REMARKS ON KATO'S SQUARE-ROOT PROBLEM

JEAN-LIN JOURNÉ

0. Introduction

Let T be a maximal accretive operator on a Hilbert space $\mathcal{H}$, with domain $\mathcal{V}$. There is a well-defined square-root of T, $T^{1/2}$ which satisfies the equation

$$T^{1/2} u = \frac{1}{\pi} \int_0^\infty \lambda^{-1/2} (T + \lambda)^{-1} T u \, d\lambda,$$

for all u in $\mathcal{V}$[1]. When T is sectorial, Kato considers the prehilbertian structure on $\mathcal{V}$ defined by

$$< u, v >_{\mathcal{V}} \; = \; < Tu, v > + < u, Tv > + < u, v >,$$

where $< \, . \, , . \, >$ denotes the scalar product on $\mathcal{H}$. Kato conjectured that the domain of $T^{1/2}$ was the completion of $\mathcal{V}$ in $\mathcal{H}$ for the prehilbertian structure defined by $< \, . \, , . \, >_{\mathcal{V}}$. This conjecture was disproved by A. McIntosh [2]. He observed that the failure of this conjecture was connected with the failure of the inequality

$$\left\| \, |A|B - B|A| \, \right\| \; \leqslant \; C \; \| \, AB - BA \, \|,$$

for general self-adjoint operators, $|A|$ being $(A^2)^{1/2}$. Since the boundedness of the first commutator of Calderón [3] is a (true) special case of the above (false) general inequality, McIntosh suggested that Kato's conjecture might be true when T is a differential operator on $\mathbf{R}^n$ of the form $\operatorname{div} A \nabla$, where A is a matrix-valued function such that, for some $\delta > 0$ and for all x in $\mathbf{R}^n$ and ξ in $\mathbf{C}^n$,

$$\operatorname{Re} < \xi, A(x)\xi > \; \geqslant \; \delta |\xi|^2.$$

It is this special case which is now known as Kato's conjecture or, more precisely, Kato's square-root problem.

The connection suspected by McIntosh turned out to be quite significant and led simultaneously to the solution of Kato's conjecture in dimension 1 and to the proof of the L^2−boundedness of the Cauchy-kernel on Lipschitz graphs [4], conjectured by Calderón and proved by himself in the case of small Lipschitz constants [5].

A simple rescaling on A shows that it is enough to consider the case where $\| A - I \|_\infty < 1$. In dimension larger than 1, Kato's conjecture has been solved when A is a small perturbation of the identity, that is if $\| A - I \|_\infty < \varepsilon_n$, where ε_n depends only on the dimension n and decays exponentially with n [6],[7]. A natural way to attack Kato's conjecture is therefore to find best possible lower bounds for ε_n. Here we obtain explicit lower bounds for ε_n, decaying like $n^{-1/2}$. In particular, if $n \leqslant 5$, one can take ε_n to be $1/3$.

Also we answer a question of A. McIntosh concerning the solution of Kato's conjecture in dimension 1. Rescaling A so that, for all x in $\mathbf{R}^n$, Re $A(x) \geqslant I$, one may assume that $\|A^{-1} - I\|_\infty < 1$. In dimension 1, the solution is obtained by expanding an operator depending on A in a series of operators depending multilinearly on the new variable $B = I - A^{-1}$, then by showing that these operators are bounded on L^2 and that the norms can be summed when $\| B \|_\infty < 1$. In dimension larger than 1, a similar procedure does not seem to bring the same kind of simplification and hence, it was suggested by A. McIntosh, as a first step to understand the higher-dimensional case, to reprove Kato's conjecture in dimension 1 without using the variable B. This is what we do at the end of this paper.

The content of the sections of this paper are as follows. In the first we introduce some notations and recall some formulas. In the second we recall some basic facts about Carleson-measures. In the third, we state a theorem about multi-linear operators, which readily implies the improvement on the order of magnitude of ε_n, and whose proof we outline. In the fourth and fifth we prove some technical estimates on kernels of operators and in the sixth we conclude the proof of the aforementioned theorem. In the seventh we show how to use various properties of the operators arising in Kato's problem to improve the lower bound for ε_n given by our theorem. In the eighth we give a direct proof of Kato's conjecture in dimension 1.

1. Notations and Formulas

The following is partly borrowed from [7].

If we make the change of variable $\lambda = t^{-2}$, we see that the formula giving $T^{1/2}$ can be rewritten as

$$T^{1/2} = \frac{2}{\pi} \int_0^\infty \frac{T}{t^2 T + I}\, dt.$$

Let $D_j = -i\frac{\partial}{\partial x_j}$, $D = \nabla$ and $D^* = (D_1, ..., D_n)$. We denote by $\{A\}$ the operator of pointwise multiplication by the matrix A. Then T can be written as $D^*\{A\}D$. Let $U = A - I$. Then $T = D^*\{U\}D + \Delta$. Let $P_t = (t^2\Delta + I)^{-1}$. Then $(t^2 T + I)^{-1}$ can be written as

$$\sum_{j \geqslant 0} (-1)^j P_t \Big[t^2 D^* \{U\} D P_t \Big]^j,$$

as long as $\| U \|_\infty < 1$. And in this case,

$$\frac{T}{t^2 T + I} = \sum_{j \geqslant 0} (-1)^j P_t \Big[t^2 D \{U\} D^* P_t \Big]^j D^* \{A\} D$$

$$= \sum_{j \geqslant 0} (-1)^j D^* P_t \Big[t^2 \{U\} DD^* P_t \Big]^j \{A\} D.$$

Let $Q_t = t D^* P_t$ and $\tilde{R}$ be the matrix of operators whose entries are $\frac{D_i D_l}{\Delta}$. Observe that $t^2 DD^* P_t = \tilde{R}(I - P_t)$. Therefore the Sobolev space $W^{1,2}$ will be in the domain of $T^{1/2}$ if the series of operators

(1)
$$\sum_{j \geqslant 0} (-1)^j \int_0^\infty Q_t^* [\{U\} \tilde{R}(I - P_t)]^j \frac{dt}{t}$$

is convergent in L^2-operator norm. That each summand is bounded with a norm dominated by $C_n{}^j \| U \|_\infty^j$ is known [6],[7].

Observe that if A is bounded and strictly accretive, $\| \lambda A - I \|_\infty < 1$, for λ small enough. Hence one can always reduce to the case where $\| U \|_\infty < 1$. Therefore Kato's conjecture will be solved if one can show that, for all $\varepsilon > 0$, each summand is bounded with a norm at most $C_\varepsilon (1 + \varepsilon)^j$ when $\| U \|_\infty < 1$.

One can also do a power series expansion in the variable $V = I - A^{-1}$.

Starting from

$$\Big(t^2 D^* \{A\} D + I \Big) t D^* = t D^* \{A\} (I + t^2 DD^*) - t D^* \{A\} \Big(I - \{A\}^{-1} \Big)$$

$$= t D^* \{A\} \Big(I - \{V\} (I + t^2 DD^*)^{-1} \Big)(I + t^2 DD^*),$$

we obtain

$$\Big(t^2 D^* \{A\} D + I \Big)^{-1} t D^* \{A\} = t D^* (I + t^2 DD^*)^{-1} \Big(I - \{V\}(I + t^2 DD^*)^{-1} \Big).$$

Let $R_t = (I + t^2 DD^*)^{-1}$. Then if $\| V \|_\infty < 1$, one obtains that $W^{1,2}$ will be the domain of $T^{1/2}$ if the series of operators

