The solution of the Kato Square Root Problem for Second Order Elliptic Operators on $\mathbb{R}^n$

Pascal Auscher
LAMFA, CNRS, FRE 2270
Université de Picardie-Jules Verne
33, rue Saint Leu
F-80039 AmiensCedex 1
email: auscher@mathinfo.u-picardie.fr

Steve Hofmann*
Department of Mathematics
University of Missouri-Columbia
Columbia, MO 65211
email: hofmann@math.missouri.edu

Michael Lacey†
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332
email: lacey@math.gatech.edu

Alan M^{c}Intosh
Centre for Mathematics and its Applications
Australian National University
Canberra, ACT 0200
email: alan@maths.anu.edu.au

Philippe Tchamitchian
Faculté des Sciences et
Techniques de St.-Jérome
Université d'Aix-Marseille III
Avenue Escadrille Normandie-Niemen
and LATP, CNRS, UMR 6632
F-13397 Marseille Cedex 20
email: tchamphi@math.u-3mrs.fr

January 31, 2001, revised July 5, 2001

dedicated to the memories of T. Kato and J. L. Lions

Abstract

We prove the Kato conjecture for elliptic operators on $\mathbb{R}^n$. More precisely, we establish that the domain of the square root of a uniformly complex elliptic operator $L = -\mathrm{div}\,(A\nabla)$ with bounded measurable coefficients in $\mathbb{R}^n$ is the Sobolev space $H^1(\mathbb{R}^n)$ in any dimension with the estimate $\|\sqrt{L}f\|_2 \sim \|\nabla f\|_2$.

1 Introduction, history and statement of the main results

Let $A = A(x)$ be an $n \times n$ matrix of complex, L^∞ coefficients, defined on $\mathbb{R}^n$, and satisfying the ellipticity (or "accretivity") condition

$$\lambda|\xi|^2 \leq \mathrm{Re}\, A\xi \cdot \xi^* \text{ and } |A\xi \cdot \zeta^*| \leq \Lambda|\xi||\zeta|, \tag{1.1}$$

for $\xi, \zeta \in \mathbb{C}^n$ and for some λ, Λ such that $0 < \lambda \leq \Lambda < \infty$. Here, $u \cdot v = u_1 v_1 + \cdots + u_n v_n$ and u^* is the complex conjugate of u so that $u \cdot v^*$ is the usual inner product in $\mathbb{C}^n$ and, therefore, $A\xi \cdot \zeta^* \equiv \sum_{j,k} a_{j,k}(x)\xi_k\, \overline{\zeta}_j$. We define a second order divergence form operator

$$Lf \equiv -\mathrm{div}(A\nabla f), \tag{1.2}$$

*supported by NSF
†supported by NSF

which we interpret in the usual weak sense via a sesquilinear form.

The accretivity condition (1.1) enables one to define a square root $L^{1/2} \equiv \sqrt{L}$ (see [23]), and a fundamental question is to determine whether one can solve the "square root problem", i.e. establish the estimate

$$\|\sqrt{L}f\|_2 \sim \|\nabla f\|_2, \tag{1.3}$$

where $\sim$ is the equivalence in the sense of norms, with constants C depending only on n, λ and Λ, and $\|f\|_2 = (\int_{\mathbb{R}^n} |f(x)|_H^2 \, dx)^{1/2}$ denotes the usual norm for functions on $\mathbb{R}^n$ valued in a Hilbert space H. We answer here this question in the affirmative.

Theorem 1.4. *For any operator as above the domain of $\sqrt{L}$ coincides with the Sobolev space $H^1(\mathbb{R}^n)$ and $\|\sqrt{L}f\|_2 \sim \|\nabla f\|_2$.*

This has been a long-standing open problem, essentially posed by Kato [23], and refined by M^{c}Intosh [28, 27]. Kato actually formulated this question for a more general class of abstract maximal accretive operators. A counterexample to the abstract problem was found by Lions [25] and, for the maximal accretive operators arising from a form by M^{c}Intosh [26]. However, it has been pointed out in [28] that, in posing the problem, Kato had been motivated by the special case of elliptic differential operators, and by the applicability of a positive result, in that special case, to the perturbation theory for parabolic and hyperbolic evolution equations. For example, the application to hyperbolic equations depends on the validity of (1.3) in a complex L^∞-neighborhood of real and symmetric matrices.

The problem has a long history, and a number of people have contributed to its solution. First, Coifman, M^{c}Intosh and Meyer [12] proved Theorem 1.4 in one dimension, simultaneously with their proof of the L^2-boundedness of the Cauchy integral along a Lipschitz curve. In fact, the two results are known to be equivalent, see [24] or [6].

The first positive results in higher dimensions exploited the same technique as had been used in one dimension, namely that of multilinear operators. Independently, Coifman, Deng and Meyer [11], and Fabes, Jerison and Kenig [18] established the square root estimate (1.3) provided $\|A - I\|_\infty \leq \epsilon(n)$. Clearly, their methods allowed one also to replace the identity matrix I by any constant accretive matrix (see [19]). David and Journé gave a different proof using the T(1) theorem [14]. Sharper bounds for the constant $\epsilon(n)$ on the order of $n^{-\frac{1}{2}}$ were obtained by Journé [22].

The multilinear expansion method also may be extended to operators with smooth (i.e., locally near constant) coefficients; in this case, one obtains an equivalence of inhomogeneous norms: M^{c}Intosh in [29] considered coefficients being multipliers in some Sobolev space H^s, $s > 0$, Fabes, Jerison, and Kenig, uniformly continuous coefficients (unpublished), and Escauriaza VMO coefficients (unpublished). In addition, Alexopoulos [1] used homogenisation techniques for real Hölder continuous coefficients with periodicity, obtaining a homogeneous estimate.

In [9], two of us proposed another method to attack the problem. This initially led to some improvement of earlier results (such as VMO improved to a bigger subspace of BMO called $ABMO$) and the observation that one could get away from the perturbation cases at the expense of imposing some structure on the matrix A. The key notes of this method are 1) the use of functional calculus and, in particular, pointwise estimates on the heat kernel 2) the reduction to a Carleson measure estimate, and 3) the introduction of a "T(b) theorem for square roots" in the spirit of the T(b) theorems for singular integrals of M^{c}Intosh and Meyer [30], and of David, Journé and Semmes [15], and based on the alternative proof of Semmes [32]. We note that those T(b) theorems were motivated by the Cauchy integral.

The control of the Carleson measure in point 2) above has been achieved very recently in two ways both exploiting the T(b) theorem for square roots. Auscher, Lewis, Hofmann and Tchamitchian [5] use an extrapolation technique for Carleson measures involving a stopping-time decomposition of the Carleson region to prove the Kato conjecture for perturbations of real symmetric operators in any dimension, which was, as mentioned earlier, one of Kato's original motivations. By a different stopping-time argument, Hofmann, Lacey and M^{c}Intosh [21] prove the Kato conjecture under a restriction of sufficient pointwise decay of the heat kernel.

Pointwise decay is available for real operators by results of Aronson [2] and in some cases for complex operators: in two dimensions by a result of Auscher, M^{c}Intosh and Tchamitchian [7] and for perturbations of real operators or even for small coefficients in BMO-norm by results of Auscher [3]. Hence, [21] solves the conjecture in two dimensions and includes, in particular, the result of [5].

But heat kernel decay may fail for complex operators: counterexamples are due to Auscher, Coulhon and Tchamitchian [4]. Thus, to solve the Kato conjecture in all dimensions it remains to remove the pointwise upper bound assumption in [21]. This is the main new contribution of this article. It turns out, as we will show, that there is enough decay in some averaged sense to carry out the reduction to a Carleson measure estimate, to develop an appropriate version of the T(b) theorem for square roots of [9] and to adapt the stopping-time argument of [21]. This decay is akin to that first proved by Gaffney [20] for the Laplace-Beltrami operator of a complete Riemannian manifold, and is valid for complex operators as in (1.1) and (1.2).

The proof of Theorem 1.4 (Sections 2-5) will be essentially self-contained assuming the basic background on functional calculus for accretive operators and on Littlewood-Paley theory. While having the added virtue of improving the paper's readability, this degree of completeness is for the most part required, as we are forced to redevelop material from [9] and [21] under necessarily weaker hypotheses. We note that the proof works for $n \geq 1$.

We shall conclude this article in Section 6, by stating some miscellaneous results concerning perturbations by lower order terms, and extensions to L^p results.

