} \tag{30}$$

where the cubes Q_i satisfy (29). In the argument to control $U_t 1$, the LHS is only integrated on $\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}$. In other words, we allow a “black hole” region $\cup \mathcal{R}_{Q_i}$ on which we know nothing provided the “bad” cubes Q_i do not cover all of Q.

Let me make a semantic digression. In French, a region $\mathcal{R}_Q$ is called “fenêtre de Carleson”, that is “Carleson window”. A very clean window lets the light through. A window which may have some dark spots but not too many of them still lets enough through. In other words, the light goes through except for some “black hole” regions.

How to get the picture given by the “lighted” region $\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}$? The answer is by a stopping-time argument.

The Carleson region $\mathcal{R}_Q$ can be partitioned as the union of rectangles

$$Q' \times]\ell(Q')/2, \ell(Q')]$$

indexed by all dyadic subcubes of Q (they are called Whitney rectangles), on which

$$(x,t) \mapsto S_t b_Q(x)$$

is the constant function

$$\frac{1}{Q'} \int_{Q'} b_Q$$

(recall that $S_t b_Q(x)$ is a dyadic average of b_Q over a dyadic cube).

Let us assume that $\int_Q b_Q = |Q|$. Let $\delta < 1$. Consider one of the dyadic children Q' of Q, that is the cubes obtained by subdividing Q with cubes with sidelength $\ell(Q)/2$. We have two options:

(i) if the average gets too small, that is

$$\mathrm{Re} \int_{Q'} b_Q \leq \delta |Q'|,$$

then stop and select that cube.

(ii) otherwise subdivide Q' and argue similarly for each dyadic children.

Keep going indefinitely and call Q_i the cubes on which b_Q has a small average.

By construction, these cubes are disjoint and one can see right away that the region $\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}$ is the region where $\mathrm{Re}(S_t b_Q)(x) \geq \delta$.

It remains to see (29). Indeed, one has

$$\sum (1-\delta)|Q_i| \leq \sum \mathrm{Re} \int_{Q_i} 1 - b_Q = -\mathrm{Re} \int_{Q \setminus \cup Q_i} 1 - b_Q \leq C|Q|^{1/2} |Q \setminus \cup Q_i|^{1/2}$$

by Cauchy-Schwarz inequality and (24). One easily concludes from there. Observe the crucial use of real parts in the above equality.

As we see, instead of asking for a pointwise lower bound $|b_Q| \geq \delta$ on Q, we only need a lower bound on the average of b_Q over Q, which is weaker.

Summarizing we have obtained the following theorem.

Theorem 31 (local T(b) theorem)**.** *Let U_t be as above. Assume that one has a family of functions $b_Q : 3Q \to \mathbb{C}$ satisfying* (24)*, $|\int_Q b_Q| \geq |Q|$ and* (27)*, then $|(U_t 1)(x)|^2 \frac{dxdt}{t}$ is a Carleson measure.*

Again, one can state variations of the statement provided one can make sense of $U_t 1$ and have the SFE for V_t or $V_t S_t$.

8. Back to square roots

We are considering a pure second order operator $L = -\operatorname{div} A \nabla$ with ellipticity constants λ and Λ on $\mathbb{R}^n$ ($\kappa = \kappa' = 0$).

Since L is maximal-accretive, a theorem of McIntosh and Yagi [15] asserts that

$$\|L^{1/2} f\|_2^2 \sim \int_0^\infty \|(I + t^2 L)^{-1} t L f\|_2^2 \frac{dt}{t}.$$

This can also be obtained using almost-orthogonality arguments. If we set

$$\theta_t F = (I + t^2 L)^{-1} t \operatorname{div}(AF)$$

for $F = (F_1, \dots, F_n)$ then we want to establish

$$\int_0^\infty \|\theta_t(\nabla f)\|_2^2 \frac{dt}{t} \leq C\|\nabla f\|_2^2. \tag{32}$$

We are therefore facing a square function estimate and we need to see what kind of estimates are available.