(2)
$$\sum_{j \geqslant 0} \int_0^\infty t D^* R_t \Big[\{V\} R_t \Big]^j \frac{dt}{t}$$

is convergent. It is easy to check that $t D^* R_t = Q_t^*$ and that $R_t = I - \tilde{R} + \tilde{R} P_t$. Hence there is a very close resemblance between the two expansions (1) and (2). In dimension 1 however, $\tilde{R} = I$ and $R_t = P_t$, so that the expansion in

the variable V is much easier to handle than in the expansion in the variable U. By multiplying A by a large number so that $\operatorname{Re} A \geqslant I$, one can reduce to the case where $\| V \|_\infty < 1$. Therefore one can try to prove Kato's conjecture by showing that for all $\varepsilon > 0$, the summands in the series (2) are dominated by $C_\varepsilon (1 + \varepsilon)^j \| V \|_\infty^j$ in operator norm. This is how it was done originally in dimension 1 [4]. In higher dimension it is not clear that one of the series (1) or (2) has an advantage over the other. One can ask in particular if there is a simple relation between their radii of convergence. In dimension 1, it is easy to see a priori that they are the same. Indeed, a function a is such that $\| a - 1 \|_\infty \leqslant \lambda_0 < 1$, if and only if $\| a^{-1} - 1/(1 - \lambda_0^2) \|_\infty \leqslant \lambda_0/(1 - \lambda_0^2)$. Therefore if Kato's conjecture is true when $\| a - 1 \|_\infty \leqslant \lambda_0 < 1$, it is also true when $\| (1 - \lambda_0^2)a^{-1} - 1 \|_\infty \leqslant \lambda_0$, and also when $\| a^{-1} - 1 \|_\infty \leqslant \lambda_0$ after rescaling. It is a little surprising, but easy to see, that this remains true in higher dimension. It is a consequence of the following fact. If one puts on $\mathrm{GL}_n(\mathbf{C})$ the distance induced by the operator-norm on $\mathbf{C}^n$,

$$d(A, B) = \max_{\xi \in \mathbf{C}^n, \|\xi\|=1} \| U\xi - V\xi \|,$$

then inversion does not in general map balls to other balls. However it does map a ball centered around a multiple of the identity to a ball of the same nature, just as in $\mathbf{C}^*$. This is what the next lemma expresses.

Lemma 1. *Let* $0 < \lambda_0 < 1$ *and* A *in* $\mathrm{GL}_n(\mathbf{C})$. *Then*

$$\| A - I \| \leqslant \lambda_0 \implies \| (1 - \lambda_0^2)A^{-1} - I \| \leqslant \lambda_0.$$

Proof: Let w be a unit vector in $\mathbf{C}^n$. We want to show

$$\| (1 - \lambda_0^2)A^{-1}w - w \| \leqslant \lambda_0.$$

Making the change of variables $v = A^{-1}w$, it is enough to show that

$$\| (1 - \lambda_0^2)v - Av \| \leqslant \lambda_0 \| Av \|.$$

But v and Av are just two vectors x and y such that $\| y - x \| \leqslant \lambda_0 \| x \|$. Therefore we can check this inequality in $\mathbf{C}$ letting $x = 1$ and $y = 1 + z$ with $|z| \leqslant \lambda_0$. After division by $|y|(1 - \lambda_0^2)$, this is equivalent to

$$|(1 + z)^{-1} - (1 - \lambda_0^2)^{-1}| \leqslant \lambda_0(1 - \lambda_0^2)^{-1}.$$

This, in turn, follows from the fact that the inversion in $\mathbf{C}^*$ maps $\mathbf{B}(1, \lambda_0)$ to $\mathbf{B}\left((1 - \lambda_0^2)^{-1}, \lambda_0(1 - \lambda_0^2)^{-1}\right)$. This proves Lemma 1. ∎

Since the inversion is an involution on $\mathrm{GL}_n(\mathbf{C})$, it follows that a ball centered around a multiple of the identity is mapped exactly onto a similar ball. From the previous lemma we see that, if A is bounded and $\operatorname{Re} A \geqslant \delta > 0$,

$$\min_{\lambda > 0} \| \lambda A - I \|_\infty = \min_{\lambda > 0} \| \lambda A^{-1} - I \|_\infty.$$

It follows that if the series (1) converges when $\| U \|_\infty < \lambda_0 < 1$, for some $\lambda_0 > 0$, then $W^{1,2}$ is the domain of $T^{1/2}$ when $\min_{\lambda > 0} \| \lambda A - I \|_\infty < \lambda_0$, and therefore when $\min_{\lambda > 0} \| \lambda A^{-1} - I \|_\infty < \lambda_0$. So the series (2) must converge when $\| V \|_\infty < \lambda_0$. And conversely.

Having observed that the radii of convergence of (1) and (2) are the same, we would like to recover this fact directly from the study of the multi-linear terms of (1) and (2), and to show that this radius is 1. As we shall see, the method we use gives two distinct values for the radii of convergence of (1) and (2), and this shows that it is not optimal.

2. Preliminaries on Carleson Measures

Let p_t denote the operator of convolution with the Poisson-kernel. Then a Carleson-measure μ on the half-space $\mathbf{R}_+^{n+1}$ is a measure for which one has the estimate

$$(3) \qquad \int_{\mathbf{R}_+^{n+1}} |p_t f(x)|^2 d\mu(x,t) \leqslant C(\mu) \| f \|_2^2$$

A necessary and sufficient condition for a measure μ to have this property is the existence of a constant $C > 0$ such that for all cubes Q in $\mathbf{R}^n$,

$$(4) \qquad \mu\Big(Q \times [0,\delta] \Big) \leqslant C|Q|,$$

where $|Q|$ and δ denote the Lebesgue measure and the side-length of Q. Let us denote by c_μ the best constant in (4). Then, for some absolute constant C, $C(\mu) \leqslant Cc_\mu$. Also, if we replace the Poisson-approximation by some other approximation of the identity $(\tilde{p}_t)_{t>0}$ satisfying appropriate estimates, then the best constant in (3) will presumably change. However, if one looks at the size of the difference

$$\int_{\mathbf{R}_+^{n+1}} \Big| (p_t f - \tilde{p}_t f)(x) \Big|^2 d\mu(x,t)$$

we see that it depends only on a constant h_μ which we define to be the best constant in the inequality

$$(5) \qquad \mu\Big(Q \times [\delta/2, \delta] \Big) \leqslant C|Q|.$$

Of course $h_\mu \leqslant c_\mu$. The point is that in general, and in working on Kato's problem in particular, one tries to estimate c_μ for measures for which one already has a good control of h_μ.

Lemma 2. *Let p_t and $\tilde{p}_t$ be as above. Then, for some $C > 0$,*

$$\int_{\mathbf{R}_+^{n+1}} \Big| (p_t f - \tilde{p}_t f)(x) \Big|^2 d\mu(x,t) < Ch_\mu \| f \|_2^2 .$$

Proof: Observe that the operators T_k defined for $k \in \mathbf{Z}$ by

$$< g, T_k f >= \int_{\mathbf{R}^n \times [2^k, 2^{k+1}]} \left[(p_t f - \tilde{p}_t f)(x) \right] \left[(p_t g - \tilde{p}_t g)(x) \right] d\mu(x, t),$$

satisfy the assumptions of the Cotlar-Stein Lemma with a constant depending only on h_μ. This proves Lemma 2. ∎

From Lemma 2 we see that

$$(6) \qquad \left[\int_{\mathbf{R}^{n+1}_+} \left| \tilde{p}_t f(x) \right|^2 d\mu(x, t) \right]^{1/2} \leqslant \left[C_0 c_\mu^{1/2} + C_1 h_\mu^{1/2} \right] \| f \|_2,$$

where C_0 is independent of the approximation of the identity but C_1, of course, is not.