We note that the Kato conjecture for higher order operators on $\mathbb{R}^n$ can also be solved. For systems on $\mathbb{R}^n$, the Kato conjecture remains open in full generality, yet the extrapolation method is extendable to perturbations of self-adjoint systems. Also the Kato conjecture for second order elliptic operators on domains with boundary conditions can be obtained. These results will be presented elsewhere.

Acknowledgments The second named author is grateful to T. Toro, and the fourth author to M. Christ, for pointing out to us that the stopping time construction of Proposition 5.7 is similar in spirit to that of [10]. We also thank X. Duong and Lixin Yan for helpful comments concerning the statement of Proposition 6.2 below.

2 Estimates for elliptic operators on $\mathbb{R}^n$

We are given an elliptic operator as in (1.2) with ellipticity constants λ and Λ in (1.1). An observation of constant use in this paper is that the operators $(1+t^2L)^{-1}$, $t\nabla(1+t^2L)^{-1}$, $(1+t^2L)^{-1}t\mathrm{div}$ and $t^2\nabla(1+t^2L)^{-1}\mathrm{div}$ are uniformly L^2 bounded with bounds depending only on n, λ and Λ. Here and in the rest of the paper, $||T||_{op}$ denotes the operator norm of an operator acting from $L^2(\mathbb{R}^n;\mathbb{C}^p)$ into $L^2(\mathbb{R}^n;\mathbb{C}^q)$ for p, q integers depending on the context. Also, we shall consistently use boldface letters to denote vector-valued functions.

In this section, we record some technical lemmata. We begin with an estimate which expresses the decay of the resolvent kernel "in the mean."

Lemma 2.1. *Let E and F be two closed sets of $\mathbb{R}^n$ and set $d = \mathrm{dist}(E, F)$, the distance between E and F. Then*

$$\int_F |(1+t^2L)^{-1}f(x)|^2\,dx \le Ce^{-\frac{d}{ct}}\int_E |f(x)|^2\,dx, \quad \mathrm{Supp}\, f \subset E,$$

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

$$\int_F |(1+t^2L)^{-1}t\mathrm{div}\,\mathbf{f}\,(x)|^2\,dx \le Ce^{-\frac{d}{ct}}\int_E |\,\mathbf{f}\,(x)|^2\,dx, \quad \mathrm{Supp}\,\mathbf{f}\, \subset E,$$

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

Proof. It suffices to obtain the inequalities for $d \ge t > 0$. The argument uses a Caccioppoli type inequality. Set $u_t = (1+t^2L)^{-1}f$. For all $v \in H^1(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} u_t v + t^2\int_{\mathbb{R}^n} A\nabla u_t\cdot\nabla v = \int_{\mathbb{R}^n} fv.$$

Taking $v = \overline{u_t}\eta^2$ with $\eta \in C_0^\infty(\mathbb{R}^n)$ supported outside of E with η positive and $\|\eta\|_\infty = 1$ and using that $\mathrm{supp}\, f \subset E$, we have

$$\int_{\mathbb{R}^n} |u_t|^2\eta^2 + t^2\int_{\mathbb{R}^n} A\nabla u_t\cdot\overline{\nabla u_t}\;\eta^2 = -2t^2\int_{\mathbb{R}^n} A(\eta\nabla u_t)\cdot\overline{u_t\nabla\eta}.$$

Using (1.1) and the inequality $2|ab| \le \varepsilon|a|^2 + \varepsilon^{-1}|b|^2$, we obtain for all $\varepsilon > 0$

$$\int_{\mathbb{R}^n} |u_t|^2\eta^2 + \lambda t^2\int_{\mathbb{R}^n} |\nabla u_t|^2\;\eta^2 \le \Lambda\varepsilon t^2\int_{\mathbb{R}^n} |\nabla u_t|^2\;\eta^2 + \Lambda\varepsilon^{-1}t^2\int_{\mathbb{R}^n} |u_t|^2|\nabla\eta|^2.$$

Choosing $\varepsilon = \frac{\lambda}{\Lambda}$ leads to

$$\int_{\mathbb{R}^n} |u_t|^2\eta^2 \le \frac{\Lambda^2t^2}{\lambda}\int_{\mathbb{R}^n} |u_t|^2|\nabla\eta|^2.$$

Replacing η by $e^{\alpha\eta} - 1$ with $\alpha = \frac{\sqrt{\lambda}}{2\Lambda t\|\nabla\eta\|_\infty}$ yields

$$\int_{\mathbb{R}^n} |u_t|^2|e^{\alpha\eta} - 1|^2 \le \frac{1}{4}\int_{\mathbb{R}^n} |u_t|^2|e^{\alpha\eta}|^2$$

so that a simple triangle inequality gives us

$$\int_{\mathbb{R}^n} |u_t|^2 |e^{\alpha\eta}|^2 \leq 4 \int_{\mathbb{R}^n} |u_t|^2 \leq 4 \int_E |f|^2.$$

Assuming furthermore $\eta = 1$ on F, we have

$$|e^\alpha|^2 \int_F |u_t|^2 \leq \int_{\mathbb{R}^n} |u_t|^2 |e^{\alpha\eta}|^2$$

and it remains to impose $\|\nabla\eta\|_\infty \sim 1/d$ to conclude for the first inequality.

Next, choose $\varepsilon = \frac{\lambda}{2\Lambda}$ and η as before to obtain

$$\int_F |t\nabla u_t|^2 \leq \int_{\mathbb{R}^n} |t\nabla u_t|^2 \eta^2 \leq \frac{2\Lambda^2 t^2}{\lambda} \int_{\mathbb{R}^n} |u_t|^2 |\nabla\eta|^2 \leq Ct^2 d^{-2} e^{-\frac{d}{ct}} \int_E |f|^2,$$

which gives us the second inequality.

The third inequality is obtained by duality from the second one applied to $L^* = -\mathrm{div}\,(A^*\nabla)$ and exchanging the roles of E and F. □

Remark. Using complex times t and a Cauchy integral, we can obtain

$$\|e^{-tL} f\|_{L^2(F)} \leq Ce^{-\frac{d^2}{ct}} \|f\|_{L^2(E)}, \quad \mathrm{Supp}\, f \subset E.$$

When L is a Laplace-Beltrami operator, this is Gaffney's estimate [20] (See also Davies [16] for an argument which adapts to our situation).

Lemma 2.2. *For any Lipschitz function f and $t > 0$,*

$$||[(1+t^2L)^{-1}, f]||_{op} \leq Ct\|\nabla f\|_\infty$$

and

$$||\nabla[(1+t^2L)^{-1}, f]||_{op} \leq C\|\nabla f\|_\infty$$

where C depend only on n, λ and Λ. Here, f denotes the operator of pointwise multiplication by f and $[,]$ is a commutator.

Proof. Write

$$\begin{aligned}[(1+t^2L)^{-1}, f] &= -(1+t^2L)^{-1}[(1+t^2L), f](1+t^2L)^{-1} \\ &= -(1+t^2L)^{-1} t^2 (\mathrm{div}\,\mathbf{b} + \tilde{\mathbf{b}} \cdot \nabla)(1+t^2L)^{-1},\end{aligned}$$

where we have set $\mathbf{b} = A\nabla f$, $\tilde{\mathbf{b}} = A^T\nabla f$ considered as operators of pointwise multiplication, so that their operator norms are controlled by $C\|\nabla f\|_\infty$. The uniform L^2 boundedness of $(1+t^2L)^{-1}$, $t\nabla(1+t^2L)^{-1}$ and $(1+t^2L)^{-1}t\mathrm{div}$ imply the first commutator estimate. Using also the L^2 boundedness of $t^2\nabla(1+t^2L)^{-1}\mathrm{div}$ yields the second one. □

By cube in $\mathbb{R}^n$, we mean a cube with sides parallel to the axes. If Q is a cube, then $|Q|$ and $\ell(Q)$ denote respectively its measure, its sidelength. We use also the notation cQ to denote the concentric cube with Q having sidelength $c\,\ell(Q)$.

Lemma 2.3. *For some C depending only on n, λ and Λ, if Q is a cube in $\mathbb{R}^n$, $t \le \ell(Q)$ and f is Lipschitz function on $\mathbb{R}^n$ then we have*

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

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

Proof. The argument will make clear that $(1+t^2L)^{-1}f$ is defined as $\sum(1+t^2L)^{-1}(f\mathcal{X}_k)$ with convergence in $L^2_{\mathrm{loc}}(\mathbb{R}^n)$, where the $\mathcal{X}_k$ is a partition of unity. It is an easy matter to verify that this definition does not depend on the particular choice of the partition.