8.1. **Elliptic estimates.** Pointwise bounds for the kernel of θ_t are false (Recall that we are merely assuming the coefficients of A to be measurable) even when the coefficients are real (where the classical Aronson-De Giorgi-Nash-Moser theory can be used). Moreover, this kernel will not be regular in its second variable.

In fact, there is no mathematical implication between the Kato problem and pointwise bounds on heat kernels and vice-versa. The pointwise bounds are just handy when we have them.

What is possible to obtain are these off-diagonal bounds in the mean.

Lemma 33. *Let E and E_0 be two closed sets of $\mathbb{R}^n$ and set $d = \mathrm{dist}(E, E_0)$, the distance between E and E_0. Then*

$$\int_E |(I + t^2L)^{-1}f(x)|^2\, dx \leq Ce^{-\frac{d}{ct}} \int |f(x)|^2\, dx, \quad \mathrm{Supp}\, f \subset E_0,$$

$$\int_E |t\nabla(I + t^2L)^{-1}f(x)|^2\, dx \leq Ce^{-\frac{d}{ct}} \int |f(x)|^2\, dx, \quad \mathrm{Supp}\, f \subset E_0,$$

$$\int_E |(I+t^2L)^{-1}t\,\mathrm{div}(AF)\,(x)|^2\, dx \leq Ce^{-\frac{d}{ct}} \int |F(x)|^2\, dx, \ \mathrm{Supp}\ F \subset E_0,$$

where $c > 0$ depends only on λ and Λ, and C on n, λ and Λ.

These bounds will be sufficient for us thanks to the theory developed for square function estimates. They are reminiscent of the bounds found by Gaffney for Laplace-Beltrami operators on manifolds.

These bounds also imply one can define in the L^2_{loc} sense the resolvent applied to functions with polynomial growth at infinity. In particular, one has

$$(I + t^2L)^{-1}(1) = 1.$$

Lemma 34. *For some C depending only on n, λ and Λ, if Q is a cube in $\mathbb{R}^n$, $t \leq \ell(Q)$ and f is Lipschitz function on $\mathbb{R}^n$ then we have*

$$\int_Q |(I + t^2L)^{-1}f - f|^2 \leq Ct^2\|\nabla f\|_\infty^2|Q|,$$

$$\int_Q |\nabla((I + t^2L)^{-1}f - f)|^2 \leq C\|\nabla f\|_\infty^2|Q|.$$

8.2. **Applying the T(1) and T(b) theorems.** Choose P_t to be here the operator of convolution by $t^{-n}p(\frac{x}{t})$ with $\int p = 1$ and $p \in C_0^\infty(B(0,1))$, where $B(0,1)$ is the unit ball.

The first thing is to apply the theory of square functions in order to reduce to a Carleson measure estimate.

We observe first that

$$(\theta_t - \theta_t P_t^2)(\nabla f) = (I - (I + t^2 L)^{-1})\frac{(I - P_t^2)f}{t}$$

so that

$$\int_0^\infty \|(\theta_t - \theta_t P_t^2)(\nabla f)\|_2^2 \frac{dt}{t} \leq 4 \int_0^\infty \|\frac{(I - P_t^2)f}{t}\|_2^2 \frac{dt}{t} = 4C\|\nabla f\|_2^2$$

where the last equality follows from Plancherel's theorem.

Now the elliptic estimates of Lemma 33 allows us to use Lemma 23 for $U_t = \theta_t P_t$.

Hence, SFE for $U_t P_t = \theta_t P_t^2$ is equivalent $|(\theta_t 1)(x)|^2 \frac{dxdt}{t}$ being a Carleson measure. Here 1 is the $n \times n$ unit matrix. Moreover, one can substitute S_t for P_t.

Summarizing we see that (32) reduces to proving that $|(\theta_t 1)(x)|^2 \frac{dxdt}{t}$ is a Carleson measure. Moreover, one has

$$\int_0^\infty \int_{\mathbb{R}^n} |(\theta_t \nabla f)(x) - (\theta_t 1)(x) \cdot (S_t \nabla f)(x)|^2 \frac{dxdt}{t} \leq C\|\nabla f\|_2^2. \tag{35}$$

Note that the product $(\theta_t 1)(x) \cdot (S_t \nabla f)(x)$ is the dot product $u_1 v_1 + \cdots + u_n v_n$ between two vectors in $\mathbb{C}^n$.