Lemma 3. *The constant C_0 can be chosen equal to 2 in all dimensions.*

Proof: By Lemma 2 we know that we can choose $\tilde{p}_t$ as we like to estimate the best C_0. By the same argument, we can replace $\tilde{p}_t f(x)$ by $S_t f(x) = m_{Q(t,x)} f(x)$, where $Q(t, x)$ is the dyadic cube containing x of size 2^k, with $2^{k-1} < t \leqslant 2^k$. It is easy to see that

$$\left[\int_{\mathbf{R}^{n+1}_+} |S_t f(x)|^2 d\mu(x, t) \right]^{1/2} \leqslant c_\mu \left\| \max_{t>0} |S_t f(x)| \right\|_2,$$

and it is a classical martingale-inequality that

$$(7) \qquad \left\| \max_{t>0} |S_t f(x)| \right\|_2 \leqslant 2\|f\|_2.$$

Together with Lemma 2 applied with S_t instead of p_t, these two inequalities readily imply Lemma 3. ∎

An obvious but useful remark is that, in the inequality

$$(8) \qquad \left[\int_{\mathbf{R}^{n+1}_+} \|\tilde{p}_t f(x)\|^2 d\mu(x, t) \right]^{1/2} \leqslant (2c_\mu^{1/2} + C_1 h_\mu^{1/2})\|f\|_2.$$

we can replace $d\mu(x, t)$ by $d\mu(x, ut)$, where $0 < u < 1$, and replace $\tilde{p}_t$ by $\tilde{p}_{\omega t}$ for some $\omega > 1$.

3. A Multilinear Estimate

Let $(A_i)_{i \in \mathbf{N}}$ be a family of matrix-valued functions satisfying $\|A_i\|_\infty \leqslant 1$. Let $(\tilde{K}_i)_{i \in \mathbf{N}}$ be a sequence of convolution operators mapping $\mathbf{C}^d$-valued functions to $\mathbf{C}^d$-valued functions. We assume that the symbols $\sigma(\tilde{K}_i)$, which are matrix-valued, are of the form $((S_{k,l}(\xi)/|\xi|^2))$ where the $S_{k,l}$'s are homogeneous polynomials of degree 2. Let $(\lambda_i)_{i \in \mathbf{N}}$ be a sequence of numbers.

For $t > 0$ and $i \in \mathbf{N}$ we define the operator $K_{i,t} = \lambda_i I + \tilde{K}_i(I - P_t)$. We assume that the symbols of the $K_{i,t}$'s and the $\tilde{K}_i$'s, which are matrix-valued, are contractions on $\mathbf{C}^d$, for all $\xi \in \mathbf{R}^d$.

Theorem 1. *For all $\epsilon > 0$ there exists $C_\epsilon > 0$ such that, for all $F \in L^2_{\mathbf{C}^d}(\mathbf{R}^d)$,*

(9)
$$\left[\int_0^\infty \|Q_t\{A_1\}K_{1,t} \dots K_{n-1,t}\{A_n\}K_{n,t}F\|_2^2 \frac{dt}{t} \right]^{1/2} \leqslant C_\epsilon(1 + 2\sqrt{d} + \epsilon)^n \|F\|_2.$$

The proof which we shall outline follows as usual from Carleson-measures estimates. The improvement over [6] and [7] comes from Lemma 3 and from the fact that for the Carleson-measures that enter into the proof, it is very easy to obtain a good control of h_μ.

If $(F_t)_{t>0}$ is a family of vector-valued or scalar-valued L^2–functions we shall define $|||F_t|||$ by

$$|||F_t||| = \left[\int_0^\infty \|F_t\|_2^2 \frac{dt}{t} \right]^{1/2}.$$

Sketch of Proof: First we reduce (9) to a similar estimate where $K_{n,t}$ is replaced by P_t. To do this we dominate

$$|||Q_t\{A_1\}K_{1,t} \dots K_{n-1,t}\{A_n\}K_{n,t}F|||$$

by

$$|||Q_t\{A_1\}K_{1,t} \dots K_{n-1,t}\{A_n\}P_t\tilde{K}_nF|||$$
$$+ \ |||Q_t\{A_1\}K_{1,t} \dots K_{n-1,t}\{A_n\}(\lambda_n I + \tilde{K}_n)F|||.$$

For all $n > 0$, X_n, Y_n and $\tilde{Y}_n$ denote the sup of these 3 quadratic expressions when $\|F\|_2 \leqslant 1$. Since $\{A_n\}(\lambda_n I + \tilde{K}_n)$ is a contraction, $\tilde{Y}_n \leqslant X_{n-1}$. Hence

(10) $$X_n \leqslant Y_n + X_{n-1}$$

So if Y_n grows at most like $(1 + 2\sqrt{d} + \epsilon)^n$ for all $\epsilon > 0$, then so does X_n.

To estimate $|||Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}\{A_n\}P_tF|||$ the classical thing to do is to decompose the expression in the norm as a sum of the type

$$\Big\{Q_t\{A_1\}K_{1,t}\dots\{A_{n-1}\}K_{n-1,t}A_n\Big\}P_tF + \text{ error term},$$

where the error is of the form L_tF with $L_t1 = 0$. The main term is estimated using a Carleson-measure estimate which itself follows from an L^2-estimate at the order $n-1$. As we shall now see, this procedure can be improved to yield better constants.

Let $(w_t)_{t>0}$ be some radial smooth approximation of the identity, such that w_1 is non-negative and supported in the unit ball. Now let $v_t = w_t * w_t$. Observe that convolution with w_t or v_t is a contraction on L^2 or L^∞. Now we dominate

$$|||Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}\{A_n\}P_tF|||$$

by the sum of the three following terms :

(11)
$$|||Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}\{A_n\}(P_t - v_{2^nt})F|||,$$

(12)
$$\Big|\Big|\Big|Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}\{A_n\}v_{2^nt}F - \Big\{Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}A_n\Big\}v_{2^nt}F\Big|\Big|\Big|,$$

and

(13)
$$\Big|\Big|\Big|\Big\{Q_t\{A_1\}K_{1,t}\dots K_{n-1,t}A_n\Big\}v_{2^nt}F\Big|\Big|\Big|.$$

The first term is less than $|||(P_t - v_{2^nt})F|||$, which is dominated by

$$|||(P_t - v_t)F||| + \sum_{k=1}^n |||(v_{2^kt} - v_{2^{k-1}t})f|||$$

and hence by $C(1 + n)\|F\|_2$.

The second term is also of the type $|||L_tF|||$ with $L_t1 = 0$ for all $t > 0$, and can be estimated directly without using induction on n. The corresponding estimate grows slower than exponentially.

To estimate (13) one has to estimate c_μ and h_μ for the Carleson-measure

$$\Big|Q_{t/2^n}\{A_1\}K_{1,t/2^n}\dots K_{n-1,t/2^n}A_n(x)\Big|^2\frac{dxdt}{t}.$$

The constant h_μ can be estimated without induction and grows slower than exponentially. To estimate c_μ we choose a cube Q of side-length δ and we want to estimate

$$\int_{x\in Q,\,t\leqslant\delta}\Big|Q_{t/2^n}\{A_1\}K_{1,t/2^n}\dots K_{n-1,t/2^n}A_n(x)\Big|^2\frac{dtdx}{t}.$$

We shall see that because of the factor 2^n, this is essentially dominated by an expression of the form

$$X_{n-1}^2 \|A_n\|_{L^2((1+\varepsilon_n)Q)}^2,$$

up to error terms, with ε_n's such that $\prod_{n=1}^{\infty}(1 + \varepsilon_n) < \infty$. Of course, here the norm of A_n has to be taken in the Hilbert-Schmidt sense and this is what introduces a factor $\sqrt{d}$. In conclusion one obtains an inequality of the form

$$Y_n \leqslant 2X_{n-1}\left((1 + \varepsilon_n)\sqrt{d}\right) + \text{error terms.}$$

Combining this with (10), one obtains

(14) $$X_n \leqslant (1 + 2\sqrt{d})X_{n-1} + \text{error terms,}$$

which implies the theorem, modulo appropriate control of the error terms.

4. Technical Preliminaries

An ingredient in most subsequent estimates is the following.

Lemma 4. *The kernels of the operators $K_{i,t}$ satisfy the inequality*

$$\|K_{i,t}(x - y)\| \leqslant \frac{C}{t^d} e^{-\|\frac{x-y}{t}\|},$$

for $\|x - y\| > t$.

This lemma follows easily from the asymptotic properties of the Fourier transform of $1/(1 + \xi^2)$, which decays exponentially as well as its derivatives. We omit the details.