By rescaling, there is no loss of generality to assume that $\ell(Q) = 1$ and that $\|\nabla f\|_\infty = 1$. Pick a partition (Q_k) of $\mathbb{R}^n$ by cubes of sidelengths 2 and with $Q_0 = 2Q$. Let $\mathcal{X}_k$ be the indicator function of Q_k. The off-diagonal estimates imply that $(1+t^2L)^{-1}(1) = 1$ in the sense that $\lim_{R\to\infty}(1+t^2L)^{-1}(\eta_R) = 1$ in $L^2_{\mathrm{loc}}(\mathbb{R}^n)$ where $\eta_R(x) = \eta(x/R)$ and η is a smooth bump function with $\eta \equiv 1$ near 0. Hence, we may write

$$(1+t^2L)^{-1}f(x) - f(x) = \sum_{k\in\mathbb{Z}^n}(1+t^2L)^{-1}((f-f(x))\mathcal{X}_k)(x) \equiv \sum_{k\in\mathbb{Z}^n} g_k(x).$$

The term for $k=0$ is nothing but $[(1+t^2L)^{-1}, f](\mathcal{X}_0)(x)$. Hence, its $L^2(Q)$-norm is controlled by $Ct\|\mathcal{X}_0\|_2$ by the first commutator estimate. The terms for $k \ne 0$ are treated using the further decomposition

$$g_k(x) = (1+t^2L)^{-1}((f-f(x_k))\mathcal{X}_k)(x) + (f(x_k)-f(x))(1+t^2L)^{-1}(\mathcal{X}_k)(x)$$

where x_k is the center of Q_k. Using the off-diagonal estimates for $(1+t^2L)^{-1}$ on sets $E = Q_k$ and $F = Q$ and the fact that f is Lipschitz, we get

$$\int_Q |g_k|^2 \leq Ct^2 e^{-\frac{|x_k|}{ct}} \|\mathcal{X}_k\|_2^2 = Ct^2 e^{-\frac{|x_k|}{ct}} 2^n |Q|.$$

The desired bound on the $L^2(Q)$-norm of $(1+t^2L)^{-1}f - f$ follows from these estimates, Minkowski's inequality and the fact that $t \leq 1$.

The proof of the second inequality uses a similar argument and is left to the reader. □

3 Reduction to a quadratic estimate

We are given an elliptic operator as in (1.2) with ellipticity constants λ and Λ in (1.1). We wish to prove *a priori* that

$$\text{(K)} \qquad \|\sqrt{L}f\|_2 \leq C\|\nabla f\|_2,$$

for f in some dense subspace of $H^1(\mathbb{R}^n)$ with C depending only on n, λ and Λ. Then (K) also holds for L^* as the hypotheses are stable under taking adjoints. Eventually, we conclude by a theorem of J.L. Lions [25] that the domain of $\sqrt{L}$ is $H^1(\mathbb{R}^n)$ and that for $f \in H^1(\mathbb{R}^n)$,

$$\|\sqrt{L}f\|_2 \sim \|\nabla f\|_2.$$

We remark that to prove (K), we may and do assume that the coefficients are C^∞ as long as we do not use this quantitatively in our estimates. This is why we shall make clear the dependance of constants. Then, one removes this assumption using a slight variant of [9, Chapter 0, Proposition 7].

To begin, we use the following resolution of the square root:

$$\sqrt{L}f = a\int_0^{+\infty} (1+t^2L)^{-3}t^3L^2f\,\frac{dt}{t},$$

where a^{-1} is the value of $\int_0^\infty (1+u^2)^{-3}u^2\,du$, and the integral converges normally in $L^2(\mathbb{R}^n)$ for $f \in C_0^\infty(\mathbb{R}^n)$ (as $C_0^\infty(\mathbb{R}^n) \subset H^4(\mathbb{R}^n) = \mathcal{D}(L^2)$ under the smoothness assumption). Take $g \in C_0^\infty(\mathbb{R}^n)$ with $\|g\|_2 = 1$. By duality and the Cauchy-Schwarz inequality

$$\text{(3.1)} \qquad |\langle\sqrt{L}f, g\rangle|^2 \leq a^2 \int_0^{+\infty} \|(1+t^2L)^{-1}tLf\|_2^2\,\frac{dt}{t} \int_0^{+\infty} \|V_t g\|_2^2\,\frac{dt}{t}$$

where $V_t = t^2 L^*(1+t^2L^*)^{-2}$. There are several ways to see that

$$\int_0^{+\infty} \|V_t g\|_2^2 \frac{dt}{t} \le C\|g\|_2^2 \le C.$$

One way is to appeal to the quadratic estimates of M^{c}Intosh and Yagi [31] since L^* has H^∞-functional calculus. Another way is to use the standard orthogonality arguments of Littlewood-Paley theory. As we shall use this again later, let us recall the method. Pick any $\psi \in C_0^\infty(\mathbb{R}^n)$ with ψ real-valued and $\int \psi = 0$ and define Q_s as the operator of convolution with $\frac{1}{s^n}\psi(\frac{x}{s})$ for $s > 0$ normalized so that

$$\int_0^{+\infty} \|Q_s g\|_2^2 \frac{ds}{s} = \|g\|_2^2.$$

Lemma 3.2. *Let $U_t : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, $t > 0$, be a family of bounded operators with $||U_t||_{op} \le 1$. If $||U_tQ_s||_{op} \le \big(\inf(\frac{t}{s}, \frac{s}{t})\big)^\alpha$, $\alpha > 0$, for a family $Q_s, s > 0$, as above, then for some constant C depending only on α,*

$$\int_0^{+\infty} \|U_t g\|_2^2 \frac{dt}{t} \le C\|g\|_2^2.$$

Proof. The argument is quite standard and follows from Schur's lemma. Details are left to the reader. □

Let us apply this to the operators V_t above which have uniform (in t) bounded extension to $L^2(\mathbb{R}^n)$. Since $V_tQ_s = -(1+t^2L^*)^{-2}t^2 \mathrm{div}\, A^*\nabla Q_s$, we have

$$||V_tQ_s||_{op} \le ||(1+t^2L^*)^{-2}t^2\mathrm{div}\, A^*||_{op}||\nabla Q_s||_{op} \le cts^{-1},$$

with c depending only on n, λ and Λ. Choose $\psi = \Delta\phi$ with $\phi \in C_0^\infty(\mathbb{R}^n)$, radial, so that, in particular, $\psi = \mathrm{div}\, \mathbf{h}$. This yields $Q_s = s\, \mathrm{div}\, \mathbf{R}_s$ with $\mathbf{R}_s$ uniformly bounded, hence

$$||V_tQ_s||_{op} \le ||t^2L^*(1+t^2L^*)^{-2}\mathrm{div}\, ||_{op}||s\mathbf{R}_s||_{op} \le ct^{-1}s,$$

with c depending only on n, λ and Λ.

Thus, the second integral in the right hand side of (3.1) is bounded, so we are reduced to proving

$$\int_0^{+\infty} \|(1+t^2L)^{-1}tLf\|_2^2 \frac{dt}{t} \le C\int_{\mathbb{R}^n} |\nabla f|^2. \tag{3.3}$$

4 Reduction to a Carleson measure estimate

We next reduce matters to checking a Carleson measure estimate. Let us first introduce some notation used throughout. Define for $\mathbb{C}^n$-valued functions $\mathbf{f} = (f_1, \dots, f_n)$,

$$\theta_t \mathbf{f} = -(1 + t^2 L)^{-1} t \partial_j (a_{j,k} f_k).$$

We use the summation convention for repeated indices. In short, we write $\theta_t = -(1 + t^2 L)^{-1} t \mathrm{div}\, A$. With this notation, (3.3) rewrites

$$\int_0^{+\infty} \|\theta_t \nabla f\|_2^2 \frac{dt}{t} \leq C \int_{\mathbb{R}^n} |\nabla f|^2. \tag{4.1}$$

Define also

$$\gamma_t(x) = (\theta_t \mathbf{1})(x) = (-(1 + t^2 L)^{-1} t \partial_j a_{j,k})(x))_{1 \leq k \leq n}$$

where $\mathbf{1}$ is the $n \times n-$identity matrix, the action of θ_t on $\mathbf{1}$ being columnwise.