Now, we want to follow the ideas of the T(b) theorem. There, the product was over the complex field $\mathbb{C}$. Since we have now the dot product on $\mathbb{C}^n$, we make a sectorial decomposition of $\mathbb{C}^n$. Let $\varepsilon > 0$ to be chosen later and cover $\mathbb{C}^n$ with a finite number depending on ε and n of cones $\mathcal{C}_w$ associated to unit vectors w in $\mathbb{C}^n$ and defined by

$$|u - (u|w)w| \leq \varepsilon\ |(u|w)|. \tag{36}$$

Here $(\ |\ )$ is the complex inner product on $\mathbb{C}^n$. It suffices to argue for each w fixed and to obtain a Carleson measure estimate for

$$\gamma_{t,w}(x) = \mathbf{1}_{\mathcal{C}_w}((\theta_t 1)(x))(\theta_t 1)(x),$$

where $\mathbf{1}_{\mathcal{C}_w}$ denotes the indicator function of $\mathcal{C}_w$.

Fix w. We are looking for the analogs of the functions b_Q. We call them f_Q. The requirements we are looking for are

$$\int_{3Q} |\nabla f_Q|^2 \leq C|Q| \tag{37}$$

$$|\int_Q \nabla f_Q| \geq \delta |Q| \tag{38}$$

$$\int_Q \int_0^{\ell(Q)} |(\theta_t \nabla f_Q)(x)|^2 \frac{dxdt}{t} \leq C|Q| \tag{39}$$

and

$$|\gamma_{t,w}(x)| \leq C|\gamma_{t,w}(x) \cdot (S_t \nabla f_Q)(x)| \tag{40}$$

on "good" regions $\mathcal{R}_Q \setminus \cup \mathcal{R}_{Q_i}$ with not too many "bad" cubes that is, $\sum |Q_i| \leq (1-\eta)|Q|$.

The novelty is the last inequality which contains some geometry.

A candidate would be $f_Q(x) = (x - x_Q|w)$ with x_Q the centre of Q, because all but the third inequality are fulfilled. Since $\theta_t \nabla = (I + t^2 L)^{-1} tL$ it is natural to approximate f_Q by applying the resolvent to f_Q:

$$f_Q^\varepsilon = (I + \varepsilon^2 \ell(Q)^2 L)^{-1} f_Q$$

where ε is our small parameter. Note that f_Q^ε is an approximation to f_Q at the scale of Q. It is defined on all of $\mathbb{R}^n$ and Lemma 34 gives us $L^2(3Q)$- estimates for $f_Q - f_Q^\varepsilon$ and its gradient.

Hence, we obtain immediately (37) and C does not depend on ε. We have

$$\theta_t \nabla f_Q^\varepsilon = (I + t^2 L)^{-1} \frac{t}{\varepsilon^2 \ell(Q)^2} (f_Q - f_Q^\varepsilon)$$

and we deduce (39).

Now, to see (38) we observe that $\nabla f_Q = w^*$ (the conjugate of w) and write

$$|\int_Q \nabla f_Q^\varepsilon| \geq \mathrm{Re}(w^*|\int_Q \nabla f_Q^\varepsilon) = |Q| - \mathrm{Re} \int_Q (w^*|\nabla(f_Q - f_Q^\varepsilon)).$$

The inequality

$$\left| \int_Q \nabla h \right| \leq C\ell(Q)^{\frac{n-1}{2}} \left(\int_Q |h|^2 \right)^{1/4} \left(\int_Q |\nabla h|^2 \right)^{1/4}$$

and Lemma 34 imply

$$\mathrm{Re} \int_Q (w^*|\nabla(f_Q - f_Q^\varepsilon)) \leq C\varepsilon^{1/2}|Q|$$

and (38) follows provided ε is small enough.