Later in the proofs, we shall not need the full force of the exponential factor. A polynomial factor of sufficiently high degree, depending on the dimension, would be enough. Such decay, however, can only come from sufficient smoothness of the symbol of $K_{i,t}$ which is $\lambda_i + (t^2\|\xi\|^2/1 + t^2\|\xi\|^2) \times \sigma(\tilde{K}_i)$. Requiring that the symbol of $K_{i,t}$ be regular enough, and in a scale-invariant way, forces $\|\xi\|^2\sigma(\tilde{K}_i)$ which is homogeneous of degree 2, to be a polynomial. This justifies our assumption on $\sigma(\tilde{K}_i)$ at least in large dimension.

A consequence of Lemma 4 is the following.

Lemma 5. *Let f and g be two L^2-functions such that $d(supp(f), supp(g)) = \delta > 0$. Then if $t \leqslant \delta/n$,*

(15) $$\left| < g, K_{1,t}\{A_1\} \ldots \{A_{n-1}\}K_{n,t}f > \right| \leqslant C^n\|f\|_2\|g\|_2 \, e^{-\frac{\delta}{2nt}},$$

for some $C > 0$.

Proof: Observe that each $K_{i,t}$ can be written as the sum of two operators $M_{i,t}$ and $N_{i,t}$, where $M_{i,t}$ has a kernel supported in the strip $\{|x - y| < \delta/n\}$ and $N_{i,t}$ has, by Lemma 4, a norm dominated by

$$\frac{1}{t^d} \int_{\|z\| > \delta/n} e^{-\frac{\|z\|}{t}} dz.$$

This quantity can be dominated by

$$(16) \qquad\qquad C(\delta/nt)^{d-1} e^{-\delta/nt},$$

and hence, $M_{i,t}$ has a norm less than $1 + C(\delta/nt)^{d-1} e^{-\delta/nt}$. Also,

$$(17) \qquad\qquad < g, M_{1,t}\{A_1\} \dots \{A_{n-1}\} M_{n,t} f > \;=\; 0$$

because of the assumptions on the supports of f and g. Using the decomposition $K_{i,t} = M_{i,t} + N_{i,t}$ for all i simultaneously, and L^2-estimates, we dominate the left hand side of (15) by

$$(18) \qquad \|f\|_2 \|g\|_2 \sum_{\substack{(\alpha_i)_{1 \leqslant i \leqslant n} \in \{0,1\}^n \\ \prod_i \alpha_i = 0}} \prod_i \left(\alpha_i + C \left(\frac{\delta}{nt} \right)^{d-1} e^{-\delta/nt} \right).$$

The restriction $\prod_i \alpha_i = 0$ follows from (17) and implies that (18) is dominated by the right hand side of (15). This proves Lemma 5. ∎

Remark: In the left hand side, of (15), one can allow t to take different values $t_1 \dots t_n$. Then in the right hand side we have to set $t = \max_i t_i$. Also we can replace one $K_{i,t}$ by Q_t.

Lemma 5 can be used to estimate truncated operators. We fix a smooth function θ defined on $\mathbf{R}^d$, equal to 1 if $\|z\| < 1$ and to 0 if $\|z\| > 2$. Notice that $\theta(x - y) = \int \hat{\theta}(\xi) e^{i<\xi, x-y>} d\xi$. For every operator T bounded on L^2, we define the truncated operator $T_{(s)}$ to be

$$\int_{\xi \in \mathbf{R}^d} \hat{\theta}(s\xi) \{e^{i<x,\xi>}\} T \{e^{-i<x,\xi>}\} s^n d\xi.$$

This new operator is automatically bounded on L^2 with a norm at most $\|T\| \|\hat{\theta}\|_1$. We denote by $(T)_s$ the difference $T - T_{(s)}$.

Lemma 6. *Let T be the operator $Q_t\{A_1\}K_{1,t}\{A_2\} \dots \{A_{n-1}\}K_{n,t}$, and let $s > 3nt$. Then for some positive constants C and c,*

$$(19) \qquad\qquad \|(T)_s\| \leqslant C^n e^{-s/cnt}.$$

Proof: We choose a function ϕ defined on $\mathbf{R}$, bounded by 1, supported in $[-1,1]$ and such that $\sum_{j \in \mathbf{Z}} \phi(x-j) = 1$. Let $u = s/(6\sqrt{d})$. For all $(j_i)_{1 \leqslant i \leqslant d} \in \mathbf{Z}^d$, we define $\Phi_{j_1 \ldots j_d}$ to be the multiplication by the function $\prod_i \phi\left(\frac{x_i}{u} - j_i\right)$. This gives the following decomposition of the identity on $\mathbf{R}^d$:

$$I = \sum_{(j_1 \ldots j_d) \in \mathbf{Z}^d} \Phi_{j_1 \ldots j_d}.$$

This, in turn, gives the following decomposition of $(T)_s$:

$$(T)_s = \sum_{k_1 \ldots k_d \in \mathbf{Z}^d} \left(\sum_{j_1 \ldots j_d \in \mathbf{Z}^d} \Phi_{j_1 \ldots j_d} T \Phi_{j_1 + k_1 \ldots j_d + k_d} \right)_s.$$

The diameter of the union of the supports of $\Phi_{j_1 \ldots j_d}$ and $\Phi_{j_1 + k_1 \ldots j_d + k_d}$ is at most $(\|k\| + 2\sqrt{d})u$. Since the kernel of $(T)_s$ vanishes in the strip $\{\|x - y\| < s\}$, the terms corresponding to $k \in \mathbf{Z}^d$ such that $(\|k\| + 2\sqrt{d})u < s$, give no contribution. For the remaining k's, we notice that $\Phi_{j_1 \ldots j_d} \Phi_{l_1 \ldots l_d} = 0$ if $|l_i - j_i| > 1$ for any i. It follows by almost-orthogonality that
(20)

$$\left\| \sum_{j_1 \ldots j_d} \Phi_{j_1 \ldots j_d} (T)_s \Phi_{j_1 + k_1 \ldots j_d + k_d} \right\| \leqslant 3^d \max_{j_1 \ldots j_d} \left\| \Phi_{j_1 \ldots j_d} (T)_s \Phi_{j_1 + k_1 \ldots j_d + k_d} \right\|.$$

Since the distance of the supports of $\Phi_{j_1 \ldots j_d}$ and $\Phi_{j_1 + k_1 \ldots j_d + k_d}$ is at least $(\|k\| - 2\sqrt{d})u$, which is more than nt if $(\|k\| + 2\sqrt{d})u > s$, we can apply Lemma 5. A summation over k yields

$$\sum_{k, (\|k\| + 2\sqrt{d})u \geqslant s} C^n e^{\frac{(\|k\| - 2\sqrt{d})u}{2nt}}.$$

as a bound for $(T)_s$. By the choice of u, this sum is dominated by

$$C^n \sum_{k, \|k\| \geqslant 4\sqrt{d}} e^{s\|k\|/24\sqrt{d}nt}.$$

This implies Lemma 6. ∎

5. Estimates for the error-terms

We first want to estimate (12). To control expressions of the form $\||L_t F\||$, where $L_t 1 = 0$ for all $t > 0$, it is enough to have a bound on $\|L_t\|$ uniform in t and some decay on $\|L_t L_{t'}^*\|$ when t/t' tends to 0 or $+\infty$.

Observe that, out of the two terms entering in the definition of L_t in (12), one is a contraction and the other is of the form $\{A_t\} v_{2^n t}$. To estimate the latter we remark that $v_{2^n t} \{A_t\}^* \{A_t\} v_{2^n t}$ is bounded on L^1 and L^∞ with a norm at most $\|v_{2^n t}(|A_t|^2)\|_\infty$. Therefore $\{A_t\} v_{2^n t}$ is bounded on L^2 with a norm dominated by $\max_x \left[m_{\mathbf{B}(x, 2^n t)} |A_t|^2 \right]^{1/2}$.