Borrowing an idea from [13], the reduction to a Carleson measure estimate and the T(b) argument will require the inequality

$$\int_{\mathbb{R}^n} \int_0^{+\infty} |\gamma_t(x) \cdot (P_t^2 \nabla g)(x) - (\theta_t \nabla g)(x)|^2 \frac{dxdt}{t} \leq C \int_{\mathbb{R}^n} |\nabla g|^2, \tag{4.2}$$

where C depends only on n, λ and Λ. Here, P_t denotes the operator of convolution with $\frac{1}{t^n} p(\frac{x}{t})$ where p is a smooth real-valued function supported in the unit ball of $\mathbb{R}^n$ with $\int p = 1$. The notation $u \cdot v$ for $u, v \in \mathbb{C}^n$ is the one in the Introduction. To prove this, we need to handle Littlewood-Paley theory just outside the classical setting.

Lemma 4.3. *Let $U_t : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, $t > 0$, be a measurable family of bounded operators with $||U_t||_{op} \leq 1$. 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 $t > 0$,*

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

(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 $L^2(B(y, t))$ and for all f and $y \in \mathbb{R}^n$,*

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

(iii) $U_t(1) = 0$ *in the sense that* $U_t(\mathcal{X}_R)$ *converges to 0 in* $L^2(B(y,t))$ *as* $R \to \infty$ *for any* y, *where* $\mathcal{X}_R$ *stands for the indicator function of the ball* $B(0,R)$.

Let Q_s, P_t, *be as above. Then for some* $\alpha > 0$ *and* C *depending on* n *and* m,

$$||U_t P_t Q_s||_{op} \le C \left(\inf \left(\frac{t}{s}, \frac{s}{t} \right) \right)^{\alpha}.$$

Proof. We first remark that $U_t^* U_t$ has a kernel satisfying

$$|K_t(x,y)| \le \frac{1}{t^n} \left(1 + \frac{|x-y|}{t} \right)^{-m}.$$

Indeed $K_t(x,y)$ is given by

$$K_t(x,y) = \int_{\mathbb{R}^n} \overline{U_t(z,x)} U_t(z,y)\, dz$$

so that the formula $(1+a+b) \le (1+a)(1+b)$ for positive a, b gives us that $\left(1 + \frac{|x-y|}{t}\right)^m |K_t(x,y)|$ is bounded by

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

from the Cauchy-Schwarz inequality and (i). Hence $U_t^* U_t$ is bounded on all L^p, $1 \le p \le +\infty$ and in particular for $p = 2$, we recover the boundedness of U_t (thus, one can drop from the hypotheses the L^2 boundedness of U_t).

For $s \le t$, by $||U_t||_{op} \le 1$ and standard Fourier analysis we have that

$$||U_t P_t Q_s||_{op} \le ||P_t Q_s||_{op} \le C \left(\frac{s}{t} \right)^{\alpha}.$$

Next, we consider $t \le s$. Since P_t has a nice kernel, $W_t = U_t^* U_t P_t$ also has an L^1 kernel. If we prove that $W_t(1) = 0$ then we can deduce from standard arguments that

$$||W_t Q_s||_{op} \le C \left(\frac{t}{s} \right)^{\alpha},$$

for $0 < \alpha < m-n$, which gives us the result as $||U_t P_t Q_s||_{op}^2 \le C ||U_t^* U_t P_t Q_s||_{op}$.

We have that $W_t(1) = U_t^* U_t(1)$. If $\varphi \in L^2(\mathbb{R}^n)$ is compactly supported (thus in $L^1(\mathbb{R}^n)$) then

$$\langle U_t^* U_t(1), \varphi \rangle = \lim_{R\to\infty} \langle U_t^* U_t(\mathcal{X}_R), \varphi \rangle = \lim_{R\to\infty} \langle U_t(\mathcal{X}_R), U_t(\varphi) \rangle.$$

Now

$$\langle U_t(\mathcal{X}_R), U_t(\varphi) \rangle = \iint U_t(\mathcal{X}_R)(x) \overline{U_t(x,y)\varphi(y)}\, dydx$$

which is, in modulus, less than a constant times

$$\left(t^{-n} \iint \left(1 + \frac{|x-y|}{t}\right)^{-2m} |U_t(\mathcal{X}_R)(x)|^2 |\varphi(y)| dydx \right)^{1/2} \|\varphi\|_1^{1/2}$$

by (i) and the Cauchy-Schwarz inequality for the measure $|\varphi(y)|dydx$. Using a covering in the x variable by a lattice of balls $B(y + ckt, t)$, $k \in \mathbb{Z}^n$, we obtain a bound

$$C \left(\sum_{k\in\mathbb{Z}^n} \int_{\mathbb{R}^n} (1 + |k|)^{-2m} c_R(y,k) |\varphi(y)| dy \right)^{1/2} \|\varphi\|_1^{1/2}$$

with $c_R(y,k) = t^{-n} \int_{B(y+ckt,t)} |U_t(\mathcal{X}_R)(x)|^2 dx$. It remains to apply the dominated convergence theorem by invoking (ii) and (iii) as R tends to ∞. □

Lemma 4.4. *Let P_t be as in Lemma 4.3. Then the operator U_t defined by $U_t \mathbf{f}\,(x) = \gamma_t(x) \cdot (P_t \mathbf{f}\,)(x) - (\theta_t P_t \mathbf{f}\,)(x)$ satisfies*

$$\int_0^{+\infty} \|U_t P_t \mathbf{f}\,\|_2^2 \frac{dt}{t} \le C \|\,\mathbf{f}\,\|_2^2$$

where C depends only on n, λ and Λ. Here the action of P_t on $\mathbf{f}$ is componentwise.

Proof. By the off-diagonal estimates of Lemma 2.1 for θ_t and the fact that p has support in the unit ball, it is easy to show that there is a constant C depending on n, λ and Λ such that for all $y \in \mathbb{R}^n$

$$\frac{1}{t^n} \int_{B(y,t)} |\gamma_t(x)|^2 \, dx \le C$$

and that the kernel of $C^{-1} U_t$ satisfies the hypotheses in Lemma 4.3. The conclusion follows from Lemma 3.2 applied to $U_t P_t$. □

We can now prove (4.2). We begin by writing

$$\gamma_t(x) \cdot (P_t^2 \nabla g)(x) - (\theta_t \nabla g)(x) = (U_t P_t \nabla g)(x) + (\theta_t (P_t^2 - I)\nabla g)(x).$$

The first term is taken care of by the above lemma. As P_t commutes with partial derivatives, we may use that $||\theta_t \nabla||_{op} = ||(1+t^2L)^{-1}tL||_{op} \leq Ct^{-1}$, so that we obtain for the second term

$$\begin{aligned}\int_{\mathbb{R}^n}\int_0^{+\infty} |(\theta_t(P_t^2 - I)\nabla g)(x)|^2 \frac{dxdt}{t} &\leq C^2 \int_{\mathbb{R}^n}\int_0^{+\infty} |((P_t^2 - I)g)(x)|^2 \frac{dt}{t^3}dx \\ &\leq C^2 c(p)\|\nabla g\|_2^2\end{aligned}$$

by the Plancherel theorem with C depending only on n, λ and Λ. This concludes the proof of (4.2).

Lemma 4.5. *The inequality (K) follows from the Carleson measure estimate*

$$\sup_Q \frac{1}{|Q|}\int_Q\int_0^{\ell(Q)} |\gamma_t(x)|^2 \frac{dxdt}{t} < \infty, \tag{4.6}$$

where the supremum is taken over all cubes in $\mathbb{R}^n$.

Proof. Indeed, (4.6) and Carleson's inequality imply

$$\int_{\mathbb{R}^n}\int_0^{+\infty} |\gamma_t(x) \cdot (P_t^2 \nabla g)(x)|^2 \frac{dxdt}{t} \leq C \int_{\mathbb{R}^n} |\nabla g|^2,$$

and together with (4.2) we deduce that (4.1) holds. □

Remark. We shall not need the easy converse that (K) implies (4.6).

To finish this section, let us state a technical lemma for later use. Let Q be a cube in $\mathbb{R}^n$, consider a collection of dyadic cubes of $\mathbb{R}^n$ that contains Q and let S_t^Q be the corresponding dyadic averaging operator:

$$S_t^Q \mathbf{f}(x) = \frac{1}{|Q'|}\int_{Q'} \mathbf{f}(y)\, dy$$

for x in the dyadic cube Q' and $\frac{1}{2}\ell(Q') < t \leq \ell(Q')$.