It remains to obtain (40). The stopping-time argument of Section 7.2 would give us a lower bound of $\mathrm{Re}(w^*|(S_t \nabla f_Q^\varepsilon)(x))$ for (x,t) in the "good" region. Given the fact that $\gamma_{t,w}(x)$ belongs to the cone $\mathcal{C}_w$ this is not enough. We also need to control $|S_t f_Q^\varepsilon(x)|$ on this "good" region.

This means that we have to introduce in the stopping-time argument a second condition: starting from Q, we subdivide Q dyadically and stop the first time that either $\mathrm{Re}\int_{Q'}(w^*|\nabla f_Q) \leq \delta|Q'|$ or $|\int_{Q'}\nabla f_Q| \geq C\varepsilon^{-1}|Q'|$ where C is appropriately chosen. As before, the union of the selected bad cubes cannot cover all of Q if ε is small enough and we are done. For details, see [1].

References

[1] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh and Ph. Tchamitchian. The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$, preprint, 2001. To appear in Annals of Math.

[2] P. Auscher, S. Hofmann, A. McIntosh and Ph. Tchamitchian. The Kato square root problem for higher order elliptic operators and systems on $\mathbb{R}^n$, *J. Evol. Equ.* **1** (2001) 361-385.

[3] P. Auscher, A. McIntosh, and A. Nahmod. Holomorphic functional calculi of operators, quadratic estimates and interpolation. *Indiana Univ. Math. J.*, 46:375–403, 1997.

[4] P. Auscher and Ph. Tchamitchian. Conjecture de Kato sur les ouverts de $\mathbb{R}$. *Rev. Mat. Iberoamericana*, 8:149–199, 1992.

[5] P. Auscher and Ph. Tchamitchian, *Square root problem for divergence operators and related topics*, Astérisque Vol. 249 (1998), Société Mathématique de France.

[6] P. Auscher and Ph. Tchamitchian, Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^2 theory, preprint, 2001. To appear in J. Ana. Math.

[7] M. Christ, *A T(b) theorem with remarks on analytic capacity and the Cauchy integral,* Colloquium Mathematicum LX/LXI (1990) 601-628.

[8] M. Christ and J.-L. Journé. Polynomial growth estimates for multilinear singular integral operators. *Acta Math.*, 159:51–80, 1987.

[9] R. Coifman, A. McIntosh, and Y. Meyer. L'intégrale de Cauchy définit un opérateur borné sur $L^2(\mathbb{R})$ pour les courbes lipschitziennes. *Ann. Math.*, 116:361–387, 1982.

[10] R. Coifman and Y. Meyer. Non-linear harmonic analysis and PDE. In E. M. Stein, editor, *Beijing Lectures in Harmonic Analysis*, volume 112 of *Annals of Math. Studies.* Princeton Univ. Press, 1986.

[11] T. Kato. *Perturbation theory for linear operators.* Springer Verlag, New York, 1966.

[12] T. Kato, *Fractional powers of dissipative operators*, J. Math. Soc. Japan **13** (1961), p. 246-274.

[13] J.-L. Lions. Espaces d'interpolation et domaines de puissances fractionnaires. *J. Math. Soc. Japan*, 14:233–241, 1962.

[14] A. McIntosh, *On the Comparability of* $A^{1/2}$ *and* $A^{*1/2}$, Proc. Amer. Math. Soc. 32 (1972), 430-434.

[15] A. McIntosh and A. Yagi. Operators of type ω without a bounded H^∞ functional calculus. In *Miniconference on Operators in Analysis*, volume 24 of *Center for Math. and Appl.*, pages 159–174, Canberra, 1989. Australian National Univ.

[16] S. Semmes. Square function estimates and the $T(b)$ Theorem. *Proc. Amer. Math. Soc.*, 110(3):721–726, 1990.

Laboratoire de mathématiques, CNRS UMR 8628, Université Paris-Sud — Bât 425, 91405 Orsay Cedex FRANCE
E-mail address: pascal.auscher@math.u-psud.fr