Lemma 7. *The operator-norm of L_t is bounded independently of n or t.*

Proof: In view of the remarks preceding Lemma 7, we see that it is enough to show that, for some $C > 0$ and all $x \in \mathbf{R}^d$,

$$(21) \qquad \left(\int_{\|x-y\|<2^n t} \|Q_t\{A_1\}K_{1,t} \ldots K_{n,t}A_n(y)\|^2 dy \right)^{1/2} \leqslant C(2^n t)^{d/2}.$$

We fix x. To prove (21), we decompose A_n as $A_0 + \sum_{j=1}^{+\infty} B_j$, where

$$A_0(z) = A_n(z)\chi_{\{\|z-x\|<(2^n+2^{n/2})t\}}$$

and for $j > 0$,

$$B_j(z) = A_n(z)\chi_{\{(2^n+2^{j-1}+\frac{n}{2})t<\|z-x\|\leqslant(2^n+2^j+\frac{n}{2})t\}}.$$

Since $Q_t\{A_1\}K_{1,t} \ldots K_{n,t}$ is a contraction on L^2, the contribution of A_0 is less than $\|A_0\|_2$, which is dominated by $C(2^n t)^{d/2}$.

To estimate the contributions of the B_j's we use Lemma 5 and the remark following its proof. Then, if a function is in L^2 with norm 1 and is supported in $B(x, 2^n t)$,

$$| < g, Q_t\{A_1\}K_{1,t} \ldots K_{n,t}B_j > | \leqslant C^n \|B_j\|_2 \left(\frac{2^{j-1+n/2}}{n} \right)^{d-1} e^{-\frac{2^{j-1+n/2}}{n}}.$$

Since $\|B_j\|_2 \leqslant C \left(2^n + 2^{j-1+n/2} \right)^{d/2} t^{d/2}$, (21) follows by summation over j. This proves Lemma 7. ∎

Remark. It follows that the measure

$$\mu = \|Q_t\{A_1\}K_{1,t} \ldots K_{n,t}A_n(y)\|^2 dy\,dt/t$$

satisfies $h_\mu < C$.

Next we turn to the decay estimate for $\|L_t L_{t'}^*\|$, when t/t' is large or small. We can assume $t < t'$, taking the adjoint if necessary.

Using the factorization of v_t as $w_t * w_t$, we see that L_t can be written as $\tilde{L}_t w_t^2$ where

$$\tilde{L}_t = Q_t\{A_1\}K_{1,t} \ldots K_{n-1,t}\{A_n\} - \left\{ Q_t\{A_1\}K_{1,t} \ldots K_{n-1,t}A_n \right\},$$

and similarly for $L_{t'}$. From the proof of Lemma 6 we see that $\tilde{L}_{t'} w_{2^n t'}$ is bounded on L^2 with a bound independent of n. Hence

$$\|L_t L_{t'}^*\| \leqslant C \|\tilde{L}_t w_{2^n t'}\|.$$

Lemma 8. $\|\tilde{L}_t w_{2^n t'}\| \leqslant C(t/t')^{1/2}$ if $t' > 4^n t$.

Proof: First we observe that, because w_1 is supported in the unit ball,

$$(22) \qquad \left(\tilde{L}_t w_{2^n t'}\right)_{2^{n+1} t'} = \left((\tilde{L}_t)_{2^n t'} w_{2^n t'}\right)_{2^{n+1} t'}.$$

We denote by V_t the operator $\tilde{L}_t w_{2^n t'}$. By Lemma 6 $\|(\tilde{L}_t)_{2^n t'}\|$ is dominated by $e^{-2^n t'/cnt}$. Hence, by (22), so is $\|(V_t)_{2^{n+1} t'}\|$. Since the kernel of $V_t - (V_t)_{2^{n+1} t'}$ is supported in $\{\|x - y\| \leqslant 2^{n+2} t'\}$, it follows that it can be decomposed in an almost-orthogonal sum of operators whose kernels are localized in x and y on balls of diameter $2^n t'$, as in the proof of Lemma 6. We just need uniform estimates for the norms of these operators, or even for their Hilbert-Schmidt norms. Equivalently, we need estimates, uniform in $x_0 \in \mathbf{R}^d$, for

$$(23) \qquad \int_{\|x-x_0\|<2^n t', \|y-x_0\|<c2^n t'} \|V_t(x,y)\|^2 dx dy.$$

Let $t_0 = 2^{n/2}(tt')^{1/2}$. For all $j = (j_1 \ldots j_d) \in \mathbf{Z}^d$, we define x_j to be the point $t_0(j_1 \ldots j_d)$. Then, we can dominate (23) by

$$(24) \qquad \sum_{\{x_j, \|x_j - x_0\| < 2^n t'\}} \left(\int_{\|x-x_j\|<\sqrt{d}t_0} \|V_t(x,y)\|^2 dx dy\right).$$

Let us rewrite $V_t(x,y)$ as

$$(25) \qquad \int \tilde{L}_t(x,z)\Big(w_{2^n t'}(z-y) - w_{2^n t'}(x_j - y)\Big) dz.$$

We use a truncation of $\tilde{L}_t$ at scale t_0. That is, we write $\tilde{L}_t$ as $(\tilde{L}_t)_{t_0} + \tilde{L}_t - (\tilde{L}_t)_{t_0}$.

For the local part we use the L^2-boundedness of $\tilde{L}_t$, treating y in (24) as a parameter, and we are reduced to controlling

$$\sum_{\{x_j, \|x_j - x_0\|<2^n t'\}} \left(\int_{y,z,\|z-x_j\|\leqslant(\sqrt{d}+2)t_0} |w_{2^n t'}(z-y) - w_{2^n t'}(x_j - y)|^2 dz dy\right).$$

This expression is easily seen to be dominated by

$$(2^n t'/t_0)^d((\sqrt{d}+2)t_0)^d(2^n t')^d \left(\frac{(\sqrt{d}+2)t_0}{(2^n t')^{d+1}}\right)^d,$$

which is equal to $C2^n t/t'$.

For the far-away part, we estimate the contributions of $w_{2^n t'}(z - y)$ and $w_{2^n t'}(x_j - y)$ separately. By Lemma 6 the first contribution is dominated by

$$Ce^{-t_0/cnt} \sum_{x_j, \|x_j - x_0\| < 2^n t'} \left(\int_{y,z, \|y - x_0\| < C2^n t'} |w_{2^n t'}(z - y)|^2 dy dz \right),$$

which is less than

$$C(2^{n/2}(t'/t))^d e^{-\frac{2^{n/2}}{n}(\frac{t'}{t})^{1/2}}.$$

The second contribution can be rewritten as

$$(25) \qquad \sum_{\{x_j, \|x_j - x_0\| < 2^n t'\}} \left(\int_{\|x - x_j\| < \sqrt{d t_0}} \left\| \int (\tilde{L}_t)_{t_0}(x, z) dz \right\|^2 dx \right)$$

$$\left(\int_y |w_{2^n t'}(x_j - y)|^2 dy \right).$$

Following the line of proof of Lemma 7, it is easy to see that the first factor in the product is dominated by $C t_0^d e^{-\sqrt{d} t_0/cnt}$, and hence, that (25) is dominated by the same quantity $\times t_0^{-d}$.

This concludes the proof of Lemma 8. ∎

We are now ready to estimate $\max_{\|F\|_2 = 1} \||L_t F\||$. By duality this is also equal to $\max_{\||f_t\|| = 1} \| \int L_t^* f_t \, dt/t \|_2$. It is easy to see that the square of this expression is less than the operator-norm on $L^2(\mathbf{R}_+, dt/t)$ of the operator of kernel $k(s, t) = \|L_s L_t^*\|$. By Lemma 7, Lemma 8 and the remark preceding it, this is less than Cn. Hence

$$(26) \qquad \||L_t F\|| < Cn^{1/2} \|F\|_2.$$

6. Proof of Theorem 1

We recall that the numbers X_n and Y_n have been defined just before (10).