Lemma 4.7. *For some C depending only on n, λ and Λ, we have*

$$\int_Q\int_0^{\ell(Q)} |\gamma_t(x) \cdot ((S_t^Q - P_t^2)\,\mathbf{f}\,)(x)|^2 \frac{dxdt}{t} \leq C \int_{\mathbb{R}^n} |\,\mathbf{f}\,|^2.$$

Proof. Of course, integration can be performed on $\mathbb{R}^n \times (0, +\infty)$. We may adapt the proof of Lemma 4.3, given the following two observations. First, the operator $U_t = (\gamma_t \cdot S_t^Q)$ is L^2 bounded from the $U_t^* U_t$ argument and the condition (i). Second, S_t^Q is an orthogonal projection. Hence,

$$||(\gamma_t \cdot S_t^Q)Q_s||_{op} = ||(\gamma_t \cdot S_t^Q)S_t^Q Q_s||_{op} \leq C||S_t^Q Q_s||_{op} \leq C\left(\frac{s}{t}\right)^\alpha .$$

The last inequality follows from the well-known fact that, for any $\alpha \in (0, 1/2)$, the dyadic averaging operator maps $L^2(\mathbb{R}^n)$ into the homogeneous Sobolev space $\dot{H}^\alpha(\mathbb{R}^n)$ with norm $Ct^{-\alpha}$ (See [9, Appendix C] for a proof). Further details are left to the reader. □

5 The T(b) argument

To obtain (4.6), we adapt the construction of [21] to verify a variant of the T(b) theorem for square roots [9, Chapter 3, Theorem 3]. Fix a cube Q, $\varepsilon \in (0, 1)$, a unit vector w in $\mathbb{C}^n$ and define a scalar-valued function

$$f^\varepsilon_{Q,w} = (1 + (\varepsilon\ell(Q))^2 L)^{-1}(\Phi_Q \cdot w^*) \tag{5.1}$$

where, denoting by x_Q the center of Q,

$$\Phi_Q(x) = x - x_Q \in \mathbb{R}^n.$$

Let us record some estimates that follow straightforwardly from Lemma 2.3 as the reader may check.

$$\int_{5Q} |f^\varepsilon_{Q,w} - \Phi_Q \cdot w^*|^2 \leq C_1 \varepsilon^2 \ell(Q)^2 |Q| \tag{5.2}$$

and

$$\int_{5Q} |\nabla(f^\varepsilon_{Q,w} - \Phi_Q \cdot w^*)|^2 \leq C_2 |Q| \tag{5.3}$$

where C_1, C_2 depend on n, λ, Λ and not on ε, Q and w. It is an important fact that the constants C_1, C_2 above are independent of ε.

The proof of (4.6) follows immediately from the combination of the next two lemmata and the rest of this section is devoted to their proofs.

Lemma 5.4. *There exists an $\varepsilon > 0$ depending on n, λ, Λ, and a finite set W of unit vectors in $\mathbb{C}^n$ whose cardinality depends on ε and n, such that*

$$\sup \frac{1}{|Q|} \int_Q \int_0^{\ell(Q)} |\gamma_t(x)|^2 \frac{dxdt}{t} \leq$$
$$C \sum_{w \in W} \sup \frac{1}{|Q|} \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (S_t^Q \nabla f_{Q,w}^\varepsilon)(x)|^2 \frac{dxdt}{t}$$

where C depends only on ε, n, λ and Λ. The suprema are taken over all cubes Q.

Lemma 5.5. *For C depending only on n, λ, Λ and $\varepsilon > 0$, we have*

$$\int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (S_t^Q \nabla f_{Q,w}^\varepsilon)(x)|^2 \frac{dxdt}{t} \leq C|Q|. \tag{5.6}$$

Proof of Lemma 5.5. We follow [9, Chapter 3]. Pick a smooth cut-off function $\mathcal{X} = \mathcal{X}_Q$ localized on $4Q$ and equal to 1 on $2Q$ with $\|\mathcal{X}\|_\infty + \ell(Q)\|\nabla\mathcal{X}\|_\infty \leq c = c(n)$. By Lemma 4.3 and (4.2), the left hand side of (5.6) is bounded by

$$C \int_{\mathbb{R}^n} |\nabla(\mathcal{X} f)|^2 + 2 \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (P_t^2 \nabla(\mathcal{X} f))(x)|^2 \frac{dxdt}{t}$$
$$\leq C \int_{\mathbb{R}^n} |\nabla(\mathcal{X} f)|^2 + 4 \int_Q \int_0^{\ell(Q)} |(\theta_t \nabla(\mathcal{X} f))(x)|^2 \frac{dxdt}{t}.$$

Our task is, therefore, to control by $C|Q|$ the last expression, where to simplify the exposition, we have set $f = f_{Q,w}^\varepsilon$.

First, it follows easily from (5.2) and (5.3) that $\int_{\mathbb{R}^n} |\nabla(\mathcal{X} f)|^2 \leq C|Q|$ with C independent of Q and w (it may depend on ε which we allow).

Next, we write

$$\theta_t \nabla(\mathcal{X} f) = (1 + t^2 L)^{-1} t(\mathcal{X} L f - \operatorname{div}(A f \nabla \mathcal{X}) - A \nabla f \cdot \nabla \mathcal{X}),$$

and treat each term in the right hand side by separate arguments.

To handle the first term, observe that $Lf = \frac{f - \Phi_Q \cdot w^*}{\varepsilon^2 \ell(Q)^2}$, so that $\int_{\mathbb{R}^n} |\mathcal{X} L f|^2 \leq C|Q|(\varepsilon \ell(Q))^{-2}$ from (5.2) with C independent of Q and w. Using the $L^2(\mathbb{R}^n)$ boundedness of $(1 + t^2 L)^{-1}$, we obtain

$$\int_Q \int_0^{\ell(Q)} |(1 + t^2 L)^{-1} t(\mathcal{X} L f))(x)|^2 \frac{dxdt}{t} \leq \int_0^{\ell(Q)} C|Q| \frac{t^2}{(\varepsilon \ell(Q))^2} \frac{dt}{t} \leq \frac{C|Q|}{\varepsilon^2}.$$

To handle the second term, use the off-diagonal estimates for the operator $(1+t^2L)^{-1}t\mathrm{div}$ with sets $F=Q$ and $E=\mathrm{supp}(f\nabla\mathcal{X})\subset 4Q\setminus 2Q$ to obtain for $\int_Q\int_0^{\ell(Q)}|((1+t^2L)^{-1}t\mathrm{div}\,(Af\nabla\mathcal{X})(x)|^2\,\frac{dxdt}{t}$ a bound

$$C\int_0^{\ell(Q)} e^{-\frac{\ell(Q)}{ct}}\,\frac{dt}{t}\int_{4Q\setminus 2Q}|Af\nabla\mathcal{X}|^2\leq C|Q|,$$

where the last integral is treated using (5.2) and $\|\nabla\mathcal{X}\|_\infty\leq C\ell(Q)^{-1}$, and C depends only on n, λ and Λ.

To handle the last term, use the L^2-boundedness of $(1+t^2L)^{-1}$ to obtain for $\int_Q\int_0^{\ell(Q)}|(1+t^2L)^{-1}t(A\nabla f\cdot\nabla\mathcal{X})(x)|^2\,\frac{dxdt}{t}$ a bound

$$\int_0^{\ell(Q)} t^2\,\frac{dt}{t}\int_{4Q\setminus 2Q}|A\nabla f\cdot\nabla\mathcal{X}|^2\leq C|Q|,$$

where again the last integral is treated using (5.3) and the bound on $\nabla\mathcal{X}$, and C depends only on n, λ and Λ. This proves Lemma 5.5. □

Proof of Lemma 5.4. The main ingredient is the following result whose proof is delayed for a moment. We note that, in retrospect, the proof of this lemma is similar in spirit to a previous argument of M. Christ [10].