We wish to prove, for all $\varepsilon > 0$, the existence of a constant C_ε such that

$$(27) \qquad X_n \leqslant C_\varepsilon (1 + 2\sqrt{d} + \varepsilon)^n.$$

Let us recall that, by the remarks in Section 3 and (26),

$$Y_n \leqslant (11) + (12) + (13) \leqslant C(n + 1) + Cn^{1/2} + (13).$$

Also, by Lemma 3 and (6), if we define Z_n to be the best constant in the inequality

$$(28) \qquad \left(\int_{x \in Q, 0 < t < \delta(Q)} \|Q_{t/2^n}\{A_1\} K_{t/2^n} \dots K_{t/2^n} A_n(x)\|^2 \frac{dx \, dt}{t} \right)^{1/2} < C|Q|.$$

then

(29) $(13) \leqslant Z_n^{1/2} + Ch_{\mu_n}^{1/2},$

where μ_n denotes the Carleson-measure appearing in (28). By the remark following the proof of Lemma 7, we know that $h_{\mu_n} < C$. Therefore,

(30) $Y_n \leqslant 2Z_n + C(n+1).$

Finally the proof of Lemma 7 shows, modulo trivial modifications, that

$$Z_n \leqslant X_{n-1}^2(1 + C2^{-n/2}) \times d.$$

The factor d comes from the fact that in estimating c_{μ_n}, one has to use the Hilbert-Schmidt norm of A_n, which can be as large as $\sqrt{d}$.

From these two inequalities, we obtain

$$Y_n \leqslant 2\sqrt{d}(1 + C2^{-n/2})X_{n-1} + C(n+1).$$

Since $X_n \leqslant X_{n-1} + Y_n$, it follows

$$X_n \leqslant (1 + 2\sqrt{d})(1 + C2^{-n/2})\, X_{n-1} + C(n+1).$$

This implies (27) and proves Theorem 1. ∎

7. Application to the Kato-problem

Theorem 1 implies that the radius of convergence of (1) or (2) is $1/(1 + 2\sqrt{d})$ at least. There are various ways to see that this is not sharp. One of them is to observe that the constant 2 in (7) can be attained only for L^2−functions which have a particular distribution-function. Since the functions to which we apply (7) are images of bounded functions restricted to cubes, their distribution-functions have properties which permit to conclude that the constant 2 can be improved in our case, thus leading to an improvement of the constant $(1 + 2\sqrt{d})$ in Theorem 1. In this section we shall be interested only in those improvements which take in account the special form of the operators $\bar{K}_i$ which occur in the series (1) or (2).

The properties of the Riesz matrix which we shall use are that it is positive and defines an orthogonal projection. We shall be working with the series (2) and show the following.

Theorem 2. *The radius of convergence of the Kato functional is at least* α^{-1}, *where* α *is the largest positive root of the equation*

$$(31) \qquad \left(X - \frac{1}{2}\right)\sqrt{X^2 - 1} = X\sqrt{d}.$$

Before sketching the proof of this theorem, let us indicate that approximate numerical values for α^{-1} in dimensions 2, 3, 4, and 5 are respectively .474, .416, .376, and .347.

Sketch of proof: For simplicity we shall proceed as if the kernels of $(I - P_t)\tilde{R}$ and Q_t were supported in $\{\|x - y\| \leqslant t\}$. This of course is not true, but, as the proof of Lemma 7 shows, this is true for all practical purposes.

The fact that $\tilde{R}$ defines an orthogonal projection yields the following improvement of (10):

$$(32) \qquad X_n \leqslant \sqrt{Y_n^2 + X_{n-1}^2}.$$

To see this we just need to observe that for a function F in L^2, $\|\tilde{F}\|_2^2 + \|(I - \tilde{R})F\|_2^2 = \|F\|_2^2$. Hence $X_n \leqslant \max_{\lambda^2 + \mu^2 = 1} \lambda X_{n-1} + \mu Y_n$, which is exactly (32).

By the positivity of $\tilde{R}$ we can write it as $\frac{I+S}{2}$, where S is a contraction. Since $K_t = I - (I - P_t)\tilde{R} = I - (I - P_t)\frac{I+S}{2}$ we can rewrite it as

$$\frac{1}{2}P_t + \frac{1}{2}\left(I - (I - P_t)S\right).$$

It follows that

$$(33) \qquad Z_n \leqslant \frac{1}{2}(1 + C2^{-n})(Z_{n-1} + \sqrt{d}X_{n-1}), + \text{negligible error terms.}$$

Note that, by (32),

$$(34) \qquad X_n \leqslant \left(\sum_{k=1}^{n} Y_n^2\right)^{1/2}.$$

Combining (30) and (34) we obtain

$$(35) \qquad Z_n \leqslant \frac{1}{2}Z_{n-1} + \sqrt{d}\left(\sum_{j=1}^{n-1} Z_j^2\right)^{1/2} + \text{negligible error terms.}$$

Let us ignore the error terms in (35). Let $C > 0$ and $\beta > 1$ be such that for $j < n, Z_j < C\beta^j$. Then $Z_n < C\beta^n$ if β is such that

$$\beta^n \geqslant \frac{1}{2}\beta^{n-1} + \sqrt{d}\frac{\beta^{n-1}}{\sqrt{1 - \frac{1}{\beta^2}}},$$

that is, if $\beta \geqslant \alpha$. This implies Theorem 2 modulo the handling of the error terms, which is trivial. ∎

Let us mention that in working with the series (1) instead of (2) one does not seem to be able to improve (10) into (34). Hence one obtains significantly worse estimates for X_n for (1) than for (2), while, as the next section will suggest, one should expect a discrepancy growing slower than exponentially.

8. A direct proof of Kato's conjecture in dimension 1

We wish to prove the following:

Theorem 3. *For all $\varepsilon > 0$ there exists a constant C_ε such that for all $n \leqslant 1$, and $a_i \in L_{\mathbf{C}}^\infty(\mathbf{R}), 1 \leqslant i \leqslant n$, and $f \in L^2(\mathbf{R})$,*

$$(36) \qquad \left\|\left|Q_t \prod_{i=1}^n (\{a_i\}(I - P_t))\right|\right\| \leqslant C_\varepsilon (1 + \varepsilon)^n \prod_{i=1}^n \|a_i\|_\infty \|f\|_2.$$

The proof we shall give clearly yields multilinear estimates. However we shall work with one single bounded function a of norm 1, and we shall denote $\{a\}$ by a. This symbol will stand for $\{a\}^j$ for any power j, for instance in (43) and (44).

Proof: The proof of Theorem 3 relies on a trivial extension of one identity of [4], namely:

$$(37) \qquad Q_t a P_t = P_t\{P_t a\}Q_t + \{Q_t a\}P_t - Q_t\{Q_t a\}Q_t.$$

Let S_t be such that $P_t S_t = P_{\alpha t}$ for some $\alpha > 0$. Then $Q_t S_t = tDP_t S_t = tDP_{\alpha t} = \frac{1}{\alpha}Q_{\alpha t}$. This, together with (37) and a rescaling immediately gives:

Lemma 9. *For all $u > 0$ and $\alpha > 0$,*

$$(38) \qquad Q_{ut} a P_{\alpha t} = \frac{u}{\alpha} P_{ut}\{P_{ut}a\}Q_{\alpha t} + \{Q_{ut}a\}P_{\alpha t} - \frac{u}{\alpha}Q_{ut}\{Q_{ut}a\}Q_{\alpha t}.$$

For all $u \in]0, 1]$, we shall prove

$$(39) \qquad \left\|\left|Q_{ut}\left(a(I - P_t)\right)^n a P_t\right|\right\| \leqslant C(1 + \varepsilon)^n \|f\|_2.$$

This clearly implies (36). The reason for introducing the parameter u will become apparent during the proof.

Let p be an integer possibly equal to 1. We write $n = qp + r = q + (p-1)q + r = q + s$, with $r < p$. Let $\alpha > 1$ and $\beta = \alpha^q$. Finally, for $\gamma > 1$, we set $\tilde{Q}_t^\gamma = P_t - P_{\gamma t}$.

We wish to reduce the study of $\left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^n aP_t\right|\right|\right|$ to the study of

$$(40) \quad \left|\left|\left|Q_{ut}a(I - P_{\alpha t})a(I - P_{\alpha^2 t})a \ldots a(I - P_{\alpha^{q-1}t})\left(a(I - P_{\beta t})\right)^{s+1} aP_{\beta t}f\right|\right|\right|,$$

where we shall be able to take advantage of the factor $\frac{1}{\alpha}$ in Lemma 9. To do this we replace each P_t by the corresponding $P_{\gamma t}$, where γ is some appropriate power of α. Starting this process from the right and using Minkowsky's inequality we dominate the left hand side of (39) by the sum of (40) and the following expressions.

$$(41) \quad \begin{cases} \left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^n a\tilde{Q}_t^\beta f\right|\right|\right| \\[2mm] \left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^{n-1} a\tilde{Q}_t^\beta aP_{\beta t}f\right|\right|\right| \\[2mm] \left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^{n-2} a\tilde{Q}_t^\beta a(I - P_{\beta t})aP_{\beta t}f\right|\right|\right| \\[2mm] \qquad\qquad\qquad \cdot \\ \qquad\qquad\qquad \cdot \\ \qquad\qquad\qquad \cdot \\ \left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^{q-1} a\tilde{Q}_t^\beta\left(a(I - P_{\beta t})\right)^s aP_{\beta t}f\right|\right|\right| \end{cases}$$

$$(42) \quad \begin{cases} \left|\left|\left|Q_{ut}\left(a(I - P_t)\right)^{q-2} a\tilde{Q}_t^{\alpha^{q-1}}\left(a(I - P_{\beta t})\right)^{s+1} aP_{\beta t}f\right|\right|\right| \\[2mm] \qquad\qquad\qquad \cdot \\ \qquad\qquad\qquad \cdot \\ \left|\left|\left|Q_{ut}a\tilde{Q}_t^\alpha a(I - P_{\alpha^2 t})a \ldots a(I - P_{\alpha^{q-1}t})\left(a(I - P_{\beta t})\right)^{s+1} aP_{\beta t}f\right|\right|\right| \end{cases}$$

To study (40) we expand each $(I - P_{\gamma t})$ and regroup the resulting 2^n terms according to the location of the first $P_{\gamma t}$ which appears when going from the left. A new application of Minkowski's inequality then shows that (40) is less

than the sum of the following expressions:

(43)

$$\left\{ \begin{array}{l} \left|\left|\left| Q_{ut}\mathbf{a}P_{\alpha t}\mathbf{a}(I - P_{\alpha^2 t})\mathbf{a}..\,\mathbf{a}(I - P_{\alpha^{q-1}t})\Big(\mathbf{a}(I - P_{\beta t})\Big)^{s+1}\mathbf{a}P_{\beta t}f \right|\right|\right| \\ \qquad\qquad . \\ \qquad\qquad . \\ \qquad\qquad . \\ \left|\left|\left| Q_{ut}\mathbf{a}P_{\alpha^{q-1}t}\Big(\mathbf{a}(I - P_{\beta t})\Big)^{s+1}\mathbf{a}P_{\beta t} \right|\right|\right| \end{array} \right.$$

(44)

$$\left\{ \begin{array}{l} \left|\left|\left| Q_{ut}\mathbf{a}P_{\beta t}\Big(\mathbf{a}(I - P_{\beta t})\Big)^{s}\mathbf{a}P_{\beta t} \right|\right|\right| \\ \qquad\qquad . \\ \qquad\qquad . \\ \qquad\qquad . \\ \left|\left|\left| Q_{ut}\mathbf{a}P_{\beta t}\mathbf{a}P_{\beta t}f \right|\right|\right| \\ \left|\left|\left| Q_{ut}\mathbf{a}P_{\beta t}f. \right|\right|\right| \end{array} \right.$$

An application of Lemma 9 permits to decompose each of these expressions in the sum of three terms, the two last of which can be handled via the usual Carleson-measure argument, thanks to the following lemma.

Lemma 10. *Let $0 < \omega_0 \ldots \leqslant \omega_n$. Then the kernel $\Gamma_t(x, y)$ of*

$$P_{\omega_0}\mathbf{a}(I - P_{\omega_1})\mathbf{a}(I - P_{\omega_2})\ldots(I - P_{\omega_{n-1}})\mathbf{a}P_{\omega_n}$$

is dominated by $C\,(\omega_n/\omega_0)^{1/2}\,(1+n)^3\,\frac{(\omega_n t)^{1/2}}{(\omega_n t)^{3/2}+|x-y|^{3/2}}$. The same is true if $P_{\omega_0 t}$ is replaced by $Q_{\omega_0 t}$.

We defer the proof of this lemma until the end of the section.

We introduce some notations for best constants in quadratic estimates, for which we shall obtain estimates by induction. These constants depend on α which, for the time being, is fixed. Recall that $u \in]0, 1]$.

The first constant $D^u_{q,s}$ is the best constant in the inequality

(45) $$(40) \leqslant C\|f\|_2.$$

The constant E^u_k, independent of α, is the best constant in the inequality

(46) $$\left|\left|\left| Q_{ut}\Big(\mathbf{a}(I - P_t)\Big)^k\mathbf{a}P_t f \right|\right|\right| \leqslant C\|f\|_2.$$

We want to derive an inequality for $D^u_{q,s}$. As we already observed, (40) is less than the sum of $(q+s+1)$ quadratic quantities, $(q-1)$ for (43) and $(s+2)$ for

(44). Each of these quantities can be dominated by three others using Lemma 9. It will follow that $D^u_{q,s}$ is dominated by the sum of $3(q+s+1)$ numbers. Corresponding to the first term in Lemma 9 we have the following $(q+s+1)$ numbers:

$$\frac{u}{\alpha} D_{q-1,s}, \frac{u}{\alpha^2} D_{q-2,s}, \cdots, \frac{u}{\alpha^{q-1}} D_{1,s}, \frac{u}{\beta} E_s, \frac{u}{\beta} E_{s-1}, \cdots, \frac{u}{\beta} E_0, \frac{u}{\beta} c_0.$$

Here c_0 denotes the best constant in the inequality

$$|||Q_t f||| \leqslant C \|f\|_2.$$

Those numbers are obtained by rescaling and using the fact that the P_t's are contractions on L^2 and L^∞. The two last terms in Lemma 9 can be regrouped in a single one. Lemma 10 and the usual Carleson-measure argument give a global estimate in $C\beta^{1/2}(1+n)^4$. In doing this, one uses that the measure $|Q_{ut}a|^2 dx dt/t$ is a Carleson-measure uniformly in u.

The previous remarks yield the following :

(47) $$D^u_{q,s} \leqslant \sum_{j=1}^{q-1} \frac{u}{\alpha^j} D_{q-j,s} + \frac{u}{\beta} c_0 + \frac{u}{\beta} \sum_{m=0}^{s} E_m + C\alpha^{q/2}(1+n)^4.$$

The strategy is to use (47) to show that if

(48) $$E^u_m \leqslant C_1(1+\varepsilon)^m$$

for some $\varepsilon > 0$ and uniformly for $m \in \mathbf{N}$ and $u \in]0,1]$, then for some $\varepsilon' < \varepsilon$,

(49) $$D^u_{q,s} \leqslant C_2(q,s)(1+\varepsilon')^q(1+\varepsilon)^s$$

where $C_2(q,s)$ has the form $C(q+s+1)^4$ and is independent of $u \in]0,1]$, This will require an appropriate choice of α and p. Then one shall show that for some $\varepsilon'' < \varepsilon, C > 0$, and uniformly in $u \in]0,1]$,

(50) $$E^u_m < C(1+\varepsilon'')^m.$$

By iterating this procedure one can make ε as small as we want, thus proving the theorem.