Proposition 5.7. *There exists a small $\varepsilon>0$ depending on n, λ and Λ, and $\eta=\eta(\varepsilon)>0$ such that for each unit vector w in $\mathbb{C}^n$ and cube Q, one can find a collection $\mathcal{S}'_w=\{Q'\}$ of non-overlapping dyadic sub-cubes of Q with the following properties*

(i) *The union of the cubes in $\mathcal{S}'_w$ has measure not exceeding $(1-\eta)|Q|$*

(ii) *If $Q''\in\mathcal{S}''_w$, the collection of all dyadic sub-cubes of Q not contained in any $Q'\in\mathcal{S}'_w$, then*

$$\frac{1}{|Q''|}\int_{Q''}\mathrm{Re}(\nabla f^\varepsilon_{Q,w}(y)\cdot w)\,dy\geq\frac{3}{4}\tag{5.8}$$

and

$$\frac{1}{|Q''|}\int_{Q''}|\nabla f^\varepsilon_{Q,w}(y)|^2\,dy\leq(4\varepsilon)^{-2}.\tag{5.9}$$

A second ingredient of a purely geometrical nature is needed.

Lemma 5.10. *Let w be a unit vector in a Hilbert space H, u, v be vectors in H and $0 < \varepsilon \leq 1$ be such that*

(i) $|u - (u \cdot w^*)w| \leq \varepsilon \; |u \cdot w^*|$,

(ii) $\mathrm{Re}(v \cdot w) \geq \frac{3}{4}$,

(iii) $|v| \leq (4\varepsilon)^{-1}$.

Then $|u| \leq 4|u \cdot v|$.

Proof. First, we deduce from (ii) that $\frac{3}{4}|u \cdot w^*| \leq |(u \cdot w^*)(v \cdot w)|$. Moreover, (i) and the triangle inequality imply that $|u| \leq (1+\varepsilon)|u \cdot w^*| \leq 2|u \cdot w^*|$. Also, by (i) and (iii), we have that $|(u - (u \cdot w^*)w) \cdot v| \leq \frac{1}{4}|u \cdot w^*|$. Hence, again by the triangle inequality, we obtain that

$$|u \cdot v| \geq |(u \cdot w^*)(v \cdot w)| - |(u - (u \cdot w^*)w) \cdot v| \geq (\frac{3}{4} - \frac{1}{4})|u \cdot w^*| \geq \frac{1}{4}|u|.$$

□

Let us continue the proof of Lemma 5.4 admitting Proposition 5.7. 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

$$|u - (u \cdot w^*)w| \leq \varepsilon \; |u \cdot w^*|. \tag{5.11}$$

It suffices to argue for each w fixed and to obtain a Carleson measure estimate for $\gamma_{t,w}(x) \equiv \mathbf{1}_{\mathcal{C}_w}(\gamma_t(x))\gamma_t(x)$, where $\mathbf{1}_{\mathcal{C}_w}$ denotes the indicator function of $\mathcal{C}_w$. Therefore, define

$$A \equiv A_w \equiv \sup_Q \frac{1}{|Q|} \int_Q \int_0^{\ell(Q)} |\gamma_{t,w}(x)|^2 \, \frac{dxdt}{t} \tag{5.12}$$

where the supremum is taken over all cubes Q. By truncating $\gamma_{t,w}(x)$ for t small and t large we may consider that this quantity is qualitatively finite. Once an *a priori* bound independent of the truncation is obtained, we can pass to the limit by monotone convergence. In order not to introduce further notation we ignore this easy step and assume that $A < +\infty$.

Now, fix a cube Q and let $Q'' \in \mathcal{S}''_w$ as defined in Proposition 5.7. Set

$$v = \frac{1}{|Q''|} \int_{Q''} \nabla f^{\varepsilon}_{Q,w}(y)\, dy \in \mathbb{C}^n.$$

Clearly, (5.8) and (5.9) in Proposition 5.7 yield (ii) and (iii) in Lemma 5.10. If $x \in Q''$ and $\frac{1}{2}\ell(Q'') < t \le \ell(Q'')$, then $v = (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)$, hence

$$|\gamma_{t,w}(x)| \le 4 \left| \gamma_{t,w}(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x) \right| \le 4 |\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)| \tag{5.13}$$

from Lemma 5.10 with $u = \gamma_{t,w}(x)$ and the definition of $\gamma_{t,w}(x)$.

The next observation is that the Carleson box $Q \times (0, \ell(Q)]$ can be partitioned into the Carleson boxes $Q' \times (0, \ell(Q')]$ for Q' describing $\mathcal{S}'_w$ and the Whitney rectangles $Q'' \times (\frac{1}{2}\ell(Q''), \ell(Q'')]$ for Q'' describing $\mathcal{S}''_w$. Hence,

$$\begin{aligned} \int_Q \int_0^{\ell(Q)} |\gamma_{t,w}(x)|^2 \frac{dxdt}{t} = \sum_{Q' \in \mathcal{S}'_w} \int_{Q'} \int_0^{\ell(Q')} |\gamma_{t,w}(x)|^2 \frac{dxdt}{t} \\ + \sum_{Q'' \in \mathcal{S}''_w} \int_{Q''} \int_{\frac{1}{2}\ell(Q'')}^{\ell(Q'')} |\gamma_{t,w}(x)|^2 \frac{dxdt}{t}. \end{aligned}$$

The first term in the right hand side is controlled by

$$\sum_{Q' \in \mathcal{S}'_w} A|Q'| \le A(1-\eta)|Q|.$$

Using (5.13), the second term is dominated by

$$\begin{aligned} 16 \sum_{Q'' \in \mathcal{S}''_w} \int_{Q''} \int_{\frac{1}{2}\ell(Q'')}^{\ell(Q'')} |\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)|^2 \frac{dxdt}{t} \\ \le 16 \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)|^2 \frac{dxdt}{t}. \end{aligned}$$

Altogether, we have obtained that $\int_Q \int_0^{\ell(Q)} |\gamma_{t,w}(x)|^2 \frac{dxdt}{t}$ is bounded by

$$A(1-\eta)|Q| + 16 \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)|^2 \frac{dxdt}{t}$$

so that dividing out by $|Q|$, taking the supremum over cubes and using the definition and the finiteness of A yield the bound

$$A \le 16\,\eta^{-1} \sup_Q \frac{1}{|Q|} \int\!\!\int_Q^{\ell(Q)}_0 |\gamma_t(x) \cdot (S_t^Q \nabla f^\varepsilon_{Q,w})(x)|^2 \, \frac{dxdt}{t}.$$

□

Proof of Proposition 5.7. We begin with a key estimate. We have

$$\left| \int_Q 1 - (\nabla f^\varepsilon_{Q,w}(x) \cdot w)\, dx \right| \le C\varepsilon^{1/2}|Q|, \tag{5.14}$$

where C depends on n, λ and Λ, but not on ε, Q and w. Indeed, we observe that $(\nabla(\Phi_Q \cdot w^*)(x) \cdot w) = |w|^2 = 1$, so that

$$1 - (\nabla f^\varepsilon_{Q,w}(x) \cdot w) = (\nabla g(x) \cdot w),$$

where $g(x) = \Phi_Q(x) \cdot w^* - f^\varepsilon_{Q,w}(x)$. Hence, (5.14) follows immediately from (5.2) and (5.3) and the application to g of the next lemma, the proof of which will be postponed to the end of this section.

Lemma 5.15. *There exists $C = C(n)$ such that for all $h \in H^1(Q)$,*

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

Continuing the proof of Proposition 5.7, we deduce from (5.14) that

$$\frac{1}{|Q|} \int_Q \mathrm{Re}(\nabla f^\varepsilon_{Q,w}(x) \cdot w)\, dx \ge \frac{7}{8}$$

provided ε is small enough. We also observe as a consequence of (5.3) that

$$\frac{1}{|Q|} \int_Q |\nabla f^\varepsilon_{Q,w}(x)|^2\, dx \le C_3,$$

with C_3 independent of ε. Now, we perform a stopping-time decomposition to select a collection $\mathcal{S}'_w$ of dyadic sub-cubes of Q which are maximal with the property that one of

$$\frac{1}{|Q'|} \int_{Q'} \mathrm{Re}(\nabla f^\varepsilon_{Q,w}(x) \cdot w)\, dx \le \frac{3}{4} \tag{5.16}$$

or

$$\frac{1}{|Q'|}\int_{Q'}|\nabla f^{\varepsilon}_{Q,w}(x)|^2\,dx \geq (4\varepsilon)^{-2} \tag{5.17}$$

holds (that is, subdivide dyadically Q and stop the first time that one of the inequalities hold). By construction, we obtain (ii) in the statement of Proposition 5.7.