Let us be more precise . Let $\varepsilon > 0$ and choose p such that

(51) $$2^{1/2p+1} < 1+\varepsilon < 2^{1/2p-2}.$$

Then we claim that if (48) holds, then (49) holds for any

(52) $$\varepsilon' > \left(\frac{2}{(1+\varepsilon)^{2p-2}}\right)^{1/3} - 1.$$

Notice that, by (51), $\left(\frac{2}{(1+\varepsilon)^{2p-2}}\right)^{1/3} - 1 < \varepsilon$. Also we claim that in (50) we can choose any ε'' such that

$$(53) \qquad \varepsilon'' > (1+\varepsilon)^{1-1/p}(1+\varepsilon')^{1/p} - 1.$$

Since $2^{1/2p+1}$ is the only fixed point of the transformation

$$x \longrightarrow \left(\frac{2}{x^{2p-2}}\right)^{1/3},$$

it follows by iteration that we can make ε as close as $2^{1/2p+1} - 1$ as we want. Then, by increasing the value of p we can get ε as close to 0 as we want. Hence, to finish the proof of the theorem, we just need to prove two claims and Lemma 10.

Proof of the first claim: We assume (48) and (51), and we want to conclude (49) for any ε' such that

$$(54) \qquad 1 + \varepsilon' > \left(\frac{2}{(1+\varepsilon)^{2p-2}}\right)^{1/3}.$$

Let $\alpha = (1+\varepsilon)^{2p-2}(1+\varepsilon')^2$. Notice that, by (54),

$$(55) \qquad \alpha(1+\varepsilon') > 2.$$

In order to deduce (49) from (48), we shall use an induction on q based on (47). Then it will be sufficient that C_2 be so that, for all $0 < k' \leqslant q$,

$$(56)$$

$$C_2(s,k')(1+\varepsilon)^s(1+\varepsilon')^{k'} \geqslant \sum_{j=1}^{k'-1} \frac{C_2(s,k'-j)}{\alpha^j}(1+\varepsilon)^s(1+\varepsilon')^{k'-j}$$

$$+ C\alpha^{k'/2}(s+k'+1)^4 + C_1\frac{s+1}{\alpha^{k'}}(1+\varepsilon)^s.$$

Observe that, by (55), $\sum_{j \geqslant 1}\left(\alpha(1+\varepsilon')\right)^{-j} < 1$. Hence, since $C_2(s,.)$ will be increasing one just needs

$$(57) \qquad C_2(s,k')(1+\varepsilon)^s(1+\varepsilon')^{k'} \geqslant C_3\alpha^{k'/2}(s+k'+1)^4 + \frac{s+1}{\alpha^{k'}}(1+\varepsilon)^s,$$

where C_3 is some constant which remains bounded if ε and ε' stay away from their minimum values and C_1 remains bounded. Let us set $C_2(s,k') = 2C_3(s+$

$k' + 1)^4$ and check that with this choice (57) is satisfied. Equivalently we need that for all $0 \leqslant k' \leqslant q$,

$$(58) \quad 2(s + k' + 1)^4(1 + \varepsilon)^s(1 + \varepsilon')^{k'} \geqslant \alpha^{k'/2}(s + k' + 1)^4 + \frac{s + 1}{\alpha^{k'}}(1 + \varepsilon)^s.$$

In (58), the second term of the right hand side is obviously less than half of the left hand side. To see that this is the same for the first term, we first observe than when $k' = q$ it follows from $\alpha^{q/2} = (1 + \varepsilon)^{(p-1)q}(1 + \varepsilon')^q \leqslant (1 + \varepsilon)^s(1 + \varepsilon')^q$. To deduce it for smaller values of k', just observe that $\alpha^{1/2} \geqslant (1 + \varepsilon')$. Hence, (58) is proved, from which follow (57) and (56). Hence (49) follows from (48) by induction, using (56) and (47). ∎

Proof of the second claim: With our choice for C_2, (49) reads as follows:

$$(59) \quad D^u_{q,s} \leqslant C(q + s + 1)^4(1 + \varepsilon)^s(1 + \varepsilon')^q.$$

We want to deduce (50) from (48) and (59), assuming that ε'' satisfies (53). The decomposition of the left hand side of (39) in (41)+(42)+(40) implies that E^u_n is dominated by the sum of $(s + 2) + (q - 1) + 1$ terms. To handle the $(s + 2)$ first ones, we decompose $\tilde{Q}^\beta_t$ as $-\int_1^\beta 2P_{ut}Q_{ut}\frac{du}{u}$. By (48), this gives a contribution of $c_0 \log \beta$ for the first one and $C_1 \log \beta(1 + \varepsilon)^j$, $0 \leqslant j \leqslant s$, for the $(s + 1)$ following. By (59) one has contributions of the form $C(q - j + s + 1)^4 j \log \alpha (1 + \varepsilon)^s(1 + \varepsilon')^{q-j}$, $1 \leqslant j \leqslant q - 1$ for the terms constituting (42) and finally $C_2(q, s)(1 + \varepsilon)^s(1 + \varepsilon')^q$ for (40). Summing these estimates, and then using (53) and the fact that $(q + 1)(p - 1) \geqslant s$ we obtain

$$E^u_n \leqslant C(q + s + 1)^6(1 + \varepsilon)^s(1 + \varepsilon')^q < C_{\varepsilon''}(1 + \varepsilon'')^n.$$

This proves the second claim. ∎

Proof of Lemma 10: An obvious estimate is

$$\Gamma_t(x, y) \leqslant \|P_{\omega_0 t}(x - \cdot)\|_2 \|P_{\omega_n t}(\cdot - y)\|_2 = Ct^{-1}(\omega_0 \omega_n)^{-1/2}.$$

This is sufficient when $|x - y| \leqslant C(1 + n)^{3/2}\omega_n t$ since in this case,

$$t^{-1}(\omega_0 \omega_n)^{-1/2} \leqslant C \left(\frac{\omega_n}{\omega_0}\right)^{1/2} (1 + n)^3 \frac{\omega_n t}{(\omega_n t)^2 + (x - y)^2} .$$

If $|x - y| \geqslant C(1 + n)^{3/2}\omega_n t$, we truncate $P_{\omega_0 t}(x - \cdot)$ and $P_{\omega_0 t}(\cdot - y)$ on balls respectively centered at x and y and of diameter $|x - y|/4$. The far-away parts have $L^2 -$ norms of the order of $|x - y|^{-3/2}\omega_0 t$ and $|x - y|^{-3/2}\omega_n t$ respectively. Hence, they will give contributions of at most $C\omega_n t|x - y|^{-3/2}(\omega_0 t)^{-1/2}$. It remains to control the contribution of the local parts. Here we use that their supports have a distance $|x - y|/4$. As we mentioned in the remark following the proof of Lemma 5, we can apply it in our situation and this gives a contribution $C^n c^{|x-y|/4\omega_n t}(\omega_0 \omega_n)^{-1/2}t^{-1}$. This concludes the proof of Lemma 10 and of Theorem 3. ∎

References

1. T. KATO, "*Perturbation Theory for Linear Operators*," Springer-Verlag, 1966.

2. A. MCINTOSH, On the comparability of $A^{1/2}$ and $A^{*1/2}$, *Proc.A.M.S.* **32** (1972), 430–434.

3. A.P. CALDERÓN, Commutators of Singular Integral Operators, *Proc. Nat. Acad. Sci. U.S.A.* **53** (1965), 1092–1099.

4. R. COIFMAN, A. MCINTOSH AND Y. MEYER, L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes, *Ann. of Math.* **116** (1982), 361–388.

5. A.P. CALDERÓN, Cauchy integrals on Lipschitz-curves and related operators, *Proc. Nat. Acad. Sci. U.S.A.* **74** (1977), 1324–1327.

6. E. FABES, D. JERISON AND C.E. KENIG, Multilinear Littlewood-Paley estimates with applications to partial differential equations, *Proc. Nat. Acad. Sci., U.S.A.* **79** (1982), 5746–5750.

7. R. COIFMAN, D.G. DENG AND Y. MEYER, Domaine de la racine carrée de certains opérateurs différentiels accrétifs, *Ann. Inst. Fourier* **33** (1983), 123–134.

Department of Mathematics
Princeton University
Fine Hall, Box 708
Princeton NJ 08544
U.S.A.