It remains to establish (i). To this end, let $B = \cup_{Q'\in\mathcal{S}'_w} Q'$. We have to show that $|B| \leq (1-\eta)|Q|$. Let B_1 (resp. B_2) consist of the union of those cubes in $\mathcal{S}'_w$ for which (5.16) (resp. (5.17)) holds. We have $|B| \leq |B_1| + |B_2|$.

The fact that the cubes in $\mathcal{S}'_w$ do not overlap yields

$$|B_2| \leq (4\varepsilon)^2 \int_Q |\nabla f^{\varepsilon}_{Q,w}(x)|^2\,dx \leq (4\varepsilon)^2 C_3 |Q|.$$

Setting $b(x) = 1 - \mathrm{Re}(\nabla f^{\varepsilon}_{Q,w}(x)\cdot w)$, we also have

$$|B_1| \leq 4\sum \int_{Q'} b(x)\,dx = 4\int_Q b(x)\,dx - 4\int_{Q\setminus B_1} b(x)\,dx$$

where the sum was taken over the cubes Q' that compose B_1. The first term in the right hand side is bounded above by $C\varepsilon^{1/2}|Q|$ by (5.14). The second term is controlled in absolute value by

$$4|Q\setminus B_1| + 4|Q\setminus B_1|^{1/2}(C_3|Q|)^{1/2} \leq 4|Q\setminus B_1| + 4C_3\varepsilon^{1/2}|Q| + \varepsilon^{-1/2}|Q\setminus B_1|.$$

Since $|Q\setminus B_1| = |Q| - |B_1|$, we obtain

$$(5+\varepsilon^{-1/2})|B_1| \leq (4 + C\varepsilon^{1/2} + \varepsilon^{-1/2})|Q|$$

which gives us $|B_1| \leq (1-\varepsilon^{1/2} + o(\varepsilon^{1/2}))|Q|$ if ε is small enough. Hence $|B| \leq (1-\eta(\varepsilon))|Q|$ with $\eta(\varepsilon) \sim \varepsilon^{1/2}$ for small ε. We have proved Proposition 5.7 modulo the truth of Lemma 5.15 which we show now. □

Proof of Lemma 5.15. For simplicity, assume that Q is the unit cube $[-1,1]^n$, the general case following by homogeneity. Set $M = \left(\int_Q |h|^2\right)^{1/2}$ and $M' = \left(\int_Q |\nabla h|^2\right)^{1/2}$. If $M \geq M'$, there is nothing to prove, so we assume $M < M'$. Take $t \in (0,1)$ and $\varphi \in C_0^\infty(Q)$ with $\varphi(x) = 1$ when $\mathrm{dist}(x,\partial Q) \geq t$ (here take the distance in the sup norm in $\mathbb{R}^n$) and $0 \le \varphi \le 1$, $\|\nabla\varphi\|_\infty \le C/t$, $C = C(n)$. Then

$$\int_Q \nabla h = \int_Q (1-\varphi)\nabla h - \int_Q h\nabla\varphi$$

and the Cauchy-Schwarz inequality gives us

$$\left|\int_Q \nabla h\right| \le C(M' t^{1/2} + M t^{-1/2}).$$

It remains to choose $t = M/M'$ to conclude the proof. □

The proof of Theorem 1.4 is complete.

6 Miscellani

As far as the Kato conjecture is concerned, lower order terms do not affect the domain of square roots by a result in [8]. See also [9, Chapter 0, Proposition 11] for a different proof. This gives us the following result.

Theorem 6.1. *Consider complex bounded measurable coefficients* $a_{\alpha\beta}$ *on* $\mathbb{R}^n$ *such that the form* $Q(f,g) = \sum_{|\alpha|,|\beta|\le 1} \int_{\mathbb{R}^n} a_{\alpha\beta}(x)\partial^\beta f(x)\partial^\alpha \bar{g}(x)\,dx$ *satisfies* $|Q(f,g)| \le \widetilde{\Lambda}\|f\|_{H^1(\mathbb{R}^n)}\|g\|_{H^1(\mathbb{R}^n)}$ *and* $\mathrm{Re}\,Q(f,f) \ge \tilde{\lambda}\|f\|^2_{H^1(\mathbb{R}^n)}$. *Then the square root of the associated maximal-accretive operator* L *has domain equal to* $H^1(\mathbb{R}^n)$ *and* $\|\sqrt{L}f\|_2 \sim \|f\|_{H^1(\mathbb{R}^n)}$ *for all* $f \in H^1(\mathbb{R}^n)$, *with constants depending only on* n, $\tilde{\lambda}$, *and* $\widetilde{\Lambda}$.

We turn to L^p estimates for homogeneous elliptic operators $L = -\mathrm{div}\,(A\nabla)$ on $\mathbb{R}^n$ with ellipticity constants λ and Λ.

Proposition 6.2. *For any such* L *with* $(1+t^2L)^{-1}$, $t > 0$, *uniformly bounded on* $L^\rho(\mathbb{R}^n)$ *for some* $\rho \in [1, \frac{n}{n-1})$, *we have for* $f \in C_0^\infty(\mathbb{R}^n)$,

$$\|L^{1/2}f\|_{\mathcal{H}^1} \le c_1\|\nabla f\|_{\mathcal{H}^1}$$

hence

$$\|L^{1/2}f\|_p \le c_p\|\nabla f\|_p,$$

for all $p \in (1,2]$. *The constant* c_p *depends only on* n, λ, Λ, p *and the uniform bound above.*

Here, $\mathcal{H}^1$ denotes the classical Hardy space. This inequality was obtained in [9, Chapter 3] under the Gaussian upper bound hypothesis for the kernel of e^{-tL} and (K) now valid. The converse $\mathcal{H}^1$-inequality was proved, assuming (K) for L^*, the Gaussian upper bound and Hölder regularity of the heat kernel in [9]. The converse L^p-inequality, $1 < p < 2$, was obtained by Duong and M^{c}Intosh [17] without regularity through a weak (1,1) estimate. Putting together these results and [9, Chapter 1] yields

Theorem 6.3. *Let L be as above. If the heat kernel $W_t(x,y)$ of e^{-tL} satisfies the pointwise upper bound*

$$|W_{t^2}(x,y)| \leq Ct^{-n}\, e^{-\left(\frac{|x-y|}{ct}\right)^2}, \tag{6.4}$$

for almost every $(x,y) \in \mathbb{R}^{2n}$ and all $t > 0$, then there exists an $\varepsilon = \varepsilon(n,\lambda,\Lambda) > 0$ such that for $f \in C_0^\infty(\mathbb{R}^n)$, if $1 < p < 2+\varepsilon$,

$$\|L^{1/2}f\|_p \sim \|\nabla f\|_p$$

and if $p \geq 2+\varepsilon$,

$$\|L^{1/2}f\|_p \leq C\|\nabla f\|_p.$$

In particular, such estimates hold for real operators and their complex perturbations, for any complex operator whose coefficients have small BMO-norm depending on dimension and ellipticity.

We mention that the existence of ε and the sharpness of the range of p's is explained in [9], as well as the density argument to allow more general f.

Proof of Proposition 6.2. We begin with the same resolution of $\sqrt{L}$ as before

$$\sqrt{L}f = a\int_0^{+\infty} t^2L(1+t^2L)^{-3}tLf\,\frac{dt}{t},$$

which is now valid for any $f \in H^1(\mathbb{R}^n)$ and we further decompose it as

$$\begin{aligned}
\sqrt{L}f &= a\int_0^\infty \widetilde{\theta}_t P_t^2(\nabla f)\,\frac{dt}{t} + a\int_0^\infty \widetilde{\theta}_t(I-P_t^2)(\nabla f)\,\frac{dt}{t}.\\
&= a\int_0^\infty \widetilde{\theta}_t P_t^2(\nabla f)\,\frac{dt}{t} + a\int_0^\infty t^2L(1+t^2L)^{-3}t^2LQ_t\,\frac{dt}{t}(\sqrt{-\Delta}f)\\
&\equiv T_1(\nabla f) + T_2(\sqrt{-\Delta}f)
\end{aligned}$$

with the following notation: $\widetilde{\theta}_t = -t^2L(1+t^2L)^{-3}t\text{div}\, A$, P_t is as in Section 3, and $Q_t = (t\sqrt{-\Delta}\,)^{-1}(I - P_t^2)$.

Using the boundedness of the Riesz transforms on $\mathcal{H}^1$ we have $\|\nabla f\|_{\mathcal{H}^1} \sim \|\sqrt{-\Delta} f\|_{\mathcal{H}^1}$ so that it suffices to establish the $\mathcal{H}^1$ boundedness of T_1 and T_2.

First, observe that T_1 and T_2 are L^2-bounded. For T_2, this follows from duality, the Cauchy-Schwarz inequality and the basic Littlewood-Paley estimates for Q_t by the Fourier transform and for $t^2L^*(1+t^2L^*)^{-3}t^2L^*$ by the same argument as in Section 3. For T_1, we use the same method by splitting $\widetilde{\theta}_t P_t^2$ as $t^2L(1+t^2L)^{-2}\theta_t P_t^2$ and using the Littlewood-Paley estimate for the adjoint of $t^2L(1+t^2L)^{-2}$ and for $\theta_t P_t^2$, the latter being a byproduct of Lemma 4.4 and Lemma 4.5.

By [9, Chapter 4, Lemma 11], duality and $T_1^*(1) = 0 = T_2^*(1)$, the $\mathcal{H}^1$-boundedness of T_1 and T_2 rely on an improved version of Hörmander's inequality for their kernels, namely

$$\left(\int_{r\le|x-y|\le 2r} |K(x,y+h) - K(x,y)|^p\, dx\right)^{1/p} \le \frac{c}{r^{n(1-1/p)}}\left(\frac{|h|}{r}\right)^{\mu} \tag{6.5}$$

for some $p > 1$, $\mu > 0$ where $y, h \in \mathbb{R}^n$ and $4|h| \le r$.

Let us prove (6.5) with $p = 2$ for $K_1(x,y)$ the kernel of T_1. First, observe that

$$\widetilde{\theta}_t = t^2L(1+t^2L)^{-2}\theta_t = (1+t^2L)^{-1}\theta_t - (1+t^2L)^{-2}\theta_t.$$

We know that $(1+t^2L)^{-1}$ and θ_t satisfy separately the off-diagonal estimates of Lemma 2.1 valid for all closed sets E, F and $t > 0$. It is easy to show that such estimates are preserved by the operator product (with different constant C and c not depending on the sets E, F and $t > 0$). Hence, $\widetilde{\theta}_t$ satisfies the off-diagonal estimates. If $U_t(x,y)$ is the kernel of $\widetilde{\theta}_t P_t^2$, then we have

$$U_t(x,y+h) - U_t(x,y) = \widetilde{\theta}_t\left(\frac{1}{t^n}\tilde{p}\left(\frac{\cdot - (y+h)}{t}\right) - \frac{1}{t^n}\tilde{p}\left(\frac{\cdot - y}{t}\right)\right)(x)$$

where $\tilde{p} = p * p$, and it follows from the off-diagonal estimates, $\operatorname{supp}\tilde{p} \subset B(0,2)$ and the regularity of $\tilde{p}$ that

$$\int_{r\le|x-y|\le 2r} |U_t(x,y+h) - U_t(x,y)|^2\, dx \le Ce^{-\frac{(r-2(t+|h|))_+}{ct}}t^{-2-n}|h|^2.$$

Hence, (6.5) with $p = 2$ follows immediately from the integral Minkowski inequality.

Next, we prove (6.5) with some $p \in (\rho, n/(n-1))$ for $K_2(x,y)$. Remark that from the hypothesis the operator $t^2L(1+t^2L)^{-3}t^2L$ is uniformly bounded on $L^\rho(\mathbb{R}^n)$ while it satisfies the L^2 off diagonal estimates. By interpolation, it satisfies the L^p off diagonal estimates for any $p \in (\rho, n/(n-1))$ (in which L^p norms replace L^2 norm). Now Q_t is the convolution operator with $t^{-n}\psi(x/t)$ where for any $\mu \in (0,1)$ and some $C \geq 0$,

$$|\psi(x)| \leq \frac{C}{|x|^{n-1}(1+|x|)^2}$$

and

$$|\psi(x+h) - \psi(x)| \leq \frac{C|h|^\mu}{|x|^{n-1+\mu}(1+|x|)^2}, \quad |h| \leq |x|/2.$$

Using this together with a further chopping of ψ by a smooth partition of unity associated to a covering by ball of radius 1, it can be shown that the kernel, $V_t(x,y)$, of $t^2L(1+t^2L)^{-3}t^2LQ_t$ satisfies

$$\int_{r \leq |x-y| \leq 2r} |V_t(x,y+h) - V_t(x,y)|^p \, dx \leq C \inf(1, e^{-\frac{r}{ct}}) t^{-n(p-1)-p\eta} |h|^{p\eta}.$$

for some $\eta > 0$ and the desired inequality for $K_2(x,y)$ follows readily.

References

[1] G. Alexopoulos. La conjecture de Kato pour les opérateurs différentiels elliptiques à coefficients périodiques. *C. R. Acad. Sci. Paris*, 312, Série 1:263–266, 1991.

[2] D. Aronson. Bounds for fundamental solutions of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.

[3] P. Auscher. Regularity theorems and heat kernels for elliptic operators, *J. London Math. Soc.* **54** (1996), pp. 284-296.

[4] P. Auscher, T. Coulhon, and Ph. Tchamitchian. Absence de principe du maximum pour certaines équations paraboliques complexes. *Coll. Math.*, 171:87–95, 1996.

[5] P. Auscher, S. Hofmann, J. Lewis and Ph. Tchamitchian. Extrapolation of Carleson measures and the analyticity of Kato's square root operator. *Acta Math.*, to appear.

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

[7] P. Auscher, A. McIntosh, and Ph. Tchamitchian. Heat kernel of complex elliptic operators and applications. *J. Funct. Anal.*, 152:22–73, 1998.

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

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

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

[11] R. Coifman, D. Deng, and Y. Meyer, *Domaine de la racine carrée de certains operateurs differentiels accrétifs*, Annales de l'Institut Fourier **33** (1983), 123-134.

[12] 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.

[13] 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.

[14] G. David and J.-L. Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. Math.*, 120:371–398, 1984.

[15] G. David, J.-L. Journé, and S. Semmes. Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoamericana*, 1:1–56, 1985.

[16] E.B. Davies. Heat kernel bounds, conservation of probability and the Feller property. *J. Anal. Math.*, 58:99–119, 1992.

[17] X.T. Duong and A. McIntosh. The L^p boundedness of Riesz transforms associated with divergence forms operators, Workshop on Analysis and Applications, Brisbane, 1997, Volume 37 of *Proceedings of the Centre for Mathematical Analysis*, ANU, Canberra, pp 15–25, 1999.

[18] E. Fabes, D. Jerison and C. Kenig, *Multilinear square functions and partial differential equations*, Amer. J. of Math. **107** (1985), 1325-1367.

[19] E. Fabes, D. Jerison and C. Kenig, *Necessary and sufficient conditions for absolute continuity of elliptic-harmonic measure*, Annals of Math. **119** (1984), 121-141.

[20] M.P. Gaffney. The conservation property for the heat equation on Riemannian manifold. *Comm. Pure & Appli. Math.*, 12:1–11, 1959.

[21] S. Hofmann, M. Lacey and A. McIntosh. The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds. Preprint, 2000.

[22] J.-L. Journé, *Remarks on the square root problem*, Pub. Math. **35** (1991), 299-321.

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

[24] C. Kenig and Y. Meyer, *The Cauchy integral on Lipschitz curves and the square root of second order accretive operators are the same*, in Recent Progress in Fourier Analysis, I. Peral, ed., North Holland Math Studies **111** (1985), 123-145.

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

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

[27] A. McIntosh, *On representing closed accretive sesquilinear forms as* $(A^{1/2}u, A^{*1/2}v)$, Collège de France Seminar, Volume III (H. Brezis and J.-L. Lions, eds.), Research Notes in Mathematics, no. 70, Pitman, 1982, pp. 252-267.

[28] A.M^{c}Intosh, *Square roots of operators and applications to hyperbolic PDE*, in the proceedings of the Miniconference on Operator Theory and PDE, CMA, The Australian National University, Canberra, 1983.

[29] A. M^{c}Intosh. Square roots of elliptic operators. *J. Funct. Anal.*, 61:307–327, 1985.

[30] A. M^{c}Intosh and Y. Meyer. Algèbres d'opérateurs définis par des intégrales singulières. *C. R. Acad. Sci. Paris*, 301, Série 1:395–397, 1985.

[31] A. M^{c}Intosh 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.

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